Die Edgeworth-Approximation zur Kalibrierung eines Lévy-Hybridmodells unter dem physikalischen Maß

Dissertation zur Erlangung des Doktorgrades der Fakultät für Mathematik und Physik der Albert-Ludwigs-Universität Freiburg im Breisgau

> vorgelegt von Manuel Polley Juni 2016

Dekan: Prof. Dr. Dietmar Kröner

Referenten: **Prof. Dr. Ernst Eberlein Prof. Dr. Kathrin Glau**

Datum der Promotion: 19. September 2016

Inhaltsverzeichnis

Bezeichnungen 3				
1	Einleitung und Motivation	5		
2	Präsentation des Modells2.1Der treibende Prozess2.2Annahmen über die Dynamik unter dem physikalischen Maβ2.3Gemeinsame Momente von relativen Zuwächsen2.4Der risikoneutrale Fall	9 9 11 17 23		
3	Kumulanten von stochastischen Integralen nach einem Lévy-Prozess und die multivariate Edgeworth-Approximation3.1Die Annahme (EM) und Kumulanten von Zufallsvektoren3.2Gemeinsame Kumulanten von stochastischen Integralen nach	25 25		
	einem Lévy-Prozess	41 46 49		
4	Eine multivariate Erweiterung des Cramér-von-Mises- und des Anderson-Darling-Tests 4.1 Ein Anpassungstest für \mathbb{R}^N -wertige Zufallsvektoren 4.2 Ein Anpassungstest für Copulas	59 59 74 88		
5	 Implementierung der Edgeworth-Approximation 5.1 Technische Lemmata	97 98 .03 .11		
6	Kalibrierung unter dem physikalischen Maß16.1Die Modellparameter16.2Das Kalibrierungsportfolio und seine Modellierung16.3Die Kalibrierungsarten1	36 .36 .39 .40		

	6.4	Kalibrierungsergebnisse für deutsche Staatsanleihen und den DAX	145
7	Zusa	ammenfassung	161
Α	Anh A.1	ang Ein alternativer Beweis für $\mathbb{P}(D^2_{\mathbb{B},\mu} < \infty) = 1$ und einige fast sichere Pfadeigenschaften des Brownschen Blatts	162 162
Literaturverzeichnis 18			

Bezeichnungen

$\mathbb{N} := \{1, 2, 3, \ldots\}$	Menge der natürlichen Zahlen ohne Null
$\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}$	Menge der natürlichen Zahlen mit Null
\mathbb{R}	Menge der reellen Zahlen
$\mathbb{R}_+ := \{ x \in \mathbb{R} \mid x \ge 0 \}$	Menge der nichtnegativen reellen Zahlen
$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$	Menge der erweiterten reellen Zahlen
C	Menge der komplexen Zahlen
$\Re(z),\Im(z)$	Realteil bzw. Imaginärteil der komplexen Zahl z
$\langle \cdot, \cdot \rangle$	Komplexe Fortsetzung des euklidischen Skalarpro-
	dukts. Achtung! Dabei handelt es sich nicht um die
	sonst übliche positiv definite Hermitesche Sesqui-
	linearform!
.	Für Vektoren: Euklidische Norm.
1 1	für Multiindices: Summe der Komponenten,
	für Mengen: Mächtigkeit
.	Natürliche Matrixnorm
A^{C}	Komplement der Menge A
$\mathcal{B}(A)$	Borelsche σ -Algebra auf A
$C^{0}([0,1]^{N},\mathbb{R})$	Menge der reellwertigen, stetigen Funktionen
	auf $[0, 1]^N$
$\mathbb{E}[X \mathcal{F}]$	Bedingte Erwartung eines Zufallsvektors X gege-
	ben die σ -Algebra \mathcal{F} . Den üblichen Konventionen
	entsprechend sind Gleichungen, in denen ein be-
	dingter Erwartungswert außerhalb eines Integrals
	vorkommt, stets als fast sichere Gleichheiten zu ver-
	stehen.
E_N	$N \times N$ -Einheitsmatrix
f_X	Wahrscheinlichkeitsdichte der Verteilung von X
F_X	Verteilungsfunktion von X
$F_{X_1 X_2}$	Bedingte Verteilungsfunktion von X_1 gegeben X_2
K_X	Kumulantenerzeugende Funktion von X
$L^p(\Omega, \mathcal{F}, \mu)$	Raum der <i>p</i> -fach integrierbaren Funktionen auf
	dem Maßraum $(\Omega, \mathcal{F}, \mu)$
$\mathcal{L}(X)$	Verteilung von X
M_X	Momenterzeugende Funktion von X
$\mathcal{U}\left([0,1]^N\right)$	Gleichverteilung auf $([0,1]^N, \mathcal{B}([0,1]^N))$
κ_{α}^{X}	Kumulante von X der Ordnung α
$\lambda^{\overline{N}}$	Lebesgue-Maß auf $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$
φ_X	Charakteristische Funktion von X

1 Einleitung und Motivation

Es ist eine in Finanzkreisen allgemein bekannte Tatsache, dass sich durch eine geschickte Diversifikation von Investitionspositionen Gewinnchancen erhöhen und/oder Verlustrisiken verringern lassen. Zur Quantifizierung derartiger Diversifikationseffekte wird ein Finanzmarktmodell benötigt, das sowohl isoliert die Dynamik der einzelnen betrachteten Finanzprodukte als auch deren Abhängigkeiten konsistent beschreibt.

Das Interesse an dieser Thematik geht weit über den Bankensektor mit seinem Investmentgeschäft hinaus. In der Versicherungsindustrie kommt ihr ebenfalls eine zentrale Bedeutung zu, beispielsweise bei der Wahl der Portfoliostrukturierung für kurzfristig verfügbare Finanzanlagen, die eine Versicherung vorhält, um diese im Schadensfall zu liquidieren und so ohne Verzögerung ihren Verpflichtungen nachzukommen.

In der vorliegenden Arbeit wird zu diesem Zweck ein Lévy-Hybridmodell eingeführt, das den Ansatz des Modells von Heath, Jarrow und Morton (1992) mit dem Ansatz des Modells von Black und Scholes (1973) und Merton (1973) kombiniert und so die gemeinsame Dynamik der Zinsstruktur und einer Aktie, also von zwei unterschiedlichen Marktsegmenten, beschreibt. Weil die Erfahrung immer wieder zeigt, dass ein Normalverteilungsansatz zur Erklärung von Beobachtungen am Finanzmarkt nicht die dafür erforderliche Flexibilität bietet, wird als treibender Prozess nicht bloß ein Wiener-Prozess, sondern ein zeitinhomogener Lévy-Prozess verwendet. Insbesondere wird so der Teil des Lévy-Hybridmodells, der die Zinsstruktur beschreibt, vom HJM-Modell zum Lévy-Forward-Rate-Modell.

Ziel der Arbeit ist die Kalibrierung des Lévy-Hybridmodells unter dem physikalischen Maß. Neben einer Maximum-Likelihood-Kalibrierung werden auch zwei alternative Kalibrierungsarten eingeführt und eingesetzt. Dabei liegt der Fokus nicht auf einzelnen Randverteilungen, sondern stattdessen auf der gemeinsamen Verteilung von mehreren Marktgrößen. Erst hierdurch wird die gewünschte quantitative Untersuchung von Diversifikationseffekten möglich.

In ihrer richtungsweisenden Arbeit zeigen Eberlein und Kluge (2007), dass das Lévy-Forward-Rate-Modell, das wie bereits erwähnt mit dem Teil des Lévy-Hybridmodells zur Beschreibung der Zinsstruktur zusammenfällt, für die isolierte Beschreibung von Dichten von Log-Returns sehr gute Ergebnisse liefert. Weiterhin zeigen Beinhofer, Eberlein, Janssen und Polley (2011), dass es für die isolierte Beschreibung der Korrelationsstruktur, ohne dabei Rücksicht auf Dichten von Randverteilungen zu nehmen, ebenfalls sehr gute Ergebnisse liefert. Durch die Betrachtung der gemeinsamen Verteilung von mehreren Marktgrößen gehen die Untersuchungen in der vorliegenden Arbeit über die bekannten Ergebnisse hinaus. Insbesondere wird so ebenfalls untersucht, ob das Lévy-Hybridmodell sowohl Dichten von Randverteilungen als auch die Korrelationsstruktur konsistent mit einem einzigen Parametersatz abbilden kann.

Neben der zusätzlichen Einbeziehung einer Aktie wird auch in technischer Hinsicht eine Verallgemeinerung zu Eberlein und Kluge (2007) vorgenommen. Die Autoren kalibrieren das Lévy-Forward-Rate-Modell unter dem physikalischen Maß an Log-Returns. Dabei verwenden sie als Treiber einen eindimensionalen GH-Prozess und als Volatilitätsstruktur eine Konstante, also die Ho-Lee-Volatilitätsstruktur. Ferner wählen sie ihre Zeitskala derart, dass die Länge der Zeitintervalle, die den Log-Returns zugrunde liegen, gleich eine Modellzeiteinheit ist. In dieser Konfiguration sind die Zufallsvariablen, durch die die Log-Returns modelliert werden, eindimensional GH-verteilt, was für die numerische Auswertbarkeit der Zielfunktion für die Maximum-Likelihood-Kalibrierung einen enormen Vorteil darstellt.

In der vorliegenden Arbeit hingegen erfolgt die Kalibrierung nicht an einzelnen Log-Returns, sondern an Log-Return-Vektoren, also Vektoren, deren Einträge Log-Returns von verschiedenen Marktgrößen sind. Als Treiber wird ein mehrdimensionaler GH-Prozess verwendet. Weiterhin wird eine allgemeinere, nicht konstante Volatilitätsstruktur in Betracht gezogen. Ferner ist die Länge der Zeitintervalle, die den Log-Return-Vektoren zugrunde liegen, nicht auf eine Modellzeiteinheit festgesetzt. Dies hat unter anderem zur Folge, dass die Zufallsvektoren, die zur Modellierung der Log-Return-Vektoren verwendet werden, im Allgemeinen nicht mehrdimensional GH-verteilt sind und auch keine geschlossene Form für ihre Dichte bekannt ist. Somit ist man bei einer Maximum-Likelihood-Kalibrierung auf numerische Verfahren zur approximativen Berechnung der Dichte angewiesen. Die sonst übliche Vorgehensweise der Fourier-Inversion der charakteristischen Funktion ist hier jedoch für eine Kalibrierung nicht anwendbar, weil sie durch das numerisch zu lösende Mehrfachintegral zu rechenintensiv wäre. Stattdessen wird die multivariate Edgeworth-Approximation als Lösungsansatz herangezogen. Hierbei wird die charakteristische Funktion eines Zufallsvektors unter Verwendung seiner Kumulanten in geeigneter Weise durch eine Funktion approximiert, deren Fourier-Inversion zu einer geschlossenen, numerisch effizient auswertbaren Form führt. Diese geschlossene Form wird als Approximation der Dichte verwendet. Es zeigt sich, dass hierdurch, zusammen mit einer aufwändigen Programmierung sowie dem Einsatz von starker Hardware, eine Maximum-Likelihood-Kalibrierung des Lévy-Hybridmodells wirklich möglich ist.

Die Arbeit ist folgendermaßen gegliedert. In Abschnitt 2 wird das Lévy-Hybridmodell vorgestellt und einige seiner Eigenschaften untersucht. Insbesondere wird auch eine Bedingung an die Driftstrukturen angegeben, die impliziert, dass das Modell direkt unter einem Martingalmaß arbeitet und somit zur Bewertung von Derivaten geeignet ist.

In Abschnitt 3 werden Eigenschaften der Klasse von Zufallsvektoren, die der Annahme (EM) genügen, untersucht. Hierunter ist auch ein Konvergenzsatz, der äquivalente Bedingungen angibt, unter denen sich die Annahme (EM) von einer konvergenten Folge von Zufallsvektoren auf ihren Grenzwert überträgt. Ferner wird eine Formel zur Berechnung der Kumulanten eines Vektors, dessen Einträge stochastische Integrale sind, entwickelt und es wird die multivariate Edgeworth-Approximation vorgestellt.

In Abschnitt 4 werden neue Anpassungstests, die multivariate Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests darstellen, eingeführt und untersucht.

Abschnitt 5 gibt einen Uberblick über die Implementierung der Edgeworth-Approximation. Hierbei werden unter anderem die dabei auftretenden Probleme diskutiert und das Konzept der Code-Generierung sowie der Einsatz von Grafikprozessoren als mögliche Lösungsansätze vorgestellt.

In Abschnitt 6 wird unter Verwendung der Erkenntnisse aus den vorhergehenden Abschnitten das Lévy-Hybridmodell unter dem physikalischen Maß an Log-Return-Vektoren kalibriert. Auf der Grundlage der zuvor behandelten Anpassungstests werden die zwei bereits erwähnten neuen Kalibrierungsarten eingeführt. Schließlich werden die Kalibrierungsergebnisse für deutsche Staatsanleihen und den DAX präsentiert und es wird ferner die Leistungsfähigkeit der neuen Kalibrierungsarten mit der der Maximum-Likelihood-Kalibrierung verglichen.

Die eigentliche Arbeit schließt mit der Zusammenfassung in Abschnitt 7.

Im Anhang A.1 wird noch ein alternativer Beweis für eine Tatsache, die im Laufe der Arbeit bereits bewiesen wird, vorgestellt. Dieser alternative Beweis stellt eine Verallgemeinerung des entsprechenden historischen Beweises von Anderson und Darling (1952) vom eindimensionalen auf den mehrdimensionalen Fall dar. Auf dem Weg dorthin werden unter anderem einige Resultate über fast sichere Pfadeigenschaften des Brownschen Blatts gezeigt, die von eigenem mathematischem Interesse sind.

An dieser Stelle möchte ich Herrn Prof. Dr. Ernst Eberlein für die Betreuung meiner Dissertation und für seine Geduld meinen ganz besonderen Dank zum Ausdruck bringen. In den Betreuungsgesprächen wirkte er immer wieder motivierend auf mich ein und durch wertvolle Diskussionen beflügelte er meine mathematische Kreativität und Phantasie. Weiterhin möchte ich ihm auch für seinen hervorragenden Vorlesungszyklus über Wahrscheinlichkeitstheorie herzlich danken, in dem ich die erforderlichen Grundlagen für meine Arbeit erlernt habe. Ganz besonders möchte ich ihm auch dafür danken, dass er mich als Doktorand angenommen hat, obwohl ich Diplom-Physiker und kein Diplom-Mathematiker bin, und auch dafür, dass er darauf vertraut hat, dass ich die Arbeit zum Abschluss bringen werde.

2 Präsentation des Modells

Das hier vorgestellte Lévy-Hybridmodell, dessen Kalibrierung unter dem physikalischen Maß im Fokus der vorliegenden Arbeit steht, beschreibt die gemeinsame Dynamik der Zinsstruktur und einer Aktie. Dabei wird die Zinsstruktur durch ein Lévy-Forward-Rate-Modell beschrieben, wie es in Eberlein und Raible (1999) eingeführt, in Eberlein und Özkan (2003) verallgemeinert und in Eberlein, Jacod und Raible (2005), Eberlein und Kluge (2006, 2007), Beinhofer, Eberlein, Janssen und Polley (2011) sowie weiteren darauf aufbauenden Veröffentlichungen untersucht wird. Sein treibender zeitinhomogener Lévy-Prozess wird mit L bezeichnet. Die Aktie wird durch eine Art Black-Scholes-Modell beschrieben, das jedoch nicht wie im klassischen Ansatz von einem Wiener-Prozess, sondern ebenfalls von L getrieben wird.

Abschnitt 2.1 ist dem treibenden zeitinhomogenen Lévy-Prozess L gewidmet. In Abschnitt 2.2 werden die Annahmen über die Dynamik unter dem physikalischen Maß formuliert. Als erste Schlussfolgerungen hieraus ergeben sich unter anderem einige Eigenschaften von Log-Returns, die für die weiteren Überlegungen und insbesondere für die angestrebte Kalibrierung essenziell sind. In Abschnitt 2.3 werden gemeinsame Momente von relativen Zuwächsen von Nullcoupon-Anleihen und der Aktie untersucht. In Abschnitt 2.4 wird schließlich eine Driftbedingung angegeben, die impliziert, dass das Modell direkt unter einem risikoneutralen Maß arbeitet.

2.1 Der treibende Prozess

Die Präsentation der für die vorliegende Arbeit relevanten Eigenschaften des treibenden Prozesses orientiert sich an Beinhofer, Eberlein, Janssen und Polley (2011, Abschnitt 2) und den darin enthaltenen Referenzen. Für eine umfassende Darstellung, in der auch die Beweise enthalten sind, siehe Kluge (2005, Abschnitt 1.3).

Seien $T^* \in (0, \infty)$ ein endlicher Zeithorizont und $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ eine vollständige stochastische Basis, die Filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T^*]}$ genüge den üblichen Bedingungen und es gelte $\mathcal{F} = \mathcal{F}_{T^*}$. Ferner sei $D \in \mathbb{N}$. Der treibende Prozess $L = (L_t)_{t \in [0,T^*]}$ sei ein *D*-dimensionaler zeitinhomogener Lévy-Prozess, also ein adaptierter Prozess mit stochastisch unabhängigen Zuwächsen und absolut stetigen Charakteristiken, die hier mit $(b_t, c_t, F_t)_{t \in [0,T^*]}$ bezeichnet werden.

Es wird angenommen, dass

$$\int_{0}^{T^{*}} \left(|b_{s}| + ||c_{s}|| + \int_{\mathbb{R}^{D}} (|x|^{2} \wedge 1) F_{s}(\mathrm{d}x) \right) \,\mathrm{d}s < \infty$$

gilt, wobei hier $||c_s|| := \sup_{\{x \in \mathbb{R}^D | |x| \le 1\}} |c_s x|$ die natürliche Matrixnorm für $D \times D$ -Matrizen ist. Ferner wird angenommen, dass $L_0 = 0$ gilt, dass alle Pfade von L càdlàg sind und dass L der folgenden Annahme ($\mathbb{E}M$) über exponentielle Momente genügt.

Definition 1 (Annahme ($\mathbb{E}\mathbb{M}$) für einen zeitinhomogenen Lévy-Prozess). Der zeitinhomogene Lévy-Prozess L genügt der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$, falls

$$\int_0^{T^*} \int_{\{\tilde{x} \in \mathbb{R}^D \mid |\tilde{x}| > 1\}} \exp\left(\langle u, x \rangle\right) \, F_s(\mathrm{d}x) \, \mathrm{d}s < \infty$$

für alle $u \in [-M, M]^D$ gilt. Ohne Beschränkung der Allgemeinheit kann dann ebenfalls angenommen werden, dass $\int_{\{\tilde{x}\in\mathbb{R}^D\mid |\tilde{x}|>1\}} \exp(\langle u,x\rangle) F_s(\mathrm{d}x) < \infty$ für alle $s \in [0, T^*]$ gilt.

Die Verwendung eines zeitinhomogenen Lévy-Prozesses als Treiber, der der Annahme (EM) genügt, ist aus theoretischer Sicht eine gewisse Restriktion. Aus der Sicht eines Praktikers stellt sie jedoch keine nennenswerte Einschränkung dar. Die Klasse der Prozesse, die dieser Annahme genügen, ist für praktische Anwendungen reichhaltig und flexibel genug.

Die Annahme ($\mathbb{E}\mathbb{M}$) impliziert $\mathbb{E}[L_t] < \infty$ für alle $t \in [0, T^*]$. Somit kann für alle $t \in [0, T^*]$ in der Lévy-Khintchine-Darstellung der charakteristischen Funktion von L_t durch eine Modifikation von b_t auf eine Abschneidefunktion verzichtet werden. Im Folgenden sollen unter $(b_t, c_t, F_t)_{t \in [0, T^*]}$ stets die Charakteristiken verstanden werden, die man erhält, wenn man keine Abschneidefunktion verwendet.

Für $s \in [0, T^*]$ wird die Funktion $\theta_s : [-M, M]^D \to \mathbb{R}$ definiert durch

$$\theta_s(z) := \langle z, b_s \rangle + \frac{1}{2} \langle z, c_s z \rangle + \int_{\mathbb{R}^D} \left(e^{\langle z, x \rangle} - 1 - \langle z, x \rangle \right) F_s(\mathrm{d}x),$$

wobei M die Konstante aus der Annahme ($\mathbb{E}M$) ist. Die Beweisidee des nun folgenden Satzes geht auf Eberlein und Raible (1999, Lemma 3.1) zurück, die ihn für den Spezialfall beweisen, dass L ein Lévy-Prozess ist. Ein Beweis für den hier angegebenen allgemeinen Fall ist in Kluge (2005, Proposition 1.9) zu finden.

Satz 2. Seien $t, T \in [0, T^*]$ mit $t \leq T$, $\varepsilon \in (0, 1)$ und

$$f: [0, T^*] \to [-(1-\varepsilon)M, (1-\varepsilon)M]^D$$

eine stetige Funktion.

Dann gilt

$$\mathbb{E}\left[\exp\left(\int_{t}^{T} f(s) \, \mathrm{d}L_{s}\right)\right] = \exp\left(\int_{t}^{T} \theta_{s}\left(f(s)\right) \, \mathrm{d}s\right).$$

2.2 Annahmen über die Dynamik unter dem physikalischen Maß

2.2.1 Die modellierten Größen und ihre Bezeichnungen

Die Zinsstruktur: Im betrachteten Modellmarkt existiert für jeden Fälligkeitszeitpunkt $T \in [0, T^*]$ eine Nullcoupon-Anleihe ohne Ausfallrisiko und mit Nominalwert 1. Ihr Preis zum Zeitpunkt $t \in [0, T]$ wird mit B(t, T) bezeichnet. Es wird angenommen, dass sich die Preise der Nullcoupon-Anleihen konsistent über eine Familie von momentanen Forward-Raten $f(t, T), 0 \leq$ $t \leq T \leq T^*$, berechnen lassen. Ferner existiert ein risikoloses Bankkonto B_t , das stetig mit der Short-Rate r(t) verzinst wird.

Hier sei angemerkt, dass die momentanen Forward-Raten f(t,T) und die Short-Rate r(t) theoretische Konstrukte sind und in realen Märkten nicht direkt beobachtet werden können.

Die Aktie: Weiter existiert im Modellmarkt eine Aktie, die keine Dividende zahlt. Ihr Preis zum Zeitpunkt $t \in [0, T^*]$ wird mit S(t) bezeichnet. Zur Vereinfachung wird ohne Beschränkung der Allgemeinheit angenommen, dass S(0) = 1 gilt. Stimmrechte auf der Hauptversammlung der Aktiengesellschaft oder andere Rechte, die aus dem Besitz der Aktie resultieren, werden aus der Betrachtung ausgeschlossen.

Die Annahme, dass die Aktie keine Dividende zahlt, dient der Vereinfachung und stellt nur eine sehr geringe Einschränkung dar. Man kann sich vorstellen, dass jede ausgezahlte Dividende sofort wieder in die Aktie reinvestiert wird, wodurch eine synthetische Aktie entsteht, die keine Dividende zahlt. Dieses Verfahren ist zumindest immer dann anwendbar, wenn zu den Zahlungszeitpunkten der Dividende ein Kauf der Aktie in dem dafür erforderlichen Umfang möglich ist.

2.2.2 Die Dynamik der modellierten Größen

Für $t, T \in [0, T^*]$ mit $t \leq T$ wird den modellierten Größen die Dynamik

$$f(t,T) = f(0,T) + \int_{0}^{t} \alpha_{B}(s,T) \,\mathrm{d}s - \int_{0}^{t} \sigma_{B}(s,T) \,\mathrm{d}L_{s}$$
(1)

$$B(t,T) = \exp\left(-\int_{t}^{T} f(t,u) \, \mathrm{d}u\right)$$
(2)

$$r(t) = f(t,t) \tag{3}$$

$$B_t = \exp\left(\int_0^t r(s) \, \mathrm{d}s\right) \tag{4}$$

$$S(t) = \exp\left(\int_0^t r(s) \, \mathrm{d}s + \int_0^t \alpha_S(s) \, \mathrm{d}s + \int_0^t \sigma_S(s) \, \mathrm{d}L_s\right)$$
(5)

unterstellt. Dabei seien die Startwerte der momentanen Forward-Raten $f(0, \cdot) : [0, T^*] \to \mathbb{R}$, die Driftstruktur der Zinsstruktur $\alpha_B : [0, T^*] \times [0, T^*] \to \mathbb{R}$, die Volatilitätsstruktur der Zinsstruktur $\sigma_B : [0, T^*] \times [0, T^*] \to \mathbb{R}^D$, die Driftstruktur der Aktie $\alpha_S : [0, T^*] \to \mathbb{R}$ und die Volatilitätsstruktur der Aktie $\sigma_S : [0, T^*] \to \mathbb{R}$ und die Volatilitätsstruktur der Aktie $\sigma_S : [0, T^*] \to \mathbb{R}^D$ deterministisch und stetig.

Bemerkung 3. Die Abhängigkeit der beiden Marktsegmente im Modell wird durch zwei ausschlaggebende Momente erzeugt. Erstens werden die Zinsstruktur und die Aktie von demselben zeitinhomogenen Lévy-Prozess L getrieben. Eine anschauliche Vorstellung hiervon ist, dass die Marktschocks, die ja durch Bewegungen des treibenden Prozesses modelliert werden, sich immer auf beide Marktsegmente auswirken. Dabei hängt das quantitative Ausmaß eines Schocks auf die unterschiedlichen Marktsegmente zusätzlich von der jeweiligen Volatilitätsstruktur ab. Zweitens enthält der Preis der Aktie in Gleichung (5) den Faktor $\exp(\int_0^t r(s) ds)$. Hierdurch erhält die Drift der Aktie eine Komponente, die von der Zinsstruktur, also dem anderen Marktsegment, abhängt.

Bemerkung 4. Als mögliche weitere Verallgemeinerung kann angenommen werden, dass $\alpha_B, \sigma_B, \alpha_S$ und σ_S nicht deterministisch sind sondern vom Zufall abhängen. Für die Implementierung werden jedoch ausschließlich deterministische Drift- und Volatilitätsstrukturen verwendet. Deshalb wird in der vorliegenden Arbeit von Anfang an unter dieser Annahme gearbeitet.

2.2.3 Erste Folgerungen aus den Annahmen

Für $s, t, T \in [0, T^*]$ werden die Abkürzungen

$$A_B(s,T) := \int_{s\wedge T}^T \alpha_B(s,u) \, \mathrm{d}u \quad \text{und} \quad A_B(s,t,T) := A_B(s,T) - A_B(s,t),$$

$$\Sigma_B(s,T) := \int_{s\wedge T}^T \sigma_B(s,u) \, \mathrm{d}u \quad \text{und} \quad \Sigma_B(s,t,T) := \Sigma_B(s,T) - \Sigma_B(s,t)$$

sowie

$$\tilde{\alpha}_S(s,t) := \alpha_S(s) + A_B(s,t)$$
 und $\tilde{\sigma}_S(s,t) := \sigma_S(s) - \Sigma_B(s,t)$

eingeführt.

Lemma 5. Für $t, T \in [0, T^*]$ mit $t \leq T$ sind der Preis einer Nullcoupon-Anleihe, das Bankkonto und der Preis der Aktie gegeben durch

$$B(t,T) = \frac{B(0,T)}{B(0,t)} \exp\left(-\int_{0}^{t} A_{B}(s,t,T) \,\mathrm{d}s + \int_{0}^{t} \Sigma_{B}(s,t,T) \,\mathrm{d}L_{s}\right),$$

$$B(t,T) = B(0,T) \exp\left(\int_{0}^{t} \left(r(s) - A_{B}(s,T)\right) \,\mathrm{d}s + \int_{0}^{t} \Sigma_{B}(s,T) \,\mathrm{d}L_{s}\right),$$

$$B_{t} = \frac{1}{B(0,t)} \exp\left(\int_{0}^{t} A_{B}(s,t) \,\mathrm{d}s - \int_{0}^{t} \Sigma_{B}(s,t) \,\mathrm{d}L_{s}\right),$$

$$S(t) = \frac{B(t,T)}{B(0,T)} \exp\left(\int_{0}^{t} \tilde{\alpha}_{S}(s,T) \,\mathrm{d}s + \int_{0}^{t} \tilde{\sigma}_{S}(s,T) \,\mathrm{d}L_{s}\right)$$

und

$$S(t) = \frac{1}{B(0,t)} \exp\left(\int_{0}^{t} \tilde{\alpha}_{S}(s,t) \,\mathrm{d}s + \int_{0}^{t} \tilde{\sigma}_{S}(s,t) \,\mathrm{d}L_{s}\right).$$

Beweis: Die ersten drei behaupteten Gleichungen sind aus Beinhofer, Eberlein, Janssen und Polley (2011, Abschnitt 3) entnommen. Die verbleibenden beiden behaupteten Gleichungen folgen hieraus unter zusätzlicher Berücksichtigung von Gleichung (5) und der Definition von $\tilde{\alpha}_S$ und $\tilde{\sigma}_S$. **Definition 6** (Forward-Preis, relativer Zuwachs, modifizierter relativer Zuwachs, Log-Return). Seien $t, T \in [0, T^*]$ und $\Delta \in \mathbb{R}_+$ mit $t \leq t + \Delta \leq T$.

Für den Forward-Preis zum Zeitpunkt t und Erfüllungszeitpunkt $t + \Delta$ werden die folgenden Abkürzungen eingeführt. Ist der Basiswert des Forward-Vertrags eine Nullcoupon-Anleihe mit Fälligkeitszeitpunkt T, so wird der Forward-Preis mit

$$\operatorname{Fwd}_{B(\cdot,T)}(t,t+\Delta) := \frac{B(t,T)}{B(t,t+\Delta)}$$

bezeichnet. Dient die Aktie als Basiswert, so wird der Forward-Preis mit

$$\operatorname{Fwd}_{S(\cdot)}(t, t + \Delta) := \frac{S(t)}{B(t, t + \Delta)}$$

bezeichnet.

Der relative Zuwachs der Nullcoupon-Anleihe mit Fälligkeitszeitpunkt T und der relative Zuwachs der Aktie, jeweils über das Zeitintervall $(t, t + \Delta]$, sind definiert durch

$$R_{B(\cdot,T)}(t,t+\Delta) := \frac{B(t+\Delta,T) - Fwd_{B(\cdot,T)}(t,t+\Delta)}{Fwd_{B(\cdot,T)}(t,t+\Delta)}$$

und

$$R_{S(\cdot)}(t, t + \Delta) := \frac{S(t + \Delta) - Fwd_{S(\cdot)}(t, t + \Delta)}{Fwd_{S(\cdot)}(t, t + \Delta)}$$

Der modifizierte relative Zuwachs der Nullcoupon-Anleihe mit Fälligkeitszeitpunkt T und der modifizierte relative Zuwachs der Aktie, jeweils über das Zeitintervall $(t, t + \Delta]$, sind definiert durch

$$\mathbf{R}_{B(\cdot,T)}^{\mathrm{mod}}(t,t+\Delta) := \mathbf{R}_{B(\cdot,T)}(t,t+\Delta) + 1 = \frac{B(t+\Delta,T)}{\mathrm{Fwd}_{B(\cdot,T)}(t,t+\Delta)}$$

und

$$\mathrm{R}^{\mathrm{mod}}_{S(\cdot)}(t,t+\Delta) := \mathrm{R}_{S(\cdot)}(t,t+\Delta) + 1 = \frac{S(t+\Delta)}{\mathrm{Fwd}_{S(\cdot)}(t,t+\Delta)}.$$

Der Log-Return der Nullcoupon-Anleihe mit Fälligkeitszeitpunkt T und der Log-Return der Aktie, jeweils über das Zeitintervall $(t, t + \Delta]$, sind definiert durch

$$LR_{B(\cdot,T)}(t,t+\Delta) := \log\left(R_{B(\cdot,T)}^{\text{mod}}(t,t+\Delta)\right) = \log\left(\frac{B(t+\Delta,T)}{\text{Fwd}_{B(\cdot,T)}(t,t+\Delta)}\right)$$

und

$$\operatorname{LR}_{S(\cdot)}(t,t+\Delta) := \log\left(\operatorname{R}_{S(\cdot)}^{\operatorname{mod}}(t,t+\Delta)\right) = \log\left(\frac{S\left(t+\Delta\right)}{\operatorname{Fwd}_{S(\cdot)}(t,t+\Delta)}\right).$$

Relative Zuwächse stellen für praktische Anwendungen ökonomisch gut interpretierbare Größen dar. Für die mathematische Untersuchung im betrachteten Modell ist jedoch die Arbeit mit modifizierten relativen Zuwächsen angenehmer. Weil sich relative Zuwächse und modifizierte relative Zuwächse lediglich um eine Konstante unterscheiden, lassen sich alle Ergebnisse über modifizierte relative Zuwächse in kanonischer Weise auf relative Zuwächse übertragen.

Das nun folgende Lemma liefert Darstellungen von modifizierten relativen Zuwächsen und Log-Returns, die für deren Untersuchung geeignet sind.

Lemma 7. Seien $t, T \in [0, T^*]$ und $\Delta \in \mathbb{R}_+$ mit $t \leq t + \Delta \leq T$.

Dann gilt

$$R_{B(\cdot,T)}^{\text{mod}}(t,t+\Delta) = \exp\left(-\int_{t}^{t+\Delta} A_B(s,t+\Delta,T)\,\mathrm{d}s\right) + \int_{t}^{t+\Delta} \Sigma_B(s,t+\Delta,T)\,\mathrm{d}L_s\right) + \left(\int_{t}^{t+\Delta} \tilde{\alpha}_S(s,t+\Delta)\,\mathrm{d}s + \int_{t}^{t+\Delta} \tilde{\sigma}_S(s,t+\Delta)\,\mathrm{d}L_s\right) + \left(\sum_{t}^{t+\Delta} \tilde{\alpha}_S(s,t+\Delta)\,\mathrm{d}s + \int_{t}^{t+\Delta} \tilde{\sigma}_S(s,t+\Delta)\,\mathrm{d}L_s\right) + \left(\sum_{t}^{t+\Delta} A_B(s,t+\Delta,T)\,\mathrm{d}s + \int_{t}^{t+\Delta} \Sigma_B(s,t+\Delta,T)\,\mathrm{d}L_s\right) + \left(\sum_{t}^{t+\Delta} X_B(s,t+\Delta,T)\,\mathrm{d}L_s\right) + \left(\sum_{t}^{t+\Delta} X_B(s,t+\Delta,$$

und

$$\mathrm{LR}_{S(\cdot)}(t,t+\Delta) = \int_{t}^{t+\Delta} \tilde{\alpha}_{S}(s,t+\Delta) \,\mathrm{d}s + \int_{t}^{t+\Delta} \tilde{\sigma}_{S}(s,t+\Delta) \,\mathrm{d}L_{s}.$$

Ferner ist jede der Zufallsvariablen $\mathrm{R}^{\mathrm{mod}}_{B(\cdot,T)}(t,t+\Delta)$, $\mathrm{R}^{\mathrm{mod}}_{S(\cdot)}(t,t+\Delta)$, $\mathrm{LR}_{B(\cdot,T)}(t,t+\Delta)$ und $\mathrm{LR}_{S(\cdot)}(t,t+\Delta)$ stochastisch unabhängig von \mathcal{F}_t und messbar bezüglich $\mathcal{F}_{t+\Delta}$. **Beweis:** Die behaupteten Gleichungen lassen sich unter Verwendung von Lemma 5 leicht nachrechnen.

Nach Glau (2010, Lemma I.6) ist für jede Funktion $f : [0, T^*] \to \mathbb{R}^D$, die Borel-messbar und beschränkt ist, der Prozess $(\int_0^t f(s) \, \mathrm{d}L_s)_{t \in [0,T^*]}$ ein zeitinhomogener Lévy-Prozess. Hieraus folgt die behauptete stochastische Unabhängigkeit von \mathcal{F}_t und die Messbarkeit bezüglich $\mathcal{F}_{t+\Delta}$.

Korollar 8. Seien $N, K \in \mathbb{N}, t_1, \ldots, t_K \in [0, T^*], \Delta_1, \ldots, \Delta_K \in \mathbb{R}_+$ mit $t_1 \leq t_1 + \Delta_1 \leq t_2 \leq t_2 + \Delta_2 \leq \ldots \leq t_K \leq t_K + \Delta_K \leq T^*$ und $T_1, \ldots, T_{N-1} \in [\max_{k \in \{1,\ldots,K\}} \Delta_k, T^*]$ mit $t_K + T_n \leq T^*$ für alle $n \in \{1, \ldots, N-1\}$. Weiterhin sei für $k \in \{1, \ldots, K\}$ der N-dimensionale Zufallsvektor X_k definiert durch

$$X_k^n := \operatorname{LR}_{B(\cdot, t_k + T_n)} (t_k, t_k + \Delta_k)$$

für $n \in \{1, ..., N-1\}$ und

$$X_k^N := \operatorname{LR}_{S(\cdot)}(t_k, t_k + \Delta_k)$$

- 1. Dann sind Zufallsvektoren X_1, \ldots, X_K stochastisch unabhängig.
- 2. Ferner sei L ein Lévy-Prozess, es gelte $\Delta_1 = \ldots = \Delta_K =: \Delta$ und die Driftstrukturen α_B und α_S sowie die Volatilitätsstrukturen σ_B und σ_S seien stationär, das heißt für alle $s, T \in [0, T^*]$ mit $s \leq T$ gelte $\alpha_B(s,T) = \alpha_B(0,T-s), \alpha_S(s) = \alpha_S(0), \sigma_B(s,T) = \sigma_B(0,T-s)$ und $\sigma_S(s) = \sigma_S(0).$

Dann sind X_1, \ldots, X_K auch identisch verteilt.

Für den Beweis der Teilaussage 1 von Korollar 8 wird der nun vorgestellte Satz 9 benötigt, der aus Polley (2010, Satz 32) entnommen ist und hier ohne Beweis angegeben wird.

Satz 9. Seien $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, \mathcal{A} eine Unter- σ -Algebra von \mathcal{F} , $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor. Ferner sei für alle $v \in \mathbb{R}^N$ die Zufallsvariable $\langle v, X \rangle$ stochastisch unabhängig von \mathcal{A} .

Dann ist auch X stochastisch unabhängig von \mathcal{A} .

Beweis von Teilaussage 1 von Korollar 8: Der Beweis wird für den Spezialfall $\alpha_B = 0$, $\alpha_S = 0$ geführt. Der allgemeine Fall für beliebige Driftstrukturen, die deterministisch und stetig sind, folgt dann direkt daraus, weil

bei einer Verschiebung um eine Konstante die stochastische Unabhängigkeit erhalten bleibt.

Sei $k \in \{1, \ldots, K\}$. Trivialerweise ist X_k messbar bezüglich $\mathcal{F}_{t_k+\Delta_k}$. Für $v = (v^1, \ldots, v^N) \in \mathbb{R}^N$ definiere $g_v : [0, T^*] \to \mathbb{R}^D$ vermöge

$$g_v(s) := \left[\sum_{n=1}^{N-1} v^n \Sigma_B(s, t_k + \Delta_k, T_n)\right] + v^N \tilde{\sigma}_S(s, t_k + \Delta_k).$$

Dann gilt $\langle v, X_k \rangle = \int_{t_k}^{t_k + \Delta} g_v(s) dL_s$ und Glau (2010, Lemma I.6) impliziert, dass die Zufallsvariable $\langle v, X_k \rangle$ stochastisch unabhängig von \mathcal{F}_{t_k} ist. Weil $v \in \mathbb{R}^N$ beliebig war, folgt aus Satz 9, dass der Zufallsvektor X_k stochastisch unabhängig von \mathcal{F}_{t_k} ist. Die stochastische Unabhängigkeit von X_1, \ldots, X_K folgt nun mit vollständiger Induktion aus Bauer (2002, Korollar 7.3).

Teilaussage 2 von Korollar 8 ergibt sich als eine direkte Folgerung aus Korollar 37, das in Abschnitt 3.2 vorgestellt wird. Der Beweis wird dort nachgeholt.

Bemerkung 10. Der treibende Prozess L sei ein Lévy-Prozess und es sei ein beliebiger der Zufallsvektoren X_1, \ldots, X_K aus Korollar 8 betrachtet. Dieser wird hier mit X bezeichnet. In jedem Eintrag von X steht ein Log-Return. Ist man lediglich am Erwartungswertvektor $\mathbb{E}[X]$ und an der Kovarianzmatrix Cov(X) interessiert, so bietet ein treibender Lévy-Prozess L gegenüber einem treibenden Wiener-Prozess keinen echten Vorteil. Verwendet man nämlich statt L einen D-dimensionalen Wiener-Prozess $B = (B_t)_{t \in [0,T^*]}$ mit $\mathbb{E}[B_{T^*}] = \mathbb{E}[L_{T^*}]$ und $\text{Cov}(B_{T^*}) = \text{Cov}(L_{T^*})$ als Treiber und behält alle anderen Modellparameter bei, so führt dies zum identischen Erwartungswertvektor $\mathbb{E}[X]$ und zur identischen Kovarianzmatrix Cov(X).

Die Verwendung eines treibenden Lévy-Prozesses L bietet gegenüber einem Wiener-Prozess B erst dann einen echten Vorteil, wenn man auch an Momenten dritter und höherer Ordnungen interessiert ist. Dies ist insbesondere dann der Fall, wenn man sich für die mehrdimensionale Wahrscheinlichkeitsdichte (sofern eine solche existiert) der gemeinsamen Verteilung der Log-Returns, aus denen X besteht, interessiert.

2.3 Gemeinsame Momente von relativen Zuwächsen

Im vorliegenden Abschnitt 2.3 wird eine Formel zur Berechnung von gemeinsamen Momenten von modifizierten relativen Zuwächsen von Nullcoupon-Anleihen und der Aktie bereitgestellt. Als Korollare ergeben sich Formeln für deren Erwartungswerte, Varianzen und Korrelationen. Die Ergebnisse sind in kanonischer Weise auf die entsprechenden (nicht modifizierten) relativen Zuwächse übertragbar.

Für ein gegebenes Portfolio, das aus Nullcoupon-Anleihen und der Aktie besteht, können auf dieser Grundlage auch ohne eine zeitintensive Monte-Carlo-Simulation die Momente der Verteilung des dazugehörigen Portfoliowertes zu einem zukünftigen Zeitpunkt berechnet werden.

Satz 11. Seien $t, T_1, T_2 \in [0, T^*]$ und $\Delta \in \mathbb{R}_+$ mit $t \leq t + \Delta \leq T_1 \leq T_2$. Ferner seien $k, l, m \in \mathbb{R}_+$ und es existiere ein $\varepsilon \in (0, 1)$, sodass

$$k\Sigma_B(s,t+\Delta,T_1) + l\Sigma_B(s,t+\Delta,T_2) + m\tilde{\sigma}_S(s,t+\Delta) \in [-(1-\varepsilon)M, (1-\varepsilon)M]^D$$

für alle $s \in [t, t+\Delta]$ gilt. Dabei ist M die Konstante aus der Annahme ($\mathbb{E}\mathbb{M}$).

Dann gilt

$$\mathbb{E}\left[\left(\mathrm{R}_{B(\cdot,T_{1})}^{\mathrm{mod}}(t,t+\Delta)\right)^{k}\left(\mathrm{R}_{B(\cdot,T_{2})}^{\mathrm{mod}}(t,t+\Delta)\right)^{l}\left(\mathrm{R}_{S(\cdot)}^{\mathrm{mod}}(t,t+\Delta)\right)^{m}\right]$$

$$=\exp\left(\int_{t}^{t+\Delta}\left[-kA_{B}(s,t+\Delta,T_{1})-lA_{B}(s,t+\Delta,T_{2})+m\tilde{\alpha}_{S}(s,t+\Delta)\right.\right.\right.$$

$$\left.+\theta_{s}\left(k\Sigma_{B}(s,t+\Delta,T_{1})+l\Sigma_{B}(s,t+\Delta,T_{2})+m\tilde{\sigma}_{S}(s,t+\Delta)\right)\right]\mathrm{d}s\right).$$

Beweis: Die Behauptung folgt aus Lemma 7 und Satz 2.

Die nun folgenden Korollare 12 und 13 sind direkte Schlussfolgerungen aus Satz 11.

Korollar 12. Scient, $T \in [0, T^*]$ und $\Delta \in \mathbb{R}_+$ mit $t \leq t + \Delta \leq T$. Ferner existiere ein $\varepsilon \in (0, 1)$, sodass $2\Sigma_B(s, t + \Delta, T) \in [-(1 - \varepsilon)M, (1 - \varepsilon)M]^D$ und $2\tilde{\sigma}_S(s) \in [-(1 - \varepsilon)M, (1 - \varepsilon)M]^D$ für alle $s \in [t, t + \Delta]$. Dann gilt

$$\mathbb{E}\left[\mathrm{R}_{B(\cdot,T)}^{\mathrm{mod}}(t,t+\Delta)\right] = \exp\left(\int_{t}^{t+\Delta} \left[-A_{B}(s,t+\Delta,T) + \theta_{s}\left(\Sigma_{B}(s,t+\Delta,T)\right)\right] \mathrm{d}s\right),$$
$$+ \theta_{s}\left(\Sigma_{B}(s,t+\Delta,T)\right) \mathrm{d}s\right),$$
$$\mathbb{E}\left[\mathrm{R}_{S(\cdot)}^{\mathrm{mod}}(t,t+\Delta)\right] = \exp\left(\int_{t}^{t+\Delta} \tilde{\alpha}_{S}(s,t+\Delta) + \int_{t}^{t+\Delta} \theta_{s}\left(\tilde{\sigma}_{S}(s,t+\Delta)\right) \mathrm{d}s\right),$$

sowie

$$\operatorname{Var}\left(\operatorname{R}_{B(\cdot,T)}^{\operatorname{mod}}(t,t+\Delta)\right) = \exp\left(-2\int_{t}^{t+\Delta} A_{B}(s,t+\Delta,T) \,\mathrm{d}s\right)$$
$$\times \left[\exp\left(\int_{t}^{t+\Delta} \theta_{s}\left(2\Sigma_{B}(s,t+\Delta,T)\right) \,\mathrm{d}s\right)\right]$$
$$-\exp\left(\int_{t}^{t+\Delta} 2\theta_{s}\left(\Sigma_{B}(s,t+\Delta,T)\right) \,\mathrm{d}s\right)\right]$$

und

$$\operatorname{Var}\left(\operatorname{R}_{S(\cdot)}^{\operatorname{mod}}(t,t+\Delta)\right) = \exp\left(2\int_{t}^{t+\Delta} \tilde{\alpha}_{S}(s,t+\Delta) \,\mathrm{d}s\right)$$
$$\times \left[\exp\left(\int_{t}^{t+\Delta} \theta_{s}\left(2\tilde{\sigma}_{S}(s,t+\Delta)\right) \,\mathrm{d}s\right) - \exp\left(\int_{t}^{t+\Delta} 2\theta_{s}\left(\tilde{\sigma}_{S}(s,t+\Delta)\right) \,\mathrm{d}s\right)\right].$$

Beweis: Die Behauptung folgt aus Satz 11 für die Spezialfälle (r, l, m) = (1, 0, 0), (r, l, m) = (0, 0, 1), (r, l, m) = (2, 0, 0) und (r, l, m) = (0, 0, 2).

Korollar 13. Die Voraussetzungen von Korollar 12 seien erfüllt. Ferner seien $T_1, T_2 \in [0, T^*]$ mit $t + \Delta \leq T_1 \leq T_2$ und die Varianzen der modifizierten relativen Zuwächse $\mathbb{R}^{\text{mod}}_{B(\cdot,T)}(t, t + \Delta)$, $\mathbb{R}^{\text{mod}}_{B(\cdot,T_1)}(t, t + \Delta)$, $\mathbb{R}^{\text{mod}}_{B(\cdot,T_2)}(t, t + \Delta)$ und $\mathbb{R}^{\text{mod}}_{S(\cdot)}(t, t + \Delta)$ seien strikt positiv. Dann gilt

$$\operatorname{Korr}\left(\operatorname{R}_{B(\cdot,T_{1})}^{\operatorname{mod}}(t,t+\Delta), \operatorname{R}_{B(\cdot,T_{2})}^{\operatorname{mod}}(t,t+\Delta)\right) = \frac{g_{1}(t,t+\Delta,T_{1},T_{2}) - g_{2}(t,t+\Delta,T_{1},T_{2})}{\sqrt{h_{B}(t,t+\Delta,T_{1})}\sqrt{h_{B}(t,t+\Delta,T_{2})}}$$

und

$$\operatorname{Korr}\left(\operatorname{R}_{B(\cdot,T)}^{\operatorname{mod}}(t,t+\Delta), \operatorname{R}_{S(\cdot)}^{\operatorname{mod}}(t,t+\Delta)\right) = \frac{g_3(t,t+\Delta,T) - g_4(t,t+\Delta,T)}{\sqrt{h_B(t,t+\Delta,T)}\sqrt{h_S(t,t+\Delta)}},$$

wobei

$$g_{1}(t, t + \Delta, T_{1}, T_{2}) = \exp\left(\int_{t}^{t+\Delta} \theta_{s} \left(\Sigma_{B}(s, t + \Delta, T_{1}) + \Sigma_{B}(s, t + \Delta, T_{2})\right) ds\right),$$

$$g_{2}(t, t + \Delta, T_{1}, T_{2}) = \exp\left(\int_{t}^{t+\Delta} \left[\theta_{s} \left(\Sigma_{B}(s, t + \Delta, T_{1})\right) + \theta_{s} \left(\Sigma_{B}(s, t + \Delta, T_{2})\right)\right] ds\right),$$

$$g_{3}(t, t + \Delta, T) = \exp\left(\int_{t}^{t+\Delta} \theta_{s} \left(\Sigma_{B}(s, t + \Delta, T) + \tilde{\sigma}_{S}(s, t + \Delta)\right) ds\right),$$

$$g_{4}(t, t + \Delta, T) = \exp\left(\int_{t}^{t+\Delta} \left[\theta_{s} \left(\Sigma_{B}(s, t + \Delta, T) + \tilde{\sigma}_{S}(s, t + \Delta)\right)\right] ds\right),$$

$$+ \theta_{s} \left(\tilde{\sigma}_{S}(s, t + \Delta)\right)\right] ds\right)$$

sowie

$$h_B(t, t + \Delta, T) = \exp\left(\int_t^{t+\Delta} \theta_s \left(2\Sigma_B(s, t + \Delta, T)\right) \, \mathrm{d}s\right) - \exp\left(\int_t^{t+\Delta} 2\theta_s \left(\Sigma_B(s, t + \Delta, T)\right) \, \mathrm{d}s\right)$$

und

$$h_{S}(t, t + \Delta) = \exp\left(\int_{t}^{t+\Delta} \theta_{s} \left(2\tilde{\sigma}_{S}(s, t + \Delta)\right) \, \mathrm{d}s\right) - \exp\left(\int_{t}^{t+\Delta} 2\theta_{s} \left(\tilde{\sigma}_{S}(s, t + \Delta)\right) \, \mathrm{d}s\right).$$

Beweis: Eine Anwendung von Satz 11 für den Spezialfall (k,l,m) = (1,0,1)liefert

$$\mathbb{E}\left[\mathrm{R}_{B(\cdot,T)}^{\mathrm{mod}}(t,t+\Delta)\,\mathrm{R}_{S(\cdot)}^{\mathrm{mod}}(t,t+\Delta)\right]$$

= $g_3(t,t+\Delta,T)\exp\left(\int_t^{t+\Delta} \left[-A_B(s,t+\Delta,T)+\tilde{\alpha}_S(s,t+\Delta)\right]\,\mathrm{d}s\right).$

Ferner erhält man unter Verwendung von Korollar $12\,$

$$\mathbb{E}\left[\mathrm{R}_{B(\cdot,T)}^{\mathrm{mod}}(t,t+\Delta)\right] \mathbb{E}\left[\mathrm{R}_{S(\cdot)}^{\mathrm{mod}}(t,t+\Delta)\right] \\ = g_4(t,t+\Delta,T) \exp\left(\int_t^{t+\Delta} \left[-A_B(s,t+\Delta,T) + \tilde{\alpha}_S(s,t+\Delta)\right] \mathrm{d}s\right),$$

$$\operatorname{Var}\left(\mathrm{R}_{B(\cdot,T)}^{\mathrm{mod}}(t,t+\Delta)\right) = h_B(t,t+\Delta,T) \exp\left(-2\int_t^{t+\Delta} A_B(s,t+\Delta,T) \,\mathrm{d}s\right)$$

und

$$\operatorname{Var}\left(\operatorname{R}^{\operatorname{mod}}_{S(\cdot)}(t,t+\Delta)\right) = h_{S}(t,t+\Delta) \exp\left(2\int_{t}^{t+\Delta} \tilde{\alpha}_{S}(s,t+\Delta) \,\mathrm{d}s\right).$$

Die zweite behauptete Gleichung folgt nun durch Einsetzen der gefundenen Ausdrücke in die Definitionsgleichung der Korrelation und Kürzen des Terms

$$\exp\left(\int_{t}^{t+\Delta} \left[-A_B(s,t+\Delta,T) + \tilde{\alpha}_S(s,t+\Delta)\right] \mathrm{d}s\right).$$

Der Nachweis für die Gültigkeit der anderen behaupteten Gleichung verläuft analog. $\hfill \Box$

Korollar 14. Es gelten die Voraussetzungen und die Bezeichnungen von Korollar 13.

Dann gilt

Korr
$$\left(B(t,T_1), B(t,T_2)\right) = \frac{g_1(0,t,T_1,T_2) - g_2(0,t,T_1,T_2)}{\sqrt{h_B(0,t,T_1)}\sqrt{h_B(0,t,T_2)}}$$

und

Korr
$$\left(B(t,T), S(t)\right) = \frac{g_3(0,t,T) - g_4(0,t,T)}{\sqrt{h_B(0,t,T)}\sqrt{h_S(0,t)}}$$

Beweis: B(0,t) und B(0,T) sind deterministisch. Ferner gilt $B(t,T) = \frac{B(0,T)}{B(0,t)} R_{B(\cdot,T)}^{\text{mod}}(0,t)$ sowie $S(t) = \frac{1}{B(0,t)} R_{S(\cdot)}^{\text{mod}}(0,t)$ und damit

$$\operatorname{Korr}\left(B\left(t,T\right),S\left(t\right)\right) = \operatorname{Korr}\left(\operatorname{R}_{B(\cdot,T)}^{\operatorname{mod}}\left(0,t\right),\operatorname{R}_{S(\cdot)}^{\operatorname{mod}}\left(0,t\right)\right).$$

Hieraus und aus Korollar 13 folgt der zweite Teil der Behauptung. Der erste Teil folgt analog. $\hfill \Box$

Bemerkung 15. Korrelationen in Lévy-Zinsstrukturmodellen wurden bereits von Beinhofer, Eberlein, Janssen und Polley (2011) untersucht. Erwartungsgemäß fällt die Formel, die die Autoren für die Korrelation der Preise von Nullcoupon-Anleihen im Lévy-Forward-Rate-Modell gefunden haben, mit der ersten Gleichung aus Korollar 14 zusammen.

Bemerkung 16. Im Gegensatz zu dem in Bemerkung 10 betrachteten Fall von Log-Returns bietet bereits bei der Untersuchung von (gemeinsamen) Momenten erster und zweiter Ordnung von (modifizierten) relativen Zuwächsen von Nullcoupon-Anleihen und der Aktie sowie von deren Preisen die Verwendung eines Lévy-Prozesses als Treiber gegenüber einem Wiener-Prozess einen echten Mehrwert. Für den Spezialfall von Korrelationen von Preisen von Nullcoupon-Anleihen wurde diese Tatsache in der bereits zuvor erwähnten Arbeit von Beinhofer, Eberlein, Janssen und Polley (2011) bei einer Kalibrierung auf der Grundlage von empirischen Korrelationen demonstriert.

2.4 Der risikoneutrale Fall

Im HJM-Modell bzw. im Lévy-Forward-Rate-Modell kann durch eine geeignete Wahl der Driftstruktur erreicht werden, dass das Modell direkt unter einem risikoneutralen Maß arbeitet. Ähnlich verhält es sich im betrachteten Lévy-Hybridmodell, das sich vom Lévy-Forward-Rate-Modell lediglich dadurch unterscheidet, dass es um eine Aktie erweitert ist. Im folgenden Satz werden Anforderungen an die Driftstrukturen formuliert, die hinreichend sind, damit sowohl die diskontierten Preisprozesse der Anheihen als auch der diskontierte Preisprozess der Aktie Martingale sind. Wie zu erwarten, fällt dabei die Bedingung an die Driftstruktur der Zinsstruktur mit der wohlbekannten Heath-Jarrow-Morton-Driftbedingung zusammen.

Satz 17. Es existiere ein $\varepsilon \in (0, 1)$, sodass die aus der Volatilitätsstruktur der Zinsstruktur berechnete Größe Σ_B und die Volatilitätsstruktur der Aktie σ_S ausschließlich Werte im Intervall $[-(1-\varepsilon)M, (1-\varepsilon)M]^D$ annehmen. Dabei ist M die Konstante aus der Annahme (EM). Ferner seien für alle $s, T \in [0, T^*]$ mit $s \leq T$ die Bedingungen

$$A_B(s,T) = \theta_s \left(\Sigma_B(s,T) \right)$$

und

$$\alpha_S(s) = -\theta_s\left(\sigma_S(s)\right)$$

erfüllt.

Dann sind für jeden Fälligkeitszeitpunkt $T \in [0, T^*]$ der diskontierte Preisprozess der dazugehörigen Nullcoupon-Anleihe $(B(t, T) / B_t)_{t \in [0,T]}$ und der diskontierte Preisprozess der Aktie $(S(t) / B_t)_{t \in [0,T^*]}$ Martingale. Insbesondere ist das zugrunde gelegte Wahrscheinlichkeitsmaß \mathbb{P} ein Martingalmaß.

Beweis: Nach Kluge (2005, Proposition 2.2) ist $(B(t,T)/B_t)_{t\in[0,T]}$ ein Martingal.

Seien $t_1, t_2 \in [0, T^*]$ mit $t_1 \leq t_2$. Glau (2010, Lemma I.6) impliziert, dass $\int_0^{t_1} \sigma_S(s) dL_s$ messbar bezüglich \mathcal{F}_{t_1} und $\int_{t_1}^{t_2} \sigma_S(s) dL_s$ stochastisch unabhängig von \mathcal{F}_{t_1} ist. Damit erhält man unter Verwendung von Satz 2 und

der Bedingung an die Driftstruktur des Aktienindex

$$\mathbb{E}\left[\frac{S(t_2)}{B_{t_2}}\middle|\mathcal{F}_{t_1}\right] = \exp\left(\int_0^{t_1} \alpha_S(s) \,\mathrm{d}s + \int_0^{t_1} \sigma_S(s) \,\mathrm{d}L_s\right) \\ \times \exp\left(\int_{t_1}^{t_2} \alpha_S(s) \,\mathrm{d}s\right) \mathbb{E}\left[\exp\left(\int_{t_1}^{t_2} \sigma_S(s) \,\mathrm{d}L_s\right)\right] \\ = \exp\left(\int_0^{t_1} \alpha_S(s) \,\mathrm{d}s + \int_0^{t_1} \sigma_S(s) \,\mathrm{d}L_s\right) \\ \times \exp\left(\int_{t_1}^{t_2} \alpha_S(s) \,\mathrm{d}s\right) \exp\left(\int_{t_1}^{t_2} \theta_s\left(\sigma_S(s)\right) \,\mathrm{d}s\right) \\ = \frac{S(t_1)}{B_{t_1}},$$

also die Behauptung.

3 Kumulanten von stochastischen Integralen nach einem Lévy-Prozess und die multivariate Edgeworth-Approximation

Für die angestrebte Kalibrierung des in Abschnitt 2 vorgestellten Modells ist es erforderlich, die Wahrscheinlichkeitsdichte und die bedingte Verteilungsfunktion eines Zufallsvektors in sehr kurzer Zeit numerisch auswerten zu können. Bei den verwendeten Kalibrierungskriterien werden dabei Zufallsvektoren betrachtet, in deren Einträgen stochastische Integrale nach einem Lévy-Prozess stehen. Die sonst übliche Methode der Fourier-Inversion der charakteristischen Funktionen der Zufallsvektoren scheidet aus, weil hierfür Mehrfachintegrale zu berechnen sind, die für die verwendeten Verteilungen numerisch ausgewertet werden müssen, und somit selbst unter Einsatz von modernen Hochleistungsrechnern noch zu viel Zeit in Anspruch nehmen würden.

Ziel des vorliegenden Abschnitts ist die Einführung der multivariaten Edgeworth-Approximation. Sie basiert auf den Kumulanten eines Zufallsvektors und ermöglicht eine sehr zeiteffiziente numerische Auswertung der (bedingten) Wahrscheinlichkeitsdichte und der (bedingten) Verteilungsfunktion. Weiterhin wird mit Satz 35 eine Formel zur Berechnung der Kumulanten von Zufallsvektoren des betrachteten Typs bereitgestellt. Diese Konzepte bilden den Ausgangspunkt für die Implementierung eines Computerprogramms, das weiter unten in Abschnitt 5 vorgestellt und in Abschnitt 6 zur Modellkalibrierung eingesetzt wird.

Für alle weiteren Untersuchungen bis zum Ende der vorliegenden Arbeit sei stets $T^* \in [1, \infty)$. Aus mathematischer Sicht ist auch die Annahme $T^* \in (0, \infty)$ hinreichend für alle Überlegungen. Die Annahme $T^* \ge 1$ dient lediglich zur Vereinfachung der Notation, weil so $1 \in [0, T^*]$ gilt und damit der Prozess $L = (L_t)_{t \in [0, T^*]}$ zum Zeitpunkt 1, also L_1 , definiert ist.

3.1 Die Annahme (EM) und Kumulanten von Zufallsvektoren

In Abschnitt 2.1 wurde in Definition 1 die Annahme ($\mathbb{E}\mathbb{M}$) für zeitinhomogene Lévy-Prozesse eingeführt. Im vorliegenden Abschnitt wird die Annahme ($\mathbb{E}\mathbb{M}$) für Zufallsvektoren formuliert und der Zusammenhang zwischen den beiden Definitionen beleuchtet.

Die Klasse der Zufallsvektoren, die der Annahme (EM) genügen, hat schöne Eigenschaften. Dabei spielen die momenterzeugende Funktion und die kumulantenerzeugende Funktion eine zentrale Rolle. Die Sätze über die momenterzeugende Funktion ähneln in ihren Aussagen entsprechenden Resultaten über die charakteristische Funktion. Die Beweise sind in einigen Fällen jedoch deutlich anspruchsvoller.

Um die vorliegende Arbeit bestmöglich in sich abgeschlossen zu gestalten, werden neben neuen Erkenntnissen unter anderem auch einige bereits bekannte Resultate über Kumulanten wiederholt.

Die wichtigsten Resultate des vorliegenden Abschnitts sind der Satz 28, das Lemma 29 sowie die Sätze 32 und 33. Während mit Lemma 29 und den Sätzen 28 und 32 wirkungsvolle Rechenregeln für die kumulantenerzeugende Funktion und für Kumulanten bereitgestellt werden, macht Satz 33 eine Aussage darüber, wann der Grenzwert einer konvergenten Folge von Zufallsvektoren, die der Annahme (EM) genügen, ebenfalls der Annahme (EM) genügt.

Definition 18 (Annahme ($\mathbb{E}\mathbb{M}$) für einen Zufallsvektor). Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor. X genügt der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$, falls

$$\mathbb{E}\left[\exp\left(\langle u, X \rangle\right)\right] < \infty$$

für alle $u \in [-M, M]^N$ gilt.

Satz 19. Seien $D \in \mathbb{N}$, $L = (L_t)_{t \in [0,T^*]}$ ein D-dimensionaler zeitinhomogener Lévy-Prozess und $M \in (0, \infty)$.

- 1. L genügt der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante M genau dann, wenn für jedes $t \in [0, T^*]$ der Zufallsvektor L_t der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante M genügt.
- 2. Falls L der Annahme (\mathbb{EM}) zur Konstante M genügt, so genügt für alle $t_1, t_2 \in [0, T^*]$ mit $t_1 < t_2$ der dazugehörige Zuwachs $L_{t_2} - L_{t_1}$ der Annahme (\mathbb{EM}) zur Konstante M.

Beweis: Teilaussage 1 ist aus Kluge (2005, Lemma 1.6) übernommen.

