Portfolio Analysis with Multivariate Normal Tempered Stable Processes and Distributions

Zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für Wirtschaftswissenschaften
des Karlsruher Institut für Technologie (KIT)
genehmigte

Dissertation

von

Dipl.-Wi.-Ing. Dirk Krause

Tag der mündlichen Prüfung: 06.12.2011
Referent: ..... Prof. Dr. S. T. Rachev
Korreferent: ..... Prof. Dr. K.-H. Waldmann

Karlsruhe, Dezember 2011

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Historical Background and Motivation

As early as 1900, Louis Bachelier [Bach00] already conceived a stochastic concept, which should become the well-known Brownian motion. It was later elaborated with mathematical rigor by Einstein and Wiener during the first half of the past century. This important class of stochastic processes in continuous time together with the corresponding normal distribution established itself in the following years as a cornerstone of the most successful models in several different areas of modern quantitative finance. These include, for instance, Markowitz' pioneering model for portfolio selection and asset allocation [Mark52], the famous Capital Asset Pricing Model (CAPM) of Sharpe [Shar64], Treynor, Lintner and Mossin, Ross' Arbitrage Pricing Theory (APT) [Ross76]. Black and Scholes [BlSc73] as well as Merton [Mert73] in their seminal works on option pricing, draw upon the geometric Brownian motion as a central assumption about stochastic stock price dynamics. With the Black-Scholes formula laying the foundation for the whole new field of financial engineering emerging from it, geometric Brownian motion as a basic building block spreads to related fields of research as different as term structure modeling or credit risk.

However, already in the early 1960ies a countermovement under Mandelbrot and his student Fama against the prevailing normal distribution hypothesis started to form. In their

investigations, repeatedly occurring violent fluctuations of observable prices for several different tradable goods, such as cotton, led them to the conclusion of strictly rejecting the Normal assumption for these price processes[1]. Instead, they found strong evidence for these distributions obeying a power law on the one hand and a present time-scaling property of the price processes on the other. Besides, with this latter characterizing property it was possible to convey the term *fractal* to the world of stochastic processes. Most unfortunately, their substantial work gained only little recognition if not even experienced harsh criticism over the following years.

Usually, financial data is present in terms of continuously compounded log-returns $r_t = \ln(P_t) - \ln(P_{t-1})$ or the respective annualized log-returns $r_t = \frac{1}{\Delta t}(\ln(P_t) - \ln(P_{t-\Delta t}))$[2], derived from corresponding price series P_t. Today, it is a widely recognized fact that financial data exhibit certain features opposing normal distributions or Brownian motion. In the following, some of these stylized facts of financial markets are listed, where one has to clearly distinguish between the cross-sectional and longitudinal perspective. Cross-sections are connected to distributions whereas the longitudinal viewpoint corresponds to process trajectories. Among these stylized facts are to be mentioned, in particular with regard to the cross-sectional perspective or corresponding distributional properties,

- Heavy-Tailedness or Excess Kurtosis: Empirical return distributions have more pronounced tails in addition to a more peaked center compared to the Gaussian assumption.

- Asymmetry or Skewness: Whilst normal distributions are always symmetric around their mean, observable returns mostly exhibit asymmetry in favor of large negative return deviations.

Whereas concerning the longitudinal view, one observes,

[1] See e.g. Mandelbrot [Mand63], [Mand67], Mandelbrot and Taylor [MaTa67] as well as Fama and Mandelbrot [Fama63].

[2] Its advantages and disadvantages in comparison to the simply compounded return rates $r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ are discussed in [Broo08, p. 6 ff.].

- Volatility Clustering: In reality, large price movements in one period tend to be followed by equally large price movements in the next period. In the same manner, after calm periods one is more likely to observe only moderate price change in the following time step. Therefore, in this sense volatility is clustered on the time scale. The ARCH[3] and GARCH[4] time series framework has been designed in order to incorporate and represent precisely these effects. This explicit consideration of non-constant volatility is able to account for a certain amount of excess kurtosis present in observed empirical asset returns. However, there remains a part which cannot be attributed to the effect of clustered volatility when using normal distributions for standardized innovations in this time series model. To compensate for this discrepancy, process innovations itself have to possess heavy tail properties.

- Price Jumps: Financial tick data is available on a rather high-frequency scale today, but is still collected in discrete time. This makes it difficult to tell whether large and sudden moves in asset prices correspond to actual jumps in the chosen continuous-time modeling or are merely due to discrete sampling. Nevertheless, it is proven evidently that empirical observations cannot be implied by diffusion processes with continuous trajectories alone but necessarily require the incorporation of jumps in the employed underlying models.

To develop the theory adequate to his observed phenomena and conceived principles, Mandelbrot recalls the work of the French mathematician Paul Lévy and others, who laid the foundations for the class of stochastic processes in continuous time named after him, compiled in his classic book [Lé48]. Mandelbrot realized that his field of study directly corresponds to the subset of *self-similar* non-Gaussian Lévy processes. Like the Brownian motion, which itself is a self-similar Lévy processes with Gaussian laws however, this class maintains the two central assumptions of independent and identically distributed (i.i.d) increments and the time-scaling property. But at the same time, it remedies the drawbacks with regard to unrealistic symmetry and mesokurtosis. Both these properties

[3] *Autoregressive Conditional Heteroscedasticity*, see Engle [Engl82].
[4] *Generalized Autoregressive Conditional Heteroscedasticity*, see Bollerslev [Boll86].

appear to be particularly desirable as they reflect the *Generalized Central Limit Theorem (GCLT)*. Considered separately, the i.i.d. property of increments, which is common to all general Lévy processes, is consistent to the *efficient market hypothesis (EMH)*[5]. The EMH states that the uniform and independent inflow of new information into the market is immediately incorporated into quoted prices. Finally, this class seemingly being adequate for capturing the mentioned observable phenomena, later came to be known as *Stable Paretian* or α-stable models. Since then, the approach experienced extensive developments and applications by researchers like Zolotarev, Samorodnitsky and Rachev, only to mention a few.

However, their advantage of improved flexibility and adaptability came at the price of some practical difficulties. The major problems when dealing with α-stable distributions are twofold. Firstly, there is in general no analytical representation of the probability density function available[6]. Secondly, depending on the value of the tail index $\alpha \in (0, 2)$, distributional moments become infinite. More precisely, moments of fractional order d only exist for $d < \alpha$. These facts impede for example estimation and computation of practical probabilities on the one hand, while on the other infinite variance and possibly non-existent expectation is evoking scepticism among practitioners.

Therefore, the next advance was aimed at eradicating the problematic behavior while preserving the already achieved advantages by modifying the extreme tails of α-stable distributions. One rather indirect approach for this is to reduce the intensity of large jumps by multiplying the Lévy density function of an α-stable process with a decreasing tempering function of sufficiently fast decay. By this operation, the result is still a Lévy process with associated infinitely divisible distributions. It allows for possible skewness while the tail behavior have been changed from heavy to semi-heavy, characterized by exponential decay with excess kurtosis instead of polynomial decay. This in turn ensures by now the existence of conventional moments for arbitrary orders. Moreover, possibly even exponential moments exist for a range of orders. This existence is particularly required for various different applications such as option pricing. Of course, this approach was not

[5] See Fama [Fama70].

[6] The only exceptions are $\alpha = 1$ and $\alpha = \frac{1}{2}$.

able to generate closed-form probability densities, which still requires a cumbersome handling by means of the characteristic function. In addition, the property of self-similarity corresponding to time-scaling vanishes. At last, by finite first and second moment and independent increments of the Lévy process, these models fall under the sphere of influence of the conventional central limit theorem. Hence, the process distributions lose their asymmetry and excess kurtosis over time and weakly converge to a Gaussian one. This is not necessarily a drawback but can rather be seen as an advantage, as return distributions tend to become more tame when measured over longer periods compared to very erratic behavior of high-frequency return data.

In summary, the advantages outweigh the difficulties, in particular when compared to the preceding α-stable model. Moreover, by their less extreme nature though not ignoring the stylized facts of the distributions under study, the present variety of one-dimensional tempered stable models[7] are additionally able to achieve a better statistical fit than the α-stable models. This makes them a suitable and promising candidate for multivariate extensions in the area of non-Gaussian modeling.

Apart from only a few exceptions, most of the areas in quantitative finance are concerned with the multidimensional modeling of price dynamics and risk, such as for example portfolio optimization or the valuation of derivatives on multiple assets. This drives the research into both realistic but yet reasonably manageable models for a high number of involved dimensions. One of the new main challenges in a multivariate perspective lies in the accurate capturing of present empirical dependence structures. The realistic description of dependence structures prove indispensable when e.g. assessing the potential of beneficial diversification in portfolio optimization or joint default events in credit risk. Although one general formulation of multivariate tempered stable models has already been introduced by Rosiński [Rosi07], its use is rather confined to merely theoretical aspects. This is mainly due to the lack of parsimonious parameterizations of the spectral measure and the corresponding tempering function. Approaches which are based on the explicit tempering of a multivariate α-stable Lévy density prove to be ineligible for this reason.

[7] These entail for example process classes like the CTS, NTS, MTS, KR or RDTS where more detailed explanations are given in the following chapters.

Only for one of the univariate classes of tempered stable models mentioned above, namely the normal tempered stable model, besides the definition by tempering and tilting a stable Lévy measure, an alternative way for construction via the method of Brownian subordination is available. Brownian subordination embodies a highly versatile tool for creating stochastic processes with semi-heavy-tailed distributions. It basically consists of a Brownian motion whose deterministic physical time flow is replaced by a new virtual process time. This virtual process time, called the subordinator, is driven by another Lévy process fulfilling a certain regularity condition. The outcome emerging from the subordination is a Lévy process again. Moreover, a range of advantages concerning various different aspects such as derivation of characteristic functions, moments or alternative representations of processes, random variables and their corresponding distributions are facilitated by this construction technique. Finally, one has to note that this approach enjoys a considerable level of acceptance, in particular among practitioners, since they do not have to give up their familiar opinion of Brownian motion for price dynamics. It is merely necessary to supplement this Gaussian framework by the assumption of a stochastic flow of virtual trading time replacing the conventional deterministic and linearly evolving physical time. Incidentally, this phenomenon of slowed and accelerated, even jumping, clocks on the trading floor, in a metaphorical sense, is confirmed by the subjective perception of many professional traders in the market[8]. This statement makes the theoretical modeling approach more plausible and demonstrates that subordination is not merely a technical convention but a very natural consequence of the attempt to incorporate present market realities into the development of models.

The most central innovation, however, is contained in the fact that the approach of Brownian subordination can be transferred to a multivariate setting while all of the mentioned advantages are preserved. This transition is accomplished by simply employing a n-dimensional Brownian motion instead of a merely univariate version of this basic stochastic process. At first sight, one might fear that the coupling of two stochastic processes in the course of subordination is creating a 'doubly' stochastic process of very intractable

[8] see Geman et al. [GeMY01].

nature which would require an arsenal of advanced and excessively technical calculation methods. This apprehension turns out to be unfounded, however. Instead, the exact opposite is the case as will be demonstrated in the course of the thesis. One has to emphasize that just this creates the actual utility with regard to practical applications of the chosen method in comparison to other alternative definitions of genuine tempered stable modeling approaches. In particular, it realizes a sufficiently flexible model setting in the multivariate case with a parsimonious parametrization involved while at the same time offering considerable level of analytic tractability. When considering the past decades from the beginning of quantitative financial modeling up to this day, this can truly be identified as one of the main factors which significantly decides over the success or failure of a mathematical model, at least with regard to practical relevance and adoption by potentials users in practice. Although the normal tempered stable process and corresponding distribution is far more complex and costly to handle than a Gaussian alternative, also not allowing for a closed form solution for most of the given problems, the essence is obvious and very similar. Because even as with the evidence becoming more and more overwhelming, convincing the majority of market participants of the inherent non-Gaussian nature of asset price dynamics, the normal distribution and Brownian motion retained their dominant role in the subsequent period. This was basically due to their capability of enabling analytical calculations like no other more realistic model, an advantage which began to crumble when computing power became affordable at a reasonable price[9].

In order to further clarify the objective and aim of the research comprised in this thesis, the following alternative approaches which are in close relationship to the multivariate normal tempered stable framework should be briefly touched upon here. This rounds out the purposeful overview and motivation presented at this point.

The concept of Brownian subordination or Gaussian mixture models, respectively, as a means to construct semi-heavy-tailed distributions and processes, has already been explored to some extent in the recent past. Among these are the variance gamma and the normal inverse Gaussian model with the generalized hyperbolic model embedding the for-

[9] A similar argumentation can be found in Greiner [Grei11]

mer two instances, although their multivariate extensions have not been studied in depth yet. Some efforts in this direction were made, but none of them could prevail on a broad range though. The current goal of this research is to generalize a genuine tempered stable process, whereas these former attempts do not fall under this specific category. This aim is motivated first and foremost by the already demonstrated flexible and precise adaptability of its univariate NTS counterparts to empirical financial data and appears the most promising for this reason.

Finally, another important concept in the context of multivariate modeling are copula functions. They are able to couple marginal distributions together with its inherent dependence structure to a multidimensional joint distribution. The main advantage lies in the given possibility of separating the marginal laws from their dependence structure, the latter represented in the copula function. In the opposite direction, the coupling via copula functions can be carried out with virtually arbitrary marginal distributions under very general conditions as there are very little restricting requirements for the inputs to be satisfied. Moreover, copula functions are able to reflect a wide range of possible dependence structures, for example non-linear and asymmetric cases, which then become incorporated in the resulting joint distribution. Therefore, they appear to be very appropriate for addressing multivariate problems which exhibit complex dependencies. On the other hand, their accurate estimation is not an easy task, especially when a parametrization with a suitable degree of freedom is chosen and the number of considered dimensions increases. In our focus, the main disadvantage, however, is that the coupling is only applicable for distributions and not for processes in general. More precisely, the copula is not preserving the infinite divisibility of distributions to be combined in this procedure. To overcome this drawback, Cont [CoTa04, ch. 5] is presenting copulas for creating multidimensional Lévy densities, they however lack the possibility for their direct estimation based on observable variables. Further textbooks on the topic have been written by Nelsen [Nels06] or Cherubini et al. [ChLV05], related publications are those of Embrechts et al. [EmLM03] and Demarta and McNeil [DeMc05].

1.2 Outline

The present thesis is organized as follows. While the objective was stated with an overview over the current status of related research in this chapter 1, chapter 2 continues with a brief introduction to Lévy processes as well as to further methods and theory concerning subordination. In chapter 3 follows a comprehensive treatment of univariate normal tempered stable models with an emphasis on the Brownian subordination perspective. Chapter 4 subsequently addresses the generalization of this concept to the case of multiple dimensions. It comprises the definition and thorough study of various forms of representation as well as central properties of the multivariate normal tempered stable process and related distributions. Their presentation is supplemented by necessary techniques for their manipulation. The primary theoretical contribution is to be found in this part. Chapter 5 is dedicated to a first practical application of the multivariate NTS in a simple model for joint stock and index returns. The emphasis lies on the development of an efficient estimation method which is then carried out based on empirical historical DJIA return data. A period of turbulent market conditions as part of the recent financial crisis is contained in this considered data set. Afterwards, the model adaptability is put to a test by means of several statistical goodness of fit tests. Based on the estimated model a portfolio optimization is conducted whose emerging strategies are subsequently compared with regard to their empirical performance in in-sample and out-of-sample tests. In chapter 6, the abilities of normal tempered stable distributions are further examined by its integration into a multivariate ARMA-GARCH time series framework of non-constant volatility. After the estimation of the entire econometric model the suitability assessment in this context is substantiated by two backtests on empirical data. It is followed by the derivation of optimal, now dynamic, portfolio strategies based on the estimated econometric model with respect to different criteria. In the remainder of this chapter, these strategies undergo empirical sample tests and performance evaluations similar to the lines of chapter 5. Chapter 7 concludes with a summary and interpretation of the main findings and results. It moreover points at directions of future research and possible further improvements.

Chapter 2

Lévy Processes and Tempered Stable Models

2.1 Lévy Processes

Lévy processes constitute a rather rich and flexible class of stochastic processes in continuous time. Paul Lévy developed the concept shortly after Norbert Wiener formalized the present notion of Brownian motion in the class of processes named after him. While Wiener based his processes on trajectories with continuous paths, although nowhere differentiable, Lévy maintained the basic principle of independent and identically distributed increments, but more generally allowing for jumps and discontinuities in their trajectories. This generalized principle is able to generate very flexible stochastic processes for a broad range of application fields but still offering a considerable degree of tractability. The mentioned flexibility of Lévy processes is reflected by the fact that by an arbitrary combination of a Gaussian diffusion, a deterministic linear trend and a variably designed jump component with finite or unbounded variation, numerous types of present stochastic processes can be covered. With the precise specification of these three components comprised in the so-called Lévy triplet each instance of a Lévy process is entirely and uniquely defined[1].

[1]Further details will be given in appendix A.1.

Apart from fundamental Poisson and compound Poisson processes the Brownian motion is contained in this class of processes, with the latter still, despite some countermovement has started to form in the recent past[2], representing a standard model in the area of financial modeling. Lévy processes, in recognition of their significance in the modern theory of stochastic processes, have already been well studied in the past. For an extensive overview over their theoretical and practical properties, we refer to a series of various existing textbooks on the topic, such as those by Bertoin [Bert05], Sato [Sato99], Applebaum [Appl05], Schoutens [Scho03], Kyprianou [Kypr06] or Barndorff-Nielsen [BNMR01].

2.2 Subordination

Besides offering the opportunity of defining a particular Lévy process by means of each single elements of the Lévy triplet, primarily the Lévy measure controlling the behavior of the involved jump component, Lévy subordination makes available a versatile approach for creating new instances of Lévy processes. In fact, subordination describes a stochastic time change of an existing stochastic process by a subordinator. The following presentation is based on the exposition of Sato [Sato99, p. 179 ff.] and Bertoin [Bert05, p. 71 ff.]

2.2.1 Subordinators

Univariate Lévy processes ${S_{(t),\, t\geq 0}}^3$ on $\mathbb{R}$ with almost surely non-decreasing trajectories are called *subordinators*. This property of their dynamics is closely related to the idea of time flow in the natural world[4]. According to Cont and Tankov [CoTa04, p. 88], this property possesses four different characterizations over the following equivalent conditions:

- Sample paths of $S_{(t)}$ are almost surely non-decreasing: $t > u \Rightarrow S_t \geq S_u$ a.s.
- $S_t \geq 0$ a.s. for some $t > 0$.

[2] We therefore refer to the exposition in chapter 1.

[3] For preventing the notation from becoming overly complicated, the time range of the processes will be omitted from now on. If not explicitly indicated otherwise, $S_{(t)}$ denotes a stochastic process $S_{(t),t\geq 0}$.

[4] With the difference that the time flow is assumed to be deterministic in the real world.

- $S_t \geq 0$ a.s. for every $t > 0$.

- The Lévy triplet $(A, \nu, m^{(0)})$ of $S_{(t)}$ has no diffusion component ($A = 0$) and a non-negative additional deterministic linear trend[5] $m^{(0)} \geq 0$. Moreover, the Lévy measure ν which has to be concentrated on the positive real line $\mathbb{R}_{++}$, i.e. $\nu((-\infty, 0]) = 0$, is of finite variation: $\int_0^\infty \min\{s, 1\}\nu(ds) < \infty$.

2.2.2 Time-changed Stochastic Processes

When a subordinator $S_{(t)}$ of the preceding section 2.2.1 is used for replacing the usual deterministic time index s of a stochastic process $Y_{(s)}$, Y is said to be *subordinated* by S. This creates a new stochastic process $X_{(t)}$

$$X_{(t)} = Y_{S_{(t)}}\,, \tag{2.1}$$

where $S_{(t)}$ represents the course of a virtual process time s for $Y_{(s)}$, embedded in the deterministic physical time flow t. One has to be aware that the process $Y_{(s)}$ is not restricted to one dimension and can therefore generally be defined on $\mathbb{R}^n$.

As long as $Y_{(s)}$ is a general Lévy process itself, the outcome $X_{(t)}$ of the subordination again satisfies the conditions of a Lévy process. It can be represented either in terms of its set of characteristic functions ϕ_{X_t} for X_t, obeying a specific structure induced by the general properties of a Lévy process, or the corresponding entire Lévy triplet (A, ν, m) of the process. In more detail, the characteristic exponent function (CEF) ψ_X of a Lévy process $X_{(t)}$ is able to describe the characteristic function (CF) of the distribution of the process variable X_t for each $t > 0$, while being time-independent itself. The characteristic function of X_t is given by

$$\phi_{X_t}(z) = E(\exp(iz^\mathsf{T} X_t)) = \exp(t \cdot \psi_X(z)) \Leftrightarrow \psi_X(z) = t^{-1} \ln \phi_{X_t}(z) \;\forall\, t > 0\,,$$

[5] A detailed description of several common conventions for compensating the jump component specified by the Lévy measure ν together with connected types of linear trend components $m^{(i)}$, $i = 0, 1, 2$ is given in appendix A.1.

for possibly multivariate arguments $z \in \mathbb{R}^n$. A similar result holds true for the moment-generating function M_{X_t}, $t > 0$, possibly defined only on a restricted feasible subset G of $\mathbb{R}^n$

$$M_{X_t}(u) = E(\exp(u^\mathsf{T} X_t)) = \exp(t \cdot l_X(u)), \ u \in G$$

which can be stated in terms of the corresponding Laplace exponent function $l_S(u) = t^{-1} \ln M_{S_t}(u) \ \forall \, t > 0$. But while the characteristic function ϕ_{X_t} and the characteristic exponent function ψ_X generally exist on the whole domain $\mathbb{R}^n$, the same is not true in general for the moment-generating function M_{X_t} and the associated Laplace exponent function l_X, as the former one is closely linked to the existence of exponential orders of X_t.

An eminently desirable property of Lévy subordination is that the CEF and therefore also the CF of the subordination outcome $X_{(t)}$ describing the distributional properties of the associated X_t can be easily obtained by means of the CEF of the process Y and the Laplace exponent function of the subordinator S. This is reflected in the following expression

$$\psi_X(z) = l_S(\psi_Y(z)) \tag{2.2}$$

$$\phi_{X_t}(z) = \exp(t \cdot \psi_X(z)) = \exp(t \cdot l_S(\psi_Y(z))) = M_S(\psi_Y(z)), \ z \in \mathbb{R}^n \tag{2.3}$$

At this point, it is important to note that although the Laplace exponent function l_S might only be defined on a restricted domain G, its expression can nevertheless be considered to be valid for the required entire range of ψ_Y. This is due to arguments of analytic continuation[6], see for example Cont and Tankov [CoTa04, p. 121].

For this considered most general case, the Lévy triplet[7] $(A_X, \nu_X, m_X^{(1)})$ of $X_{(t)}$ can be derived from the Lévy triplet $(A, \nu, m^{(1)})$ of $Y_{(s)}$ and the constrained one of the subordinator

[6] Note that complex arguments u of $l_S(u)$ do not pose a problem in any case. Analytic continuation becomes relevant in the connection with $\Re(u) \notin G$.

[7] The convention connected to $m^{(1)}$ compensates small jumps of magnitude less than 1. By the defining properties, this particular convention is always valid, making $m^{(1)}$ a universal trend specification.

$(A_S, \nu_S, m_S^{(0)}) = (0, \rho, b)$ satisfying the above requirements, by

$$A_X = bA\,, \tag{2.4}$$

$$\nu_X(B) = b\nu(B) + \int_0^\infty P_{Y_s}(B)\rho(ds)\,, \quad \forall\, B \in \mathcal{B}(\mathbb{R}^n)\,, \tag{2.5}$$

$$m_X^{(1)} = bm^{(1)} + \int_0^\infty \rho(ds) \int_{|y|\leq 1} y P_{Y_s}(dy)\,. \tag{2.6}$$

$P_{Y_s}(B) = P(Y_s \in B)$ is the probability measure of the process variable Y_s at time s on the Borelian σ-algebra $\mathcal{B}(\mathbb{R}^n)$ of $\mathbb{R}^n$.

Both general schemes presented here will be put to use for particular cases in the forthcoming sections and in chapter 4.

2.2.3 Brownian Subordination

Since the above general setting demonstrates the potential possibilities while not being very constructive at the same time, a less comprehensive situation should be considered in this section. For this purpose, the *subordination object* or *subordinant* $Y_{(s)}$ is chosen to be a, possibly multivariate, Brownian motion $B_{(s)}$. In a consistently multivariate perspective, the Brownian motion is entirely described by a real vector $\gamma \in \mathbb{R}^n$ and a $n \times n$ covariance matrix Σ, Σ therefore having to be positive semi-definite. While γ specifies the deterministic drift over time, Σ is responsible for the diffusion behavior.

At a later stage the characteristic function ϕ_{B_s} of B_s

$$\phi_{B_s}(z) = \exp(s \cdot \underbrace{(i\gamma^\mathsf{T} z - \frac{1}{2} z^\mathsf{T} \Sigma z)}_{\psi_B(z)})\,,\ z \in \mathbb{R}^n \tag{2.7}$$

is required, where ψ_B denotes the time-independent characteristic exponent function of $B_{(s)}$. Moreover, the Lévy triplet is given by $(A, \nu, m^{(1)})$ or $(A, \nu, m^{(0)})$, respectively, with

$$A = \Sigma$$

$$\nu(B) = 0\,, \quad \forall\, B \in \mathcal{B}(\mathbb{R}^n)$$

$$m^{(1)} = m^{(0)} = \gamma$$

The Lévy triplet demonstrates that the Brownian motion as a purely Gaussian process incorporates, besides a deterministic trend and a diffusion component, no jumps. Therefore, both representations $m^{(1)}$ and $m^{(0)}$ are identical. For the further processing in the context of Brownian subordination it is of relevance that B_s at time s is normally distributed, $B_s \sim \mathcal{N}^n(s\gamma, s\Sigma)$, with n-dimensional multivariate density function

$$f_{B_s}(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|s\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - s\gamma)^{\mathsf{T}}(s\Sigma)^{-1}(x - s\gamma)\right)\,, \tag{2.8}$$

as long as Σ is strictly positive definite.

The Brownian motion still serves as the standard model for price dynamics in many fields of mathematical finance, which is, above all, due to its analytical tractability. This enabled, inter alia, the development of a widely applicable stochastic differential calculus, see for example Karatzas and Shreve [KaSh00] or Øksendal [Økse07]. It provides the theoretical foundation on which modern financial engineering is built upon. For these particular reasons, Brownian motion has become the most well-established and accepted stochastic model among practitioners in the financial industry, although its inappropriateness for describing empirical market behavior has been sufficiently proven to this date[8].

Many of the later considered applications of Brownian subordination do not necessarily require the handling of the Lévy triplet or the characteristic function, respectively. These elaborate procedures can often be circumvented by concerning the following decompositions, which make the Brownian subordination an attractive approach from an analytical point of view

$$Y_{(t)} = B_{S_{(t)}} \tag{2.9}$$

[8]The central arguments with associated references are given in chapter 1.

$$= S_{(t)}\gamma + Z_{S_{(t)}}\ ,\ Z_{(s)} \equiv \mathrm{BM}(0, \Sigma)$$

$$= S_{(t)}\gamma + \mathrm{chol}(\Sigma) W_{S_{(t)}}\ ,\ W_{(s)} \equiv \mathrm{BM}(0, I)$$

$$= S_{(t)}\gamma + \mathrm{diag}(\sigma)\mathrm{chol}(\mathsf{P}) W_{S_{(t)}}\ .$$

In this notation, I is the identity matrix of dimension n, $\mathrm{BM}(\cdot\,,\cdot)$ denotes a Brownian motion with the corresponding parameter values, while $\mathrm{chol}(\cdot)$ and $\mathrm{diag}(\cdot)$ stands for the Cholesky decomposition or the diagonal matrix generated by a n-dimensional vector, respectively. In the last equation, the covariance matrix is decomposed into its contained vector of variances σ and the correlation matrix P, $\Sigma = \mathrm{diag}(\sigma)\mathsf{P}\mathrm{diag}(\sigma)$.

Brownian subordination provides a powerful means for representing, illustrating and understanding general Lévy processes in an arbitrary number of dimensions on the one hand. As it was shown by Clark [Clar73], every such Lévy process can be constructed by subordinating a Brownian motion with a particular Lévy subordinator of the type presented in section 2.2.1. Hence, every Lévy process can be understood as a well-established Brownian motion, but with the only specialty of elapsing in a virtual process time embodied by a stochastic subordinator. Nevertheless, with this additional assumption regarding the modified time flow[9], it is still possible to essentially maintain the Brownian motion in this way, even when having to strongly reject Gaussianity for observable prices. On the other hand, decomposing general Lévy processes in such a manner will yield reasonable subordinator processes in only very few instances.

The other way around, subordination will be used in the next section as a constructive tool for defining a particular class of Lévy processes. This particular approach will enable the use of all the helpful features subordination is offering to the modeler.

[9] As the time flow is ruled by a Lévy process of some regularity, this does not even appear to be overly artificial.

2.3 Tempering Stable Processes

2.3.1 The α-stable Process

The class of α-stable processes is comprising the set of all self-similar Lévy processes[10]. These α-stable processes, as being purely non-Gaussian Lévy processes, are essentially specified by their Lévy density $\nu(x)$, given in the univariate case by

$$\nu(x) = \frac{c_+}{|x|^{\alpha+1}} \mathbb{1}_{>0}(x) + \frac{c_-}{|x|^{\alpha+1}} \mathbb{1}_{<0}(x)\,, \tag{2.10}$$

with parameters $\alpha \in (0,2)$, $c_+, c_- \geq 0$, and the linear trend component $m^{(1)} \in \mathbb{R}$. The polynomial decay of the Lévy density function $\nu(x)$ is then responsible for the problematic non-existence of moments of fractional order greater or equal than α. For a complete treatment of α-stable processes and distributions it can be referred to Samorodnitsky and Taqqu [SaTa00] and Rachev and Mittnik [RaMF05] with the area of finance in view. This particular feature, in addition to the fact that no closed-form expression for the related probability density function exists, renders its practical handling quite burdensome. Further required details on α-stable processes including the characteristic function of their distributions are given in appendix A.2.

2.3.2 Tempered Stable Processes

To remedy at least some of these disadvantages, one might extenuate the above polynomial tails of the α-stable Lévy density by the multiplication of an exponentially decreasing *tempering function*, which in turn is primarily being responsible for generating finite moments of all orders. This approach introduced by Koponen [Kopo95] as *truncated Lévy flights*, Boyarchenko and Levendorskiĭ [BoLe00] as *KoBoL* and finally by Carr et al. [CGMY02] as *CGMY*, has been generally termed as *Tempered Stable* processes in the literature. Since then, a broad range of various types of tempered stable processes, characterized by their individual classes of tempering functions, have been proposed. Rosiński [Rosi07] and Bianchi

[10] Less generally speaking, the Brownian motion connected to the stability index $\alpha = 2$ also possesses the property of self-similarity but is usually excluded from the actual class of α-stable processes.

et al. [BRKF10a] have both individually developed separate frameworks for generalizing the concept of tempering stable or infinitely divisible distributions, respectively.

One of the most basic instances of TS processes is the *Classical Tempered Stable (CTS)* process[11], whose details shall be presented in the following example for demonstrating the general approach. It employs, as already Koponen did in his pioneering work, an exponential tempering function. Moreover, a suitable subordinator will be derived from the CTS process for later use.

$$\nu(x) = \frac{c_+ \exp(-\lambda_+|x|)}{|x|^{\alpha+1}} \mathbb{1}_{>0}(x) + \frac{c_- \exp(-\lambda_-|x|)}{|x|^{\alpha+1}} \mathbb{1}_{<0}(x)\,, \tag{2.11}$$

where $\mathbb{1}_{>0}(x)$ represents the indicator function for strictly positive real numbers. The additional parameters $\lambda_+, \lambda_- > 0$ serve to control the exponential tempering on the positive and negative area of the real support, respectively.

[11] Incidentally, CTS lies at the core of the aforementioned KoBoL and CGMY models.

Chapter 3

Univariate Normal Tempered Stable Processes

Besides the Classical Tempered Stable class, univariate Normal Tempered Stable (NTS) processes suggested by Barndorff-Nielsen and Levendorskiĭ [BNLe01] were one of the first implementations of the tempered stable principle. Moreover, amongst the broad variety of tempered stable processes existing at the present time, normal tempered stable models are the only ones whose initial definition resort to Brownian subordination and its numerous advantages.

3.1 The Subordinator

In order to lead to a suitable subordinator, the Lévy density of a CTS from equation (2.11) has to be concentrated on the positive real line first

$$\nu(x) = \frac{C \exp(-\lambda x)}{x^{\alpha+1}} \cdot \mathbb{1}_{>0}(x)\,, \tag{3.1}$$

by just eliminating the negative branch of the support. The parameters C and λ both have to be strictly positive, i.e. $C, \lambda > 0$. The parameter α is the index of stability of the

α-stable process, where $\alpha \in (0, 2)$[1]. For $\alpha \in [1, 2)$ however, the jumps are not of finite variation, contradicting the requirements of a valid subordinator, which is why they have to be excluded from the range of feasible parameter values. On the other side, for $\alpha < 0$ the total mass of the Lévy measure on either large jumps and small jumps is bounded, the latter is due to exponential tempering. Although an existing Lévy process would be specified, it is only of finite activity, however. This corresponds to a compound Poisson process, which moves by discrete jumps only, a behavior considered unrealistic for the virtual trading time in view.

To complete the specification of its entire Lévy triplet, it is furthermore necessary to state that no Gaussian diffusion component is present ($A = 0$). As the Lévy density in equation (3.1) implies a process of finite variation, it is appropriate to choose the valid uncompensated version of the jump process. This choice enables a direct view on the effects of the linear trend component $m^{(0)}$ on the properties of the trajectories. In order obtain a parsimonious parametrization for the subordinator, $m^{(0)}$ is set to zero.

The kind of process defined so far implies the following characteristic function for the distribution of their process variables S_t

$$\phi_{S_t}(z) = \exp(t \cdot \psi_S(z)) \tag{3.2}$$

$$\psi_S(z) = C\Gamma(-\alpha)\left[(\lambda - iz)^\alpha - \lambda^\alpha\right], \tag{3.3}$$

where $\Gamma(\cdot)$ denotes the conventional Gamma function. This CF and CEF can be obtained from applying the Lévy-Khintchine formula[2] to the Lévy triplet of the preliminary subordinator specified so far. The Lévy-Khintchine formula provides a universal link between the representation of a Lévy process by means of its Lévy triplet and will be recurrently employed for many of the subsequent calculations. Note however that the CEF only corresponds to the expression in above equation (3.3) if one excludes $\alpha = 0$ from the feasible

[1] In the definition of an α-stable Lévy density, see equation (2.10), $\alpha \geq 2$ does not yield a valid Lévy process, as the Lévy density is no longer square-integrable around 0.

[2] For a definition of the Lévy-Khintchine formula and further details see appendix A.1. The derivation of CF and CEF in the case of general CTS processes is performed in appendix A.3.

parameter range, which therefore should be the case for the further presentation.

At this stage, a useful implicit method for calculating distributional moments from characteristic functions is going to be introduced, which will be further drawn upon in the forthcoming sections. First of all it should be noted that the necessary existence of all moments of order k in this case can be inferred from the following existence criterion[3]

$$\int_{\mathbb{R}} |x|^k \nu(x) dx \overset{!}{<} \infty \,,$$

which is based on the Lévy density function $\nu(x)$.