Aus der stochastischen Unabhängigkeit der Zuwächse von L folgt für $t_1, t_2 \in [0, T^*]$ mit $t_1 < t_2$ und $u \in [-M, M]^D$

$$\mathbb{E}\left[\exp\left(\langle u, L_{t_2}\rangle\right)\right] = \mathbb{E}\left[\exp\left(\langle u, L_{t_2} - L_{t_1}\rangle\right)\right] \mathbb{E}\left[\exp\left(\langle u, L_{t_1}\rangle\right)\right].$$

Teilaussage 1 des vorliegenden Satzes impliziert $\mathbb{E}\left[\exp\left(\langle u, L_{t_2}\rangle\right)\right] < \infty$ und $\mathbb{E}\left[\exp\left(\langle u, L_{t_1}\rangle\right)\right] < \infty$. Hieraus folgt Teilaussage 2.

Satz und Definition 20 (Momenterzeugende Funktion und kumulantenerzeugende Funktion). Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}M$) zur Konstante $M \in (0, \infty)$ genügt.

Dann ist die Funktion $M_X : \{ \tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N \} \to \mathbb{C}$ mit

$$M_X(u) := \mathbb{E}\left[\exp(\langle u, X \rangle)\right]$$

holomorph. Ferner existiert eine eindeutig bestimmte holomorphe Funktion $K_X: \{ \tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N \} \rightarrow \mathbb{C} \text{ mit } K_X(0) = 0 \text{ und}$

$$M_X(u) = \exp\left(K_X(u)\right)$$

 $\label{eq:full_full} \textit{für alle } u \in \left\{ \tilde{u} \in \mathbb{C}^N \big| \Re(\tilde{u}) \in (-M,M)^N \right\}.$

Für $u \in (-M, M)^N$ stimmt M_X mit der momenterzeugenden Funktion von X überein und K_X stimmt mit der kumulantenerzeugenden Funktion von X überein. Aus diesem Grund werden in der vorliegenden Arbeit M_X und K_X ebenfalls als momenterzeugende Funktion bzw. als kumulantenerzeugende Funktion von X bezeichnet.

Beweis: Sei $u = (u^1, \ldots, u^N) \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\}$. Der Nachweis für $\mathbb{E}[|\exp(\langle u, X \rangle)|] < \infty$ ist trivial.

Sei nun zusätzlich $n \in \{1, \ldots, N\}$. Wähle $p_1 \in (1, \infty)$ mit $p_1 \Re(u) \in (-M, M)^N$ und definiere $q_1 := \frac{p_1}{p_1 - 1} \in (1, \infty)$. Wähle weiter $q_2 \in (q_1, \infty)$ und $\delta \in (0, \infty)$, sodass die Bedingung $q_1 q_2 \delta < M$ erfüllt ist. Setze schließlich noch $p_2 := \frac{q_2}{q_2 - 1} \in (1, \infty)$.

Dann gilt $\frac{1}{p_1} + \frac{1}{q_1} = 1$ und $\frac{1}{p_2} + \frac{1}{q_2} = 1$ sowie $\mathbb{E}\left[\exp\left(\langle p_1 \Re(u), X \rangle\right)\right] < \infty$ und $\mathbb{E}\left[\exp\left(q_1 q_2 \delta \left| X^n \right|\right)\right] < \infty$. Weiterhin lässt sich unter Verwendung von $\lim_{|x|\to\infty} |x|^{p_2q_1} / \exp\left(\frac{M}{2}|x|\right) = 0$ leicht zeigen, dass auch $\mathbb{E}\left[|X^n|^{p_2q_1}\right] < \infty$ gilt. Durch zweifache Anwendung der Hölderschen Ungleichung erhält man

$$\mathbb{E}\left[\left|\exp\left(\langle u, X \rangle\right) \sum_{k=1}^{\infty} \frac{h^{k-1}(X^{n})^{k}}{k!}\right|\right] \\
\leq \mathbb{E}\left[\exp\left(\langle \Re(u), X \rangle\right) |X^{n}| \exp\left(|hX^{n}|\right)\right] \\
\leq \mathbb{E}\left[\exp\left(\langle \Re(u), X \rangle\right) |X^{n}| \exp\left(\delta |X^{n}|\right)\right] \\
\leq \left(\mathbb{E}\left[\left(\exp\left(\langle \Re(u), X \rangle\right)\right)^{p_{1}}\right]\right)^{\frac{1}{p_{1}}} \left(\mathbb{E}\left[\left(|X^{n}| \exp\left(\delta |X^{n}|\right)\right)^{q_{1}}\right]\right)^{\frac{1}{q_{1}}} \\
\leq \left(\mathbb{E}\left[\left(\exp\left(\langle \Re(u), X \rangle\right)\right)^{p_{1}}\right]\right)^{\frac{1}{p_{1}}} \\
\times \left(\mathbb{E}\left[|X^{n}|^{p_{2}q_{1}}\right]\right)^{\frac{1}{p_{2}q_{1}}} \left(\mathbb{E}\left[\left(\exp\left(\delta |X^{n}|\right)\right)^{q_{1}q_{2}}\right]\right)^{\frac{1}{q_{1}q_{2}}} \\
= \left(\mathbb{E}\left[\exp\left(\langle p_{1}\Re(u), X \rangle\right)\right]\right)^{\frac{1}{p_{1}}} \\
\times \left(\mathbb{E}\left[|X^{n}|^{p_{2}q_{1}}\right]\right)^{\frac{1}{p_{2}q_{1}}} \left(\mathbb{E}\left[\exp\left(q_{1}q_{2}\delta |X^{n}|\right)\right]\right)^{\frac{1}{q_{1}q_{2}}} \\
< \infty$$

für alle $h \in \mathbb{C}$ mit $|h| < \delta$. Für den Spezialfall h = 0 folgt hieraus insbesondere $\mathbb{E}[X^n \exp(\langle u, X \rangle)] \in \mathbb{C}$. Weiter berechnet man mit dem Satz von der majorisierten Konvergenz

$$\lim_{\substack{h \to 0, \\ h \in \mathbb{C}}} \frac{\mathbb{E}\left[\exp\left(\left\langle (u^{1}, \dots, u^{n-1}, u^{n} + h, u^{n+1}, \dots, u^{N}), X\right\rangle\right)\right] - \mathbb{E}\left[\exp\left(\left\langle u, X\right\rangle\right)\right]}{h}$$
$$= \lim_{\substack{h \to 0, \\ h \in \mathbb{C}}} \mathbb{E}\left[\exp\left(\left\langle u, X\right\rangle\right) \sum_{k=1}^{\infty} \frac{h^{k-1}(X^{n})^{k}}{k!}\right]$$
$$= \mathbb{E}\left[X^{n} \exp\left(\left\langle u, X\right\rangle\right)\right].$$

Also ist der Schnitt

$$(-M, M) \times i\mathbb{R} \ni z \mapsto \mathbb{E}\left[\exp\left(\left\langle (u^1, \dots, u^{n-1}, z, u^{n+1}, \dots, u^N), X\right\rangle\right)\right]$$

holomorph. Der Satz von Hartogs (1906, § 4. Fall von n Veränderlichen) impliziert nun, dass M_X holomorph ist. Damit ist die erste Behauptung bewiesen.

Zur Definition der Funktion $K_X : \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\} \to \mathbb{C}$ wähle $\delta \in (0, M)$, sodass $M_X(B_\delta(0)) \subset (0, \infty) \times i(-\pi, \pi)$ gilt. Dabei bezeichnet $B_\delta(0) := \{\tilde{u} \in \mathbb{C}^N | |\tilde{u}| < \delta\}$ den offenen Ball um den Ursprung mit Radius δ . Definiere nun $f : B_\delta(0) \to \mathbb{C}$ mit $f(u) := \log(M_X(u))$, wobei hier log der Hauptwert des komplexen Logarithmus ist. Dann ist f holomorph und es gilt f(0) = 0.

Sei $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\}$. Falls $u \in B_{\delta}(0)$, setze $K_X(u) := f(u)$. Falls $u \notin B_{\delta}(0)$, wähle einen stetigen Weg $\gamma : [0, 1] \to \mathbb{C}^N$ mit $\gamma(0) := 0$ und $\gamma(1) = u$. Dann ist $\Gamma : [0, 1] \to \mathbb{C} \setminus \{0\}, \Gamma(s) := (M_X \circ \gamma)(s)$, ebenfalls ein stetiger Weg und es gilt $\Gamma(0) = 1 \in \mathbb{C} \setminus (-\mathbb{R}_+)$. Wähle nun $K_X(u)$ als die analytische Fortsetzung des Hauptwertes des komplexen Logarithmus entlang Γ .

Sei $\tilde{\gamma} : [0,1] \to \mathbb{C}^N$ ein weiterer stetiger Weg mit $\tilde{\gamma}(0) := 0$ und $\tilde{\gamma}(1) = u$, $\tilde{\Gamma} : [0,1] \to \mathbb{C} \setminus \{0\}, \ \tilde{\Gamma}(s) := (M_X \circ \tilde{\gamma})(s)$, sowie $h : [0,1] \times [0,1] \to \mathbb{C} \setminus \{0\}$ mit $h_t(s) := M_X((1-t)\gamma(s) + t\tilde{\gamma}(s))$. Dann ist $\tilde{\Gamma}$ ebenfalls ein stetiger Weg mit $\tilde{\Gamma}(0) = \Gamma(0)$ und $\tilde{\Gamma}(1) = \Gamma(1)$. Außerdem ist h eine Homotopie zwischen Γ und $\tilde{\Gamma}$ und der Hauptwert des komplexen Logarithmus lässt sich für jedes $t \in [0,1]$ entlang des stetigen Weges h_t analytisch fortsetzen. Also ist nach dem Monodromiesatz die Definition von $K_X(u)$ unabhängig von der Wahl von Γ und somit ist K_X wohldefiniert. Ferner ist klar, dass die so definierte Funktion K_X holomorph ist.

Sei nun $\tilde{K}_X : \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\} \to \mathbb{C}$ eine weitere holomorphe Funktion mit $\tilde{K}_X(0) = 0$ und $M_X(u) = \exp(\tilde{K}_X(u))$ für alle $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\}$. Dann stimmt die Einschränkung $\tilde{K}_X|_{B_{\delta}(0)}$ mit $f = K_X|_{B_{\delta}(0)}$ überein. Weil $B_{\delta}(0)$ eine nichtleere offene Menge ist, folgt aus dem Eindeutigkeitssatz für mehrdimensionale holomorphe Funktionen aus Scheidemann (2005, Conclusion 1.2.12.2), dass $\tilde{K}_X = K_X$ gilt. Damit ist auch die zweite Behauptung bewiesen.

Bemerkung 21. Die Definition von K_X als die Funktion

$$\left\{\tilde{u} \in \mathbb{C}^N \middle| \Re(\tilde{u}) \in (-M, M)^N \right\} \ni u \mapsto \log\left(M_X(u)\right)$$

mag vielleicht auf den ersten Blick naheliegender erscheinen. Diese Vorgehensweise hat jedoch den folgenden Nachteil. $M_X(u)$ kann für $\Im(u) \neq 0$ auch einen komplexen Wert annehmen und falls $\log(M_X(u))$ eindeutig sein soll, muss man sich auf einen Zweig des komplexen Logarithmus festlegen. Es gibt jedoch keinen Zweig, der auf ganz $\mathbb{C}\setminus\{0\}$ holomorph ist und es gibt Beispiele von Zufallsvariablen X, für die es keinen Zweig gibt, für den $\log \circ M_X$ holomorph ist.

Weil die Holomorphie von K_X jedoch für die weiteren Uberlegungen eine zentrale Rolle spielt, wird in der vorliegenden Arbeit der in Satz und Definition 20 präsentierte Weg zur Definition von K_X gewählt, der die gewünschte Eigenschaft der Holomorphie garantiert. **Definition 22** (Kumulante). Seien $N \in \mathbb{N}$, $X \text{ ein } \mathbb{R}^N$ -wertiger Zufallsvektor und $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ ein Multiindex.

Falls $\mathbb{E}[|X^{\alpha}|] < \infty$, so ist die Kumulante κ_{α}^{X} von X der Ordnung α definiert durch

$$\kappa_{\boldsymbol{\alpha}}^{X} := \left. \frac{1}{\mathbf{i}^{|\boldsymbol{\alpha}|}} \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial y^{\boldsymbol{\alpha}}} \log\left(\varphi_{X}(y)\right) \right|_{y=0},$$

wobei $\varphi_X : \mathbb{R}^N \to \mathbb{C}$ die charakteristische Funktion von X und log der Hauptwert des komplexen Logarithmus ist.

Satz 23. Set X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt.

Dann existieren für jeden Multiindex $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ das dazugehörige Moment $\mathbb{E}[X^{\boldsymbol{\alpha}}]$ und die dazugehörige Kumulante $\kappa_{\boldsymbol{\alpha}}^X$ in \mathbb{R} und es gilt

$$\kappa_{\alpha}^{X} = \left. \frac{\partial^{|\alpha|}}{\partial u^{\alpha}} K_{X}(u) \right|_{u=0}.$$

Beweis: Nach Satz 20 ist die Funktion M_X holomorph. Insbesondere ist sie damit auf ihrem gesamten Definitionsbereich beliebig oft stetig partiell differenzierbar. Hieraus und aus $M_X\left((-M, M)^N\right) \subset \mathbb{R}_+ \setminus \{0\}$ folgt $\mathbb{E}\left[X^{\alpha}\right] = \frac{\partial^{|\alpha|}}{\partial u^{\alpha}} M_X(u) \Big|_{u=0} \in \mathbb{R}.$

Es existiert eine Umgebung $U \subset \mathbb{R}^N$ von 0 mit $\log(\varphi_X(y)) = K_X(iy)$ für alle $y \in U$, wobei log wieder der Hauptwert des komplexen Logarithmus ist. Hieraus folgt zusammen mit der Kettenregel die behauptete Gleichung $\kappa_{\alpha}^X = \frac{\partial^{|\alpha|}}{\partial u^{\alpha}} K_X(u)|_{u=0}$. Der Nachweis für $\kappa_{\alpha}^X \in \mathbb{R}$ verläuft analog zum Nachweis der entsprechenden Aussage über das Moment $\mathbb{E}[X^{\alpha}]$.

Korollar 24 (Potenzreihenentwicklung der kumulantenerzeugenden Funktion). Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt.

Dann gilt

$$K_X(u) = \sum_{k=1}^{\infty} \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_0^N, \\ |\boldsymbol{\alpha}| = k}} \frac{\kappa_{\boldsymbol{\alpha}}^X}{\boldsymbol{\alpha}!} u^{\boldsymbol{\alpha}}$$

für alle $u \in \{ \tilde{u} = (\tilde{u}^1, \dots, \tilde{u}^N) \in \mathbb{C}^N | |\tilde{u}^n| < M, n \in \{1, \dots, N\} \}.$

Beweis: Die Aussage des Korollars folgt aus dem Satz über die Potenzreihenentwicklung für mehrdimensionale holomorphe Funktionen, wie er beispielsweise in Scheidemann (2005, Corollary 1.5.9 (Taylor expansion)) zu finden ist. $\hfill \Box$

Das nun folgende Lemma handelt davon, wie man aus den Kumulanten einer mehrdimensionalen Verteilung die Kumulanten ihrer Randverteilungen erhält. Die Kumulanten werden einfach an die Randverteilungen vererbt.

Lemma 25. Seien $N \in \mathbb{N} \setminus \{1\}$, $n \in \{1, \ldots, N-1\}$ und $X = (X^1, \ldots, X^N)$ ein \mathbb{R}^N -wertiger Zufallsvektor. Ferner sei $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^N) \in \mathbb{N}_0^N$ ein Multiindex mit $\alpha^m = 0$ für alle $m \in \{n + 1, \ldots, N\}$ und es gelte $\mathbb{E}[|X^{\boldsymbol{\alpha}}|] < \infty$.

Dann existiert die Kumulante $\kappa_{(\alpha^1,\ldots,\alpha^n)}^{(X^1,\ldots,X^n)}$ von (X^1,\ldots,X^n) der Ordnung $(\alpha^1,\ldots,\alpha^n)$ und es gilt $\kappa_{(\alpha^1,\ldots,\alpha^n)}^{(X^1,\ldots,X^n)} = \kappa_{\alpha}^X$.

Beweis: Unter Verwendung von Definition 22 erhält man

$$\begin{aligned} \kappa_{(\alpha^{1},\dots,X^{n})}^{(X^{1},\dots,X^{n})} &= \left. \frac{1}{\mathbf{i}^{|\boldsymbol{\alpha}|}} \left(\frac{\partial}{\partial y^{1}} \right)^{\alpha^{1}} \dots \left(\frac{\partial}{\partial y^{n}} \right)^{\alpha^{n}} \\ &\times \log\left(\mathbb{E}\left[\exp\left(\mathbf{i}\left\langle (y^{1},\dots,y^{n}), (X^{1},\dots,X^{n})\right\rangle \right) \right] \right) \Big|_{(y^{1},\dots,y^{n})=0} \\ &= \left. \frac{1}{\mathbf{i}^{|\boldsymbol{\alpha}|}} \left(\frac{\partial}{\partial y^{1}} \right)^{\alpha^{1}} \dots \left(\frac{\partial}{\partial y^{n}} \right)^{\alpha^{n}} \log\left(\mathbb{E}\left[\exp\left(\mathbf{i}\left\langle y,X\right\rangle \right) \right] \right) \right|_{(y^{1},\dots,y^{N})=0} \\ &= \left. \kappa_{\boldsymbol{\alpha}}^{X} \end{aligned}$$

und damit die Behauptung.

Beim Hamburger Momentproblem wird die Fragestellung untersucht, ob es zu einer gegebenen Folge $(m_k)_{k \in \mathbb{N}_0}$ in \mathbb{R} ein Borel-Maß μ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ gibt mit $m_k = \int x^k d\mu(x)$ für alle $k \in \mathbb{N}_0$. Ferner wird untersucht, ob μ im Fall der Existenz eindeutig ist. In diesem Zug wird unter anderem gezeigt, dass es Beispiele gibt, für die die Lösung nicht eindeutig ist. Somit ist die Verteilung eines Zufallsvektors im Allgemeinen auch nicht durch die Angabe seiner Momente eindeutig festgelegt.

Falls der Zufallsvektor jedoch der Annahme (EM) zu einer strikt positiven Konstante genügt, so ist seine Verteilung durch seine Momente bereits eindeutig bestimmt. Das nun folgende Korollar impliziert, dass seine Verteilung ebenfalls durch seine sämtlichen Kumulanten eindeutig bestimmt ist.

Korollar 26. Sei $N \in \mathbb{N}$. Weiterhin sei X_1 ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M_1 \in (0, \infty)$ genügt und X_2 sei ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M_2 \in (0, \infty)$ genügt. Ferner gelte $\kappa_{\alpha}^{X_1} = \kappa_{\alpha}^{X_2}$ für alle Multiindices $\alpha \in \mathbb{N}_0^N$.

Dann sind X_1 und X_2 identisch verteilt.

Beweis: Definiere $M := M_1 \wedge M_2$. Unter Verwendung von Korollar 24 erhält man, dass die Funktionen K_{X_1} und K_{X_2} auf dem nichtleeren offenen Polyzylinder $\{\tilde{u} = (\tilde{u}^1, \ldots, \tilde{u}^N) \in \mathbb{C}^N | |\tilde{u}^n| < M, n \in \{1, \ldots, N\}\}$ übereinstimmen. Somit impliziert der Eindeutigkeitssatz für mehrdimensionale holomorphe Funktionen aus Scheidemann (2005, Conclusion 1.2.12.2), dass K_{X_1} und K_{X_2} auf der Menge $\{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\}$ übereinstimmen und damit insbesondere auch die charakteristischen Funktionen $\varphi_{X_1}(y) = \exp(K_{X_1}(iy))$ und $\varphi_{X_2}(y) = \exp(K_{X_2}(iy))$ für alle $y \in \mathbb{R}^N$ übereinstimmen. Hieraus folgt die Behauptung. \Box

Applebaum (2009, Lemma 1.3.2) zeigt die bekannte Tatsache, dass für einen \mathbb{R}^{D} -wertigen Lévy-Prozess $L = (L_t)_{t \in [0,T^*]}$ für jedes $y \in \mathbb{R}^{D}$ die Abbildung $[0,T^*] \ni s \mapsto \varphi_{L_s}(y)$ stetig ist. Im Beweis des weiter unten vorgestellten Satzes 28 wird jedoch ein deutlich stärkeres Resultat benötigt. Dort ist die Stetigkeit der Abbildung $[0,T^*] \ni s \mapsto M_{L_s}(u_0)$ für jedes $u_0 \in \{\tilde{u} \in \mathbb{C}^{D} | \Re(\tilde{u}) \in (-M, M)^D\}$ erforderlich. Hiervon handelt das nun präsentierte Lemma 27. Sein Beweis orientiert sich an der Argumentation von Applebaum, verallgemeinert diese jedoch an einigen entscheidenden Stellen.

Lemma 27. Seien $D \in \mathbb{N}$ und L ein D-dimensionaler Lévy-Prozess, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt.

Dann ist für jedes fest gewählte $u_0 \in \{\tilde{u} \in \mathbb{C}^D | \Re(\tilde{u}) \in (-M, M)^D\}$ die Abbildung $[0, T^*] \ni s \mapsto M_{L_s}(u_0)$ stetig.

Beweis: Sei $u_0 \in \left\{ \tilde{u} \in \mathbb{C}^D | \Re(\tilde{u}) \in (-M, M)^D \right\}$ fest vorgegeben. Ferner seien $s_1, s_2 \in [0, T^*]$ mit $s_1 < s_2$ und $\varepsilon > 0$.

Falls $\mathbb{P}(\{L_{s_2-s_1} \neq 0\}) = 0$ gilt, so ist der Beweis trivial. Im Folgenden wird der Fall $\mathbb{P}(\{L_{s_2-s_1} \neq 0\}) > 0$ behandelt.

Wähle $q \in (1,\infty)$ mit $q\Re(u_0) \in (-M,M)^D$ und definiere $p := \frac{q}{q-1} \in (1,\infty)$. Dann gilt $\frac{1}{p} + \frac{1}{q} = 1$. Aus der Stetigkeit von $\mathbb{R}^D \ni x \mapsto \exp(\langle u_0, x \rangle) - 1$ und aus $\mathbb{E}\left[\exp\left(\langle \Re(u_0), L_{s_1} \rangle\right)\right] \in (0,\infty)$ folgt, dass es ein $\delta_1 \in (0,\infty)$ gibt mit

$$|\exp(\langle u_0, x \rangle) - 1| < \frac{\varepsilon}{2\mathbb{E}\left[\exp\left(\langle \Re(u_0), L_{s_1} \rangle\right)\right]}$$

für alle $x \in \mathbb{R}^D$ mit $|x| < \delta_1$. Weiterhin gilt

$$S := \left(\mathbb{E}\left[\left| \exp\left(\langle u_0, L_{s_2 - s_1} \rangle \right) - 1 \right|^q \right] \right)^{\frac{1}{q}} \in (0, \infty).$$

Nach Kluge (2005, Lemma 1.3) ist L stochastisch stetig und somit existiert ein $\delta_2 \in (0, \infty)$ mit

$$\mathbb{P}\left(\{|L_t| \ge \delta_1\}\right) < \left(\frac{\varepsilon}{2S\mathbb{E}\left[\exp\left(\langle \Re(u_0), L_{s_1}\rangle\right)\right]}\right)^p$$

für alle $t \in [0, T^*] \cap (0, \delta_2)$.

Falls $s_2 - s_1 < \delta_2$ gilt, so berechnet man unter Verwendung der stochastischen Unabhängigkeit und der Stationarität der Zuwächse von L sowie der Hölderschen Ungleichung

$$\begin{split} \left| M_{L_{s_{2}}}(u_{0}) - M_{L_{s_{1}}}(u_{0}) \right| \\ &\leq \mathbb{E} \left[\left| \exp\left(\langle u_{0}, L_{s_{1}} \rangle\right) \right| \left| \exp\left(\langle u_{0}, L_{s_{2}} - L_{s_{1}} \rangle\right) - 1 \right| \right] \\ &= \mathbb{E} \left[\left| \exp\left(\langle u_{0}, L_{s_{1}} \rangle\right) \right| \right] \mathbb{E} \left[\left| \exp\left(\langle u_{0}, L_{s_{2} - s_{1}} \rangle\right) - 1 \right| \right] \\ &\quad + \mathbb{E} \left[\mathbb{E} \left[\exp\left(\langle \Re(u_{0}), L_{s_{1}} \rangle\right) \right] \left(\mathbb{E} \left[\mathbb{1}_{\{|L_{s_{2} - s_{1}}| \ge \delta_{1}\}} \left| \exp\left(\langle u_{0}, L_{s_{2} - s_{1}} \rangle\right) - 1 \right| \right] \right) \\ &< \mathbb{E} \left[\exp\left(\langle \Re(u_{0}), L_{s_{1}} \rangle\right) \right] \left(\frac{\varepsilon}{2\mathbb{E} \left[\exp\left(\langle \Re(u_{0}), L_{s_{1}} \rangle\right) \right]} \\ &\quad + \left(\mathbb{E} \left[\left| \mathbb{1}_{\{|L_{s_{2} - s_{1}}| \ge \delta_{1}\}} \right|^{\frac{1}{p}} \right) \mathbb{E} \left[\left| \exp\left(\langle u_{0}, L_{s_{2} - s_{1}} \rangle\right) - 1 \right|^{q} \right] \right)^{\frac{1}{q}} \right) \\ &= \frac{\varepsilon}{2} + S\mathbb{E} \left[\exp\left(\langle \Re(u_{0}), L_{s_{1}} \rangle\right) \right] \left(\mathbb{P} \left(\{|L_{s_{2} - s_{1}}| \ge \delta_{1}\}\right) \right)^{\frac{1}{p}} \\ &< \frac{\varepsilon}{2} + S\mathbb{E} \left[\exp\left(\langle \Re(u_{0}), L_{s_{1}} \rangle\right) \right] \frac{\varepsilon}{2S\mathbb{E} \left[\exp\left(\langle \Re(u_{0}), L_{s_{1}} \rangle\right) \right]} \\ &= \varepsilon. \end{split}$$

Hieraus folgt die Behauptung.

Für die Momente erster und zweiter Ordnung eines Lévy-Prozesses L gelten die Zusammenhänge $\mathbb{E}[L_t] = t\mathbb{E}[L_1]$ und $\operatorname{Cov}(L_t) = t\operatorname{Cov}(L_1)$ für alle $t \in [0, T^*]$. Es gibt jedoch keinen entsprechenden allgemein gültigen linearen Zusammenhang der Form $\mathbb{E}[L_t^{\alpha}] = t\mathbb{E}[L_1^{\alpha}]$ für Momente einer Ordnung α mit $|\alpha| \geq 3$. Für Kumulanten hingegen besteht ein derartiger Zusammenhang. Deshalb ist gerade bei der Untersuchung von Lévy-Prozessen die Arbeit mit Kumulanten in vielen Fällen deutlich angenehmer als die Arbeit mit Momenten.

Satz 28. Seien $D \in \mathbb{N}$ und L ein D-dimensionaler Lévy-Prozess, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt. Ferner sei $t \in [0, T^*]$.

Dann gilt $K_{L_t}(u) = tK_{L_1}(u)$ für alle $u \in \left\{ \tilde{u} \in \mathbb{C}^D \middle| \Re(\tilde{u}) \in (-M, M)^D \right\}$ und $\kappa_{\alpha}^{L_t} = t\kappa_{\alpha}^{L_1}$ für alle Multiindices $\alpha \in \mathbb{N}_0^D$.

Beweis: Applebaum (2009, Theorem 1.3.3) zeigt, dass zwischen dem Lévy-Symbol ψ_{L_t} von L_t und dem Lévy-Symbol ψ_{L_1} von L_1 der Zusammenhang $\psi_{L_t}(y) = t\psi_{L_1}(y)$ für alle $y \in \mathbb{R}^D$ besteht. Dafür verwendet er unter anderem sein Lemma 1.3.2, das, wie bereits erwähnt, lediglich die Stetigkeit von $[0, T^*] \ni s \mapsto \varphi_{L_s}(y)$ für alle $y \in \mathbb{R}^D$ garantiert. Seine Argumentation lässt sich ohne nennenswerte Modifikationen auf die vorliegende erste Behauptung von $K_{L_t}(u) = tK_{L_1}(u)$ für alle $u \in \{\tilde{u} \in \mathbb{C}^D | \Re(\tilde{u}) \in (-M, M)^D\}$ übertragen. Hierfür wird jedoch die Stetigkeit von $[0, T^*] \ni s \mapsto M_{L_s}(u_0)$ für jedes $u_0 \in \{\tilde{u} \in \mathbb{C}^D | \Re(\tilde{u}) \in (-M, M)^D\}$ benötigt. Diese wird von Lemma 27 geliefert.

Die zweite Behauptung folgt unmittelbar aus der ersten.

Lemma 29. 1. Seien $N \in \mathbb{N}$ und X_1, X_2 zwei \mathbb{R}^N -wertige Zufallsvektoren, die der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügen.

Dann genügt die Summe X_1+X_2 der Annahme (\mathbb{EM}) zur Konstante $\frac{M}{2}$. Sind zusätzlich X_1 und X_2 stochastisch unabhängig, so genügt X_1+X_2 der Annahme (\mathbb{EM}) zur Konstante M und es gilt

$$M_{X_1+X_2}(u) = M_{X_1}(u)M_{X_2}(u)$$

sowie

$$K_{X_1+X_2}(u) = K_{X_1}(u) + K_{X_2}(u)$$

jeweils für alle $u \in \left\{ \tilde{u} \in \mathbb{C}^N \middle| \Re(\tilde{u}) \in (-M, M)^N \right\}.$

2. Seien $D, N \in \mathbb{N}$, $A = (a_{n,d})_{n \in \{1,...,N\}, d \in \{1,...,D\}} \in \mathbb{R}^{N \times D}$ eine Matrix und X ein \mathbb{R}^D -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt. Ferner sei

$$M^* := \frac{M}{\max_{d \in \{1,\dots,D\}} \sum_{n=1}^{N} |a_{n,d}|} > 0,$$

wobei hier $M/0 := \infty$ gesetzt wird.

Falls $M^* < \infty$, so genügt der \mathbb{R}^N -wertige Zufallsvektor AX der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante M^* . Falls $M^* = \infty$, so genügt AX der Annahme ($\mathbb{E}\mathbb{M}$) zu jeder Konstante aus $(0,\infty)$. In beiden Fällen gilt außerdem

$$M_{AX}(u) = M_X \left(A^T u \right)$$
sowie

$$K_{AX}(u) = K_X \left(A^T u \right)$$

jeweils für alle $u \in \left\{ \tilde{u} \in \mathbb{C}^N \middle| \Re(\tilde{u}) \in (-M^*, M^*)^N \right\}.$

Beweis: Der Beweis von Teilaussage 1 ist trivial und wird ausgelassen. Es wird lediglich der Nachweis für Teilaussage 2 präsentiert.

Falls $M^* < \infty$, so sei $\tilde{M} := M^*$ und falls $M^* = \infty$, so sei $\tilde{M} \in (0, \infty)$. Dann berechnet man für $u = (u^1, \ldots, u^N) \in [-\tilde{M}, \tilde{M}]^N$ und $d \in \{1, \ldots, D\}$

$$\left|\sum_{n=1}^{N} a_{n,d} u^n\right| \le \sum_{n=1}^{N} |a_{n,d}| |u^n| \le M.$$

Hieraus folgt $A^T u \in [-M, M]^D$ und damit

$$M_{AX}(u) = \mathbb{E}\left[\exp\left(\langle u, AX \rangle\right)\right] = \mathbb{E}\left[\exp\left(\langle A^T u, X \rangle\right)\right] = M_X\left(A^T u\right) < \infty.$$

Also genügt AX der Annahme (\mathbb{EM}) zur Konstante M.

Man prüft leicht nach, dass es eine offene Umgebung von $0 \in \mathbb{C}^N$ gibt, auf der K_{AX} und $\{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M^*, M^*)^N\} \ni u \mapsto K_X(A^T u)$ übereinstimmen. Zusammen mit dem Eindeutigkeitssatz für mehrdimensionale holomorphe Funktionen aus Scheidemann (2005, Conclusion 1.2.12.2) folgt hieraus $K_{AX}(u) = K_X(A^T u)$ und damit auch $M_{AX}(u) = M_X(A^T u)$ für alle $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M^*, M^*)^N\}$. \Box

Lemma 30. Seien $D, N, K \in \mathbb{N}$, $B = (b_{d,n})_{d \in \{1,\ldots,D\},n \in \{1,\ldots,N\}} \in \mathbb{R}^{D \times N}$ eine Matrix, $U \subset \mathbb{R}^D$ eine offene Teilmenge, $g : U \to \mathbb{R}$, $y \mapsto g(y)$, eine K-fach stetig partiell differenzierbare Funktion und $(i_1,\ldots,i_K) \in \{1,\ldots,N\}^K$ ein K-Tupel von Indices.

Dann gilt für alle $u \in \{\tilde{u} \in \mathbb{R}^N | B\tilde{u} \in U\}$

$$\frac{\partial^K}{\partial u^{i_1}\cdots\partial u^{i_K}}\Big(g(Bu)\Big)=\sum_{d_1,\dots,d_K=1}^D\left(\prod_{k=1}^K b_{d_k,i_k}\right)\left(\frac{\partial^K g}{\partial y^{d_1}\cdots\partial y^{d_K}}\right)(Bu).$$

Beweis: Der Nachweis ist trivial und kann in kanonischer Weise durch vollständige Induktion über K und unter Verwendung der mehrdimensionalen Kettenregel geführt werden.

Definition 31 (Multiindex zu gegebenem Indexvektor und Indexvektor zu gegebenem Multiindex).

- 1. Seien $N, K \in \mathbb{N}$. Zu einem K-Tupel $(i_1, \ldots, i_K) \in \{1, \ldots, N\}^K$ von Indices bezeichnet $\boldsymbol{\alpha}_N(i_1, \ldots, i_K) = (\alpha^1, \ldots, \alpha^N) \in \mathbb{N}_0^N$ den Multiindex mit den Einträgen $\alpha^n = |\{\tilde{k} \in \{1, \ldots, K\} | i_{\tilde{k}} = n\}|$ für $n \in \{1, \ldots, N\}$. $\boldsymbol{\alpha}_N(i_1, \ldots, i_K)$ heißt der zu (i_1, \ldots, i_K) dazugehörige Multiindex.
- 2. Set $N \in \mathbb{N}$. Zu einem gegebenen Multiindex $\boldsymbol{\alpha} = (\alpha^1, \dots, \alpha^N) \in \mathbb{N}_0^N$ bezeichnet $\boldsymbol{I}_{\boldsymbol{\alpha}} = (I_{\boldsymbol{\alpha}}^1, \dots, I_{\boldsymbol{\alpha}}^{|\boldsymbol{\alpha}|}) \in \mathbb{N}^{|\boldsymbol{\alpha}|}$ dasjenige dazugehörige $|\boldsymbol{\alpha}|$ -Tupel von Indices, dessen Einträge monoton wachsend sind, also

$$\boldsymbol{I}_{\boldsymbol{\alpha}} := (\underbrace{1, \dots, 1}_{\alpha^{1}-\mathrm{mal}}, \dots, \underbrace{n, \dots, n}_{\alpha^{n}-\mathrm{mal}}, \dots, \underbrace{N, \dots, N}_{\alpha^{N}-\mathrm{mal}}).$$

Sind X_1 und X_2 stochastisch unabhängige Zufallsvektoren derselben Dimension, so gilt $\mathbb{E} [X_1 + X_2] = \mathbb{E} [X_1] + \mathbb{E} [X_2]$ und $\operatorname{Cov}(X_1 + X_2) = \operatorname{Cov}(X_1) + \operatorname{Cov}(X_2)$. Für Momente dritter und höherer Ordnung sind die Zusammenhänge jedoch deutlich komplizierter. Ferner ist die Berechnung von Momenten dritter und höherer Ordnung einer linearen Abbildung AX eines Zufallsvektors X ebenfalls recht aufwändig. Anders verhält es sich hingegen mit Kumulanten. Die Berechnung von Kumulanten einer Summe von stochastisch unabhängigen Zufallsvektoren ist trivial. Ferner lässt sich für einen gegebenen Multiindex α die Kumulante κ_{α}^{AX} von AX der Ordnung α als Linearkombination der Kumulanten $\kappa_{\overline{\alpha}}^{\overline{\alpha}}$ für geeignete Multiindices $\widetilde{\alpha}$ mit $|\widetilde{\alpha}| = K := |\alpha|$ darstellen. Aus diesem Grund ist bei der Arbeit mit Summen von stochastisch unabhängigen Zufallsvektoren und linearen Abbildungen davon die Betrachtung von Kumulanten häufig angenehmer als die Betrachtung von Momenten.

Satz 32. 1. Seien $N \in \mathbb{N}$ und X_1, X_2 stochastisch unabhängige \mathbb{R}^N -wertige Zufallsvektoren, die der Annahme ($\mathbb{E}\mathbb{M}$) zu einer beliebigen strikt positiven Konstante genügen. Ferner sei $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ ein Multiindex.

Dann gilt für die Kumulante $\kappa_{\alpha}^{X_1+X_2}$ von X_1+X_2 der Ordnung α

$$\kappa_{\alpha}^{X_1+X_2} = \kappa_{\alpha}^{X_1} + \kappa_{\alpha}^{X_2}$$

2. Seien $D, N, K \in \mathbb{N}$, $A = (a_{n,d})_{n \in \{1,...,N\}, d \in \{1,...,D\}} \in \mathbb{R}^{N \times D}$ eine Matrix, $(i_1, \ldots, i_K) \in \{1, \ldots, N\}^K$ ein K-Tupel von Indices und X ein \mathbb{R}^D -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zu einer beliebigen strikt positiven Konstante genügt.

Dann gilt für die Kumulante $\kappa_{\boldsymbol{\alpha}_N(i_1,\ldots,i_K)}^{AX}$ des Zufallsvektors AX der Ordnung $\boldsymbol{\alpha}_N(i_1,\ldots,i_K)$

$$\kappa_{\boldsymbol{\alpha}_N(i_1,\dots,i_K)}^{AX} = \sum_{d_1,\dots,d_K=1}^D \left(\prod_{k=1}^K a_{i_k,d_k}\right) \kappa_{\boldsymbol{\alpha}_D(d_1,\dots,d_K)}^X.$$
 (6)

Insbesondere gilt

$$\kappa_{\alpha}^{cX} = c^K \kappa_{\alpha}^X.$$

für alle $c \in \mathbb{R}$ und jeden Multiindex $\boldsymbol{\alpha} \in \mathbb{N}_0^D$ mit $|\boldsymbol{\alpha}| = K$.

Beweis: Lemma 29 impliziert, dass sowohl $X_1 + X_2$ als auch AX ebenfalls der Annahme ($\mathbb{E}M$) zu einer strikt positiven Konstante genügen und damit nach Satz 23 alle Kumulanten von $X_1 + X_2$ und von AX existieren.

Weiter erhält man unter Verwendung der Teilaussage 1 von Lemma 29

$$\kappa_{\boldsymbol{\alpha}}^{X_1+X_2} = \left. \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial u^{\boldsymbol{\alpha}}} K_{X_1+X_2}(u) \right|_{u=0} = \left. \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial u^{\boldsymbol{\alpha}}} \left[K_{X_1}(u) + K_{X_2}(u) \right] \right|_{u=0} = \kappa_{\boldsymbol{\alpha}}^{X_1} + \kappa_{\boldsymbol{\alpha}}^{X_2},$$

also Teilaussage 1 des Satzes.

Aus Teilaussage 2 von Lemma 29 in Kombination mit Lemma 30 für $g(y) := K_X(y)$ und $B = (b_{d,n})_{d \in \{1,...,D\}, n \in \{1,...,N\}} := A^T$ folgt

$$\kappa_{\boldsymbol{\alpha}_{N}(i_{1},\dots,i_{K})}^{AX} = \frac{\partial^{K}}{\partial u^{i_{1}}\cdots\partial u^{i_{K}}} \left(g(Bu)\right)\Big|_{u=0}$$

$$= \sum_{d_{1},\dots,d_{K}=1}^{D} \left(\prod_{k=1}^{K} b_{d_{k},i_{k}}\right) \left(\frac{\partial^{K}g}{\partial y^{d_{1}}\cdots\partial y^{d_{K}}}\right) (Bu)\Big|_{u=0}$$

$$= \sum_{d_{1},\dots,d_{K}=1}^{D} \left(\prod_{k=1}^{K} a_{i_{k},d_{k}}\right) \kappa_{\boldsymbol{\alpha}_{D}(d_{1},\dots,d_{K})}^{X},$$

also die erste Gleichung von Teilaussage 2 des Satzes. Die letzte behauptete Gleichung folgt als Spezialfall für D = N und $A = cE_D$. Dabei ist E_D die $D \times D$ Einheitsmatrix.

Der nächste Satz ist ein zentrales Resultat über konvergente Folgen von Zufallsvektoren, die der Annahme (EM) genügen. Er macht eine Aussage darüber, unter welchen Bedingungen der Grenzwert bezüglich der Konvergenz in Verteilung ebenfalls der Annahme (EM) genügt und wie es in diesem Fall um das Konvergenzverhalten der dazugehörigen Folge von momenterzeugenden Funktionen bestellt ist. Der Satz stellt eine Erweiterung des Stetigkeitssatzes von Lévy-Cramér dar. **Satz 33.** Seien $N \in \mathbb{N}$ und $(X_R)_{R \in \mathbb{N}}$ eine Folge von \mathbb{R}^N -wertigen Zufallsvektoren, die in Verteilung gegen einen \mathbb{R}^N -wertigen Zufallsvektor X konvergiert. Ferner sei $M \in (0, \infty)$ und für alle $R \in \mathbb{N}$ genüge X_R der Annahme ($\mathbb{E}M$) zu jeder Konstante aus (0, M).

Dann sind die folgenden drei Eigenschaften äquivalent.

- 1. Es existiert eine Funktion $S : (-M, M)^N \to \mathbb{R}$ mit $M_{X_R}(u) \leq S(u)$ für alle $u \in (-M, M)^N$ und alle $R \in \mathbb{N}$.
- 2. X genügt der Annahme (\mathbb{EM}) zu jeder Konstante aus (0, M) und es gilt

$$\lim_{R \to \infty} M_{X_R}(u) = M_X(u)$$

 $\label{eq:full_full} \textit{für alle } u \in \big\{ \tilde{u} \in \mathbb{C}^N \big| \Re(\tilde{u}) \in (-M,M)^N \big\}.$

3. X genügt der Annahme (\mathbb{EM}) zu jeder Konstante aus (0, M) und es gilt

$$\lim_{R \to \infty} M_{X_R}(u) = M_X(u)$$

 $f \ddot{u}r \ alle \ u \in (-M, M)^N.$

Beweis: $1. \Rightarrow 2$. Der Standard-Beweis, der in den einschlägigen Lehrbüchern zum Nachweis des Stetigkeitssatzes von Lévy-Cramér verwendet wird, nutzt die Tatsache aus, dass die Abbildung $\mathbb{R} \ni y \mapsto \exp(iy)$ beschränkt ist. Zum Nachweis von Satz 33 ist jedoch die Abbildung $\mathbb{C} \ni z \mapsto \exp(z)$ zu betrachten. Diese ist nicht beschränkt. Somit funktioniert der Standard-Beweis des Stetigkeitssatzes von Lévy-Cramér nicht und es sind tiefergehende Überlegungen erforderlich.

Seien $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\}, x = (x^1, \dots, x^N) := \Re(u) \in (-M, M)^N$ und $y = (y^1, \dots, y^N) := \Im(u)$. Dann gilt für $R \in \mathbb{N}$

$$\mathbb{E}\left[\left|\exp\left(\langle u, X_R \rangle\right)\right|\right] = \mathbb{E}\left[\exp\left(\langle x, X_R \rangle\right)\right] < \infty$$

und damit

$$M_{X_R}(u) = \mathbb{E}\left[\left[\exp\left(\langle x, X_R \rangle\right) \cos\left(\langle y, X_R \rangle\right)\right]^+\right] \\ -\mathbb{E}\left[\left[\exp\left(\langle x, X_R \rangle\right) \cos\left(\langle y, X_R \rangle\right)\right]^-\right] \\ +i\mathbb{E}\left[\left[\exp\left(\langle x, X_R \rangle\right) \sin\left(\langle y, X_R \rangle\right)\right]^+\right] \\ -i\mathbb{E}\left[\left[\exp\left(\langle x, X_R \rangle\right) \sin\left(\langle y, X_R \rangle\right)\right]^-\right].$$
(7)

Zunächst wird die gleichgradige Integrierbarkeit der Familie

$$\mathcal{C} := \left\{ \left[\exp\left(\langle x, X_R \rangle \right) \cos\left(\langle y, X_R \rangle \right) \right]^+ \middle| R \in \mathbb{N} \right\}$$

nachgewiesen. $\mathcal{C} \subset L^1(\Omega, \mathcal{F}, \mathbb{P})$ ist klar. Sei nun

$$\alpha := \min_{n \in \{1, \dots, N\}} \frac{1}{2} \left(M / |x^n| - 1 \right) \wedge 1,$$

wobei hier $M/0 := \infty$ gesetzt wird. Dann ist $\alpha > 0$, die Funktion $G : \mathbb{R}_+ \to \mathbb{R}, t \mapsto G(t) := t^{1+\alpha}$, ist nichtnegativ, monoton wachsend und konvex und es gilt $\lim_{t\to\infty} G(t)/t = \infty$. Ferner gilt $(1+\alpha)x \in (-M, M)^N$ und damit

$$\sup_{R \in \mathbb{N}} \mathbb{E} \left[G \left(\left[\exp \left(\langle x, X_R \rangle \right) \cos \left(\langle y, X_R \rangle \right) \right]^+ \right) \right] \\ \leq \sup_{R \in \mathbb{N}} \mathbb{E} \left[\exp \left(\left\langle (1 + \alpha) x, X_R \rangle \right) \right] \\ \leq S((1 + \alpha) x) \\ < \infty.$$

Somit liefert der Satz von de la Vallée-Poussin die gleichgradige Integrierbarkeit von \mathcal{C} .

Sei nun $\varepsilon > 0$. Weil \mathcal{C} gleichgradig integrierbar ist, existiert ein $K \in \mathbb{R}_+$ mit

$$\mathbb{E}\left[(f \circ X_R)\mathbb{1}_{(f \circ X_R \ge K)}\right] < \varepsilon$$

für alle $R \in \mathbb{N}$, wobei $f : \mathbb{R}^N \to \mathbb{R}_+$ definiert ist als

$$f(v) := \left[\exp\left(\langle x, v \rangle\right) \cos\left(\langle y, v \rangle\right)\right]^+.$$

Mit $f_K := f \wedge K$ erhält man

$$\begin{split} \mathbb{E}\left[f \circ X_{R}\right] &= \mathbb{E}\left[(f \circ X_{R})\mathbb{1}_{\left(f \circ X_{R} < K\right)}\right] + \mathbb{E}\left[(f \circ X_{R})\mathbb{1}_{\left(f \circ X_{R} \geq K\right)}\right] \\ &= \mathbb{E}\left[f_{K} \circ X_{R}\right] - \mathbb{E}\left[K\mathbb{1}_{\left(f \circ X_{R} \geq K\right)}\right] + \mathbb{E}\left[(f \circ X_{R})\mathbb{1}_{\left(f \circ X_{R} \geq K\right)}\right] \\ &< \mathbb{E}\left[f_{K} \circ X_{R}\right] + \varepsilon \end{aligned}$$

für alle $R \in \mathbb{N}$. Weil f_K stetig und nach oben beschränkt ist und ferner f stetig und nach unten beschränkt ist, liefert das Portemanteau Theorem

$$\limsup_{R \to \infty} \mathbb{E} \left[f \circ X_R \right] \leq \limsup_{R \to \infty} \mathbb{E} \left[f_K \circ X_R \right] + \varepsilon$$
$$\leq \mathbb{E} \left[f_K \circ X \right] + \varepsilon$$
$$\leq \mathbb{E} \left[f \circ X \right] + \varepsilon$$
$$\leq \liminf_{R \to \infty} \mathbb{E} \left[f \circ X_R \right] + \varepsilon.$$

Weil $\varepsilon > 0$ beliebig war, folgt hieraus $\lim_{R \to \infty} \mathbb{E} [f \circ X_R] = \mathbb{E} [(f \circ X)]$, also $\lim_{R \to \infty} \mathbb{E} \left[\left[\exp \left(\langle x, X_R \rangle \right) \cos \left(\langle y, X_R \rangle \right) \right]^+ \right] = \mathbb{E} \left[\left[\exp \left(\langle x, X \rangle \right) \cos \left(\langle y, X \rangle \right) \right]^+ \right].$ Analog zeigt man die Konvergenz der anderen Terme auf der rechten Seite von Gleichung (7). Somit ist

$$\lim_{R \to \infty} M_{X_R}(u) = M_X(u).$$

gezeigt. Insbesondere folgt für $u \in (-M, M)^N$

$$M_X(u) = \lim_{R \to \infty} M_{X_R}(u) \le S(u) < \infty,$$

also genügt X der Annahme ($\mathbb{E}M$) zu jeder Konstante aus (0, M).

Die Implikation $2. \Rightarrow 3.$ ist trivial.

$$3. \Rightarrow 1.$$
 Sei $u \in (-M, M)^N$. Aus $\lim_{R \to \infty} M_{X_R}(u) = M_X(u) < \infty$ folgt
 $S(u) := \sup_{R \in \mathbb{N}} M_{X_R}(u) < \infty$

und damit die Behauptung.

Der nun vorgestellte letzte Satz von Abschnitt 3.1 wird in der vorliegenden Arbeit zwar nirgends angewandt. Er ist jedoch von eigenem mathematischen Interesse und wird deshalb an dieser Stelle präsentiert. Er macht eine Aussage darüber, unter welchen Zusatzbedingungen an eine Folge von Zufallsvektoren der Schluss von der Konvergenz in Verteilung auf die Konvergenz aller Momente und Kumulanten zulässig ist.

Satz 34. Es gelten die Voraussetzungen von Satz 33 und es sei eine (und damit alle) der äquivalenten Eigenschaften 1. 2. und 3. erfüllt. Ferner existiere eine kompakte Umgebung $K \subset \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M, M)^N\}$ von 0 derart, dass für alle $u \in K$ die Folge $(|M_{X_R}(u) - M_X(u)|)_{R \in \mathbb{N}}$ monoton fallend ist.

Dann gilt für alle Multiindices $\boldsymbol{\alpha} \in \mathbb{N}_0^N$

$$\lim_{R \to \infty} \mathbb{E}\left[X_R^{\alpha}\right] = \mathbb{E}\left[X^{\alpha}\right]$$

und

$$\lim_{R \to \infty} \kappa_{\alpha}^{X_R} = \kappa_{\alpha}^X$$

Beweis: Weil sich die Kumulante der Ordnung $\boldsymbol{\alpha}$ als Polynom der Momente der Ordnungen $\tilde{\boldsymbol{\alpha}} \in \mathbb{N}_0^N$ mit $\tilde{\boldsymbol{\alpha}} \leq \boldsymbol{\alpha}$ darstellen lässt, genügt der Nachweis der Konvergenz der Momente.

Der Satz von Dini impliziert, dass die Funktionenfolge $(|M_{X_R} - M_X|)_{R \in \mathbb{N}}$ auf K gleichmäßig gegen die Nullfunktion konvertiert. Hieraus folgt, dass die Funktionenfolge $(M_{X_R})_{R \in \mathbb{N}}$ auf K gleichmäßig gegen M_X konvergiert. Zusammen mit dem Weierstraßschen Konvergenzsatz für mehrdimensionale holomorphe Funktionen aus Scheidemann (2005, Theorem 1.4.20) erhält man, dass die Folge $\left(\frac{\partial^{|\alpha|}}{\partial u^{\alpha}}M_{X_R}\right)_{R \in \mathbb{N}}$ der partiellen Ableitungen im Inneren von Kgleichmäßig auf den kompakten Teilmengen gegen die entsprechende partielle Ableitung $\frac{\partial^{|\alpha|}}{\partial u^{\alpha}}M_X$ konvergiert. Weil 0 ein Element des Inneren von K ist, folgt insbesondere

$$\lim_{R \to \infty} \mathbb{E} \left[X_R^{\alpha} \right] = \lim_{R \to \infty} \left. \frac{\partial^{|\alpha|}}{\partial u^{\alpha}} M_{X_R}(u) \right|_{u=0} = \left. \frac{\partial^{|\alpha|}}{\partial u^{\alpha}} M_X(u) \right|_{u=0} = \mathbb{E} \left[X^{\alpha} \right],$$

also die Behauptung.

3.2 Gemeinsame Kumulanten von stochastischen Integralen nach einem Lévy-Prozess

Das Hauptergebnis des Abschnitts ist Satz 35, der eine Formel für die explizite Berechnung von Kumulanten von Zufallsvektoren, deren Komponenten stochastische Integrale nach einem Lévy-Prozess sind, bereitstellt. Dieser spielt bei der angestrebten Kalibrierung des Modells aus Abschnitt 2 eine zentrale Rolle.

Am Ende des Abschnitts wird, wie bereits angekündigt, der Beweis von Teilaussage 2 von Korollar 8 aus Abschnitt 2.2 nachgeholt.

Satz 35. Seien $D, N \in \mathbb{N}$, $M \in (0, \infty)$ und $L = (L_t)_{t \in [0,T^*]}$ ein D-dimensionaler Lévy-Prozess, der der Annahme ($\mathbb{E}\mathbb{M}$) zu jeder Konstante aus (0, M) genügt. Weiterhin seien $f_1, \ldots, f_N : [0, T^*] \to \mathbb{R}^D$ stetige Funktionen, $\Delta_1, \Delta_2 \in [0, T^*]$ mit $\Delta_1 < \Delta_2$ und X sei der \mathbb{R}^N -wertige Zufallsvektor mit den Komponenten

$$X^n := \int_{\Delta_1}^{\Delta_2} f_n(s) \,\mathrm{d}L_s$$

 $f \ddot{u} r \ n \in \{1, \dots, N\}.$

1. Sei ferner

$$M^* := \frac{M}{\max_{s \in [\Delta_1, \Delta_2]} \max_{d \in \{1, \dots, D\}} \sum_{n=1}^N |f_n^d(s)|} > 0,$$

wobei hier $M/0 := \infty$ gesetzt wird.

Dann genügt X der Annahme ($\mathbb{E}\mathbb{M}$) zu jeder Konstante aus $(0, M^*)$ und es gilt

$$K_X(u) = \int_{\Delta_1}^{\Delta_2} K_{L_1} \left([F(s)]^T u \right) \, \mathrm{d}s$$

für alle $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M^*, M^*)^N\}$, wobei die matrixwertige Funktion $F : [0, T^*] \to \mathbb{R}^{N \times D}$ definiert ist durch

$$F(s) := \begin{pmatrix} f_1^1(s) & \dots & f_1^D(s) \\ \vdots & \ddots & \vdots \\ f_N^1(s) & \dots & f_N^D(s) \end{pmatrix}$$

2. Seien zusätzlich $K \in \mathbb{N}$, $(i_1, \ldots, i_K) \in \{1, \ldots, N\}^K$ ein K-Tupel von Indices und $\boldsymbol{\alpha}_N(i_1, \ldots, i_K) \in \mathbb{N}_0^N$ der gemäß Definition 31 dazugehörige Multiindex. Dann gilt für die Kumulante $\kappa_{\boldsymbol{\alpha}_N(i_1,\ldots,i_K)}^X$ von X der Ordnung $\boldsymbol{\alpha}_N(i_1,\ldots,i_K)$

$$\kappa_{\alpha_N(i_1,\dots,i_K)}^X = \sum_{d_1,\dots,d_K=1}^D \kappa_{\alpha_D(d_1,\dots,d_K)}^{L_1} \int_{\Delta_1}^{\Delta_2} \left(\prod_{k=1}^K f_{i_k}^{d_k}(s)\right) \, \mathrm{d}s.$$

Beweis: Betrachtet sei die Folge $(X_R)_{R\in\mathbb{N}}$ von \mathbb{R}^N -wertigen Zufallsvektoren mit den Komponenten

$$X_R^n := \sum_{r=0}^{R-1} \left\langle f_n\left(\Delta_1 + r\frac{\Delta_2 - \Delta_1}{R}\right), L_{\Delta_1 + (r+1)\frac{\Delta_2 - \Delta_1}{R}} - L_{\Delta_1 + r\frac{\Delta_2 - \Delta_1}{R}} \right\rangle$$

für $n \in \{1, \ldots, N\}$ und $R \in \mathbb{N}$.

Mit der matrixwertigen Funktion F erhält man

$$X_R = \sum_{r=0}^{R-1} F\left(\Delta_1 + r \frac{\Delta_2 - \Delta_1}{R}\right) \left(L_{\Delta_1 + (r+1)\frac{\Delta_2 - \Delta_1}{R}} - L_{\Delta_1 + r\frac{\Delta_2 - \Delta_1}{R}}\right) \tag{8}$$

für $R \in \mathbb{N}$.

Aus Protter (2005, Part II, Theorem 21) folgt, dass die Folge $(X_R)_{R\in\mathbb{N}}$ stochastisch und damit auch in Verteilung gegen X konvergiert. Somit impliziert der Stetigkeitssatz von Lévy-Cramér, dass die Folge $(\varphi_{X_R})_{R\in\mathbb{N}}$ der charakteristischen Funktionen der X_R auf ganz \mathbb{R}^N punktweise gegen die charakteristische Funktion φ_X von X konvergiert. Hieraus lässt sich jedoch nicht ohne Weiteres schlussfolgern, dass X der Annahme ($\mathbb{E}M$) zu einer Konstante aus $(0, M^*)$ genügt. Für den vorliegenden Fall ist die Aussage des Stetigkeitssatzes von Lévy-Cramér also nicht stark genug. Deshalb wird der nun vorgestellte Weg beschritten.

Sei $\tilde{M} \in (0, M^*)$. Weiterhin sei für $R \in \mathbb{N}$ und $r \in \{0, \dots, R-1\}$

$$M_{R,r}^{*} := \frac{M}{\max_{d \in \{1,...,D\}} \sum_{n=1}^{N} \left| f_{n}^{d} \left(\Delta_{1} + r \frac{\Delta_{2} - \Delta_{1}}{R} \right) \right|},$$

wobei hier wieder $M/0 := \infty$ gesetzt wird.

Falls $M_{R,r}^* < \infty$, so sei $\tilde{M}_{R,r} := M_{R,r}^*$ und falls $M_{R,r}^* = \infty$, so sei $\tilde{M}_{R,r} \in (\tilde{M}, \infty)$. In beiden Fällen gilt $\tilde{M} \in (0, \tilde{M}_{R,r})$. Satz 19 impliziert, dass der Zuwachs $L_{\Delta_1+(r+1)} \frac{\Delta_2-\Delta_1}{R} - L_{\Delta_1+r} \frac{\Delta_2-\Delta_1}{R}$ der Annahme (EM) zu jeder Konstante aus (0, M) genügt. Hieraus lässt sich nun unter Verwendung von Teilaussage 2 von Lemma 29 schlussfolgern, dass der Zufallsvektor

$$F\left(\Delta_1 + r\frac{\Delta_2 - \Delta_1}{R}\right) \left(L_{\Delta_1 + (r+1)\frac{\Delta_2 - \Delta_1}{R}} - L_{\Delta_1 + r\frac{\Delta_2 - \Delta_1}{R}}\right)$$

der Annahme ($\mathbb{E}\mathbb{M}$) zu jeder Konstante aus $(0, \tilde{M}_{R,r})$ genügt. Somit genügt dieser Zufallsvektor insbesondere auch der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante \tilde{M} . Weiterhin folgt aus der stochastischen Unabhängigkeit der Summanden auf der rechten Seite von Gleichung (8) zusammen mit Teilaussage 1 von Lemma 29, dass X_R ebenfalls der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante \tilde{M} genügt. Insbesondere genügt X_R damit auch der Annahme ($\mathbb{E}\mathbb{M}$) zu jeder Konstante aus $(0, \tilde{M})$.

Sei nun $u \in (-\tilde{M}, \tilde{M})^N$. Dann gilt $[F(s)]^T u \in (-M, M)^D$ für alle $s \in [\Delta_1, \Delta_2]$ und weil die Funktion

$$[\Delta_1, \Delta_2] \ni s \mapsto K_{L_1}\left([F(s)]^T u \right) \in \mathbb{R}$$

stetig ist, existiert ein $K_{L_1}^{\max}(u) \in \mathbb{R}_+$ mit

$$K_{L_1}\left(\left[F\left(s\right)\right]^T u\right) \le K_{L_1}^{\max}(u)$$

für alle $s \in [\Delta_1, \Delta_2]$. Aus der stochastischen Unabhängigkeit und der Stationarität der Zuwächse von L, Lemma 29 sowie Satz 28 folgt damit für die momenterzeugende Funktion M_{X_R} von X_R

$$M_{X_{R}}(u) = \prod_{r=0}^{R-1} M_{\left(L_{\Delta_{1}+(r+1)} \frac{\Delta_{2}-\Delta_{1}}{R} - L_{\Delta_{1}+r} \frac{\Delta_{2}-\Delta_{1}}{R}\right)} \left(\left[F\left(\Delta_{1}+r \frac{\Delta_{2}-\Delta_{1}}{R}\right) \right]^{T} u \right)$$

$$= \prod_{r=0}^{R-1} M_{L_{\frac{\Delta_{2}-\Delta_{1}}{R}}} \left(\left[F\left(\Delta_{1}+r \frac{\Delta_{2}-\Delta_{1}}{R}\right) \right]^{T} u \right)$$

$$= \exp \left[\sum_{r=0}^{R-1} \frac{\Delta_{2}-\Delta_{1}}{R} K_{L_{1}} \left(\left[F\left(\Delta_{1}+r \frac{\Delta_{2}-\Delta_{1}}{R}\right) \right]^{T} u \right) \right]$$

$$\leq \exp \left[\sum_{r=0}^{R-1} \frac{\Delta_{2}-\Delta_{1}}{R} K_{L_{1}}^{\max}(u) \right]$$

$$= \exp \left[(\Delta_{2}-\Delta_{1}) K_{L_{1}}^{\max}(u) \right] =: S(u) < \infty$$

$$(9)$$

für alle $R \in \mathbb{N}$.

Damit sind alle Voraussetzungen für die Implikation 1. \Rightarrow 2. von Satz 33 erfüllt. Also genügt X der Annahme ($\mathbb{E}\mathbb{M}$) zu jeder Konstante aus $(0, \tilde{M})$ und es gilt $\lim_{R\to\infty} M_{X_R}(u) = M_X(u)$ für alle $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-\tilde{M}, \tilde{M})^N \}$. Weiterhin erhält man zusammen mit Gleichung (9)

$$M_X(u) = \exp\left[\lim_{R \to \infty} \sum_{r=0}^{R-1} \frac{\Delta_2 - \Delta_1}{R} K_{L_1} \left(\left[F\left(\Delta_1 + r \frac{\Delta_2 - \Delta_1}{R}\right) \right]^T u \right) \right] \\ = \exp\left[\int_{\Delta_1}^{\Delta_2} K_{L_1} \left([F(s)]^T u \right) ds \right]$$

und damit

$$K_X(u) = \int_{\Delta_1}^{\Delta_2} K_{L_1} \left([F(s)]^T u \right) \,\mathrm{d}s \tag{10}$$

 $\text{für alle } u \in \big\{\tilde{u} \in \mathbb{C}^N \big| \Re(\tilde{u}) \in (-\tilde{M}, \tilde{M})^N \big\}.$

Weil $\tilde{M} \in (0, M^*)$ beliebig war, folgt damit, dass X der Annahme ($\mathbb{E}M$) zu jeder Konstante aus $(0, M^*)$ genügt und dass Gleichung (10) auch für alle $u \in \{\tilde{u} \in \mathbb{C}^N | \Re(\tilde{u}) \in (-M^*, M^*)^N\}$ gilt. Damit ist die erste Teilaussage des Satzes gezeigt.

Zusammen mit Lemma 30 für $B := [F(s)]^T$ und $g(y) := K_{L_1}(y)$ folgt

weiter

$$\begin{aligned} \kappa_{\boldsymbol{\alpha}_{N}(i_{1},...,i_{K})}^{X} &= \int_{\Delta_{1}}^{\Delta_{2}} \frac{\partial^{K}}{\partial u^{i_{1}}\cdots\partial u^{i_{K}}} \left(K_{L_{1}}\left([F(s)]^{T} u\right)\right) \,\mathrm{d}s \Big|_{u=0\in\mathbb{R}^{N}} \\ &= \int_{\Delta_{1}}^{\Delta_{2}} \sum_{d_{1},...,d_{K}=1}^{D} \left(\prod_{k=1}^{K} f_{i_{k}}^{d_{k}}(s)\right) \left(\frac{\partial^{K}K_{L_{1}}}{\partial y^{d_{1}}\cdots\partial y^{d_{K}}}\right) \left([F(s)]^{T} u\right) \,\mathrm{d}s \Big|_{u=0\in\mathbb{R}^{N}} \\ &= \int_{\Delta_{1}}^{\Delta_{2}} \sum_{d_{1},...,d_{K}=1}^{D} \left(\prod_{k=1}^{K} f_{i_{k}}^{d_{k}}(s)\right) \left(\frac{\partial^{K}K_{L_{1}}}{\partial y^{d_{1}}\cdots\partial y^{d_{K}}}\right) (v) \Big|_{v=0\in\mathbb{R}^{D}} \,\mathrm{d}s \\ &= \sum_{d_{1},...,d_{K}=1}^{D} \kappa_{\boldsymbol{\alpha}_{D}(d_{1},...,d_{K})}^{L_{1}} \int_{\Delta_{1}}^{\Delta_{2}} \left(\prod_{k=1}^{n} f_{i_{k}}^{d_{k}}(s)\right) \,\mathrm{d}s, \end{aligned}$$

also die zweite Teilaussage des Satzes.

Bemerkung 36. Im vorstehenden Satz 35 wird als Voraussetzung lediglich gefordert, dass der Lévy-Prozess L der Annahme ($\mathbb{E}M$) zu jeder Konstante aus (0, M) genügt. Diese Anforderung ist schwächer als die Bedingung, dass L der Annahme ($\mathbb{E}M$) zur Konstante M genügt.

Doch selbst wenn L sogar der Annahme ($\mathbb{E}M$) zur Konstante M genügt, so lässt sich aus der Argumentation im Beweis trotzdem nicht schlussfolgern, dass X der Annahme ($\mathbb{E}M$) zur Konstante M^* genügt.

Korollar 37. Seien $D, N, K \in \mathbb{N}, f_1, \ldots, f_N : [0, T^*] \to \mathbb{R}^D$ stetige Funktionen und $L = (L_t)_{t \in [0,T^*]}$ ein D-dimensionaler Lévy-Prozess, der der Annahme ($\mathbb{E}\mathbb{M}$) zu einer strikt positiven Konstante genügt. Weiterhin seien $\Delta \in \mathbb{R}_+, t_1 \ldots, t_K \in [0, T^* - \Delta]$ und für $k \in \{1, \ldots, K\}$ sei X_k der \mathbb{R}^N wertige Zufallsvektor mit den Komponenten

$$X_k^n = \int_{t_k}^{t_k+\Delta} f_n(s-t_k) \,\mathrm{d}L_s$$

 $f \ddot{u} r \ n \in \{1, \dots, N\}.$

Dann sind die Zufallsvektoren X_1, \ldots, X_K identisch verteilt.

Beweis: Satz 35 impliziert, dass alle X_k der Annahme ($\mathbb{E}M$) zu einer strikt positiven Konstante genügen und dass für jeden fest gewählten Multiindex $\alpha \in \mathbb{N}_0^N$ die Kumulanten $\kappa_{\alpha}^{X_k}$ der X_k für alle $k \in \{1, \ldots, K\}$ übereinstimmen. Somit folgt die Behauptung aus Korollar 26.

Nun steht das Rüstzeug zur Verfügung, um den Beweis von Teilaussage 2 von Korollar 8 aus Abschnitt 2.2 zu erbringen.

Beweis von Teilaussage 2 von Korollar 8: Weil die Drift- und Volatilitätsstrukturen stationär sind, sind auch die Funktionen A_B und $\tilde{\alpha}_S$ sowie Σ_B und $\tilde{\sigma}_S$ stationär. Damit gilt für $k \in \{1, \ldots, K\}$ und $n \in \{1, \ldots, N-1\}$

$$\begin{aligned} \mathrm{LR}_{B(\cdot,t_{k}+T_{n})}(t_{k},t_{k}+\Delta) \\ &= -\int_{t_{k}}^{t_{k}+\Delta} A_{B}(s,t_{k}+\Delta,t_{k}+T_{n}) \,\mathrm{d}s + \int_{t_{k}}^{t_{k}+\Delta} \Sigma_{B}(s,t_{k}+\Delta,t_{k}+T_{n}) \,\mathrm{d}L_{s} \\ &= -\int_{0}^{\Delta} A_{B}(s,\Delta,T_{n}) \,\mathrm{d}s + \int_{t_{k}}^{t_{k}+\Delta} \Sigma_{B}(s-t_{k},\Delta,T_{n}) \,\mathrm{d}L_{s} \end{aligned}$$

sowie

$$LR_{S(\cdot)}(t_{k}, t_{k} + \Delta) = \int_{t_{k}}^{t_{k}+\Delta} \tilde{\alpha}_{S}(s, t_{k} + \Delta) ds + \int_{t_{k}}^{t_{k}+\Delta} \tilde{\sigma}_{S}(s, t_{k} + \Delta) dL_{s}$$
$$= \int_{0}^{\Delta} \tilde{\alpha}_{S}(s, \Delta) ds + \int_{t_{k}}^{t_{k}+\Delta} \tilde{\sigma}_{S}(s - t_{k}, \Delta) dL_{s}.$$

Die Behauptung folgt nun aus Korollar 37.

3.3 Multivariate Hermite-Polynome

Ein weiteres erforderliches Hilfsmittel auf dem Weg zur Einführung der multivariaten Edgeworth-Approximation sind multivariate Hermite-Polynome. Sie werden hier zusammen mit den für die weiteren Untersuchungen relevanten Eigenschaften kurz vorgestellt.

Definition 38 (Multivariates Hermite-Polynom). Seien $N \in \mathbb{N}$, $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ ein Multiindex und $A \in \mathbb{R}^{N \times N}$. Das multivariate Hermite-Polynom $H_{\boldsymbol{\alpha}}(\cdot, A)$: $\mathbb{R}^N \to \mathbb{R}$ ist definiert durch

$$H_{\alpha}(x,A) := \exp\left(\frac{1}{2}\langle x,Ax\rangle\right) \left[\left(-\frac{\partial}{\partial x}\right)^{\alpha} \exp\left(-\frac{1}{2}\langle x,Ax\rangle\right)\right].$$
(11)

Die Beweise der nun vorgestellten Lemmata 39 und 40 sind trivial und werden ausgelassen.

Lemma 39. Seien $N \in \mathbb{N}$ und $\boldsymbol{\alpha} = (\alpha^1, \dots, \alpha^N) \in \mathbb{N}_0^N$ ein Multiindex.

Dann gilt

$$H_{\alpha}(x, E_N) = \prod_{n=1}^N He_{\alpha^n}(x^n)$$

für alle $x = (x^1, \ldots, x^N) \in \mathbb{R}^N$. Hierbei ist $E_N \in \mathbb{R}^{N \times N}$ die Einheitsmatrix und für $k \in \mathbb{N}_0$ bezeichnet He_k das k-te eindimensionale Hermite-Polynom zur Gewichtungsfunktion $\mathbb{R} \ni y \mapsto e^{-y^2/2}$, also $He_k : \mathbb{R} \to \mathbb{R}$ mit

$$He_k(y) := (-1)^k e^{y^2/2} \left(\frac{\mathrm{d}^k}{\mathrm{d}y^k} e^{-y^2/2} \right).$$

Lemma 40. Seien $N \in \mathbb{N}$, $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ ein Multiindex und $\Sigma \in \mathbb{R}^{N \times N}$ eine strikt positiv definite Matrix.

Dann gilt

$$\frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} e^{-\mathrm{i}\langle \tilde{y}, x \rangle} (\mathrm{i}\tilde{y})^{\alpha} \varphi_{\mathcal{N}(0,\Sigma)}(\tilde{y}) \,\mathrm{d}\tilde{y} = H_{\alpha}(x, \Sigma^{-1}) \phi_{\mathcal{N}(0,\Sigma)}(x)$$

für alle $x \in \mathbb{R}^N$. Hierbei bezeichnet $\mathcal{N}(0, \Sigma)$ die N-dimensionale Normalverteilung mit Erwartungswertvektor 0 und Kovarianzmatrix Σ , $\varphi_{\mathcal{N}(0,\Sigma)}$ ist die dazugehörige charakteristische Funktion und $\phi_{\mathcal{N}(0,\Sigma)}$ ist die dazugehörige Wahrscheinlichkeitsdichte.

Satz 41. Seien $N \in \mathbb{N}$, $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^N$ zwei Multiindices mit $|\boldsymbol{\alpha}| \neq |\boldsymbol{\beta}|$ und $\Sigma \in \mathbb{R}^{N \times N}$ eine strikt positiv definite Matrix.

Dann sind $H_{\alpha}(\cdot, \Sigma^{-1})$ und $H_{\beta}(\cdot, \Sigma^{-1})$, aufgefasst als Funktionen auf dem Maßraum ($\mathbb{R}^{N}, \mathcal{B}(\mathbb{R}^{N}), \mathcal{N}(0, \Sigma)$), orthogonal bezüglich des L²-Skalarprodukts, also

$$\int_{\mathbb{R}^N} H_{\alpha}(\tilde{x}, \Sigma^{-1}) H_{\beta}(\tilde{x}, \Sigma^{-1}) \phi_{\mathcal{N}(0, \Sigma)}(\tilde{x}) \, \mathrm{d}\tilde{x} = 0.$$

Beweis: Seien zunächst $\alpha, \beta \in \mathbb{N}_0^N$ zwei beliebige Multiindices. Unter Verwendung des Integraltransformationssatzes und unter Ausnutzung der Orthogonalität von eindimensionalen Hermite-Polynomen, also

$$\int_{\mathbb{R}} He_{k_1}(\tilde{y}) He_{k_2}(\tilde{y}) \phi_{\mathcal{N}(0,1)} \,\mathrm{d}\tilde{y} = \begin{cases} 0 & \text{für } k_1 \neq k_2 \\ k_1! & \text{für } k_1 = k_2 \end{cases}$$

für $k_1,k_2\in\mathbb{N}_0,$ erhält man zusammen mit Lemma 39

$$\int_{\mathbb{R}^N} H_{\boldsymbol{\alpha}}(\Sigma^{-\frac{1}{2}}\tilde{x}, E_N) H_{\boldsymbol{\beta}}(\Sigma^{-\frac{1}{2}}\tilde{x}, E_N) \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) \, \mathrm{d}\tilde{x}$$

$$= \int_{\mathbb{R}^N} H_{\boldsymbol{\alpha}}(\tilde{y}, E_N) H_{\boldsymbol{\beta}}(\tilde{y}, E_N) \phi_{\mathcal{N}(0,E_N)}(\tilde{y}) \, \mathrm{d}\tilde{y}$$

$$= \begin{cases} 0 & \text{für } \boldsymbol{\alpha} \neq \boldsymbol{\beta} \\ \boldsymbol{\alpha}! & \text{für } \boldsymbol{\alpha} = \boldsymbol{\beta} \end{cases}.$$

Eine Anwendung von Lemma 30 für $B := \Sigma^{-\frac{1}{2}} = (\Sigma_{n_1,n_2}^{-\frac{1}{2}})_{n_1,n_2 \in \{1,\dots,N\}}$ und $g(y) := \phi_{\mathcal{N}(0,E_N)}(y)$ ergibt

$$H_{\alpha}(x, \Sigma^{-1})\phi_{\mathcal{N}(0,\Sigma)}(x)$$

$$= \frac{1}{\sqrt{\det(\Sigma)}}(-1)^{|\alpha|} \left(\frac{\partial}{\partial x}\right)^{\alpha} \phi_{\mathcal{N}(0,E_{N})}(\Sigma^{-\frac{1}{2}}x)$$

$$= \frac{1}{\sqrt{\det(\Sigma)}}(-1)^{|\alpha|}$$

$$\times \sum_{n_{1},\dots,n_{K}=1}^{N} \left(\prod_{k=1}^{|\alpha|} \Sigma_{n_{k},I_{\alpha}^{k}}^{-\frac{1}{2}}\right) \left(\left(\frac{\partial}{\partial y}\right)^{\alpha_{N}(n_{1},\dots,n_{|\alpha|})} \phi_{\mathcal{N}(0,E_{N})}\right)(\Sigma^{-\frac{1}{2}}x)$$

$$= \sum_{n_{1},\dots,n_{K}=1}^{N} \left(\prod_{k=1}^{|\alpha|} \Sigma_{n_{k},I_{\alpha}^{k}}^{-\frac{1}{2}}\right) H_{\alpha_{N}(n_{1},\dots,n_{|\alpha|})}(\Sigma^{-\frac{1}{2}}x,E_{N})\phi_{\mathcal{N}(0,\Sigma)}(x).$$

Hierbei sind $I_{\alpha} = (I_{\alpha}^1, \ldots, I_{\alpha}^{|\alpha|})$ und $\alpha_N(n_1, \ldots, n_K)$ für $(n_1, \ldots, n_K) \in \{1, \ldots, N\}^K$ die in Definition 31 eingeführten Größen. Aus $\phi_{\mathcal{N}(0,\Sigma)} > 0$ folgt

$$H_{\alpha}(x, \Sigma^{-1}) = \sum_{n_1, \dots, n_K = 1}^N \left(\prod_{k=1}^{|\alpha|} \Sigma_{n_k, I_{\alpha}^k}^{-\frac{1}{2}} \right) H_{\alpha_N(n_1, \dots, n_{|\alpha|})}(\Sigma^{-\frac{1}{2}}x, E_N)$$

für alle $x \in \mathbb{R}^N$.

Sei nun $|\boldsymbol{\alpha}| \neq |\boldsymbol{\beta}|$. Dann gilt $\boldsymbol{\alpha}_N(n_1, \dots, n_{|\boldsymbol{\alpha}|}) \neq \boldsymbol{\alpha}_N(m_1, \dots, m_{|\boldsymbol{\beta}|})$ für alle $(n_1, \dots, n_{|\boldsymbol{\alpha}|}) \in \{1, \dots, N\}^{|\boldsymbol{\alpha}|}$ und $(n_1, \dots, n_{|\boldsymbol{\beta}|}) \in \{1, \dots, N\}^{|\boldsymbol{\beta}|}$. Hieraus folgt

$$\begin{split} &\int_{\mathbb{R}^N} H_{\boldsymbol{\alpha}}(\tilde{x}, \Sigma^{-1}) H_{\boldsymbol{\beta}}(\tilde{x}, \Sigma^{-1}) \phi_{\mathcal{N}(0, \Sigma)}(\tilde{x}) \, \mathrm{d}\tilde{x} \\ &= \sum_{n_1, \dots, n_{|\boldsymbol{\alpha}|}=1}^N \sum_{m_1, \dots, m_{|\boldsymbol{\beta}|}=1}^N \left(\prod_{k=1}^{|\boldsymbol{\alpha}|} \Sigma_{n_k, I_{\boldsymbol{\alpha}}^k}^{-\frac{1}{2}} \right) \left(\prod_{k=1}^{|\boldsymbol{\beta}|} \Sigma_{m_k, I_{\boldsymbol{\beta}}^k}^{-\frac{1}{2}} \right) \\ &\times \int_{\mathbb{R}^N} H_{\boldsymbol{\alpha}_N(n_1, \dots, n_{|\boldsymbol{\alpha}|})}(\Sigma^{-\frac{1}{2}} \tilde{x}, E_N) H_{\boldsymbol{\alpha}_N(m_1, \dots, m_{|\boldsymbol{\beta}|})}(\Sigma^{-\frac{1}{2}} \tilde{x}, E_N) \phi_{\mathcal{N}(0, \Sigma)}(\tilde{x}) \, \mathrm{d}\tilde{x} \\ &= 0, \end{split}$$

also die Behauptung.

3.4 Die multivariate Edgeworth-Approximation

Der vorliegende Abschnitt ist der multivariaten Edgeworth-Approximation gewidmet. Grob gesprochen beruht sie auf der Idee, einen Teil der charakteristischen Funktion eines Zufallsvektors in geeigneter Weise durch ein Taylor-Polynom zu approximieren um anschließend durch Fourier-Inversion hieraus eine approximative Dichte zu erhalten. Verglichen mit der numerischen Fourier-Inversion der (nicht approximierten) charakteristischen Funktion lässt sich so der Rechenaufwand auf einen Bruchteil reduzieren.

Die Argumentation orientiert sich an der üblichen Vorgehensweise, wie sie beispielsweise in Hall (1992) oder in Blinnikov und Moessner (1998) für den eindimensionalen Fall vorgestellt wird. Durch die zusätzliche Einbeziehung der Annahme (EM) wird diese jedoch mit Satz 46 in Bezug auf das Konvergenzverhalten der vorgenommenen Approximation präzisiert. Ferner wird zusätzlich zur Edgeworth-Approximation der Dichte und der Verteilungsfunktion auch ein Ansatz für eine Edgeworth-Approximation der bedingten Dichte und der bedingten Verteilungsfunktion eingeführt.

In Abschnitt 3.4.1 werden noch die letzten dafür erforderlichen technischen Lemmata abgehandelt, bevor dann in Abschnitt 3.4.2 schließlich die multivariate Edgeworth-Approximation präsentiert wird.

3.4.1 Technische Lemmata

Definition 42 (T_S) . Set $S \in \mathbb{N}$. Die Menge T_S von Multiindices ist definiert durch

$$T_S := \left\{ \boldsymbol{\eta} = (\eta^1, \dots, \eta^S) \in \mathbb{N}_0^S \left| \sum_{s=1}^S s \eta^s = S \right\}.$$

Lemma 43 (Folgerung aus der Formel von Faà di Bruno). Seien $R_f, R_g > 0$, $D_f := \{\tilde{z} \in \mathbb{C} | |\tilde{z}| < R_f\}, D_g := \{\tilde{z} \in \mathbb{C} | |\tilde{z}| < R_g\} \text{ und } f : D_f \to \mathbb{C},$ $g : D_g \to \mathbb{C}$ zwei holomorphe Funktionen mit $g(D_g) \subset D_f$ und g(0) = 0. Ferner seien $f(z) = \sum_{m=0}^{\infty} a_m z^m, z \in D_f, und g(z) = \sum_{m=1}^{\infty} b_m z^m, z \in D_g,$ die Potenzreihendarstellungen von f und g mit Entwicklungspunkt $z_0 = 0$.

Dann ist die Verkettung $f \circ g$ holomorph auf D_g und es gilt

$$(f \circ g)(z) = \sum_{S=0}^{\infty} c_S z^S$$

für alle $z \in D_g$. Dabei ist $c_0 = a_0$ und

$$c_{S} = \sum_{\boldsymbol{\eta} \in T_{S}} \binom{|\boldsymbol{\eta}|}{\boldsymbol{\eta}} a_{|\boldsymbol{\eta}|} \prod_{\substack{s=1,\\ \eta^{s} \ge 1}}^{S} b_{s}^{\eta^{s}}$$

für $S \in \mathbb{N}$.

Im Spezialfall $f(z) = e^z$ gilt $c_0 = 1$ und

$$c_S = \sum_{oldsymbol{\eta} \in T_S} rac{1}{oldsymbol{\eta}!} \prod_{\substack{s=1,\ \eta^s \geq 1}}^S b_s^{\eta^s}$$

für $S \in \mathbb{N}$.

Lemma 43 ist eine direkte Folgerung aus der Formel von Faà di Bruno.

Lemma 44. Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0,\infty)$ genügt. Weiterhin seien $u \in \left\{ \tilde{u} = (\tilde{u}^1, \dots, \tilde{u}^N) \in \mathbb{C}^N | |\tilde{u}^n| < M, n \in \{1, \dots, N\} \right\}$ und

$$R := \min_{n \in \{1, \dots, N\}} \frac{1}{2} \left(1 + \frac{M}{|u^n|} \right),$$

wobei hier $M/0 := \infty$ gesetzt wird. Ferner seien $U := \{\tilde{z} \in \mathbb{C} | |\tilde{z}| < R\}$ und $h: U \to \mathbb{C}$ mit

$$h(z) := \exp\left(\sum_{m=1}^{\infty} b_m z^m\right),$$
$$b_m := \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_0^N, \\ |\boldsymbol{\alpha}| = m+2}} \frac{\kappa_{\boldsymbol{\alpha}}^X}{\boldsymbol{\alpha}!} u^{\boldsymbol{\alpha}}$$
(12)

wobei

für $m \in \mathbb{N}$.

Dann gilt $R \in (1, \infty]$. Ferner ist h holomorph auf U und es gilt

$$h(z) = 1 + \sum_{S=1}^{\infty} \mathcal{P}_S z^S$$

für alle $z \in U$, wobei

$$\mathcal{P}_S := \sum_{\boldsymbol{\eta} \in T_S} \frac{1}{\boldsymbol{\eta}!} \prod_{\substack{s=1,\\ \boldsymbol{\eta}^s \ge 1}}^S b_s^{\boldsymbol{\eta}^s} \tag{13}$$

für $S \in \mathbb{N}$.

Beweis: Für u = 0 erhält man $R = \infty$. Ansonsten ist dieser Fall trivial.

Nun wird der Fall $u \neq 0$ behandelt. Direkt aus der Definition von R folgt $R \in (1, \infty)$. Weiter erhält man $\Re(Ru) \in (-M, M)^N$ und damit $K_X(Ru) \in \mathbb{C}$. Zusammen mit Lemma 24 folgt hieraus

$$\sum_{m=1}^{\infty} b_m R^m = \frac{1}{R^2} \left[\left(\sum_{k=1}^{\infty} \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_{0,n}^N \\ |\boldsymbol{\alpha}| = k}} \frac{\kappa_{\boldsymbol{\alpha}}^X}{\boldsymbol{\alpha}!} (Ru)^{\boldsymbol{\alpha}} \right) - \left(\sum_{k=1}^2 \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_{0,n}^N \\ |\boldsymbol{\alpha}| = k}} \frac{\kappa_{\boldsymbol{\alpha}}^X}{\boldsymbol{\alpha}!} u^{\boldsymbol{\alpha}} R^k \right) \right]$$
$$= \frac{1}{R^2} \left[K_X(Ru) - \left(\sum_{k=1}^2 \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_{0,n}^N \\ |\boldsymbol{\alpha}| = k}} \frac{\kappa_{\boldsymbol{\alpha}}^X}{\boldsymbol{\alpha}!} u^{\boldsymbol{\alpha}} R^k \right) \right] \in \mathbb{C}.$$

Somit ist die Reihe $\sum_{m=1}^{\infty} b_m R^m$ konvergent und der Konvergenzradius der Potenzreihe $\sum_{m=1}^{\infty} b_m z^m$ ist größer oder gleich R. Insbesondere ist die Funktion $U \ni z \mapsto \sum_{m=1}^{\infty} b_m z^m$ holomorph. Damit sind alle Voraussetzungen von Lemma 43 erfüllt und dessen Anwendung liefert die Behauptung.

Lemma 45. Seien $N, S \in \mathbb{N}$, $\eta \in T_S$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, dessen Kumulanten κ_{α}^X jeder Ordnung $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \leq S + 2$ existieren. Für $u \in \mathbb{C}^N$ und $m \in \{1, \ldots, S\}$ sei ferner b_m durch Gleichung (12) aus Lemma 44 definiert. Dann gilt

$$\prod_{\substack{s=1,\\\eta^s\geq 1}}^S b_s^{\eta^s} = \sum_{\substack{\boldsymbol{\alpha_1,\ldots,\alpha_{|\boldsymbol{\eta}|}\in\mathbb{N}_0^N,\\|\boldsymbol{\alpha_r}|=I_{\boldsymbol{\eta}}^r+2 \text{ für }r\in\{1,\ldots,|\boldsymbol{\eta}|\}}} \left(\prod_{r=1}^{|\boldsymbol{\eta}|} \frac{\kappa_{\boldsymbol{\alpha_r}}^X}{\boldsymbol{\alpha_r}!}\right) u^{\sum_{r=1}^{|\boldsymbol{\eta}|} \boldsymbol{\alpha_r}},$$

wobei $\mathbf{I}_{\boldsymbol{\eta}} = (I_{\boldsymbol{\eta}}^1, \dots, I_{\boldsymbol{\eta}}^{|\boldsymbol{\eta}|})$ das in Definition 31 eingeführte $|\boldsymbol{\eta}|$ -Tupel von Indices ist. Insbesondere ist $\prod_{s=1,\eta^s \ge 1}^{S} b_s^{\eta^s}$ ein Polynom in u vom Grad $2|\boldsymbol{\eta}| + S$. Ferner ist die in Gleichung (13) aus Lemma 44 definierte Größe \mathcal{P}_S ein Polynom in u vom Grad 3S.

Beweis: Für $u \in \mathbb{C}^N$ berechnet man

$$\begin{split} \prod_{\substack{s=1,\\\eta^s \ge 1}}^{S} b_s^{\eta^s} &= \prod_{r=1}^{|\eta|} b_{I_{\eta}^r} \\ &= \prod_{r=1}^{|\eta|} \sum_{\substack{\alpha \in \mathbb{N}_0^N,\\|\alpha| = I_{\eta}^r + 2}} \frac{\kappa_{\alpha}^X}{\alpha!} u^{\alpha} \\ &= \sum_{\substack{\alpha_1, \dots, \alpha_{|\eta|} \in \mathbb{N}_0^N,\\|\alpha_r| = I_{\eta}^r + 2 \text{ für } r \in \{1, \dots, |\eta|\}}} \prod_{r=1}^{|\eta|} \frac{\kappa_{\alpha_r}^X}{\alpha_r!} u^{\alpha_r} \\ &= \sum_{\substack{\alpha_1, \dots, \alpha_{|\eta|} \in \mathbb{N}_0^N,\\|\alpha_r| = I_{\eta}^r + 2 \text{ für } r \in \{1, \dots, |\eta|\}}} \left(\prod_{r=1}^{|\eta|} \frac{\kappa_{\alpha_r}^X}{\alpha_r!} \right) u^{\sum_{r=1}^{|\eta|} \alpha_r} \end{split}$$

Damit ist die erste Behauptung gezeigt. Aus

$$\sum_{r=1}^{|\boldsymbol{\eta}|} |\boldsymbol{\alpha}_r| = \sum_{r=1}^{|\boldsymbol{\eta}|} \left(I_{\boldsymbol{\eta}}^r + 2 \right) = 2|\boldsymbol{\eta}| + \sum_{s=1}^{S} s \eta^s = 2|\boldsymbol{\eta}| + S$$

folgt die zweite Behauptung. Mit max $\{|\boldsymbol{\eta}| | \boldsymbol{\eta} \in T_S\} = S$ und dem soeben Gezeigten erhält man schließlich auch die letzte Behauptung.

3.4.2 Einführung der multivariaten Edgeworth-Approximation

Satz 46. Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt. Ferner seien $\mathbb{E}[X] = 0$ und $\Sigma := \operatorname{Cov}(X).$

Dann gilt

$$M_X(u) = e^{\frac{1}{2}\langle u, \Sigma u \rangle} \\ \times \left(1 + \sum_{S=1}^{\infty} \sum_{\eta \in T_S} \frac{1}{\eta!} \sum_{\substack{\alpha_1, \dots, \alpha_{|\eta|} \in \mathbb{N}_0^N, \\ |\alpha_r| = I_\eta^r + 2 \text{ für } r \in \{1, \dots, |\eta|\}}} \left(\prod_{r=1}^{|\eta|} \frac{\kappa_{\alpha_r}^X}{\alpha_r!} \right) u^{\sum_{r=1}^{|\eta|} \alpha_r} \right)$$

für alle $u \in \left\{ (\tilde{u}^1, \dots, \tilde{u}^N) \in \mathbb{C}^N \middle| |\tilde{u}^n| < M, n \in \{1, \dots, N\} \right\}.$

Beweis: Sei $u \in \{(\tilde{u}^1, \ldots, \tilde{u}^N) \in \mathbb{C}^N | |\tilde{u}^n| < M, n \in \{1, \ldots, N\}\}$. Aus $\mathbb{E}[X] = 0$ folgt $\kappa_{\alpha}^X = 0$ für alle Multiindices $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| = 1$ und somit erhält man zusammen mit Lemma 24

$$K_X(u) := \sum_{k=2}^{\infty} \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_0^N, \\ |\boldsymbol{\alpha}| = k}} \frac{\kappa_{\boldsymbol{\alpha}}^X}{\boldsymbol{\alpha}!} u^{\boldsymbol{\alpha}}.$$

Aus Lemma 44 folgt für die dort definierten $b_m, m \in \mathbb{N}, \mathcal{P}_S, S \in \mathbb{N}$, und R

$$\exp\left(\sum_{m=1}^{\infty} b_m z^m\right) = 1 + \sum_{S=1}^{\infty} \mathcal{P}_S z^S$$

für alle $z \in U := \{ \tilde{z} \in \mathbb{C} | |\tilde{z}| < R \}.$ R > 1 impliziert $1 \in U$. Somit erhält man

$$\exp\left(K_X(u) - \frac{1}{2} \langle u, \Sigma u \rangle\right) = \exp\left(\sum_{k=3}^{\infty} \sum_{\substack{\alpha \in \mathbb{N}_0^N, \\ |\alpha| = k}} \frac{\kappa_{\alpha}^X}{\alpha!} u^{\alpha}\right)$$
$$= \exp\left(\sum_{m=1}^{\infty} b_m z^m\right)\Big|_{z=1}$$
$$= 1 + \sum_{S=1}^{\infty} \mathcal{P}_S z^S\Big|_{z=1}$$
$$= 1 + \sum_{S=1}^{\infty} \mathcal{P}_S$$
$$= 1 + \sum_{S=1}^{\infty} \sum_{\eta \in T_S} \frac{1}{\eta!} \prod_{\substack{s=1, \\ \eta^s \ge 1}}^S b_s^{\eta^s}.$$

Aus $M_X(u) = e^{\frac{1}{2}\langle u, \Sigma u \rangle} \exp\left(K_X(u) - \frac{1}{2}\langle u, \Sigma u \rangle\right)$ folgt zusammen mit Lemma 45 die Behauptung.

Im Folgenden wird angenommen, dass die Voraussetzungen und die Bezeichnungen von Satz 46 gelten. Weiterhin sei Σ strikt positiv definit und die charakteristische Funktion φ_X von X sei integrierbar. Ferner sei $S_{max} \in \mathbb{N}$ fest vorgegeben.

Satz 46 impliziert, dass

$$\begin{aligned} \varphi_X(y) &= M_X(\mathrm{i}y) \\ &= e^{-\frac{1}{2}\langle y, \Sigma y \rangle} \\ &\times \left(1 + \sum_{S=1}^{\infty} \sum_{\eta \in T_S} \frac{1}{\eta!} \sum_{\substack{\alpha_1, \dots, \alpha_{|\eta|} \in \mathbb{N}_0^N, \\ |\alpha_r| = I_{\eta}^r + 2 \text{ für } r \in \{1, \dots, |\eta|\}} \left(\prod_{r=1}^{|\eta|} \frac{\kappa_{\alpha_r}^X}{\alpha_r!} \right) (\mathrm{i}y)^{\sum_{r=1}^{|\eta|} \alpha_r} \right) \end{aligned}$$

für alle $y \in (-M, M)^N$ gilt. Bricht man diese unendliche Reihe nach S_{max} Summanden ab und verwendet den so erhaltenen Funktionsausdruck für alle $y \in \mathbb{R}^N$, so erhält man die Funktion $\tilde{\varphi}_X^{S_{max}} : \mathbb{R}^N \to \mathbb{C}$ mit

$$\begin{split} \tilde{\varphi}_X^{S_{max}}(y) &= e^{-\frac{1}{2}\langle y, \Sigma y \rangle} \\ & \times \left(1 + \sum_{S=1}^{S_{max}} \sum_{\eta \in T_S} \frac{1}{\eta!} \sum_{\substack{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{|\eta|} \in \mathbb{N}_0^N, \\ |\boldsymbol{\alpha}_r| = I_{\eta}^r + 2 \text{ für } r \in \{1, \dots, |\eta|\}}} \left(\prod_{r=1}^{|\eta|} \frac{\kappa_{\boldsymbol{\alpha}_r}^X}{\boldsymbol{\alpha}_r!} \right) (\mathrm{i}y)^{\sum_{r=1}^{|\eta|} \boldsymbol{\alpha}_r} \right). \end{split}$$

Die Kernidee der Edgeworth-Approximation ist nun, die charakteristische Funktion φ_X von X auf ganz \mathbb{R}^N durch die Funktion $\tilde{\varphi}_X^{S_{max}}$ zu approximieren und hieraus durch Fourier-Inversion eine approximative Dichte zu berechnen.

Unter Verwendung von Lemma 40 erhält man für $x \in \mathbb{R}^N$

$$f_{X}(x) = \frac{1}{(2\pi)^{N}} \int_{\mathbb{R}^{N}} e^{-i\langle \tilde{y}, x \rangle} \varphi_{X}(\tilde{y}) d\tilde{y}$$

$$\approx \frac{1}{(2\pi)^{N}} \int_{\mathbb{R}^{N}} e^{-i\langle \tilde{y}, x \rangle} \tilde{\varphi}_{X}^{S_{max}}(\tilde{y}) d\tilde{y}$$

$$= \phi_{\mathcal{N}(0,\Sigma)}(x) + \sum_{S=1}^{S_{max}} \sum_{\eta \in T_{S}} \frac{1}{\eta!} \sum_{\substack{\boldsymbol{\alpha}_{1}, \dots, \boldsymbol{\alpha}_{|\eta|} \in \mathbb{N}_{0}^{N}, \\ |\boldsymbol{\alpha}_{r}| = I_{\eta}^{r} + 2 \text{ für } r \in \{1, \dots, |\eta|\}}} \left(\prod_{r=1}^{|\eta|} \frac{\kappa_{\boldsymbol{\alpha}_{r}}^{X}}{\boldsymbol{\alpha}_{r}!} \right)$$

$$\times \frac{1}{(2\pi)^{N}} \int_{\mathbb{R}^{N}} e^{-i\langle \tilde{y}, x \rangle} (i\tilde{y})^{\sum_{r=1}^{|\eta|} \boldsymbol{\alpha}_{r}} \varphi_{\mathcal{N}(0,\Sigma)}(\tilde{y}) d\tilde{y}$$

$$= \phi_{\mathcal{N}(0,\Sigma)}(x) \left(1 + \sum_{S=1}^{S_{max}} \mathcal{K}_{X}^{S}(x) \right).$$

Wie bereits in Lemma 40 bezeichnet hierbei $\mathcal{N}(0, \Sigma)$ die *N*-dimensionale Normalverteilung mit Erwartungswertvektor 0 und Kovarianzmatrix Σ , $\varphi_{\mathcal{N}(0,\Sigma)}$ ist die dazugehörige charakteristische Funktion und $\phi_{\mathcal{N}(0,\Sigma)}$ ist die dazugehörige Wahrscheinlichkeitsdichte. Ferner wurde hier bereits die Abkürzung $\mathcal{K}_X^S(x)$ verwendet, die erst in Gleichung (14) definiert wird.

Als Ergebnis der vorstehenden Approximationsüberlegungen wird die folgende Definition eingeführt.

Definition 47 (Multivariate Edgeworth-Approximation). Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt. Weiterhin sei $\mathbb{E}[X] = 0$ und die Kovarianzmatrix $\Sigma := \operatorname{Cov}(X)$ sei strikt positiv definit. Ferner seien $n \in \{2, \ldots, N\}, m \in \{1, \ldots, n-1\}$ und $S_{max} \in \mathbb{N}_0$.

1. Für $S \in \mathbb{N}$ ist der S-te Edgeworth-Korrekturterm definiert als die Funktion $\mathcal{K}_X^S : \mathbb{R}^N \to \mathbb{R}$ mit

$$\mathcal{K}_X^S(x) := \sum_{\boldsymbol{\eta} \in T_S} \frac{1}{\boldsymbol{\eta}!} \sum_{\substack{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{|\boldsymbol{\eta}|} \in \mathbb{N}_0^N, \\ |\boldsymbol{\alpha}_r| = I_{\boldsymbol{\eta}}^r + 2 \text{ für } r \in \{1, \dots, |\boldsymbol{\eta}|\}}} \left(\prod_{r=1}^{|\boldsymbol{\eta}|} \frac{\kappa_{\boldsymbol{\alpha}_r}^N}{\boldsymbol{\alpha}_r!} \right) H_{\sum_{r=1}^{|\boldsymbol{\eta}|} \boldsymbol{\alpha}_r}(x, \Sigma^{-1}).$$
(14)

2. Die Edgeworth-Approximation $f_X^{S_{max}} : \mathbb{R}^N \to \mathbb{R}$ der Dichte von X mit

 S_{max} Korrekturtermen ist definiert als

$$f_X^{S_{max}}(x) := \phi_{\mathcal{N}(0,\Sigma)}(x) \left(1 + \sum_{S=1}^{S_{max}} \mathcal{K}_X^S(x) \right).$$

3. Die Edgeworth-Approximation $F_X^{S_{max}} : \mathbb{R}^N \to \mathbb{R}$ der Verteilungsfunktion von X mit S_{max} Korrekturtermen ist definiert als

$$F_X^{S_{max}}(x) := \int_{(-\infty,x]} f_X^{S_{max}}(\tilde{x}) \,\mathrm{d}\tilde{x}.$$

4. Die Edgeworth-Approximation $f_{X^1,...,X^m|X^{m+1},...,X^n}^{S_{max}}$: $\mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}$ der bedingten Dichte von $(X^1,...,X^m)$ gegeben $(X^{m+1},...,X^n)$ mit S_{max} Korrekturtermen ist definiert als

$$f_{X^{1},\dots,X^{m}|X^{m+1},\dots,X^{n}}^{S_{max}}\left(x^{1},\dots,x^{m}|x^{m+1},\dots,x^{n}\right)$$
$$:=\frac{f_{(X^{1},\dots,X^{n})}^{S_{max}}\left(x^{1},\dots,x^{n}\right)}{f_{(X^{m+1},\dots,X^{n})}^{S_{max}}\left(x^{m+1},\dots,x^{n}\right)}.$$

5. Die Edgeworth-Approximation $F_{X^1,...,X^m|X^{m+1},...,X^n}^{S_{max}}$: $\mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}$ der bedingten Verteilungsfunktion von $(X^1,...,X^m)$ gegeben $(X^{m+1},...,X^n)$ mit S_{max} Korrekturtermen ist definiert als

$$F_{X^{1},\dots,X^{m}|X^{m+1},\dots,X^{n}}^{S_{max}}\left(x^{1},\dots,x^{m}|x^{m+1},\dots,x^{n}\right) := \int_{-\infty}^{x^{1}}\dots\int_{-\infty}^{x^{m}} f_{X^{1},\dots,X^{m}|X^{m+1},\dots,X^{n}}^{S_{max}}\left(\tilde{x}^{1},\dots,\tilde{x}^{m}|x^{m+1},\dots,x^{n}\right) \,\mathrm{d}\tilde{x}^{m}\dots\,\mathrm{d}\tilde{x}^{1}.$$

Bemerkung 48. Ist X multivariat normalverteilt, so gilt $\kappa_{\alpha}^{X} = 0$ für alle $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \geq 3$. In diesem Fall gilt $\mathcal{K}_{S} = 0$ für alle $S \in \mathbb{N}$ und die Edgeworth-Approximation der (bedingten) Dichte bzw. der (bedingten) Verteilungsfunktion fällt mit der entsprechenden exakten Größe zusammen.

Bemerkung 49. Zur Definition der Edgeworth-Approximation $f_X^{S_{max}}$ der Dichte von X mit S_{max} Korrekturtermen ist es nicht unbedingt erforderlich, dass X der Annahme ($\mathbb{E}M$) zu einer gewissen strikt positiven Konstante genügt. Der Ausdruck für $f_X^{S_{max}}$ ist bereits dann wohldefiniert, wenn die Kumulanten κ_{α}^X von X für alle Ordnungen $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| \leq S_{max} + 2$ existieren. Falls X der Annahme ($\mathbb{E}\mathbb{M}$) zur Konstante $M \in (0, \infty)$ genügt, so impliziert Satz 46, dass die Folge $(\tilde{\varphi}_X^{S_{max}})_{S_{max}\in\mathbb{N}}$ auf jeder kompakten Teilmenge von $(-M, M)^N$ gleichmäßig gegen die charakteristische Funktion φ_X von X konvergiert. Somit stellt $\tilde{\varphi}_X^{S_{max}}$ für ein hinreichend groß gewähltes S_{max} zumindest in der Nähe des Ursprungs eine zufriedenstellende Approximation für φ_X dar. Falls X jedoch nicht der Annahme ($\mathbb{E}\mathbb{M}$) genügt, so ist nicht ohne Weiteres klar, ob eine Approximation von φ_X durch $\tilde{\varphi}_X^{S_{max}}$ außerhalb des Ursprungs überhaupt zu einer zufriedenstellenden Näherung führt. Und weil $f_X^{S_{max}}$ durch Fourier-Inversion aus $\tilde{\varphi}_X^{S_{max}}$ entsteht, ist die Güte der Approximation von φ_X durch $\tilde{\varphi}_X^{S_{max}}$ essentiell für die Güte der gesamten Edgeworth-Approximation. Deshalb wurde in der vorliegenden Arbeit die Annahme ($\mathbb{E}\mathbb{M}$) in die Voraussetzungen von Definition 47 mit aufgenommen.

Bemerkung 50. Weiter unten in Abschnitt 6.4 wird sich zeigen, dass die Edgeworth-Approximation für finanzmathematische Anwendungen, die nicht auf einer Normalverteilungshypothese basieren, ein fruchtbares Konzept ist. Im Allgemeinen ist $f_X^{S_{max}}$ jedoch keine echte Wahrscheinlichkeitsdichte. Blinnikov und Moessner (1998) geben Beispiele, in denen $f_X^{S_{max}}$ auch negative Werte annimmt.

Satz 51. Unter den Voraussetzungen und mit den Bezeichnungen von Definition 47 gilt

$$\int_{\mathbb{R}^N} f_X^{S_{max}}(\tilde{x}) \,\mathrm{d}\tilde{x} = 1.$$

Falls also $f_X^{S_{max}}(x) \ge 0$ für alle $x \in \mathbb{R}^N$ vorliegt, so ist $f_X^{S_{max}}$ eine Wahrscheinlichkeitsdichte.

Beweis: Für den Multiindex $\boldsymbol{\beta} := (0, \dots, 0) \in \mathbb{N}_0^N$ gilt $H_{\boldsymbol{\beta}}(x, \Sigma^{-1}) = 1$. Ferner treten auf der rechten Seite der Definitionsgleichung (14) für den *S*-ten Edgeworth-Korrekturterm ausschließlich multivariate Hermite-Polynome $H_{\boldsymbol{\alpha}}(\cdot, \Sigma^{-1})$ zu Multiindices $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| \neq |\boldsymbol{\beta}| = 0$ auf. Somit impliziert Satz 41

$$\begin{split} \int_{\mathbb{R}^{N}} \mathcal{K}_{X}^{S}(\tilde{x}) \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) \, \mathrm{d}\tilde{x} \\ &= \sum_{\eta \in T_{S}} \frac{1}{\eta!} \sum_{\substack{\boldsymbol{\alpha}_{1}, \dots, \boldsymbol{\alpha}_{|\eta|} \in \mathbb{N}_{0}^{N}, \\ |\boldsymbol{\alpha}_{r}| = I_{\eta}^{r} + 2 \text{ für } r \in \{1, \dots, |\eta|\}} \left(\prod_{r=1}^{|\eta|} \frac{\kappa_{\boldsymbol{\alpha}_{r}}^{X}}{\boldsymbol{\alpha}_{r}!} \right) \\ &\times \int_{\mathbb{R}^{N}} H_{\sum_{r=1}^{|\eta|} \boldsymbol{\alpha}_{r}}(x, \Sigma^{-1}) H_{\boldsymbol{\beta}}(x, \Sigma^{-1}) \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) \, \mathrm{d}\tilde{x} \\ &= 0 \end{split}$$

für alle $S\in\mathbb{N}.$ Hieraus folgt die Behauptung.

58

4 Eine multivariate Erweiterung des Cramérvon-Mises- und des Anderson-Darling-Tests

Gegeben sei eine Stichprobe von mehreren stochastisch unabhängigen und identisch verteilten Realisierungen einer Zufallsgröße und es sei auf der Grundlage eines statistischen Tests zu prüfen, ob die Beobachtungen einer angenommenen Wahrscheinlichkeitsverteilung folgen.

Für univariate Beobachtungen sind viele statistische Tests für Fragestellungen dieser Art bekannt und hervorragend erforscht. Für multivariate Beobachtungen sind einige Tests auf eine multivariate Normalverteilung bekannt. Für den Fall, dass die angenommene multivariate Wahrscheinlichkeitsverteilung keine Normalverteilung ist, sind jedoch erst wenige statistische Tests bekannt und etabliert.

Im vorliegenden Abschnitt 4 werden neue Verfahren für diese Fragestellung vorgestellt und untersucht. In Abschnitt 4.1 wird ein Anpassungstest für \mathbb{R}^N -wertige Zufallsvektoren eingeführt. Abschnitt 4.2 ist einem Anpassungstest für Copulas gewidmet, wobei ein besonderes Augenmerk auf das Konvergenzverhalten der Teststatistik für immer größer werdende Stichprobengrößen gerichtet ist. In Abschnitt 4.3 werden schließlich noch für zwei Spezialfälle geschlossene Formen zur Auswertung der Teststatistik hergeleitet.

Die Ergebnisse werden in Abschnitt 6 eingesetzt, in dem das Modell aus Abschnitt 2 unter dem physikalischen Maß kalibriert wird. Dort wird die Anpassungsgüte des kalibrierten Modells an die historischen Marktdaten untersucht. Ferner werden auf der Basis der hier vorgestellten Verfahren zwei neue Kalibrierungskriterien eingeführt.

4.1 Ein Anpassungstest für \mathbb{R}^N -wertige Zufallsvektoren

Im vorliegenden Abschnitt wird eine allgemeine Teststatistik eingeführt, die eine multivariate Erweiterung der Cramér-von-Mises-Teststatistik und der Anderson-Darling-Teststatistik darstellt.

Im Anschluss werden zwei Spezialfälle hiervon, die kanonische multivariate Erweiterung und die multivariate Rosenblatt-Erweiterung, untersucht. Definition und Satz 54 sowie Satz 60 implizieren, dass sich in diesen beiden Fällen das Testproblem derart transformieren lässt, dass es sich um Spezialfälle des im darauffolgenden Abschnitt 4.2 behandelten Falls für Copulas handelt.

Der Abschnitt schließt mit einer Diskussion von einigen für die Anwendung relevanten Eigenschaften der kanonischen multivariaten Erweiterung und der multivariaten Rosenblatt-Erweiterung.

4.1.1 Das Szenario

Das nun vorgestellte Szenario gelte im gesamten Abschnitt 4.1 in dem Sinne, dass die folgenden Annahmen und Bezeichnungen ohne weitere Erwähnung in jeder Definition, jedem Lemma und jedem Satz als Voraussetzung verwendet werden.

Sei $N \in \mathbb{N}$ und X_0 ein \mathbb{R}^N -wertiger Zufallsvektor, dessen Verteilung eine stetige Lebesgue-Dichte $f_0 := f_{X_0}$ mit $f_0(x) \in (0, \infty)$ für alle $x \in \mathbb{R}^N$ besitzt. $F_0 := F_{X_0}$ bezeichne die Verteilungsfunktion von X_0 . Ferner sei $(X_k)_{k \in \mathbb{N}}$ eine Folge von stochastisch unabhängigen und identisch verteilten \mathbb{R}^N -wertigen Zufallsvektoren. Für eine gegebene Stichprobengröße $K \in \mathbb{N}$ sei die

Nullhypothese
$$H_0$$
: $F_{X_k} = F_0$ für $k \in \mathbb{N}$.

gegen die

Alternative
$$H_1$$
: Es existiert eine Verteilungsfunktion
 $F_1 : \mathbb{R}^N \to [0,1] \text{ mit } F_1 \neq F_0 \text{ und}$
 $F_{X_k} = F_1 \text{ für } k \in \mathbb{N}.$

zu testen.

Im Fall N = 1 hat sich die Verwendung der Teststatistik

$$K \int_{\mathbb{R}} \frac{\left[\frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{X_k \le \tilde{x}\}} - F_0(\tilde{x})\right]^2}{\left[F_0(\tilde{x}) \left(1 - F_0(\tilde{x})\right)\right]^{\mu}} \, \mathrm{d}F_0(\tilde{x}) \tag{15}$$

mit $\mu \in \{0, 1\}$ bewährt, siehe hierzu die bahnbrechenden Arbeiten von Anderson und Darling (1952, 1954). Dabei erhält man für $\mu = 0$ die Cramérvon-Mises-Teststatistik und für $\mu = 1$ erhält man die Anderson-Darling-Teststatistik.

In der vorliegenden Arbeit wird nun als möglicher Ansatz für eine multivariate Erweiterung des Cramér-von-Mises-Tests bzw. des Anderson-Darling-Tests die Teststatistik

$$A_{T,K,\mu}^{2} := K \int_{\mathbb{R}^{N}} \frac{\left[\mathbb{F}_{T,K}(T(\tilde{x})) - F_{T(X_{0})}(T(\tilde{x}))\right]^{2}}{\left[F_{T(X_{0})}(T(\tilde{x}))\left(1 - F_{T(X_{0})}(T(\tilde{x}))\right)\right]^{\mu}} \,\mathrm{d}F_{0}(\tilde{x})$$
(16)

betrachtet. Hierbei seien $\mu \in \mathbb{R}_+$ und $T : \mathbb{R}^N \to \mathbb{R}^N$ eine Transformation derart, dass die Abbildung $T' : \mathbb{R}^N \to T(\mathbb{R}^N), T'(x) = T(x)$, mit eingeschränkter Zielmenge ein Diffeomorphismus ist. Ferner bezeichne $\mathbb{F}_{T,K} : \mathbb{R}^N \to [0,1],$

$$\mathbb{F}_{T,K}(u) := \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{T(X_k) \le u\}},$$

die empirische Verteilungsfunktion der Stichprobe, die aus den ersten K Beobachtungen der Folge $(T(X_k))_{k\in\mathbb{N}}$ besteht und $F_{T(X_0)}$ bezeichne die Verteilungsfunktion von $T(X_0)$.

Der folgende Satz impliziert, dass es sich bei $A_{T,K,\mu}^2$ tatsächlich um eine multivariate Erweiterung der univariaten Cramér-von-Mises-Teststatistik bzw. der univariaten Anderson-Darling-Teststatistik handelt.

Satz 52. Sei N = 1. Dann stimmt für $\mu = 0$ die in Gleichung (16) definierte Teststatistik $A_{T,K,\mu}^2$ mit der des univariaten Cramér-von-Mises-Tests überein. Für $\mu = 1$ stimmt sie mit der des univariaten Anderson-Darling-Tests überein.

Beweis: Als eindimensionaler Diffeomorphismus ist T' streng monoton und damit ist trivialerweise auch T streng monoton. Der Fall für ein streng monoton wachsendes T ist trivial. Falls T streng monoton fallend ist, gilt $\{T(X_k) \leq T(x)\} = \{X_k \geq x\}$ für alle $x \in \mathbb{R}$ und $k \in \mathbb{N}_0$. Hieraus folgt

$$\mathbb{F}_{T,K}(T(x)) = 1 - \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{X_k \le x\}} + \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{X_k = x\}}$$

und

$$F_{T(X_0)}(T(x)) = 1 - F_0(x) + \mathbb{P}\left(\{X_0 = x\}\right) = 1 - F_0(x)$$

für alle $x \in \mathbb{R}$. Für $\omega \in \Omega$ erhält man

$$\begin{aligned} &A_{T,K,\mu}^{2}(\omega) \\ &= K \int_{\mathbb{R}} \frac{\left[\frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{X_{k} \leq \tilde{x}\}}(\omega) - \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{X_{k} = \tilde{x}\}}(\omega) - F_{0}(\tilde{x})\right]^{2}}{\left[(1 - F_{0}(\tilde{x})) F_{0}(\tilde{x})\right]^{\mu}} f_{X_{0}}(\tilde{x}) \,\mathrm{d}\tilde{x} \\ &= K \int_{\mathbb{R}} \frac{\left[\frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{X_{k} \leq \tilde{x}\}}(\omega) - F_{0}(\tilde{x})\right]^{2}}{\left[F_{0}(\tilde{x}) \left(1 - F_{0}(\tilde{x})\right)\right]^{\mu}} \,\mathrm{d}F_{0}(\tilde{x}). \end{aligned}$$

Hierbei wurde beim zweiten Umformungsschritt der Integrand auf der Lebesgue-Nullmenge $\{X_1(\omega), \ldots, X_K(\omega)\} \subset \mathbb{R}$ abgeändert. \Box

Bemerkung 53. Die wählbaren Komponenten für die in Gleichung (16) definierte Teststatistik $A_{T,K,\mu}^2$ sind $\mu \in \mathbb{R}_+$ und die Transformation T. Bei ihrer konkreten Wahl sind die folgenden drei Kriterien von zentraler Bedeutung und gegeneinander abzuwägen.

- 1. Der Test soll in den gewünschten Bereichen trennscharf sein. Welche Bereiche das sind, hängt vom Anwendungsfall ab.
- 2. Die Teststatistik soll für eine gegebene Stichprobe X_1, \ldots, X_K mit vertretbarem Aufwand auswertbar sein.
- 3. Die relevanten Werte der Verteilungsfunktion der Teststatistik sollen mit vertretbarem Aufwand berechenbar sein.

4.1.2 Die kanonische multivariate Erweiterung

Es wird nun der Spezialfall der kanonischen multivariaten Erweiterung betrachtet. Wie bereits in der Einleitung angekündigt, wird in Definition und Satz 54 nachgewiesen, dass sich in diesem Fall das betrachtete Testproblem derart transformieren lässt, dass es ein Spezialfall des in Abschnitt 4.2 behandelten Falls für Copulas ist.

Definition und Satz 54 (Kanonische multivariate Erweiterung). Für jedes $n \in \{1, ..., N\}$ sei die Lebesgue-Dichte $f_{X_0^n}$ der dazugehörigen eindimensionalen Randverteilung stetig.

Sowohl für $T = id_{\mathbb{R}^N}$ als auch für T = U, wobei $U : \mathbb{R}^N \to \mathbb{R}^N$ mit $U^n(x) := F_{X_0^n}(x^n)$ für $n \in \{1, \ldots, N\}$, heißt die in Gleichung (16) definierte Teststatistik $A_{T,K,\mu}^2$ die kanonische multivariate Erweiterung der univariaten Teststatistik in Gleichung (15).

Sei ferner $C_0 : [0,1]^N \to [0,1]$ die eindeutig bestimmte, zu F_0 dazugehörige Copula.

Dann gelten die folgenden Aussagen.

- 1. $C_0(u) \in (0,1)$ für alle $u \in (0,1)^N$.
- 2. $(U(X_k))_{k\in\mathbb{N}}$ ist eine Folge von stochastisch unabhängigen und identisch verteilten $[0,1]^N$ -wertigen Zufallsvektoren. Falls die Nullhypothese H_0 wahr ist, so gilt $F_{U(X_k)} = C_0$ für alle $k \in \mathbb{N}$. Falls die Alternative H_1 wahr ist, so existiert eine Verteilungsfunktion $C_1 : [0,1]^N \to [0,1]$ mit $C_1 \neq C_0$ und $F_{U(X_k)} = C_1$ für alle $k \in \mathbb{N}$.
- 3. Sowohl für $T = id_{\mathbb{R}^N}$ als auch für T = U gilt

$$A_{T,K,\mu}^{2} = K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{U,K}(\tilde{u}) - C_{0}(\tilde{u})\right]^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu}} \,\mathrm{d}C_{0}(\tilde{u}).$$
(17)

Beweis: Zunächst ist nachzuweisen, dass $\mathrm{id}_{\mathbb{R}^N}$ und U als Transformation T verwendet werden dürfen. Trivialerweise ist $\mathrm{id}_{\mathbb{R}^N}$ ein Diffeomorphismus. Weil f_0 stetig und strikt positiv auf ganz \mathbb{R}^N ist und weil alle eindimensionalen Randverteilungen von X_0 stetige Lebesgue-Dichten besitzen, ist auch $U': \mathbb{R}^N \to (0,1)^N$ mit U'(x) := U(x) ein Diffeomorphismus. Ferner gilt

$$|\det(J_U(x))| = \prod_{n=1}^N f_{X_0^n}(x^n)$$

für alle $x \in \mathbb{R}^N$. Hierbei bezeichnet J_U die Jacobi-Matrix von U.

Für $u \in (0,1)^N$ gilt $C_0(u) = F_0\left(F_{X_0^1}^{-1}(u^1), \dots, F_{X_0^N}^{-1}(u^N)\right) \in (0,1)$. Damit ist Teilaussage 1 nachgewiesen.

Aus der Messbarkeit von U und aus $U(\mathbb{R}^N) = (0,1)^N$ folgt, dass $(U(X_k))_{k\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen und identisch verteilten $[0,1]^N$ -wertigen Zufallsvektoren ist.

Falls die Nullhypothese H_0 wahr ist, so folgt $F_{U(X_k)} = C_0$ für $k \in \mathbb{N}$ direkt aus der Definition von C_0 .

Falls die Alternative H_1 wahr ist, so existiert eine Verteilungsfunktion $F_1 : \mathbb{R}^N \to [0, 1]$ mit $F_1 \neq F_0$ und $F_{X_k} = F_1$ für $k \in \mathbb{N}$. Definiere $C_1 := F_{U(X_1)}$. Falls $C_1 = C_0$, so gilt $F_1(x) = (C_1 \circ U)(x) = (C_0 \circ U)(x) = F_0(x)$ für alle $x \in \mathbb{R}^N$. Das ist aber ein Widerspruch zu $F_1 \neq F_0$. Also gilt $C_1 \neq C_0$ und Teilaussage 2 ist bewiesen.

Weil die Verteilung von X_0 eine Lebesgue-Dichte besitzt, besitzt C_0 ebenfalls eine Lebesgue-Dichte c_0 . Diese hat die Eigenschaft

$$f_0(x) = c_0 (U(x)) \prod_{n=1}^N f_{X_0^n}(x^n)$$

für alle $x \in \mathbb{R}^N$. Ferner gilt $F_{\mathrm{id}_{\mathbb{R}^N}(X_0)}(\mathrm{id}_{\mathbb{R}^N}(x)) = F_0(x) = F_{U(X_0)}(U(x)) = C_0(U(x))$ und $\mathbb{F}_{\mathrm{id}_{\mathbb{R}^N},K}(\mathrm{id}_{\mathbb{R}^N}(x)) = \mathbb{F}_{U,K}(U(x))$ für alle $x \in \mathbb{R}^N$. Hieraus folgt unter Verwendung des Integraltransformationssatzes

$$\begin{split} &K \int_{\mathbb{R}^{N}} \frac{\left[\mathbb{F}_{\mathrm{id}_{\mathbb{R}^{N}},K}(\mathrm{id}_{\mathbb{R}^{N}}(\tilde{x})) - F_{\mathrm{id}_{\mathbb{R}^{N}}(X_{0})}(\mathrm{id}_{\mathbb{R}^{N}}(\tilde{x})) \right]^{2}}{\left[F_{\mathrm{id}_{\mathbb{R}^{N}}(X_{0})}(\mathrm{id}_{\mathbb{R}^{N}}(\tilde{x})) \left(1 - F_{\mathrm{id}_{\mathbb{R}^{N}}(X_{0})}(\mathrm{id}_{\mathbb{R}^{N}}(\tilde{x})) \right) \right]^{\mu}} \, \mathrm{d}F_{0}(\tilde{x}) \\ &= K \int_{\mathbb{R}^{N}} \frac{\left[\mathbb{F}_{U,K}\left(U(\tilde{x}) \right) - F_{U(X_{0})}\left(U(\tilde{x}) \right) \right]^{2}}{\left[F_{U(X_{0})}\left(U(\tilde{x}) \right) \left(1 - F_{U(X_{0})}\left(U(\tilde{x}) \right) \right) \right]^{\mu}} \, \mathrm{d}F_{0}(\tilde{x}) \\ &= K \int_{\mathbb{R}^{N}} \frac{\left[\mathbb{F}_{U,K}\left(U(\tilde{x}) \right) - C_{0}\left(U(\tilde{x}) \right) \right]^{2}}{\left[C_{0}\left(U(\tilde{x}) \right) \left(1 - C_{0}\left(U(\tilde{x}) \right) \right) \right]^{\mu}} c_{0}\left(U(\tilde{x}) \right) \prod_{n=1}^{N} f_{X_{0}^{n}}(\tilde{x}^{n}) \, \mathrm{d}\tilde{x} \\ &= K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{U,K}(\tilde{u}) - C_{0}(\tilde{u}) \right]^{2}}{\left[C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right) \right]^{\mu}} \, \mathrm{d}C_{0}(\tilde{u}). \end{split}$$

Damit ist auch Teilaussage 3 nachgewiesen.

Bemerkung 55. Die Zusatzvoraussetzung in Definition und Satz 54, dass die eindimensionalen Randverteilungen von X_0 stetige Lebesgue-Dichten besitzen, ist nicht überflüssig, wie das folgende Beispiel zeigt. Für N = 2 und $f_0 : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f_0(x,y) := \exp\left(-2\left[\pi x e^{\frac{1}{2}y^2}\right]^2\right)$$

gilt $\int_{\mathbb{R}^2} f_0(\tilde{x}, \tilde{y}) d\tilde{x} d\tilde{y} = 1$. Ferner ist f_0 stetig und strikt positiv und erfüllt damit alle Voraussetzungen, die in Abschnitt 4.1.1 an die Lebesgue-Dichte

der Verteilung von X_0 gestellt werden. Die eindimensionale Randverteilung der ersten Komponente hat jedoch die Lebesgue-Dichte

$$\mathbb{R} \ni x \mapsto f_{X_0^1}(x) = \int_{\mathbb{R}} f_0(x, \tilde{y}) \,\mathrm{d}\tilde{y} = \int_{\mathbb{R}} \exp\left(-2\left[\pi \, x \, e^{\frac{1}{2}\tilde{y}^2}\right]^2\right) \,\mathrm{d}\tilde{y} \in \overline{\mathbb{R}}.$$

Diese ist an der Stelle $x_0 = 0$ nicht endlich und damit insbesondere auch nicht stetig bezüglich der Standardtopologie auf \mathbb{R} . Hieraus folgt, dass die Abbildung U' mit eingeschränkter Zielmenge kein Diffeomorphismus ist.

Derartige technische Schwierigkeiten können umgangen werden, indem man die Anforderungen an die Transformation T lockert und den Integraltransformationssatz in einer Version mit schwächeren Voraussetzungen verwendet. Eine derartige Verallgemeinerung geht jedoch über den Rahmen der vorliegenden Arbeit hinaus und wird hier nicht weiter untersucht.

4.1.3 Die multivariate Rosenblatt-Erweiterung

/ \

Die nun betrachtete multivariate Rosenblatt-Erweiterung basiert auf der Idee, eine Transformation T zu verwenden, für die die Verteilung der transformierten Zufallsvektoren $T(X_k)$, $k \in \mathbb{N}$, möglichst einfach ist. Wie bereits angekündigt, gelingt mit Satz 60 auch für diesen Fall der Nachweis, dass sich das Testproblem derart transformieren lässt, dass es ein Spezialfall des in Abschnitt 4.2 behandelten Falls für Copulas ist.

Definition 56 (Rosenblatt-Transformation). Set σ eine Permutation von $\{1, \ldots, N\}$.

Die Rosenblatt-Transformation $R_{\sigma} = (R^1_{\sigma}, \dots, R^N_{\sigma}) : \mathbb{R}^N \to [0, 1]^N$ zur Permutation σ ist definiert durch

Dabei ist $F_{X_0^{\sigma(1)}}$ die Verteilungsfunktion von $X_0^{\sigma(1)}$ und für $n \in \{2, ..., N\}$ ist $F_{X_0^{\sigma(n)}|X_0^{\sigma(1)},...,X_0^{\sigma(n-1)}}$ die bedingte Verteilungsfunktion von $X_0^{\sigma(n)}$ gegeben $(X_0^{\sigma(1)},...,X_0^{\sigma(n-1)})$.

Anmerkung: Die Bezeichnungen R_{σ} wurden zu Ehren von Murray Rosenblatt gewählt.

Der nun folgende Satz geht auf Rosenblatt (1952) zurück, der ihn für den Spezialfall $\sigma = \mathrm{id}_{\{1,\ldots,N\}}$ beweist. Die Verallgemeinerung auf den hier vorgestellten Fall für beliebige Permutationen σ von $\{1,\ldots,N\}$ folgt hieraus durch Umbenennung der Indices. Der Satz impliziert, dass vermöge der Rosenblatt-Transformation R_{σ} ein Zufallsvektor $R_{\sigma}(X_0)$ entsteht, dessen Verteilung sehr einfach ist. Diese Verteilung ist unabhängig von der gewählten Permutation σ und sogar auch unabhängig von F_0 , vergleiche hierzu auch die Diskussion in Abschnitt 4.1.4.