Then, cumulants of S_t are reflected by the derivatives of CEF $\psi_S(z)$ of corresponding k-th order at the point $z = 0$ in the following way

$$c_{S_t}^{(k)} = \frac{1}{i^k} \left. \frac{\partial^k \ln \phi_{S_t}(z)}{\partial z^k} \right|_{z=0} = \frac{t}{i^n} \left. \frac{\partial^k \psi_S(z)}{\partial z^k} \right|_{z=0} . \tag{3.4}$$

With the corresponding derivatives of the subordinators CEF

$$\frac{\partial^k \psi_S(z)}{\partial z^k} = i^k C\Gamma(n-\alpha) \left[(\lambda - iz)^{\alpha - n} \right] \,,\ k \geq 1 \,, \tag{3.5}$$

we have that

$$c_{S_t}^{(k)} = t \cdot \left[C\Gamma(n-\alpha)\lambda^{\alpha-n} \right] \,,\ n \geq 1 \,.$$

From this sequence of cumulants, any moment of integer order $k \in \mathbb{N}$ can be easily calculated, which should be omitted for a general general case here. Only the case of $k = 1$, which is indeed of use for the further presentation, should be considered explicitly here. Although general cumulants are aggregations of moments of different orders, $c_{S_t}^{(1)}$ immediately corresponds to the first moment and expectation $E(S_t)$

$$c_{S_t}^{(1)} = E(S_t) = t \cdot C\Gamma(1-\alpha)\lambda^{\alpha-1} \,.$$

[3] see Cont and Tankov [CoTa04, ch. 3]

To avoid later redundancies concerning the generated processes[4], one furthermore restricts the set of feasible subordinators to those possessing a unit rate of expectation, i.e. $E(S_t) = 1$. This can be achieved by adjusting the value of parameter C appropriately

$$E(S_t) = t \cdot C\Gamma(1-\alpha)\lambda^{\alpha-1} \stackrel{!}{=} 1$$

$$\Rightarrow C \stackrel{!}{=} \tilde{C} = \frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha)}\,.$$

This chosen way of ensuring the required expectation rate of 1 leaves the two parameters $\alpha \in (0,1), \lambda > 0$ remaining. Because the parameter values of α of the CTS subordinator, in contrast to the former initial α-stable process, is restricted to $(0,1)$, we change to a new parameter $a = 2\alpha$ for the polynomial decay of the stable Lévy density involved, resulting in a to be within the original boundaries $a \in (0,2)$ of the α-stable Lévy density. Moreover, this should also help to clearly distinguish the difference between a conventional CTS process and the new subclass of CTS processes used for subordination defined here, which i.a. have a restricted range for the values of the stability index.

In summarizing, the elaborated subordinator based on the general CTS process has the following properties and representations. First of all, the subordinator S_t is a purely non-Gaussian Lévy process with infinite activity but finite variation. Because of the latter fact, it can be represented with the Lévy density $\rho_S(s)$ of an uncompensated jump component

$$\rho_S(s) = \frac{\tilde{C}\exp(-\lambda s)}{s^{1+\frac{a}{2}}} \cdot \mathbb{1}_{>0}(s) = \frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})} \cdot \frac{\exp(-\lambda s)}{s^{1+\frac{a}{2}}} \cdot \mathbb{1}_{>0}(s)\,. \tag{3.6}$$

As no additional linear trend component is added, the Lévy triplet (A_S, ρ_S, b_S) of the process $S_{(t)}$ therefore is simply

$$(A_S, \rho_S, b_S) = (0, \rho_S, 0)\,.$$

[4]By redundancies we refer to the fact that, as will be shown later, it would be possible to create the same subordination process in a number of ways. This is achieved by varying C in a particular ratio to the parameter values of the Brownian motion.

The Lévy triplet, besides its use for the derivation and verification of some important properties of the CTS subordination process, will mainly be further utilized for determining central characteristics of the subordination result in section 3.5.

Applying the Lévy-Khintchine formula to the above Lévy triplet yields the characteristic exponent function in a straightforward way[5]

$$\psi_S(z) = \tilde{C}\Gamma\left(-\frac{a}{2}\right)\left[(\lambda - iz)^{\frac{a}{2}} - \lambda^{\frac{a}{2}}\right] = \frac{\lambda^{1-\frac{a}{2}}}{\Gamma\left(1-\frac{a}{2}\right)}\Gamma\left(-\frac{a}{2}\right)\left[(\lambda - iz)^{\frac{a}{2}} - \lambda^{\frac{a}{2}}\right]$$

$$= \frac{2\lambda}{a}\left[1 - \left(1 - \frac{iz}{\lambda}\right)^{\frac{a}{2}}\right], \quad z \in \mathbb{R}\,.$$

Hence, the actual characteristic functions ϕ_{S_t} of S_t for $t > 0$ follows by

$$\phi_{S_t}(z) = \exp(t \cdot \psi_S(z)) = \exp\left[t \cdot \frac{2\lambda}{a}\left[1 - \left(1 - \frac{iz}{\lambda}\right)^{\frac{a}{2}}\right]\right]. \tag{3.7}$$

As the Lévy density of the jumps, being the only non-trivial component of the pure-jump process, is concentrated on the positive real line only, as well as therefore the probability mass of S_t at any time $t > 0$, valid arguments u of the moment-generating function M_{S_t} are not bounded towards $-\infty$. For the specified CTS subordinator, it actually holds that $G = (-\infty, \lambda]$. The associated Laplace exponent function can be directly obtained from the CEF

$$l_S(u) = \psi_S(-iu) = \frac{2\lambda}{a}\left[1 - \left(1 - \frac{u}{\lambda}\right)^{\frac{a}{2}}\right], \quad u \in G\,,$$

whose analytical expression is allowed to be extended to purely imaginary arguments for $u \in \mathbb{F}$. This, after all, yields the MGF

$$M_{S_t}(u) = \exp\left[t \cdot l_S(u)\right] = \phi_{S_t}(-iu)$$

$$= \exp\left[\frac{2\lambda t}{a}\left[1 - \left(1 - \frac{u}{\lambda}\right)^{\frac{a}{2}}\right]\right], \quad u \in G\,.$$

[5] Single steps of the calculation are performed in Cont and Tankov [CoTa04, p. 121]

In fact, it has to be again noted that the actual G is not of any relevance for later subordination purposes.

After discussing the process dynamics by means of the Lévy triplet, the distributional properties of S_t for all points in physical time $t > 0$ should be presented here in form of distributional moments. Again, their calculation can be performed by the entire sequence of cumulants for different orders, obtained by the method presented in equation (3.4)

$$c_{S_t}^{(k)} = \frac{1}{i^k} \cdot \frac{\partial^k \ln(\phi_{S_t}(z))}{\partial z^k}\bigg|_{z=0} = \frac{t}{i^k} \cdot \frac{\partial^k \cdot \psi_S(z)}{\partial z^k}\bigg|_{z=0}, \quad k \in \mathbb{N}$$

$$= t\lambda^{1-k} \prod_{j=1}^{k-1} \left(j - \frac{a}{2}\right), \quad d = 1, 2, \ldots$$

The existing moments of all orders of $S_t, t > 0$ ensures the required smoothness of $\psi_{S_t}(z)$ around 0. Note at this point that, like as for every Lévy process, the cumulants are evolving strictly linear in t. Corresponding central statistics are in a further step directly concluded from these cumulants by

$$E(S_t) = c_{S_t}^{(1)} = t \tag{3.8}$$

$$Var(S_t) = c_{S_t}^{(2)} = \frac{t}{\lambda}\left(1 - \frac{a}{2}\right) \tag{3.9}$$

$$s(S_t) = \frac{c_{S_t}^{(3)}}{\left[c_{S_t}^{(2)}\right]^{3/2}} = \frac{t\left(1-\frac{a}{2}\right)\left(2-\frac{a}{2}\right)\lambda^{3/2}}{\lambda^2 t^{3/2}\left(1-\frac{a}{2}\right)^{3/2}} = \frac{\left(2-\frac{a}{2}\right)}{\sqrt{\lambda t}\sqrt{1-\frac{a}{2}}} \tag{3.10}$$

$$e(S_t) = \frac{c_{S_t}^{(4)}}{\left[c_{S_t}^{(2)}\right]^{2}} = \frac{t\left(1-\frac{a}{2}\right)\left(2-\frac{a}{2}\right)\left(3-\frac{a}{2}\right)\lambda^2}{\lambda^3 t^2\left(1-\frac{a}{2}\right)^2} = \frac{\left(2-\frac{a}{2}\right)\left(3-\frac{a}{2}\right)}{\lambda t \cdot \left(1-\frac{a}{2}\right)} \tag{3.11}$$

Expectation $E(S_t)$ as a measure of location and the variance $Var(S_t)$ as a measure of dispersion, both grow strictly linear in time t, therefore being unbounded. In contrast, the skewness $s(S_t)$ and excess kurtosis $e(S_t)$, as being rescaled central moments of S_t, embody scale independent measures of asymmetry and heavy-tailedness, respectively, and

both converge to 0 as $t \to \infty$. This behavior corresponds to the more general fact that every Lévy process with finite moment of first an second order is under the rule of the *Central Limit Theorem*. As these conditions are fulfilled for the CTS subordinator under consideration, its distributions do more and more converge to those of a corresponding univariate Brownian motion.

This already gives an adequate impression of the stochastic progress of the virtual time S_t at the point t measured in physical time. The first four cumulants and their associated central statistics illustrate how the two process parameters (a, λ), each responsible for polynomial or exponential decay of the Lévy density function, respectively, affect the distribution of the subordinator. The probability density function (PDF) f_{S_t} of the random variable S_t, which would provide a more thorough understanding of the distribution characteristics cannot be given in an analytical expression, however. The reason for this is that the inversion of the characteristic function based on a one-dimensional Fourier transformation possesses no closed-form solution. Various numerical techniques designed for obtaining at least point-wise numerical approximations of function values of $f_{S_t}(x)$ will be presented in the context of a direct application in section 5.3.

An important remark concerning the connection to two other well established Lévy processes follows here. The defined CTS subordinator includes the Inverse Gaussian (IG) process for $a = 1 \in (0, 2)$. This particular process has already been employed as a subordinating process for the univariate and multivariate Brownian motion in the past, yielding the Normal Inverse Gaussian (NIG) model. Another major instance is the Gamma process obtained in the limiting case $a \to 0$, which in the same way has been employed to create the Variance Gamma (VG) model.[6]. The reason for excluding the case $a = 0$ from the CTS subordinator lies in the fact that, although the parameterized term for the Lévy density in equation (2.11) remains valid for the Gamma process, the further processing of this Lévy density function with the Lévy-Khintchine formula towards the characteristic functions yields a different expression compared to the one in equation (3.7) for the CTS with $a \in (0, 2)$. Recall that $a = 0$ represents the lowest a for which the Lévy density in

[6] An extensive overview over the development of Variance Gamma and Normal Inverse Gaussian models as well as their applications to finance and other fields are given in section 3.8.

equation (3.6) still implies a process of infinite activity, while $a < 0$ produces compound Poisson-like processes with inappropriate properties with regard to the modeling of virtual time flows. Both Inverse Gaussian and Gamma distributions belong to the general class of hyperbolic distributions, making it possible to state their probability density function in closed form. However, as already pointed out, this is not given in general for the case $a \in (0,2)\backslash\{1\}$ in the definition of the considered CTS subordinator process.

3.2 The Subordination Object: Univariate Brownian Motion

As object $Y_{(s)}$ for the subordination, the Brownian motion $B_{(s)}$ is chosen. The univariate Brownian motion considered here comprises two individual scalar parameters $\gamma \in \mathbb{R}$ and $\sigma^2 > 0$. Brownian motions are the only purely Gaussian processes in the entire set of Lévy processes, their Lévy triplet (A_B, ν_B, m_B) only incorporate a diffusion component $A_B = \sigma^2$, no jump component $\nu_B(x) = 0$ and a universal[7] linear deterministic trend component $m_B = \gamma$.

This results in all $B_s\,,\ s > 0$ being normally distributed, $B_s \sim \mathcal{N}(s\gamma, s\sigma^2)$. For purposes of later subordination, the corresponding probability density function

$$f_B(s;x) = \frac{1}{\sigma\sqrt{2s\pi}} \exp\left(-\frac{(x-s\gamma)^2}{2s\sigma^2}\right)$$

and characteristic function[8]

$$\phi_{B_s}(z) = \exp\left(s(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right)$$

$$\psi_B(z) = s^{-1} \ln \phi_{B_t}(z) = i\gamma z - \frac{1}{2}\sigma^2 z^2\,, \quad z \in \mathbb{R}$$

[7] As no jump component is present, all compensation conventions are admissible and yield the same external trend coefficient $m_B = m_B^{(0)} = m_B^{(1)} = m_B^{(2)}$, which is noted here for the sake of completeness.

[8] Both PDF and CF of a normally distributed random variable can for example be found in Feller [Fell68], [Fell71].

are of closer relevance.

3.3 Construction

The univariate *Normal Tempered Stable (NTS)* process $X_{(t)}$ in the univariate case is defined by replacing the physical time t of a univariate Brownian motion $B_{(s)}$ with drift parameter $\gamma \in \mathbb{R}$ and diffusion parameter $\sigma^2 > 0$ from the above section 3.2 by a stochastic time flow according to the CTS subordinator $S_{(t)}$ with parameters $a \in (0, 2)$ and $\lambda > 0$ from section 3.1. In fact, this is implemented by means of the familiar subordination scheme presented equations (2.1) and (2.9)

$$X_{(t)} = B_{S_{(t)}} + \Delta_t = \delta \cdot t + \gamma S_{(t)} + \sigma W_{S_{(t)}}\,, \quad W_{(s)} \equiv BM(0,1)\,,$$

which is generalized in the sense of incorporating an additional deterministic linear trend in physical time, Δ_t. The processes $B_{(s)}$ and $S_{(t)}$ are mutually independent and Δ_t is completely specified by its associated coefficient $\delta \in \mathbb{R}$ in the above equation.

Moreover, $W_{(s)}$ denotes a standard Wiener process on time scale $s \in \mathbb{R}_+$, which is equivalent to a Brownian motion with standardized parameters, $BM(0,1)$. Because the focus is still on the entire process instead of only single distributions, the stochastic time index $S_{(t)}$ cannot be separated from the Brownian motion yet, as it will be the case for distributions in a further section 3.7. This separation will prove to be useful for the moments of X_t, especially for the multivariate case still to follow in chapter 4 and in particular for related distributions.

For the ease of subsequent operations, the deterministic drift coefficient is set to $\delta = (\mu - \gamma)$, determined by γ coming from the Brownian motion and an additional implicit parameter $\mu \in \mathbb{R}$. This particular construction yields

$$\begin{aligned} X_{(t)} &= \delta \cdot t + \gamma S_{(t)} + \sigma W_{S_{(t)}} = (\mu - \gamma) \cdot t + \gamma S_{(t)} + \sigma W_{S_{(t)}} \\ &= \mu t + \gamma (S_{(t)} - t) + \sigma W_{S_{(t)}}\,. \end{aligned} \tag{3.12}$$

The main advantage is that $(S_{(t)} - t)$ is becoming a driftless process governed by γ. Equation (3.12) defines the *univariate Normal Tempered Stable (NTS) process* $X_{(t)}$ with parameter tuple $(a, \lambda, \gamma, \sigma^2, \mu)$.

3.4 Characteristic Functions of the NTS Process

Determining the characteristic exponent function as a central means of representation for the created univariate NTS process does not necessarily require its entire Lévy triplet and the further processing by means of the Lévy-Khintchine formula.

Instead, one particular advantage of constructing the NTS process via generalized subordination lies in the availability of simple formula for characteristic exponent functions and characteristic functions of X_t

$$\psi_X(z) = l_S(\psi_B(z)) + \psi_\Delta(z) \tag{3.13}$$

$$\phi_{X_t}(z) = M_{S_t}(\psi_B(z)) \cdot \phi_{\Delta_t}(z) = \exp(t \cdot \psi_X(z)) \tag{3.14}$$

similar to equations (2.2) and (2.3). This proves particularly helpful, when for further issues only distributional properties of the subordination outcome is of relevance while detailed information of the process dynamics in form of the entire Lévy triplet is not required. Filling in the corresponding components from sections 3.1 and 3.2 into equations (3.13) and (3.14) leads to

$$\psi_X(z) = \frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right)^2\right] + i(\mu - \gamma)z\,, \quad z \in \mathbb{R} \tag{3.15}$$

$$\phi_{X_t}(z) = \exp\left[t \cdot \left[\frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right)^2\right] + i(\mu - \gamma)z\right]\right]\,, \quad z \in \mathbb{R}\,. \tag{3.16}$$

3.5 Lévy Triplet of the NTS Process

Nevertheless, besides the distributional properties obtained from the characteristic functions in the above section 3.4, the knowledge of the entire Lévy triplet of $X_{(t)}$ gives additional valuable insights in the structure and certain aspects of the process behavior, therefore it should be covered in this treatise. Special applications, for example in the area of stochastic calculus and financial option pricing, necessarily require this detailed kind of representation.

The basic methods for determining the entire Lévy triplet of a subordinated Lévy process given in equations (2.4) – (2.6) and are now carried out for this more specific context.

The Lévy triplet $(A_X, \nu_X, m_X^{(2)})$ of the resulting Lévy process $X_{(t)}$ contains the coefficient of the Gaussian diffusion component

$$A_X = m_S^{(0)} \cdot A_B = 0 \cdot \sigma^2 = 0\,.$$

Hence, as well as the subordinator $S_{(t)}$ constitutes a purely non-Gaussian process so does $X_{(t)}$. An existing diffusion component of the subordination object would only be conveyed to the subordination result by a deterministic trend of the subordinator not being present in this case.