Satz 57. Sei σ eine Permutation von $\{1, \ldots, N\}$ und R_{σ} die dazugehörige Rosenblatt-Transformation. Dann ist der Zufallsvektor $R_{\sigma}(X_0)$ gleichverteilt auf $[0, 1]^N$.

Lemma 58. Sei σ eine Permutation von $\{1, \ldots, N\}$ und R_{σ} die dazugehörige Rosenblatt-Transformation. Ferner habe die Lebesgue-Dichte f_0 der Verteilung von X_0 die folgenden zusätzlichen Eigenschaften.

- 1. f_0 ist stetig partiell differenzierbar.
- 2. Für jedes $n \in \{1, \ldots, N\}$ gilt

$$f_{\left(X_0^{\sigma(1)},\dots,X_0^{\sigma(n)}\right)}\left(x^{\sigma(1)},\dots,x^{\sigma(n)}\right) < \infty$$

für alle $x \in \mathbb{R}^n$. Dabei ist $f_{\left(X_0^{\sigma(1)}, \dots, X_0^{\sigma(n)}\right)}$ die Wahrscheinlichkeitsdichte von $\left(X_0^{\sigma(1)}, \dots, X_0^{\sigma(n)}\right)$.

3. Für jedes $n \in \{1, ..., N-1\}$ existient eine Funktionen $g_n \in L^1(\mathbb{R}^{N-n}, \mathbb{R})$ mit

$$\left|\frac{\partial}{\partial x^{\sigma(m)}}f_{\left(X_{0}^{\sigma(1)},\ldots,X_{0}^{\sigma(N)}\right)}\left(x^{\sigma(1)},\ldots,x^{\sigma(N)}\right)\right| \leq g_{n}\left(x^{\sigma(n+1)},\ldots,x^{\sigma(N)}\right)$$

für alle $m \in \{1, \ldots, n\}$ und $x \in \mathbb{R}^N$.

Dann ist die Abbildung $R'_{\sigma} : \mathbb{R}^N \to (0,1)^N$, $R'_{\sigma}(x) := R_{\sigma}(x)$, mit eingeschränkter Zielmenge ein Diffeomorphismus. Ferner gilt

$$\det\left(J_{R_{\sigma}}(x)\right) = f_0(x)$$

für alle $x \in \mathbb{R}^N$. Dabei bezeichnet $J_{R_{\sigma}}$ die Jacobi-Matrix von R_{σ} .

Beweis: Der Beweis wird in die Schritte 1) bis 4) unterteilt.

1) Behauptung: R'_{σ} ist bijektiv.

Weil X_0 eine strikt positive Lebesgue-Dichte besitzt, sind die Verteilungsfunktion

$$\mathbb{R} \ni x^{\sigma(1)} \mapsto F_{X_0^{\sigma(1)}}\left(x^{\sigma(1)}\right)$$

und für $n \in \{2, ..., N\}$ und $(x^{\sigma(1)}, ..., x^{\sigma(n-1)}) \in \mathbb{R}^{n-1}$ die bedingten Verteilungsfunktionen

$$\mathbb{R} \ni x^{\sigma(n)} \mapsto F_{X_0^{\sigma(n)} | X_0^{\sigma(1)}, \dots, X_0^{\sigma(n-1)}} \left(x^{\sigma(n)} | x^{\sigma(1)}, \dots, x^{\sigma(n-1)} \right)$$

streng monoton wachsend und damit injektiv. Betrachtet man diese Funktionen als Funktionen mit der Zielmenge (0, 1), so sind sie sogar bijektiv.

Seien $x_1, x_2 \in \mathbb{R}^N$ mit $x_1 \neq x_2$ und $n_0 \in \{1, \dots, N\}$ sei der erste Index mit $x_1^{\sigma(n_0)} \neq x_2^{\sigma(n_0)}$. Falls $n_0 = 1$, so gilt

$$R_{\sigma}^{\sigma(1)}(x_1) = F_{X_0^{\sigma(1)}}\left(x_1^{\sigma(1)}\right) \neq F_{X_0^{\sigma(1)}}\left(x_2^{\sigma(1)}\right) = R_{\sigma}^{\sigma(1)}(x_2)$$

und falls $n_0 \in \{2, \ldots, N\}$, so gilt

$$\begin{aligned} R_{\sigma}^{\sigma(n_{0})}(x_{1}) &= F_{X_{0}^{\sigma(n_{0})}|X_{0}^{\sigma(1)},\dots,X_{0}^{\sigma(n_{0}-1)}} \left(\left. x_{1}^{\sigma(n_{0})} \right| x_{1}^{\sigma(1)},\dots,x_{1}^{\sigma(n_{0}-1)} \right) \\ &\neq F_{X_{0}^{\sigma(n_{0})}|X_{0}^{\sigma(1)},\dots,X_{0}^{\sigma(n_{0}-1)}} \left(\left. x_{2}^{\sigma(n_{0})} \right| x_{2}^{\sigma(1)},\dots,x_{2}^{\sigma(n_{0}-1)} \right) \\ &= R_{\sigma}^{\sigma(n_{0})}(x_{2}). \end{aligned}$$

Also ist R'_{σ} injectiv.

Für $u \in (0, 1)^N$ definiere $x \in \mathbb{R}^N$ rekursiv durch

$$x^{\sigma(1)} := \left(F_{X_0^{\sigma(1)}}\right)^{-1} (u^{\sigma(1)})$$

und

$$x^{\sigma(n)} := \left(F_{X_0^{\sigma(n)} | X_0^{\sigma(1)}, \dots, X_0^{\sigma(n-1)}} \left(\cdot \left| x_1^{\sigma(1)}, \dots, x_1^{\sigma(n-1)} \right. \right) \right)^{-1} \left(u^{\sigma(n)} \right)$$

für $n \in \{2, \ldots, N\}$. Dann gilt $R'_{\sigma}(x) = u$, also ist R'_{σ} auch surjektiv.

2) Behauptung: R_{σ} ist stetig partiell differenzierbar.

Für $n \in \{2, ..., N\}$ ist die Dichte von $\left(X_0^{\sigma(1)}, \ldots, X_0^{\sigma(n-1)}\right)$ gegeben durch

Aus den Eigenschaften von f_0 folgt, dass $f_{(X^{\sigma(1)},...,X^{\sigma(n-1)})}$, aufgefasst als Abbildung mit Definitionsmenge \mathbb{R}^N , stetig partiell differenzierbar ist und dass $f_{\left(X^{\sigma(1)},\ldots,X^{\sigma(n-1)}\right)}\left(x^{\sigma(1)},\ldots,x^{\sigma(n-1)}\right) > 0 \text{ für alle } x \in \mathbb{R}^{N} \text{ gilt.}$

Weiter gilt für $n \in \{1, \ldots, N\}$

$$\int_{-\infty}^{x^{\sigma(n)}} f_{\left(X_{0}^{\sigma(1)},\ldots,X_{0}^{\sigma(n)}\right)}\left(x^{\sigma(1)},\ldots,x^{\sigma(n-1)},\tilde{x}^{\sigma(n)}\right) \,\mathrm{d}\tilde{x}^{\sigma(n)}$$

$$= \int_{-\infty}^{x^{\sigma(n)}} \int \ldots \int f_{\left(X_{0}^{\sigma(1)},\ldots,X_{0}^{\sigma(N)}\right)}\left(x^{\sigma(1)},\ldots,x^{\sigma(n-1)},\tilde{x}^{\sigma(n)},\ldots,\tilde{x}^{\sigma(N)}\right)$$

$$\times \,\mathrm{d}\tilde{x}^{\sigma(N)}\ldots \,\mathrm{d}\tilde{x}^{\sigma(n)}$$

Aus den Eigenschaften von f_0 folgt, dass auch dieser Ausdruck, aufgefasst als Abbildung mit Definitionsmenge \mathbb{R}^N , stetig partiell differenzierbar ist.

Somit ist

$$R_{\sigma}^{\sigma(1)}(x) = \int_{-\infty}^{x^{\sigma(1)}} f_{X_0^{\sigma(1)}} \left(\tilde{x}^{\sigma(1)} \right) \, \mathrm{d}\tilde{x}^{\sigma(1)}$$

stetig partiell differenzierbar und für $n \in \{2, \ldots, N\}$ ist

$$R_{\sigma}^{\sigma(n)}(x) = \frac{\int\limits_{-\infty}^{x^{\sigma(n)}} f_{\left(X_{0}^{\sigma(1)}, \dots, X_{0}^{\sigma(n)}\right)}\left(x^{\sigma(1)}, \dots, x^{\sigma(n-1)}, \tilde{x}^{\sigma(n)}\right) \, \mathrm{d}\tilde{x}^{\sigma(n)}}{f_{\left(X_{0}^{\sigma(1)}, \dots, X_{0}^{\sigma(n-1)}\right)}\left(x^{\sigma(1)}, \dots, x^{\sigma(n-1)}\right)}$$

ebenfalls stetig partiell differenzierbar.

3) Behauptung: det $(J_{R_{\sigma}}(x)) = f_0(x)$ für alle $x \in \mathbb{R}^N$. Sei $x \in \mathbb{R}^N$. Ferner seien $P_{\sigma}, P_{\sigma^{-1}} \in \mathbb{R}^{N \times N}$ die Permutationsmatrizen, die zu den Permutationen σ und σ^{-1} gehören, und $A = (a_{i,j})_{i,j \in \{1,...,N\}} :=$ $P_{\sigma}J_{R_{\sigma}}(x)P_{\sigma^{-1}}.$

Aus det $(P_{\sigma}) = \det(P_{\sigma^{-1}}) \in \{-1,1\}$ folgt det $(J_{R_{\sigma}}(x)) = \det(A)$. Zum Nachweis der Behauptung über die Funktionaldeterminante von $R_{\sigma}(x)$ genügt also die Berechnung von det(A).

Für $i, j \in \{1, \ldots, N\}$ mit i < j gilt

$$a_{1,j} = \frac{\partial R_{\sigma}^{\sigma(1)}}{\partial x^{\sigma(j)}}(x) = \frac{\partial}{\partial x^{\sigma(j)}} F_{X_0^{\sigma(1)}}\left(x^{\sigma(1)}\right) = 0$$

und

$$a_{i,j} = \frac{\partial R_{\sigma}^{\sigma(i)}}{\partial x^{\sigma(j)}}(x) = \frac{\partial}{\partial x^{\sigma(j)}} F_{X_0^{\sigma(i)}|X_0^{\sigma(1)},\dots,X_0^{\sigma(i-1)}}\left(x^{\sigma(i)} \middle| x^{\sigma(1)},\dots,x^{\sigma(i-1)}\right) = 0.$$

Die Matrix A ist also eine untere Dreiecksmatrix und somit ist ihre Determinante das Produkt ihrer Hauptdiagonalelemente. Man berechnet

$$a_{1,1} = \frac{\partial R_{\sigma}^{\sigma(1)}}{\partial x^{\sigma(1)}}(x) = f_{X_0^{\sigma(1)}}\left(x^{\sigma(1)}\right)$$

und

$$a_{n,n} = \frac{\partial R_{\sigma}^{\sigma(n)}}{\partial x^{\sigma(n)}}(x) = f_{X_0^{\sigma(n)}|X_0^{\sigma(1)},\dots,X_0^{\sigma(n-1)}}\left(x^{\sigma(n)} \,|\, x^{\sigma(1)},\dots,x^{\sigma(n-1)}\right)$$

für $n \in \{2, \ldots, N\}$. Hieraus folgt

$$det(A) = \prod_{n=1}^{N} a_{n,n}$$

$$= f_{X_0^{\sigma(1)}} \left(x^{\sigma(1)} \right) \prod_{n=2}^{N} f_{X_0^{\sigma(n)} | X_0^{\sigma(1)}, \dots, X_0^{\sigma(n-1)}} \left(x^{\sigma(n)} | x^{\sigma(1)}, \dots, x^{\sigma(n-1)} \right)$$

$$= f_{\left(X_0^{\sigma(1)}, \dots, X_0^{\sigma(N)} \right)} \left(x^{\sigma(1)}, \dots, x^{\sigma(N)} \right)$$

$$= f_0(x).$$

4) Behauptung: R'_{σ} be
sitzt eine stetig partiell differenzierbare Umkehrabbildung.

Aus $\det(J_{R'_{\sigma}}(x)) = \det(J_{R_{\sigma}}(x)) = f_0(x) > 0$ folgt, dass $J_{R'_{\sigma}}(x)$ für alle $x \in \mathbb{R}^N$ invertierbar ist. Zusammen mit der bereits nachgewiesenen Bijektivität folgt aus dem Satz von der Umkehrabbildung, dass R'_{σ} eine stetig partiell differenzierbare Umkehrabbildung besitzt. \Box

Lemma 58 impliziert insbesondere, dass unter den dort angegebenen Voraussetzungen die Rosenblatt-Transformation R_{σ} als Transformation T in der Teststatistik $A_{T,K,\mu}^2$ zulässig ist. Hierdurch ist die folgende Definition gerechtfertigt.

Definition 59 (Multivariate Rosenblatt-Erweiterung). Es gelten die Voraussetzungen und Bezeichnungen von Lemma 58. Für $T = R_{\sigma}$, heißt die in Gleichung (16) definierte Teststatistik $A_{T,K,\mu}^2 = A_{R_{\sigma},K,\mu}^2$ die zur Permutation σ dazugehörige multivariate Rosenblatt-Erweiterung der univariaten Teststatistik in Gleichung (15). Satz 60. Es gelten die Voraussetzungen und Bezeichnungen von Lemma 58.

Dann gelten die folgenden Aussagen.

- 1. $(R_{\sigma}(X_k))_{k\in\mathbb{N}}$ ist eine Folge von stochastisch unabhängigen und identisch verteilten $[0,1]^N$ -wertigen Zufallsvektoren. Falls die Nullhypothese H_0 wahr ist, so gilt $F_{R_{\sigma}(X_k)} = F_{\mathcal{U}([0,1]^N)}$ für alle $k \in \mathbb{N}$. Falls die Alternative H_1 wahr ist, so existiert eine Verteilungsfunktion $C_1 : [0,1]^N \to [0,1]$ mit $C_1 \neq F_{\mathcal{U}([0,1]^N)}$ und $F_{R_{\sigma}(X_k)} = C_1$ für alle $k \in \mathbb{N}$.
- 2. Es gilt

$$A_{R_{\sigma},K,\mu}^{2} = K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{R_{\sigma},K}(\tilde{u}) - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right]^{2}}{\left[F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\left(1 - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right)\right]^{\mu}} \,\mathrm{d}\tilde{u}.$$
 (18)

Hierbei bezeichnet $F_{\mathcal{U}([0,1]^N)}$ die Verteilungsfunktion der Gleichverteilung auf $[0,1]^N$.

Beweis: Aus Lemma 58 folgt $R_{\sigma}(\mathbb{R}^N) = (0, 1)^N$. Ferner impliziert Satz 57, dass, falls die Nullhypothese H_0 wahr ist, $F_{R_{\sigma}(X_k)} = F_{\mathcal{U}([0,1]^N)}$ für alle $k \in \mathbb{N}$ gilt. Hiermit verläuft der Nachweis von Teilaussage 1 analog zum Nachweis von Teilaussage 2 von Definition und Satz 54.

Teilaussage 2 folgt direkt aus Lemma 58 und dem Integraltransformationssatz. $\hfill \Box$

Bemerkung 61. Es wäre auch möglich gewesen, Gleichung (18) als Definitionsgleichung für die multivariate Rosenblatt-Erweiterung $A_{R_{\sigma},K,\mu}^2$ zu verwenden und dabei nicht die hohen Anforderungen an die Dichte f_0 der Verteilung von X_0 , die in den Voraussetzungen von Lemma 58 aufgeführt sind, zu stellen. Bei dieser Vorgehensweise wäre die Teststatistik in Gleichung (18) im Allgemeinen jedoch kein Spezialfall der Teststatistik in Gleichung (16) gewesen.

Ferner stellen die hohen Anforderungen an f_0 für die Zwecke der vorliegenden Arbeit keine Einschränkung dar, weil alle hier betrachteten Dichten, in deren Zusammenhang die Rosenblatt-Transformation angewandt wird, diesen Anforderungen genügen.
4.1.4 Diskussion der kanonischen multivariaten Erweiterung und der multivariaten Rosenblatt-Erweiterung

Auswertbarkeit der Teststatistik für eine gegebene Stichprobe: Im Fall der kanonischen multivariaten Erweiterung ist ein *N*-fach-Integral auszuwerten. Falls hierfür keine geschlossene Form bekannt ist und man auf numerische Verfahren angewiesen ist, kann das für nicht allzu kleine N zu einem Rechenaufwand führen, der nicht in vertretbarer Zeit zu bewältigen ist. Falls das Integral in seiner transformierten Form in Gleichung (17) einfacher auszuwerten ist, ist trotzdem für jedes Element $X_k, k \in \{1, \ldots, K\}$, der beobachteten Stichprobe der Wert $U(X_k)$ zu berechnen.

In Abschnitt 4.3 wird für $\mu \in \{0, 1\}$ eine geschlossene Form für das Integral in Gleichung (18) berechnet. Somit reduziert sich im Fall der multivariaten Rosenblatt-Erweiterung mit $\mu \in \{0, 1\}$ das Problem der Auswertung der Teststatistik im Wesentlichen auf die Berechnung der $R_{\sigma}(X_k)$ für $k \in \{1, \ldots, K\}$. Wie anspruchsvoll und aufwändig deren Berechnung ist, hängt vom Anwendungsfall ab.

Berechenbarkeit der relevanten Werte der Verteilungsfunktion der Teststatistik: In Abschnitt 4.2 wird Satz 72 bewiesen. Dieser impliziert, dass unter der Nullhypothese H_0 für geeignete $\mu \in \mathbb{R}_+$ sowohl im Fall der kanonischen multivariaten Erweiterung als auch im Fall der multivariaten Rosenblatt-Erweiterung die Teststatistik für $K \to \infty$ in Verteilung konvergiert. Es sei nun ein festes $\mu \in \mathbb{R}_+$ betrachtet, für das die besagte Verteilungskonvergenz vorliegt.

Bei der kanonischen multivariaten Erweiterung hängt dabei die Verteilung der Teststatistik und ihre Grenzverteilung für $K \to \infty$ im Allgemeinen von F_0 ab. Somit müssen für jeden Anwendungsfall die relevanten Werte der Verteilungsfunktion der Teststatistik bzw. der Grenzverteilung eigens dafür berechnet werden. Abhängig von K, N und F_0 kann das mit erheblichem Aufwand verbunden sein.

Im Fall der multivariaten Rosenblatt-Erweiterung ist unter der Nullhypothese H_0 die Verteilung der Teststatistik und ihre Grenzverteilung für $K \to \infty$ unabhängig von F_0 . Ferner sind diese Verteilungen auch unabhängig von der verwendeten Permutation σ . Die Verteilung der Teststatistik hängt lediglich von K und N ab und die Grenzverteilung hängt ausschließlich von N ab. Somit sind, bei vorgegebenen K und N, die Quantile der Teststatistik für jeden Anwendungsfall identisch und müssen nur einmal berechnet werden. Für hinreichend große K können diese durch die Quantile der Grenzverteilung approximiert werden. Hierin liegt eine große Stärke der multivariaten Rosenblatt-Erweiterung gegenüber anderen multivariaten Tests, bei denen die Quantile der Teststatistik unter der Nullhypothese H_0 von F_0 abhängen.

Willkür in Bezug auf die Definition des Begriffs der Verteilungsfunktion: Sei X ein \mathbb{R}^N -wertiger Zufallsvektor, dessen Verteilung eine Lebesgue-Dichte besitzt. Für $(i_1, \ldots, i_N) \in \{0, 1\}^N$ sei $F_X^{(i_1, \ldots, i_N)} : \mathbb{R}^N \to [0, 1]$ mit

$$F_X^{(i_1,\dots,i_N)}(x) := \mathbb{P}\left(\bigcap_{n=1}^N \left\{ (-1)^{i_n} X^n \le (-1)^{i_n} x^n \right\} \right).$$
(19)

Wenn man für jedes $n \in \{1, \ldots, N\}$ die Zielmenge \mathbb{R} der *n*-ten Komponente X^n von X als isotrop annehmen kann, so ist keine dieser Funktionen in irgendeiner mathematischen Art und Weise gegenüber den anderen ausgezeichnet. Somit ist die Wahl von $F_X := F_X^{(0,\ldots,0)}$ als die Verteilungsfunktion von X willkürlich. Man hätte genauso gut jede andere Funktion $F_X^{(i_1,\ldots,i_N)}$ mit $(i_1,\ldots,i_N) \in \{0,1\}^N$ als die Verteilungsfunktion von X definieren können. Es gibt also 2^N gleichberechtigte Möglichkeiten, den Begriff der Verteilungsfunktion zu definieren.

Sei $x_0 \in \mathbb{R}^N$. Für jedes $(i_1, \ldots, i_N) \in \{0, 1\}^N$ enthält der dazugehörige Wert $F_X^{(i_1,\ldots,i_N)}(x_0)$ eine andere Information. Diese Informationen sind jedoch nicht komplett unabhängig. Im Allgemeinen ermöglicht die Kenntnis von $2^N - 1$ dieser Werte die Berechnung des verbleibenden 2^N -ten Wertes. Somit ist nur im Spezialfall N = 1 die Definition des Begriffs der Verteilungsfunktion eindeutig in dem Sinne, dass die Kenntnis eines einzigen Wertes $F_X^{(i_1)}(x_0)$, $i_1 \in \{0, 1\}$, bereits die vollständige Information über alle anderen Werte, im eindimensionalen Fall also lediglich $F_X^{(1-i_1)}(x_0)$, enthält.

Für ein gegebenes $(i_1, \ldots, i_N) \in \{0, 1\}^N$ bezeichne

$$A_{T,K,\mu}^{2,(i_1,\dots,i_N)} = K \int_{\mathbb{R}^N} \frac{\left[\mathbb{F}_{T,K}^{(i_1,\dots,i_N)}(T(\tilde{x})) - F_{T(X_0)}^{(i_1,\dots,i_N)}(T(\tilde{x})) \right]^2}{\left[F_{T(X_0)}^{(i_1,\dots,i_N)}(T(\tilde{x})) \left(1 - F_{T(X_0)}^{(i_1,\dots,i_N)}(T(\tilde{x})) \right) \right]^{\mu}} \, \mathrm{d}F_0(\tilde{x})$$
(20)

die Teststatistik, die man erhält, wenn man in der Definitionsgleichung (16) die Verteilungsfunktion $F_{T(X_0)}$ durch $F_{T(X_0)}^{(i_1,\ldots,i_N)}$ und die empirische Verteilungsfunktion $\mathbb{F}_{T,K}$ durch $\mathbb{F}_{T,K}^{(i_1,\ldots,i_N)} : \mathbb{R}^N \to [0,1],$

$$\mathbb{F}_{T,K}^{(i_1,\dots,i_N)}(u) := \frac{1}{K} \sum_{k=1}^K \mathbb{1}_{\bigcap_{n=1}^N \{(-1)^{i_n} T^n(X_k) \le (-1)^{i_n} u^n\}},$$

ersetzt. Im Allgemeinen hängt sowohl der Wert der Teststatistik $A_{T,K,\mu}^{2,(i_1,\ldots,i_N)}$ für eine gegebene Stichprobe als auch ihre Verteilung unter der Nullhypothese H_0 von (i_1,\ldots,i_N) ab. Somit kann es vorkommen, dass das Ergebnis

eines Tests entscheidend davon abhängt, welche der 2^N gleichberechtigten Möglichkeiten für die Definition des Begriffs der Verteilungsfunktion gewählt wird.

Der nächste Satz impliziert, dass für eine fest gewählte Permutation σ die dazugehörige multivariate Rosenblatt-Erweiterung im folgenden Sinne unabhängig von dieser Problematik ist.

Satz 62. Es gelten die Voraussetzungen und Bezeichnungen von Lemma 58. Ferner seien $(i_1, \ldots, i_N) \in \{0, 1\}^N$ und

$$R_{\sigma}^{(i_1,\dots,i_N)} = (R_{\sigma}^{1,(i_1,\dots,i_N)},\dots,R_{\sigma}^{N,(i_1,\dots,i_N)}) : \mathbb{R}^N \to [0,1]^N$$

die zu σ dazugehörige Rosenblatt-Transformation, die man erhält, wenn man die Definition des Begriffs der bedingten Verteilungsfunktion analog zu Gleichung (19) anpasst, also

$$R_{\sigma}^{\sigma(1),(i_1,\dots,i_N)}(x) := \mathbb{P}\left(\left\{(-1)^{i_{\sigma(1)}} X_0^{\sigma(1)} \le (-1)^{i_{\sigma(1)}} x^{\sigma(1)}\right\}\right)$$

und

$$\begin{aligned} R^{\sigma(n),(i_1,\dots,i_N)}_{\sigma}(x) &:= \\ \mathbb{P}\left(\left\{ (-1)^{i_{\sigma(n)}} X_0^{\sigma(n)} \le (-1)^{i_{\sigma(n)}} x^{\sigma(n)} \right\} \middle| X_0^{\sigma(1)} = x^{\sigma(1)},\dots, X_0^{\sigma(n-1)} = x^{\sigma(n-1)} \right) \\ f \ddot{u} r \ n \in \{2,\dots,N\}. \end{aligned}$$

Dann gilt

$$A_{R_{\sigma}^{(i_1,\dots,i_N)},K,\mu}^{2,(i_1,\dots,i_N)} = A_{R_{\sigma},K,\mu}^2.$$

Hierbei bezeichnet $A^{2,(i_1,\ldots,i_N)}_{R^{(i_1,\ldots,i_N)}_{\sigma},K,\mu}$ die in Gleichung (20) definierte Teststatistik, die man für $T = R^{(i_1,\ldots,i_N)}_{\sigma}$ erhält.

Beweis: Sei $x \in \mathbb{R}^N$. Für $n \in \{1, ..., N\}$ und $k \in \mathbb{N}$ gilt

$$R_{\sigma}^{\sigma(n),(i_1,\dots,i_N)}(x) = \begin{cases} R_{\sigma}^{\sigma(n)}(x) & \text{für } i_{\sigma(n)} = 0\\ 1 - R_{\sigma}^{\sigma(n)}(x) & \text{für } i_{\sigma(n)} = 1 \end{cases}$$

und damit

$$\begin{split} & \left\{ (-1)^{i_n} R_{\sigma}^{n,(i_1,\dots,i_N)}(X_k) \leq (-1)^{i_n} R_{\sigma}^{n,(i_1,\dots,i_N)}(x) \right\} \\ & = \begin{cases} \{R_{\sigma}^n(X_k) \leq R_{\sigma}^n(x)\} & \text{für } i_n = 0 \\ \{-(1-R_{\sigma}^n(X_k)) \leq -(1-R_{\sigma}^n(x))\} & \text{für } i_n = 1 \end{cases} \\ & = \end{cases} \\ & = \{R_{\sigma}^n(X_k) \leq R_{\sigma}^n(x)\} \,. \end{split}$$

Hiermit berechnet man weiter

$$\mathbb{F}_{R_{\sigma}^{(i_{1},...,i_{N})},K}^{(i_{1},...,i_{N})}(R_{\sigma}^{(i_{1},...,i_{N})}(x)) = \mathbb{F}_{R_{\sigma},K}(R_{\sigma}(x))$$

sowie

$$\begin{aligned} F_{R_{\sigma}^{(i_{1},...,i_{N})}(X_{0})}^{(i_{1},...,i_{N})}(X_{0})}(R_{\sigma}^{(i_{1},...,i_{N})}(x)) \\ &= \mathbb{P}\left(\bigcap_{n=1}^{N}\left\{(-1)^{i_{n}}R_{\sigma}^{n,(i_{1},...,i_{N})}(X_{0}) \leq (-1)^{i_{n}}R_{\sigma}^{n,(i_{1},...,i_{N})}(x)\right\}\right) \\ &= \mathbb{P}\left(\bigcap_{n=1}^{N}\left\{R_{\sigma}^{n}(X_{k}) \leq R_{\sigma}^{n}(x)\right\}\right) \\ &= F_{R_{\sigma}(X_{0})}(R_{\sigma}(x)). \end{aligned}$$

Hieraus folgt $A^{2,(i_1,\ldots,i_N)}_{R^{(i_1,\ldots,i_N)}_{\sigma},K,\mu} = A^2_{R_{\sigma},K,\mu}$, also die Behauptung.

Willkür in Bezug auf die Permutation σ der Rosenblatt-Transformation: Die Verteilung der Teststatistik $A_{R_{\sigma},K,\mu}^2$ ist unabhängig von der Wahl der Permutation σ der Rosenblatt-Transformation. Für eine gegebene Stichprobe von Realisierungen $(x_k)_{k \in \{1,...,K\}}$ der Zufallsvektoren $(X_k)_{k \in \{1,...,K\}}$ kann jedoch der konkrete Wert der Teststatistik von σ abhängen. Sind σ_1 und σ_2 zwei Permutationen von $\{1, \ldots, N\}$ mit $\sigma_1 \neq \sigma_2$ und ist ein Signifikanzniveau vorgegeben, so kann es passieren, dass der Test mit der Teststatistik $A_{R_{\sigma_1},K,\mu}^2$ ablehnt und der Test mit der Teststatistik $A_{R_{\sigma_2},K,\mu}^2$ nicht ablehnt. Ist keine der Permutationen durch den betrachteten Anwendungsfall ausgezeichnet, sollte man sich vorher überlegen, welche Entscheidung in diesem Fall zu treffen ist.

4.2 Ein Anpassungstest für Copulas

Nun wird ein Testproblem behandelt, das dem aus dem vorhergehenden Abschnitt sehr ähnlich ist. Der Unterschied ist lediglich, dass der Wertebereich der betrachteten Zufallsvektoren das Intervall $[0, 1]^N$ ist und dass die Verteilungsfunktion unter der Nullhypothese H_0 eine Copula ist. Bei der Untersuchung der betrachteten Teststatistik wird ihr Erwartungswert berechnet und es wird ihr Konvergenzverhalten für immer größer werdende Stichprobengrößen erforscht. Hierbei stellt sich heraus, dass die Teststatistik unter der Nullhypothese H_0 in Verteilung gegen eine Zufallsvariable mit demselben Erwartungswert konvergiert und dass die Teststatistik unter der Alternative H_1 fast sicher gegen ∞ konvergiert.

4.2.1 Das Szenario

Das nun vorgestellte Szenario gelte im gesamten Abschnitt 4.2 in dem Sinne, dass die folgenden Annahmen und Bezeichnungen ohne weitere Erwähnung in jeder Definition, jedem Lemma und jedem Satz als Voraussetzung verwendet werden.

Sei $N \in \mathbb{N}, C_0 : [0,1]^N \to [0,1]$ eine Copula und $(U_k)_{k \in \mathbb{N}}$ eine Folge von stochastisch unabhängigen und identisch verteilten $[0,1]^N$ -wertigen Zufallsvektoren. Zu testen sei die

Nullhypothese
$$H_0$$
: $F_{U_k} = C_0$ für $k \in \mathbb{N}$.

gegen die

Alternative
$$H_1$$
: Es existiert eine Verteilungsfunktion
 $C_1 : [0,1]^N \to [0,1] \text{ mit } C_1 \neq C_0 \text{ und}$
 $F_{U_h} = C_1 \text{ für } k \in \mathbb{N}.$

Die nun vorgestellten Überlegungen basieren auf einer Betrachtung der Folge $(\mathbb{F}_K)_{K\in\mathbb{N}}$, deren Glieder folgendermaßen definiert sind. Für $K \in \mathbb{N}$ sei $\mathbb{F}_K(u) : [0,1]^N \to [0,1],$

$$\mathbb{F}_{K}(u) := \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{U_{k} \le u\}},$$

die empirische Verteilungsfunktion der Stichprobe, die aus den ersten K Beobachtungen der Folge $(U_k)_{k \in \mathbb{N}}$ besteht.

4.2.2 Definition der Teststatistik $A_{K,\mu}^2$ und der Zufallsvariable $D_{\mathbb{B},\mu}^2$ Definition 63 (*C*-Brownsche Brücke). Sei $C : [0,1]^N \to [0,1]$ eine stetige Copula.

Eine C-Brownsche Brücke $\mathbb{B} = (\mathbb{B}_u)_{u \in [0,1]^N}$ ist ein zentrierter Gauß-Prozess mit der Kovarianzfunktion

$$[0,1]^N \times [0,1]^N \ni (u_1, u_2) \mapsto K_C(u_1, u_2) := C(u_1 \wedge u_2) - C(u_1) C(u_2)$$

und dessen N-parametrige Pfade fast sicher stetig sind.

Im Spezialfall $C = F_{\mathcal{U}([0,1]^N)}$ nennt man \mathbb{B} eine N-parametrige Brownsche Brücke.

Definition 64 $(A_{K,\mu}^2 \text{ und } D_{\mu}^2)$. Die Copula C_0 sei stetig und es gelte $C_0(u) \in (0,1)$ für alle $u \in (0,1)^N$. Ferner sei \mathbb{B} eine C_0 -Brownsche Brücke.

Für $K \in \mathbb{N}$ und $\mu \in \mathbb{R}_+$ sind die Teststatistik $A^2_{K,\mu}$ und die Zufallsvariable $D^2_{\mathbb{B},\mu}$ definiert durch

$$A_{K,\mu}^{2} := K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u})\right]^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu}} \,\mathrm{d}C_{0}(\tilde{u})$$
(21)

und

$$D^{2}_{\mathbb{B},\mu} := \int_{(0,1)^{N}} \frac{\left(\mathbb{B}_{\tilde{u}}\right)^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu}} \,\mathrm{d}C_{0}(\tilde{u}).$$
(22)

Der Spezialfall $\mu = 0$ heißt Cramér-von-Mises-Fall und der Spezialfall $\mu = 1$ wird Anderson-Darling-Fall genannt.

Anmerkung: Die Bezeichnungen $A_{K,\mu}^2$ und $D_{\mathbb{B},\mu}^2$ wurden zu Ehren von Theodore Wilbur Anderson und Donald Allan Darling gewählt.

Bemerkung 65. Sei *G* ein Wahrscheinlichkeitsmaß auf $((0, 1)^N, \mathcal{B}((0, 1)^N))$. Alle Ergebnisse des vorliegenden Abschnitts 4.2 bleiben entsprechend gültig, wenn man an allen Stellen, an denen C_0 als Integrator verwendet wird, stattdessen nach *G* integriert. Die entsprechend modifizierten Definitionsgleichungen (21) und (22) von $A_{K,\mu}^2$ und $D_{\mathbb{B},\mu}^2$ lauten in diesem Fall

$$A_{K,\mu}^{2} := K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u})\right]^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu}} G(\mathrm{d}\tilde{u})$$

und

$$D^{2}_{\mathbb{B},\mu} := \int_{(0,1)^{N}} \frac{(\mathbb{B}_{\tilde{u}})^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu}} G(\mathrm{d}\tilde{u}).$$

Hierdurch bekommt man weitere Wahlmöglichkeiten, für welche Bereiche in $(0,1)^N$ die Abweichungen zwischen \mathbb{F}_K und C_0 erhöhtes bzw. verringertes Gewicht erhalten sollen.

4.2.3 Erwartungswert von $A_{K,\mu}^2$ und $D_{\mathbb{B},\mu}^2$

Definition 66. Eine Copula $C : [0,1]^N \to [0,1]$ genügt der Annahme (A) zur Konstante $\nu_{\mathbb{A}} \in (0,\infty)$, falls die folgenden Bedingungen erfüllt sind.

- 1. C ist stetig.
- 2. Es gilt $C(u) \in (0,1)$ für alle $u \in (0,1)^N$.
- 3. Es gilt

$$\int_{(0,1)^N} \frac{1}{\left[C(\tilde{u})(1-C(\tilde{u}))\right]^{\nu}} \,\mathrm{d}C(\tilde{u}) < \infty$$

für alle $\nu \in [0, \nu_{\mathbb{A}})$.

Satz 67. Die Unabhängigkeitscopula, also die Verteilungsfunktion $F_{\mathcal{U}([0,1]^N)}$ der Gleichverteilung auf $[0,1]^N$, genügt der Annahme (\mathbb{A}) zur Konstante $\nu_{\mathbb{A}} = 1$.

Der Beweis von Satz 67 ist trivial und wird ausgelassen.

Satz 68. Die Nullhypothese H_0 sei wahr. Ferner genüge die Copula C_0 der Annahme (A) zur Konstante $\nu_{\mathbb{A}} \in (0, \infty)$ und \mathbb{B} sei eine C_0 -Brownsche Brücke.

Dann gilt

$$\mathbb{E}\left[A_{K,\mu}^{2}\right] = \mathbb{E}\left[D_{\mathbb{B},\mu}^{2}\right] = \int_{(0,1)^{N}} \frac{1}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu-1}} \, \mathrm{d}C_{0}(\tilde{u}) < \infty$$

für alle $K \in \mathbb{N}$ und $\mu \in [0, \nu_{\mathbb{A}} + 1)$.

Im Anderson-Darling-Fall, also im Spezialfall $\mu = 1$, gilt

$$\mathbb{E}\left[A_{K,1}^2\right] = \mathbb{E}\left[D_{\mathbb{B},1}^2\right] = 1$$

für alle $K \in \mathbb{N}$.

Beweis: Für $u \in (0,1)^N$ und $k \in \{1,\ldots,K\}$ ist die Zufallsvariable $\mathbb{1}_{\{U_k \leq u\}}$ Bernoulli-verteilt mit Erfolgswahrscheinlichkeit $C_0(u)$. Weil die Zufallsvektoren U_1, \ldots, U_K stochastisch unabhängig sind, ist die Zufallsvariable $K\mathbb{F}_K(u)$ binomialverteilt mit K Versuchen und Erfolgswahrscheinlichkeit $C_0(u)$. Hieraus folgt zusammen mit dem Satz von Fubini

$$\mathbb{E}\left[A_{K,\mu}^{2}\right] = \int_{(0,1)^{N}} \frac{\mathbb{E}\left[\left[K\mathbb{F}_{K}(\tilde{u}) - KC_{0}(\tilde{u})\right]^{2}\right]}{K\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu}} dC_{0}(\tilde{u})$$
$$= \int_{(0,1)^{N}} \frac{1}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right)\right]^{\mu-1}} dC_{0}(\tilde{u})$$
$$< \infty.$$

Eine weitere Anwendung des Satzes von Fubini ergibt

$$\mathbb{E} \left[D_{\mathbb{B},\mu}^2 \right] = \int_{(0,1)^N} \frac{\mathbb{E} \left[(\mathbb{B}_{\tilde{u}})^2 \right]}{\left[C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{\mu}} \, \mathrm{d}C_0(\tilde{u}) \\ = \int_{(0,1)^N} \frac{1}{\left[C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{\mu-1}} \, \mathrm{d}C_0(\tilde{u}),$$

und damit die Behauptung.

4.2.4 Konvergenz von $A_{K,\mu}^2$ für $K \to \infty$ unter der Nullhypothese H_0 und unter der Alternative H_1

Die nun folgende Definition der schwachen Konvergenz ist aus Vaart und Wellner (1996, Definition 1.3.3) übernommen.

Definition 69 (Schwache Konvergenz). Sei \mathbb{D} ein metrischer Raum. Für $K \in \mathbb{N}$ seien $(\Omega_K, \mathcal{A}_K, \mathbb{P}_K)$ ein Wahrscheinlichkeitsraum und $\mathbb{X}_K : \Omega_K \to \mathbb{D}$ eine beliebige Abbildung. Ferner seien $(\Omega_0, \mathcal{A}_0, \mathbb{P}_0)$ ein weiterer Wahrscheinlichkeitsraum und $\mathbb{X} : \Omega_0 \to \mathbb{D}$ eine $\mathcal{A}_0 - \mathcal{B}(\mathbb{D})$ -messbare Abbildung, wobei $\mathcal{B}(\mathbb{D})$ die Borelsche σ -Algebra auf \mathbb{D} bezeichnet.

Die Folge $(\mathbb{X}_K)_{K \in \mathbb{N}}$ konvergiert schwach gegen \mathbb{X} , falls

$$\lim_{K \to \infty} \mathbb{E}^*_{\mathbb{P}_K} \left[H(\mathbb{X}_K) \right] = \mathbb{E}_{\mathbb{P}_0} \left[H(\mathbb{X}) \right]$$
(23)

für alle stetigen, beschränkten Funktionen $H : \mathbb{D} \to \mathbb{R}$ gilt. Für $K \in \mathbb{N}$ bezeichnet hierbei $\mathbb{E}^*_{\mathbb{P}_K} [\cdot]$ den äußeren Erwartungswert bezüglich \mathbb{P}_K und $\mathbb{E}_{\mathbb{P}_0} [\cdot]$ bezeichnet den Erwartungswert bezüglich \mathbb{P}_0 .

Satz 70. Die Nullhypothese H_0 sei wahr und die Copula C_0 sei stetig. $\ell^{\infty}([0,1]^N)$ bezeichne den Raum der beschränkten, reellwertigen Funktionen auf $[0,1]^N$, ausgestattet mit der Supremumsnorm, und $\mathcal{B}(\ell^{\infty}([0,1]^N))$ bezeichne die Borelsche σ -Algebra auf $\ell^{\infty}([0,1]^N)$. Für $K \in \mathbb{N}$ sei ferner $\mathbb{Y}_K : \Omega \to \ell^{\infty}([0,1]^N)$,

$$[0,1]^N \ni u \mapsto \mathbb{Y}_K(u)(\omega) := \sqrt{K} \left(\mathbb{F}_K(u)(\omega) - C_0(u) \right) \in \mathbb{R},$$

der zu \mathbb{F}_K dazugehörige empirische Prozess.

Dann existiert eine auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega_0, \mathcal{A}_0, \mathbb{P}_0)$ definierte C_0 -Brownsche Brücke \mathbb{B} mit den folgenden Eigenschaften.

- 1. \mathbb{B} , aufgefasst als Abbildung $\mathbb{B} : \Omega_0 \to \ell^{\infty}([0,1]^N)$, ist $\mathcal{A}_0 \mathcal{B}(\ell^{\infty}([0,1]^N)) messbar$.
- 2. Die Verteilung $\mathbb{P}_0^{\mathbb{B}}$ von \mathbb{B} , aufgefasst als Borel-Maß auf $(\ell^{\infty}([0,1]^N), \mathcal{B}(\ell^{\infty}([0,1]^N)))$, ist straff.
- 3. Die Folge $(\mathbb{Y}_K)_{K \in \mathbb{N}}$ konvergiert für $K \to \infty$ schwach gegen \mathbb{B} .

Beweis: Sei $\mathcal{F}_{[0,1]^N} := \{ \mathbb{1}_{(-\infty,t]} | t \in [0,1]^N \}$ und $\ell^{\infty} (\mathcal{F}_{[0,1]^N})$ bezeichne den Raum der beschränkten, reellwertigen Funktionen auf $\mathcal{F}_{[0,1]^N}$, ausgestattet mit der Supremumsnorm. Nach Vaart und Wellner (1996, Example 2.5.4) ist die Familie

$$\mathcal{F} := \left\{ \mathbb{1}_{(-\infty,t]} \middle| t \in \mathbb{R}^N \right\}$$

eine universelle Donsker-Klasse. Also ist auch $\mathcal{F}_{[0,1]^N} \subset \mathcal{F}$ eine universelle Donsker-Klasse. Identifiziert man für $t \in [0,1]^N$ die Funktion $\mathbb{1}_{(-\infty,t]}$ mit t und identifiziert man ferner $\ell^{\infty}(\mathcal{F}_{[0,1]^N})$ mit $\ell^{\infty}([0,1]^N)$, so folgt aus der Definition einer Donsker-Klasse, dass ein $\mathbb{B} : \Omega_0 \to \ell^{\infty}([0,1]^N)$ existiert, das $\mathcal{A}_0 - \mathcal{B}(\ell^{\infty}([0,1]^N))$ -messbar ist, eine straffe Verteilung $\mathbb{P}_0^{\mathbb{B}}$ besitzt und gegen das die Folge $(\mathbb{Y}_K)_{K \in \mathbb{N}}$ schwach konvergiert.

Aus $\mathcal{F}_{[0,1]^N} \subset L^2(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathcal{L}(U_1))$ folgt, dass \mathbb{B} ein zentrierter Gauß-Prozess mit der Kovarianzfunktion

$$[0,1]^N \times [0,1]^N \ni (u_1, u_2) \mapsto K_{C_0}(u_1, u_2) := C_0(u_1 \wedge u_2) - C_0(u_1) C_0(u_2)$$

ist. Nach Segers (2012, Section 2) impliziert die Straffheit von $\mathbb{P}_0^{\mathbb{B}}$ zusammen mit der Stetigkeit von K_{C_0} , dass eine Version von \mathbb{B} existiert, deren Pfade fast sicher stetig sind. Diese Version von \mathbb{B} ist eine C_0 -Brownsche Brücke mit den gewünschten Eigenschaften.

Bemerkung 71. In Chibisov (1965), Billingsley (1968, Chapter 18, Seiten 150-153) und Vaart und Wellner (1996, Problem 1.7.3) wird für den Spezialfall N = 1 ein Beispiel konstruiert, in dem sowohl die empirische Verteilungsfunktion \mathbb{F}_1 der ersten Beobachtung als auch der dazugehörige empirische Prozess \mathbb{Y}_1 , jeweils aufgefasst als Abbildungen mit Definitionsmenge Ω und Zielmenge $\mathbb{D}([0,1])$, nicht \mathcal{A} - $\mathcal{B}(\mathbb{D}([0,1]))$ -messbar sind. Hierbei bezeichnet $\mathbb{D}([0,1]) \subset \ell^{\infty}([0,1])$ den Raum der càdlàg-Funktionen auf [0,1], ausgestattet mit der Supremumsnorm, und $\mathcal{B}(\mathbb{D}([0,1]))$ bezeichnet die Borelsche σ -Algebra auf $\mathbb{D}([0,1])$. Somit ist in der Situation von Satz 70 auf der linken Seite der Definitionsgleichung (23) der schwachen Konvergenz wirklich die Verwendung des äußeren Erwartungswertes erforderlich und dieser kann im Allgemeinen nicht durch den Erwartungswert ersetzt werden. **Satz 72.** Die Nullhypothese H_0 sei wahr. Ferner genüge die Copula C_0 der Annahme (A) zur Konstante $\nu_{\mathbb{A}} \in [1, \infty)$ und \mathbb{B} sei der schwache Grenzwert der Folge $(\mathbb{Y}_K)_{K \in \mathbb{N}}$ aus Satz 70.

Falls $\mu \in [0, \frac{\nu_{\mathbb{A}}}{2} + 1)$, so konvergiert die Folge $(A_{K,\mu}^2)_{K \in \mathbb{N}}$ für $K \to \infty$ in Verteilung gegen $D_{\mathbb{B},\mu}^2$.

Beweis: Für $\mu, \zeta \in \mathbb{R}_+$ und $K \in \mathbb{N}$ definiere

$$A_{K,\mu,\zeta}^{2} := K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u})\right]^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right) + \zeta\right]^{\mu}} \, \mathrm{d}C_{0}(\tilde{u})$$

und

$$D^{2}_{\mathbb{B},\mu,\zeta} := \int_{(0,1)^{N}} \frac{(\mathbb{B}_{\tilde{u}})^{2}}{\left[C_{0}(\tilde{u})\left(1 - C_{0}(\tilde{u})\right) + \zeta\right]^{\mu}} \, \mathrm{d}C_{0}(\tilde{u}).$$

Für $\zeta \in (0,\infty)$ gilt trivialerweise $A^2_{K,\mu,\zeta} < \infty$ und $D^2_{\mathbb{B},\mu,\zeta} < \infty$ jeweils fast sicher.

Der Beweis wird in die Teile 1) bis 6) unterteilt.

1) Behauptung: Für alle $\mu \in \mathbb{R}_+$ und $\zeta \in (0, \infty)$ konvergiert die Folge $(A^2_{K,\mu,\zeta})_{K\in\mathbb{N}}$ in Verteilung gegen $D^2_{\mathbb{B},\mu,\zeta}$.

Die Argumentation für den vorliegenden Teil 1) des Beweises orientiert sich an Genest, Huang und Dufour (2013, Beweis von Proposition 1) und passt diesen auf den hier vorliegenden Fall eines deterministischen Integrators an. Dabei schließen die Autoren in ihrem Beweis den Fall $\zeta = 0$ aus.

Genest, Huang und Dufour definieren die Funktion $\Psi : C^0([0,1]^N, \mathbb{R}) \to C^0([0,1]^N, \mathbb{R})$ mit

$$\Psi(g) := \frac{g^2}{\left[C_0 \left(1 - C_0\right) + \zeta\right]^{\mu}}$$

und zeigen, dass sie stetig ist. Dabei statten sie den Raum $C^0([0,1]^N, \mathbb{R})$ mit der Supremumsnorm aus. Ferner ist die Abbildung $C^0([0,1]^N, \mathbb{R}) \ni g \mapsto \int_{(0,1)^N} g(\tilde{u}) \, \mathrm{d}C_0(\tilde{u}) \in \mathbb{R}$ ebenfalls stetig. Somit ist auch die Verkettung $C^0([0,1]^N, \mathbb{R}) \ni g \mapsto \int_{(0,1)^N} \Psi(g(\tilde{u})) \, \mathrm{d}C_0(\tilde{u})$ stetig. Es gilt

$$A_{K,\mu,\zeta}^2 = \int_{(0,1)^N} \Psi(\mathbb{Y}_K(\tilde{u})) \,\mathrm{d}C_0(\tilde{u})$$

für $K \in \mathbb{N}$. Also folgt aus Satz 70 zusammen mit Vaart und Wellner (1996, Theorem 1.3.6 (Continuous mapping)), dass die Folge $(A_{K,\mu,\zeta}^2)_{K\in\mathbb{N}}$ für $K \to \infty$ schwach gegen

$$D^2_{\mathbb{B},\mu,\zeta} = \int_{(0,1)^N} \Psi(\mathbb{B}_{\tilde{u}}) \,\mathrm{d}C_0(\tilde{u})$$

konvergiert. Weil für reellwertige Zufallsvariablen die schwache Konvergenz und die Konvergenz in Verteilung zusammenfallen, folgt die Behauptung.

2) Behauptung: Für $p \in [1, 2)$ ist die Folge

$$\left(\int_{(0,1)^N} \left[K \frac{\left[\mathbb{F}_K(\tilde{u}) - C_0(\tilde{u})\right]^2}{C_0(\tilde{u}) \left(1 - C_0(\tilde{u})\right)} \right]^p \, \mathrm{d}C_0(\tilde{u}) \right)_{K \in \mathbb{N}}$$

gleichgradig integrierbar.

Für $K \in \mathbb{N}$ und $u \in (0,1)$ ist die Zufallsvariable $K\mathbb{F}_K(u)$ binomialverteilt mit K Versuchen und Erfolgswahrscheinlichkeit $C_0(u)$. Eine einfache Rechnung ergibt

$$\mathbb{E}\left[\left|K\left[\mathbb{F}_{K}(u) - C_{0}(u)\right]\right|^{4}\right] \leq 4K^{2}C_{0}(u)(1 - C_{0}(u)).$$

Definiere $\gamma := \frac{1}{p} + \frac{1}{2} > 1$. Dann ist die Funktion $G : \mathbb{R}_+ \to \mathbb{R}$, $G(t) = t^{\gamma}$, nichtnegativ, monoton wachsend, konvex und es gilt $\lim_{t\to\infty} \frac{G(t)}{t} = \infty$. Unter Verwendung des Satzes von Fubini und zweifacher Anwendung der Jensenschen Ungleichung berechnet man

$$\begin{split} \sup_{K \in \mathbb{N}} & \mathbb{E} \left[G \left(\int_{(0,1)^N} \left[K \frac{\left[\mathbb{F}_K(\tilde{u}) - C_0(\tilde{u}) \right]^2}{C_0(\tilde{u}) \left(1 - C_0(\tilde{u})\right)} \right]^p \, \mathrm{d}C_0(\tilde{u}) \right) \right] \\ \leq & \sup_{K \in \mathbb{N}} \int_{(0,1)^N} \frac{\mathbb{E} \left[|K \left[\mathbb{F}_K(\tilde{u}) - C_0(\tilde{u}) \right] |^{2(1+\frac{p}{2})} \right]}{\left[K C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{1+\frac{p}{2}}} \, \mathrm{d}C_0(\tilde{u}) \\ \leq & \sup_{K \in \mathbb{N}} \int_{(0,1)^N} \frac{\left(\mathbb{E} \left[|K \left[\mathbb{F}_K(\tilde{u}) - C_0(\tilde{u}) \right] |^4 \right] \right)^{\frac{1}{2} + \frac{p}{4}}}{\left[K C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{1+\frac{p}{2}}} \, \mathrm{d}C_0(\tilde{u}) \\ \leq & 4^{\frac{1}{2} + \frac{p}{4}} \int_{(0,1)^N} \frac{1}{\left[C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{\frac{1}{2} + \frac{p}{4}}} \, \mathrm{d}C_0(\tilde{u}) \\ < & \infty. \end{split}$$

Hierbei wurde ausgenutzt, dass $\frac{1}{2} + \frac{p}{4} \in [0, 1) \subset [0, \nu_{\mathbb{A}})$ gilt. Die Behauptung folgt nun aus dem Satz von de la Vallée-Poussin.

3) Behauptung: Sei $\mu \in \left[0, \frac{\nu_{\mathbb{A}}}{2} + 1\right)$. Falls $\mu \in [0, 1]$, so sei ferner $I := \mathbb{R}_+$, und falls $\mu \in \left[0, \frac{\nu_{\mathbb{A}}}{2} + 1\right) \setminus [0, 1]$, so sei $I := \left[0, \frac{\nu_{\mathbb{A}}}{\mu - 1}\right)$. Dann gilt

$$\lim_{\zeta \searrow 0} \int_{(0,1)^N} \left[\frac{1 - \left[\frac{C_0(\tilde{u})(1 - C_0(\tilde{u}))}{C_0(\tilde{u})(1 - C_0(\tilde{u})) + \zeta} \right]^{\mu}}{\left[C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{\mu - 1}} \right]^q \, \mathrm{d}C_0(\tilde{u}) = 0$$

für alle $q \in I$.

Der Fall $\mu \in [0,1]$ ist trivial. Im Folgenden wird der Fall $\mu \in [0, \frac{\nu_{\mathbb{A}}}{2} + 1) \setminus [0,1]$ behandelt. Für $u \in (0,1)^N$ gilt $C_0(u) (1 - C_0(u)) \in (0,1)$ und damit

$$\lim_{\zeta \searrow 0} \left[\frac{1 - \left[\frac{C_0(u)(1 - C_0(u))}{C_0(u)(1 - C_0(u)) + \zeta} \right]^{\mu}}{\left[C_0(u) \left(1 - C_0(u) \right) \right]^{\mu - 1}} \right]^q = 0.$$

Weil C_0 der Annahme (A) zur Konstante $\nu_{\mathbb{A}}$ genügt und weil $q(\mu - 1) \in [0, \nu_{\mathbb{A}})$ gilt, folgt für die Funktion $g : (0, 1)^N \to \mathbb{R}$,

$$g(u) := \frac{1}{\left[C_0(u)\left(1 - C_0(u)\right)\right]^{q(\mu-1)}},$$

 $\int_{(0,1)^N} g(\tilde{u}) \, \mathrm{d}C_0(\tilde{u}) < \infty$. Ferner gilt

$$\left[\frac{1 - \left[\frac{C_0(u)(1 - C_0(u))}{C_0(u)(1 - C_0(u)) + \zeta}\right]^{\mu}}{\left[C_0(u)\left(1 - C_0(u)\right)\right]^{\mu - 1}}\right]^q \le g(u)$$

für alle $u \in (0,1)^N$ und $\zeta \ge 0$. Somit folgt die Behauptung aus dem Satz von der majorisierten Konvergenz.

4) Behauptung: Sei $\mu \in [0, \frac{\nu_{\mathbb{A}}}{2} + 1)$ und $f : \mathbb{R} \to \mathbb{R}$ eine stetige, beschränkte Funktion mit kompaktem Träger. Dann gilt

$$\lim_{\zeta \searrow 0} \mathbb{E}\left[f\left(A_{K,\mu,\zeta}^{2}\right)\right] = \mathbb{E}\left[f\left(A_{K,\mu}^{2}\right)\right]$$

für alle $K \in \mathbb{N}$. Ferner gilt

$$\lim_{K \to \infty} \lim_{\zeta \searrow 0} \mathbb{E} \left[f \left(A_{K,\mu,\zeta}^2 \right) \right] = \lim_{\zeta \searrow 0} \lim_{K \to \infty} \mathbb{E} \left[f \left(A_{K,\mu,\zeta}^2 \right) \right].$$

Nach Satz 68 gilt $\mathbb{E}\left[A_{K,\mu}^2\right] < \infty$ und damit $A_{K,\mu}^2 < \infty$ fast sicher. Somit ist $f\left(A_{K,\mu}^2\right)$ fast sicher definiert.

Sei nun $\varepsilon > 0$. Ferner sei $S := \sup_{x \in \mathbb{R}} |f(x)|$. Falls S = 0, so ist die Behauptung trivial. Im Folgenden wird der Fall S > 0 behandelt.

Weil f gleichmäßig stetig ist, existiert ein $\delta > 0$, sodass $|f(x_1) - f(x_2)| < \frac{\varepsilon}{2}$ für alle $x_1, x_2 \in \mathbb{R}$ mit $|x_1 - x_2| < \delta$.

Falls $\mu \in [0, 1]$, definiere $p := \frac{3}{2}$ und q := 3. Dann gilt $q \in \mathbb{R}_+$. Falls $\mu \in [0, \frac{\nu_{\mathbb{A}}}{2} + 1) \setminus [0, 1]$, definiere $p := \frac{2(\mu - 1)}{\nu_{\mathbb{A}}} + 1$ und $q := \frac{\nu_{\mathbb{A}}}{2(\mu - 1)} + 1$. Dann gilt $q \in \left(2, \frac{\nu_{\mathbb{A}}}{\mu - 1}\right)$. In beiden Fällen gilt jedoch $p \in (1, 2)$ und $\frac{1}{p} + \frac{1}{q} = 1$.

Aus Teil 2) des vorliegenden Beweises folgt, dass es ein $b_0 \in [1, \infty)$ gibt mit

$$\sup_{K \in \mathbb{N}} \mathbb{E} \left[\mathbbm{1}_{B_K} \int_{(0,1)^N} \left[K \frac{\left[\mathbb{F}_K(\tilde{u}) - C_0(\tilde{u}) \right]^2}{C_0(\tilde{u}) \left(1 - C_0(\tilde{u})\right)} \right]^p \, \mathrm{d}C_0(\tilde{u}) \right] < \frac{\varepsilon}{4S}$$

wobei

$$B_K := \left\{ \int_{(0,1)^N} \left[K \frac{\left[\mathbb{F}_K(\tilde{u}) - C_0(\tilde{u}) \right]^2}{C_0(\tilde{u}) \left(1 - C_0(\tilde{u})\right)} \right]^p \, \mathrm{d}C_0(\tilde{u}) > b_0 \right\}$$

für $K \in \mathbb{N}$. Insbesondere gilt $\mathbb{P}(B_K) < \frac{\varepsilon}{4S}$ für alle $K \in \mathbb{N}$.

Aus Teil 3) des vorliegenden Beweises folgt, dass ein $\zeta_0 > 0$ existiert mit

$$\int_{(0,1)^N} \left[\frac{1 - \left[\frac{C_0(\tilde{u})(1 - C_0(\tilde{u}))}{C_0(\tilde{u})(1 - C_0(\tilde{u})) + \zeta} \right]^{\mu}}{\left[C_0(\tilde{u}) \left(1 - C_0(\tilde{u}) \right) \right]^{\mu - 1}} \right]^q \, \mathrm{d}C_0(\tilde{u}) < \left(\frac{\delta}{b_0^{\frac{1}{p}}} \right)^q$$

für alle $\zeta \in (0, \zeta_0)$.

Sei nun $K \in \mathbb{N}$ und $\zeta \in (0, \zeta_0)$. Dann berechnet man unter Verwendung

der Hölderschen Ungleichung

$$\begin{split} \left| A_{K,\mu}^{2} - A_{K,\mu,\zeta}^{2} \right| \\ &\leq \int_{(0,1)^{N}} K \left| \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u}) \right]^{2}}{\left[C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right) \right]^{\mu}} - \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u}) \right]^{2}}{\left[C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right) + \zeta \right]^{\mu}} \right| \, \mathrm{d}C_{0}(\tilde{u}) \\ &= \int_{(0,1)^{N}} \left| K \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u}) \right]^{2}}{C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right)} \right| \left| \frac{1 - \left[\frac{C_{0}(\tilde{u})(1 - C_{0}(\tilde{u}))}{C_{0}(\tilde{u})(1 - C_{0}(\tilde{u})) + \zeta} \right]^{\mu}}{\left[C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right) \right]^{\mu-1}} \right| \, \mathrm{d}C_{0}(\tilde{u}) \\ &\leq \left(\int_{(0,1)^{N}} \left[K \frac{\left[\mathbb{F}_{K}(\tilde{u}) - C_{0}(\tilde{u}) \right]^{2}}{C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right)} \right]^{p} \, \mathrm{d}C_{0}(\tilde{u}) \right)^{\frac{1}{p}} \\ &\times \left(\int_{(0,1)^{N}} \left[\frac{1 - \left[\frac{C_{0}(\tilde{u})(1 - C_{0}(\tilde{u}))}{C_{0}(\tilde{u})(1 - C_{0}(\tilde{u})) + \zeta} \right]^{\mu}}{\left[C_{0}(\tilde{u}) \left(1 - C_{0}(\tilde{u}) \right) \right]^{\mu-1}} \right]^{q} \, \mathrm{d}C_{0}(\tilde{u}) \right)^{\frac{1}{q}}. \end{split}$$

Also gilt $|A_{K,\mu}^2 - A_{K,\mu,\zeta}^2| < \delta$ auf B_K^C . Aufgrund der bereits erwähnten gleichmäßigen Stetigkeit von f folgt damit $|f(A_{K,\mu}^2) - f(A_{K,\mu,\zeta}^2)| < \frac{\varepsilon}{2}$ auf B_K^C .

Zusammen ergibt sich

$$\sup_{K \in \mathbb{N}} \left| \mathbb{E} \left[f \left(A_{K,\mu}^{2} \right) \right] - \mathbb{E} \left[f \left(A_{K,\mu,\zeta}^{2} \right) \right] \right|$$

$$\leq \sup_{K \in \mathbb{N}} \left(\mathbb{E} \left[\mathbb{1}_{B_{K}^{C}} \left| f \left(A_{K,\mu}^{2} \right) - f \left(A_{K,\mu,\zeta}^{2} \right) \right| \right] \right.$$

$$\left. + \mathbb{E} \left[\mathbb{1}_{B_{K}} \left| f \left(A_{K,\mu}^{2} \right) - f \left(A_{K,\mu,\zeta}^{2} \right) \right| \right] \right)$$

$$< \frac{\varepsilon}{2} + 2S \frac{\varepsilon}{4S}$$

$$= \varepsilon$$

für alle $\zeta \in (0, \zeta_0)$.

Hieraus folgt $\lim_{\zeta \searrow 0} \mathbb{E} \left[f\left(A_{K,\mu,\zeta}^2\right) \right] = \mathbb{E} \left[f\left(A_{K,\mu}^2\right) \right]$ und der Grenzwert wird gleichmäßig bezüglich $K \in \mathbb{N}$ angenommen. Teil 1) des vorliegenden Beweises impliziert, dass $\lim_{K\to\infty} \mathbb{E} \left[f\left(A_{K,\mu,\zeta}^2\right) \right] = \mathbb{E} \left[f\left(D_{\mathbb{B},\mu,\zeta}^2\right) \right]$ für alle $\zeta > 0$ gilt. Die Behauptung folgt nun aus dem Satz über das Vertauschen von Grenzwerten, wie er in Heuser (1992, Satz 107.2) zu finden ist.

5) Behauptung: Für alle $\mu \in [0, \nu_{\mathbb{A}} + 1)$ und jede stetige, beschränkte Function $f : \mathbb{R} \to \mathbb{R}$ gilt

$$\lim_{\zeta \searrow 0} \mathbb{E}\left[f\left(D_{\mathbb{B},\mu,\zeta}^{2}\right)\right] = \mathbb{E}\left[f\left(D_{\mathbb{B},\mu}^{2}\right)\right].$$

Aus $\mu \in [0, \nu_{\mathbb{A}} + 1)$ und Satz 68 folgt $\mathbb{E} \left[D^2_{\mathbb{B}, \mu} \right] < \infty$ und damit $D^2_{\mathbb{B}, \mu} < \infty$ fast sicher. Somit ist $f \left(D^2_{\mathbb{B}, \mu} \right)$ fast sicher definiert.

Sei nun $\omega \in \{D^2_{\mathbb{B},\mu} < \infty\}$. Dann gilt für alle $u \in (0,1)^N$

$$0 \le \frac{\left(\mathbb{B}_{u}(\omega)\right)^{2}}{\left[C_{0}(u)\left(1 - C_{0}(u)\right) + \zeta\right]^{\mu}} \nearrow \frac{\left(\mathbb{B}_{u}(\omega)\right)^{2}}{\left[C_{0}(u)\left(1 - C_{0}(u)\right)\right]^{\mu}}$$

für $\zeta \searrow 0$. Somit folgt aus dem Satz von der monotonen Konvergenz $\lim_{\zeta\searrow 0} D^2_{\mathbb{B},\mu,\zeta}(\omega) = D^2_{\mathbb{B},\mu}(\omega)$. Die Behauptung folgt nun aus der Stetigkeit und der Beschränktheit von f sowie dem Satz von der majorisierten Konvergenz.

6) Behauptung: Sei $\mu \in [0, \frac{\nu_{\mathbb{A}}}{2} + 1)$. Dann konvergiert die Folge $(A_{K,\mu}^2)_{K \in \mathbb{N}}$ für $K \to \infty$ in Verteilung gegen $D_{\mathbb{B},\mu}^2$.

Sei $f : \mathbb{R} \to \mathbb{R}$ stetig, beschränkt und mit kompaktem Träger. Dann folgt aus den Teilen 4), 1) und 5) des vorliegenden Beweises

$$\lim_{K \to \infty} \mathbb{E} \left[f \left(A_{K,\mu}^2 \right) \right] = \lim_{K \to \infty} \lim_{\zeta \searrow 0} \mathbb{E} \left[f \left(A_{K,\mu,\zeta}^2 \right) \right]$$
$$= \lim_{\zeta \searrow 0} \lim_{K \to \infty} \mathbb{E} \left[f \left(A_{K,\mu,\zeta}^2 \right) \right]$$
$$= \lim_{\zeta \searrow 0} \mathbb{E} \left[f \left(D_{\mathbb{B},\mu,\zeta}^2 \right) \right]$$
$$= \mathbb{E} \left[f \left(D_{\mathbb{B},\mu}^2 \right) \right].$$

Also konvergiert die Folge $(A_{K,\mu}^2)_{K\in\mathbb{N}}$ für $K \to \infty$ vag gegen $D_{\mathbb{B},\mu}^2$. Aus $\lim_{K\to\infty} \mathbb{P}\left(\{A_{K,\mu}^2\in\mathbb{R}\}\right) = 1 = \mathbb{P}\left(\{D_{\mathbb{B},\mu}^2\in\mathbb{R}\}\right)$ und Bauer (1968, Satz 45.7) folgt schließlich die Behauptung.

Bemerkung 73. Im vorstehenden Beweis wird $\mathbb{P}\left(D_{\mathbb{B},\mu}^2 < \infty\right) = 1$ für geeignete $\mu \in \mathbb{R}_+$ und alle $N \in \mathbb{N}$ direkt aus Satz 68 geschlussfolgert. Anderson und Darling (1952) untersuchen das vorliegende Testproblem für den Spezialfall N = 1 und damit $C_0 = F_{\mathcal{U}([0,1])}$. Auch wenn sie für ihren Nachweis der Verteilungskonvergenz von $A_{K,\mu}^2$ gegen $D_{\mathbb{B},\mu}^2$ für $K \to \infty$ unter der Nullhypothese H_0 einen anderen Zugang wählen, nutzen auch sie die Tatsache aus, dass $D_{\mathbb{B},\mu}^2$ fast sicher endlich ist. Dies zeigen sie unter Verwendung des Gesetzes des iterierten Logarithmus für die Brownsche Bewegung. In Abschnitt A.1 im Anhang wird ein alternativer Beweis für $\mathbb{P}\left(D_{\mathbb{B},\mu}^2 < \infty\right) = 1$

für den Spezialfall $C_0 = F_{\mathcal{U}([0,1]^N)}$ vorgestellt, der die Argumentation von Anderson und Darling auf den Fall $N \in \mathbb{N}$ verallgemeinert. Dabei werden unter anderem auch einige fast sichere Pfadeigenschaften des Brownschen Blatts gezeigt, die von eigenem mathematischen Interesse sind.

Zum Testen der Nullhypothese H_0 gegen die Alternative H_1 ist die Teststatistik $A_{K,\mu}^2$ stark konsistent im folgenden Sinne.

Satz 74. Die Nullhypothese H_0 sei falsch. Stattdessen sei die Alternative H_1 wahr und für die Verteilungsfunktion C_1 gelte zusätzlich

$$\int_{(0,1)^N} 1_{\{C_1 \neq C_0\}}(\tilde{u}) \, \mathrm{d}C_0(\tilde{u}) > 0.$$

Ferner sei $\mu \in \mathbb{R}_+$.

Dann gilt

$$\lim_{K\to\infty}A_{K,\mu}^2=\infty$$

fast sicher.

Beweis: Aus $\int_{(0,1)^N} 1_{\{C_1 \neq C_0\}}(\tilde{u}) dC_0(\tilde{u}) > 0$ folgt, dass es ein $\delta > 0$ gibt mit

$$\int_{(0,1)^N} \mathbf{1}_{\{|C_1 - C_0| \ge \delta\}}(\tilde{u}) \, \mathrm{d}C_0(\tilde{u}) > 0.$$

Sei $\omega \in \{\lim_{K\to\infty} \sup_{\tilde{u}\in(0,1)^N} |\mathbb{F}_K(\tilde{u}) - C_1(\tilde{u})| = 0\}$. Dann gibt es ein $K_0(\omega) \in \mathbb{N}$, sodass $\sup_{\tilde{u}\in(0,1)^N} |\mathbb{F}_K(\tilde{u})(\omega) - C_1(\tilde{u})| < \frac{\delta}{2}$ für alle $K \in \mathbb{N}$ mit $K \ge K_0(\omega)$. Hieraus folgt

$$|\mathbb{F}_{K}(u)(\omega) - C_{0}(u)| \ge |C_{1}(u) - C_{0}(u)| - \sup_{\tilde{u} \in (0,1)^{N}} |\mathbb{F}_{K}(\tilde{u})(\omega) - C_{1}(\tilde{u})| \ge \frac{\delta}{2}$$

für alle $K \ge K_0(\omega)$ und $u \in \{|C_1 - C_0| \ge \delta\}$ und damit

$$\begin{aligned} A_{K,\mu}^{2}(\omega) &\geq K \int_{(0,1)^{N}} \mathbb{1}_{\{|C_{1}-C_{0}| \geq \delta\}}(\tilde{u}) \left[\mathbb{F}_{K}(\tilde{u})(\omega) - C_{0}(\tilde{u})\right]^{2} \mathrm{d}C_{0}(\tilde{u}) \\ &\geq K \frac{\delta^{2}}{4} \int_{(0,1)^{N}} \mathbb{1}_{\{|C_{1}-C_{0}| \geq \delta\}}(\tilde{u}) \mathrm{d}C_{0}(\tilde{u}) \\ &\to \infty \end{aligned}$$

für $K \to \infty$.

Wie bereits im Beweis von Satz 70 dargelegt, ist die Familie $\mathcal{F}_{[0,1]^N} = \{\mathbbm{1}_{(-\infty,t]} | t \in [0,1]^N\}$ eine universelle Donsker-Klasse und damit insbesondere auch eine \mathbb{P} -Glivenko-Cantelli-Klasse unter der Alternative H_1 . Unter Verwendung der dort gemachten Identifikationen von $\mathbbm{1}_{(-\infty,t]}$ mit $t, t \in [0,1]^N$, und $\ell^{\infty}(\mathcal{F}_{[0,1]^N})$ mit $\ell^{\infty}([0,1]^N)$ erhält man

$$\mathbb{P}\left(\left\{\lim_{K\to\infty}\sup_{\tilde{u}\in(0,1)^N}|\mathbb{F}_K(\tilde{u})-C_1(\tilde{u})|=0\right\}\right)=1$$

Hieraus folgt $\lim_{K\to\infty} A_{K,\mu}^2 = \infty$ fast sicher, also die Behauptung.

Anderson und Darling (1952, 1954) bestimmen für den Fall N = 1 und $\mu \in \{0, 1\}$ jeweils die charakteristische Funktion von $D^2_{\mathbb{B},\mu}$ unter der Nullhypothese H_0 und benutzen diese als Ausgangspunkt zur approximativen Berechnung gewisser Werte der Verteilungsfunktion $F_{D^2_{\mathbb{B},\mu}}$ von $D^2_{\mathbb{B},\mu}$. In der vorliegenden Arbeit wird auf eine exakte Berechnung der charakteristischen Funktion von $A^2_{K,\mu}$ oder $D^2_{\mathbb{B},\mu}$ und anschließende Fourier-Inversion verzichtet. Stattdessen wird für $N \geq 2$ die Verteilungsfunktion von $A^2_{K,\mu}$ mithilfe einer Monte-Carlo-Simulation geschätzt. In Tabelle 1 sind für $N \in \{2,3,4\}$, $C_0 = F_{\mathcal{U}([0,1]^N)}$ und $\mu \in \{0,1\}$ die auf einer Grundlage von je $5 \cdot 10^4$ Realisierungen geschätzten kritischen Werte für einige Signifikanzniveaus angegeben. Für $N \in \{2,3\}$ wird K = 253 gewählt. Für N = 4 werden hiervon abweichende Werte für K verwendet, um die Rechenzeit für die Monte-Carlo-Simulationen in einem vertretbaren Rahmen zu halten. Bei der Erstellung der Tabelle kamen unter anderem die geschlossenen Formen für $A^2_{K,\mu}$ zum Einsatz, die nun in Abschnitt 4.3 hergeleitet werden.

Test-	N	K	Signifikanzniveau				
statistik			$15 \ \%$	10~%	5~%	2,5~%	1~%
$A_{K,0}^2$	2	253	0,216	0,255	0,328	0,404	0,505
	3	253	0,128	0,148	0,187	0,226	0,277
	4	204	0,069	0,080	0,097	0, 117	0,142
$A_{K,1}^2$	2	253	1,470	1,695	2,081	2,523	3,111
,	3	253	1,355	1,528	1,851	2,183	2,676
	4	64	1,354	1,577	2,025	2,529	3,328

Tabelle 1: Durch Monte-Carlo-Simulation geschätzte kritische Werte der Teststatistiken $A_{K,0}^2$ und $A_{K,1}^2$ für $N \in \{2, 3, 4\}$ und ausgewählte Werte für K

4.3 Geschlossene Formen für die Teststatistik $A^2_{K,\mu}$ im Cramér-von-Mises- und im Anderson-Darling-Fall

4.3.1 Voraussetzungen und Bezeichnungen

Sei $N \in \mathbb{N}$. Ferner seien $K \in \mathbb{N}$ und $(u_k)_{k \in \{1,...,K\}}$ eine gegebene Stichprobe von Realisierungen der $[0,1]^N$ -wertigen Zufallsvektoren $(U_k)_{k \in \{1,...,K\}}$. Im vorliegenden Abschnitt 4.3 werden für $\mu \in \{0,1\}$, also den Cramér-von-Mises-Fall bzw. den Anderson-Darling-Fall, und $C_0 = F_{\mathcal{U}([0,1]^N)}$ geschlossene Formen für die in Gleichung (21) definierte Teststatistik, also

$$A_{K,\mu}^{2} = K \int_{(0,1)^{N}} \frac{\left[\mathbb{F}_{K}(\tilde{u}) - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right]^{2}}{\left[F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\left(1 - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right)\right]^{\mu}} \,\mathrm{d}\tilde{u},$$
(24)

berechnet. Hierfür werden die folgenden Bezeichnungen eingeführt.

- 1. Für $n \in \{1, \ldots, N\}$ sei $(u_{(k)}^n)_{k \in \{1, \ldots, K\}}$ die zu $(u_k^n)_{k \in \{1, \ldots, K\}}$ gehörige geordnete Stichprobe. Ferner sei $u_{(0)}^n := 0$ und $u_{(K+1)}^n := 1$.
- 2. Für $(k_1, \ldots, k_N) \in \{0, \ldots, K+1\}^N$ sei

$$F_K(k_1,\ldots,k_N) := \mathbb{F}_K\left(u_{(k_1)}^1,\ldots,u_{(k_N)}^N\right) = \frac{1}{K} \sum_{k=1}^K \mathbb{1}_{\left\{u_k \le \left(u_{(k_1)}^1,\ldots,u_{(k_N)}^N\right)\right\}}$$

die empirische Verteilungsfunktion der realisierten Stichprobe $(u_k)_{k \in \{1,...,K\}}$, ausgewertet an der Stelle $(u_{(k_1)}^1, \ldots, u_{(k_N)}^N)$.

Für $(k_1, ..., k_N) \in \{0, ..., K\}^N$ erhält man damit $\mathbb{F}_K(u) = F_K(k_1, ..., k_N)$ für alle $u \in \left((u_{(k_1)}^1, ..., u_{(k_N)}^N), (u_{(k_1+1)}^1, ..., u_{(k_N+1)}^N) \right).$

Bemerkung 75. Hier sei noch einmal daran erinnert, dass sich nach Satz 60 im Fall der multivariaten Rosenblatt-Erweiterung aus Abschnitt 4.1.3 die Auswertung der Teststatistik $A_{R_{\sigma},K,\mu}^2$ auf die Auswertung von $A_{K,\mu}^2$ für den hier betrachteten Spezialfall von $C_0 = F_{\mathcal{U}([0,1]^N)}$ transformieren lässt. Somit sind die im vorliegenden Abschnitt 4.3 gefundenen geschlossenen Formen für $A_{K,\mu}^2$ insbesondere auch zur Auswertung von $A_{R_{\sigma},K,\mu}^2$ geeignet. Voraussetzung hierfür ist natürlich, dass die Berechnung der Rosenblatt-Transformationen $R_{\sigma}(x_k)$ der Beobachtungen $u_k, k \in \{1, \ldots, K\}$, gelingt.

4.3.2 Berechnung einer geschlossenen Form für $A_{K,0}^2$

Satz 76. Es gelten die Voraussetzungen und die Bezeichnungen aus Abschnitt 4.3.1. Ferner liege der Cramér-von-Mises-Fall $\mu = 0$ vor.

Dann gilt die folgende geschlossene Form zur Berechnung des Wertes der Teststatistik in Gleichung (24)

$$A_{K,0}^{2} = K \sum_{(k_{1},\dots,k_{N})\in\{1,\dots,K\}^{N}} \left[\left(F_{K}(k_{1},\dots,k_{N})\right)^{2} \prod_{n=1}^{N} \left(u_{(k_{n}+1)}^{n}-u_{(k_{n})}^{n}\right) - \left(\frac{1}{2}\right)^{N-1} F_{K}(k_{1},\dots,k_{N}) \prod_{n=1}^{N} \left(\left[u_{(k_{n}+1)}^{n}\right]^{2}-\left[u_{(k_{n})}^{n}\right]^{2}\right)\right] + K \sum_{(k_{1},\dots,k_{N})\in\{0,\dots,K\}^{N}} \left(\frac{1}{3}\right)^{N} \prod_{n=1}^{N} \left(\left[u_{(k_{n}+1)}^{n}\right]^{3}-\left[u_{(k_{n})}^{n}\right]^{3}\right).$$

Beweis: Durch eine Unterteilung des Integrationsintervalls erhält man

$$A_{K,0}^{2} = K \sum_{\substack{(k_{1},\dots,k_{N})\in\{0,\dots,K\}^{N}\\ \times \int_{u_{(k_{1})}^{1}}^{u_{(k_{1}+1)}^{1}} \dots \int_{u_{(k_{N})}^{N}}^{u_{(k_{N}+1)}^{N}} \left[F_{K}(k_{1},\dots,k_{N}) - \prod_{n=1}^{N} \tilde{u}^{n}\right]^{2} d\tilde{u}^{N} \dots d\tilde{u}^{1}.$$

Die weitere Berechnung ist trivial.

4.3.3 Berechnung einer geschlossenen Form für $A_{K,1}^2$

Die Berechnung einer geschlossenen Form für $A_{K,1}^2$ ist selbst im Spezialfall $C_0 = F_{\mathcal{U}([0,1]^N)}$ recht aufwändig. Hierfür werden einige Resultate über den Polylogarithmus benötigt, der folgendermaßen definiert ist.

Definition 77 (Polylogarithmus). Für $N \in \mathbb{N}_0$ ist der Polylogarithmus der Ordnung N definiert als die Abbildung $\operatorname{Li}_N : [0,1] \to \overline{\mathbb{R}}$,

$$\operatorname{Li}_N(x) := \sum_{k=1}^{\infty} \frac{x^k}{k^N}.$$
(25)

Bemerkung 78. Es sind wesentlich allgemeinere Definitionen des Polylogarithmus möglich und auch in der Literatur üblich. Die Menge der zulässigen Ordnungen kann als Z gewählt werden und es ist ebenfalls eine analytische Fortsetzung als holomorphe Funktion möglich. Seinen Namen erhält der Polylogarithmus von den beiden Eigenschaften $\text{Li}_1(x) = -\log(1-x)$ und $\text{Li}_{N+1}(x) = \int_0^x \frac{\text{Li}_N(\tilde{x})}{\tilde{x}} d\tilde{x}$. Die erste dieser Eigenschaften folgt sofort aus der Potenzreihendarstellung des Logarithmus. Die zweite Eigenschaft ist ein Spezialfall von Teilaussage 4 des nun folgenden Lemmas 79, in dem die für die vorliegende Arbeit relevanten Eigenschaften des Polylogarithmus zusammengefasst sind. Um die Arbeit weitestgehend in sich abgeschlossen zu gestalten, wird der Beweis dieser bekannten Resultate hier ebenfalls kurz skizziert.

Lemma 79 (Eigenschaften des Polylogarithmus). Sei $N \in \mathbb{N}_0$. Ferner sei I := [0, 1), falls $N \in \{0, 1\}$, und I := [0, 1], falls N > 1.

1. Die Potenzreihe in der Definitionsgleichung (25) des Polylogarithmus konvergiert für alle $x \in I$ und die Abbildung $I \ni x \mapsto \text{Li}_N(x) \in \mathbb{R}$, also die Abbildung $I \ni x \mapsto \text{Li}_N(x)$ mit \mathbb{R} statt $\overline{\mathbb{R}}$ als Zielmenge, ist stetig.

2. Es gilt
$$\operatorname{Li}_0(x) = \begin{cases} \frac{x}{1-x} & \text{für } x \in [0,1) \\ \infty & \text{für } x = 1 \end{cases}$$

3. $\operatorname{Li}_{N+1}(x)$ ist differencies and [0,1) und es gilt

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{Li}_{N+1}(x) = \begin{cases} \frac{1}{x}\mathrm{Li}_N(x) & \text{für } x \in (0,1)\\ 1 & \text{für } x = 0 \end{cases}.$$

4. Für
$$a, x \in [0, 1]$$
 gilt $\int_{0}^{x} \frac{1}{\tilde{x}} \operatorname{Li}_{N}(a\tilde{x}) d\tilde{x} = \operatorname{Li}_{N+1}(ax)$

Beweis: Das Quotientenkriterium liefert die absolute Konvergenz der Reihe in Gleichung (25) sogar für alle $x \in (-1, 1)$. Für N > 1 und x = 1 erhält man die konvergente Reihe $\sum_{k=1}^{\infty} \frac{1}{k^N}$. Zusammen mit dem Abelschen Grenzwertsatz folgt Teilaussage 1.

Für N = 0 und $x \in [0, 1)$ ist die Potenzreihe in der Definitionsgleichung (25) des Polylogarithmus eine geometrische Reihe, für x = 1 divergiert die Reihe bestimmt gegen ∞ . Hieraus folgt Teilaussage 2.

Aus der absoluten Konvergenz der Potenzreihe in Gleichung (25) für alle $N \in \mathbb{N}_0$ und $x \in (-1, 1)$ folgt die Differenzierbarkeit von $\operatorname{Li}_{N+1}(x)$ auf [0, 1) und die Ableitung ergibt sich durch gliedweise Differentiation, also

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{Li}_{N+1}(x) = \sum_{k=1}^{\infty} k \ \frac{x^{k-1}}{k^{N+1}} = \begin{cases} \frac{1}{x} \sum_{k=1}^{\infty} \frac{x^k}{k^N} & \text{für } x \in (0,1) \\ 1 & \text{für } x = 0 \end{cases}$$

Hieraus folgt Teilaussage 3.

Teilaussage 4 folgt direkt aus den Teilaussagen 1, 2 und 3, sowie aus $\text{Li}_{N+1}(0) = 0$. Dabei ist zu beachten, dass im Fall N = 0 und a = x = 1 beide Seiten den Wert ∞ annehmen.

Lemma 80. Seien $N \in \mathbb{N}$, $u = (u^1, \dots, u^N) \in [0, 1]^N$ und $a \in (0, 1]$.

Dann gilt

$$\int_0^{u^1} \dots \int_0^{u^N} \frac{1}{1 - a \prod_{n=1}^N \tilde{u}^n} \, \mathrm{d}\tilde{u}^N \dots \, \mathrm{d}\tilde{u}^1 = \frac{1}{a} \mathrm{Li}_N \left(a \prod_{n=1}^N u^n \right).$$

Beweis: Falls es ein $n \in \{1, ..., N\}$ mit $u^n = 0$ gibt, so ist die Aussage trivial. Der Nachweis für den Fall $u^n > 0$ für alle $n \in \{1, ..., N\}$ wird durch vollständige Induktion über N geführt.

Induktionsanfang: ${\cal N}=1.$ Unter Verwendung der Teilaussagen 2 und 4 von Lemma 79 folgt

$$\int_0^{u^1} \frac{1}{1 - a\tilde{u}^1} \,\mathrm{d}\tilde{u}^1 = \frac{1}{a} \int_0^{u^1} \frac{1}{\tilde{u}^1} \,\frac{a\tilde{u}^1}{1 - a\tilde{u}^1} \,\mathrm{d}\tilde{u}^1 = \frac{1}{a} \mathrm{Li}_1 \left(au^1 \right).$$

Induktionsschritt: $N \rightarrow N + 1$. Mit dem Satz von Fubini folgt

$$\begin{split} &\int_{0}^{u^{1}} \dots \int_{0}^{u^{N+1}} \frac{1}{1-a \prod_{n=1}^{N+1} \tilde{u}^{n}} \, \mathrm{d}\tilde{u}^{N+1} \dots \, \mathrm{d}\tilde{u}^{1} \\ &= \int_{0}^{u^{N+1}} \left[\int_{0}^{u^{1}} \dots \int_{0}^{u^{N}} \frac{1}{1-a \tilde{u}^{N+1} \prod_{n=1}^{N} \tilde{u}^{n}} \, \mathrm{d}\tilde{u}^{N} \dots \, \mathrm{d}\tilde{u}^{1} \right] \, \mathrm{d}\tilde{u}^{N+1} \\ &= \int_{0}^{u^{N+1}} \frac{1}{a \tilde{u}^{N+1}} \mathrm{Li}_{N} \left(a \tilde{u}^{N+1} \prod_{n=1}^{N} \tilde{u}^{n} \right) \, \mathrm{d}\tilde{u}^{N+1} \\ &= \frac{1}{a} \int_{0}^{u^{N+1}} \frac{1}{\tilde{u}^{N+1}} \mathrm{Li}_{N} \left(\left[a \prod_{n=1}^{N} \tilde{u}^{n} \right] u^{N+1} \right) \, \mathrm{d}\tilde{u}^{N+1} \\ &= \frac{1}{a} \mathrm{Li}_{N+1} \left(a \prod_{n=1}^{N+1} \tilde{u}^{n} \right). \end{split}$$

Dabei wurde beim zweiten Umformungsschritt die Induktionsvoraussetzung verwendet und bei der letzten Umformung wurde Teilaussage 4 von Lemma 79 angewandt. $\hfill \Box$

Korollar 81. Sei $N \in \mathbb{N}$.

1. Scien $u_1 = (u_1^1, \dots, u_1^N), u_2 = (u_2^1, \dots, u_2^N) \in [0, 1]^N$ mit $u_1 \le u_2$. Im Fall N = 1 sei ferner $u_1 \ne 1$. Dann gilt

$$\int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} \frac{1}{1 - \prod_{n=1}^N \tilde{u}^n} \, \mathrm{d}\tilde{u}^N \dots \, \mathrm{d}\tilde{u}^1$$
$$= \sum_{\iota_1, \dots, \iota_N \in \{1, 2\}} (-1)^{\sum_{n=1}^N \iota_n} \, \mathrm{Li}_N \left(\prod_{n=1}^N u_{\iota_n}^n\right)$$

2. Scien $u_1, u_2 \in (0, 1]^N$ mit $u_1 \leq u_2$ und $C \in [0, 1]$. Im Fall N = 1 sci ferner $u_1 \neq 1$. Und im Fall N = 1, $u_2 = 1$ sci ferner $C \neq 1$. Dann gilt

$$\int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} \frac{\left[C - \prod_{n=1}^N \tilde{u}^n\right]^2}{\prod_{n=1}^N \tilde{u}^n \left(1 - \prod_{n=1}^N \tilde{u}^n\right)} d\tilde{u}^N \dots d\tilde{u}^1 = C^2 \prod_{n=1}^N \log\left(\frac{u_2^n}{u_1^n}\right) + (1 - C)^2 \sum_{\iota_1, \dots, \iota_N \in \{1, 2\}} (-1)^{\sum_{n=1}^N \iota_n} \operatorname{Li}_N\left(\prod_{n=1}^N u_{\iota_n}^n\right) - \prod_{n=1}^N (u_2^n - u_1^n)$$

3. Für $u_1, u_2 \in (0, 1]^N$ mit $u_1 \le u_2$ gilt

$$\int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} \frac{1 - \prod_{n=1}^N \tilde{u}^n}{\prod_{n=1}^N \tilde{u}^n} \, \mathrm{d}\tilde{u}^N \dots \, \mathrm{d}\tilde{u}^1 = \prod_{n=1}^N \log\left(\frac{u_2^n}{u_1^n}\right) - \prod_{n=1}^N \left(u_2^n - u_1^n\right).$$

•

Beweis: Zum Nachweis von Teilaussage 1 sei $f:[0,1]^N \to \mathbb{R}_+$ mit

$$f(u) := \begin{cases} \frac{1}{1-\prod_{n=1}^{N} u^n} & \text{für } u \in [0,1]^N \setminus \{(1,\dots,1)\} \\ 0 & \text{für } u = (1,\dots,1) \end{cases}$$

Dann ist f stetig auf $[0,1]^N \setminus \{1,\ldots,1\}$ und die Abbildung

$$\mathcal{B}([0,1]^N) \ni A \mapsto \nu(A) := \int_{[0,1]^N} \mathbb{1}_A(\tilde{u}) f(\tilde{u}) \,\mathrm{d}\tilde{u} = \left(f\lambda^N|_{[0,1]^N}\right)(A)$$

ist ein Maß. Hierbei bezeichnet $\mathcal{B}([0,1]^N)$ die Borelsche σ -Algebra auf $[0,1]^N$ und $\lambda^N|_{[0,1]^N}$ bezeichnet die Spur des Lebesgue-Maßes auf $([0,1]^N, \mathcal{B}([0,1]^N))$.

Im Fall N = 1 folgt Teilaussage 1 aus dem Hauptsatz der Differentialund Integralrechnung und Lemma 80, wobei im Fall $u_2 = 1$ das Integral als uneigentliches Integral aufzufassen ist und den Wert ∞ annimmt.

Im Fall N > 1 gilt $\nu([0,1]^N) = \text{Li}_N(1) < \infty$ und Teilaussage 1 folgt aus Elstrodt (2005, Kapitel II, Satz 3.6) und Lemma 80.

Als nächstes wird Teilaussage 2 nachgewiesen. Durch einfaches Nachrechnen erhält man

$$\frac{[C-x]^2}{x(1-x)} = C^2 \frac{1}{x} + (1-C)^2 \frac{1}{1-x} - 1$$

für alle $x \in \mathbb{R} \setminus \{0, 1\}$. Hieraus folgt zusammen mit Teilaussage 1

$$\begin{split} &\int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} \frac{\left[C - \prod_{n=1}^N \tilde{u}^n\right]^2}{\prod_{n=1}^N \tilde{u}^n \left(1 - \prod_{n=1}^N \tilde{u}^n\right)} \,\mathrm{d}\tilde{u}^N \dots \,\mathrm{d}\tilde{u}^1 \\ &= C^2 \int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} \frac{1}{\prod_{n=1}^N \tilde{u}^n} \,\mathrm{d}\tilde{u}^N \dots \,\mathrm{d}\tilde{u}^1 \\ &+ (1 - C)^2 \int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} \frac{1}{1 - \prod_{n=1}^N \tilde{u}^n} \,\mathrm{d}\tilde{u}^N \dots \,\mathrm{d}\tilde{u}^1 \\ &- \int_{u_1^1}^{u_2^1} \dots \int_{u_1^N}^{u_2^N} 1 \,\mathrm{d}\tilde{u}^N \dots \,\mathrm{d}\tilde{u}^1 \\ &= C^2 \prod_{n=1}^N \log\left(\frac{u_2^n}{u_1^n}\right) \\ &+ (1 - C)^2 \sum_{\iota_1, \dots, \iota_N \in \{1, 2\}} (-1)^{\sum_{n=1}^N \iota_n} \,\mathrm{Li}_N\left(\prod_{n=1}^N u_{\iota_n}^n\right) - \prod_{n=1}^N (u_2^n - u_1^n) \,, \end{split}$$

also die behauptete Teilaussage 2.

Der Nachweis von Teilaussage 3 ist trivial.

In den Voraussetzungen des nun folgende Satzes 82 wird die Zusatzannahme getroffen, dass $u_k \in (0,1)^N$ für alle $k \in \{1,\ldots,K\}$ gilt. Diese Zusatzannahme wird dadurch gerechtfertigt, dass im betrachteten Spezialfall von $C_0 = F_{\mathcal{U}([0,1]^N)}$ unter der Nullhypothese H_0 aus Abschnitt 4.2.1 alle U_k , $k \in \mathbb{N}$, gleichverteilt auf $[0,1]^N$ sind und damit

$$\mathbb{P}\left(\bigcap_{k\in\mathbb{N}} \{U_k\in(0,1)^N\}\right) = 1$$

vorliegt.

Satz 82. Es gelten die Voraussetzungen und die Bezeichnungen aus Abschnitt 4.3.1 und es liege der Anderson-Darling-Fall $\mu = 1$ vor. Ferner sei $u_k \in (0,1)^N$ für alle $k \in \{1,\ldots,K\}.$

Dann gilt die folgende geschlossene Form zur Berechnung des Wertes der Teststatistik in Gleichung (24)

$$A_{K,1}^{2} = K \left[-1 + \sum_{(k_{1},\dots,k_{N})\in\{1,\dots,K\}^{N}} (F_{K}(k_{1},\dots,k_{N}))^{2} \prod_{n=1}^{N} \log\left(\frac{u_{(k_{n}+1)}^{n}}{u_{(k_{n})}^{n}}\right) + \sum_{(k_{1},\dots,k_{N})\in\{0,\dots,K\}^{N}\setminus\{K,\dots,K\}} (1 - F_{K}(k_{1},\dots,k_{N}))^{2} \times \sum_{\iota_{1},\dots,\iota_{N}\in\{1,2\}} (-1)^{\sum_{n=1}^{N}\iota_{n}} \operatorname{Li}_{N}\left(\prod_{n=1}^{N} u_{(k_{n}-1+\iota_{n})}^{n}\right) \right].$$

Beweis: Durch eine Unterteilung des Integrationsintervalls erhält man

$$A_{K,1}^{2} = K \sum_{\substack{(k_{1},\dots,k_{N})\in\{0,\dots,K\}^{N}\\ \times \int_{u_{(k_{1})}^{1}}^{u_{(k_{1}+1)}^{1}} \dots \int_{u_{(k_{N})}^{N}}^{u_{(k_{N}+1)}^{N}} \frac{\left[F_{K}(k_{1},\dots,k_{N}) - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right]^{2}}{F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\left(1 - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right)} \, \mathrm{d}\tilde{u}^{N} \dots \, \mathrm{d}\tilde{u}^{1}.$$

Falls $(k_1, \ldots, k_N) \in \{0, \ldots, K\}^N \setminus \{1, \ldots, K\}^N$, so gilt $F_K(k_1, \ldots, k_N) = 0$ und zusammen mit Teilaussage 1 von Korollar 81 folgt

$$\begin{split} &u_{(k_{1}+1)}^{1} \dots \int_{u_{(k_{N})}^{N}}^{u_{(k_{N}+1)}^{N}} \frac{\left[F_{K}(k_{1},\dots,k_{N})-F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right]^{2}}{F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\left(1-F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right)} \,\mathrm{d}\tilde{u}^{N}\dots \,\mathrm{d}\tilde{u}^{1} \\ &= \int_{u_{(k_{1})}}^{u_{(k_{1})}^{1}} \dots \int_{u_{(k_{N})}}^{u_{(k_{N})}^{N}} \left[\frac{1}{1-F_{\mathcal{U}([0,1]^{N})}(\tilde{u})}-1\right] \,\mathrm{d}\tilde{u}^{N}\dots \,\mathrm{d}\tilde{u}^{1} \\ &= \sum_{\iota_{1},\dots,\iota_{N}\in\{1,2\}} (-1)^{\sum_{n=1}^{N}\iota_{n}} \operatorname{Li}_{N}\left(\prod_{n=1}^{N}u_{(k_{n}-1+\iota_{n})}^{n}\right) - \prod_{n=1}^{N}\left(u_{(k_{n}+1)}^{n}-u_{(k_{n})}^{n}\right) \\ &= (1-F_{K}(k_{1},\dots,k_{N}))^{2}\sum_{\iota_{1},\dots,\iota_{N}\in\{1,2\}} (-1)^{\sum_{n=1}^{N}\iota_{n}} \operatorname{Li}_{N}\left(\prod_{n=1}^{N}u_{(k_{n}-1+\iota_{n})}^{n}\right) \\ &-\prod_{n=1}^{N}\left(u_{(k_{n}+1)}^{n}-u_{(k_{n})}^{n}\right). \end{split}$$

Falls $(k_1, \ldots, k_N) \in \{1, \ldots, K\}^N \setminus \{K, \ldots, K\}$, so gilt $u_{(k_n)}^n > 0$ für alle $n \in \{1, \ldots, N\}$ und $F_K(k_1, \ldots, k_N) \in [0, 1)$. Zusammen mit Teilaussage 2 von Korollar 81 berechnet man

$$\begin{split} & \int_{u_{(k_{1}+1)}}^{u_{(k_{1}+1)}^{1}} \dots \int_{u_{(k_{N})}^{N}}^{u_{(k_{N}+1)}^{N}} \frac{\left[F_{K}(k_{1},\dots,k_{N}) - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right]^{2}}{F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\left(1 - F_{\mathcal{U}([0,1]^{N})}(\tilde{u})\right)} \, \mathrm{d}\tilde{u}^{N} \dots \, \mathrm{d}\tilde{u}^{1} \\ & = \left(F_{K}(k_{1},\dots,k_{N})\right)^{2} \prod_{n=1}^{N} \log\left(\frac{u_{(k_{n}+1)}^{n}}{u_{(k_{n})}^{n}}\right) \\ & + \left(1 - F_{K}(k_{1},\dots,k_{N})\right)^{2} \sum_{\iota_{1},\dots,\iota_{N} \in \{1,2\}} (-1)^{\sum_{n=1}^{N} \iota_{n}} \operatorname{Li}_{N}\left(\prod_{n=1}^{N} u_{(k_{n}-1+\iota_{n})}^{n}\right) \\ & - \prod_{n=1}^{N} \left(u_{(k_{n}+1)}^{n} - u_{(k_{n})}^{n}\right). \end{split}$$

Falls $(k_1, ..., k_N) = \{K, ..., K\}$, so gilt $u_{(k_n)}^n > 0$ für alle $n \in \{1, ..., N\}$

und $F_K(k_1, \ldots, k_N) = 1$. Zusammen mit Teilaussage 3 von Korollar 81 folgt

Zusammengefasst ergibt sich

$$A_{K,1}^{2} = K \left[\sum_{(k_{1},...,k_{N})\in\{1,...,K\}^{N}} (F_{K}(k_{1},...,k_{N}))^{2} \prod_{n=1}^{N} \log\left(\frac{u_{(k_{n}+1)}^{n}}{u_{(k_{n})}^{n}}\right) + \sum_{(k_{1},...,k_{N})\in\{0,...,K\}^{N}\setminus\{K,...,K\}} (1 - F_{K}(k_{1},...,k_{N}))^{2} \times \sum_{\iota_{1},...,\iota_{N}\in\{1,2\}} (-1)^{\sum_{n=1}^{N}\iota_{n}} \operatorname{Li}_{N}\left(\prod_{n=1}^{N} u_{(k_{n}-1+\iota_{n})}^{n}\right) - \sum_{(k_{1},...,k_{N})\in\{0,...,K\}^{N}} \prod_{n=1}^{N} (u_{(k_{n}+1)}^{n} - u_{(k_{n})}^{n}) \right].$$

Aus

$$\sum_{(k_1,\dots,k_N)\in\{0,\dots,K\}^N} \prod_{n=1}^N \left(u_{(k_n+1)}^n - u_{(k_n)}^n \right) = 1$$

folgt schließlich die Behauptung.

5 Implementierung der Edgeworth-Approximation

Die Implementierung der Edgeworth-Approximation erfolgte in MATLAB und CUDA und wurde in der Form einer Toolbox umgesetzt. Dabei ist eine Toolbox eine Zusammenstellung von logisch zusammengehörenden Dateien und Teilprogrammen.

MATLAB ist eine Programmiersprache der vierten Generation und bietet dem Anwender komfortable Möglichkeiten, numerische und symbolische Berechnungen mit relativ wenigen Code-Zeilen zu implementieren.

CUDA ist eine Programmiertechnik, mit der Programmteile auf einen oder mehrere Grafikprozessoren (GPUs) ausgelagert werden können, was den Vorteil hat, dass durch den Einsatz von GPUs zusätzliche Rechenkapazität zur Verfügung steht. Dabei arbeitet eine GPU im Allgemeinen bei hochgradig parallelisierbaren Programmabläufen signifikant schneller als ein normaler Prozessor (CPU). Ein Nachteil ist jedoch, dass die Programmierung mit CUDA sehr kompliziert und aufwändig ist, weil der Quellcode exakt an die Hardwarearchitektur der verwendeten GPU angepasst sein muss.

Die Edgeworth-Toolbox wurde im Rahmen der vorliegenden Arbeit zum Zweck der Modellkalibrierung entwickelt. Sie ist so konzipiert, dass numerische Operationen, die sowohl rechenintensiv als auch zeitkritisch sind, unter Verwendung von CUDA auf zwei GPUs ausgelagert werden. Alle anderen Operationen werden in MATLAB auf der CPU ausgeführt. Verglichen mit der Rechengeschwindigkeit, die bei ausschließlicher Verwendung einer CPU erreicht wird, wird durch den zusätzlichen Einsatz der beiden GPUs eine Geschwindigkeitsverbesserung in der Größenordnung von ca. 10 realisiert. So wurde die Rechenzeit für die durchgeführten Kalibrierungen auf ca. 1/10 reduziert. Diese betrug jedoch immer noch mehrere Wochen.

Der Quellcode der Edgeworth-Toolbox ist sehr komplex und umfangreich und umfasst einige zehntausend Zeilen. Deshalb werden im vorliegenden Abschnitt 5 nur die wichtigsten Eckpunkte zur Implementierung grob skizziert. Die vollständige Programmarchitektur inklusive der Klassenstrukturen sowie alle Funktionalitäten und Algorithmen sind in einem Handbuch ausführlich dargestellt.

In Abschnitt 5.1 werden einige technischen Lemmata abgehandelt, die zur Aufbereitung von gewissen Teilrechnungen dienen. In ihrer jeweiligen transformierten Form sind diese Teilrechnungen besser zur Programmierung geeignet oder mit weniger numerischem Rechenaufwand zu lösen. In Abschnitt 5.2 werden die wesentlichen programmiertechnischen Schwierigkeiten kurz angedeutet und die Code-Generierung mit der Symbolic Math Toolbox von MATLAB als Lösungsansatz vorgestellt. Ferner wird die Funktionsweise einiger spezieller Programme der Edgeworth-Toolbox skizziert. Hierbei handelt es sich um Programme, die selbst Programme schreiben.

Thema von Abschnitt 5.3 ist die konkrete Implementierung von Programmen zur approximativen Berechnung der Dichte, der Verteilungsfunktion und der bedingten Verteilungsfunktion eines Zufallsvektors X. Unter anderem wird aufgezeigt, warum es zweckdienlicher ist, die Edgeworth-Approximation nicht direkt auf X anzuwenden, sondern stattdessen X zunächst durch eine geeignete affine Transformation zu standardisieren und auf den so entstandenen Zufallsvektor dann die Edgeworth-Approximation anzuwenden. Weiterhin werden die hierfür umgesetzten Programmablaufpläne vorgestellt.

In Abschnitt 5.4 werden die Programme zur approximativen Berechnung der Log-Likelihood, der Rosenblatt-Transformation sowie der Teststatistiken der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests inklusive ihrer Programmablaufpläne präsentiert. Diese greifen bei ihren Berechnungen auf die im vorherigen Abschnitt 5.3 beschriebenen Programme zurück.

Abschnitt 5.5 geht schließlich auf die Stärken und Schwächen von GPUs ein und an welchen Stellen der Edgeworth-Toolbox GPUs gewinnbringend eingesetzt werden.

5.1 Technische Lemmata

Lemma 83. Seien $N \in \mathbb{N}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor. Ferner seien $A \in \mathbb{R}^{N \times N}$ eine invertierbare Matrix, $b \in \mathbb{R}^N$ und der Zufallsvektor A(X - b) habe die Lebesgue-Dichte $f_{A(X-b)}$.

Dann hat X die Lebesgue-Dichte $f_X : \mathbb{R}^N \to \mathbb{R}$,

$$f_X(x) = |\det(A)| f_{A(X-b)} (A(x-b)).$$

Lemma 84. Seien $N \in \mathbb{N}$ und $\Sigma \in \mathbb{R}^{N \times N}$ eine symmetrische und strikt positiv definite Matrix.

Dann gilt

$$\int_{(-\infty,x]} H_{\alpha}(\tilde{x},\Sigma^{-1}) \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) \,\mathrm{d}\tilde{x} = \left(-\frac{\partial}{\partial x}\right)^{\alpha} \Phi_{\mathcal{N}(0,\Sigma)}(x)$$

für alle $x \in \mathbb{R}^N$ und $\boldsymbol{\alpha} \in \mathbb{N}_0^N$. Hierbei ist $H_{\boldsymbol{\alpha}}(\cdot, \Sigma^{-1})$ das in Abschnitt 3.3 eingeführte multivariate Hermite-Polynom.

Im Spezialfall N = 1 gilt

$$\left(-\frac{\partial}{\partial x}\right)^{\alpha}\Phi_{\mathcal{N}(0,1)}(x) = \begin{cases} \Phi_{\mathcal{N}(0,1)}(x) & \text{für } \alpha = 0\\ -He_{\alpha-1}(x)\phi_{\mathcal{N}(0,1)}(x) & \text{für } \alpha \ge 1 \end{cases}$$

für alle $x \in \mathbb{R}$ und $\alpha \in \mathbb{N}_0$, wobei $He_{\alpha-1}$ das $\alpha - 1$ -te eindimensionale Hermite-Polynom zur Gewichtungsfunktion $\mathbb{R} \ni y \mapsto e^{-y^2/2}$ ist.

Die Beweise der Lemmata 83 und 84 sind trivial und werden ausgelassen.

Lemma 85. Sei $\Sigma \in \mathbb{R}^{2 \times 2}$ eine symmetrische und strikt positiv definite Matrix mit inverser Matrix $\Sigma^{-1} = (\Sigma_{n_1,n_2}^{-1})_{n_1,n_2 \in \{1,2\}}$.

Dann gilt

$$\frac{\partial}{\partial y} \Phi_{\mathcal{N}(0,\Sigma)}(x,y) = 2^{-\frac{3}{2}} \sqrt{\frac{\det(\Sigma^{-1})}{\pi \Sigma_{1,1}^{-1}}} \exp\left(-\frac{1}{2} \left[\Sigma_{2,2}^{-1} - \frac{\left(\Sigma_{2,1}^{-1}\right)^2}{\Sigma_{1,1}^{-1}}\right] y^2\right) \left[1 + \operatorname{erf}(t)\right]$$

für alle $(x, y) \in \mathbb{R}^2$, wobei

$$t := \sqrt{\frac{\sum_{1,1}^{-1}}{2}} x + \frac{\sum_{2,1}^{-1}}{\sqrt{2\sum_{1,1}^{-1}}} y.$$

Ferner bezeichnet $\operatorname{erf} : \mathbb{R} \to \mathbb{R}$ mit

$$\operatorname{erf}(t) := \frac{2}{\sqrt{\pi}} \int_0^t e^{-\tilde{t}^2} \,\mathrm{d}\tilde{t}$$

die Fehlerfunktion.

Beweis: Eine einfache Rechnung ergibt

$$-\frac{1}{2}(\tilde{x},y)\Sigma^{-1}\begin{pmatrix}\tilde{x}\\y\end{pmatrix} = -\left(\sqrt{\frac{\Sigma_{1,1}^{-1}}{2}}\,\tilde{x} + \frac{\Sigma_{2,1}^{-1}}{\sqrt{2\,\Sigma_{1,1}^{-1}}}\,y\right)^2 - \frac{1}{2}\left[\Sigma_{2,2}^{-1} - \frac{\left(\Sigma_{2,1}^{-1}\right)^2}{\Sigma_{1,1}^{-1}}\right]y^2$$

für alle $(\tilde{x}, y) \in \mathbb{R}^2$. Hiermit und unter Verwendung der linearen Substitution

$$\tilde{t} = \sqrt{\frac{\Sigma_{1,1}^{-1}}{2}}\,\tilde{x} + \frac{\Sigma_{2,1}^{-1}}{\sqrt{2\,\Sigma_{1,1}^{-1}}}\,y$$

erhält man

also die Behauptung.

Lemma 86. Seien $N \in \mathbb{N} \setminus \{1\}$ und $X = (X^1, \ldots, X^N)^T$ ein \mathbb{R}^N -wertiger Zufallsvektor, dessen Verteilung die stetige Lebesgue-Dichte $f_X : \mathbb{R}^N \to \mathbb{R}$ besitzt. Weiterhin sei $x = (x^1, \ldots, x^N)^T \in \mathbb{R}^N$ derart, dass

$$\int_{-\infty}^{\infty} f_X\left((\tilde{x}^1, x^2, \dots, x^N)^T\right) \, \mathrm{d}\tilde{x}^1 < \infty$$

gilt. Ferner seien $A = (a_{n_1,n_2})_{n_1,n_2 \in \{1,\dots,N\}}$ eine Matrix mit $a_{1,1} > 0$ und $a_{n_1,1} = 0$ für $n_1 \in \{2,\dots,N\}$, $b = (b^1,\dots,b^N)^T \in \mathbb{R}^N$, $Y = (Y^1,\dots,Y^N)^T := A(X-b)$ und $y = (y^1,\dots,y^N)^T := A(x-b)$.

Dann gilt

$$F_{X^1|X^2,\dots,X^N}(x^1|x^2,\dots,x^N) = F_{Y^1|Y^2,\dots,Y^N}(y^1|y^2,\dots,y^N).$$

Beweis: Es sei $A_{1,1} := (a_{n_1+1,n_2+1})_{n_1,n_2 \in \{1,\ldots,N-1\}} \in \mathbb{R}^{(N-1) \times (N-1)}$ die Matrix, die aus A durch Streichung der ersten Zeile und ersten Spalte hervorgeht. Hiermit gilt $(Y^2, \ldots, Y^N)^T = A_{1,1}(X^2 - b^2, \ldots, X^N - b^N)^T$ und $(y^2, \ldots, y^N)^T = A_{1,1} \left(x^2 - b^2, \ldots, x^N - b^N\right)^T$. Mit Lemma 83 und der linearen Substitution $\tilde{y}^1 = a_{1,1} \tilde{x}^1 - a_{1,1} b^1 + \sum_{n=2}^N a_{1,n} (x^n - b^n)$ berechnet man

$$\int_{-\infty}^{x^{1}} f_{X} \left((\tilde{x}^{1}, x^{2}, \dots, x^{N})^{T} \right) d\tilde{x}^{1}$$

$$= |\det(A)| \int_{-\infty}^{x^{1}} f_{Y} \left(A \left((\tilde{x}^{1}, x^{2}, \dots, x^{N})^{T} - b \right) \right) d\tilde{x}^{1}$$

$$= \frac{|\det(A)|}{a_{1,1}} \int_{-\infty}^{y^{1}} f_{Y} \left((\tilde{y}^{1}, y^{2}, \dots, y^{N})^{T} \right) d\tilde{y}^{1}$$

sowie

$$\int_{-\infty}^{\infty} f_X \left((\tilde{x}^1, x^2, \dots, x^N)^T \right) d\tilde{x}^1$$

= $\frac{|\det(A)|}{a_{1,1}} \int_{-\infty}^{\infty} f_Y \left((\tilde{y}^1, y^2, \dots, y^N)^T \right) d\tilde{y}^1$
= $\frac{|\det(A)|}{a_{1,1}} f_{(Y^2, \dots, Y^N)^T} \left((y^2, \dots, y^N)^T \right),$

wobei $f_Y : \mathbb{R}^N \to \mathbb{R}$ und $f_{(Y^2, \dots, Y^N)^T} : \mathbb{R}^{N-1} \to \mathbb{R}$ die Lebesgue-Dichten der Zufallsvektoren Y und $(Y^2, \dots, Y^N)^T$ sind. Daraus folgt

$$F_{X^{1}|X^{2},...,X^{n}}(x^{1}|x^{2},...,x^{N}) = \frac{\int_{-\infty}^{x^{1}} f_{X}\left((\tilde{x}^{1},x^{2},...,x^{N})^{T}\right) d\tilde{x}^{1}}{\int_{-\infty}^{\infty} f_{X}\left((\tilde{x}^{1},x^{2},...,x^{N})^{T}\right) d\tilde{x}^{1}}$$
$$= \frac{\int_{-\infty}^{y^{1}} f_{Y}\left((\tilde{y}^{1},y^{2},...,y^{n})^{T}\right) d\tilde{y}^{1}}{f_{(Y^{2},...,Y^{N})^{T}}\left((y^{2},...,y^{n})^{T}\right)}$$
$$= F_{Y^{1}|Y^{2},...,Y^{n}}(y^{1}|y^{2},...,y^{n}),$$

also die Behauptung.

Lemma 87. Seien $N \in \mathbb{N} \setminus \{1\}$ und X ein \mathbb{R}^N -wertiger Zufallsvektor, der der Annahme ($\mathbb{E}\mathbb{M}$) zu einer strikt positiven Konstante genügt und für den $\mathbb{E}[X] = 0$ sowie $\operatorname{Cov}(X) = E_N$ gilt. Weiterhin sei $x \in \mathbb{R}^N$ und für $S \in \mathbb{N}$ seien

$$\mathcal{Z}_{X}^{S}(x^{1},\ldots,x^{N}) := \sum_{\boldsymbol{\eta}\in T_{S}} \frac{1}{\boldsymbol{\eta}!} \sum_{\substack{\boldsymbol{\alpha}_{1},\ldots,\boldsymbol{\alpha}_{|\boldsymbol{\eta}|}\in\mathbb{N}_{0}^{N},\\|\boldsymbol{\alpha}_{r}|=I_{\boldsymbol{\eta}}^{r}+2 \ f\ddot{\boldsymbol{u}}r \ r\in\{1,\ldots,|\boldsymbol{\eta}|\}}} \left(\prod_{r=1}^{|\boldsymbol{\eta}|} \frac{\kappa_{\boldsymbol{\alpha}_{r}}^{X}}{\boldsymbol{\alpha}_{r}!}\right) \\
\times \left(\prod_{n=2}^{N} He_{\sum_{r=1}^{|\boldsymbol{\eta}|} \alpha_{r}^{n}}(x^{n})\right) \left(-\frac{\partial}{\partial x^{1}}\right)^{\sum_{r=1}^{|\boldsymbol{\eta}|} \alpha_{r}^{1}} \Phi_{\mathcal{N}(0,1)}(x^{1}) \quad (26)$$

sowie

$$\mathcal{N}_{X}^{S}(x^{2},\ldots,x^{N}) := \sum_{\boldsymbol{\eta}\in T_{S}} \frac{1}{\boldsymbol{\eta}!} \sum_{\substack{\boldsymbol{\alpha}_{1}^{1}=\ldots=\boldsymbol{\alpha}_{|\boldsymbol{\eta}|}^{1}=0, |\boldsymbol{\alpha}_{r}|=I_{\boldsymbol{\eta}}^{r}+2 \text{ für } r\in\{1,\ldots,|\boldsymbol{\eta}|\}}} \left(\prod_{r=1}^{|\boldsymbol{\eta}|} \frac{\kappa_{\boldsymbol{\alpha}_{r}}^{X}}{\boldsymbol{\alpha}_{r}!}\right) \times \left(\prod_{n=2}^{N} He_{\sum_{r=1}^{|\boldsymbol{\eta}|} \alpha_{r}^{n}}(x^{n})\right).$$
(27)

Ferner sei $S_{max} \in \mathbb{N}_0$ vorgegeben.

Dann gilt für die Edgeworth-Approximation der bedingten Verteilungsfunktion von X^1 gegeben (X^2, \ldots, X^N) mit S_{max} Korrekturtermen

$$F_{X^{1}|X^{2},...,X^{N}}^{S_{max}}\left(x^{1} \mid x^{2},...,x^{N}\right) = \frac{\Phi_{\mathcal{N}(0,1)}(x^{1}) + \sum_{S=1}^{S_{max}} \mathcal{Z}_{X}^{S}(x^{1},...,x^{N})}{1 + \sum_{S=1}^{S_{max}} \mathcal{N}_{X}^{S}(x^{2},...,x^{N})}.$$

Beweis: Eine einfache Rechnung ergibt

$$\frac{1}{\phi_{\mathcal{N}(0,E_{N-1})}(x^{2},\ldots,x^{N})} \int_{-\infty}^{x^{1}} f_{X}^{S_{max}} \left(\tilde{x}^{1},x^{2},\ldots,x^{N}\right) \,\mathrm{d}\tilde{x}^{1}$$
$$= \Phi_{\mathcal{N}(0,1)}(x^{1}) + \sum_{S=1}^{S_{max}} \mathcal{Z}_{X}^{S}(x^{1},\ldots,x^{N})$$

und

$$\frac{f_{(X^2,\dots,X^N)}^{S_{max}}\left(x^2,\dots,x^N\right)}{\phi_{\mathcal{N}(0,E_{N-1})}(x^2,\dots,x^N)} = 1 + \sum_{S=1}^{S_{max}} \mathcal{N}_X^S(x^2,\dots,x^N).$$

Hierbei sind $f_X^{S_{max}}$ und $f_{(X^2,...,X^N)}^{S_{max}}$ die Edgeworth-Approximationen der Dichten von X und $(X^2,...,X^N)$, jeweils mit S_{max} Korrekturtermen. Die Behauptung ergibt sich nun aus der Definition der Edgeworth-Approximation der bedingten Verteilungsfunktion und durch Kürzen von $\phi_{\mathcal{N}(0,E_{N-1})}(x^2,...,x^N)$.

5.2 Code-Generierung mit der Symbolic Math Toolbox

Von hier an bis zum Ende des Abschnitts 5 seien $D, N \in \mathbb{N}$ und $L = (L_t)_{t \in [0,T^*]}$ ein *D*-dimensionaler Lévy-Prozess, der der Annahme ($\mathbb{E}\mathbb{M}$) zu einer strikt positiven Konstante genügt. Weiterhin seien $f_1, \ldots, f_N : [0, T^*] \rightarrow \mathbb{R}^D$ stetige Funktionen, $\Delta_1, \Delta_2 \in [0, T^*]$ mit $\Delta_1 < \Delta_2$ und X sei der \mathbb{R}^N -wertige Zufallsvektor mit den Komponenten

$$X^n := \int_{\Delta_1}^{\Delta_2} f_n(s) \, \mathrm{d}L_s$$

für $n \in \{1, \ldots, N\}$. Ferner sei die Kovarianzmatrix $\Sigma := \text{Cov}(X)$ strikt positiv definit.

Abgesehen von der Annahme, dass Σ strikt positiv definit ist, sind das die Voraussetzungen von Satz 35.

Der Weg zur approximativen Berechnung der Dichte, der Verteilungsfunktion und der bedingten Verteilungsfunktion von X unter Verwendung der Edgeworth-Approximation, so wie er im nächsten Abschnitt 5.3 beschrieben wird und im Rahmen der Edgeworth-Toolbox implementiert wurde, setzt sich aus mehreren Einzelrechnungen zusammen. Die Herausforderung hierbei liegt nicht in der Komplexität der Einzelrechnungen, weil jede von ihnen aus mathematischer Sicht in kanonischer Weise abgearbeitet werden kann. Stattdessen besteht die Herausforderung darin, dass eine wirklich sehr große Anzahl derartiger Einzelrechnungen auszuführen ist. Hierfür werden nun zwei Beispiele angegeben.

Beispiel 1: Als ein Zwischenschritt sind unter Verwendung von Teilaussage 2 von Satz 35 für jedes $K \in \{1, \ldots, S_{max} + 2\}$, wobei $S_{max} \in \mathbb{N}_0$ die verwendete Anzahl an Edgeworth-Korrekturtermen ist, die Kumulanten κ_{α}^X für alle Multiindices $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| = K$ zu berechnen. In Tabelle 2 ist die Anzahl dieser Multiindices in Abhängigkeit von N und K dargestellt. Für jedes solche $\boldsymbol{\alpha}$ wird $\kappa_{\boldsymbol{\alpha}}^{X}$ als Summe aus D^{K} Summanden berechnet. Dabei erscheint in jedem dieser Summanden ein bestimmtes Integral. Tabelle 3 zeigt in Abhängigkeit von D, N und K die Anzahl dieser Integrale, die zur Berechnung der $\kappa_{\boldsymbol{\alpha}}^{X}$ für alle $\boldsymbol{\alpha} \in \mathbb{N}_{0}^{N}$ mit $|\boldsymbol{\alpha}| = K$ bestimmt werden müssen. Eine numerische Integration scheidet aus, weil diese bei einer Kalibrierung zu einer Rechenzeit in der Größenordnung von Jahren führen würde. Für die verwendeten Integranden f_{1}, \ldots, f_{N} , auf die in Abschnitt 6 genauer eingegangen wird, existieren für die Integrale jedoch geschlossene Formen. Deren analytische Berechnung und anschließende Programmierung sind zwar prinzipiell möglich, aber sehr aufwändig. Ferner kommt bei einer Programmierung von Hand das Risiko von Tippfehlern hinzu, das gerade bei langen Formeln ein ernstzunehmendes Problem darstellt.

N	K = 1	K = 2	K = 3	K = 4	K = 5	K = 6
1	1	1	1	1	1	1
2	2	3	4	5	6	7
3	3	6	10	15	21	28
4	4	10	20	35	56	84
5	5	15	35	70	126	210

D	N	K = 1	K = 2	K = 3	K = 4	K = 5	K = 6
1	1	1	1	1	1	1	1
	2	2	3	4	5	6	7
	3	3	6	10	15	21	28
	4	4	10	20	35	56	84
	5	5	15	35	70	126	210
2	1	2	4	8	16	32	64
	2	4	12	32	80	192	448
	3	6	24	80	240	672	1792
	4	8	40	160	560	1792	5376
	5	10	60	280	1120	4032	13440
3	1	3	9	7	81	243	729
	2	6	27	108	405	1458	5103
	3	9	54	270	1215	5103	20412
	4	12	90	540	2835	13608	61236
	5	15	135	945	5670	30618	153090

Tabelle 2: Anzahl der Multi
indices $\pmb{\alpha} \in \mathbb{N}_0^N$ mit $|\pmb{\alpha}| = K$

Tabelle 3: Anzahl der Integrale bei Berechnung der $\kappa^X_{\pmb{\alpha}}$ für alle $\pmb{\alpha}\in\mathbb{N}_0^N$ mit $|\pmb{\alpha}|=K$

Beispiel 2: Bei der approximativen Berechnung der Dichte von X sind in einem weiteren Rechenschritt die Edgeworth-Korrekturterme \mathcal{K}_X^S auszuwerten und bei der approximativen Berechnung der bedingten Verteilungsfunktion von X sind die Edgeworth-Korrekturterme \mathcal{Z}_X^S und \mathcal{N}_X^S der bedingten Verteilungsfunktion auszuwerten. Die Berechnungen erfolgen dabei jeweils für alle $S \in \{1, \ldots, S_{max}\}$. Betrachtet man die Definitionsgleichungen (14), (26) und (27) von \mathcal{K}_X^S , \mathcal{Z}_X^S und \mathcal{N}_X^S und nimmt man ferner an, dass die nume-rischen Werte von x, Σ^{-1} und aller darin auftretenden Kumulanten bekannt sind, so hat jeder der auftretenden Summanden eine recht einfache Struktur. Bei den Summanden in \mathcal{K}_X^S und \mathcal{N}_X^S handelt es sich lediglich um Polynome. Bei den Summanden von \mathcal{Z}_X^S taucht noch ein zusätzlicher Faktor auf, der ohne Schwierigkeiten numerisch auswertbar ist. Eine algorithmische Berechnung der konkreten Form der Summanden in \mathcal{K}_X^S , \mathcal{Z}_X^S und \mathcal{N}_X^S ist problemlos möglich. Für nicht allzu kleine $N, S \in \mathbb{N}$ ist die Anzahl der Summanden jedoch sehr groß, siehe hierzu Tabelle 4. Falls die jeweiligen konkreten Formen bei jedem Funktionsaufruf eigens dafür berechnet würden, so führte das zu einer unnötig langen Rechenzeit. Stattdessen ist es sinnvoll, die konkreten Formen nur einmal zu berechnen und dann bei allen Funktionsaufrufen auf die bereits berechneten Formen zurückzugreifen.

N	S = 1	S=2	S=3	S = 4	S = 5	S = 6
1	1	2	3	5	7	11
2	4	21	90	392	1.606	6.629
3	10	115	1.171	11.963	119.981	$1,20\cdot 10^6$
4	20	435	8.756	176.429	$3,53\cdot10^6$	$7,07\cdot 10^7$
5	35	1.295	45.451	$1,60 \cdot 10^{6}$	$5,59 \cdot 10^{7}$	$1,96\cdot 10^9$

Tabelle 4: Anzahl Summanden im Edgeworth-Korrekturter
m \mathcal{K}^S_X der Dichte und in den Edgeworth-Korrekturterme
n \mathcal{Z}^S_X und \mathcal{N}^S_X der bedingten Verteilungsfunktion

Zur Lösung der dargestellten Probleme wird das Konzept der Code-Generierung eingesetzt. Für gewisse Formeln, die zwar aus mathematischer Sicht eine einfache Struktur haben, aber sehr lang sind, wurden Programme geschrieben, die Programme zur numerischen Auswertung dieser Formeln schreiben. Diese generierten Programme werden in einer zentralen Kartei gespeichert. Jedes Mal, wenn eine der besagten Formeln numerisch auszuwerten ist, wird dann auf die zentrale Kartei zugegriffen und das entsprechende Programm aufgerufen. Insbesondere müssen diese Programme nur einmal und nicht bei jeder numerischen Auswertung der betrachteten Formel erneut generiert werden.

Die Symbolic Math Toolbox von MATLAB bietet die erforderlichen Funktionalitäten aus der symbolischen Mathematik. Im Gegensatz zu numerischen Berechnungen stehen in der symbolischen Mathematik keine konkreten numerischen Zahlenwerte im Fokus. Stattdessen stehen Terme im Vordergrund, die Variablen im mathematischen Sinne enthalten. Ein derartiger Term wird symbolischer Ausdruck genannt. Die Edgeworth-Toolbox macht dabei insbesondere von den folgenden MATLAB-Befehlen Gebrauch.