Recall from equation (2.5), the most general form of the determining formula for the jump measure ν_X in the univariate case

$$\nu_X(\xi) = b_S \cdot \nu_B(\xi) + \int_0^\infty P_{B_s}(\xi)\rho(ds)\,, \quad \forall\, \xi \in \mathcal{B}(\mathbb{R})\,,$$

is given in terms of masses $\nu_X(\xi)$ of appropriate subsets $\xi \in \mathcal{B}(\mathbb{R})$ of $\mathbb{R}$. A similar situation is present for the Brownian motion and the subordinator, therefore allowing the probability measure P_{B_s} to be replaced by its likewise well-defined Gaussian density function f_{B_s}, while $\rho(ds)$ becomes $\rho_S(s)ds$. Details on the following derivation are additionally given in

appendix B.1, p. 172.

$$\nu_X(x) = \frac{\nu_X(dx)}{dx} = m_S^{(0)}\nu_B(x) + \int_0^\infty f_B(s;x)\rho_S(s)ds$$

$$= 0 + \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma\sqrt{s}} \exp\left(-\frac{(x-\gamma s)^2}{2\sigma^2 s}\right) \cdot \frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})} \frac{\exp(-\lambda s)}{s^{\frac{a}{2}+1}}$$

$$= \frac{\sqrt{2}\lambda^{1-\frac{a}{2}}\left(\gamma^2+2\sigma^2\lambda\right)^{\frac{a+1}{4}}}{\sqrt{\pi}\sigma\Gamma(1-\frac{a}{2})} \exp\left(\frac{x\gamma}{\sigma^2}\right) \frac{K_{\frac{a+1}{2}}\left(\frac{|x|\sqrt{\gamma^2+2\sigma^2\lambda}}{\sigma^2}\right)}{|x|^{\frac{a+1}{2}}}. \tag{3.17}$$

K_d denotes the modified Bessel function of the second kind and of order d. Note that this Bessel function now embodies the tempering function of a re-emerging α-stable Lévy density. A more detailed discussion of the present structure of the Lévy density function $\nu_X(x)$ follows below.

As X_t possesses finite moments of every order k, especially for $k = 1$ corresponding to the expectation, the jump component can be considered as evolving under full compensation of either small and large jumps. Moreover, Δ_t is contributing an additional trend to the pure subordination component B_{S_t}, which has to be incorporated in the following expression for the corresponding external trend coefficient $m_X^{(2)}$

$$m_X^{(2)} = m_S^{(0)} \cdot m_B^{(2)} + \int_0^\infty \left[\int_{\mathbb{R}} x f_B(s;x)dx\right] \rho_S(s)ds + (\mu - \gamma)$$

$$= 0 + \int_0^\infty E(B_s)\rho_S(s)ds + (\mu - \gamma)$$

$$= \int_0^\infty \gamma s \frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})} \frac{\exp(-\lambda s)}{s^{1+\frac{a}{2}}} ds + (\mu - \gamma)$$

$$= \frac{\gamma\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})} \int_0^\infty \exp(-\lambda s) s^{-\frac{a}{2}} ds + (\mu - \gamma)$$

$$= \frac{\gamma\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})} \int_0^\infty \exp(-t) \left(\frac{1}{\lambda}\right)^{-\frac{a}{2}} t^{-\frac{a}{2}} \frac{1}{\lambda} dt + (\mu - \gamma) \quad | \, t := \gamma s$$

$$= \frac{\gamma\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})}\lambda^{\frac{a}{2}-1}\int_0^\infty \exp(-t)t^{-\frac{a}{2}}dt + (\mu-\gamma)$$

$$= \frac{\gamma\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})}\lambda^{\frac{a}{2}-1}\Gamma(1-\frac{a}{2}) + (\mu-\gamma) = \mu\,.$$

The possible trend representation by $m_X^{(2)}$ in this case has the advantage of containing the entire drift of the process. This provides the opportunity of verifying the above result by comparing $m_X^{(2)}$ to the expectation $E(X_t)$ of X_t in equation (3.20).

The correspondence to an alternative means of defining the univariate NTS process will further be treated in section 3.7.4. This definition is based on an explicit specification of the Lévy density beforehand, which is constructed by exponential tilting of the Lévy density of a symmetric *Modified Tempered Stable (MTS)* process.

A further method for a cross-check of this result would be to convert the preceding Lévy triplet, which has emerged from the subordination, to the corresponding CEF of $X_{(t)}$, by application of the Lévy-Khintchine formula in its appropriate version for the case of full compensation

$$\psi_X(z) = im_X^{(2)}z - \frac{1}{2}A^2z^2 + \int_{\mathbb{R}\setminus\{0\}}(\exp(izx)-1-x)\nu_X(x)dx$$

$$= i\mu z + \int_{\mathbb{R}\setminus\{0\}}(\exp(izx)-1-x)\frac{\sqrt{2}\lambda^{1-\frac{a}{2}}\left(\mu^2+2\sigma^2\lambda\right)^{\frac{a+1}{4}}}{\sqrt{\pi}\sigma\Gamma(1-\frac{a}{2})}\exp\left(\frac{x\mu}{\sigma^2}\right)$$

$$\cdot\frac{K_{\frac{a+1}{2}}\left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right)}{|x|^{\frac{a+1}{2}}}dx \tag{3.18}$$

$$\stackrel{!}{=} \frac{2\lambda}{a}\left[1-\left(1-\frac{1}{\lambda}(i\gamma z-\frac{1}{2}\sigma^2z^2)\right)^{\frac{a}{2}}\right] + i(\mu-\gamma)z\,. \tag{3.19}$$

For the above result to be valid, the expression in equation (3.18) has to be equivalent to the characteristic exponent function given in equation (3.19). An actual implementation

of this calculation is waived here. After substituting $K_{\frac{a+1}{2}}$ by its integral representation the remaining procedure can be completed by following the steps given in Kim [Kim05, p. 30 ff.].

3.6 Cumulants and Moments

This section deals with the derivation of time-dependent moments of the process variable X_t. This is achieved by determining the cumulants from derivatives of k-th order of the characteristic exponent function associated with the process $X_{(t)}$. Thereafter, all information related to moments and other central statistics can be extracted from this sequence of cumulants.

In the following equations the first four derivatives of the characteristic exponent function[9] of the process $X_{(t)}$ with parameter vector $(a, \lambda, \gamma, \sigma^2, \mu)$ after its scalar argument z are presented.

$$\frac{\partial \psi_X(z)}{\partial z} = (i\gamma - \sigma^2 z)\left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-1} + i(\mu - \gamma)$$

$$\frac{\partial^2 \psi_X(z)}{\partial z^2} = -\sigma^2\left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-1} + \frac{1}{\lambda}(1 - \frac{a}{2})(i\gamma - \sigma^2 z)^2\left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-2}$$

$$\begin{aligned}\frac{\partial^3 \psi_X(z)}{\partial z^3} = &-3 \cdot \frac{\sigma^2}{\lambda}\left(1 - \frac{a}{2}\right)(i\gamma - \sigma^2 z) \cdot \left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-2} \\ &+ \frac{1}{\lambda^2}(1 - \frac{a}{2})(2 - \frac{a}{2})(i\gamma - \sigma^2 z)^3 \cdot \left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-3}\end{aligned}$$

$$\begin{aligned}\frac{\partial^4 \psi_X(z)}{\partial z^4} = &\, 3\frac{\sigma^4}{\lambda}(1 - \frac{a}{2})\left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-2} \\ &- 6\frac{\sigma^2}{\lambda^2}(1 - \frac{a}{2})(2 - \frac{a}{2})(i\gamma - \sigma^2 z)^2\left[1 - \frac{1}{\lambda}(i\gamma z - \frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-3}\end{aligned}$$

[9] see equation (3.15)

$$+\frac{1}{\lambda^3}(1-\frac{a}{2})(2-\frac{a}{2})(3-\frac{a}{2})(i\gamma-\sigma^2 z)^4\left[1-\frac{1}{\lambda}(i\gamma z-\frac{1}{2}\sigma^2 z^2)\right]^{\frac{a}{2}-4}$$

The actual cumulants are obtained from the corresponding derivatives of the now time-dependent *characteristic exponent (CE)* or *cumulant generating function (CGF)* $\Psi_{X_t}(z)$ of the random variable X_t, which is emerging from the time-independent CEF of the process $X_{(t)}$ by $\Psi_{X_t}(z) = \psi_X(z) \cdot t$. The possibility of isolating a specific time-independent term CEF, which implies the characteristic functions of every X_t, $t > 0$, is a distinguishing feature of the inherent structure of Lévy processes, compounded by i.i.d. increments. It holds that

$$c_{X_t}^{(1)} = \frac{1}{i}\frac{\partial \Psi_{X_t}(0)}{\partial z} = \frac{t}{i}\frac{\partial \psi_X(0)}{\partial z} = \mu t$$

$$c_{X_t}^{(2)} = \frac{1}{i^2}\frac{\partial^2 \Psi_{X_t}(0)}{\partial z^2} = \frac{t}{i^2}\frac{\partial^2 \psi_X(0)}{\partial z^2} = \left[\sigma^2 + \frac{\gamma^2}{\lambda}(1-\frac{a}{2})\right]\cdot t$$

$$c_{X_t}^{(3)} = \frac{1}{i^3}\frac{\partial^3 \Psi_{X_t}(0)}{\partial z^3} = \frac{t}{i^3}\frac{\partial^3 \psi_X(0)}{\partial z^3} = \left[3\frac{\sigma^2\gamma}{\lambda}(1-\frac{a}{2}) + \frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2})\right]\cdot t$$

$$c_{X_t}^{(4)} = \frac{1}{i^4}\frac{\partial^4 \Psi_{X_t}(0)}{\partial z^4} = \frac{t}{i^4}\frac{\partial^4 \psi_X(0)}{\partial z^4}$$

$$= \left[3\frac{\sigma^4}{\lambda}(1-\frac{a}{2}) + 6\frac{\sigma^2\gamma^2}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2}) + \frac{\gamma^4}{\lambda^3}(1-\frac{a}{2})(2-\frac{a}{2})(3-\frac{a}{2})\right]\cdot t\,.$$

As raw moments are more difficult to interpret and still obscure a clear view on the specific features of the distribution of X_t, various standard indicators are presented here instead

$$E(X_t) = c_{X_t}^{(1)} = \mu t \tag{3.20}$$

$$Var(X_t) = c_{X_t}^{(2)} = \left[\sigma^2 + \frac{1}{\lambda}(1-\frac{a}{2})\gamma^2\right]\cdot t \tag{3.21}$$

$$s(X_t) = \frac{c^{(3)}_{X_t}}{\left[c^{(2)}_{X_t}\right]^{3/2}} = \frac{\left[3\frac{\sigma^2\gamma}{\lambda}(1-\frac{a}{2}) + \frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2})\right]}{\left[\sigma^2 + \frac{\gamma^2}{\lambda}(1-\frac{a}{2})\right]^{3/2}} \cdot \frac{1}{\sqrt{t}} \tag{3.22}$$

$$e(X_t) = \frac{c^{(4)}_{X_t}}{\left[c^{(2)}_{X_t}\right]^{2}} = \frac{\left[3\frac{\sigma^4}{\lambda}(1-\frac{a}{2}) + 6\frac{\sigma^2\gamma^2}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2}) + \frac{\gamma^4}{\lambda^3}(1-\frac{a}{2})(2-\frac{a}{2})(3-\frac{a}{2})\right]}{\left[\sigma^2 + \frac{\gamma^2}{\lambda}(1-\frac{a}{2})\right]^{2}} \cdot \frac{1}{t} \tag{3.23}$$

One notes again that both skewness $s(X_t)$ and excess kurtosis $e(X_t)$ converge to 0 when the physical process time $t \to \infty$. This, in the wider sense, is implied by the fact that, as Lévy processes in addition to i.i.d. increments are having their first two integer moments finite, the central limits theorem applies. In consequence, the standardized process random variable $\frac{X_t - E(X_t)}{\sqrt{Var(X_t)}}$ converges in distribution to a standard Gaussian one for $t \to \infty$. Therefore, in contrast to self-similar α-stable Lévy processes with infinite moments $k \geq \alpha$, $\alpha \in (0,2)$, X_t becomes more Gaussian-like in the course of process time. In other words, the univariate NTS process lies in the domain of attraction of the 2-stable self-similar Brownian motion again. For more details on self-similar α-stable processes and distributions, respectively, see Samorodnitsky and Taqqu [SaTa00] or Zolotarev [Zolo86]. However, this vanishing of skewness and excess kurtosis over time is not a disadvantage with regard to its appropriateness for modeling of stock returns per se, although already Mandelbrot and his followers[10] in their seminal work have strongly rejected the Gaussian hypothesis for those stock returns. Early studies[11] indicate that the distribution of such stock returns become more and more symmetric and mesokurtic when measured over longer time intervals compared to short measurement horizons.

Constructing the process via Brownian subordination implies that the distribution of X_t is a s-mixture or a weighted overlay of normal distributions B_s, weighted with the mixing distribution of S_t, respectively. This has two main effects on the central statistics to be

[10] See for example Mandelbrot [Mand63], [Mand67], Mandelbrot and Taylor [MaTa67] as well as Fama [Fama63], [Fama65].

[11] See for example Upton and Shannon [UpSh79].

mentioned here: First, the variance is a composition of the form

$$Var(X_t) = E(S_t) \cdot Var(B_1) + Var(S_t) \cdot E^2(B_1)\,,$$

which will become significant later on for a specific standardization procedure. Secondly, the parameter γ controls the offset rate in the overlay and is therefore responsible for adding possible skewness to the distribution

- symmetric for $\gamma = 0$,
- left-skewed for $\gamma < 0$,
- right-skewed for $\gamma > 0$,

which can moreover be deduced from equation (3.22).

3.7 The Univariate NTS Distribution

Whilst e.g. for option pricing in continuous time the detailed dynamics of the stochastic process is indispensable, for many other applications in the area of finance such as different topics in financial econometrics and portfolio management it is, by contrast, sufficient to solely focus on distributional properties of the NTS model. Thus, in this section the Normal Tempered Stable *distribution*, based on a specific marginal[12] distribution X_t of the NTS process is defined.

The *Normal Tempered Stable (NTS)* distribution is defined as the distribution of the process variable X_t of a NTS process $X_{(t)}$ in section 3.3 with parameters $(a, \lambda, \gamma, \sigma^2, \mu)$ after unit time $(t = 1)$, i.e. X_1. This distribution will be further denoted as $NTS(a, \lambda, \gamma, \sigma^2, \mu)$,

[12]While still dealing with merely univariate processes, the term 'marginal' refers to a timely cross-section X_t of the process $X_{(t)}$, whereas the entire process embodies in a formal sense the whole family of $\mathcal{F}_\tau$-measurable random variables X_τ , $\tau \geq 0$ on a filtered probability space $(\Omega, \mathcal{F}_\tau, P)$. When being focused on multivariate distributions at a later stage, 'marginal' will be used again in its more common meaning as a projection on one single dimension.

$a \in (0,2)$, $\lambda, \sigma^2 > 0$ and $\gamma, \mu \in \mathbb{R}$, with parameter set in accordance with the corresponding NTS process.

When focusing on the distribution only, the subordination scheme in equation (3.12) can be further extended and at the same time simplified to

$$\begin{aligned} X &= \mu + \gamma(S_1 - 1) + \sigma \cdot W_{(S_1)} \\ &= \mu + \gamma(S - 1) + \sigma \cdot \sqrt{S} \cdot W \,. \end{aligned} \tag{3.24}$$

Note that for the case when dealing with the random variable S_1, the time index will be further omitted, becoming S in this notation. Again, $W_{(s)}$ is denoting a standard Wiener process $BM(0,1)$, whereas $W \sim \mathcal{N}(0,1)$ stands for a standard normal distribution. Due to the self-similarity of the Wiener process with stability index $\alpha = 2$, the stochastic subordinator S_1 can be dissolved from the time index, resulting in a factor $S^{1/\alpha} = \sqrt{S}$. Hence, the subordination construction can be reduced to a multiplication of two independent random variables S and W. Although this useful relation is not of any practical importance as yet, it will, however, turn out to be exceedingly helpful in the multivariate case in chapter 4 when approaching calculation of moments and for several other tasks involved with the theoretical handling and practical application of such distributions.

In the univariate case so far, characteristic functions and distributional moments can be directly obtained by setting $t = 1$ in the corresponding equations (3.16), (3.20) – (3.23) for the marginal process distributions X_t in the preceding sections.

In the following sections 3.7.1 and 3.7.2 some important properties of this NTS distribution together with their formal verification are derived.

3.7.1 Convolution

Here, the properties of the NTS distribution concerning their convolution are studied in greater detail. For this purpose, one considers two such independently NTS distributed

random variables $X^{(1)}$ and $X^{(2)}$

$$X^{(1)} \sim NTS(a_1, \lambda_1, \gamma_1, \sigma_1^2, \mu_1)$$

$$X^{(2)} \sim NTS(a_2, \lambda_2, \gamma_2, \sigma_2^2, \mu_2)\,.$$

If one moreover defines the sum $Y = X^{(1)} + X^{(2)}$, it holds for the characteristic function[13] ϕ_Y of Y in general

$$\phi_Y(z) = \phi_{X^{(1)}+X^{(2)}}(z) = \phi_{X^{(1)}}(z) \cdot \phi_{X^{(2)}}(z) \;\Rightarrow\; \Psi_Y(z) = \Psi_{X^{(1)}+X^{(2)}}(z) = \Psi_{X^{(1)}}(z) + \Psi_{X^{(2)}}(z)$$

and for the NTS distribution in particular

$$\Psi_Y(z) = \Psi_{X^{(1)}+X^{(2)}}(z) = \frac{2\lambda_1}{a_1}\left[1 - \left(1 - \frac{1}{\lambda_1}(i\gamma_1 z - \frac{1}{2}\sigma_1^2 z^2)\right)^{\frac{a_1}{2}}\right] + i(\mu_1 - \gamma_1)z$$

$$+ \frac{2\lambda_2}{a_2}\left[1 - \left(1 - \frac{1}{\lambda_2}(i\gamma_2 z - \frac{1}{2}\sigma_2^2 z^2)\right)^{\frac{a_2}{2}}\right] + i(\mu_2 - \gamma_2)z$$

The above expression can be consolidated into a characteristic exponent of NTS structure if and only if the parameter values a_1, a_2 are equal ($a_1 = a_2 = a$) and the parameter values of $\lambda, \gamma, \sigma^2$ each have a fixed ratio

$$\frac{\lambda_1}{\lambda_2} = \frac{\gamma_1}{\gamma_2} = \frac{\sigma_1^2}{\sigma_2^2} = \frac{1}{\kappa}\,,$$

yielding

$$\Psi_Y(z) = \frac{2\lambda_1}{a}\left[1 - \left(1 - \frac{1}{\lambda_1}(i\gamma_1 z - \frac{1}{2}\sigma_1^2 z^2)\right)^{\frac{a}{2}}\right] + i(\mu_1 - \gamma_1)z$$

$$+ \frac{2\kappa\lambda_1}{a}\left[1 - \left(1 - \frac{1}{\kappa\lambda_1}(i\kappa\gamma_1 z - \frac{1}{2}\kappa\sigma_1^2 z^2)\right)^{\frac{a}{2}}\right] + i(\mu_2 - \kappa\gamma_1)z$$

[13] Note at this point, that the *characteristic exponent function (CEF)* is only valid for Lévy processes while it gets replaced by the *characteristic exponent (CE)* or the *cumulant generating function (CGF)* $\Psi_Y(z) = \ln \phi_Y(z)$ for the corresponding *infinitely-divisible (ID)* distributions.

$$= (1+\kappa)\frac{2\lambda_1}{a}\left[1-\left(1-\frac{1}{\lambda_1}(i\gamma_1 z-\frac{1}{2}\sigma_1^2 z^2)\right)^{\frac{a}{2}}\right]+i(\mu_1+\mu_2-(1+\kappa)\gamma_1)z\,.$$

Substituting the matching terms in the last equation gives rise to the new NTS parameters associated with the convolution Y. Summarizing the above operations verifies a partial closedness of the NTS distribution under convolutions in the sense that the convolution of two distributions $NTS(a,\lambda,\gamma,\sigma^2,\mu_1)$ and $NTS(a,\kappa\lambda,\kappa\gamma,\kappa\sigma^2,\mu_2)$, $\kappa>0$ leads to another $NTS(a,(1+\kappa)\lambda,(1+\kappa)\gamma,(1+\kappa)\sigma^2,(\mu_1+\mu_2))$ distribution.