- diff: Differentiation eines symbolischen Ausdrucks
- int: Integration eines symbolischen Ausdrucks
- simplify: Algebraische Vereinfachung eines symbolischen Ausdrucks
- matlabFunction: Konvertierung von symbolischen Ausdrücken in MATLAB-Code

Im Folgenden werden die Programme der Edgeworth-Toolbox, die Programme zur numerischen Auswertung von langen Formeln schreiben, kurz beschrieben. Dabei greifen die Programme, die die Programme zur Auswertung von \mathcal{K}_X^S , \mathcal{Z}_X^S und \mathcal{N}_X^S schreiben, und das Programm, das die Programme zur Auswertung von $\int_{(-\infty,x]} \mathcal{K}_X^S(\tilde{x}) \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) d\tilde{x}$ schreibt, auf einen Algorithmus zur Berechnung der Lösungsmenge T_S der diophantischen Gleichung aus Definition 42 zurück, der in Blinnikov und Moessner (1998, Appendix C) beschrieben ist und im Rahmen der vorliegende Arbeit in MATLAB implementiert wurde.

5.2.1 Das Programm

Erstelle_m_file_fuer_KumulantenTreibenderProzess

Die Eingabeargumente sind der Namen der Verteilung, von der der treibende Lévy-Prozess L erzeugt wird, die Dimension D und zusätzlich eine Zahl $K \in \mathbb{N}$. Unterstützt werden dabei die Normalverteilung, die GH-Verteilung, die NIG-Verteilung und die VG-Verteilung.

Unter Verwendung von diff werden symbolisch alle Ableitungen K-ter Ordnung der kumulantenerzeugenden Funktion K_{L_1} berechnet. Diese werden an der Stelle u = 0 ausgewertet und mit simplify algebraisch vereinfacht. Aus den so erhaltenen symbolischen Ausdrücken wird mit matlabFunction ein m-file generiert, das aus den numerischen Werten der Parameter von $\mathcal{L}(L_1)$ die erwünschten numerischen Werte der Kumulanten $\kappa_{\alpha}^{L_1}$ von L_1 der Ordnungen $\boldsymbol{\alpha} \in \mathbb{N}_0^D$ mit $|\boldsymbol{\alpha}| = K$ berechnet.
5.2.2 Das Programm

Erstelle_m_file_fuer_KumulantenIntegralProzess

Es gelten die Voraussetzungen und die Bezeichnungen von Satz 35, Teilaussage 2.

Als Eingabeargumente verlangt die Funktion symbolische Ausdrücke für die Integranden $f_1, \ldots, f_N : [0, T^*] \to \mathbb{R}^D$ und eine Zahl $K \in \mathbb{N}$.

Unter Verwendung von int werden hieraus symbolische Ausdrücke für die Kumulanten κ_{α}^{X} von X der Ordnungen $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| = K$ berechnet und mit simplify algebraisch vereinfacht. Aus den so erhaltenen symbolischen Ausdrücken wird mit matlabFunction ein m-file generiert, das aus den numerischen Werten der Integrationsgrenzen Δ_{1} und Δ_{2} , der Kumulanten $\kappa_{\alpha}^{L_{1}}$ von L_{1} der Ordnungen $\alpha \in \mathbb{N}_{0}^{D}$ mit $|\alpha| = K$ und der Parameter der Integranden f_{1}, \ldots, f_{N} die erwünschten numerischen Werte der Kumulanten κ_{α}^{X} von X der Ordnungen $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| = K$ berechnet.

5.2.3 Das Programm Erstelle_m_file_fuer_EdgeworthCdfKorrekturterm

Es gelten die Voraussetzungen und die Bezeichnungen von Definition 47.

Als Eingabeargumente verlangt die Funktion die Dimension N und eine Zahl $S \in \mathbb{N}$. Dabei werden für das Eingabeargument N die Werte 1 und 2 unterstützt.

Unter Verwendung von Teilaussage 1 von Lemma 84 erhält man

$$\int_{(-\infty,x]} \mathcal{K}_{X}^{S}(\tilde{x})\phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) \,\mathrm{d}\tilde{x}$$

$$= \sum_{\boldsymbol{\eta}\in T_{S}} \frac{1}{\boldsymbol{\eta}!} \sum_{\substack{\boldsymbol{\alpha}_{1},\dots,\boldsymbol{\alpha}_{|\boldsymbol{\eta}|}\in\mathbb{N}_{0}^{N},\\|\boldsymbol{\alpha}_{r}|=I_{\boldsymbol{\eta}}^{r}+2 \text{ für } r\in\{1,\dots,|\boldsymbol{\eta}|\}} \left(\prod_{r=1}^{|\boldsymbol{\eta}|} \frac{\kappa_{\boldsymbol{\alpha}_{r}}^{X}}{\boldsymbol{\alpha}_{r}!}\right) \left(-\frac{\partial}{\partial x}\right)^{\sum_{r=1}^{|\boldsymbol{\eta}|} \boldsymbol{\alpha}_{r}} \Phi_{\mathcal{N}(0,\Sigma)}(x).$$
(28)

Für diesen Term wird ein symbolischer Ausdruck berechnet. Dabei bereitet der Fall N = 1 keine nennenswerten Schwierigkeiten. Dagegen ist der Fall N = 2 deutlich komplizierter.

Für $N \in \mathbb{N}$ sind in MATLAB effiziente Algorithmen zur numerischen Auswertung der Verteilungsfunktion $\Phi_{\mathcal{N}(0,\Sigma)}$ der N-dimensionalen Normalverteilung bereits vorprogrammiert. Für $N \geq 2$ ist hierfür jedoch keine sogenannte spezielle symbolische Funktion hinterlegt. Stattdessen muss in symbolischen Ausdrücken die Darstellung

$$\Phi_{\mathcal{N}(0,\Sigma)}(x^1,\ldots,x^N) = \int_{-\infty}^{x^1} \ldots \int_{-\infty}^{x^N} \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}^1,\ldots,\tilde{x}^N) \,\mathrm{d}\tilde{x}^N \ldots \,\mathrm{d}\tilde{x}^1$$

mit

$$\phi_{\mathcal{N}(0,\Sigma)}(x^1,\dots,x^N) = \sqrt{\frac{\det(\Sigma^{-1})}{(2\pi)^N}} \exp\left((x^1,\dots,x^N)\Sigma^{-1}(x^1,\dots,x^N)^T\right)$$

verwendet werden, wobei hier die Integrationen unter Verwendung von int realisiert werden. Die gewählte Darstellung für $\phi_{\mathcal{N}(0,\Sigma)}$ hat den Vorteil, dass darin lediglich symbolische Variablen für x^n , $n \in \{1, \ldots, N\}$, sowie für Σ_{n_1,n_2}^{-1} , $n_1, n_2 \in \{1, \ldots, N\}$, auftauchen und somit keine eigenen symbolischen Variablen für Σ_{n_1,n_2} , $n_1, n_2 \in \{1, \ldots, N\}$, berechnet oder eingeführt werden müssen.

Für N = 2 und $\boldsymbol{\alpha} = (\alpha^1, \alpha^2) \in \{(\tilde{\alpha}^1, 0) | \tilde{\alpha}^1 \in \mathbb{N}\} \cup \{(0, \tilde{\alpha}^2) | \tilde{\alpha}^2 \in \mathbb{N}\}$ findet MATLAB unter Verwendung von diff, int und simplify keinen symbolischen Ausdruck für

$$\left(-\frac{\partial}{\partial x},-\frac{\partial}{\partial y}\right)^{\alpha}\Phi_{\mathcal{N}(0,\Sigma)}(x,y),$$

der eine anschließende effiziente numerische Auswertung für gegebene numerische Werte von $(x, y) \in \mathbb{R}^2$ und von $\Sigma^{-1} \in \mathbb{N}^{N \times N}$ ermöglicht. Das Problem wird folgendermaßen gelöst. Für $\boldsymbol{\alpha} = (0, \alpha^2)$ mit $\alpha^2 \in \mathbb{N}$ gilt

$$\left(-\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}\right)^{\alpha} \Phi_{\mathcal{N}(0,\Sigma)}(x,y) = -\left(-\frac{\partial}{\partial y}\right)^{\alpha^2 - 1} \frac{\partial}{\partial y} \Phi_{\mathcal{N}(0,\Sigma)}(x,y).$$

Nun wird $\frac{\partial}{\partial y} \Phi_{\mathcal{N}(0,\Sigma)}(x, y)$ durch die in Lemma 85 gefundene Darstellung ersetzt. Für die darin enthaltene Fehlerfunktion erf : $\mathbb{R} \to \mathbb{R}$ ist eine spezielle symbolische Funktion in MATLAB hinterlegt. Hierdurch ist MATLAB unter Verwendung von **diff** in der Lage, einen symbolischen Ausdruck für $(-\frac{\partial}{\partial x}, -\frac{\partial}{\partial y})^{\alpha} \Phi_{\mathcal{N}(0,\Sigma)}(x, y)$ zu finden, der eine anschließende effiziente numerische Auswertung für gegebene numerische Werte von $(x, y) \in \mathbb{R}^2$ und von $\Sigma^{-1} \in \mathbb{R}^{N \times N}$ ermöglicht. Diese Auswertung erfolgt dann unter Aufruf des in MATLAB bereits vorprogrammierten Algorithmus zur numerischen Auswertung der Fehlerfunktion.

Für $\boldsymbol{\alpha} = (\alpha^1, \alpha^2) \in \mathbb{N}^2$, also $\alpha^1 > 0$ und $\alpha^2 > 0$, erhält man zusammen mit der Definitionsgleichung (11) für das multivariate Hermite-Polynom

$$\left(-\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}\right)^{\alpha} \Phi_{\mathcal{N}(0,\Sigma)}(x,y) = H_{(\alpha^{1}-1,\alpha^{2}-1)}(x,\Sigma^{-1})\phi_{\mathcal{N}(0,\Sigma)}(x,y).$$

Die Berechnung von $H_{(\alpha^1-1,\alpha^2-1)}(x,\Sigma^{-1})$ erfolgt dann durch einen eigens dafür programmierten Algorithmus.

Der insgesamt erhaltene symbolische Ausdruck für den in Gleichung (28) dargestellten Term wird mit **simplify** algebraisch vereinfacht. Anschließend wird mit **matlabFunction** ein m-file generiert, das aus den numerischen Werten von x, der Einträgen von Σ^{-1} und der Kumulanten κ_{α}^{X} von X der Ordnungen $\alpha \in \mathbb{N}_{0}^{0}$ mit $|\alpha| \in \{3, \ldots, S+2\}$ den erwünschten numerischen Wert von $\int_{(-\infty,x]} \mathcal{K}_{X}^{S}(\tilde{x}) \phi_{\mathcal{N}(0,\Sigma)}(\tilde{x}) d\tilde{x}$ berechnet.

5.2.4 Das Programm Erstelle_m_file_fuer_Kumulanten_X_standardisiert

Es gelten die Voraussetzungen und die Bezeichnungen von Satz 32, Teilaussage 2. Zusätzlich wird angenommen, dass D = N sowie $A = \text{Chol}(\text{Cov}(X)^{-1})$ gilt. Insbesondere ist A damit eine obere Dreiecksmatrix.

Als Eingabeargumente verlangt die Funktion die Dimension N und eine Zahl $K \in \mathbb{N}$.

Es werden symbolische Ausdrücke für die Kumulanten κ_{α}^{AX} von AX der Ordnungen $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| = K$ berechnet und mit simplify algebraisch vereinfacht. Hierbei besteht die Vereinfachung lediglich darin, dass diejenigen Summanden, deren Wert 0 ist, herausgestrichen werden. Aus den so erhaltenen symbolischen Ausdrücken wird mit matlabFunction ein m-file generiert, das aus den numerischen Werten der Kumulanten κ_{α}^{X} von X der Ordnungen $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| = K$ und der Einträgen der oberen Dreiecksmatrix Adie erwünschten numerischen Werte der Kumulanten κ_{α}^{AX} von AX der Ordnungen $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| = K$ berechnet.

5.2.5 Das Programm Erstelle_m_file_fuer_EdgeworthKorrekturtermStandardisiert

Es gelten die Voraussetzungen und die Bezeichnungen von Definition 47. Zusätzlich wird angenommen, dass $\Sigma := \text{Cov}(X) = E_N$ gilt.

Als Eingabeargumente verlangt die Funktion die Dimension N und eine Zahl $S \in \mathbb{N}.$

Es wird ein symbolischer Ausdruck für den S-ten Edgeworth-Korrekturterm \mathcal{K}_X^S berechnet und mit simplify algebraisch vereinfacht. Anschließend wird hieraus mit matlabFunction ein m-file generiert, das aus den numerischen Werten der Kumulanten κ_{α}^X von X der Ordnungen $\alpha \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| \in \{3, \dots, S+2\}$ und von x den erwünschten numerischen Wert des S-ten Edgeworth-Korrekturterms $\mathcal{K}_X^S(x)$ berechnet.

5.2.6 Das Programm

Erstelle_m_file_fuer_EdgeworthBedingteCdfStandardisiert

Es gelten die Voraussetzungen und die Bezeichnungen von Definition 47. Zusätzlich wird angenommen, dass $\Sigma := \text{Cov}(X) = E_N$ gilt.

Als Eingabeargumente verlangt die Funktion die Dimension N und eine Zahl $S \in \mathbb{N}$.

Es werden symbolische Ausdrücke für die Größen \mathcal{Z}_X^S und \mathcal{N}_X^S aus Lemma 87 berechnet. Dabei wird der Ausdruck $(-\frac{\partial}{\partial x^1})^{\alpha} \Phi_{\mathcal{N}(0,1)}(x^1)$ unter Verwendung von Teilaussage 2 von Lemma 84 berechnet. Die so erhaltenen symbolischen Ausdrücke für \mathcal{Z}_X^S und \mathcal{N}_X^S werden mit simplify algebraisch vereinfacht. Anschließend werden mit matlabFunction zwei m-files generiert, die aus den numerischen Werten von x und der Kumulanten κ_{α}^X von X der Ordnungen $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S+2\}$ die erwünschten numerischen Werte von $\mathcal{Z}_X^S(x)$ und $\mathcal{N}_X^S(x)$ berechnen.

5.2.7 Laufzeiten der Programme, die Programme schreiben

Bei allen hier vorgestellten Programmen, die Programme schreiben, hängt die benötigte Laufzeit sehr stark von den übergebenen Eingabeargumenten ab. Während die Laufzeiten aller anderen dieser Programme sich in einem vertretbaren Rahmen bewegen, stellt die Laufzeit des Programms Erstelle_m_file_fuer_KumulantenIntegralProzess einen beschränkenden Faktor für die durchgeführten Untersuchungen dar. In Abschnitt 6 wird auf die Kalibrierung des Modells aus Abschnitt 2 eingegangen. Für die dabei auftretenden Integranden betrug die reine Laufzeit zur Erstellung der dazugehörigen Programme zur Berechnung der gemeinsamen Kumulanten der stochastischen Integrale insgesamt ca. zwei Wochen. Die Engpässe liegen beim simplify-Befehl zur algebraischen Vereinfachung der gefundenen symbolischen Ausdrücke sowie beim matlabFunction-Befehl, mit dem schließlich das m-file erstellt wird. Das größte der so erstellten m-files ist 8.280 kB groß.

Einen beschränkenden Faktor stellt diese große Laufzeit insofern dar, dass keine komplizierteren als die verwendeten Volatilitätsstrukturen für die Kalibrierung gewählt werden konnten, weil die Erstellung der dazugehörigen Programme zu lange gedauert hätte. Die verwendeten Volatilitätsstrukturen bieten jedoch die erforderliche Flexibilität. Es konnte jedoch nicht untersucht werden, ob allgemeinere Volatilitätsstrukturen gegenüber den verwendeten überhaupt noch einen Mehrwert bringen.

Wie bereits erwähnt, ist die Edgeworth-Toolbox so konzipiert, dass die Programme, die von Programmen geschrieben werden, nicht bei jeder numerischen Berechnung erneut sondern nur einmal erstellt werden müssen. Nachdem die Erstellung all dieser Programme abgeschlossen war, stellte die dafür benötigte Laufzeit für die im Anschluss ausgeführten numerischen Berechnungen bei der Kalibrierung keinen beschränkenden Faktor mehr dar.

5.3 Approximative Berechnung der Dichte, der Verteilungsfunktion und der bedingten Verteilungsfunktion von stochastischen Integralen

Ziel des vorliegenden Abschnitts ist die Entwicklung von effektiven Verfahren zur approximativen Berechnung der Dichte, der Verteilungsfunktion und der bedingten Verteilungsfunktion des Zufallsvektors X aus Abschnitt 5.2, dessen Einträge stochastische Integrale nach einem Lévy-Prozess sind. Hierzu soll die Edgeworth-Approximation mit $S_{max} \in \mathbb{N}$ Korrekturtermen eingesetzt werden. Bei den Untersuchungen steht insbesondere auch der damit verbundene Rechenaufwand im Fokus.

In Definition 47 wird die Edgeworth-Approximation der Dichte, der Verteilungsfunktion, der bedingten Dichte und der bedingten Verteilungsfunktion von X eingeführt. Die Zusatzvoraussetzung $\mathbb{E}[X] = 0$ stellt dabei kein nennenswertes Problem dar, weil sich diese durch den Übergang von X zu $X - \mathbb{E}[X]$ stets herstellen lässt.

Die multivariaten Hermite-Polynome $H_{\alpha}(\cdot, \Sigma^{-1})$ bereiten jedoch Probleme. Es sind zwar Algorithmen zur Berechnung ihrer konkreten Form bekannt, die deutlich effizienter als die Verwendung der Definitionsgleichung (11) sind. Ein Beispiel hierfür ist die Methode, die im Beweis von Satz 41 hergeleitet wird. Dennoch wird ihre Berechnung mit wachsendem $|\alpha|$ sehr schnell recht aufwändig.

Weiterhin tauchen für nicht allzu kleine $N, S \in \mathbb{N}$ im Edgeworth-Korrekturterm \mathcal{K}_X^S in Gleichung (14) sowie in den Termen \mathcal{Z}_X^S und \mathcal{N}_X^S in den Gleichungen (26) und (27), die zur Berechnung der bedingten Verteilungsfunktion eingesetzt werden, sehr viele verschiedene multivariate Hermite-Polynome auf. Die Berechnung der Form der $H_{\alpha}(\cdot, \Sigma^{-1})$ für alle auftretenden Multiindices α ist zwar aufwändig, aber prinzipiell möglich. Dennoch ist dann die numerische Auswertung von $\mathcal{K}_X^S(x), \mathcal{Z}_X^S(x)$ und $\mathcal{N}_X^S(x)$ für eine gegebene Kombination von numerischen Werten von $x \in \mathbb{R}^N$ und von $\Sigma^{-1} \in \mathbb{R}^{N \times N}$ recht zeitintensiv. Zusammengefasst ist eine direkte Anwendung der Edgeworth-Approximation auf X bzw. auf $X - \mathbb{E}[X]$ lediglich für die Berechnung einiger weniger approximativer Werte geeignet. Im Rahmen der vorliegenden Arbeit wird diese direkte Anwendung nur für die Erstellung von Schaubildern der approximativen Verteilungsfunktion von ein- und zweidimensionalen Randverteilungen eingesetzt. Für eine Kalibrierung ist die direkte Anwendung jedoch nicht praktikabel, weil dieser Ansatz mit einem zu großen Rechenaufwand und damit mit einer zu großen Laufzeit verbunden ist.

Bei der approximativen Berechnung der Dichte und der bedingten Verteilungsfunktion wird dieses Problem umgangen, indem die Edgeworth-Approximation nicht auf X direkt angewandt wird. Stattdessen wird X zunächst durch eine geeignete affine Transformation standardisiert und auf den so entstandenen Zufallsvektor wird dann die Edgeworth-Approximation angewandt. Hierdurch reduziert sich das Problem der multivariaten Hermite-Polynomen auf die Berechnung der Form und auf die numerische Auswertung von multivariaten Hermite-Polynomen zur Einheitsmatrix, also $H_{\alpha}(\cdot, E_N)$. Diese Aufgabe lässt sich unter Verwendung von Lemma 39 mit einem vertretbaren Aufwand lösen.

Es werden nun die Programmablaufpläne präsentiert, die bei der approximativen Berechnung der Dichte, der Verteilungsfunktion und der bedingten Verteilungsfunktion durchlaufen werden. Dabei sind lediglich die nichttrivialen Operationen mit einer Nummer versehen und lediglich diese werden in eigenen Paragraphen genauer erläutert. Ferner werden die mathematischen Ideen beleuchtet, die zur Entscheidung für das verwendete Design geführt haben.

Bemerkung 88. Die in der vorliegenden Arbeit dargestellten Programmablaufplänen haben nicht den Anspruch, dass sie der eigens hierfür eingeführten DIN 66001 (Sinnbilder für Datenfluss- und Programmablaufpläne) genügen. Ihr Zweck ist es lediglich, dem Leser einen anschaulichen Überblick über die Rechenabläufe in der Edgeworth-Toolbox zu geben ohne dabei zu tief ins programmiertechnische Detail zu gehen.

5.3.1 Berechnung der Kumulanten von X

Die Berechnung der Kumulanten von X erfolgt nach dem in Abbildung 1 dargestellten Programmablaufplan.

Vorbedingung: Die numerischen Werte der Integrationsgrenzen Δ_1 und Δ_2 , der Parameter von $\mathcal{L}(L_1)$ und der Parameter der Integranden f_1, \ldots, f_N



Abbildung 1: Programmablaufplan zur Berechnung der Kumulanten von X

seien gegeben.

Operation 1: Zur Berechnung der Kumulanten $\kappa_{\alpha}^{L_1}$ von L_1 der Ordnungen $\alpha \in \mathbb{N}_0^D$ mit $|\alpha| \in \{1, \ldots, S_{max} + 2\}$ werden die Programme eingesetzt, die vom Programm

${\tt Erstelle_m_file_fuer_KumulantenTreibenderProzess}$

aus Abschnitt 5.2.1 geschrieben wurden.

Operation 2: Teilaussage 2 von Satz 35 bildet den Ausgangspunkt für die Berechnung der Kumulanten κ_{α}^{X} von X der Ordnungen $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{1, \ldots, S_{max}+2\}$. Die Operation wird mit den Programmen ausgeführt, die vom Programm

${\tt Erstelle_m_file_fuer_KumulantenIntegralProzess}$

aus Abschnitt 5.2.2 geschrieben wurden.

Ausgabe: Ausgegeben werden die numerischen Werte des Erwartungswertvektors $\mathbb{E}[X]$, der Kovarianzmatrix Σ und der Kumulanten κ_{α}^{X} für $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$.

5.3.2 Approximative Berechnung der Verteilungsfunktion von X

Die Edgeworth-Toolbox unterstützt die approximative Berechnung der Verteilungsfunktion für $N \in \{1, 2\}$. In Abbildung 2 ist der dazugehörige Programmablaufplan dargestellt.

Vorbedingung: Die numerischen Werte des Vektors $\mathbb{E}[X]$, der Matrix Σ und der Kumulanten κ_{α}^{X} für $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ seien bereits nach dem in Abbildung 1 dargestellten Programmablaufplan berechnet. Ferner sei der numerische Wert von $x \in \mathbb{R}^{N}$ gegeben.

Operation 3: Die Operation wird mit den Programmen ausgeführt, die vom Programm

${\tt Erstelle_m_file_fuer_EdgeworthCdfKorrekturterm}$

aus Abschnitt 5.2.3 geschrieben wurden.



Abbildung 2: Programmablaufplan zur approximativen Berechnung der Verteilungsfunktion von X

Operation 4: In Definition 47 wird die Edgeworth-Approximation lediglich für Zufallsvektoren mit $\mathbb{E}[X] = 0$ definiert. Falls $\mathbb{E}[X] \neq 0$ gilt, so wird hier in kanonischer Weise $F_{X-\mathbb{E}[X]}^{S_{max}}(x-\mathbb{E}[X])$ als Approximation für $F_X(x)$ gewählt.

Ausgabe: Ausgegeben wird der numerische Wert der Approximation von $F_X(x)$.

5.3.3 Berechnung der Kumulanten von $\operatorname{Chol}(\Sigma^{-1})(X - \mathbb{E}[X])$

Das nun vorgestellte Lemma bildet die mathematische Grundlage für die bereits angekündigte affine Transformation zur Standardisierung von X.

Lemma 89. Es bezeichne $\operatorname{Chol}(\Sigma^{-1})$ die Cholesky-Zerlegung von Σ^{-1} , also die eindeutig bestimmte obere Dreiecksmatrix aus $\mathbb{R}^{N \times N}$ mit der Eigenschaft $\operatorname{Chol}(\Sigma^{-1})^T \operatorname{Chol}(\Sigma^{-1}) = \Sigma^{-1}$. Ferner sei der N-dimensionale Zufallsvektor Y definiert durch

$$Y := \operatorname{Chol}(\Sigma^{-1})(X - \mathbb{E}[X]).$$

Dann gilt $\mathbb{E}[Y] = 0$ und $\operatorname{Cov}(Y) = E_N$.

Der Beweis von Lemma 89 ist trivial und wird ausgelassen.

Bemerkung 90. Im Fall N = 1 ist

$$Y = \frac{X - \mathbb{E}\left[X\right]}{\sqrt{\operatorname{Var}(X)}}$$

die einzige Zufallsvariable mit $\mathbb{E}[Y] = 0$ und $\operatorname{Var}(Y) = 1$, die aus der Zufallsvariable X durch eine affine Transformation hervorgeht.

Für $N \ge 2$ gibt es mehrere Möglichkeiten, den Zufallsvektor X zu standardisieren. Neben dem Zufallsvektor Y aus Lemma 89 berechnet man beispielsweise für den Zufallsvektor

$$\tilde{Y} := \Sigma^{-\frac{1}{2}} (X - \mathbb{E}[X])$$

ebenfalls $\mathbb{E}[\tilde{Y}] = 0$ und $\operatorname{Cov}(\tilde{Y}) = E_N$. Hierbei steht $\Sigma^{-\frac{1}{2}}$ für die Quadratwurzel der Matrix Σ^{-1} .

Die affine Transformation aus Lemma 89 zur Standardisierung von X ist dadurch ausgezeichnet, dass die Transformationsmatrix $\operatorname{Chol}(\Sigma^{-1})$ eine obere Dreiecksmatrix ist. Hieraus resultiert eine Reduktion des Aufwands zur Berechnung der Kumulanten des standardisierten Zufallsvektors. Dieser Effekt wird im Paragraph über die Operation 5 genauer beleuchtet. Ferner sind so die Voraussetzungen von Lemma 86 erfüllt und dessen Anwendung spielt eine zentrale Rolle bei der approximativen Berechnung der bedingten Verteilungsfunktion.

In Abbildung 3 ist der Programmablaufplan zur Berechnung der Kumulanten von $\operatorname{Chol}(\Sigma^{-1})(X - \mathbb{E}[X])$ dargestellt. Hierbei wird berücksichtigt, dass

$$\kappa^{\operatorname{Chol}(\Sigma^{-1})X}_{\pmb{\alpha}} = \kappa^{\operatorname{Chol}(\Sigma^{-1})(X - \mathbb{E}[X])}_{\pmb{\alpha}}$$

für alle $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| \geq 2$ gilt. Somit genügt zur Berechnung der gewünschten Kumulanten $\kappa_{\boldsymbol{\alpha}}^{\operatorname{Chol}(\Sigma^{-1})(X-\mathbb{E}[X])}$ von $\operatorname{Chol}(\Sigma^{-1})(X-\mathbb{E}[X])$ der Ordnungen $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| \in \{3, \ldots, S_{max} + 2\}$ die Berechnung der entsprechenden Kumulanten $\kappa_{\boldsymbol{\alpha}}^{\operatorname{Chol}(\Sigma^{-1})X}$ von $\operatorname{Chol}(\Sigma^{-1})X$.



Abbildung 3: Programmablaufplan zur Berechnung der Kumulanten des Zufallsvektors $\operatorname{Chol}(\Sigma^{-1})(X - \mathbb{E}[X])$

Vorbedingung: Die numerischen Werte der Matrix Σ und der Kumulanten κ_{α}^{X} für $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ seien bereits nach dem in Abbildung 1 dargestellten Programmablaufplan berechnet.

Operation 5: Teilaussage 2 von Satz 32 ist die mathematische Grundlage zur Berechnung der Kumulanten $\kappa_{\alpha}^{\operatorname{Chol}(\Sigma^{-1})X}$ von $\operatorname{Chol}(\Sigma^{-1})X$ der Ordnungen $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S_{max} + 3\}$. Für $K \in \mathbb{N}, \ \alpha \in \mathbb{N}_0^N$ mit $|\alpha| = K$ und eine voll besetzte Matrix $A \in$

Für $K \in \mathbb{N}$, $\boldsymbol{\alpha} \in \mathbb{N}_0^N$ mit $|\boldsymbol{\alpha}| = K$ und eine voll besetzte Matrix $A \in \mathbb{R}^{N \times N}$ hat die Summe auf der rechten Seite von Gleichung (6) insgesamt N^K Summanden. Diese Anzahl ist für nicht allzu kleine N und K sehr groß. Für die spezielle Wahl $A = (a_{n_1,n_2})_{n_1,n_2 \in \{1,\ldots,N\}} := \text{Chol}(\Sigma^{-1})$ ist A eine obere Dreiecksmatrix. Das hat zur Folge, dass einige der Summanden verschwinden. Wie viele Summanden das sind, hängt dabei von $\boldsymbol{\alpha}$ ab. Unter der Annahme, dass alle Einträge auf und oberhalb der Hauptdiagonalen von A ungleich 0 sind und mit der Definition

$$Anzahl(\boldsymbol{\alpha}) := \left| \left\{ (n_1, \dots, n_K) \in \{1, \dots, N\}^K \middle| \prod_{k=1}^K a_{I_{\boldsymbol{\alpha}}^k, n_k} \neq 0 \right\} \right|,$$

wobei hier $I_{\alpha} = (I_{\alpha}^1, \ldots, I_{\alpha}^K)$ das in Definition 31 eingeführte K-Tupel von Indices ist, erhält man zum Beispiel für $N = |\alpha| = 3$ die in Tabelle 5 aufgeführten Anzahlen an nicht verschwindenden Summanden.

α	Anzahl($\boldsymbol{\alpha}$)
(3, 0, 0)	27
(2,1,0)	18
(2,0,1)	9
(1,2,0)	12
(1,1,1)	6
(1, 0, 2)	3
(0,3,0)	8
(0, 2, 1)	4
(0,1,2)	2
(0, 0, 3)	1

Tabelle 5: Anzahl der nicht verschwindenden Summanden bei der Berechnung der Kumulanten von AX für $N = |\alpha| = 3$ und eine obere Dreiecksmatrix A in Abhängigkeit von α

Weil für jedes $K \in \{3, \ldots, S_{max} + 2\}$ die Kumulanten κ_{α}^{AX} für alle Multindices $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| = K$ zu berechnen sind, gibt ein Vergleich von N^K

und der durchschnittlichen Anzahl an Summanden

$$\frac{1}{\left|\left\{\tilde{\boldsymbol{\alpha}}\in\mathbb{N}_{0}^{N}\middle||\tilde{\boldsymbol{\alpha}}|=K\right\}\right|}\sum_{\boldsymbol{\alpha}\in\{\tilde{\boldsymbol{\alpha}}\in\mathbb{N}_{0}^{N}\mid|\tilde{\boldsymbol{\alpha}}|=K\}}Anzahl(\boldsymbol{\alpha})$$

ein gutes Gefühl dafür, wie sich der Rechenaufwand beim Übergang von einer voll besetzten Matrix zu einer oberen Dreiecksmatrix verringert. In Tabelle 6 sind für verschiedene N und K die Anzahl der Summanden für eine vollbesetzte Matrix und die durchschnittliche Anzahl an Summanden für eine obere Dreiecksmatrix aufgetragen.

	$ \alpha = 3$		$ \boldsymbol{\alpha} = 4$		$ \boldsymbol{\alpha} = 5$		$ \boldsymbol{\alpha} = 6$	
N	v	0	v	0	v	0	V	0
1	1	1	1	1	1	1	1	1
2	8	3, 8	16	6, 2	32	10, 5	64	18, 1
3	27	9	81	20, 1	243	46	729	108, 0
4	64	17, 5	256	48, 6	1.024	138, 8	4.096	406, 0
5	125	30	625	99, 3	3.125	337, 5	15.625	1.174, 9

Tabelle 6: Anzahl der Summanden bei der Berechnung der Kumulanten von AX für eine voll besetzte Matrix A (Teilspalten v) und durchschnittliche Anzahl der nicht verschwindenden Summanden für eine obere Dreiecksmatrix A (Teilspalten o)

Die vorliegende Operation 5 wird mit den Programmen ausgeführt, die vom Programm

aus Abschnitt 5.2.4 geschrieben wurden.

Ausgabe: Ausgegeben werden die numerischen Werte der Matrix A = Chol (Σ^{-1}) und der Kumulanten κ_{α}^{AX} für $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max}+2\}$.

5.3.4 Berechnung von *x*_{standardisiert}

Der numerische Wert von

$$x_{standardisiert} := A(X - \mathbb{E}[X])$$

mit $A = \text{Chol}(\Sigma^{-1})$ wird nach dem in Abbildung 4 dargestellten Programmablaufplan berechnet.



Abbildung 4: Programmablaufplan zur Berechnung von $x_{standardisiert}$

Vorbedingung: Die numerischen Werte des Vektors $\mathbb{E}[X]$ und der Matrix $A = \text{Chol}(\Sigma^{-1})$ seien bereits nach den Programmablaufplänen in den Abbildungen 1 und 3 berechnet. Ferner sei der numerische Wert von $x \in \mathbb{R}^N$ gegeben.

Die Berechnung ist trivial und bereitet keine Probleme.

Ausgabe: Ausgegeben wird der numerische Wert von $x_{standardisiert}$.

5.3.5 Approximative Berechnung der Dichte von X

Die Berechnung der Approximation von $f_X(x)$ erfolgt nach dem in Abbildung 5 dargestellten Programmablaufplan.

Vorbedingung: Die numerischen Werte der Matrix $A = \text{Chol}(\Sigma^{-1})$ und der Kumulanten κ_{α}^{AX} für $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ seien bereits nach dem Programmablaufplan in Abbildung 3 berechnet. Ferner sei der numerische Wert von $x_{standardisiert}$ nach dem in Abbildung 4 dargestellten Programmablaufplan ebenfalls bereits berechnet.

Operationen 6 und 7: Der S-te Edgeworth-Korrekturterm $\mathcal{K}^{S}_{A(X-\mathbb{E}[X])}$ und die Edgeworth-Approximation $f^{S_{max}}_{A(X-\mathbb{E}[X])}$ der Dichte eines zentrierten Zufallsvektors werden in den Teilen 1 und 2 von Definition 47 eingeführt. Die numerischen Werte der $\mathcal{K}^{S}_{A(X-\mathbb{E}[X])}$ für $S \in \{1, \ldots, S_{max}\}$ werden mit den



Abbildung 5: Programmablaufplan zur approximativen Berechnung der Dichte von \boldsymbol{X}

Programmen berechnet, die vom Programm

Erstelle_m_file_fuer_EdgeworthKorrekturtermStandardisiert

aus Abschnitt 5.2.5 geschrieben wurden.

Hierbei wird ausgenutzt, dass $\kappa_{\alpha}^{AX} = \kappa_{\alpha}^{A(X-\mathbb{E}[X])}$ für alle dabei auftreten-den Multiindices α , also $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$, gilt.

Operation 8: Als approximativer Wert für $f_X(x)$ wird die Größe

 $|\det(A)| f_{A(X-\mathbb{E}[X])}^{S_{max}}(x_{standardisiert})$

verwendet. Diese Wahl ist durch Lemma 83 motiviert.

Ausgabe: Ausgegeben wird der numerische Wert der Approximation von $f_X(x)$.

5.3.6Approximative Berechnung der bedingten Verteilungsfunktion

Abbildung 6 zeigt den Programmablaufplan für die approximative Berechnung der bedingten Verteilungsfunktion von X^1 gegeben X^2, \ldots, X^N .

Vorbedingung: Die numerischen Werte der Kumulanten κ_{α}^{AX} für $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ seien bereits nach dem in Abbildung 3 dargestellten Programmablaufplan berechnet. Ferner sei der numerische Wert für $x_{standardisiert}$ nach dem Programmablaufplan in Abbildung 4 ebenfalls bereits berechnet.

Operation 9: \mathcal{Z}_X^S und \mathcal{N}_X^S sind die Größen aus Lemma 87. Die Operation wird mit den Programmen ausgeführt, die vom Programm

Erstelle_m_file_fuer_EdgeworthBedingteCdfStandardisiert

aus Abschnitt 5.2.6 geschrieben wurden.

Auch hier wird ausgenutzt, dass $\kappa_{\alpha}^{AX} = \kappa_{\alpha}^{A(X-\mathbb{E}[X])}$ für alle dabei auftre-tenden Multiindices α , also $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$, gilt.



Abbildung 6: Programmablaufplan zur approximativen Berechnung der bedingten Verteilungsfunktion von X^1 gegeben X^2, \ldots, X^N

Operation 10: Als approximativer Wert für $F_{X^1|X^2,...,X^N}(x^1|x_2,...,x^N)$ wird die Größe $F_{Y^1|Y^2,...,Y^N}^{S_{max}}(y^1|y^2,...,y^N)$ mit $Y = (Y^1,...,Y^N)^T := A(X - \mathbb{E}[X])$ und $(y^1,...,y^N) := x_{standardisiert}$ verwendet. Diese Wahl ist motiviert durch Lemma 86, nach dem

$$F_{X^1|X^2,\dots,X^N}(x^1|x_2,\dots,x^N) = F_{Y^1|Y^2,\dots,Y^N}(y^1|y^2,\dots,y^N)$$

gilt. Dabei sind die Anforderungen an A erfüllt, weil A eine obere Dreiecksmatrix ist.

Ausgabe: Ausgegeben wird der numerische Wert der Approximation von $F_{X^1|X^2,...,X^N}(x^1|x_2,...,x^N)$.

5.4 Approximative Berechnung der Log-Likelihood und der Teststatistiken der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests

Betrachtet wird der \mathbb{R}^N -wertige Zufallsvektor X aus Abschnitt 5.2. Die Integrationsgrenzen werden zu einem Vektor $\xi_{Integrationsgrenzen} := (\Delta_1, \Delta_2) \in$ $\Xi_{Integrationsgrenzen} := \{(\tilde{\Delta}_1, \tilde{\Delta}_2) \in [0, T^*] | \tilde{\Delta}_1 < \tilde{\Delta}_2\}$ zusammengefasst. Weiterhin wird angenommen, dass die Verteilung $\mathcal{L}(L_1)$ von einem Parametervektor $\xi_{Integrator} \in \Xi_{Integrator}$ und die Integranden f_1, \ldots, f_N von einem Parametervektor $\xi_{Integranden} \in \Xi_{Integranden}$ abhängen. Hierbei seien $\Xi_{Integrator}$ und $\Xi_{Integranden}$ geeignete Parametermengen. Definiere nun

$$\xi := (\xi_{Integrationsgrenzen}, \xi_{Integrator}, \xi_{Integranden})$$

und

$$\Xi := \Xi_{Integrationsgrenzen} \times \Xi_{Integrator} \times \Xi_{Integranden}$$

Ferner seien $K \in \mathbb{N}$ und x_1, \ldots, x_K eine gegebene Stichprobe von stochastisch unabhängigen und identisch verteilten Realisierungen eines \mathbb{R}^N -wertigen Zufallsvektors.

Das Design der nun vorgestellten Algorithmen ist so gewählt, dass der Rechenaufwand möglichst gering ist. Dies wird unter anderem dadurch erreicht, dass Teilrechnungen, die mehrfach und mit denselben numerischen Werten für die Eingabeargumente auftreten, nicht mehrfach ausgeführt werden. Stattdessen werden diese Teilrechnungen nur einmal durchgeführt und das Ergebnis wird dann an allen relevanten Stellen verwendet. Für Teilrechnungen, die ansonsten für jede Beobachtung aus der Stichprobe ausgeführt worden wären, reduziert sich so der Rechenaufwand auf 1/K des ursprünglichen Volumens. Bei einer Stichprobengröße von K = 253 führt das beispielsweise zu einem Speedup von über 200, wobei hier der entstehende Overhead bereits berücksichtigt ist.

5.4.1 Approximative Berechnung der Log-Likelihood

Die Log-Likelihood zur Stichprobe x_1, \ldots, x_K ist die Funktion $\ell : \Xi \to \mathbb{R}$ mit

$$\ell(\xi) := \sum_{k=1}^{K} f(x_k, \xi),$$
(29)

wobei hier $f(\cdot,\xi): \mathbb{R}^N \to \mathbb{R}$ die Wahrscheinlichkeitsdichte der Verteilung von X ist, die für den Parametervektor $\xi \in \Xi$ vorliegt.

Für ein gegebenes $\xi \in \Xi$ wird nun zur approximativen Berechnung der Log-Likelihood $\ell(\xi)$ die exakte Wahrscheinlichkeitsdichte $f(\cdot, \xi)$ durch die dazugehörige Approximation aus Abschnitt 5.3.5 ersetzt.

Abbildung 7 zeigt den Programmablaufplan, der den dazugehörigen Funktionalitäten der Edgeworth-Toolbox zugrunde liegt. Hierbei wird ausgenutzt, dass im betrachteten Szenario davon ausgegangen wird, dass x_1, \ldots, x_K eine Stichprobe von identisch verteilten Realisierungen ist. Der Programmablaufplan in Abbildung 1 zur Berechnung von $\mathbb{E}[X]$, Σ und κ_{α}^X für $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ sowie der Programmablaufplan in Abbildung 3 zur Berechnung von $A = \text{Chol}(\Sigma^{-1})$ und κ_{α}^{AX} für $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ werden nicht für jedes $k \in \{1, \ldots, K\}$ eigens durchlaufen. Stattdessen werden sie jeweils nur einmal durchlaufen und die gefundenen numerischen Werte für die Ausgaben werden als Grundlage für die weiteren Berechnung für alle $k \in \{1, \ldots, K\}$ verwendet. Lediglich der Programmablaufplan in Abbildung 4 zur Berechnung von $x_{standardisiert}$ und der Programmablaufplan in Abbildung 5 zur Berechnung der Approximation der Dichte von X werden für jedes $k \in \{1, \ldots, K\}$ eigens durchlaufen.

5.4.2 Approximative Berechnung der Rosenblatt-Transformation

Zusätzlich zu den bereits definierten Größen sei ferner eine Permutation σ von $\{1, \ldots, N\}$ fest vorgegeben und R_{σ} bezeichne die dazugehörige Rosenblatt-Transformation aus Definition 56.

Die Darstellung zur approximativen Berechnung der Rosenblatt-Transformation wird in zwei Schritte unterteilt. Zunächst wird für ein gegebenes $n \in \{1, \ldots, N\}$ das Vorgehen zur approximativen Berechnung der $\sigma(n)$ -ten Komponente vorgestellt. Abbildung 8 zeigt den dazugehörigen Programmablaufplan. Hierauf wird dann im nächsten Schritt bei der approximativen Berechnung der gesamten Rosenblatt-Transformation zurückgegriffen. Der dazugehörige Programmablaufplan ist in Abbildung 9 dargestellt.

Zur Abkürzung werden für $n \in \{1, \ldots, N\}$ die Bezeichnungen $X_n := (X^{\sigma(n)}, X^{\sigma(1)}, \ldots, X^{\sigma(n-1)})^T$, $\Sigma_n := \operatorname{Cov}(X_n)$, $A_n := \operatorname{Chol}(\Sigma_n^{-1})$, $x_{n,k} := (x_k^{\sigma(n)}, x_k^{\sigma(1)}, \ldots, x_k^{\sigma(n-1)})^T$ und $x_{n,k,standardisiert} := A_n(x_{n,k} - \mathbb{E}[X_n])$ verwendet.



Abbildung 7: Programmablaufplan zur approximativen Berechnung der Log-Likelihood



Abbildung 8: Programmablaufplan zur approximativen Berechnung der $\sigma(n)$ -ten Komponente $R_{\sigma}^{\sigma(n)}$ der Rosenblatt-Transformation der Beobachtungen x_1, \ldots, x_K

Vorbedingung zur approximativen Berechnung der $\sigma(n)$ -ten Komponente der Rosenblatt-Transformation: Es sei $n \in \{1, \ldots, N\}$ vorgegeben. Ferner wird angenommen, dass für den zugrunde gelegten Parametervektor ξ die numerischen Werte des Erwartungswertvektors $\mathbb{E}[X]$, der Kovarianzmatrix Σ und der Kumulanten κ_{α}^{X} für alle $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ bereits berechnet sind.

Operation 11: Die Berechnung der Kumulanten $\kappa_{\alpha}^{X_n}$ für alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ erfolgt auf der Grundlage von Lemma 25.

Bei der approximativen Berechnung der $\sigma(n)$ -ten Komponente der Rosenblatt-Transformation, so wie sie im Programmablaufplan in Abbildung 8 dargestellt ist, wird die Annahme ausgenutzt, dass x_1, \ldots, x_K eine Stichprobe von identisch verteilten Realisierungen ist. Deshalb wird der Programmablaufplan in Abbildung 3, der zur Berechnung der numerischen Werte von $A_n := \operatorname{Chol}(\Sigma_n^{-1})$ und von $\kappa_{\alpha}^{A_n X_n}$ für $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ verwendet wird, nur einmal durchlaufen. Die gefundenen numerischen Werte te werden dann als Grundlage für die weiteren Berechnungen für alle $k \in \{1, \ldots, K\}$ verwendet. Lediglich der Programmablaufplan in Abbildung 4, mit dem die numerischen Werte der $x_{n,k,standardisiert}$ berechnet werden, und der Programmablaufplan in Abbildung 6, mit dem die Approximationen der $F_{X^{\sigma(n)}|X^{\sigma(1)},\ldots,X^{\sigma(n-1)}}(x_k^{\sigma(n)}|x_k^{\sigma(1)},\ldots,x_k^{\sigma(n-1)})$ berechnet werden, werden für jedes $k \in \{1,\ldots,K\}$ eigens durchlaufen. Die so berechneten approximativen Werte für die $F_{X^{\sigma(n)}|X^{\sigma(1)},\ldots,X^{\sigma(n-1)}}(x_k^{\sigma(n)}|x_k^{\sigma(1)},\ldots,x_k^{\sigma(n-1)})$ werden als Approximationen für die $R_{\sigma}^{\sigma(n)}(x_k)$ verwendet.

Bei der approximativen Berechnung der Rosenblatt-Transformation gemäß dem Programmablaufplan in Abbildung 9 wird ebenfalls die Annahme ausgenutzt, dass x_1, \ldots, x_K eine Stichprobe von identisch verteilten Realisierungen ist. Deshalb werden die numerischen Werte von $\mathbb{E}[X]$, von Σ und von κ_{α}^X für $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| \in \{3, \ldots, S_{max} + 2\}$ nur einmal berechnet. Diese Werte werden dann als Grundlage für die Berechnung der Approximation der $\sigma(n)$ -ten Komponente der Rosenblatt-Transformation gemäß dem Programmablaufplan in Abbildung 8 für alle $n \in \{1, \ldots, N\}$ verwendet.



Abbildung 9: Programmablaufplan zur approximativen Berechnung der Rosenblatt-Transformation R_{σ} der Beobachtungen x_1, \ldots, x_K

5.4.3 Approximative Berechnung der Teststatistiken der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Misesund des Anderson-Darling-Tests

Wie bereits in Abschnitt 5.4.2 sei zusätzlich zu den bereits definierten Größen eine Permutation σ von $\{1, \ldots, N\}$ fest vorgegeben und R_{σ} bezeichne die dazugehörige Rosenblatt-Transformation. Die multivariate Rosenblatt-Erweiterung wurde in Definition 59 eingeführt. Für $\mu = 0$ erhält man die multivariate Rosenblatt-Erweiterung des Cramér-von-Mises-Tests und für $\mu = 1$ erhält man die multivariate Rosenblatt-Erweiterung des Anderson-Darling-Tests. Nach Satz 60 lässt sich die Auswertung der dazugehörigen Teststatistiken $A_{R_{\sigma},K,0}^2$ und $A_{R_{\sigma},K,1}^2$ auf den in Abschnitt 4.2 behandelten Fall für Copulas transformieren und somit können ihre Werte unter Verwendung der in Abschnitt 4.3 hergeleiteten geschlossenen Formen berechnet werden. Hierfür ist jedoch die Berechnung der Rosenblatt-Transformationen $u_1 :=$ $R_{\sigma}(x_1), \ldots, u_K := R_{\sigma}(x_K)$ der Realisierungen x_1, \ldots, x_K erforderlich.

Bei der approximativen Berechnung der Teststatistiken $A_{R_{\sigma},K,0}^2$ und $A_{R_{\sigma},K,1}^2$ werden die darin auftretenden exakten Werte für die Rosenblatt-Transformation $R_{\sigma}(\cdot)$ durch die dazugehörige Approximation aus dem Abschnitt 5.4.2 ersetzt.

Die Teststatistik $A^2_{R_{\sigma},K,0}$ wird unter Verwendung der Formel aus Satz 76 ausgewertet. Zur Berechnung des Wertes der Teststatistik $A^2_{R_{\sigma},K,1}$ wird die Formel aus Satz 82 verwendet.

Für die numerische Auswertung des darin auftretenden Polylogarithmus $[0,1] \ni x \mapsto \operatorname{Li}_N(x)$ ist in MATLAB zwar bereits das Programm polylog vorprogrammiert. Dieses Programm hat jedoch eine sehr lange Laufzeit. Weil in der Formel aus Satz 82 der Polylogarithmus an sehr vielen Stellen auszuwerten ist, ist polylog für diese Arbeit ungeeignet. Deshalb wurde ein eigenes Programm geschrieben, das in deutlich kürzerer Zeit die numerischen Werte des Polylogarithmus mit der erforderlichen Genauigkeit berechnet. Für $N \in \{0,1\}$ verwendet es die bekannten geschlossenen Formen für den Polylogarithmus. Für N > 2 sind keine geschlossenen Formen bekannt und es wird in Abhängigkeit vom numerischen Wert von x eine Berechnungsart gewählt. Für x = 1 wird der entsprechende Wert der ζ -Funktion verwendet, der für die relevanten Werte von N mit hinreichend großer Genauigkeit hardgecoded im Quellcode hinterlegt ist. Für $x \in (\frac{1}{4}, 1)$ erfolgt die Berechnung unter Verwendung der Potenzreihe aus Wood (1992, Formel (9.5)) und für $x \in [0, \frac{1}{4}]$ erfolgt die Auswertung unter Verwendung der Potenzreihe aus der Definitionsgleichung (25). Hierbei werden die Potenzreihen jeweils nach endlich vielen Summanden abgebrochen. Dabei hängt die Anzahl der berücksichtigten Summanden vom numerischen Wert von x ab und ist jeweils so gewählt, dass der numerische Wert der Ausgabe der Funktion um weniger als 10^{-11} vom wahren Wert des Polylogarithmus abweicht.

5.5 Der Einsatz von Grafikprozessoren

Im Fokus der Programmierung steht die Kalibrierung des Modells aus Abschnitt 2. Hierbei wird derjenige Satz von Modellparametern gesucht, für den eine gewisse Zielfunktion einen optimalen Wert annimmt. Das Auffinden des Optimums ist ein nichtlineares Optimierungsproblem, das hier durch iterative Verfahren gelöst wird. Dabei wird die Zielfunktion an vielen verschiedenen Stellen ausgewertet. Die Rechenzeit für den gesamten Optimierungsprozess hängt somit kritisch von der Rechenzeit für jeden Funktionsaufruf der Zielfunktion ab. Weil die Anzahl der benötigten Funktionsaufrufe sehr groß ist, kommt es bei der Auswertung der Zielfunktion durchaus auf jede hundertstel Sekunde an.

Bei den Kalibrierungen, die in Abschnitt 6 vorgestellt werden, werden drei verschiedene Zielfunktionen betrachtet. Die erste ist die Approximation der Log-Likelihood aus Abschnitt 5.4.1. Die zweite und die dritte hängen von den Approximationen der Teststatistiken der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests aus Abschnitt 5.4.3 ab.

Eine notwendige Voraussetzung, um die Kalibrierungen in vertretbarer Zeit ausführen zu können, ist also eine Kombination aus einer angemessenen Hardware und Programmen, mit der sich die besagten Größen in sehr kurzer Zeit numerisch berechnen lassen.

Wie es bereits in der Einleitung zum Abschnitt 5 erwähnt wurde, nutzt die Edgeworth-Toolbox neben dem Prozessor (CPU) ebenfalls auch zwei zusätzliche Grafikprozessoren (GPUs), um hierdurch weitere Rechenkapazitäten zu erschließen.

5.5.1 Charakteristiken von Prozessoren und Grafikprozessoren

Quellcode, der für die Ausführung auf einem Prozessor geschrieben wurde, ist im Allgemeinen nicht auf einem Grafikprozessor ausführbar. Sollen also Programmteile durch Grafikprozessoren abgearbeitet werden, so muss hierfür eigener Quellcode geschrieben werden, der für die Ausführung auf einer GPU geeignet ist. In einigen Fällen genügt es sogar nicht einmal, den für eine CPU geschriebenen Quellcode in eine Programmiersprache für Grafikprozessoren zu übersetzen. Der Grund hierfür ist, dass sich die Hardwarearchitektur einer GPU und die Hardwarearchitektur einer CPU grundlegend unterscheiden.

Vereinfacht ausgedrückt ist die Hardwarearchitektur einer CPU darauf optimiert, einen bis wenige Threads parallel zu bearbeiten. Dafür wird jeder dieser Threads schnell abgearbeitet.

Im Gegensatz dazu ist die Hardwarearchitektur einer GPU nicht darauf ausgelegt, einzelne Threads möglichst schnell abzuarbeiten. Stattdessen ist sie darauf optimiert, an sehr vielen Threads parallel zu arbeiten. Dabei sollte es möglichst wenig Interaktion zwischen den einzelnen Threads geben. Ferner sind GPUs insbesondere auf Gleitkommaarithmetik spezialisiert.

Jeder einzelne Thread hat also auf einer GPU eine deutlich höhere Laufzeit als auf einer CPU. Weil aber auf einer GPU so viele Threads parallel bearbeitet werden können, ist die gesamte Leistungsfähigkeit einer GPU deutlich höher als die einer CPU. Als Beispiel werden hier die Kenndaten der für die Kalibrierung verwendeten Hardware angegeben. Die Leistung für numerische Berechnungen wird in der Einheit Floating Point Operations Per Second (FLOPS) gemessen. Der verwendete XEON E5-1650 Prozessor hat für den Datentyp double eine theoretische Spitzenleistung von 153, 6 GFLOPS. Die beiden verwendeten GPUs vom Modell NVIDIA Geforce GTX Titan haben laut Herstellerangaben für diesen Datentyp jeweils eine Leistung von 1311, 7 GFLOPS. Bei seiner Markteinführung im Jahr 2013 war dieses Modell die stärkste Desktop-Grafikkarte am Markt.

Für eine CPU stellt es keinen nennenswerten Nachteil dar, wenn die zu bearbeitenden Threads voneinander abhängen oder aufeinander aufbauen und somit nacheinander, also seriell, ausgeführt werden müssen. Anders verhält es sich bei einer GPU. Sie kann ihre gewaltige Leistung nur dann vollständig entfalten, wenn sie hinreichend viele Threads zur parallelen Bearbeitung zugeteilt bekommt. Ansonsten bleiben Teile ihrer Rechenkapazität ungenutzt. Dieser Punkt spielt beim Design von Algorithmen, die auf einer GPU abgearbeitet werden sollen, eine zentrale Rolle.

Ein weiterer Unterschied zwischen den Hardwarearchitekturen einer CPU und einer GPU liegt in der physischen Anordnung des Arbeitsspeichers. Bei ihren Berechnungen nutzt die CPU den Hauptspeicher des Computers. Dieser ist physisch von der CPU getrennt. Bei jedem Zugriff müssen die erforderlichen Daten über den Front Side Bus, die Northbridge der Hauptplatine und den Memory Bus zwischen der CPU und dem Hauptspeicher hin und her transferiert werden. Dies hat den Vorteil, dass jeder Thread auf den gesamten Hauptspeicher zugreifen kann. Der Nachteil ist jedoch die Geschwindigkeit. Bei der verwendeten Quad-Channel-fähigen Hauptplatine und den vier DDR3-1600-SDRAM Modulen wird so eine maximale Datenübertragungsrate zwischen CPU und Hauptspeicher von 51,8 GB/s erreicht, was für aufwändige numerische Berechnungen unzureichend ist.

Im Gegensatz dazu ist bei einer GPU der Grafikspeicher physisch direkt bei den Recheneinheiten verbaut und es gibt eine ausgeklügelte Hierarchie der unterschiedlichen Bereiche des Grafikspeichers. Laut Herstellerangaben wird so bei jeder der verwendeten NVIDIA Geforce GTX Titan GPUs eine Datenübertragungsrate zwischen den Recheneinheiten und dem Grafikspeicher von 288,4 GB/s erreicht. Es ist jedoch zu beachten, dass bei der Auslagerung einer Teilrechnung auf eine GPU die dafür erforderlichen Daten zunächst vom Hauptspeicher in den Grafikspeicher transferiert werden müssen. Ferner müssen die Ergebnisse nach Abschluss der Teilrechnung wieder in den Hauptspeicher des Computers zurück transferiert werden. Dies erfolgt über den Memory Bus, die Northbridge der Hauptplatine und den PCI Express, wobei hier die verwendete Hardware eine Datenübertragungsrate von etwa 15,8 GB/s zu jeder der beiden GPUs erreicht. Beim Design von Programmen, die Teilrechnungen auf eine GPU auslagern, sollte also darauf geachtet werden, dass zwischen Haupt- und Grafikspeicher nur möglichst geringe Datenmengen ausgetauscht werden.

Monte-Carlo-Simulationen, numerische Integrationen, numerische Faltungen und das numerische Lösen von partiellen Differentialgleichungen sind Beispiele von Aufgabentypen, für die eine GPU deutlich besser geeignet ist als eine CPU. Dagegen sind symbolische Berechnungen für die Ausführung auf einer GPU völlig ungeeignet. Für diese bietet eine CPU eine deutlich bessere Performance.

Wie groß der Speedup ist, den man durch den Einsatz einer GPU tatsächlich erhält, hängt von der konkreten Aufgabenstellung sowie von den verwendeten Algorithmen ab und lässt sich im Allgemeinen nur schwer abschätzen.

5.5.2 Einsatz von Grafikprozessoren in der Edgeworth-Toolbox

Bei der Entwicklung der Edgeworth-Toolbox wurden diejenigen Teilprogramme in CUDA übersetzt bzw. mit eigens dafür entwickelten Algorithmen in CUDA implementiert, bei denen der Einsatz von GPUs einen deutlichen Speedup erwarten lässt und die ferner auch zeitkritisch für die Auswertung der Zielfunktionen für die Kalibrierungen sind. Auf diese Teilprogramme wird nun eingegangen.

Bei der approximativen Berechnung der Log-Likelihood wird der Programmablaufplan in Abbildung 7 durchlaufen. Hierbei werden das Unterprogramm, das durch den Programmablaufplan in Abbildung 4 beschrieben wird, und das Unterprogramm, das durch den Programmablaufplan in Abbildung 5 beschrieben wird, für jede einzelne Beobachtung aus der Stichprobe, also K-mal, ausgeführt.

Bei der approximativen Berechnung der Teststatistiken der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests aus Abschnitt 5.4.3 werden als ein Teilschritt die Rosenblatt-Transformationen aller Beobachtungen aus der Stichprobe berechnet. Hierzu wird der Programmablaufplan in Abbildung 9 durchlaufen, in dem für jede Komponente der Rosenblatt-Transformation als Unterprogramm der Programmablaufplan in Abbildung 8 aufgerufen wird. Dabei wiederum werden das Unterprogramm, das durch den Programmablaufplan in Abbildung 4 beschrieben wird, und das Unterprogramm, das durch den Programmablaufplan in Abbildung 6 beschrieben wird, für jede einzelne Beobachtung aus der Stichprobe, also K-mal, ausgeführt.

In jedem dieser vier Fälle sind die K Durchläufe des jeweiligen Unterprogrammes unabhängig voneinander und somit liegt ein sogenannter embarrassingly parallel workload vor. Ferner wird bei jedem einzelnen Durchlauf des jeweiligen Unterprogramms derselbe Quellcode ausgeführt. Die verschiedenen Durchläufe unterscheiden sich lediglich in den numerischen Werten der Eingabeargumente.

Für diese Art von Problemstellung ist der Einsatz von Grafikprozessoren hervorragend geeignet. Im vorliegenden Fall war es sogar hinreichend, den MATLAB-Code, der von den in Abschnitt 5.2 beschriebenen Programmen geschrieben wurde, in CUDA zu übersetzen und über eine eigens dafür programmierte Schnittstelle in das Gesamtprogramm einzubinden. Die Übersetzung erfolgte größtenteils von Hand unter intensiver Ausnutzung der **replace**-Funktionalität der verwendeten integrierten Entwicklerumgebung für CUDA. Als Ergebnis wurden für die betrachteten Programmabschnitte Speedups im Bereich von 20 bis 150 realisiert.

Bei der approximativen Berechnung der multivariaten Teststatistiken aus Abschnitt 5.4.3 sind weiterhin die Formeln aus den Sätzen 76 und 82 auszuwerten. Die darin vorkommenden Summen haben sehr viele Summanden. Allein bei den Summen über die Indices $(k_1, \ldots, k_N) \in \{1, \ldots, K\}^N$ sind das K^N Summanden, was beispielsweise für N = 4 und K = 204 zu $1, 7 \cdot 10^9$ Summanden führt. Bei Verwendung des Datentyps double haben allein die numerischen Werte dieser Summanden bereits ein Datenvolumen von 13,9 GB.

Dieses enorme Volumen an Gleitkommaarithmetik legt den Einsatz von GPUs nahe. Die Schwierigkeit im vorliegenden Fall liegt jedoch bei der Berechnung der relevanten Werte $F_K(k_1, \ldots, k_N)$ der empirischen Verteilungsfunktion, die in nahezu jedem Summanden auftritt.

Prinzipiell ist es denkbar, jeden Summanden in einem eigenen Thread auszuwerten, der unabhängig von allen anderen Threads ist, und somit auch die numerischen Werte der $F_K(k_1, \ldots, k_N)$ unabhängig voneinander zu berechnen. Im mehrdimensionalen Fall ist die unabhängige Auswertung jedes einzelnen $F_K(k_1, \ldots, k_N)$ jedoch mit einem unerwartet hohen Rechenaufwand verbunden, sodass dieser Weg die Auswertung der Teststatistiken in vertretbarer Zeit ausschließt.

Verwendet man zur Auswertung der $F_K(k_1, \ldots, k_N)$ hingegen einen Algorithmus, der die Tatsache ausnutzt, dass zwischen den $F_K(k_1, \ldots, k_N)$ Abhängigkeiten bestehen, so lässt sich das Pensum an durchzuführenden Rechenoperationen auf einen Bruchteil reduzieren. Dies führt jedoch dazu, dass sich so die einzelnen Summanden nicht mehr unabhängig voneinander auswerten lassen und man Interaktionen zwischen den Threads akzeptieren muss. Zum einen erhöht das die Komplexität bei der Programmierung und zum anderen geht bei den Interaktionen zwischen den Threads wertvolle Zeit verloren. Ferner führen derartige Algorithmen zu einem erhöhten Bedarf an Arbeitsspeicher. Bei einigen Standardalgorithmen, die die Abhängigkeiten zwischen den $F_K(k_1, \ldots, k_N)$ vollständig ausnutzen, ist es sogar erforderlich, alle $F_K(k_1,\ldots,k_N)$ für $(k_1,\ldots,k_N) \in \{1,\ldots,K\}^N$ gleichzeitig im Arbeitsspeicher vorzuhalten. Bei dem oben angeführte Zahlenbeispiel führt das allein für die numerischen Werte der $F_K(k_1,\ldots,k_N)$ zu einem Speicherbedarf von 13,9 GB, was die 6 GB, die bei einer NVIDIA Geforce GTX Titan GPU zur Verfügung stehen, übersteigt.

Für die Edgeworth-Toolbox wurden deshalb eigene Algorithmen zur Auswertung der Teststatistiken entwickelt, in CUDA implementiert und über eine Schnittstelle in das Gesamtprogramm eingebunden. Sie stellen einen Mittelweg zwischen einer möglichst hohen Ausnutzung der Abhängigkeiten der $F_K(k_1, \ldots, k_N)$ und einer möglichst geringen Interaktion zwischen den einzelnen Threads zur Berechnung der Summanden dar. Sie sind genau auf die verwendeten GPUs zugeschnitten und nutzen deren Rechenkapazität vollständig aus. Als Ergebnis wurden für die betrachteten Programmabschnitte Speedups im Bereich von 5 bis 100 realisiert.