To give an illustrative interpretation of this relationship, one has to note that the restriction on the parameter values is equivalent with both respective NTS processes underlying the above distributions having a Lévy density of the same shape. In other words, the second Lévy measure emerges from first one by multiplying each mass involved by a positive factor $\kappa>0$, which can be interpreted as a simple mass scaling. This manipulation in turn causes a scaling of the jump intensities of the whole jump spectrum with factor κ.

3.7.2 Linear Transformation

Another important manipulation in conjunction with univariate NTS distributions is the linear transformation. For this purpose, one considers a NTS distributed random variable $X\sim NTS(a,\lambda,\gamma,\sigma^2,\mu)$ and two coefficients $\alpha\neq 0$, $\beta\in\mathbb{R}$, which define a second random variable $Y=\alpha X+\beta$. From the universal linear transformation formula[14]

$$\phi_Y(z)=\phi_{\alpha X+\beta}(z)=\exp(i\beta z)\cdot\phi_X(\alpha z)\;\Rightarrow\;\Psi_Y(z)=\Psi_{\alpha X+\beta}(z)=i\beta z+\Psi_X(\alpha z)$$

we get by inserting the corresponding items from equation (3.16)

$$\Psi_Y(z)=iz\beta+i(\mu-\gamma)\alpha z+\frac{2\lambda}{a}\left[1-\left(1-\frac{1}{\lambda}(i\gamma\alpha z-\frac{1}{2}\sigma^2(\alpha z)^2)\right)^{\frac{a}{2}}\right]$$

$$= iz((\alpha\mu+\beta)-\alpha\gamma)+\frac{2\lambda}{a}\left[1-\left(1-\frac{1}{\lambda}(i\alpha\gamma z-\frac{1}{2}(\alpha\sigma)^2 z^2)\right)^{\frac{a}{2}}\right]\,.$$

[14] see e.g. Lukacs [Luka70]

Subsequently, similar to the approach in the preceding section, matching terms in the above equation are consolidated into new parameter values according to the proper structure of NTS cumulant generating functions. Thus, one easily recognizes that the resulting random variable Y is again NTS distributed, $Y \sim NTS(a, \lambda, \tilde{\gamma}, \tilde{\sigma}^2, \tilde{\mu})$, with modified parameters

$$\tilde{\gamma} = \alpha\gamma\,,\ \tilde{\sigma}^2 = \alpha^2\sigma^2 \tag{3.25}$$

$$\tilde{\mu} = \alpha\mu + \beta \tag{3.26}$$

while parameter values of (a, λ) stay unaffected by the linear transformation.

From the differentiated effects of such a linear transformation on the single distribution parameters one is able to conclude the following interpretation concerning the underlying NTS process. First of all, it is easily concluded that the dynamics of the subordinator involved remains unaltered. Secondly, the support of the corresponding Brownian motion is stretched by the coefficient α. Note that this is not equivalent to a time scaling of the Brownian motion with factor α^2, as the Brownian motion involved here is not a driftless one in general. At last, the additional external trend

$$\tilde{\delta} = (\tilde{\mu} - \tilde{\gamma}) = \alpha\mu + \beta - \alpha\gamma = \alpha(\mu - \gamma) + \beta = \alpha\delta + \beta$$

is likewise linearly transformed with coefficients α and β. Linear transformations constitute a helpful tool for certain applications in chapter 5, e.g. for performing linear standardizations of univariate NTS distributions.

Concerning a possible time scaling of the underlying process, one has to finally remark the following relation. Such a time scaling is only achieved with admissible convolutions presented in section 3.7.1, which are equivalent to a corresponding mass scaling of the Lévy measure. By looking at the possible manipulations of parameter values made available by either feasible convolution and linear transformation, it becomes obvious that a change in NTS process speed cannot be achieved by linear transformations and vice versa.

3.7.3 Standardized Univariate NTS Distributions

Especially when modeling situations, where the considered data is assumed to have zero mean and unit variance, it is suitable to have a standardized version of the univariate NTS distribution at hand with fewer parameters involved.

In practical modeling situations it is very often recommended to remove the individual scales inherent in the modeled data. A simple manipulation to achieve this is a linear transformation such that first and second moments are set to zero and one, respectively[15]. The aim is to, while keeping the information of the once present inherent scales, be able to use a distribution for the standardized situation with a reduced number of parameters. This approach can be seen in analogy to e.g. the normal distribution, where standardization leads from a once 2-parameter distribution class to a unique instance. Hence, standardization eliminates two free parameters, one for first and second moment, each, where both are constituting the scale of the distribution.

In principle, the following three basic approaches are available for performing such a standardization or restriction of the class of univariate NTS distributions to those instances with zero mean and unit variance, respectively:

1.) Linear transformation: X undergoes a linear transformation, the effects on the parameters can be described by equations (3.25) – (3.26).

2.) Mass scaling: Besides adjusting the external trend of the underlying process by $\mu = 0$, the jump intensities of all jump sizes of the underlying NTS process are uniformly rescaled in order to achieve a unit variance of the distribution $X \equiv X_1$ in $t = 1$.

3.) Parameter manipulation: This constitutes a rather arbitrary way of establishing restriction dependencies between selected parameters or groups of parameters in order to obtain standardized distributions, thereby reducing the number of free dimensions in the parameter set.

[15] This operation is commonly comprehended as *linear standardization* and will be referred to in a later context.

While the first two approaches are based on natural manipulations, the latter is rather arbitrary and can be designed with regard to a good level of tractability in mind. In particular, the third approach will be carried out as the two former ones would prove to be rather complicated with regard to the evolving parametrization. In general it is expedient to focus on the class of emerging standardized distributions and their associated parametrization than on the actual process of standardization itself.

Thus considering a purely parameter-based technique as the most promising approach, an obvious first step is to set $E(X) = 0$ by choosing $\mu = \tilde{\mu} = 0$. This makes sense, as this parameter μ is directly linked to the expectation of X. For the variance, however, no such single parameter is available. In fact, according to equation (3.21), it holds that

$$Var(X) = \sigma^2 + \frac{2-a}{2\lambda}\gamma^2 \tag{3.27}$$

for the variance of a $NTS(a, \lambda, \gamma, \sigma^2, \mu)$ distributed random variable X. Recall that, as can be clearly recognized from the above equation, there are two distinct sources of variance. One is the diffusion rate σ^2 of the subordinated Brownian motion, the other one is its squared trend rate γ^2, multiplied by the variance of the employed CTS subordinator given by equation (3.9).

Looking at equation (3.27), the simplest way of standardizing the variance by adjusting only one single other parameter appears to be to set

$$\sigma = \tilde{\sigma} = \sqrt{1 - \frac{2-a}{2\lambda}\gamma^2}\,.$$

In order to ensure a feasible implied Brownian motion $B_{(s)}$, i.e. $\tilde{\sigma} > 0$,

$$|\gamma| < \sqrt{\frac{2\lambda}{2-a}} \tag{3.28}$$

has to hold. Descriptively, this means that the amount of variance caused by the stochastic trend driven by γ and the subordinators stochastic dynamics, which has to be properly less than 1, is supplemented by an appropriate quantity of additional diffusion rate $\tilde{\sigma}^2$.

Note that not every instance in the NTS class is linked to a corresponding standard NTS distribution by this standardization procedure. Cases where such a standardization is not feasible are characterized by violation of the restriction in equation (3.28).

For further tasks to be accomplished, it makes sense to recognize these specific values of parameters as functions of the remaining free distribution parameters (a, λ, γ)[16]

$$\tilde{\mu}(a, \lambda, \gamma) = 0$$

$$\tilde{\sigma}(a, \lambda, \gamma) = \sqrt{1 - \frac{2-a}{2\lambda}\gamma^2}\,.$$

This immediately leads to the definition of the *standard Normal Tempered Stable (stdNTS)* distribution, by

$$\begin{aligned} stdNTS(a, \lambda, \gamma) &\equiv NTS\left(a, \lambda, \gamma, \tilde{\sigma}^2(a, \lambda, \gamma), \tilde{\mu}(a, \lambda, \gamma)\right) \\ &\equiv NTS\left(a, \lambda, \gamma, 1 - \frac{2-a}{2\lambda}\gamma^2, 0\right), \end{aligned} \tag{3.29}$$

where again $a \in (0, 2)$, $\lambda > 0$ and $\gamma \in \mathbb{R}$ in the limits given by eq. (3.28).

As basically all consequences of this procedure or definition of stdNTS could be derived in a straightforward manner by inserting the relations laid down in equation (3.29), only the most central aspects shall be presented here. In particular, the characteristic function of a standard NTS random variable $\tilde{X} \sim stdNTS(a, \lambda, \gamma)$ is simply determined by

$$\begin{aligned} \phi_{\tilde{X}}(z) &= \exp\left[\frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i\gamma z + \frac{1}{2}\tilde{\sigma}^2(a, \lambda, \gamma)z^2)\right)^{\frac{a}{2}}\right] + i(\tilde{\mu}(a, \lambda, \gamma) - \gamma)z\right] \\ &= \exp\left[\frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i\gamma z + \frac{1}{2}(1 - \frac{\gamma^2}{\lambda}(1 - \frac{a}{2}))z^2)\right)^{\frac{a}{2}}\right] - i\gamma z\right]. \end{aligned}$$

[16] As long as they obey the restriction in eq. (3.28).

The first four standardized central moments of $\tilde{X}$ are

$$E(\tilde{X}) = \tilde{\mu}(a,\lambda,\gamma) = 0$$

$$Var(\tilde{X}) = \tilde{\sigma}^2(a,\lambda,\gamma) + \frac{\gamma^2}{\lambda}(1-\frac{a}{2}) = 1 - \frac{\gamma^2}{\lambda}(1-\frac{a}{2}) + \frac{\gamma^2}{\lambda}(1-\frac{a}{2}) = 1 = c_2(\tilde{X})$$

$$s(\tilde{X}_1) = \underbrace{\frac{c_3(\tilde{X}_1)}{c_2(\tilde{X}_1)^{3/2}}}_{=1} = \left[3\frac{\tilde{\sigma}^2(a,\lambda,\gamma)\gamma}{\lambda}(1-\frac{a}{2}) + \frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2})\right]$$

$$= \left[3(1-\frac{\gamma^2}{\lambda}(1-\frac{a}{2}))\frac{\gamma}{\lambda}(1-\frac{a}{2}) + \frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2})\right]$$

$$= 3\frac{\gamma}{\lambda}(1-\frac{a}{2}) - 3\frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})^2 + \frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2})$$

$$= 3\frac{\gamma}{\lambda}(1-\frac{a}{2}) + \frac{\gamma^3}{\lambda^2}(1-\frac{a}{2})(a-1) = \frac{\gamma}{\lambda}(1-\frac{a}{2})\left[3 + \frac{\gamma^2}{\lambda}(a-1)\right]$$

$$e(\tilde{X}) = \underbrace{\frac{c_4(\tilde{X})}{c_2(\tilde{X})^{\,2}}}_{=1} = 3\frac{\tilde{\sigma}^4(a,\lambda,\gamma)}{\lambda}(1-\frac{a}{2}) + 6\frac{\tilde{\sigma}(a,\lambda,\gamma)^2\gamma^2}{\lambda^2}(1-\frac{a}{2})(2-\frac{a}{2})$$

$$+ \frac{\gamma^4}{\lambda^3}(1-\frac{a}{2})(2-\frac{a}{2})(3-\frac{a}{2})$$

$$= 3\frac{1}{\lambda}(1-\frac{\gamma^2}{\lambda}(1-\frac{a}{2}))^2(1-\frac{a}{2}) + 6\frac{\gamma^2}{\lambda^2}(1-\frac{\gamma^2}{\lambda}(1-\frac{a}{2}))(1-\frac{a}{2})(2-\frac{a}{2})$$

$$+ \frac{\gamma^4}{\lambda^3}(1-\frac{a}{2})(2-\frac{a}{2})(3-\frac{a}{2}),$$

the first two reflecting the standardization procedure just carried out.

3.7.4 Alternative Representation and Parameter Correspondence

For the sake of completeness, an alternative way for defining a NTS process and distribution is presented here. Without giving all the details involved, the existing correspondence on the level of parameterizations in both directions, is mainly pointed out. Another useful effect of this bidirectional relationship is the equivalence of these two definitions becoming clearly recognizable thereby.

While the NTS-framework developed so far is based on a process introduced by Barndorff-Nielsen and Levendorskiĭ [BNLe01], also employing Brownian motion as a construction tool for the NTS process[17], Kim et al. [KRCB08] hit upon a complemental possibility of defining this purely non-Gaussian Lévy process in an equivalent way. For this purpose, they used exponential tilting of the Lévy measure of a symmetric *Modified Tempered Stable (MTS)* process. The MTS process being a Lévy process comprised in Rosiński's [Rosi07] class of tempered stable Lévy processes was first proposed by Kim [Kim05] and has been further elaborated by Kim et. al. [KRCB08].

The general $MTS(\alpha, \mathcal{H}_+, \mathcal{H}_-, \theta_+, \theta_-, \eta)$ model refers to a purely non-Gaussian Lévy process specified by the Lévy measure[18]

$$\begin{aligned}\nu(dx) &= \nu_{MTS}(x)dx \\ &= \left(\frac{\mathcal{H}_+ \theta_+^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\theta_+ x)}{x^{\frac{\alpha+1}{2}}} \cdot \mathbb{1}_{x>0} + \frac{\mathcal{H}_- \theta_-^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\theta_- |x|)}{|x|^{\frac{\alpha+1}{2}}} \cdot \mathbb{1}_{x<0} \right) dx\,, \end{aligned} \tag{3.30}$$

with $\alpha \in (0,2)$, $\mathcal{H}_+, \mathcal{H}_-, \theta_+, \theta_- > 0$. The deterministic linear trend component is specified by a parameter $\eta \in \mathbb{R}$.

As the input for the tilting procedures serves a *symmetric* version of the MTS process *(sMTS)* with $\mathcal{H}_+ = \mathcal{H}_- = \mathcal{H} > 0$ and $\theta_+ = \theta_- = \theta > 0$. The corresponding Lévy density function $\nu_{sMTS}(x)$ is multiplied with an additional exponential function controlled by a

[17] Barndorff-Nielsen and Levendorskiĭ only consider the restricted subordination component $B_{S_{(t)}}$ without a general additional deterministic drift term.

[18] see Kim [Kim05, p. 26].

parameter β. This directly results in a new density function

$$\begin{aligned}\nu_{Tilt}(x) &= \exp(\beta x)\nu_{sMTS}(x) \\ &= \exp(\beta x)\mathcal{H}\theta^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\theta|x|)|x|^{-\frac{\alpha+1}{2}} \mathbb{1}_{x\neq 0}\,, \end{aligned} \tag{3.31}$$

while maintaining the original trend component η when compensating jumps of the entire jump spectrum. In order to be a valid Lévy density, the following condition has to be satisfied

$$\int_{|x|>1} \exp(\beta x)\nu_{sMTS}(x)dx < \infty\,,$$

which requires the parameter β to be restricted to $|\beta| < \theta$ due to the present structure of $\nu_{sMTS}(x)$ given in equation (3.30).

The above expression in equation (3.31) closely resembles the one familiar from equation (3.17) for the Lévy density function of the initially defined NTS process. Furthermore, this process additionally possesses a deterministic linear trend specified by coefficient $\eta \equiv m^{(2)}$. Processing these two non-trivial components of the Lévy triplet in the Lévy-Khintchine formula[19] leads to the following characteristic exponent function of $X_{(t)}$

$$\begin{aligned}\psi_X(z) &= iz\eta + \int_{\mathbb{R}\setminus\{0\}} (\exp(izx) - 1 - izx)\exp(\beta x)\mathcal{H}\theta^{\frac{\alpha+1}{2}} K_{\frac{\alpha+1}{2}}(\theta|x|)|x|^{-\frac{\alpha+1}{2}} dx \\ &= iz\eta - iz2^{\frac{(1-\alpha)}{2}}\sqrt{\pi}\mathcal{H}\Gamma\left(1-\frac{\alpha}{2}\right)\left[\beta\left(\theta^2-\beta^2\right)^{\frac{\alpha}{2}-1}\right] \\ &\quad + 2^{-\frac{(\alpha+1)}{2}}\sqrt{\pi}\mathcal{H}\Gamma\left(\frac{\alpha}{2}\right)\left[\left(\theta^2-(\beta+iz)^2\right)^{\frac{\alpha}{2}} - \left(\theta^2-\beta^2\right)^{\frac{\alpha}{2}}\right] \end{aligned} \tag{3.32}$$

Note that performing the above calculation requires steps similar to the ones suggested in appendix A.3 or in Kim [Kim05, p. 30 ff.] and is almost the only available possibility of obtaining the central characteristic exponent function.