Durch die vorgenommenen Auslagerungen von Teilrechnungen auf die GPUs konnten im Gesamtprogramm für die Auswertung der Zielfunktionen insgesamt Speedups in der Größenordnung von ca. 10 realisiert werden. Trotzdem betrug die reine Rechenzeit für die in Abschnitt 6 vorgestellten Kalibrierungen immer noch mehrere Wochen. Somit sind die Kalibrierungen erst durch den Einsatz der GPUs in vertretbarer Zeit möglich geworden.

6 Kalibrierung unter dem physikalischen Maß

Die Vorarbeit aus den vorangehenden Abschnitten erlaubt nun die angestrebte Kalibrierung des Hybridmodells aus Abschnitt 2. Im vorliegenden Abschnitt wird der hierfür beschrittene Weg beschrieben. Dabei wird zunächst in Abschnitt 6.1 auf die Modellparameter eingegangen. Abschnitt 6.2 ist den Marktgrößen gewidmet, deren historische Realisierungen als Datengrundlage für die Kalibrierung verwendet werden. In Abschnitt 6.3 werden die Kalibrierungsarten vorgestellt. Neben dem wohlbekannten Maximum-Likelihood-Kriterium werden mit dem Cramér-von-Mises- und dem Anderson-Darling-Kriterium zwei neue Kalibrierungsarten eingeführt. In Abschnitt 6.4 werden schließlich die Ergebnisse der Kalibrierung für deutsche Staatsanleihen und den DAX präsentiert.

6.1 Die Modellparameter

Bei der Kalibrierung werden für die dafür benötigten exogenen Größen die nun vorgestellten parametrischen Formen verwendet.

6.1.1 Der treibende Lévy-Prozess L

Als treibender Prozess wird ein *D*-dimensionaler verallgemeinerter hyperbolischer Lévy-Prozess (*GH*-Prozess) verwendet. Die Klasse der *D*-dimensionalen *GH*-Prozesse wird durch die Klasse der *D*-dimensionalen *GH*-Verteilungen erzeugt. Das ist eine Verteilungsklasse mit den Parametern $\lambda \in \mathbb{R}, \alpha, \delta \in \mathbb{R}_+ \setminus \{0\}, \beta, \mu \in \mathbb{R}^D$ und einer strikt positiv definiten Matrix $\Delta \in \mathbb{R}^{D \times D}$ mit det(Δ) = 1, sodass $\alpha^2 - \langle \beta, \Delta \beta \rangle > 0$ erfüllt ist. Für einen gegebenen Parametersatz ist die dazugehörige momenterzeugende Funktion $M_{GH(\lambda,\alpha,\beta,\delta,\mu,\Delta)}$ gegeben durch

$$M_{GH(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(u) = e^{\langle u,\mu \rangle} \left(\frac{\alpha^2 - \langle \beta, \Delta\beta \rangle}{\alpha^2 - \langle \beta + u, \Delta(\beta + u) \rangle} \right)^{\frac{\lambda}{2}} \times \frac{K_\lambda \left(\delta \sqrt{\alpha^2 - \langle \beta + u, \Delta(\beta + u) \rangle} \right)}{K_\lambda \left(\delta \sqrt{\alpha^2 - \langle \beta, \Delta\beta \rangle} \right)}$$

für alle $u \in \mathbb{R}^D$ mit $\alpha^2 - \langle \beta + u, \Delta(\beta + u) \rangle > 0$. Dabei ist K_{λ} die modifizierte Besselfunktion dritten Grades mit Index λ . Insbesondere genügt die Verteilung und damit auch der von ihr erzeugte *GH*-Prozess der Annahme ($\mathbb{E}M$) zu einer strikt positiven Konstante. Die Parameter der Verteilung von L_1 , die bekanntlich bereits die Verteilung von ganz L eindeutig festlegen, werden zu einem Vektor $\vartheta_{Treiber}$ zusammengefasst. Die Menge der in Betracht gezogenen Parametervektoren wird mit $\Theta_{Treiber}$ bezeichnet.

Für weitere Details zu GH-Verteilungen, ihren Spezialfällen und ihren Grenzwerten siehe Hammerstein (2010).

Zu Vergleichszwecken wird ferner auch ein *D*-dimensionaler Wiener-Prozess $W = (W_t)_{t \in [0,T^*]}$ als Treiber betrachtet.

6.1.2 Die Driftstruktur der Zinsstruktur α_B und die Driftstruktur des Aktienindex α_S

Für die Driftstrukturen wird hier keine parametrische Form festgelegt. Für das Folgende wird lediglich angenommen, dass α_B und α_S stationär sind und damit

$$\int_{t}^{t+\Delta} A_B(s,t+\Delta,t+T) \,\mathrm{d}s = \int_{0}^{\Delta} A_B(s,\Delta,T) \,\mathrm{d}s$$

und

$$\int_{t}^{t+\Delta} \tilde{\alpha}_{S}(s,t+\Delta) \,\mathrm{d}s = \int_{0}^{\Delta} \tilde{\alpha}_{S}(s,\Delta) \,\mathrm{d}s$$

für alle $t, T \in [0, T^*]$ und $\Delta \in \mathbb{R}_+$ mit $t \le t + \Delta \le t + T \le T^*$ gilt.

6.1.3 Die Volatilitätsstruktur der Zinsstruktur σ_B und die Volatilitätsstruktur des Aktienindex σ_S

Als Volatilitätsstruktur der Zinsstruktur $\sigma_B : [0, T^*] \to \mathbb{R}^D$ wird die *D*dimensionale *abcd*-Formel gewählt. Für jedes $i \in \{1, \ldots, D\}$ wird also

$$\sigma_B^i(s,T) := (a^i(T-s) + d^i)e^{-b^i(T-s)} + c^i$$

gesetzt, wobei $a^i, b^i, c^i, d^i \in \mathbb{R}$ mit $a^i, b^i, c^i \geq 0$ und $d^i + c^i \geq 0$. Mit dieser Wahl ist σ_B stationär, es gilt also $\sigma_B(s, T) = \sigma_B(0, T-s)$ für alle $s, T \in [0, T^*]$ mit $s \leq T$.

Für $a^i = c^i = 0$ erhält man in der *i*-ten Komponente von σ_B die Vasiček-Volatilitätsstruktur als einen Spezialfall. Ferner ist auch die Ho-Lee-Volatilitätsstruktur ein Spezialfall der *abcd*-Formel. Diese erhält man für $a^i = d^i = 0$ und $b^i \ge 0$ beliebig, wobei dann der Wert von b^i keinen Einfluss hat.

Für den Aktienindex wird $\sigma_S : [0, T^*] \to \mathbb{R}^D, \sigma_S(s) := \sigma_{S,0}, \text{ mit } \sigma_{S,0} \in \mathbb{R}^D$ als Volatilitätsstruktur verwendet. Trivialerweise ist σ_S stationär, das heißt, es gilt $\sigma_S(s) = \sigma_S(0)$ für alle $s \in [0, T^*]$.

Die Parameter von σ_B und σ_S werden zu einem Vektor $\vartheta_{Volatilitätsstrukturen} \in \Theta_{Volatilitätsstrukturen}$ zusammengefasst, wobei $\Theta_{Volatilitätsstrukturen}$ die Menge der in Betracht gezogenen Vektoren ist.

6.1.4 Der Zeitskalierungsparameter

In der vorliegenden Arbeit wird zwischen der Modellzeiteinheit und der Zeiteinheit der realen Welt unterschieden. Der Zeitskalierungsparameter $\vartheta_{Zeitskalierung}$ ist eine Zahl aus $\Theta_{Zeitskalierung} := (0, \infty)$ und gibt an, wie vielen Jahren eine Modellzeiteinheit entspricht.

Der Hintergrund für die Einführung dieses zusätzlichen Parameters ist der folgende. Bei der Verwendung eines GH-Prozesses als Treiber ist L_1 stets Ddimensional GH-verteilt. Weil die Klasse der GH-Verteilungen jedoch nicht stabil unter Faltung ist, ist L_t für alle $t \in (0, T^*] \setminus \{1\}$ im Allgemeinen nicht GH-verteilt. Falls man also nicht zwischen Modellzeiteinheiten und Zeiteinheiten der realen Welt unterscheidet, so fällt die Länge der Zeitintervalle, über die die Zuwächse des treibenden Prozesses GH-verteilt sind, mit der gewählten Zeiteinheit der realen Welt zusammen und ist damit willkürlich. Durch die Unterscheidung zwischen Modellzeiteinheiten und Zeiteinheiten der realen Welt und die damit verbundene Einführung eines Zeitskalierungsparameters fällt diese Willkür weg, weil dann die Länge der Zeitintervalle (nun gemessen in Zeiteinheiten der realen Welt), über die der Zuwachs des treibenden Prozesses GH-verteilt ist, eben nicht mehr von der Wahl der Zeiteinheit der realen Welt abhängt.

Falls der treibende Lévy-Prozess aus einer Prozessklasse gewählt wird, die von einer faltungsstabilen Klasse von unbegrenzt teilbaren Verteilungen erzeugt wird, so ist der Zeitskalierungsparameter überflüssig. Für den Wiener-Prozess ist das beispielsweise der Fall.

6.1.5 Zusammenfassung der Modellparameter zu einem Vektor

Der Zeitskalierungsparameter $\vartheta_{Zeitskalierung}$, die Parameter $\vartheta_{Treiber}$ der Verteilung von L_1 (wobei hier der Zeitpunkt 1 als gemessen in Modellzeiteinheiten zu verstehen ist) sowie die Parameter $\vartheta_{Volatilitätsstrukturen}$ der Volatilitätsstrukturen der Zinsstruktur σ_B und des Aktienindex σ_S werden zu einem Vektor

$$\vartheta := (\vartheta_{Zeitskalierung}, \vartheta_{Treiber}, \vartheta_{Volatilitätsstrukturen})$$

zusammengefasst. Ferner wird die Menge der in Betracht gezogenen Parametervektoren mit

$$\Theta := \Theta_{Zeitskalierung} \times \Theta_{Treiber} \times \Theta_{Volatilitätsstrukturen} \subset \mathbb{R}^p$$

bezeichnet. Hierbei ist $p \in \mathbb{N}$.

6.2 Das Kalibrierungsportfolio und seine Modellierung

Ohne Beschränkung der Allgemeinheit seien der Nullpunkt der Zeitskala in der realen Welt und der Nullpunkt der Zeitskala im Modell so gewählt, dass der Zeitpunkt 0 in der realen Welt dem Zeitpunkt 0 im Modell entspricht. Aufgrund des Zeitskalierungsparameters fällt eine Zeiteinheit für die reale Welt im Allgemeinen jedoch nicht mit einer Modellzeiteinheit zusammen. Das hat zur Folge, dass in einem Modell, das den Zeitskalierungsparameter $\vartheta_{Zeitskalierung}$ verwendet, für $t \in \mathbb{R}_+$, $\Delta \in (0, \infty)$ und $T \in (\Delta, \infty)$ die Log-Returns $\mathrm{LR}_{B(\cdot,t+T)}(t,t+\Delta)$ und $\mathrm{LR}_{S(\cdot)}(t,t+\Delta)$, hier interpretiert als Größen der realen Welt, durch die Zufallsvariablen

$$\mathrm{LR}_{B\left(\cdot,\frac{t+T}{\vartheta_{Zeitskalierung}}\right)}\left(\frac{t}{\vartheta_{Zeitskalierung}},\frac{t+\Delta}{\vartheta_{Zeitskalierung}}\right)$$

und

$$\operatorname{LR}_{S(\cdot)}\left(\frac{t}{\vartheta_{Zeitskalierung}}, \frac{t+\Delta}{\vartheta_{Zeitskalierung}}\right)$$

modelliert werden. Deshalb ist es für die weiteren Überlegungen wichtig, stets zwischen den Größen der realen Welt und den Zufallsvariablen, die zu ihrer Modellierung herangezogen werden, präzise zu unterscheiden. Während Größen der realen Welt unabhängig von $\vartheta_{Zeitskalierung}$ sind, hängt es sehr wohl von $\vartheta_{Zeitskalierung}$ ab, durch welche Zufallsvariablen, also Modellgrößen, sie modelliert werden.

Die Gesamtheit der Marktgrößen der realen Welt, deren historische Realisierungen als Datengrundlage für die Kalibrierung dienen, wird als Kalibrierungsportfolio bezeichnet. Seien nun $N, K \in \mathbb{N}$. Weiterhin seien $\Delta \in (0, \infty)$ und $T_1, \ldots, T_{N-1} \in (\Delta, \infty)$ und es gelte $T_{n_1} \neq T_{n_2}$ für $n_1, n_2 \in \{1, \ldots, N-1\}$ mit $n_1 \neq n_2$. Ferner seien $t_1, \ldots, t_K \in [0, \infty)$ mit $t_1 + \Delta \leq t_2, \ldots, t_{K-1} + \Delta \leq t_K$. Als Kalibrierungsportfolio wird die Familie von Vektoren

$$\left(\operatorname{LR}_{B(\cdot,t_{k}+T_{1})}(t_{k},t_{k}+\Delta),\ldots,\operatorname{LR}_{B(\cdot,t_{k}+T_{N-1})}(t_{k},t_{k}+\Delta),\operatorname{LR}_{S(\cdot)}(t_{k},t_{k}+\Delta)\right),$$

 $k \in \{1, \ldots, K\}$, gewählt, wobei hier die auftretenden Log-Returns als (synthetische) Größen der realen Welt zu verstehen sind. Weil in jedem Eintrag dieser Vektoren ein Log-Return steht, werden sie hier als Log-Return-Vektoren bezeichnet. Die historischen Realisierungen der Log-Return-Vektoren aus dem Kalibrierungsportfolio werden im Folgenden mit x_1, \ldots, x_K bezeichnet.

Der Zeitskalierungsparameter $\vartheta_{Zeitskalierung}$ gehört zu den Parametern, die bei der Kalibrierung geschätzt werden. Das bedeutet unter anderem, dass vor der Kalibrierung und auch während des Kalibrierungsprozesses noch nicht klar ist, durch welche Zufallsvektoren X_1, \ldots, X_K , also durch welche Modellgrößen, die Log-Return-Vektoren im Kalibrierungsportfolio, die ja Größen der realen Welt sind, modelliert werden. Dies steht erst dann fest, wenn im Rahmen der Kalibrierung die Schätzung des numerischen Wertes von $\vartheta_{Zeitskalierung}$ abgeschlossen ist.

6.3 Die Kalibrierungsarten

Für jedes $\vartheta_{Zeitskalierung} \in \Theta_{Zeitskalierung}$, für das $(t_K + T_n)/\vartheta_{Zeitskalierung} \leq T^*$ für alle $n \in \{1, \ldots, N-1\}$ gilt, genügen im dazugehörigen Modell die Zufallsvektoren X_1, \ldots, X_K , die zur Modellierung der Log-Return-Vektoren aus dem Kalibrierungsportfolio herangezogen werden, den Voraussetzungen von Teilaussage 2 von Korollar 8 und sind damit stochastisch unabhängig und identisch verteilt. Dabei ist $T^* > 0$ der endliche Zeithorizont des Modells. Im Folgenden wird stets davon ausgegangen, dass T^* hinreichend groß ist, damit diese Eigenschaft erfüllt ist.

Hiermit ist x_1, \ldots, x_K eine Stichprobe von stochastisch unabhängigen und identisch verteilten Realisierungen und das Problem kann auf die Situation des in Abschnitt 5.4 betrachteten Zufallsvektors X zurückgeführt werden, wobei der Parametervektor $\xi = (\xi_{Integrationsgrenzen}, \xi_{Integrator}, \xi_{Integranden})$ von X eine Funktion der Modellparameter ϑ ist. Dabei ist zu beachten, dass das Paar $(\Delta_1, \Delta_2) = \xi_{Integrationsgrenzen}$ der Integrationsgrenzen von X sowie die Modellrestlaufzeiten, die im Parametervektor $\xi_{Integranden}$ der Integranden von X kodifiziert sind, vom Zeitskalierungsparameter $\vartheta_{Zeitskalierung}$ abhängen. Zur Hervorhebung der Abhängigkeit des Vektors ξ vom Vektor ϑ wird die Schreibweise $\xi = \xi(\vartheta)$ verwendet und es ist jeweils aus zum Zusammenhang zu erkennen, ob es sich bei ξ um einen Vektor $\xi \in \Xi$ oder um die Abbildung $\xi : \Theta \to \Xi$ handelt. Diese Art der Notation ist zwar formal nicht korrekt, sie trägt jedoch zur Vereinfachung der Darstellung bei und Verwechslungen sind ausgeschlossen. Bei der Kalibrierung wird nun derjenige Satz von Modellparametern gesucht, für den das dazugehörige Modell die tatsächlich beobachteten Realisierungen der Log-Return-Vektoren aus dem Kalibrierungsportfolio nach gewissen Kriterien am besten beschreibt. Dies geschieht durch die Optimierung gewisser Zielfunktionen, die in den Abschnitten 6.3.2 und 6.3.3 vorgestellt werden. Zuvor wird jedoch in Abschnitt 6.3.1 die Parametrisierung des Parameterraums behandelt. In Abschnitt 6.3.4 wird schließlich noch auf einige zusätzliche Nebenbedingungen, die bei der Kalibrierung angewandt werden, eingegangen.

6.3.1 Kalibrierung als nichtlineare Optimierung und Parametrisierung des Parameterraums

Die Kalibrierung ist ein nichtlineares Optimierungsproblem, das hier durch iterative Verfahren gelöst wird. Sobald ein Kalibrierungskriterium und damit eine Zielfunktion $Z: \Theta \to \mathbb{R}$ gewählt ist, lässt es sich formulieren als

maximiere/minimiere	$Z(\vartheta)$ für $\vartheta\in\Theta$
unter den Nebenbedingungen	$g_i(\vartheta) \leq 0$ für $i \in \{1, \dots, l_U\}$
und	$h_j(\vartheta) = 0$ für $j \in \{1, \ldots, l_G\}$

für geeignete nichtlineare Nebenbedingungen in Ungleichungsform $g_i : \Theta \to \mathbb{R}, i \in \{1, \ldots, l_U\}$, und nichtlineare Nebenbedingungen in Gleichungsform $h_j : \Theta \to \mathbb{R}, j \in \{1, \ldots, l_G\}$, wobei $l_U, l_G \in \mathbb{N}_0$.

Eine direkte Optimierung über $\vartheta \in \Theta$ ist zwar prinzipiell möglich, hat jedoch einige Nachteile. Im Parametervektor ϑ sind die Informationen über den Parameter Δ des treibenden GH-Prozesses enthalten. Δ ist eine symmetrische und strikt positiv definite Matrix mit $\det(\Delta) = 1$. Eine numerisch effizient auswertbare Kodifizierung dieser Anforderungen an Δ ist im vorliegenden Fall nicht trivial und für die Verarbeitung durch ein Computerprogramm, das einen Optimierungsalgorithmus ausführt, ungeeignet. Weiterhin sind hier in manchen Fällen die partiellen Ableitungen von Z nach den unterschiedlichen Einträgen von ϑ von komplett unterschiedlichen Größenordnungen, was bei einer endlichen Rechengenauigkeit zu unerwünschten Effekten führen kann.

Empirische Untersuchungen zeigen, dass die in MATLAB vorprogrammierten Optimierungsalgorithmen am besten funktionieren, wenn

- die Menge, über die optimiert wird, möglichst einfach, bestenfalls ein mehrdimensionales Intervall, ist,
- die Anzahl der nichtlinearen Nebenbedingungen in Ungleichungsform möglichst gering ist,

- keine nichtlinearen Nebenbedingungen in Gleichungsform vorliegen und
- alle partiellen Ableitungen der Zielfunktion von einer ähnlichen Größenordnung sind.

Deshalb wird in der vorliegenden Arbeit bei der Kalibrierung die nichtlineare Optimierung nicht direkt über die Menge Θ ausgeführt. Stattdessen wird Θ in geeigneter Weise parametrisiert.

Es wird ein Tripel (q_Z, R_Z, ϑ_Z) gewählt, das aus einer Zahl $q_Z \in \mathbb{N}$, einem mehrdimensionalen Intervall $R_Z \subset \mathbb{R}^{q_Z}$ und einer Abbildung $\vartheta_Z : R_Z \to \Theta$ besteht. Die Wahl erfolgt mit dem Ziel, die folgenden Kriterien möglichst gut zu erfüllen.

- q_Z ist möglichst klein.
- Für alle $x \in R_Z$ erfüllt $\vartheta_Z(x)$ alle nichtlinearen Nebenbedingungen in Gleichungsform, also $h_j(\vartheta_Z(x)) = 0$ für alle $j \in \{1, \ldots, l_G\}$.
- Die zulässige Menge

$$\{\vartheta \in \Theta | g_i(\vartheta) \le 0, \ i \in \{1, \dots, l_U\}, \ h_i(\vartheta) = 0, j \in \{1, \dots, l_G\}\}$$

wird von $\vartheta_Z(R_Z)$ vollständig überdeckt.

• $Z \circ \vartheta_Z$ ist stetig partiell differenzierbar und alle partiellen Ableitungen erster Ordnung liegen in einer ähnlichen Größenordnung.

Es sei angemerkt, dass ϑ_Z nicht injektiv zu sein braucht und es sich deshalb im Allgemeinen nicht um eine Karte im strengen Sinne der Differentialgeometrie handelt.

Hierdurch lässt sich das nichtlineare Optimierungsproblem umformulieren zu

maximiere/minimiere $Z(\vartheta_Z(x))$ für $x \in R_Z$ unter den Nebenbedingungen $g_i(\vartheta_Z(x)) \leq 0$ für $i \in \{1, \dots, l_U\}$.

Insbesondere sind so aufgrund der Wahl der Parametrisierung ϑ_Z keine nichtlinearen Nebenbedingungen in Gleichungsform erforderlich. Ist das Ergebnis dieser nichtlinearen Optimierung der Vektor $x_0 \in R_Z$, so wird der Parametervektor $\vartheta_Z(x_0)$ als Ergebnis der Kalibrierung gewählt.
6.3.2 Maximum-Likelihood-Kalibrierung

Bei der Kalibrierung nach dem Maximum-Likelihood-Kriterium wird der Schätzer

$$\hat{\vartheta} = \operatorname*{arg\,max}_{\vartheta \in \Theta} \ell(\xi(\vartheta))$$

verwendet. Bei der hierfür auszuführenden nichtlinearen Optimierung wird die Zielfunktion $Z(\vartheta) = \ell(\xi(\vartheta))$ nach der in Abschnitt 6.3.1 beschriebenen Methode optimiert. Hierbei wird die Log-Likelihood ℓ nach dem in Abschnitt 5.4.1 beschriebenen Verfahren approximiert.

Bemerkung 91. Bei einer Kalibrierung nach dem Maximum-Likelihood-Kriterium ist es nicht notwendigerweise erforderlich, dass jedem der Log-Return-Vektoren im Kalibrierungsportfolio ein Zeitintervall derselben Länge Δ zugrunde liegt. Stattdessen ist es prinzipiell denkbar, für jeden der KVektoren ein eigens $\Delta_k \in (0, \infty), k \in \{1, \ldots, K\}$, als Länge des entsprechenden Zeitintervalls zuzulassen. Hierdurch kann beispielsweise abgebildet werden, dass in gewissen Zinsberechnungsmethoden (engl. day count convention) je zwei aufeinanderfolgende Handelstage nicht immer denselben zeitlichen Abstand haben.

Eine Auswirkung hiervon ist jedoch, dass dann im Modell x_1, \ldots, x_K im Allgemeinen keine Stichprobe von identisch verteilten Realisierungen mehr ist. Eine approximative Berechnung der Log-Likelihood ist dann zwar immer noch möglich. Aber im Programmablaufplan in Abbildung 7 genügt es dann nicht mehr, die Unterprogramme, die durch die Programmablaufpläne in den Abbildungen 1 und 3 beschrieben werden, lediglich einmal zu durchlaufen. Stattdessen müssen diese für jedes Element der Menge $\{\Delta_k | k \in \{1, \ldots, K\}\}$ eigens durchlaufen werden. Hieraus resultiert ein deutlich höherer Rechenaufwand.

Derartige Verallgemeinerungen werden in der vorliegenden Arbeit nicht weiter untersucht.

6.3.3 Cramér-von-Mises- und Anderson-Darling-Kalibrierung

Sei $\mu \in \{0, 1\}$. Für jede Permutation σ von $\{1, \ldots, N\}$ gibt es eine dazugehörige Teststatistik $A_{R_{\sigma},K,\mu}^2$. Für die Zwecke der Kalibrierung wird keine dieser Teststatistiken gegenüber den anderen als in irgendeiner Art und Weise ausgezeichnet angesehen. Deshalb wird nicht bloß der Wert von $A_{R_{\sigma},K,\mu}^2$ für eine einzige willkürlich gewählte Permutation σ betrachtet. Stattdessen wird

$$\hat{\vartheta} = \operatorname*{arg\,min}_{\vartheta \in \Theta} \sum_{\{\sigma:\{1,\dots,N\} \to \{1,\dots,N\} \mid \sigma \text{ ist eine Permutation}\}} \left(A_{R_{\sigma,K,\mu}}^2(\xi(\vartheta))\right)^2$$

als Schätzer gewählt. Für $\mu = 0$ liegt das Cramér-von-Mises-Kriterium vor und für $\mu = 1$ erhält man das Anderson-Darling-Kriterium. Um die Abhängigkeit der Teststatistiken vom Parametervektor $\xi \in \Xi$ des Zufallsvektors Xzu unterstreichen, wird hier $A_{R_{\sigma},K,\mu}^2(\xi)$ statt bloß $A_{R_{\sigma},K,\mu}^2$ geschrieben. Bei der nichtlineare Optimierung wird die Zielfunktion

$$Z(\vartheta) = \sum_{\{\sigma:\{1,\dots,N\}\to\{1,\dots,N\}\mid\sigma\text{ ist eine Permutation}\}} \left(A_{R_{\sigma},K,\mu}^2(\xi(\vartheta))\right)^2$$

nach der in Abschnitt 6.3.1 beschriebenen Methode optimiert. Hierbei werden die Teststatistiken $A^2_{R_{\sigma},K,\mu}$ nach dem in Abschnitt 5.4.3 beschriebenen Verfahren approximiert.

6.3.4 Zusätzliche Modifikationen des Optimierungsproblems

Zusätzlich zu den Nebenbedingungen, die garantieren, dass der Optimierungsalgorithmus lediglich zulässige Parametervektoren $\vartheta \in \Theta$ in Betracht zieht, werden bei allen Kalibrierungskriterien noch die folgenden beiden Modifikationen mit einbezogen.

Größte akzeptable Kondition von Σ : Auf dem Weg zur approximativen Berechnung der Log-Likelihood ℓ bzw. der Teststatistiken $A^2_{R_{\sigma},K,\mu}$ ist als ein Zwischenschritt die Matrix $\operatorname{Chol}(\Sigma^{-1})$, wobei $\Sigma = \operatorname{Cov}(X)$, zu berechnen. Aufgrund der endlichen Rechengenauigkeit können hierbei in numerisch ungünstig gelagerten Extremfällen unerwünschte Ungenauigkeiten auftreten. Die Erfahrung zeigt, dass im Rahmen des nichtlinearen Optimierungsprozesses bei der Kalibrierung derartige ungünstige Fälle wirklich regelmäßig vorkommen und zu unerwünschten Effekten führen. Deshalb erhalten alle Vektoren von Modellparametern $\vartheta \in \Theta$, für die die Kondition Cond (Σ) der Kovarianzmatrix des dazugehörigen Zufallsvektors X einen gewissen Maximalwert überschreitet, einen Malus. Dieser Maximalwert ist so gewählt, dass er erstens um einige Größenordnungen größer als die Kondition der empirischen Kovarianzmatrix der betrachteten Stichprobe x_1, \ldots, x_K ist und dass zweitens für jede Matrix Σ , deren Kondition den Maximalwert unterschreitet, von einer zufriedenstellenden Genauigkeit für die berechneten numerischen Werte für Einträge von $\operatorname{Chol}(\Sigma^{-1})$ ausgegangen werden kann.

Positivitätsprüfung: Wie bereits in Bemerkung 50 erwähnt, kann es vorkommen, dass die Edgeworth-Approximation der Dichte auch negative Werte annimmt und somit keine Wahrscheinlichkeitsdichte ist. Derartige Fälle sind unerwünscht. Um sie so gut es geht auszuschließen, werden bei der Kalibrierung zwei weitere nichtlineare Nebenbedingung in Ungleichungsform verwendet. Erstens muss die approximative Dichte von X, ausgewertet an der Stelle x_k , für jedes $k \in \{1, \ldots, K\}$ strikt positiv sein. Zweitens wird ein gewisser Bereich um $0 = \mathbb{E} \left[\text{Chol}(\Sigma^{-1})(X - \mathbb{E}[X]) \right]$ mit einem dichten Gitter von Testpunkten überdeckt und die Edgeworth-Approximation der Dichte des Zufallsvektors $\text{Chol}(\Sigma^{-1})(X - \mathbb{E}[X])$ muss auf jedem dieser Testpunkte ebenfalls strikt positiv sein. Ist mindestens eine dieser Bedingungen verletzt, so wird der entsprechende Parametervektor $\vartheta \in \Theta$ als unzulässig gewertet.

6.4 Kalibrierungsergebnisse für deutsche Staatsanleihen und den DAX

6.4.1 Konfiguration bei der Kalibrierung

Der Datensatz und das Kalibrierungsportfolio: Der verwendete Datensatz für die Zinsstruktur besteht aus Tageswerten von Svensson-Parametern. Diese wurden von der Deutschen Bundesbank aus den Kursdaten von börsennotierten deutschen Staatsanleihen geschätzt. Für Details zur Svensson-Methode zur Schätzung von Forward-Raten siehe Svensson (1994). Hieraus werden die Preise von synthetischen Nullcoupon-Anleihen mit einem Nominalwert von 1 berechnet.

Für die Aktie werden Tagesschlusskurse des DAX verwendet. Diese wurden von der Firma Bloomberg bezogen.

Als historischer Kalibrierungszeitraum wird das Krisenjahr 2008 gewählt. Es ist aus finanzmathematischer Sicht besonders beachtenswert, weil sich in diesem Jahr das einschneidende Ereignis der Insolvenz der Großbank Lehmann Brothers ereignete, die bis kurz vor ihrem Insolvenzantrag vom 15. September 2008 noch als "too big to fail" galt. Das Jahr 2008 hatte 254 Handelstage, an denen deutsche Staatsanleihen und der DAX an der Frankfurter Wertpapierbörse gehandelt wurden. Hieraus werden für das Kalibrierungsportfolio K = 253 Log-Return-Vektoren konstruiert, denen disjunkte Zeitintervalle mit einer jeweiligen Länge von einem Handelstag zugrunde liegen. Für die Staatsanleihen werden die beiden Restlaufzeiten 3 Jahre und 10 Jahre berücksichtigt. Somit hat jeder dieser Log-Return-Vektoren N = 3 Einträge.

Zeitmessung in der realen Welt: Als Zeiteinheit der realen Welt wird das Jahr gewählt. Wie jedoch bereits diskutiert, hat diese Wahl keinen Einfluss auf die Kalibrierung.

Zur Vereinfachung werden die 254 Handelstage als zeitlich äquidistant angesehen. Das bedeutet, dass je zwei aufeinander folgende Handelstage einen zeitlichen Abstand von 1/254 Jahren haben.

Dimension des treibenden Prozesses und Einschränkungen des Parameterraums Θ : Als Dimension des treibenden GH-Prozesses $L = (L_t)_{t \in [0,T^*]}$ wird D = 2 gewählt. Zum Vergleich werden ebenfalls Kalibrierungen mit einem 2-dimensionalen Wiener-Prozess $W = (W_t)_{t \in [0,T^*]}$ als Treiber ausgeführt. Dabei wird der Parameterraum $\Theta_{Treiber}$ derart eingeschränkt, dass $\mathbb{E}[L_t] = \mathbb{E}[W_t] = 0$ für alle $t \in [0, T^*]$ gilt.

Das Hauptaugenmerk liegt nicht auf den Momenten erster Ordnung der Log-Return-Vektoren. Stattdessen liegt es auf den Momenten zweiter und höherer Ordnung. Deshalb werden für die Kalibrierung die Log-Return-Vektoren aus dem Kalibrierungsportfolio durch Subtraktion des empirischen Mittelwertvektors noch zentriert und für die Driftstrukturen der Zinsstruktur und der Aktie wird $\alpha_B = \alpha_S = 0$ verwendet. Die anschauliche Interpretation hiervon ist, dass die Driftstrukturen der Zinsstruktur und der Aktie derart gewählt werden, dass die Erwartungswertvektoren der identisch verteilten Modell-Log-Return-Vektoren und der empirische Mittelwertvektor der Log-Return-Vektoren aus dem Kalibrierungsportfolio zusammenfallen.

Ferner werden für die $\Theta_{Volatilitätsstrukturen}$ -Parameter c^1 , a^2 und d^2 der 2-dimensionalen *abcd*-Formel, die als Volatilitätsstruktur der Zinsstruktur eingesetzt wird, die Einschränkungen $c^1 = 0$, $a^2 = 0$ und $d^2 = -c^2$ vorgenommen. Hierdurch wird der Rechenaufwand deutlich verringert, während das Modell immer noch eine hervorragende Flexibilität bietet.

Anzahl S_{max} der verwendeten Korrekturterme für die Edgeworth-Approximation: Im Fall eines treibenden GH-Lévy-Prozesses werden bei allen eingesetzten Kalibrierungsarten für die Edgeworth-Approximation $S_{max} = 2$ Korrekturterme verwendet. Somit werden bei der approximativen Berechnung der Dichte und der bedingten Verteilungsfunktion alle Kumulanten κ_{α}^{X} einer Ordnung $\alpha \in \mathbb{N}_{0}^{N}$ mit $|\alpha| \leq S_{max} + 2 = 4$ berücksichtigt. Empirische Untersuchungen zeigen, dass hieraus bereits eine Approximationsgüte resultiert, die für die Zwecke der vorliegenden Arbeit hinreichend ist. Ferner bleibt der Rechenaufwand in einem vertretbaren Rahmen.

Im Fall eines treibenden Wiener-Prozesses sind die zur Kalibrierung herangezogenen Log-Return-Vektoren multivariat normalverteilt. Somit ist hier für die Berechnung der Dichte und der bedingten Verteilungsfunktion keine Edgeworth-Approximation erforderlich, vergleiche hierzu auch Bemerkung 48.

6.4.2 Vergleich der Ergebnisse für einen treibenden GH-Prozess und einen treibenden Wiener-Prozess

Das Modell, das von einem GH-Prozess getrieben wird (GH-Lévy-Modell) und das Modell, das von einem Wiener-Prozess getrieben wird (Brownsches Modell), werden beide je nach dem Maximum-Likelihood-, dem Cramér-von-Mises- und dem Anderson-Darling-Kriterium aus Abschnitt 6.3 kalibriert. Hieraus ergeben sich sechs geschätzte Modelle.

Es werden nun Schaubilder von ein- und zweidimensionalen Randverteilungen präsentiert, die man für die Maximum-Likelihood-Kalibrierung erhält. In Abbildung 10 sind für die eindimensionalen Randverteilungen die Histogramme sowie die Modell-Dichten dargestellt. In Abbildung 11 sind Kerndichteschätzungen der zweidimensionalen Randverteilungen aufgetragen. Dabei diente die Arbeit von Duong und Hazelton (2005) als Vorlage für die Implementierung der verwendeten Funktionalitäten zur multivariaten Kerndichteschätzung nach dem unbiased cross-validation Kriterium. In den Abbildungen 12, 13 und 14 zeigen die Schaubilder in der oberen Zeile die dazugehörigen Modell-Dichten, die Schaubilder in der mittleren Zeile zeigen die dazugehörigen Abweichungen von den Kerndichteschätzungen als 3-D surface plot und die Schaubilder in der unteren Zeile zeigen dieselben Abweichungen als scaled data image. Das bedeutet, die Werte der Abweichungen werden durch Farben dargestellt. Abgesehen von Abbildung 11 sind in allen Abbildungen auf der linken Seite die Schaubilder zum GH-Lévy-Modell dargestellt während auf der rechten Seite die Schaubilder zum Brownschen Modell zu sehen sind.

Bei allen dargestellten Randverteilungen sind die folgenden drei Tendenzen erkennbar. Verglichen mit den empirischen Daten haben die Modell-Dichten erstens zu wenig Masse im Zentrum und zweitens zu viel Masse in den Flanken. Drittens erkennt man, dass die Modell-Dichten tendenziell zu wenig Masse in den Tails haben, um die beobachteten Ausreißer zufriedenstellend zu erklären. All diese Abweichungen sind beim Brownschen Modell sichtbar stärker ausgeprägt als beim GH-Lévy-Modell.

Im vorliegenden Fall ist eine grafische Darstellung der Dichten von Randverteilungen zur Untersuchung der Anpassungsgüte keine angemessene Vorgehensweise. Bei ein- und zweidimensionalen Randverteilungen werden viele relevante Informationen getilgt. Bei den dargestellten Randverteilungen zeichnen sich zwar bereits Unterschiede in der Anpassungsgüte zwischen dem GH-Lévy-Modell und dem Brownschen Modell ab. Diese Schaubilder vermögen jedoch nicht aufzuzeigen, wie schwerwiegend die Unterschiede wirklich



Abbildung 10: Vergleich Dichten eindimensionale Randverteilungen für GH-Lévy-Modell und Brownsches Modell für Maximum-Likelihood-Kalibrierung



Abbildung 11: Empirische Dichten (Kerndichteschätzungen) zweidimensionale Randverteilungen



Abbildung 12: Vergleich Dichten zweidimensionale Randverteilung (Anleihen mit Restlaufzeiten 3 Jahre und 10 Jahre) für GH-Lévy-Modell und Brownsches Modell für Maximum-Likelihood-Kalibrierung



Abbildung 13: Vergleich Dichten zweidimensionale Randverteilung (Anleihe mit Restlaufzeit 3 Jahre und Aktie) für GH-Lévy-Modell und Brownsches Modell für Maximum-Likelihood-Kalibrierung



Abbildung 14: Vergleich Dichten zweidimensionale Randverteilung (Anleihe mit Restlaufzeit 10 Jahre und Aktie) für GH-Lévy-Modell und Brownsches Modell für Maximum-Likelihood-Kalibrierung

Kalibrierungsart	GH-Lévy-Modell	Brownsches Modell
Maximum-Likelihood	2.826, 9	2.800, 0
Cramér-von-Mises	2.776, 1	2.722, 6
Anderson-Darling	2.821, 1	2.767, 1

Tabelle 7: Ergebnisse Log-Likelihood für alle Kalibrierungsarten

sind. Die Betrachtung der dreidimensionalen Dichte ist zweckdienlicher. Diese lässt sich jedoch mit unbewegten Bildern nicht darstellen. In einem Film, der die dreidimensionale Dichte als eine Folge von zweidimensionalen Schnitten visualisiert und der ebenfalls im Rahmen der vorliegenden Arbeit erstellt wurde, sind die Unterschiede besser erkennbar. Am besten geeignet ist jedoch eine Betrachtung der Likelihood sowie die Ausführung von statistischen Tests.

In Tabelle 7 sind die Werte der Log-Likelihood für alle Kalibrierungsarten angegeben. Die Tabellen 8 und 9 zeigen die Testergebnisse der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests für die Maximum-Likelihood-Kalibrierung. Die Tabellen 10 und 11 zeigen die entsprechenden Werte für die Cramér-von-Mises-Kalibrierung, während in den Tabellen 12 und 13 diese Werte für die Anderson-Darling-Kalibrierung aufgetragen sind. Nach der Definition 59 gibt es für jede Permutation σ von $\{1, 2, 3\}$ eine multivariate Rosenblatt-Erweiterung des Cramér-von-Mises- und des Anderson-Darling-Tests. Von diesen Permutationen wird keine gegenüber den anderen als durch den Anwendungsfall ausgezeichnet betrachtet. Man vergleiche hierzu auch die Diskussion in Abschnitt 4.1.4 zur Willkür in Bezug auf die Permutation σ der Rosenblatt-Transformation. Deshalb werden die Werte der Teststatistiken sowie die dazugehörigen *p*-Werte jeweils für alle 3! = 6 Permutationen angegeben.

Für die folgende Diskussion wird ein Signifikanz
niveau von 5 %zugrunde gelegt.

Maximum-Likelihood-Kalibrierung: Die Zielfunktion hat für das GH-Lévy-Modell einen höheren und damit günstigeren Wert als für das Brownsche Modell. Hierbei ist zu bedenken, dass als Zielfunktion der Logarithmus der Likelihood verwendet wird. Somit bedeutet die vorliegende scheinbar gering ausfallende Differenz von 26,92 in der Log-Likelihood einen Unterschied von $4, 9 \cdot 10^{11}$ in der Likelihood.

Für das GH-Lévy-Modell lehnt kein einziger der 12 durchgeführten statistischen Tests ab. Für das Brownsche Modell hingegen lehnt lediglich die

Permutation	GH-Lévy-Modell		Brownsches Modell	
σ	$A^2_{R_{\sigma},K,0}$	<i>p</i> -Wert	$A^2_{R_{\sigma},K,0}$	$p ext{-Wert}$
(1, 2, 3)	0,101	0,269	0,231	0,023
(1, 3, 2)	0,099	0,282	0,229	0,024
(2, 1, 3)	0,157	0,085	0,268	0,012
(2, 3, 1)	0,165	0,074	0,278	0,010
(3, 1, 2)	0,109	0,228	0,256	0,015
(3, 2, 1)	0,158	0,084	0,274	0,011

Tabelle 8: Testergebnisse multivariate Rosenblatt-Erweiterung des Cramérvon-Mises-Tests für Maximum-Likelihood-Kalibrierung

Permutation	GH-Lévy-Modell		Brownsches Modell	
σ	$A^2_{R_{\sigma},K,1}$	$p ext{-Wert}$	$A^2_{R_{\sigma},K,1}$	$p ext{-Wert}$
(1, 2, 3)	1,333	0,159	2,217	0,024
(1, 3, 2)	1,254	0, 192	2,119	0,029
(2, 1, 3)	1,121	0,270	1,775	0,059
(2, 3, 1)	1,187	0,228	1,878	0,048
(3, 1, 2)	1,341	0,156	2,242	0,023
(3, 2, 1)	1,128	0,265	1,940	0,042

Tabelle 9: Testergebnisse multivariate Rosenblatt-Erweiterung des Anderson-Darling-Tests für Maximum-Likelihood-Kalibrierung

multivariate Rosenblatt-Erweiterung des Anderson-Darling-Tests zur Permutation $\sigma = (2, 1, 3)$ als einziger Test nicht ab.

Zusammengefasst schneidet das GH-Lévy-Modell gegenüber dem Brownschen Modell in allen Kriterien deutlich besser ab.

Cramér-von-Mises-Kalibrierung: Weil bei der Kalibrierung derjenige Parametersatz gesucht wird, für den die Summe der Quadrate der Teststatistiken aller multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises-Tests minimal ist, wird diese Summe systematisch unterschätzt und somit werden die dazugehörigen *p*-Werte tendenziell überschätzt. Somit sollten diese nicht als Grundlage für eine Testentscheidung herangezogen werden. Stattdessen sollten lediglich die Werte der Zielfunktion für die beiden Modelle miteinander verglichen werden.

Ein Vergleich zeigt, dass das GH-Lévy-Modell einen niedrigeren und damit besseren Wert für die Zielfunktion aufweist als das Brownsche Modell.

Permutation	GH-Lévy-Modell		Brownsches Modell	
σ	$A^2_{R_{\sigma},K,0}$	$p ext{-Wert}$	$A^2_{R_{\sigma},K,0}$	$p ext{-Wert}$
(1, 2, 3)	0,036	0,973	0,050	0,825
(1, 3, 2)	0,036	0,971	0,047	0,864
(2, 1, 3)	0,036	0,969	0,048	0,844
(2, 3, 1)	0,037	0,963	0,048	0,856
(3, 1, 2)	0,037	0,964	0,045	0,892
(3, 2, 1)	0,035	0,977	0,044	0,905

Tabelle 10: Testergebnisse multivariate Rosenblatt-Erweiterung des Cramérvon-Mises-Tests für Cramér-von-Mises-Kalibrierung

Permutation	GH-Lévy-Modell		Brownsches Modell	
σ	$A^2_{R_{\sigma},K,1}$	$p ext{-Wert}$	$A^2_{R_{\sigma},K,1}$	$p ext{-Wert}$
(1, 2, 3)	1,980	0,039	3,262	0,003
(1, 3, 2)	1,926	0,043	2,879	0,007
(2, 1, 3)	0,690	0,772	0,941	0,430
(2, 3, 1)	0,744	0,691	1,006	0,364
(3, 1, 2)	2,463	0,015	3,801	0,001
(3, 2, 1)	0,957	0,414	1,190	0,226

Tabelle 11: Testergebnisse multivariate Rosenblatt-Erweiterung des Anderson-Darling-Tests für Cramér-von-Mises-Kalibrierung

Erwartungsgemäß fällt für beide Modelle die Log-Likelihood kleiner aus als ihr entsprechender Wert bei der Maximum-Likelihood-Kalibrierung. Der Grund hierfür ist der folgende. Die multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises-Tests geben Abweichungen im Zentrum ein relativ hohes Gewicht. Deshalb führt die Cramér-von-Mises-Kalibrierung zu Modellen, bei denen die Verteilung im Zentrum besonders gut an die Beobachtungen angepasst ist. Dabei fallen Abweichungen in den Tails nur relativ wenig ins Gewicht. Zur Veranschaulichung sind in Abbildung 15 die Modell-Dichten für die zweidimensionale Randverteilung der Log-Returns der Anleihen mit den Restlaufzeiten 3 Jahre und 10 Jahre für beide Modelle aufgetragen. Im Gegensatz zu den entsprechenden Schaubildern in Abbildung 12 ist die Dichte zum Zentrum hin zusammengequetscht. So bleibt für die Tails kaum noch Masse zur Verfügung und die dort angesiedelten Datenpunkte erhalten eine deutlich geringere Likelihood.

Ein Vergleich der Log-Likelihoods zeigt, dass auch in diesem Kriterium das GH-Lévy-Modell besser abschneidet als das Brownsche Modell.



Abbildung 15: Dichten zweidimensionale Randverteilung (Anleihen mit Restlaufzeiten 3 Jahre und 10 Jahre) für GH-Lévy-Modell und Brownsches Modell für Cramér-von-Mises-Kalibrierung

Auch wenn bei den *p*-Werten aller multivariaten Rosenblatt-Erweiterungen des Anderson-Darling-Tests das Brownsche Modell weit hinter dem *GH*-Lévy-Modell zurückliegt, so sind die sechs Testentscheidungen für beide Modelle identisch. Die Tests zu den Permutationen (1, 2, 3), (1, 3, 2) sowie (3, 1, 2) lehnen ab und die Tests zu den Permutationen (2, 1, 3), (2, 3, 1) sowie (3, 2, 1) lehnen nicht ab.

Falls man also bei identischer Testentscheidung einen höheren p-Wert gegenüber einem niedrigeren bevorzugt, so schneidet das GH-Modell in allen Kriterien besser ab als das Brownsche Modell.

Anderson-Darling-Kalibrierung: Für alle sechs Permutationen ist der Wert der Teststatistik der dazugehörigen multivariaten Rosenblatt-Erweiterung des Anderson-Darling-Tests beim GH-Lévy-Modell niedriger und damit günstiger als beim Brownschen Modell. Wie bereits bei den p-Werten der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises-Tests bei der Cramér-von-Mises-Kalibrierung werden auch hier bei der Anderson-Darling-Kalibrierung die p-Werten der multivariaten Rosenblatt-Erweiterungen des Anderson-Darling-Tests tendenziell überschätzt und sollten nicht als Grundlage für eine Testentscheidung verwendet werden. Trotzdem sei hier das Folgende hervorgehoben. Obwohl bei der Kalibrierung derjenige Parametersatz gesucht wird, für den die Summe der Quadrate der Teststatistiken minimal ist, so ist für das Brownsche Modell dennoch der zur Permutation (3, 1, 2)dazugehörige p-Wert unter dem Signifikanzniveau von 5 %. Ferner sind die

p-Werte, die zu den Permutationen (1, 2, 3) und (1, 3, 2) gehören, gerade knapp über dem Signifikanzniveau. Beim GH-Lévy-Modell hingegen sind die p-Werte zu allen sechs Permutationen klar über dem Signifikanzniveau.

Permutation	GH-Lévy-Modell		Brownso	thes Modell
σ	$A^2_{R_{\sigma},K,0}$	$p ext{-Wert}$	$A^2_{R_{\sigma},K,0}$	$p ext{-Wert}$
(1, 2, 3)	0,063	0,633	0,091	0,338
(1, 3, 2)	0,065	0,600	0,091	0,338
(2, 1, 3)	0,084	0,395	0,088	0,362
(2, 3, 1)	0,095	0,305	0,094	0,315
(3, 1, 2)	0,075	0,482	0,094	0,311
(3, 2, 1)	0,101	0,271	0,090	0,340

Tabelle 12: Testergebnisse multivariate Rosenblatt-Erweiterung des Cramérvon-Mises-Tests für Anderson-Darling-Kalibrierung

Permutation	GH-Lévy-Modell		Brownsches Modell	
σ	$A^2_{R_{\sigma},K,1}$	$p ext{-Wert}$	$A^2_{R_{\sigma},K,1}$	$p ext{-Wert}$
(1, 2, 3)	1,169	0,238	1,750	0,063
(1, 3, 2)	1,182	0,231	1,779	0,059
(2, 1, 3)	0,748	0,685	0,918	0,457
(2, 3, 1)	0,823	0,577	0,996	0,373
(3, 1, 2)	1,360	0, 149	1,924	0,043
(3, 2, 1)	0,890	0,489	1,032	0,340

Tabelle 13: Testergebnisse multivariate Rosenblatt-Erweiterung des Anderson-Darling-Tests für Anderson-Darling-Kalibrierung

Für beide Modelle fällt die Log-Likelihood niedriger aus als ihr entsprechender Wert bei der Maximum-Likelihood-Kalibrierung. Bemerkenswert hierbei ist jedoch, dass die Log-Likelihood für das *GH*-Lévy-Modell bei der Anderson-Darling-Kalibrierung deutlich höher und damit günstiger ist als die Log-Likelihood für das Brownsche Modell bei der Maximum-Likelihood-Kalibrierung.

Für beide Modelle lehnt keine der multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises-Tests ab. Für die Permutationen (2, 3, 1) und (3, 2, 1) ist der *p*-Wert für das Brownsche Modell sogar höher als für das *GH*-Lévy-Modell. Für alle anderen Permutationen sind die *p*-Werte für das *GH*-Modell höher.

Abgesehen von den p-Werten von multivariaten Rosenblatt-Erweiterungen des Cramér-von-Mises-Tests zu zwei Permutationen, die aber dennoch für beide Modelle zu identischen Testentscheidungen führen, schneidet das GH-Modell in allen anderen Kriterien besser ab als das Brownsche Modell.

6.4.3 Erfahrungen und Vergleich der Kalibrierungskriterien

Für die drei Kalibrierungskriterien wurden numerische Experimente mit unterschiedlichen Konfigurationen durchgeführt. Dabei wurden der treibende Prozess, die Volatilitätsstruktur der Zinsstruktur sowie das Kalibrierungsportfolio und damit auch der Datensatz variiert. Dabei wurden die folgenden Beobachtungen gemacht.

Das Maximum-Likelihood-Kriterium liefert für Datensätze ohne Ausreißer recht zuverlässig hervorragende Ergebnisse. Doch bereits wenige große Ausreißer können einen enormen Einfluss haben und zu unerwünschten Effekten führen.

Das Cramér-von-Mises-Kriterium führt zu einer hervorragenden Anpassungsgüte im Zentrum der Verteilung. Dafür zeigen sich Schwächen bei der Zuverlässigkeit der Anpassungsgüte in den Tails.

Das Anderson-Darling-Kriterium liefert mit großer Zuverlässigkeit gute Ergebnisse sowohl im Zentrum als auch in den Flanken und in den Tails. Ferner ist es weniger empfindlich gegenüber Ausreißern als das Maximum-Likelihood-Kriterium.

Bei der nichtlinearen Optimierung im Rahmen des Kalibrierungsprozesses kommen die Cramér-von-Mises- und die Anderson-Darling-Kalibrierung mit wesentlich weniger Iterationen aus als die Maximum-Likelihood-Kalibrierung. Dafür ist die Auswertung der Zielfunktion bei der Cramér-von-Mises- und der Anderson-Darling-Kalibrierung deutlich aufwändiger und zeitintensiver, sodass die Rechenzeit für den gesamten Kalibrierungsprozess bei der Maximum-Likelihood-Kalibrierung am kleinsten ist. Mit Abstand am größten ist die Rechenzeit für die Anderson-Darling-Kalibrierung.

6.4.4 Ergebnisse mit anderen Volatilitätsstrukturen

Das GH-Lévy-Modell, das man bei einer Kalibrierung unter der weiter oben gemachten Einschränkung $c^1 = 0$, $a^2 = 0$ und $d^2 = -c^2$ erhält, wird im Folgenden als das Referenzmodell bezeichnet. Bei weiteren Kalibrierungen mit anderen Einschränkungen an die Parameter der 2-dimensionalen *abcd*-Formel findet man unter anderem die folgenden Ergebnisse.

• Die Einschränkung $a^1 = c^1 = a^2 = c^2 = 0$ liefert in beiden Komponenten der Volatilitätsstruktur der Zinsstruktur σ_B eine Vasiček-

Permutation	GH-Lévy-Modell			
σ	$A^2_{R_{\sigma},K,0}$	p-Wert	$A^2_{R_{\sigma},K,1}$	p-Wert
(1, 2, 3)	0,059	0,695	0,909	0,467
(1, 3, 2)	0,054	0,758	0,828	0,570
(2, 1, 3)	0,078	0,455	0,764	0,661
(2, 3, 1)	0,076	0,470	0,693	0,767
(3, 1, 2)	0,059	0,688	0,827	0,572
(3, 2, 1)	0,078	0,454	0,706	0,749

Tabelle 14: Testergebnisse multivariate Rosenblatt-Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests für die Parametereinschränkung $a^2 = c^2 = 0$ und Maximum-Likelihood-Kalibrierung

Volatilitätsstruktur. Die Anpassungsgüte des so geschätzten Modells liegt weit hinter der des Referenzmodells zurück und ist unbefriedigend.

- Die Einschränkung $a^2 = c^2 = 0$ führt dazu, dass in der ersten Komponente von σ_B die eindimensionale *abcd*-Formel ohne Einschränkungen und in der zweiten Komponente von σ_B eine Vasiček-Volatilitätsstruktur steht. Die Anpassungsgüte des so nach dem Maximum-Likelihood-Kriterium geschätzten *GH*-Lévy-Modells ist sogar noch besser als die des Referenzmodells, siehe Tabelle 14. Ein Nachteil ist jedoch, dass hierdurch bei der Kalibrierung für σ_B insgesamt sechs und nicht wie beim Referenzmodell lediglich fünf Parameter zu schätzen sind. Auch wenn es nur ein Parameter mehr ist, resultiert hieraus ein spürbar erhöhter Rechenaufwand. Deshalb wird hier auf eine Cramér-von-Mises und eine Anderson-Darling-Kalibrierung verzichtet.
- Es wäre interessant zu untersuchen, ob die Kalibrierung eines GH-Lévy-Modells ohne Einschränkung an die Parameter der 2-dimensionalen *abcd*-Formel in σ_B gegenüber dem Modell mit der Einschränkung $a^2 = c^2 = 0$ überhaupt noch einen Mehrwert liefert. Bereits bei den ersten Voruntersuchungen zeigt sich, dass der damit verbundene Rechenaufwand selbst für eine Maximum-Likelihood-Kalibrierung zu groß ist, um ihn mit der Edgeworth-Toolbox in ihrer aktuellen Version (Stand Juli 2016) und der verwendeten Hardware in vertretbarer Zeit bewältigen zu können.

6.4.5 Fazit

Bei allen durchgeführten Untersuchungen zur Anpassungsgüte zeigt sich klar und deutlich die Überlegenheit eines treibenden GH-Prozesses gegenüber einem treibenden Wiener-Prozess. Dafür ist die Kalibrierung eines GH-Lévy-Modells mit einem deutlich höheren Rechenaufwand verbunden als die Kalibrierung des entsprechenden Brownschen Modells.

Durch eine Auslagerung von weiteren Teilrechnungen auf Grafikprozessoren und eine Übersetzung von zeitkritischen Codepassagen, die nicht zur Abarbeitung durch einen Grafikprozessor geeignet sind, in eine Programmiersprache wie beispielsweise C/C++, die für numerische Berechnungen eine bessere Performance als MATLAB bietet, kann die Rechengeschwindigkeit weiter erhöht werden. Ferner führt der technische Fortschritt dazu, dass immer stärkere Hardware verfügbar wird. Somit lässt sich der Nachteil des erhöhten Rechenaufwands, der aus der Verwendung eines *GH*-Prozesses als Treiber resultiert, relativieren bzw. er relativiert sich mit der Zeit von selbst.

Mit dem Cramér-von-Mises- und dem Anderson-Darling-Kriterium stehen zwei neue Kalibrierungsverfahren zur Verfügung. Falls man lediglich an einer guten Anpassungsgüte im Zentrum interessiert ist und Abweichungen in den Tails akzeptieren kann, ist die Cramér-von-Mises-Kalibrierung eine echte Alternative zur Maximum-Likelihood-Kalibrierung. Falls man hingegen an einer akkuraten Anpassungsgüte über den gesamten Wertebereich hinweg interessiert ist, so ist die Anderson-Darling-Kalibrierung eine hervorragende Alternative zur Maximum-Likelihood-Kalibrierung. Insbesondere hat die Anderson-Darling-Kalibrierung gegenüber der Maximum-Likelihood-Kalibrierung den Vorteil, dass sie eine geringere Anfälligkeit gegenüber großen Ausreißern im Datensatz zeigt.

7 Zusammenfassung

Es wurde ein Hybridmodell zur Beschreibung der gemeinsamen Dynamik der Zinsstruktur und einer Aktie eingeführt, das von einem zeitinhomogenen Lévy-Prozess getrieben wird. Es wurden Eigenschaften der Klasse der Zufallsvektoren, die der Annahme ($\mathbb{E}M$) genügen, untersucht. Weiterhin wurden neue Anpassungstest für Zufallsvektoren und Copulas eingeführt, die multivariate Erweiterungen des Cramér-von-Mises- und des Anderson-Darling-Tests darstellen. Das Konvergenzverhalten der Teststatistiken für immer größer werdende Stichprobengrößen wurde unter der Nullhypothese H_0 und der Alternative H_1 bestimmt. Ferner wurden auf der Grundlage dieser multivariaten Anpassungstests mit dem Cramér-von-Mises- und dem Anderson-Darling-Kriterium zwei neue Kalibrierungsarten eingeführt.

Das Hybridmodell wurde nach dem Maximum-Likelihood-, dem Cramérvon-Mises- und dem Anderson-Darling-Kriterium an Log-Return-Vektoren kalibriert. Hierbei wurde zum einen ein zweidimensionaler GH-Prozess und zum anderen ein zweidimensionaler Wiener-Prozess als Treiber verwendet. Die bei der Kalibrierung zu berechnenden Werte der Wahrscheinlichkeitsdichte und der bedingten Verteilungsfunktion wurden unter Verwendung der multivariaten Edgeworth-Approximation approximiert. Ein Vergleich der Kalibrierungsergebnisse zeigte, dass ein treibender GH-Prozesses zu einer signifikant besseren Anpassungsgüte an die empirischen Daten führt, als das bei einem treibenden Wiener-Prozess der Fall ist.

Durch die Verwendung eines Kalibrierungsportfolios, das nicht bloß einzelne Log-Returns, sondern stattdessen Log-Return-Vektoren enthält, wurde gezeigt, dass das Hybridmodell zur Beschreibung der gemeinsamen Dynamik mehrerer Marktgrößen hervorragend geeignet ist. Insbesondere konnten mit einem einzigen Parametersatz sowohl Randverteilungen als auch Korrelationen präzise und konsistent abgebildet werden.

Ein Vergleich der Leistungsfähigkeit der Kalibrierungsarten zeigte, dass das Cramér-von-Mises- und das Anderson-Darling-Kriterium probate Alternativen zum Maximum-Likelihood-Kriterium sind. Dabei lieferte das Cramérvon-Mises-Kriterium eine hervorragende Anpassungsgüte im Zentrum der Verteilung, wobei es jedoch eine Schwäche in der Anpassungsgüte in den Tails zeigte. Das Anderson-Darling-Kriterium lieferte eine gute Anpassungsgüte über den gesamten Wertebereich hinweg. Ferner zeigte es sich weniger empfindlich gegenüber Ausreißern als das Maximum-Likelihood-Kriterium.

Die Probleme bei der Implementierung und ihre Lösung durch das Konzept der Code-Generierung wurden diskutiert. Der gewaltige Rechenaufwand für die Kalibrierungen wurde durch den zusätzlichen Einsatz von zwei Grafikprozessoren bewältigt.

A Anhang

A.1 Ein alternativer Beweis für $\mathbb{P}(D^2_{\mathbb{B},\mu} < \infty) = 1$ und einige fast sichere Pfadeigenschaften des Brownschen Blatts

Seien $N \in \mathbb{N}$, $\mathbb{B} = (\mathbb{B}_u)_{u \in [0,1]^N}$ eine N-parametrige Brownsche Brücke und $\mu \in [0, 2)$. Eine Kombination der Sätze 67 und 68 impliziert, dass

$$\int_{(0,1)^N} \frac{\left(\mathbb{B}_{\tilde{u}}\right)^2}{\left[\left(\prod_{n=1}^N \tilde{u}^n\right) \left(1 - \prod_{n=1}^N \tilde{u}^n\right)\right]^{\mu}} \,\mathrm{d}\tilde{u} < \infty$$

fast sicher gilt. Im vorliegenden Abschnitt A.1 wird ein alternativer Beweis für diese Tatsache vorgestellt, der fast sichere Pfadeigenschaften der Brownschen Brücke ausnutzt und damit eine Verallgemeinerung der Argumentation von Anderson und Darling (1952) vom Spezialfall N = 1 auf den vorliegenden allgemeinen Fall von $N \in \mathbb{N}$ darstellt. Auf dem Weg dorthin werden mit den Sätzen 98 und 100, dem Korollar 101, dem Satz 103 sowie der Bemerkung 104 fast sichere Pfadeigenschaften des Brownschen Blatts nachgewiesen, die von eigenem mathematischen Interesse sind. Dabei ist das Brownsche Blatt folgendermaßen definiert.

Definition 92 (Brownsches Blatt). Sei $N \in \mathbb{N}$. Ein N-parametriges Brownsches Blatt $\mathbb{W} = (\mathbb{W}_u)_{u \in [0,1]^N}$ ist ein zentrierter Gauß-Prozess mit der Kovarianzfunktion

$$[0,1]^N \times [0,1]^N \ni (u_1, u_2) \mapsto K_{\mathbb{W}}(u_1, u_2) := \prod_{n=1}^N (u_1^n \wedge u_2^n)$$

und dessen N-parametrige Pfade fast sicher stetig sind.

Für den gesamten restlichen Abschnitt A.1 seien stets $N \in \mathbb{N}$ und \mathbb{W} ein N-parametriges Brownsches Blatt. Diese Annahmen werden ohne erneute Erwähnung in jeder Definition, jedem Lemma, jedem Satz, jedem Korollar und jeder Bemerkung stillschweigend als Voraussetzung verwendet.

Lemma 93. Seien $m \in \{1, ..., N - 1\}$ und $u_0^n \in (0, 1]$ für alle $n \in \{m + 1, ..., N\}$.

Dann ist der Prozess
$$\mathbb{W}^{\left(u_{0}^{m+1},...,u_{0}^{N}\right)} = \left(\mathbb{W}^{\left(u_{0}^{m+1},...,u_{0}^{N}\right)}_{\left(u^{1},...,u^{m}\right)}\right)_{\left(u^{1},...,u^{m}\right)\in[0,1]^{m}} mit$$

$$\mathbb{W}_{(u^{1},\dots,u^{m})}^{\left(u_{0}^{m+1},\dots,u_{0}^{N}\right)} := \left(\prod_{n=m+1}^{N} u_{0}^{n}\right)^{-2} \mathbb{W}_{\left(u^{1},\dots,u^{m},u_{0}^{m+1},\dots,u_{0}^{N}\right)}$$

für $(u^1, \ldots, u^m) \in [0, 1]^m$ ein m-parametriges Brownsches Blatt.

Der Nachweis der Behauptungen in Lemma 93 ist trivial und wird ausgelassen.