This is due to the fact that this version of the NTS process is explicitly defined by its

[19] see appendix A.1, eq. (A.2).

Lévy density function, developed by tilting the symmetric MTS Lévy density function by multiplying an additional exponential function, as performed in above equation (3.31). Thus far, there is no such convenient determination formula available as is the case for a Lévy subordination scheme[20], where only simple substitutions of arguments were required to be performed. Nevertheless, another representation based on the Esscher transform

$$\psi_X(z) = \psi_{sMTS}(z-i\beta) - \underbrace{\psi_{sMTS}(-i\beta)}_{=l_{sMTS}(\beta)} + iz\left[m_X - \underbrace{m_{sMTS}}_{=0} - \int_{\mathbb{R}} y(\exp(\beta y)-1)\nu_{sMTS}(y)dy\right], \tag{3.33}$$

becomes available here[21]. This is again only mentioned for the sake of completeness as the calculations involved in solving the above integral in equation (3.33) remain burdensome. As a matter of course, the obtained result is equivalent to the expression equation (3.32), however.

From this complete representation of the Lévy process underlying the distribution in both definitions, one is able to derive the following parameter correspondence, which is the main purpose of this section. First, the parameters of the tilt approach to the NTS distribution in terms of the parameter values of the subordination approach are given.

$$\begin{aligned}
\alpha &= a \\
\mathcal{H} &= \frac{\sqrt{2}\lambda^{1-\frac{a}{2}}\sigma^a}{\sqrt{\pi}\Gamma(1-\frac{a}{2})} \\
\theta &= \sqrt{\left(\frac{\gamma}{\sigma^2}\right)^2 + \frac{2\lambda}{\sigma^2}} = \frac{\sqrt{\left(\frac{\gamma}{\sigma}\right)^2 + 2\lambda}}{\sigma} = \frac{\sqrt{\gamma^2 + 2\sigma^2\lambda}}{\sigma^2} \\
\beta &= \frac{\gamma}{\sigma^2} \\
\eta &= \mu
\end{aligned}$$

Some of the above relations can be found in Rachev et al. [RKBF11].

[20] See equations (3.13) and (3.14).
[21] See Kim et. al. [KRCB08] and the reference therein.

Consequentially, the relations in the opposite direction are as follows.

$$a = \alpha$$

$$\lambda = 2^{-\frac{(a+1)}{2}} \left[\mathcal{H}\sqrt{\pi}\Gamma\left(1-\frac{a}{2}\right)\right] \left(\theta^2 - \beta^2\right)^{\frac{a}{2}}$$

$$\gamma = 2^{\frac{1-a}{2}} \left[\mathcal{H}\sqrt{\pi}\Gamma\left(1-\frac{a}{2}\right)\right] \beta \left(\theta^2 - \beta^2\right)^{\frac{a}{2}-1}$$

$$\sigma = 2^{\frac{1-a}{4}} \left[\mathcal{H}\sqrt{\pi}\Gamma\left(1-\frac{a}{2}\right)\right]^{\frac{1}{2}} \left(\theta^2 - \beta^2\right)^{\frac{a}{4}-\frac{1}{2}}$$

$$\mu = \eta$$

Further details concerning this derivation are given in appendix B.2.

These two parameter transformations in both directions together with appropriate parameter ranges show that every NTS distribution has a valid representation in both definitions, which is a way to verify the equivalence of the two approaches compared in this context.

3.8 Motivation and Literature Review

The reason for having presented the conventional univariate normal tempered stable model in such extensive detail at this point is threefold. Firstly, the basic principle of various of the above methods such as for linear transformation or for the standardization procedure are maintained in the multivariate case, while the univariate perspective serves to illustrate its inherent mode of operation. Secondly, for later estimation and analysis of portfolios under the assumption of multivariate normal tempered stable models, one exploits the later to be proven fact, that linear combinations of dimensions likewise result in an univariate NTS distribution. Thirdly, besides linear combinations also marginal distributions of the multivariate NTS model turn out to be univariate NTS as well. Hence, most of the aspects presented and illustrated in this part possess significant further relevance in addition to their immediate application in the multivariate analysis.

In the recent past, Brownian subordination has become a popular approach for explicitly constructing univariate tempered stable (TS) processes and other alternative one-dimensional Lévy processes with semi-heavy tails. The reason for this appreciation is that the outcome of Brownian subordination has particular advantages with regard to its analytical tractability compared to those classes of purely non-Gaussian Lévy processes which are defined by an explicit tempering function for the α-stable Lévy measure. The latter group encompasses the yet familiar classes of *Classical Tempered Stable (CTS)* and *Normal Tempered Stable (NTS)* processes. Furthermore this group includes the more recent developments such as the *Modified Tempered Stable (MTS)*, the *Rapidly Decreasing Tempered Stable (RDTS)* as well as the *Kim-Rachev Tempered Stable (KR)* Lévy processes. A comprehensive overview with respect to the distributional properties can be found in the book of Rachev et al. [RKBF11, ch. 3.1], for the processes itself one refers to the original publications to be discussed below.

Among the first implementations of the concept of tempering stable distributions and processes were the Truncated Lévy Flight by Koponen [Kopo95], the KoBoL model by Boyarchenko and Levendorskiĭ [BoLe00] and the CGMY model by Carr et al. [CGMY02], which are closely related and the separate Normal Tempered Stable model introduced by Barndorff-Nielsen and Levendorskiĭ [BNLe01]. The models of the former group, of which the KoBoL model represents the most flexible one, have been unified under the term *Classical Tempered Stable* during the last decade. Since then, a greater variety of tempered stable models defined by the explicit specification of their associated tempering function emerged. One of them is the *Kim-Rachev* tempered stable model initiated by Kim et al. [KRBF08b], [KRBF08a], where [KRBF08b] moreover demonstrates the estimation methods involved and formulates a risk-neutral market model based on the KR process. In addition, [KRBF08a] covers a first application of the KR distribution for the innovations of a GARCH time series model and compares the results to alternative tempered stable distributions. Their individual estimates are assessed later on with regard to their performance in a back-test for their asset return forecasting capability. A comparing overview over its properties in relation to the full group of tempered stable classes presented here,

together with auxiliary methods for determining the distribution and risk measures, can be found in Kim et al. [KRBF09]. The possibilities for simulating KR random variables for an application in financial Monte Carlo techniques is finally studied by Bianchi et al. in [BRKF10b].

It is important to note at this point that among the tempered stable models covered in this summary, only KR and CTS are contained in Rosiński's categorization[22], while MTS, NTS and RDTS are not. His attempt of characterizing possibly multivariate tempered stable models is based on a completely monotonic tempering function for the α-stable spectral measure. To remedy this bothering incompleteness, Bianchi et al. [BRKF10a] initiated a modified category of tempered stable model classes termed *Tempered Infinitely Divisible (TID)* models, which has no overlap with the Rosiński definition and is able to capture MTS and RDTS, while NTS still fits none of the two categories. This last representative in the universe of tempered stable model classes was brought up by Barndorff-Nielsen and Levendorskiĭ [BNLe01] and plays a central role in this study. Although NTS is not contained in any of the two formal definitions of tempered models so far, it clearly exposes characteristics of tempered stable Lévy processes, which can be observed in equation (3.17) in section 3.5. Therefore, it is located at the interface of tempered stable models presented above and Lévy models based on Brownian subordination, whose historical development and interrelations are presented below, as it possesses features of both sides.

The MTS was first described by Kim [Kim05] in his PhD thesis. Its further development is achieved by Kim et al. in [KRCB08], where a change of measure theorem and a connection to the NTS by exponentially tilted Lévy measures is substantiated. An application in the context of GARCH models for financial econometrics in [KRBM$^+$11] as well as to option pricing presented in this same publication is thoroughly studied in [KRBF08a] and [KRCB09]. Moreover, [KRBF08a] finds empirical evidence for improvements compared to the application of α-stable distributions.

The last important class of tempered stable models is the *Rapidly Decreasing Tempered Stable* process mentioned by Kim et al. in [KRBF10]. In particular, besides the discussion

[22] See Rosiński [Rosi07].

of possibilities for simulation, the paper points to the excellent suitability of RDTS for option pricing. The reason for this is the existence of exponential moments up to some fractional order k, a fact which is not shared by other alternative TS instances. RDTS-GARCH models for forecasting volatile markets are recently studied by Kim et al. in [KRBM$^+$11].

Finally, a concise overview of the central properties of the five mentioned tempered stable model types together with helpful approximation techniques for distributions and risk measures based on some transformation of their characteristic function is given by Kim et al. in [KRBF09] as well as in a recently published first textbook on the general topic by Rachev et al. [RKBF11].

Basically, there are two other central classes of Lévy processes with semi-heavy tails created by Brownian subordination in the univariate case. This is the *Variance Gamma (VG)* process on the one hand and the *Normal Inverse Gaussian (NIG)* process on the other hand, together with their associated distributions. In the case of the variance gamma process, the subordinator is similar to the one from CTS class in section 3.1. The only difference is in the choice of $a = 0$, which was excluded from the range of valid parameter values in the course of NTS construction. Nonetheless, $a = 0$ implies a feasible Lévy density in equation (3.1), which in turn generates a subordinator contained in the class of *Gamma* processes, which explains the term 'variance gamma'. The VG processes were initially defined and developed by Madan and Seneta [MaSe87] and further studied by Madan et al. [MaCC98].

When choosing $a = 1$ for the CTS subordinator, a process of the *Inverse Gaussian (IG)* class evolves. Using this process as an input for the Brownian subordination leads to the class of *Normal Inverse Gaussian* processes, where the term 'normal' again indicates its origin from Brownian subordination in combination with an IG subordinator. In contrast to the limiting case of VG processes, every NIG process constitutes a particular instance within the class of NTS processes. Hence, the univariate NTS process developed in this chapter is an attempt to unite both processes in a significantly extended and more flexible framework while maintaining a reasonable degree of analytical tractability, for example

regarding estimation techniques. NIG has its origins in the work of Barndorff-Nielsen [Barn97], [Barn98] and Rydberg [Rydb97]. After the turn of the millenium, authors in the field of electrical and communications engineering began to utilize the ideas and concepts of NIG processes in their research, such as Salberg et al. [SSØH01] and Hansen and Øigård [HaØi01].

Moreover, the Gamma and the Inverse Gaussian processes are distributed according to laws which for their part are contained in class of *Generalized Inverse Gaussian (GIG)* distributions. This not true for every other instance of the CTS subordinator since GIG distributions offer the advantage of allowing for an explicit representation of their probability density function. With this desirable property at hand, every Brownian subordination obeying an GIG distributed subordinator possesses a closed-form distribution itself[23]. The set of these subordination results is termed *Generalized Hyperbolic (GH)* model and obviously embeds the VG and NIG. Due to this beneficial property of the GIG with regard to analytical tractability one is able to construct efficient methods for financial and econometric analysis. Practical implementations have already been delivered by a number of authors, among them Eberlein [Eber01], Eberlein et al. [EbKP98] and Eberlein and Raible [EbRa01] as well as Rydberg [Rydb99]. The less general VG and NIG models have been applied by Madan et al. [MaCC98], Madan and Seneta [MaSe90] and Carr and Madan [CaMa98] concerning the VG or Rydberg [Rydb97] and Barndorff-Nielsen [Barn97], [Barn98] for the NIG, respectively.

[23]For comprehension of this statement it is helpful to recall that the Brownian subordination is a mean-variance mixture or superimposition of Normal distributions with the PDF serving as a weighting density.

Chapter 4

The Multivariate Normal Tempered Stable Model

While a detailed study of the univariate Normal Tempered Stable process and distribution and their various alternative representations was given in the preceding chapter, the task of generalizing the framework to an arbitrary number of dimensions is addressed in this chapter. With this in mind, the representation of the NTS model by Brownian subordination appears to be the most appropriate for conveying the inherent structure of the NTS process to a multivariate setting. As will become clear in the subsequent sections of this chapter, Brownian subordination possesses a number of advantages over a definition via exponentially tilted multivariate tempered stable Lévy measures, above all concerning analytical tractability.

4.1 The Multivariate Normal Tempered Stable Process

In order to make a transition from a univariate to a multivariate NTS environment while maintaining an ideally simple structure one further relies on the already familiar CTS subordinator from section 3.1. This subordinator then serves as a stochastic time change to each single component of a now multivariate Brownian motion with possible dependence

between these dimensions. This setting will be termed *Type I* in contrast to a more flexible approach, where each dimension of the Brownian motion possesses his own independent CTS subordinator, termed *Type II*. Type II processes, although offering improved flexibility compared to Type I processes, come at the cost of loosing analytical tractability for most of the related manipulations. In addition, their characteristic exponent functions can no longer be stated in closed form. For this reason it is to this day still an open question whether they constitute Lévy processes with associated distributions of the corresponding process variables being therefore infinitely divisible at all[1]. Hence, if not explicitly stated otherwise, it is referred to as multivariate models of Type I in the further text.

4.1.1 The Multivariate Subordination Object

In contrast to the subordination object for the univariate case, namely a one-dimensional Brownian motion, it is now being replaced by its general form in n dimensions. For later incorporation into various determination formulas, the Brownian motion $B_{(s)}$ has to be represented as a multivariate n-dimensional Lévy process which is purely Gaussian, i.e. it contains, besides a diffusion and a trend component, no jumps. Hence, all three different compensation rules[2] are valid which implies the corresponding $m_B^{(0)}$, $m_B^{(1)}$ and $m_B^{(2)}$ to be all existent and having the same value. This external linear trend component is further denoted by $m_B^{(1)} = \gamma$, now being an arbitrary real n-dimensional vector, i.e. $\gamma \in \mathbb{R}^n$. Thus, the only remaining non-trivial component of the Lévy triplet $(A_B, \nu_B, m_B^{(1)})$ of the multivariate Brownian motion, as the jump component $\nu_B \equiv 0$, is the diffusion part A_B. It is made up by a feasible $(n \times n)$ covariance matrix[3] Σ, controlling the dispersion as well as the correlation between single process dimensions. The complete Lévy triplet is ultimately given by $(A_B, \nu_B, m_B^{(1)}) = (\Sigma, 0, \gamma)$.

The characteristic exponent function can again be derived[4] from this Lévy triplet and the

[1] This remark is made by Eberlein and Madan in [EbMa10].

[2] This includes an uncompensated jump component, associated to $m^{(0)}$. Recall from appendix A.1 that the rule of compensating only jumps of small magnitude ($m^{(1)}$) is the only universally valid one, while the rule of compensating the entire jump spectrum is associated to the variable $m^{(2)}$.

[3] Therefore having to be positive semi-definite.

[4] References for the multivariate versions of Brownian motion and Normal distributions are e.g. the

Lévy-Khintchine formula, yielding

$$\psi_B(z) = i\gamma^\mathsf{T} z - \frac{1}{2} z^\mathsf{T} \Sigma z = i\langle \gamma, z \rangle - \frac{1}{2} z^\mathsf{T} \Sigma z \,, \quad z \in \mathbb{R}^n \,,$$

and furthermore the characteristic function of the multivariate random variable B_s by the structure of the Lévy process

$$\phi_{B_s}(z) = \exp\left[s \cdot \left(i\gamma^\mathsf{T} z - \frac{1}{2} z^\mathsf{T} \Sigma z \right) \right] \,, \quad z \in \mathbb{R}^n \,.$$

Note that B_s is n-dimensional Normal distributed, $B_s \sim \mathcal{N}^n(s\gamma, s\Sigma_B)$, therefore also the argument $z \in \mathbb{R}^n$ is becoming an n-dimensional real vector now. The multivariate Normal probability density function (PDF) of B_s has closed form

$$f_B(s; x) = \frac{1}{(2s\pi)^{\frac{n}{2}} \, |\Sigma|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2s} (x - s\gamma)^\mathsf{T} \Sigma^{-1} (x - s\gamma) \right) \,, \quad x \in \mathbb{R}^n \,, \tag{4.1}$$

and will serve as an input for several determination formulas in the context of the considered Brownian subordination. Note however that above equation (4.1) is only valid for non-degenerate Brownian motions with strictly positive definite Σ, that is why this condition should be assumed to be satisfied.

4.1.2 Process Construction

A n-dimensional *multivariate Normal Tempered stable (MVNTS) process* $X_{(t)}$ is defined by subordinating a n-dimensional multivariate Brownian motion $B_{(s)}$ from section 4.1.1 by a familiar CTS subordinator $S_{(t)}$ from section 3.1, making up the actual subordination part $Y_{(t)}$. $S_{(t)}$ and $B_{(s)}$ are independent. Furthermore, an additional deterministic linear trend process $\Delta_t = \delta \cdot t$, $\delta \in \mathbb{R}^n$ is incorporated, giving

$$X_{(t)} = (X^{(1)}_{(t)}, X^{(2)}_{(t)}, \ldots, X^{(n)}_{(t)})^\mathsf{T} = Y_{(t)} + \Delta_{(t)} = B_{S_{(t)}} + \delta \cdot t \,. \tag{4.2}$$

textbooks of Mörters and Peres [MöPe10] or Anderson [Ande03], respectively, where the following representations could be taken from.

As in the univariate case, δ is set to be

$$\delta = \mu - \gamma \,,$$

and is controlled by the free parameter $\mu \in \mathbb{R}^n$. This is again chosen in order to obtain a parametrization for the external trend which is directly linked to the expectation, a fact to be illustrated later.

Therefore, the above equation (4.2) defines a process specified by the parameters $\alpha \in (0, 2)$, $\lambda > 0$ of the CTS subordinator, $\gamma \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$, with Σ being positive definite[5], of the subordination object. Finally, $\mu \in \mathbb{R}^n$ determines the external trend in combination with γ. This process, which is further denoted by $MVNTS(\alpha, \lambda, \gamma, \Sigma, \mu)$ is again a Lévy process, which can be justified by the subordinative combination of a Gaussian Lévy process and a proper Lévy subordinator.

Equation (4.2) can be extended to the following decomposition, where the stochastic trend γ and the entire variance-covariance structure is extracted from the Brownian motion. This is done by a linear mapping $\mathbb{R}^n \to \mathbb{R}^n$ with the lower triangle Cholesky factor of Σ, denoted by $\text{chol}(\Sigma)$.

$$\begin{aligned} X_{(t)} &= B_{S_{(t)}} + \delta \cdot t = \gamma S_{(t)} + Z_{S_{(t)}} + (\mu - \gamma)t \,,\ Z_{(s)} \equiv BM(0, \Sigma) \\ &= \mu t + \gamma(S_{(t)} - t) + \text{chol}(\Sigma) W_{S_{(t)}} \,,\ W_{(s)} \equiv BM(0, I) \,. \end{aligned} \tag{4.3}$$

There, $Z_{(t)}$ denotes a centered Brownian motion with covariance matrix Σ, whereas $W_{(t)}$ is a n-dimensional process with each dimension being an independent standard Wiener process.

An even more powerful representation with regard to future tasks can be achieved when further decomposing the covariance matrix Σ into its inherent vector of standard deviations

[5]For the sake of further computability, we exclude degenerated positive semi-definite matrices Σ, although they would also generate feasible processes $Y_{(t)}$.

$\sigma \in \mathbb{R}^n_{++}$ and correlation matrix[6] P

$$\Sigma = \mathrm{diag}(\sigma)\mathsf{P}\mathrm{diag}(\sigma)\,, \tag{4.4}$$

resulting in

$$\begin{aligned} X_{(t)} &= \mu t + \gamma(S_{(t)} - t) + \mathrm{chol}(\Sigma)W_{S_{(t)}} \\ &= \mu t + \gamma(S_{(t)} - t) + \mathrm{diag}(\sigma)\mathrm{chol}(\mathsf{P})W_{S_{(t)}}\,, \end{aligned}$$

where the factors in the last product can be commutated arbitrarily. Hence, using this decomposition another possible parametrization of the MVNTS process can be established, which is $MVNTS(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$. It will primarily ease the standardization procedure of associated MVNTS distributions introduced in section 4.2.1.

4.1.3 Lévy Triplet

For gaining an understanding of the properties of the MVNTS process $X_{(t)}$ which goes beyond those properties of its distributions of X_t, one has to consider the associated Lévy triplet (A_X, ν_X, m_X) in full extent. Only this Lévy triplet is able to fully reflect the instantaneous process dynamics in continuous time, which, in consequence, is responsible for the properties of the set of characteristic functions and distributions of X_t. The other way round, it is impossible in general to reconstruct the detailed representation of the process dynamics contained in the Lévy triplet from e.g. the characteristic exponent function $\psi_X(z)$, containing all the involved characteristic functions $\phi_{X_t}(z)$, and the i.i.d.-property of increments in the context of a Lévy process.

With $b_S = m_S^{(0)}$, one is able to determine each single component of the Lévy triplet $(A_Y, \nu_Y, m_Y^{(2)})$ of the subordination part $Y_{(t)}$ by applying the universal subordination for-

[6] Remark: P is the capital Rho from the Greek alphabet and is not to be confused with a probability measure P.

mula given by equations (2.4) – (2.6)

$$A_Y = b_S \cdot A_B = 0 \cdot \Sigma = 0 \tag{4.5}$$

$$\nu_Y(y) = b_S \cdot \nu_B(y) + \int_0^\infty f_B(s;y)\rho_S(s)ds = \int_0^\infty f_B(s;y)\rho_S(s)ds \tag{4.6}$$

$$\begin{aligned} m_Y^{(2)} &= b_S \cdot m_B^{(2)} + \int_0^\infty E(B_s)\rho_S(s)ds\ , \quad E(B_s) = \gamma s \\ &= \int_0^\infty \gamma s \rho_S(s)ds\,. \end{aligned} \tag{4.7}$$

The diffusion component of $Y_{(t)}$ is $A_Y = 0$ and thus also $A_X = 0$ as no further diffusion component is added to $Y_{(t)}$ by Δ_t. The calculations for the jump and the trend component are carried out in the following paragraphs.

4.1.3.1 Lévy Measure

As the Lévy measure ν_Y is absolutely continuous by [CoTa04, th. 4.3, p. 113 f.], its associated density $\nu_Y(y)$ is existing and given by the term in the above equation (4.6). Solving the integral involves the following steps.

$$\begin{aligned} \nu_Y(y) &= \int_0^\infty f_B(s;y)\rho_S(s)ds \\ &= \int_0^\infty \frac{1}{(2\pi)^{\frac{n}{2}}\,|s\Sigma|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(y-s\gamma)^\intercal (s\Sigma)^{-1}(y-s\gamma)\right) \cdot \frac{\lambda^{1-\frac{a}{2}}\exp(-\lambda s)}{\Gamma(1-\frac{a}{2})s^{\frac{a}{2}+1}}ds \\ &= \int_0^\infty \frac{1}{(2\pi)^{\frac{n}{2}}\,[s^n\,|\Sigma|]^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(y-s\gamma)^\intercal \frac{1}{s}\Sigma^{-1}(y-s\gamma)\right) \cdot \frac{\lambda^{1-\frac{a}{2}}\exp(-\lambda s)}{\Gamma(1-\frac{a}{2})s^{\frac{a}{2}+1}}ds \\ &= \frac{\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})} \int_0^\infty \exp\left(-\frac{1}{2}(y-s\gamma)^\intercal \frac{1}{s}\Sigma^{-1}(y-s\gamma) - \lambda s\right) \cdot s^{\frac{a+n}{2}+1}ds \end{aligned}$$

$$= \int_0^\infty \exp\left(-\frac{1}{2s}\left[y^\mathsf{T}\Sigma^{-1}y - y^\mathsf{T}\Sigma^{-1}\gamma s - \gamma^\mathsf{T}\Sigma^{-1}ys + \gamma^\mathsf{T}\Sigma^{-1}\gamma s^2\right] - \lambda s\right) \cdot s^{\frac{a+n}{2}+1} ds$$

$$\cdot \frac{\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})}\,.$$

Note that for the sake of compact presentation, the covariance matrix $\Sigma = \mathrm{diag}(\sigma)\mathsf{P}\mathrm{diag}(\sigma)$ is employed here instead of its decomposed representation.

Due to the term $\gamma^\mathsf{T}\Sigma^{-1}ys$ being a scalar it is identical to its transpose

$$\left(\gamma^\mathsf{T}\Sigma^{-1}ys\right)^\mathsf{T} = y^\mathsf{T}\left(\Sigma^{-1}\right)^\mathsf{T}\gamma s = y^\mathsf{T}\Sigma^{-1}\gamma s\,.$$

Moreover, Σ is symmetric as well as its inverse $\Sigma^{-1} \Rightarrow \left(\Sigma^{-1}\right)^\mathsf{T} = \Sigma^{-1}$, which is used in the following to aggregate the present cross terms.

$$\nu_Y(y) = \int_0^\infty \exp\left(-\frac{1}{2s}\left[y^\mathsf{T}\Sigma^{-1}y - 2y^\mathsf{T}\Sigma^{-1}\gamma s + \gamma^\mathsf{T}\Sigma^{-1}\gamma s^2\right] - \lambda s\right) \cdot s^{\frac{a+n}{2}+1} ds$$

$$\cdot \frac{\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})}$$

$$= \int_0^\infty \exp\left(-\left(\frac{1}{2}\gamma^\mathsf{T}\Sigma^{-1}\gamma + \lambda\right)s - \left(\frac{1}{2}y^\mathsf{T}\Sigma^{-1}y\right)\cdot\frac{1}{s}\right) \cdot s^{\frac{a+n}{2}+1} ds$$

$$\cdot \frac{\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})} \exp\left(x^\mathsf{T}\Sigma^{-1}\gamma\right) \tag{4.8}$$

The substitution $t = \left(\frac{1}{2}\gamma^\mathsf{T}\Sigma^{-1}\gamma + \lambda\right)s$ delivers

$$\int_0^\infty \exp\left(-\left(\frac{1}{2}\gamma^\mathsf{T}\Sigma^{-1}\gamma + \lambda\right)s - \left(\frac{1}{2}y^\mathsf{T}\Sigma^{-1}y\right)\cdot\frac{1}{s}\right) \cdot s^{\frac{a+n}{2}+1} ds$$

$$= \int_0^\infty \exp\left(-t - \left(\frac{1}{2}y^\mathsf{T}\Sigma^{-1}y\right)\left(\frac{1}{2}\gamma^\mathsf{T}\Sigma^{-1}\gamma + \lambda\right)\cdot\frac{1}{t}\right) \cdot \left(\frac{t}{\frac{1}{2}\gamma^\mathsf{T}\Sigma^{-1}\gamma + \lambda}\right)^{-\frac{a+n}{2}+1}$$

$$\cdot \left(\frac{1}{2}\gamma^{\mathsf{T}}\Sigma^{-1}\gamma + \lambda\right)^{-1} dt$$

$$= \left(\frac{1}{2}\gamma^{\mathsf{T}}\Sigma^{-1}\gamma + \lambda\right)^{\frac{a+n}{2}} \int_0^\infty \exp\left(-t - \left(\frac{1}{4}\left(y^{\mathsf{T}}\Sigma^{-1}y\right)\left(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma\right) + \frac{1}{2}\lambda y^{\mathsf{T}}\Sigma^{-1}y\right)\cdot\frac{1}{t}\right)$$

$$\cdot t^{-\frac{a+n}{2}+1} dt\,.$$

The integral resembles the integral representation of the modified Bessel function $K_p(z)$ of the second kind and order p with[7]

$$\frac{z^2}{4} = \frac{1}{4}\left(y^{\mathsf{T}}\Sigma^{-1}y\right)\left(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma\right) + \frac{1}{2}\lambda y^{\mathsf{T}}\Sigma^{-1}y \Rightarrow z = \left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma) + 2\lambda}$$

and

$$p + 1 = \frac{a+n}{2} + 1 \Rightarrow p = \frac{a+n}{2}$$

Then,

$$\int_0^\infty \exp\left(-t - \left(\frac{1}{4}\left(y^{\mathsf{T}}\Sigma^{-1}y\right)\left(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma\right) + \frac{1}{2}\lambda y^{\mathsf{T}}\Sigma^{-1}y\right)\cdot\frac{1}{t}\right)\cdot t^{-\frac{a+n}{2}+1} dt$$

$$= 2^{\frac{a+d}{2}+1}\left[\left(x^{\mathsf{T}}\Sigma^{-1}x\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma) + 2\lambda}\right]^{-\frac{a+n}{2}} \cdot K_{\frac{a+n}{2}}\left(\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma) + 2\lambda}\right).$$

Inserting these particular results into equation (4.8) gives

$$\nu_Y(y) = \frac{\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})}\exp\left(y^{\mathsf{T}}\Sigma^{-1}\gamma\right)\left(\frac{1}{2}\gamma^{\mathsf{T}}\Sigma^{-1}\gamma + \lambda\right)^{\frac{a+n}{2}} 2^{\frac{a+n}{2}+1}$$

$$\cdot \left[\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma) + 2\lambda}\right]^{-\frac{a+n}{2}} \cdot K_{\frac{a+n}{2}}\left(\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma) + 2\lambda}\right)$$

[7] For an extensive treatment of Bessel functions and a description of their properties, see Bell [Bell68].

$$= \frac{2\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})} \exp\left(y^{\mathsf{T}}\Sigma^{-1}\gamma\right)\left(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma+2\lambda\right)^{\frac{a+n}{2}}\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{-\frac{a+n}{4}}$$

$$\cdot\left(\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma)+2\lambda}\right)^{-\frac{a+n}{2}} \cdot K_{\frac{a+n}{2}}\left(\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma)+2\lambda}\right),$$

and finally

$$\nu_Y(y) = \frac{2\lambda^{1-\frac{a}{2}}}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}\,\Gamma(1-\frac{a}{2})} \exp\left(y^{\mathsf{T}}\Sigma^{-1}\gamma\right)\left(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma+2\lambda\right)^{\frac{a+n}{4}}$$

$$\cdot\frac{K_{\frac{a+n}{2}}\left(\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{1}{2}}\sqrt{(\gamma^{\mathsf{T}}\Sigma^{-1}\gamma)+2\lambda}\right)}{\left(y^{\mathsf{T}}\Sigma^{-1}y\right)^{\frac{a+n}{4}}}.$$

As was true for the diffusion component, $\Delta_{(t)}$ is contributing no additional jumps, the Lévy density function $\nu_X(x)$ of the entire MVNTS process $X_{(t)}$ is therefore equal to the one of $Y_{(t)}$, $\nu_Y(x)$.

4.1.3.2 Trend Component

Solving the equation for the remaining trend component is done by the following manipulations

$$m_Y^{(2)} = \int_0^\infty \gamma s\rho_S(s)ds = \int_0^\infty \gamma s\frac{\lambda^{1-\frac{a}{2}}\exp(-\lambda s)}{\Gamma(1-\frac{a}{2})s^{\frac{a}{2}+1}}ds = \gamma\frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})}\int_0^\infty \exp(-\lambda s)s^{-\frac{a}{2}}ds$$

$$= \gamma\frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})}\int_0^\infty \exp(-t)\left(\lambda^{-1}t\right)^{-\frac{a}{2}}\lambda^{-1}dt \qquad \text{(Substitution: } t=\lambda s\text{)}$$

$$= \gamma\frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})}\lambda^{\frac{a}{2}-1}\int_0^\infty \exp(-t)t^{-\frac{a}{2}}dt = \gamma\frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})}\lambda^{\frac{a}{2}-1}\Gamma\left(1-\frac{a}{2}\right) = \gamma$$

When looking at the structure given in equation 4.3, this result seems to appear rather natural. In this case, $\Delta_{(t)}$ has to be explicitly incorporated, which evokes a non-trivial

effect on the trend component. Ultimately, it can be concluded that

$$m_X^{(2)} = m_Y^{(2)} + (\mu - \gamma) = \gamma + (\mu - \gamma) = \mu\,.$$

Again, this last result could be verified by comparing it to the multivariate expectation of X_t in equation (4.13) and the fact that $m_X^{(2)}$ due to full compensation encompasses the entire drift of the process $X_{(t)}$. With this last remaining trend component, the derivation of the entire Lévy triplet of $X_{(t)}$ is completed.

4.1.4 Characteristic Exponent Function

When only the distributional properties of the MVNTS process are of relevance, the entire situation is sufficiently captured by the characteristic functions $\phi_{X_t}(z)$. These can be derived in a simple fashion for every point in time t, $t > 0$ from the time-independent characteristic exponent function $\psi_X(z)$, being time-independent and universal for the whole process $X_{(t)}$. This characteristic exponent function can either be generally derived from the entire Lévy triplet (a_X, ν_X, m_X) associated with the MVNTS process $X_{(t)}$ by carrying out the multivariate Lévy-Khintchine formula

$$\psi_X(z) = iz^\mathsf{T} m_X^{(2)} - \frac{1}{2} z^\mathsf{T} A_X z + \int_{\mathbb{R}^n \setminus \{0\}} (\exp(iz^\mathsf{T} x) - 1 - iz^\mathsf{T} x)\nu_X(x)dx\,,\ z \in \mathbb{R}^n$$

or, in the context of subordinations, by means of the universal subordination formula presented in equation (2.3). While the former approach is rather burdensome in the presence of n dimensions, an alternative and straightforward approach is available. It only requires some simple input items which are present in closed-form representations

$$\begin{aligned}\psi_X(z) &= l_S(\psi_B(z)) + \psi_\Delta(z)\,,\ z \in \mathbb{R}^n \\ &= \frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}\left(i\gamma^\mathsf{T} z - \frac{1}{2} z^\mathsf{T} \Sigma z\right)\right)^{\frac{a}{2}}\right] + i(\mu - \gamma)^\mathsf{T} z\,.\end{aligned}$$

The following transition to the, then time-dependent, characteristic function of the individual distribution of a MVNTS process random variable X_t is simply made by

$$\phi_{X_t}(z) = \exp\left[t \cdot \psi_X(z)\right] = \exp\left[\Psi_{X_t}(z)\right] \tag{4.9}$$

$$= \exp\left[t \cdot \left(\frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}\left(i\gamma^{\mathsf{T}}z - \frac{1}{2}z^{\mathsf{T}}\Sigma z\right)\right)^{\frac{a}{2}}\right] + i(\mu-\gamma)^{\mathsf{T}}z\right)\right],\ z \in \mathbb{R}^n\,. \tag{4.10}$$

The corresponding moment generating function $M_{X_t}(u)$ is not necessarily required for further issues. For the sake of completeness it is mentioned here that

$$M_{X_t}(u) = \phi_{X_t}(-iu)\,,\ u \in G = (-\infty, \lambda]^n$$

on this feasible domain G.

The characteristic function could be an initial point for the determination of probability density and cumulated distribution functions[8] of X_t, $t > 0$ of the MVNTS process variable $X_{(t)}$. For general distributions, this can be universally achieved by appropriate CF inversion theorems[9]. However, these inversion theorems, as for not being analytically solvable in the considered setting, require numerical integration methods in an arbitrary number of dimensions n. Hence, an alternative approach for tackling this task will be presented in section 4.2.1.3 which aims at reducing the complexity of the calculations and thereby enabling efficient numerical approximations.

4.1.5 Moments

As was argued in the preceding section, the probability distributions of process variables X_t in form of their corresponding PDF or CDF, cannot be stated in a closed solution and the analytic expression of the characteristic functions offers little illustrative insight with regard to the actual shape of the distribution. Thus, at this point the first four cumulants

[8] Further abbreviated by PDF and CDF, respectively.
[9] see e.g. Shephard [Shep91a], [Shep91b]

of the distribution together with the related statistical measures should be given. They try to fill the gap and should at least allow for a rough impression of the joint distribution, namely of the shape of marginals and their dependence structure, when no closed form representation is available.