A.1.1 Verhalten des Brownschen Blatts in der Nähe der Startwerte

Für $\gamma \in \mathbb{R}_+$ wird der Prozess $\mathbb{V}(\gamma) = (\mathbb{V}_u(\gamma))_{u \in [0,1]^N}$ mit

$$\mathbb{V}_{u}(\gamma) := \begin{cases} \frac{\left(\mathbb{W}_{\left(u^{1},\dots,u^{N}\right)}\right)^{2}}{\left[\prod\limits_{n=1}^{N}u^{n}\right]^{\gamma}} & \text{für } u \in (0,1]^{N} \\ 0 & \text{für } u \in [0,1]^{N} \backslash (0,1]^{N} \end{cases}$$
(30)

betrachtet und untersucht, für welche γ die *N*-parametrigen Pfade dieses Prozesses fast sicher stetig sind und für welche γ sie fast sicher nicht beschränkt sind. Beachtet man, dass fast sicher $\mathbb{W}_u = 0$ für alle $u \in [0, 1]^N \setminus (0, 1]^N$ gilt, so ist die anschauliche Interpretation hiervon die Fragestellung, wie schnell ein Brownsches Blatt seine Startwerte verlässt.

Im Spezialfall N = 1 ist das Brownsche Blatt W ein Wiener-Prozess und man erhält unter Verwendung des klassischen Gesetzes des iterierten Logarithmus, dass die einparametrigen Pfade von $\mathbb{V}(\gamma)$ für $\gamma \in [0, 1)$ fast sicher stetig und für $\gamma \in [1, \infty)$ fast sicher nicht beschränkt sind.

Es wird nun der allgemeinen Fall mit $N \in \mathbb{N}$ betrachtet. Lemma 93 impliziert, dass für $n_0 \in \{1, \ldots, N\}$ und $u^n \in (0, 1]$ für $n \in \{1, \ldots, N\} \setminus \{n_0\}$ der Prozess

$$[0,1] \ni u^{n_0} \mapsto \left(\prod_{n \in \{1,\dots,N\} \setminus \{n_0\}} u^n\right)^{-\frac{1}{2}} \mathbb{W}_{(u^1,\dots,u^N)}$$

ein Wiener-Prozess ist.

Somit sind für $\gamma \in [1, \infty)$ die einparametrigen Pfade des Schnitts

$$[0,1] \ni u^{n_0} \mapsto \mathbb{V}_{(u^1,\dots,u^N)}(\gamma)$$

fast sicher nicht beschränkt. Man kann sogar zeigen, dass für jedes fest gewählte $u_0 \in [0,1]^N \setminus (0,1]^N$ und jede Umgebung U_{u_0} von u_0 die eingeschränkten Pfade $U_{u_0} \ni u \mapsto \mathbb{V}_u(\gamma)$ fast sicher nicht beschränkt sind. Insbesondere sind damit die *N*-parametrigen Pfade von $\mathbb{V}(\gamma)$ fast sicher nicht beschränkt.

Für $\gamma \in [0,1)$ hingegen sind die einparametrigen Pfade des Schnitts $[0,1] \ni u^{n_0} \mapsto \mathbb{V}_{(u^1,\ldots,u^N)}(\gamma)$ fast sicher stetig. Die Nullmenge, auf der die Pfade der Schnitte nicht stetig sind, kann jedoch von den überabzählbar vielen Möglichkeiten für die Wahl von $n_0 \in \{1,\ldots,N\}$ und $u^n \in (0,1]$, $n \in \{1,\ldots,N\} \setminus \{n_0\}$, abhängen. Somit ist eine direkte Schlussfolgerung, dass die N-parametrigen Pfade von $\mathbb{V}(\gamma)$ fast sicher stetig sind, nicht zulässig.

Ein zentrales Ergebnis des vorliegenden Abschnitts A.1.1 ist das Korollar 101, welches besagt, dass es tatsächlich eine Nullmenge $\Omega_1 \subset \Omega$ gibt, auf deren Komplement die *N*-parametrigen Pfade von $\mathbb{V}(\gamma)$ für alle $\gamma \in [0, 1)$ stetig sind. Die Untersuchung wird begonnen mit einer kurzen Erinnerung an die Definition des Limes superior und des Limes inferior für eine Funktion, die von mehreren reellen Variablen abhängt.

Definition 94 (Limes superior und Limes inferior). Sei $f : \mathbb{R}^N \to \mathbb{R}$. Der Limes superior für $(\tilde{x}^1, \ldots, \tilde{x}^N) \searrow 0$ und der Limes inferior für $(\tilde{x}^1, \ldots, \tilde{x}^N) \searrow 0$ sind definiert durch

$$\limsup_{(\tilde{x}^1,\dots,\tilde{x}^N)\searrow 0} f(\tilde{x}) := \lim_{\tilde{\delta}\searrow 0} \sup_{\tilde{x}\in(0,\tilde{\delta})^N} f(\tilde{x})$$

und

$$\liminf_{(\tilde{x}^1,\dots,\tilde{x}^N)\searrow 0} f(\tilde{x}) := \lim_{\tilde{\delta}\searrow 0} \inf_{\tilde{x}\in(0,\tilde{\delta})^N} f(\tilde{x}).$$

Die hier angegebene Version des Gesetzes des iterierten Logarithmus folgt direkt aus Paranjape und Park (1973, Theorem 4), das dort als Zimmerman's Theorem bezeichnet wird.

Satz 95 (Gesetz des iterierten Logarithmus für das Brownsche Blatt). Es gilt

$$\limsup_{(\tilde{u}^1,\dots,\tilde{u}^N)\searrow 0} \frac{\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}}{\sqrt{2N\left(\prod_{n=1}^N \tilde{u}^n\right)\log\left(\log\left(\prod_{n=1}^N \frac{1}{\tilde{u}^n}\right)\right)}} = 1$$

und

$$\liminf_{(\tilde{u}^1,\dots,\tilde{u}^N)\searrow 0} \frac{\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}}{\sqrt{2N\left(\prod_{n=1}^N \tilde{u}^n\right)\log\left(\log\left(\prod_{n=1}^N \frac{1}{\tilde{u}^n}\right)\right)}} = -1$$

jeweils fast sicher.

Das nun folgende Lemma ist aus Paranjape und Park (1973, Theorem 2) übernommen.

Lemma 96. Set $u \in [0,1]^N$ und $z \in \mathbb{R}_+$.

Dann gilt

$$\mathbb{P}\left(\left\{\sup_{\tilde{u}\in[0,u]}\mathbb{W}_{\tilde{u}}\geq z\right\}\right)\leq 2^{N}\mathbb{P}\left(\{\mathbb{W}_{u}\geq z\}\right).$$

Lemma 97. Set $\gamma \in (1, \infty)$ und $\Gamma \in \left(\frac{1}{\gamma - 1}, \infty\right)$.

Dann gilt

$$\sum_{k=1}^{\infty} \sum_{s=0}^{\infty} \frac{1}{\left[k^{\Gamma} + s\right]^{\gamma}} < \infty.$$

Beweis: Die Doppelreihe lässt sich zerlegen als

$$\sum_{k=1}^{\infty} \sum_{s=0}^{\infty} \frac{1}{[k^{\Gamma} + s]^{\gamma}} = \sum_{k=1}^{\infty} \frac{1}{k^{\Gamma\gamma}} + \sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \frac{1}{[k^{\Gamma} + s]^{\gamma}}.$$

Aus $\Gamma\gamma > 1$ folgt $\sum_{k=1}^{\infty} \frac{1}{k^{\Gamma\gamma}} < \infty$. Ferner gilt nach dem Integralvergleichskriterium für Reihen

$$\sum_{s=1}^{\infty} \frac{1}{[k^{\Gamma} + s]^{\gamma}} \leq \int_{1}^{\infty} \frac{1}{[k^{\Gamma} + (\tilde{s} - 1)]^{\gamma}} d\tilde{s}$$
$$= \frac{1}{(\gamma - 1)k^{\Gamma(\gamma - 1)}}$$

für alle $k \in \mathbb{N}.$ Zusammen mit $\Gamma(\gamma-1) > 1$ folgt hieraus

$$\sum_{k=1}^{\infty} \sum_{s=1}^{\infty} \frac{1}{[k^{\Gamma} + s]^{\gamma}} \le \frac{1}{\gamma - 1} \sum_{k=1}^{\infty} \frac{1}{k^{\Gamma(\gamma - 1)}} < \infty$$

und damit die Behauptung.

Während das klassische Gesetz des iterierten Logarithmus für das Brownsche Blatt, das in Satz 95 angegeben ist, lediglich eine Aussage über das fast sichere Verhalten von W in einer Umgebung des Punktes $(0, \ldots, 0) \in [0, 1]^N$ macht, geben die nun folgenden Sätze 98 und 100 sowie das anschließende Korollar 101 einen tieferen Einblick in das fast sichere Verhalten von W in der Nähe des "N-dimensionalen linken, unteren Rands" $[0, 1]^N \setminus (0, 1]^N$ des Parameterraums $[0, 1]^N$.

Satz 98. Set $N \in \mathbb{N}$ mit $N \geq 2$. Ferner seten $\mathcal{M} \subset \{1, \ldots, N\}$ mit $|\mathcal{M}| \in \{1, \ldots, N-1\}$ und $a^n, b^n \in (0, 1]$ mit $a^n \leq b^n$ für $n \in \mathcal{M}$.

Dann gilt

$$\lim_{(\tilde{u}^n)_{n\in\mathcal{M}}\searrow 0} \sup_{\substack{\tilde{u}^n\in[a^n,b^n]\\f\ddot{u}r\ n\in\mathcal{M}^C}} \frac{\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}}{\left[2\left|\mathcal{M}\right|\left(\prod_{n=1}^N \tilde{u}^n\right)\log\left(\log\left(\prod_{n\in\mathcal{M}}\frac{1}{\tilde{u}^n}\right)\right)\right]^{\frac{1}{2}} = 1 \quad (31)$$

und

$$\lim_{(\tilde{u}^n)_{n\in\mathcal{M}}\searrow 0} \inf_{\substack{\tilde{u}^n\in[a^n,b^n]\\ f\ddot{u}r\ n\in\mathcal{M}^C}} \frac{\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}}{\left[2\left|\mathcal{M}\right|\left(\prod_{n=1}^N \tilde{u}^n\right)\log\left(\log\left(\prod_{n\in\mathcal{M}}\frac{1}{\tilde{u}^n}\right)\right)\right]^{\frac{1}{2}} = -1 \quad (32)$$

jeweils fast sicher.

Beweis: Weil ein *N*-parametriges Brownsches Blatt durch Permutation seiner Parameter wieder in ein *N*-parametriges Brownsches Blatt übergeht, kann ohne Beschränkung der Allgemeinheit angenommen werden, dass $\mathcal{M} = \{1, \ldots, m\}$ gilt, wobei $m := |\mathcal{M}|$. Es wird lediglich die fast sichere Gültigkeit von Gleichung (31) gezeigt. Die fast sichere Gültigkeit von Gleichung (32) lässt sich analog nachweisen.

Zur Abkürzung der Schreibweise wird die Funktion $G:(0,e^{-1})\times\mathbb{R}_+\to\mathbb{R}$ mit

$$G(x,y) := \sqrt{2xy \log\left(\log\left(\frac{1}{x}\right)\right)}$$
(33)

eingeführt. G ist auf $(0, e^{-\frac{1}{z}}) \times \mathbb{R}_+$ in beiden Argumenten streng monoton wachsend. Hierbei ist $z \approx 0,5671$ der Wert der Lambertschen W-Funktion, ausgewertet an der Stelle 1, also die eindeutige reelle Lösung der Gleichung

 $ze^z = 1.$

Nach Lemma 93 ist für jedes $(u_0^{m+1}, \ldots, u_0^N) \in (0, 1]^{N-m}$ der dort definierte Prozess $\mathbb{W}^{(u_0^{m+1}, \ldots, u_0^N)}$ ein *m*-parametriges Brownsches Blatt. Somit folgt aus Satz 95

$$\mathbb{P}\left(\left\{\limsup_{(\tilde{u}^1,\ldots,\tilde{u}^m)\searrow 0}\frac{\mathbb{W}_{\left(\tilde{u}^1,\ldots,\tilde{u}^m,u_0^{m+1},\ldots,u_0^N\right)}}{\sqrt{m}G\left(\prod_{n=1}^m\tilde{u}^n,\prod_{n=m+1}^Nu_0^n\right)}=1\right\}\right)=1$$

und damit

$$\mathbb{P}\left(\left\{ \limsup_{\substack{(\tilde{u}^{1},\ldots,\tilde{u}^{m})\searrow 0 \\ f\ddot{u}r \ n\in\{m+1,\ldots,N\}}} \sup_{\substack{\tilde{u}^{n}\in[a^{n},b^{n}] \\ f\ddot{u}r \ n\in\{m+1,\ldots,N\}}} \frac{\mathbb{W}_{(\tilde{u}^{1},\ldots,\tilde{u}^{N})}}{\sqrt{m}G\left(\prod_{n=1}^{m}\tilde{u}^{n},\prod_{n=m+1}^{N}\tilde{u}^{n}\right)} \ge 1\right\}\right) = 1.$$
(34)

Der Nachweis von

$$\mathbb{P}\left(\left\{\lim_{(\tilde{u}^{1},\ldots,\tilde{u}^{m})\searrow 0} \sup_{\substack{\tilde{u}^{n}\in[a^{n},b^{n}]\\f\ddot{u}r\ n\in\{m+1,\ldots,N\}}} \frac{\mathbb{W}_{(\tilde{u}^{1},\ldots,\tilde{u}^{N})}}{\sqrt{m}G\left(\prod_{n=1}^{m}\tilde{u}^{n},\prod_{n=m+1}^{N}\tilde{u}^{n}\right)} \leq 1\right\}\right) = 1$$
(35)

wird indirekt geführt. In Schritt 1) wird die Annahme formuliert, die in den Schritten 2) bis 4) zum Widerspruch geführt wird.

1) Formulierung der Annahme Es wird angenommen, dass ein
 $\varepsilon \in (0,1)$ existiert mit

$$\mathbb{P}\left(\left\{\underset{\substack{(\tilde{u}^{1},\ldots,\tilde{u}^{m})\searrow 0 \\ f\ddot{u}r \ n\in\{m+1,\ldots,N\}}}{\lim \sup} \frac{\mathbb{W}_{(\tilde{u}^{1},\ldots,\tilde{u}^{N})}}{\sqrt{m}G\left(\prod_{n=1}^{m}\tilde{u}^{n},\prod_{n=m+1}^{N}\tilde{u}^{n}\right)} > 1+\varepsilon\right\}\right) > 0.$$
(36)

2) Einführung der betrachteten Objekte

Definiere $K := \left(1 + \frac{\varepsilon}{2}\right)^{\frac{1}{N-m}} - 1$ und $\delta := \min_{n \in \{m+1,\dots,N\}} Ka^n$. Dann gilt K > 0 und damit auch $\delta > 0$. Aus Gleichung (36) folgt, dass ein Teilintervall

$$\left[\left(a_0^{m+1}, \dots, a_0^N \right), \left(b_0^{m+1}, \dots, b_0^N \right) \right] \subset \left[\left(a^{m+1}, \dots, a^N \right), \left(b^{m+1}, \dots, b^N \right) \right]$$

existiert mit $b_0^n - a_0^n \leq \delta$ für $n \in \{m + 1, \dots, N\}$ und $\mathbb{P}(A_0) > 0$, wobei

$$A_0 := \left\{ \limsup_{\substack{(\tilde{u}^1, \dots, \tilde{u}^m) \searrow 0 \\ f \ddot{u}^n \in \{m+1, \dots, N\}}} \sup_{\substack{\tilde{u}^n \in [a_0^n, b_0^n] \\ f \ddot{u}^n \ n \in \{m+1, \dots, N\}}} \frac{\mathbb{W}_{(\tilde{u}^1, \dots, \tilde{u}^N)}}{\sqrt{m} G\left(\prod_{n=1}^m \tilde{u}^n, \prod_{n=m+1}^N \tilde{u}^n\right)} > 1 + \varepsilon \right\}.$$

Wähle $q \in \left(\left(1 - \frac{\varepsilon}{2}\right)^{\frac{1}{m}}, 1 \right)$. Dann gilt

$$\gamma := (1+\varepsilon)^2 q^m \left(\prod_{n=m+1}^N \frac{a_0^n}{b_0^n}\right) > \frac{(1+\varepsilon)^2 \left(1-\frac{\varepsilon}{2}\right)}{1+\frac{\varepsilon}{2}} > 1.$$
(37)

Ferner definiere für $k \in \mathbb{N}$ und $\boldsymbol{\eta} = (\eta^1, \dots, \eta^m) \in \mathbb{N}_0^m$

$$A_{k,\eta} := \begin{cases} \sup_{\tilde{u}^n \in (q^{k+\eta^n+1}, q^{k+\eta^n}]} & \sup_{\tilde{u}^n \in [a_0^n, b_0^n] \\ f \ddot{u}r \ n \in \{1, \dots, m\} & f \ddot{u}r \ n \in \{m+1, \dots, N\} \end{cases}$$

$$\frac{\mathbb{W}_{(\tilde{u}^1,\ldots,\tilde{u}^N)}}{\sqrt{m}G\left(\prod\limits_{n=1}^m q^{k+\eta^n+1},\prod\limits_{n=m+1}^N a_0^n\right)} > 1+\varepsilon \Biggr\}\,,$$

 $A_k := \bigcup_{\eta \in \mathbb{N}_0^m} A_{k,\eta}, \ \Gamma := \left\lfloor \frac{1}{\gamma - 1} \right\rfloor + 1 \text{ und } B_k := A_{k^{\Gamma}}, \text{ wobei } \lfloor \cdot \rfloor \text{ die Abrundungsfunktion ist.}$

3) Behauptung: Es gilt $\mathbb{P}\left(\liminf_{k\to\infty}B_k\right)>0.$ Se
i $\omega\in A_0.$ Es gilt $q^k\searrow 0$ für $k\to\infty$ und damit nach Definition von
 A_0

$$\lim_{k \to \infty} \sup_{\substack{\tilde{u}^n \in \{0, q^k\} \\ f \ddot{u}r \ n \in \{1, \dots, m\} }} \sup_{\substack{\tilde{u}^n \in [a_0^n, b_0^n] \\ f \ddot{u}r \ n \in \{1, \dots, m\} }} \frac{\mathbb{W}_{(\tilde{u}^1, \dots, \tilde{u}^N)}(\omega)}{\sqrt{m}G\left(\prod_{n=1}^m \tilde{u}^n, \prod_{n=m+1}^N \tilde{u}^n\right)} > 1 + \varepsilon.$$

Somit existiert ein $k_0(\omega) \in \mathbb{N}$ mit $q^{mk_0(\omega)} \prod_{n=m+1}^N b_0^n < e^{-\frac{1}{z}}$ und

$$\sup_{\substack{\tilde{u}^n \in \left(0, q^k\right) \\ \tilde{u}^n \in \left\{n, \dots, m\right\} \\ f \ddot{u}^n n \in \left\{n, \dots, m\right\}}} \sup_{\substack{\tilde{u}^n \in \left[a_0^n, b_0^n\right] \\ f \ddot{u}^n n \in \left\{m+1, \dots, N\right\}}} \frac{\mathbb{W}_{\left(\tilde{u}^1, \dots, \tilde{u}^N\right)}(\omega)}{\sqrt{m} G\left(\prod_{n=1}^m \tilde{u}^n, \prod_{n=m+1}^N \tilde{u}^n\right)} > 1 + \varepsilon$$

für alle $k \in \mathbb{N}$ mit $k \ge k_0(\omega)$. Wie bereits weiter oben im Beweis ist dabei z der Wert der Lambertschen W-Funktion, ausgewertet an der Stelle 1.

Sei nun $k \in \mathbb{N}$ mit $k \geq k_0(\omega)$. Dann existieren $(u_k^1, \ldots, u_k^m) \in (0, q^k)^m$ und $u_k^n \in [a_0^n, b_0^n]$ für $n \in \{m + 1, \ldots, N\}$ mit

$$\frac{\mathbb{W}_{\left(u_{k}^{1},\ldots,u_{k}^{N}\right)}(\omega)}{\sqrt{m}G\left(\prod_{n=1}^{m}u_{k}^{n},\prod_{n=m+1}^{N}u_{k}^{n}\right)} > 1+\varepsilon$$

Aus $(0, q^k)^m \subset \bigcup_{\eta \in N_0^m} \prod_{n=1}^m \left(q^{k+\eta^n+1}, q^{k+\eta^n} \right)$ folgt weiter, dass es einen Multiindex $\boldsymbol{\beta}(k) = (\beta^1(k), \dots, \beta^m(k)) \in \mathbb{N}_0^m$ gibt mit $u_k^n \in \left(q^{k+\beta^n(k)+1}, q^{k+\beta^n(k)} \right)$ für $n \in \{1, \dots, m\}$. Weil G auf $(0, e^{-\frac{1}{z}}) \times \mathbb{R}_+$ in beiden Argumenten streng monoton wachsend ist, gilt damit auch

$$\frac{\mathbb{W}_{\left(u_{k}^{1},\dots,u_{k}^{N}\right)}(\omega)}{\sqrt{m}G\left(\prod_{n=1}^{m}q^{k+\beta^{n}(k)+1},\prod_{n=m+1}^{N}a_{0}^{n}\right)} \geq \frac{\mathbb{W}_{\left(u_{k}^{1},\dots,u_{k}^{N}\right)}(\omega)}{\sqrt{m}G\left(\prod_{n=1}^{m}u_{k}^{n},\prod_{n=m+1}^{N}u_{k}^{n}\right)} > 1+\varepsilon.$$

Hieraus folgt $\omega \in A_{k,\beta(k)}$ und damit $\omega \in A_k$.

Somit gilt $\omega \in B_k$ für alle $k \in \mathbb{N}$ mit $k \ge (k_0(\omega))^{\frac{1}{\Gamma}}$ und damit auch

$$A_0 \subset \liminf_{k \to \infty} B_k.$$

Hieraus folgt die behauptete Ungleichung $\mathbb{P}(\liminf_{k\to\infty} B_k) > 0.$

4) Behauptung: Es gilt $\mathbb{P}(\limsup_{k\to\infty} B_k) = 0.$ Wähle $k_0 \in \mathbb{N}$ hinreichend groß, sodass $\sqrt{2\gamma m \log(\log(q^{-mk_0}))} \geq \frac{1}{\sqrt{2\pi}}$ gilt. Seien nun $k \in \mathbb{N}$ mit $k \geq k_0$ und $\boldsymbol{\eta} = (\eta^1, \dots, \eta^m) \in \mathbb{N}_0^m$. Dann erhält man unter Verwendung von Lemma 96

$$\begin{split} \mathbb{P}(A_{k,\eta}) \\ &\leq \mathbb{P}\left(\left\{\sup_{\substack{\bar{u}^{n} \in \left(0, q^{k+\eta^{n}}\right] \\ f\bar{u}^{n} \in \left\{1, \dots, m\right\} \\ f\bar{u}^{n} n \in \left\{1, \dots, m\right\} \\ \hline \frac{\mathbb{W}_{\left(\bar{u}^{1}, \dots, \bar{u}^{N}\right)}}{\sqrt{m}G\left(\prod_{n=1}^{m} q^{k+\eta^{n}+1}, \prod_{n=m+1}^{N} a_{0}^{n}\right)} > 1 + \varepsilon\right\} \right) \\ &\leq 2^{N} \mathbb{P}\left(\left\{\frac{\mathbb{W}_{\left(q^{k+\eta^{1}}, \dots, q^{k+\eta^{m}}, b_{0}^{m+1}, \dots, b_{0}^{N}\right)}}{\sqrt{m}G\left(\prod_{n=1}^{m} q^{k+\eta^{n}+1}, \prod_{n=m+1}^{N} a_{0}^{n}\right)} > 1 + \varepsilon\right\} \right) \\ &= 2^{N} \left[1 - \Phi_{\mathcal{N}(0,1)}\left(\sqrt{2\gamma m \log\left(\log\left(q^{-\sum_{n=1}^{m} (k+\eta^{n}+1)\right)\right)}\right)}\right]. \end{split}$$

Für $x \ge \frac{1}{\sqrt{2\pi}}$ gilt die Abschätzung $1 - \Phi(x) \le \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{1}{2}x^2} \le e^{-\frac{1}{2}x^2}$. Aus $k \ge k_0$ folgt $\sqrt{2\gamma m \log \left(\log \left(q^{-\sum_{n=1}^m (k+\eta^n+1)}\right)\right)} \ge \frac{1}{\sqrt{2\pi}}$ und damit

$$\mathbb{P}(A_{k,\eta}) \leq 2^{N} \exp\left(-\gamma m \log\left(\log\left(q^{-\sum_{n=1}^{m}(k+\eta^{n}+1)}\right)\right)\right)$$
$$= \frac{2^{N}}{[\log\left(q^{-1}\right)]^{\gamma m}} \left[m(k+1) + \sum_{n=1}^{m}\eta^{n}\right]^{-\gamma m}.$$

Hiermit berechnet man weiter

$$\mathbb{P}(A_k) \leq \sum_{\boldsymbol{\eta} \in \mathbb{N}_0^m} \mathbb{P}(A_{k,\boldsymbol{\eta}}) \\
\leq \frac{2^N}{[\log(q^{-1})]^{\gamma m}} \sum_{\boldsymbol{\eta} \in \mathbb{N}_0^m} \left[\left(m(k+1) + \sum_{n=1}^m \eta^n \right) \right]^{-\gamma m} \\
= \frac{2^N}{[\log(q^{-1})]^{\gamma m}} \sum_{S \in \mathbb{N}_0} \binom{S+m-1}{m-1} [m(k+1)+S]^{-\gamma m} \\
\leq \frac{2^N}{[\log(q^{-1})]^{\gamma m}} \frac{1}{(m-1)!} \sum_{S \in \mathbb{N}_0} \frac{1}{[k+S]^{\gamma}}.$$

Hierbei wurde die Tatsache verwendet, dass es zu einem gegebenen $S \in \mathbb{N}_0$ genau $\binom{S+m-1}{S} = \binom{S+m-1}{m-1}$ verschiedene Multiindices $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^m) \in \mathbb{N}_0^m$ gibt mit $\sum_{n=1}^m \alpha^n = S$. Mit $l_0 := \left\lceil k_0^{\frac{1}{\Gamma}} \right\rceil$, wobei $\lceil \cdot \rceil$ die Aufrundungsfunktion ist, folgt unter Verwendung von Lemma 97

$$\sum_{k=l_0}^{\infty} \mathbb{P}(B_k) \le \frac{2^N}{\left[\log\left(q^{-1}\right)\right]^{\gamma m}} \frac{1}{(m-1)!} \sum_{k=1}^{\infty} \sum_{S \in \mathbb{N}_0} \frac{1}{\left[k^{\Gamma} + S\right]^{\gamma}} < \infty.$$

Somit impliziert das Borel-Cantelli-Lemma $\mathbb{P}(\limsup_{k\to\infty} B_k) = 0$, was im Widerspruch zur in Schritt 3) gezeigten Ungleichung $\mathbb{P}(\liminf_{k\to\infty} B_k) > 0$ steht. \Box

Bemerkung 99. Der vorstehende Satz ist inspiriert von Zimmerman (1972, Theorem 3). Einige ihrer Beweisideen, die in dem von ihr betrachteten Spezialfall N = 2 funktionieren, funktionieren im Fall N > 2 nicht mehr und wurden hier durch tiefergehende Überlegungen ersetzt.

Zimmerman formuliert für ein zweiparametriges Brownsches Blatt $\mathbb{W} = (\mathbb{W}_{(s,t)})_{(s,t)\in\mathbb{R}^2_+}$ und $a, b \in (0,\infty)$ mit $a \leq b$ die Behauptung

$$\mathbb{P}\left(\left\{\limsup_{s\to\infty} \sup_{t\in[a,b]} \frac{\mathbb{W}_{(s,t)}}{\left[2st\log\left(\log\left(st\right)\right)\right]^{\frac{1}{2}}} = 1\right\}\right) = 1.$$

Dabei wählt Zimmerman als Parameterraum des zweiparametrigen Brownschen Blatts nicht wie in der vorliegenden Arbeit das Intervall $[0, 1]^2$, sondern den gesamten Quadranten \mathbb{R}^2_+ . Bei ihrer Behauptung liegt jedoch ein Fehler vor. Richtig ist

$$\mathbb{P}\left(\left\{\limsup_{s\to\infty}\sup_{t\in[a,b]}\frac{\mathbb{W}_{(s,t)}}{\left[2st\log\left(\log\left(s\right)\right)\right]^{\frac{1}{2}}}=1\right\}\right)=1,$$

also ohne den Faktor t im Argument des iterierten Logarithmus. Zum einen lässt sich Zimmermans Argumentation nahezu unverändert auf den Nachweis von

$$\mathbb{P}\left(\left\{\limsup_{s\to\infty}\sup_{t\in[a,b]}\frac{\mathbb{W}_{(s,t)}}{\left[2st\log\left(\log\left(s\right)\right)\right]^{\frac{1}{2}}}<1+\varepsilon\right\}\right)=1$$

übertragen. Zum anderen argumentiert sie jedoch zum Nachweis von

$$\mathbb{P}\left(\left\{\limsup_{s\to\infty}\sup_{t\in[a,b]}\frac{\mathbb{W}_{(s,t)}}{\left[2st\log\left(\log\left(st\right)\right)\right]^{\frac{1}{2}}}>1-\varepsilon\right\}\right)=1,$$

dass dies aus dem Gesetz des iterierten Logarithmus für einen einparametrigen Wiener-Prozess folge. Das stimmt aber nicht, denn für t > 0 ist $\left(t^{-\frac{1}{2}} \mathbb{W}_{(s,t)}\right)_{s \in \mathbb{R}_+}$ ein einparametriger Wiener-Prozess und folglich gilt stattdessen

$$\mathbb{P}\left(\left\{\limsup_{s\to\infty}\sup_{t\in[a,b]}\frac{t^{-\frac{1}{2}}\mathbb{W}_{(s,t)}}{\left[2s\log\left(\log\left(s\right)\right)\right]^{\frac{1}{2}}}>1-\varepsilon\right\}\right)=1.$$

Satz 100. Es existiert eine Nullmenge $\Omega_0 \subset \Omega$, sodass es für jedes Tripel $(\varepsilon, \gamma, \omega) \in (0, \infty) \times (0, 1) \times \Omega_0^C$ ein $\delta(\varepsilon, \gamma, \omega) > 0$ gibt mit

$$\sup_{\tilde{u}\in(0,1]^N\setminus[\delta(\varepsilon,\gamma,\omega),1]^N}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{\gamma}} < 1+\varepsilon.$$
(38)

Beweis: Der Beweis ist in die Schritte 1) bis 6) unterteilt. Bevor in Schritt 6) die eigentliche Behauptung des Satzes gezeigt wird, wird zunächst als Vorarbeit in den Schritten 1) bis 5) nachgewiesen, dass für jedes fest gewählte Paar $(\varepsilon_0, \gamma_0) \in (0, \infty) \times (0, 1)$ eine Nullmenge $\Omega_0(\varepsilon_0, \gamma_0)$ existiert, sodass es für jedes $\omega \in (\Omega_0(\varepsilon_0, \gamma_0))^C$ ein $\delta(\varepsilon_0, \gamma_0, \omega) > 0$ gibt mit

$$\sup_{\tilde{u}\in(0,1]^N\setminus[\delta(\varepsilon_0,\gamma_0,\omega),1]^N}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{\gamma_0}}<1+\varepsilon_0.$$

Hierzu sei $\varepsilon_1 > 0$. *G* sei die in Gleichung (33) definierte Funktion. Ferner werden für $\mathcal{M} \subset \{1, \ldots, N\}$ mit $\mathcal{M} \neq \emptyset$ und $\mu, \nu \in (0, 1]$ die Mengen

$$A_{\mathcal{M},\mu,\nu} := \begin{cases} \begin{cases} \sup_{\substack{\tilde{u}^n \in (0,\mu) \\ f \ddot{u}^n n \in \mathcal{M} \\ f \ddot{u}^n n \in \{1,...,N\} \\ \end{cases}} \frac{\left(\mathbb{W}_{\left(\tilde{u}^1,...,\tilde{u}^N \right)} \right)^2}{2N \left[\prod_{n=1}^N \tilde{u}^n \right]^{\gamma_0}} < 1 + \varepsilon_0 \\ \\ \begin{cases} \sup_{\substack{\tilde{u}^n \in (0,\mu) \\ f \ddot{u}^n n \in \{1,...,N\} \\ f \ddot{u}^n n \in \{1,...,N\} \\ \end{cases}} \frac{\left(\mathbb{W}_{\left(\tilde{u}^1,...,\tilde{u}^N \right)} \right)^2}{2N \left[\prod_{n=1}^N \tilde{u}^n \right]^{\gamma_0}} < 1 + \varepsilon_0 \\ \end{cases} \\ f \ddot{u}^n \mathcal{M} = \{1,...,N\} \end{cases}$$

und

$$B_{\mathcal{M},\mu,\nu} := \begin{cases} \begin{cases} \sup_{\substack{\tilde{u}^n \in (0,\mu) \\ f\ddot{u}r \ n \in \mathcal{M} \ f\ddot{u}r \ n \in \mathcal{M}^C \ }} \sup_{\substack{\tilde{u}^n \in [\nu,1] \\ f\ddot{u}r \ n \in \mathcal{M}^C \ }} \frac{\left(\mathbb{W}_{\left(\tilde{u}^1,\dots,\tilde{u}^N\right)}\right)^2}{|\mathcal{M}| \left[G\left(\prod_{n \in \mathcal{M}} \tilde{u}^n, \prod_{n \in \mathcal{M}^C} \tilde{u}^n\right)\right]^2} < 1 + \varepsilon_0 \\ & \text{für } \mathcal{M} \neq \{1,\dots,N\} \\ \begin{cases} \sup_{\substack{\tilde{u}^n \in (0,\mu) \\ f\ddot{u}r \ n \in \{1,\dots,N\} \ }} \frac{\left(\mathbb{W}_{\left(\tilde{u}^1,\dots,\tilde{u}^N\right)}\right)^2}{N \left[G\left(\prod_{n \in \mathcal{M}} \tilde{u}^n,1\right)\right]^2} < 1 + \varepsilon_0 \\ \end{cases} \\ & \text{für } \mathcal{M} = \{1,\dots,N\} \end{cases}$$

definiert. Dabei sei angemerkt dass die Mengen $A_{\{1,\dots,N\},\mu} := A_{\{1,\dots,N\},\mu,\nu}$ und $B_{\{1,\dots,N\},\mu} := B_{\{1,\dots,N\},\mu,\nu}$ nicht von ν abhängen.

1) Behauptung: Sei $m \in \{1, \ldots, N\}$. Dann existiert ein $\mu(m) \in (0, 1]$, sodass $B_{\mathcal{M},\tilde{\mu},\nu} \subset A_{\mathcal{M},\tilde{\mu},\nu}$ für alle $\mathcal{M} \subset \{1, \ldots, N\}$ mit $|\mathcal{M}| = m, \tilde{\mu} \in (0, \mu(m)]$ und $\nu \in (0, 1]$ gilt.

Die Regel von de l'Hospital liefert $\lim_{x \searrow 0} x^{1-\gamma_0} \log \left(\log \left(\frac{1}{x} \right) \right) = 0$. Also existiert ein $x_0 \in (0, e^{-1})$ mit $x \log \left(\log \left(\frac{1}{x} \right) \right) \leq x^{\gamma_0}$ für alle $x \in (0, x_0)$. Definiere $\mu(m) := \sqrt[m]{x_0}$.

Seien nun $\mathcal{M} \subset \{1, \ldots, N\}$ mit $|\mathcal{M}| = m$, $\tilde{\mu} \in (0, \mu(m)]$ und $\nu \in (0, 1]$. Dann gilt

$$0 < |\mathcal{M}| \left[G\left(\prod_{n \in \mathcal{M}} u^n, \prod_{n \in \mathcal{M}^C} u^n\right) \right]^2 \le 2N \left[\prod_{n=1}^N u^n\right]^{\gamma_0}$$

für alle $u = (u^1, \ldots, u^N) \in (0, 1]^N$ mit $u^n \in (0, \tilde{\mu})$ für $n \in \mathcal{M}$ und $u^n \in [\nu, 1]$ für $n \in \mathcal{M}^C$. Hieraus folgt $B_{\mathcal{M}, \tilde{\mu}, \nu} \subset A_{\mathcal{M}, \tilde{\mu}, \nu}$.

In den nun folgenden Schritten 2) und 3) wird durch eine Rückwärtsrekursion eine endliche, monoton fallende Folge $(\delta_N \dots, \delta_1)$ konstruiert.

2) Rekursionsanfang: Konstruktion von δ_N

Es wird ein $\delta_N \in (0, 1]$ konstruiert mit $\mathbb{P}(A_{\{1, \dots, N\}, \delta_N}) > 1 - \varepsilon_1$.

Für $0 < \mu_1 \leq \mu_2 \leq 1$ gilt $A_{\{1,\dots,N\},\mu_1} \supset A_{\{1,\dots,N\},\mu_2}$ und $B_{\{1,\dots,N\},\mu_1} \supset B_{\{1,\dots,N\},\mu_2}$. Ferner folgt aus Satz 95 $\lim_{\tilde{\mu} \searrow 0} \mathbb{P}\left(B_{\{1,\dots,N\},\tilde{\mu}}\right) = 1$. Somit existiert ein $\mu_N^1 \in (0,1]$ mit $\mathbb{P}\left(B_{\{1,\dots,N\},\tilde{\mu}}\right) > 1 - \varepsilon_1$ für alle $\tilde{\mu} \in (0,\mu_N^1]$. Aus Schritt 1) des vorliegenden Beweises folgt, dass es ein $\mu_N^2 \in (0,1]$ gibt mit $B_{\{1,\dots,N\},\tilde{\mu}} \subset A_{\{1,\dots,N\},\tilde{\mu}}$ für alle $\tilde{\mu} \in (0,\mu_N^2]$. Für $\delta_N := \mu_N^1 \land \mu_N^2$ gilt damit

 $\mathbb{P}\left(A_{\{1,\ldots,N\},\delta_N}\right) \geq \mathbb{P}\left(B_{\{1,\ldots,N\},\delta_N}\right) > 1 - \varepsilon_1.$

3) Rekursionsschritt für $m \in \{1, ..., N-1\}$: Konstruktion von δ_m $\delta_{m+1} \in (0, 1]$ sei bereits konstruiert. Es wird ein $\delta_m \in (0, \delta_{m+1}]$ konstruiert mit $\mathbb{P}(A_{\mathcal{M}, \delta_m, \delta_{m+1}}) > 1 - \varepsilon_1$ für alle $\mathcal{M} \subset \{1, ..., N\}$ mit $|\mathcal{M}| = m$.

Sei $\mathcal{M} \subset \{1, \ldots, N\}$ mit $|\mathcal{M}| = m$. Für $0 < \mu_1 \leq \mu_2 \leq 1$ gilt $A_{\mathcal{M},\mu_1,\delta_{m+1}} \supset A_{\mathcal{M},\mu_2,\delta_{m+1}}$ und $B_{\mathcal{M},\mu_1,\delta_{m+1}} \supset B_{\mathcal{M},\mu_2,\delta_{m+1}}$. Unter Verwendung von Satz 98 erhält man $\lim_{\tilde{\mu} \searrow 0} \mathbb{P}\left(B_{\mathcal{M},\tilde{\mu},\delta_{m+1}}\right) = 1$. Also existiert ein $\mu_m^0(\mathcal{M}) \in (0,1]$ mit $\mathbb{P}\left(B_{\mathcal{M},\tilde{\mu},\delta_{m+1}}\right) > 1 - \varepsilon_1$ für alle $\tilde{\mu} \in (0,\mu_m^0(\mathcal{M})]$. Definiere nun

$$\mu_m^1 := \min_{\substack{\mathcal{M} \subset \{1, \dots, N\}\\ mit \mid \mathcal{M} \mid = m}} \mu_m^0(\mathcal{M}) \in (0, 1].$$

Aus Schritt 1) des vorliegenden Beweises folgt, dass es ein $\mu_m^2 \in (0, 1]$ gibt mit $B_{\mathcal{M},\tilde{\mu},\delta_{m+1}} \subset A_{\mathcal{M},\tilde{\mu},\delta_{m+1}}$ für alle $\tilde{\mu} \in (0, \mu_m^2]$ und $\mathcal{M} \subset \{1, \ldots, N\}$ mit $|\mathcal{M}| = m$.

Für $\delta_m := \mu_m^1 \wedge \mu_m^2 \wedge \delta_{m+1}$ gilt damit

$$\mathbb{P}\left(A_{\mathcal{M},\delta_{m},\delta_{m+1}}\right) \geq \mathbb{P}\left(B_{\mathcal{M},\delta_{m},\delta_{m+1}}\right) > 1 - \varepsilon_{1}$$

für alle $\mathcal{M} \subset \{1, \ldots, N\}$ mit $|\mathcal{M}| = m$.

4) Behauptung: Für

$$R := (0, \delta_N)^N \cup \bigcup_{m=1}^{N-1} \bigcup_{\substack{\mathcal{M} \subset \{1, \dots, N\}\\mit \mid \mathcal{M} \mid = m}} \left[\prod_{n \in \mathcal{M}} (0, \delta_m) \times \prod_{n \in \mathcal{M}^C} [\delta_{m+1}, 1] \right].$$

gilt $(0,1]^N \setminus [\delta_1,1]^N \subset R.$

Sei $u_0 = (u_0^1, \ldots, u_0^N) \in (0, 1]^N \setminus [\delta_1, 1]^N$. Ohne Beschränkung der Allgemeinheit kann angenommen werden, dass $u_0^1 \leq \ldots \leq u_0^N$ gilt. Hieraus folgt $u_0^1 \in (0, \delta_1)$.

Für $n \in \{1, \ldots, N\}$ seien $\mathcal{M}_n := \{\tilde{n} \in \{1, \ldots, N\} | u_0^{\tilde{n}} \in (0, \delta_n)\}$ und $m_n := |\mathcal{M}_n|$. Aus der Monotonie von u_0 folgt $\mathcal{M}_n = \{1, \ldots, m_n\}$ für $n \in \{1, \ldots, N\}$ und damit $1 \le m_1 \le \ldots \le m_N \le N$.

Es können nun die folgenden beiden Fälle auftreten.

1. Fall: $m_N = N$ Dann gilt $u_0 \in (0, \delta_N)^N \subset R$. 2. Fall: $m_N < N$

Sei $n_0 \in \{0, \ldots, N-1\}$ der erste Index mit $m_{n_0+1} < n_0 + 1$. Aus $u_0^1 \in (0, \delta_1)$ folgt $m_1 \geq 1$ und damit $n_0 \geq 1$. Nach Konstruktion von n_0 gilt ferner $n_0 \leq m_{n_0} \leq m_{n_0+1} < n_0 + 1$, also $m_{n_0} = m_{n_0+1} = n_0$. Hieraus folgt $u_0^n \in (0, \delta_{n_0})$ für $n \in \{1, \ldots, n_0\}$ und $u_0^{n_0+1} \notin (0, \delta_{n_0+1})$. Unter Verwendung der Monotonie von u_0 erhält man $u_0^n \in [\delta_{n_0+1}, 1]$ für $n \in \{n_0+1, \ldots, N\}$ und damit $u_0 \in \prod_{n \in \{1, \ldots, n_0\}} (0, \delta_{n_0}) \times \prod_{n \in \{n_0+1, \ldots, N\}} [\delta_{n_0+1}, 1] \subset R$.

5) Behauptung: Es existiert eine Nullmenge $\Omega_0(\varepsilon_0, \gamma_0) \subset \Omega$, sodass es für jedes $\omega \in (\Omega_0(\varepsilon_0, \gamma_0))^C$ ein $\delta(\varepsilon_0, \gamma_0, \omega) > 0$ gibt mit

$$\sup_{\tilde{u}\in(0,1]^N\setminus[\delta(\varepsilon_0,\gamma_0,\omega),1]^N}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{\gamma_0}}<1+\varepsilon_0.$$

Definiere

$$C := \left\{ \sup_{\tilde{u} \in (0,1]^N \setminus [\delta_1,1]^N} \frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)} \right)^2}{2N \left[\prod_{n=1}^N \tilde{u}^n \right]^{\gamma_0}} < 1 + \varepsilon_0 \right\}$$

und

$$A := A_{\mathcal{M}, \delta_N} \cap \bigcap_{m=1}^{N-1} \bigcap_{\substack{\mathcal{M} \subset \{1, \dots, N\}\\ mit \ |\mathcal{M}|=m}} A_{\mathcal{M}, \delta_m, \delta_{m+1}}.$$

Sei $\omega \in A.$ Wegen $(0,1]^N \backslash [\delta_1,1]^N \subset R$ erhält man

$$\sup_{\tilde{u}\in(0,1]^N\setminus[\delta_1,1]^N}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{\gamma_0}}\leq \sup_{\tilde{u}\in R}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{\gamma_0}}<1+\varepsilon_0,$$

also $\omega \in C$ und damit $A \subset C$. Hieraus folgt

$$\mathbb{P}(C) \ge 1 - \left[\mathbb{P}\left(A_{\mathcal{M},\delta_{N}}^{C}\right) + \sum_{m=1}^{N-1} \sum_{\substack{\mathcal{M} \subset \{1,\dots,N\}\\mit \ |\mathcal{M}|=m}} \mathbb{P}\left(A_{\mathcal{M},\delta_{m},\delta_{m+1}}^{C}\right) \right] \ge 1 - 2^{N} \varepsilon_{1}.$$

Weil

$$C \subset \left\{ \text{Es existiert ein } \delta > 0 \text{ mit } \sup_{\tilde{u} \in (0,1]^N \setminus [\delta,1]^N} \frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}\right)^2}{2N \left[\prod_{n=1}^N \tilde{u}^n\right]^{\gamma_0}} < 1 + \varepsilon_0 \right\}$$

und weil $\varepsilon_1 > 0$ beliebig war, existiert eine Nullmenge $\Omega_0(\varepsilon_0, \gamma_0) \subset \Omega$ mit den oben genannten Eigenschaften.

6) Behauptung: Es existiert eine Nullmenge $\Omega_0 \subset \Omega$ mit der im Satz formulierten Eigenschaft.

Aus dem bisher Bewiesenen folgt, dass für jedes $k \in \mathbb{N}$ mit $k \geq 2$ eine Nullmenge $\Omega_0\left(\frac{1}{k}, 1 - \frac{1}{k}\right) \subset \Omega$ existiert, sodass es für jedes $\omega \in \left(\Omega_0\left(\frac{1}{k}, 1 - \frac{1}{k}\right)\right)^C$ ein $\delta\left(\frac{1}{k}, 1 - \frac{1}{k}, \omega\right) > 0$ gibt mit

$$\sup_{\tilde{u}\in(0,1]^N\setminus\left[\delta\left(\frac{1}{k},1-\frac{1}{k},\omega\right),1\right]^N}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\ldots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{1-\frac{1}{k}}}<1+\frac{1}{k}$$

Die Menge

$$\Omega_0 := \bigcup_{k \in \mathbb{N}} \Omega_0 \left(\frac{1}{k}, 1 - \frac{1}{k} \right)$$

ist eine Nullmenge, die die im Satz behauptete Eigenschaft besitzt.

Korollar 101. Für $\gamma \in \mathbb{R}_+$ sei $\mathbb{V}(\gamma) = (\mathbb{V}_u(\gamma))_{u \in [0,1]^N}$ der in Gleichung (30) definierte Prozess, also

$$\mathbb{V}_{u}(\gamma) := \begin{cases} \frac{\left(\mathbb{W}_{\left(u^{1},\ldots,u^{N}\right)}\right)^{2}}{\left[\prod\limits_{n=1}^{N}u^{n}\right]^{\gamma}} & f\ddot{u}r \ u \in (0,1]^{N}\\ 0 & f\ddot{u}r \ u \in [0,1]^{N} \backslash (0,1]^{N} \end{cases}$$

Es existiert eine Nullmenge $\Omega_1 \subset \Omega$, sodass auf Ω_1^C für jedes $\gamma \in [0, 1)$ die N-parametrigen Pfade von $\mathbb{V}(\gamma)$ stetig sind.

Beweis: Nach Definition des Brownschen Blatts existiert eine Nullmenge $N \subset \Omega$, sodass auf N^C alle Pfade von W stetig sind. Ferner bezeichne Ω_0 die Nullmenge aus Satz 100. Im Folgenden wird gezeigt, dass Nullmenge die $\Omega_1 := N \cup \Omega_0$ die gewünschte Eigenschaft besitzt.

Sei $\omega \in \Omega_1^C$. Der Nachweis der Stetigkeit von $[0,1]^N \ni u \mapsto V_u(\gamma)(\omega)$ für $\gamma = 0$ ist trivial. Sei nun $\gamma \in (0,1)$. Aus der Stetigkeit von $\mathbb{W}(\omega)$ folgt, dass $[0,1]^N \ni u \mapsto V_u(\gamma)(\omega)$ auf $(0,1]^N$ stetig ist. Sei nun $u_0 \in [0,1]^N \setminus (0,1]^N$. Nach Satz 100 existiert für $\mu := \frac{1}{2}(\gamma + 1) \in (0,1)$ und $\varepsilon := 1$ ein $\delta(\varepsilon, \mu, \omega)$ mit

$$\sup_{\tilde{u}\in(0,1]^N\setminus[\delta(\varepsilon,\mu,\omega),1]^N}\frac{\left(\mathbb{W}_{(\tilde{u}^1,\dots,\tilde{u}^N)}(\omega)\right)^2}{2N\left[\prod_{n=1}^N\tilde{u}^n\right]^{\mu}}<1+\varepsilon.$$

Zusammen mit $\mu - \gamma > 0$ folgt hieraus

$$0 \leq \lim_{\substack{\tilde{u} \to u_0 \\ mit \ \tilde{u} \in [0,1]^N}} \mathbb{V}_{\tilde{u}}(\gamma)(\omega)$$

$$\leq 2N(1+\varepsilon) \lim_{\substack{\tilde{u} \to u_0 \\ mit \ \tilde{u} \in [0,1]^N}} \left[\prod_{n=1}^N \tilde{u}^n\right]^{\mu-\gamma}$$

$$= 0.$$

Also ist $[0,1]^N \ni u \mapsto \mathbb{V}_u(\gamma)(\omega)$ auch in u_0 stetig. Hieraus folgt die Behauptung.

A.1.2 Der alternative Beweis für $\mathbb{P}\left(D^2_{\mathbb{B},\mu} < \infty\right) = 1$

Lemma 102. Seien $N \in \mathbb{N}$ mit $N \geq 2$ und $\mu \in [0, 2)$. Dann gilt

$$\int_{(0,1)^N} \frac{1}{\left[1 - \prod_{n=1}^N \tilde{u}^n\right]^{\mu}} \,\mathrm{d}\tilde{u} < \infty.$$

Beweis: Weil $\left[1 - \prod_{n=1}^{N} u^n\right]^{-\mu} \leq \left[1 - u^1 u^2\right]^{-\mu}$ für alle $u \in (0, 1)^N$ gilt, genügt es, den Fall N = 2 zu betrachten. Zur besseren Lesbarkeit wird (x, y) statt (u^1, u^2) geschrieben.

Für $\mu_1, \mu_2 \in [0, 2)$ mit $\mu_1 \leq \mu_2$ gilt $[1 - xy]^{-\mu_1} \leq [1 - xy]^{-\mu_2}$ für alle $x, y \in (0, 1)$. Somit genügt der Nachweis für den Fall $\mu \in (1, 2)$.

Die Substitution $x = \sin(t)$ liefert

$$\int_{0}^{1} \int_{0}^{1} [1 - xy]^{-\mu} dx dy = \int_{0}^{1} \int_{0}^{\frac{\pi}{2}} \cos(t) [1 - \sin(t)y]^{-\mu} dt dy$$
$$= \frac{1}{\mu - 1} \int_{0}^{1} \frac{1 - [1 - y]^{\mu - 1}}{y} \frac{1}{[1 - y]^{\mu - 1}} dy$$
$$\leq \frac{1}{\mu - 1} \int_{0}^{1} \frac{1}{[1 - y]^{\mu - 1}} dy$$
$$< \infty.$$

Hierbei wurde $0 < 1 - [1 - y]^{\mu - 1} < y$ für alle $y \in (0, 1)$ verwendet.

Satz 103. Set $n \in \mathbb{N}$ mit $N \geq 2$ und $\mu \in [0, 2)$.

Dann gilt

$$\int_{(0,1)^N} \frac{\left(\mathbb{W}_{\tilde{u}}\right)^2}{\left[\left(\prod_{n=1}^N \tilde{u}^n\right) \left(1 - \prod_{n=1}^N \tilde{u}^n\right)\right]^{\mu}} \,\mathrm{d}\tilde{u} < \infty$$

fast sicher.

Beweis: Seien Ω_1 die Nullmenge aus Korollar 101, $\omega \in \Omega_1^C$ und $\gamma \in (0 \lor \mu - 1, 1)$. Dann existiert ein $K \in \mathbb{R}_+$ mit

$$0 \leq \frac{\left(\mathbb{W}_{u}(\omega)\right)^{2}}{\left[\prod_{n=1}^{N} \tilde{u}^{n}\right]^{\gamma}} \leq K$$

für alle $u \in (0,1)^N.$ Unter Verwendung von $\mu-\gamma < 1$ sowie Lemma 102 berechnet man hiermit

$$\int_{(0,1)^{N}} \frac{\left(\mathbb{W}_{\tilde{u}}(\omega)\right)^{2}}{\left[\left(\prod_{n=1}^{N} \tilde{u}^{n}\right)\left(1-\prod_{n=1}^{N} \tilde{u}^{n}\right)\right]^{\mu}} d\tilde{u}$$

$$\leq K \int_{(0,1)^{N}} \frac{1}{\left[\prod_{n=1}^{N} \tilde{u}^{n}\right]^{\mu-\gamma} \left[1-\prod_{n=1}^{N} \tilde{u}^{n}\right]^{\mu}} d\tilde{u}$$

$$\leq K 2^{\mu} \int_{(0,1)^{N} \setminus \left(\frac{1}{2},1\right)^{N}} \frac{1}{\left[\prod_{n=1}^{N} \tilde{u}^{n}\right]^{\mu-\gamma}} d\tilde{u} + K 2^{N(\mu-\gamma)} \int_{\left(\frac{1}{2},1\right)^{N}} \frac{1}{\left[1-\prod_{n=1}^{N} \tilde{u}^{n}\right]^{\mu}} d\tilde{u}$$

$$< \infty$$

und erhält die Behauptung.

Bemerkung 104. Sei $\mu \in [1, 2)$. Interessanterweise hängt es von N ab, ob die Zufallsvariable

$$\int_{(0,1)^N} \frac{\left(\mathbb{W}_{\tilde{u}}\right)^2}{\left[\left(\prod_{n=1}^N \tilde{u}^n\right) \left(1 - \prod_{n=1}^N \tilde{u}^n\right)\right]^{\mu}} \,\mathrm{d}\tilde{u}$$

fast sicher endlich ist oder fast sicher unendlich ist. FürN=1lässt sich leicht zeigen, dass sie fast sicher unendlich ist. Für $N\geq 2$ besagt Satz 103,
dass sie fast sicher endlich ist. Der Grund für diese Tatsache ist nicht etwa das Verhalten von



auf $U \cap (0,1)^N$, wobe
i $U \subset [0,1]^N$ eine Umgebung des "N-dimensionalen linken, unteren Rands" $[0,1]^N \setminus (0,1]^N$ des Parameterraums sei. Stattdessen ist der Grund von deterministischer Natur. Es ist die Geschwindigkeit, mit der $1 - \prod_{n=1}^N u^n$ für $(u^1, \ldots, u^N) \nearrow (1, \ldots, 1)$ gegen 0 konvergiert. Für $N \ge 2$ findet diese Konvergenz hinreichend langsam statt, so dass das Integral zumindest für alle $\omega \in \Omega_1^C$ endlich ist. Dabei ist Ω_1 die Nullmenge aus Korollar 101.

Satz 105. Seien $\mathbb{B} = (\mathbb{B}_u)_{u \in [0,1]^N}$ eine N-parametrige Brownsche Brücke und $\mu \in [0,2)$.

Dann gilt

$$\int_{(0,1)^N} \frac{(\mathbb{B}_{\tilde{u}})^2}{\left[\left(\prod_{n=1}^N \tilde{u}^n\right) \left(1 - \prod_{n=1}^N \tilde{u}^n\right)\right]^{\mu}} \,\mathrm{d}\tilde{u} < \infty$$
(39)

fast sicher.

Beweis: Anderson und Darling (1952) zeigen unter Verwendung des Gesetzes des iterierten Logarithmus für den einparametrigen Wiener-Prozess, dass im Fall N = 1 und $\mu = 1$ das Integral in Gleichung (39) fast sicher endlich ist. Ihre Argumentation lässt sich ohne nennenswerte Schwierigkeiten auf den Fall $\mu \in [0, 2)$ übertragen. Im Folgenden wird der Fall $N \ge 2$ betrachtet.

Sei Z eine standardnormalverteilte Zufallsvariable, die stochastisch unabhängig von \mathbb{B} ist. Es lässt sich leicht zeigen, dass mit dem Prozess $\mathbb{W}^{\mathbb{B},Z} = (\mathbb{W}_{u}^{\mathbb{B},Z})_{u \in [0,1]^{N}}$,

$$\mathbb{W}_{u}^{\mathbb{B},Z} := \mathbb{B}_{u} + \left(\prod_{n=1}^{N} u^{n}\right) Z,$$

ein N-parametriges Brownsches Blatt vorliegt und dass

$$\mathbb{B}_{u} = \mathbb{W}_{u}^{\mathbb{B},Z} - \left(\prod_{n=1}^{N} u^{n}\right) \mathbb{W}_{(1,\dots,1)}^{\mathbb{B},Z}$$

für alle $u \in [0,1]^N$ gilt.

Durch eine einfache Rechnung erhält man

$$\int_{(0,1)^N} \left| \frac{\left(\prod_{n=1}^N \tilde{u}^n\right) \mathbb{W}_{(1,\dots,1)}^{\mathbb{B},Z} \mathbb{W}_{\tilde{u}}^{\mathbb{B},Z}}{\left[\left(\prod_{n=1}^N \tilde{u}^n\right) \left(1 - \prod_{n=1}^N \tilde{u}^n\right) \right]^{\mu}} \right| d\tilde{u} < \infty$$

und

$$\int_{(0,1)^N} \frac{\left(\prod_{n=1}^N \tilde{u}^n\right)^2 \left(\mathbb{W}_{(1,\dots,1)}^{\mathbb{B},Z}\right)^2}{\left[\left(\prod_{n=1}^N \tilde{u}^n\right) \left(1 - \prod_{n=1}^N \tilde{u}^n\right)\right]^{\mu}} \, \mathrm{d}\tilde{u} < \infty$$

jeweils fast sicher. Aus Satz 103 und

$$\left(\mathbb{B}_{u}\right)^{2} = \left(\mathbb{W}_{u}^{\mathbb{B},Z}\right)^{2} - 2\left(\prod_{n=1}^{N} u^{n}\right)\mathbb{W}_{(1,\dots,1)}^{\mathbb{B},Z}\mathbb{W}_{u}^{\mathbb{B},Z} + \left(\prod_{n=1}^{N} u^{n}\right)^{2}\left(\mathbb{W}_{(1,\dots,1)}^{\mathbb{B},Z}\right)^{2}$$

für $u \in [0,1]^N$ folgt die Behauptung.

Literatur

- T. W. Anderson und D. A. Darling. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, 23(2): 193–212, 1952.
- T. W. Anderson und D. A. Darling. A Test of Goodness of Fit. Journal of the American Statistical Association, 49(268): 765–769, 1954.
- D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge studies in advanced mathematics. Cambridge University Press, 2. Auflage, 2009.
- H. Bauer. Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie, Band 1. De Gruyter, 1968.
- H. Bauer. Wahrscheinlichkeitstheorie. De Gruyter, 5. Auflage, 2002.
- M. Beinhofer, E. Eberlein, A. Janssen und M. Polley. Correlations in Lévy interest rate models. *Quantitative Finance*, 11(9): 1315–1327, 2011.
- P. Billingsley. *Convergence of Probability Measures*. Wiley series in probability and mathematical statistics. Wiley Verlag, 1968.
- F. Black und M. S. Scholes. The Pricing of Options and Corporate Liabilities. Journal of Political Economy, 81(3): 637–54, 1973.
- S. Blinnikov und R. Moessner. Expansions for nearly Gaussian distributions. Astronomy and Astrophysics Supplement, 130: 193–205, 1998.
- D. M. Chibisov. An Investigation of the Asymptotic Power of the Tests of Fit. Theory of Probability and Its Applications, 10(3): 421–437, 1965.
- T. Duong und M. L. Hazelton. Cross-validation Bandwidth Matrices for Multivariate Kernel Density Estimation. Scandinavian Journal of Statistics, 32(3): 485–506, 2005.
- E. Eberlein, J. Jacod und S. Raible. Lévy term structure models: Noarbitrage and completeness. *Finance and Stochastics*, 9(1): 67–88, 2005.
- E. Eberlein und W. Kluge. Exact pricing formulae for caps and swaptions in a Lévy term structure model. *Journal of Computational Finance*, 9(2): 99–125, 2006.
- E. Eberlein und W. Kluge. Calibration of Lévy term structure models. In M. C. Fu, R. A. Jarrow, J.-Y. J. Yen und R. J. Elliott, Herausgeber, Advances in Mathematical Finance. Birkhäuser Verlag, 2007.

- E. Eberlein und S. Raible. Term structure models driven by general Lévy processes. *Mathematical Finance*, 9(1): 31–53, 1999.
- E. Eberlein und F. Ozkan. The defaultable Lévy term structure: Ratings and restructuring. *Mathematical Finance*, 13(2): 277–300, 2003.
- J. Elstrodt. Maß- und Integrationstheorie. Springer Verlag, 4. Auflage, 2005.
- C. Genest, W. Huang und J.-M. Dufour. A regularized goodness-of-fit test for copulas. *Journal de la Société Française de Statistique*, 154(1): 64–77, 2013.
- K. Glau. Feynman-Kac-Darstellungen zur Optionspreisberechnung in Lévy-Modellen. Dissertation, Universität Freiburg, Fakultät für Mathematik und Physik, 2010.
- P. Hall. The Bootstrap and Edgeworth Expansion. Springer Verlag, 1992.
- E. A. von Hammerstein. *Generalized hyperbolic distributions: Theory and applications to CDO pricing*. Dissertation, Universität Freiburg, Fakultät für Mathematik und Physik, 2010.
- F. M. Hartogs. Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten. *Mathematische Annalen*, 62(1): 1–88, 1906.
- D. Heath, R. Jarrow und A. J. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1): 77–105, 1992.
- H. Heuser. Lehrbuch der Analysis Teil 1. Vieweg+Teubner Verlag, 10. Auflage, 1992.
- W. Kluge. Time-inhomogeneous Lévy processes in interest rate and credit risk models. Dissertation, Universität Freiburg, Fakultät für Mathematik und Physik, 2005.
- R. C. Merton. Theory of Rational Option Pricing. Bell Journal of Economics, 4(1): 141–183, 1973.
- S.R. Paranjape und C. Park. Laws of iterated logarithm of multiparameter wiener processes. *Journal of Multivariate Analysis*, 3(1): 132 136, 1973.

- M. Polley. Parameterschätzung im Lévy-Forward-Rate-Modell auf der Basis von empirischen Korrelationen. Diplomarbeit, Albert-Ludwigs-Universität Freiburg, 2010.
- P. E. Protter. Stochastic Integration and Differential Equations. Springer Verlag, 2005.
- M. Rosenblatt. Remarks on a Multivariate Transformation. The Annals of Mathematical Statistics, 23(3): 470–472, 1952.
- V. Scheidemann. Introduction to Complex Analysis in Several Variables. Birkhäuser Verlag, 2005.
- J. Segers. Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, 18(3): 764–782, 2012.
- L. E. O. Svensson. Estimating and interpreting forward interest rates: Sweden 1992 1994. Working Paper 4871, National Bureau of Economic Research, 1994.
- A. W. van der Vaart und J. Wellner. Weak Convergence and Empirical Processes: With Applications to Statistics. Springer Verlag, 1996.
- D. C. Wood. The Computation of Polylogarithms. Technical Report 15-92, University of Kent, Computing Laboratory, 1992.
- G. J. Zimmerman. Some Sample Function Properties of the Two-parameter Gaussian Process. *The Annals of Mathematical Statistics*, 43(4): 1235– 1246, 1972.