Before this is practically done, one has to mention the following. When considering a multivariate environment, it is helpful to distinguish between moments or cumulants, respectively, of marginal distributions and those of the joint distribution. For the latter, moments and cumulants of k-th order form tensors, their definitions as well as a detailed description can be found in e.g. Bilodeau and Brenner [BiBr99] or McCullagh [McCu87]. With regard to the calculation of cumulants in particular, one notices that the method remains unalteredly based on derivatives of the cumulant generating function around 0. But as cumulants and even raw moments are influenced by several distributional features simultaneously, one would be interested in finding a multivariate analogy to the statistics skewness and kurtosis. Henze and Klar in [HeKl02] give an overview over present definitions for these two statistics for multivariate distributions of n-dimensional random vectors. However, by their scalar nature they are obviously not able to fully represent the features of the multivariate distribution in a complete fashion. Nevertheless, they can be used as an indicator for deviation of Gaussianity of the distribution.

Note furthermore, that the multivariate central moment of second order ($k = 2$) is the familiar covariance matrix which captures the dispersion of each dimension in addition to the linear dependence structure. For the critical features of multivariate financial data, however, like asymmetry, heavy-tails and non-linear dependence structures, corresponding tensors for $k \geq 3$ would be required, which are less straightforward to obtain and to display in a compact way manner.

Therefore, the presentation of moments is reasonably restricted to $k < 3$ in the joint multivariate case. Due to this restriction, the conventional derivatives technique[10] can be moreover circumvented by relying on a suitable extended decomposition of the random vector based on the subordination scheme in equation (4.3). For fixed $t > 0$, this equation

[10]The method is presented in Lukacs [Luka70].

can be transformed into

$$X_t = \mu t + \gamma(S_t - t) + \text{diag}(\sigma)\text{chol}(\mathsf{P})W_{S_t}\,,\ W_{(s)} \equiv BM(0, I)$$

$$= \mu t + \gamma(S_t - t) + \text{diag}(\sigma)\text{chol}(\mathsf{P})\sqrt{S_t}W\,, \quad W \sim \mathcal{N}^n(0, I) \tag{4.11}$$

$$= \mu t + \gamma(S_t - t) + \sqrt{S_t}Z\,, \quad Z \sim \mathcal{N}^n(0, \Sigma)\,. \tag{4.12}$$

In essence, when considering static distributions instead of the dynamic process perspective, the stochastic time index S_t can be once again dissolved from the subordinated Brownian motion, which is self-similar[11] with scale coefficient $\alpha = \frac{1}{2}$, and turned into a factor $\sqrt{S_t}$. In this context, W denotes a n-dimensional Normal distributed random variable $W \sim \mathcal{N}^n(0, I)$ with standardized marginals and pairwise independence between any two dimensions. At last, the expression in equation (4.12) with unseparated covariance is further utilized here for purpose of compact display.

The case $k = 1$ corresponds to the conventional vector of expectation

$$E(X_t) = E\left[\mu t + \gamma(S_t - t) + \sqrt{S_t}Z\right]$$

$$= \mu t + \gamma E(S_t - t) + E(\sqrt{S_t}Z)$$

$$= \mu t + \gamma\left[E(S_t) - t\right] + E(\sqrt{S_t}) \cdot E(Z)$$

$$= \mu t + \gamma\left[t - t\right] + E(\sqrt{S_t}) \cdot 0 = \mu t \tag{4.13}$$

calculated by exploiting the linearity the expectation operator.

For $k = 2$, as already mentioned above, we face the matrix $Cov(X_t)$ of pairwise covariances of dimensions of X_t, containing the elements $Cov(X_t^{(k)}, X_t^{(l)})$. They can be determined by applying the familiar manipulation rules for covariances combined with the fact that S_t

[11] The property of self-similarity of stable Lévy processes is explained in appendix A.2.

and Z are independent.

$$\begin{aligned} Cov(X_t^{(k)}, X_t^{(l)}) &= Cov(\mu_k t + \gamma_k(S_t - t) + \sqrt{S_t}Z_k, \mu_l t + \gamma_l(S_t - t) + \sqrt{S_t}Z_l) \\ &= Cov(\gamma_k S_t + \sqrt{S_t}Z_k, \gamma_l S_t + \sqrt{S_t}Z_l) \\ &= \underbrace{Cov(\gamma_k S_t, \gamma_l S_t)}_{①} + \underbrace{Cov(\gamma_k S_t, \sqrt{S_t}Z_l)}_{②} + \underbrace{Cov(\sqrt{S_t}Z_k, \gamma_l S_t)}_{③} \\ &\quad + \underbrace{Cov(\sqrt{S_t}Z_k, \sqrt{S_t}Z_l)}_{④} \end{aligned}$$

For the four above terms ① – ④ we get

$$\text{ad ①} \ : \ Cov(\gamma_k S_t, \gamma_l S_t) = \gamma_k \gamma_l Var(S_t) = \gamma_k \gamma_l \frac{1}{\lambda}(1 - \frac{a}{2}) \cdot t$$

$$\begin{aligned} \text{ad ②} \ : \ Cov(\gamma_k S_t, \sqrt{S_t}Z_l) &= \gamma_k \left[E(S_t^{3/2} Z_l) - E(S_t) E(\sqrt{S_t} Z_l) \right] \\ &= \gamma_k \left[E(S_t^{3/2}) \underbrace{E(Z_l)}_{=0} - E(S_t) E(\sqrt{S_t}) \underbrace{E(Z_l)}_{=0} \right] = 0 \end{aligned}$$

$$\text{ad ③} \ : Cov(\sqrt{S_t}Z_k, \gamma_l S_t) = 0 \quad \text{(symmetric to ②)}$$

$$\begin{aligned} \text{ad ④} \ : \ Cov(\sqrt{S_t}Z_k, \sqrt{S_t}Z_l) &= E(S_t Z_k Z_l) - E(\sqrt{S_t}Z_k) E(\sqrt{S_t}Z_l) \\ &= \underbrace{E(S_t)}_{=t} E(Z_k Z_l) - E(\sqrt{S_t}) \underbrace{E(Z_k)}_{=0} E(\sqrt{S_t}) \underbrace{E(Z_l)}_{=0} \\ &= t \cdot E(Z_k Z_l) \ , \quad E(Z_k Z_l) = Cov(Z_k, Z_l) + \underbrace{E(Z_k)E(Z_l)}_{=0} \\ &= t \cdot Cov(Z_k, Z_l) = t \cdot \varsigma_{k,l} \ , \end{aligned}$$

with $\Sigma = (\varsigma_{k,l})_{k,l=1,\dots,n}$ and $\sigma_k = \sqrt{\varsigma_{k,k}}$, $k = 1, \dots, n$, respectively.

To sum up the results of the above calculations, we have that

$$E(X_t) = \mu \cdot t$$

$$Var(X_t^{(k)}) = Cov(X_t^{(k)}, X_t^{(k)}) = t \cdot \left[\frac{1}{\lambda} \gamma_k^2 (1 - \frac{a}{2}) + \varsigma_{k,k} \right] \tag{4.14}$$

$$Cov(X_t^{(k)}, X_t^{(l)}) = t \cdot \frac{1}{\lambda} \gamma_k \gamma_l (1 - \frac{a}{2}) + t \cdot \varsigma_{k,l} \, .$$

In matrix notation, the covariance matrix $Cov(X_t)$ can be displayed in a more compact manner as

$$Cov(X_t) = \left(Cov(X_t^{(k)}, X_t^{(l)}) \right)_{k,l=1,\ldots,n} = t \left[\frac{1}{\lambda} (1 - \frac{a}{2}) \gamma \gamma^\mathsf{T} + \Sigma \right] \tag{4.15}$$

If one abstains from the joint distribution of X_t and focuses on the marginal distributions of $X_t^{(k)}$ only, it is useful to note that they obey a univariate NTS distribution with parameter tuple $(a, \lambda, \gamma_k, \varsigma_{k,k}, \mu_k)$. This becomes apparent by looking at the subordination scheme in equation (4.11) and further recalling the fact that every dimension of $B_{(s)}$ is subordinated by the same process S_t. Thus, their univariate moments up to order $k = 4$ and associated statistics such as mean, variance, skewness and excess kurtosis follow from the above parameter values and equations (3.22) – (3.23).

4.2 The Multivariate Normal Tempered Stable Distributions

For creating the multivariate Normal Tempered Stable distribution, the convention already familiar from the univariate setting is followed again. Recall from section that for defining the actual NTS distribution, the NTS process variable X_1 after a unit time interval was employed. In the same manner, we first consider a n-dimensional random vector $Y = (Y_1, Y_2, \ldots, Y_n)^\mathsf{T}$ being generated by a multivariate MVNTS process $X_{(t)}$ with proper

tuple of corresponding parameter values $(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$[12] at time $t = 1$, i.e. $Y \overset{d}{\sim} X_1$[13].

Then, the distributional properties of Y can be obtained by setting $t = 1$ in equations (4.10) and (4.13) – (4.15). Besides the characteristic function, we have in particular

$$E(Y) = \mu$$

$$Cov(Y) = \left[\frac{2-a}{2\lambda}\gamma\gamma^{\mathsf{T}} + \Sigma\right] = \frac{2-a}{2\lambda}\gamma\gamma^{\mathsf{T}} + \mathrm{diag}(\sigma)\mathsf{P}\mathrm{diag}(\sigma)\,.$$

The subsequent procedure might appear rather counterintuitive at first sight, that is why we deliver the following motivation beforehand. In contrast to the commonly chosen approach of starting with the definition of a more general class of distributions first and afterwards turning to a standardized version by setting certain constraints onto the parameter set involved, we decide to proceed just the other way round. Specifically, we first define a multivariate standard Normal Tempered Stable distribution before we expand the definition by translating and rescaling this standardized distribution by means of a simple linear transformation. The reason for this seemingly unconventional approach will become evident when turning to parameter estimation and further application in the framework of an ARMA-GARCH time series model in chapters 5 and 6.

4.2.1 The Multivariate Standard Normal Tempered Stable Distribution

With this in mind, an analogous approach as for the univariate NTS distribution is performed for Y, where the external deterministic trend component controlled by μ and the diffusion component of the Brownian motion specified by σ^2 were adjusted, ceteris paribus, in order to meet the required conditions of a standardized distribution.

[12] Feasible ranges of parameter values were also given in section 4.1.2.

[13] $\overset{d}{\sim}$ indicates equivalence in distribution.

4.2.1.1 Construction

This corresponds to setting

$$\mu = \tilde{\mu} := (0, \ldots, 0)^{\mathsf{T}},$$

which centers the distribution of Y around 0. The next task concerns ensuring that marginal distributions $Y_k, k = 1, \ldots, n$ of Y have unit variance. In order to do so in a systematical way the current values of a, λ, γ are initially assumed to be arbitrary but constant, while only the diffusion matrix Σ belonging to the Brownian motion is set as a design variable. There, the advantage of factorizing the covariance matrix Σ into its inherent correlations P and standard deviations σ, $\Sigma = \mathrm{diag}(\sigma)\mathsf{P}\mathrm{diag}(\sigma)$, $\sigma \in \mathbb{R}^n_{++}$, P a feasible $(n \times n)$-matrix, performed in section 4.1.2, equation (4.4), ultimately becomes apparent. In other words, the scale is separated from the scale-independent correlation structure for the Brownian motion B. Otherwise it would be a rather difficult task to properly choose $\tilde{\Sigma}$ in an unequivocal way, such that $Cov(Y)$ retains only values of 1 on its diagonal, as it still offers a high number of degrees of freedom and the property of positive definiteness has to be maintained during this operation.

A proposition for achieving this in a unique way is the possibility of now being able to adjust the single components of σ in an appropriate manner while keeping P constant. To be more precise, σ is set to $\tilde{\sigma} = (\tilde{\sigma}_1, \tilde{\sigma}_2, \ldots, \tilde{\sigma}_n)$ with

$$\tilde{\sigma}_k = \sqrt{1 - \frac{2-a}{2\lambda}\gamma_k^2}\,,\ k = 1, \ldots, n\,, \tag{4.16}$$

For having $\tilde{\sigma}_k$ well-defined in (4.16), again

$$|\gamma_k| < \sqrt{\frac{2\lambda}{2-a}} \tag{4.17}$$

has to hold for every $k = 1, \ldots, n$, in order to ensure

$$\frac{2-a}{2\lambda}\gamma_k^2 \overset{!}{\in} (0, 1)\,.$$

With furthermore $t = 1$, the random variable $\tilde{X}$ is subsequently constructed by inserting the corresponding $(\tilde{\mu}, \tilde{\sigma})$ into equation (4.11)

$$\begin{aligned}\tilde{X} &= \tilde{\mu} + \gamma(S_1 - 1) + \text{diag}(\tilde{\sigma})\sqrt{S_1}V \\ &= \gamma(S - 1) + \text{diag}(\tilde{\sigma})\sqrt{S}V \\ &= \gamma(S - 1) + \text{diag}(\tilde{\sigma})\sqrt{S}\text{chol}(\mathsf{P})W \qquad (4.18) \\ V \sim \mathcal{N}^n(0, \mathsf{P})\,,\ W &\sim \mathcal{N}^n(0, I)\,,\end{aligned}$$

where the time index of S_1 will be omitted from now on in this context. In (4.18), P has to be a valid $(n \times n)$ correlation matrix and $\gamma \in \mathbb{R}^n$. Again, S is the previous subordinator from the CTS class with parameters (a, λ)[14]. Finally, $\tilde{X}$ is said to be *multivariate standard Normal Tempered Stable* distributed, which is denoted by $stdMVNTS(a, \lambda, \gamma, \mathsf{P})$.

4.2.1.2 Moments and Characteristic Function

Consequentially, considering the specified parameter values $(\tilde{\mu}, \tilde{\sigma})$ to be included in the parameter tuple of a corresponding MVNTS process, in addition to $t = 1$, equations (4.13) – (4.15) exemplarily states the central moments of first and second order of $\tilde{X} \sim stdMVNTS(a, \lambda, \gamma, \mathsf{P})$

$$E(\tilde{X}) = (0, \ldots, 0)^{\mathsf{T}} \qquad (4.19)$$

$$Var(\tilde{X}) = (1, \ldots, 1)^{\mathsf{T}} \qquad (4.20)$$

$$Cov(\tilde{X}) = (Cov(\tilde{X}_j, \tilde{X}_k))_{j,k=1,\ldots,n} = \text{diag}(\tilde{\sigma})\mathsf{P}\text{diag}(\tilde{\sigma}) + \frac{2-a}{2\lambda}\gamma\gamma^{\mathsf{T}}\,, \qquad (4.21)$$

[14] see section 3.1.

where the expressions for skewness and excess kurtosis could likewise be obtained from equation (3.22) and (3.23), however.

The same applies to the characteristic function in equation (4.10), yielding

$$\phi_{X_t}(z) = \exp\left(\frac{2\lambda}{a}\left[1-\left(1-\frac{1}{\lambda}\left(i\gamma^{\mathsf{T}}z-\frac{1}{2}z^{\mathsf{T}}\mathrm{diag}(\tilde{\sigma})\mathsf{P}\mathrm{diag}(\tilde{\sigma})z\right)\right)^{\frac{a}{2}}\right]-i\gamma^{\mathsf{T}}z\right),\ z\in\mathbb{R}^n$$

$$\tilde{\sigma} = \left[\sqrt{1-\frac{2-a}{2\lambda}\gamma_k^2}\right]_{k=1,\ldots,n}^{\mathsf{T}}.$$

4.2.1.3 PDF and CDF

Although the characteristic function of the multivariate standard NTS is known in closed form, it possesses only very limited utility for the task of determining probability density function (PDF) $f_{\tilde{X}}(x)$ and cumulative distribution function (CDF) $F_{\tilde{X}}(x)$ of stdMVNTS distributions, both of which have no closed form themselves. As inversion formulas of the characteristic function, here only the equation for the PDF[15],

$$f_{\tilde{X}}(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{(-iz^{\mathsf{T}}x)}\phi_{\tilde{X}}(z)dz \tag{4.22}$$

are moreover hard to evaluate numerically in n dimensions, a different approach is chosen here. Thereby, one exploits the fact that, caused by the construction through subordination, $\tilde{X}$ is a mixture of multivariate Normal distributions composed by the distribution of subordinator S.

By this superimposition we have that both the multivariate PDF and CDF of $\tilde{X}$ can be expressed as

$$f_{\tilde{X}}(x) = \int_0^\infty f_{\bar{B}}(s;x)f_S(s)ds \tag{4.23}$$

[15]See e.g. Shephard [Shep91a], [Shep91b]. The corresponding formula for the multivariate CDF can be found therein.

$$F_{\tilde{X}}(x) = P(\tilde{X}_1 \leq x_1, \ldots, \tilde{X}_n \leq x_n) = \int_0^\infty G_{\bar{B}}(s;x) f_S(s) ds\,, \tag{4.24}$$

respectively. The PDF of S denoted by f_S serves as the weighting density for the above overlay. It is generated by the one-dimensional inversion of its characteristic function $\phi_S(z)$ with the Fourier transform

$$f_S(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-izs)\phi_S(z) dz\,, \tag{4.25}$$

which does not possess an analytical solution either. The term $G_{\bar{B}}(s;x)$ stands for the CDF in point x of a n-dimensional Normal distribution generated by the shifted Brownian motion $\bar{B}_{(s)} = B_{(s)} - \gamma$ at the time s

$$G_{\bar{B}}(s;x) \;=\; \int_{-\infty}^{\frac{x_1-(s-1)\gamma_1}{\sqrt{s}\tilde{\sigma}_1}} \ldots \int_{-\infty}^{\frac{x_n-(s-1)\gamma_n}{\sqrt{s}\tilde{\sigma}_n}} f_Z(z_1, \ldots, z_n) dz_1 \ldots dz_n\,, \tag{4.26}$$

derived from the PDF f_Z of a corresponding multivariate standard Normal distribution with correlation matrix equal to P. The shifting of the Brownian motion is due to $\mu = \tilde{\mu} = 0$ in the case of stdMVNTS distributions.

In the same way, $f_{\bar{B}}(s;x)$ denotes a multivariate Normal density function of $\bar{B}_s$, finally resulting in the PDF of the stdMVNTS distributed random vector $\tilde{X}$

$$f_{\tilde{X}}(x) = M_{a,\lambda,\gamma,\mathsf{P}} \int_{\mathbb{R}} \exp(-izs)\phi_S(z) \int_0^\infty s^{-\frac{n}{2}} \exp\left(-\frac{1}{2s}(x - \gamma s + \gamma))^{\mathsf{T}} L (x - \gamma s + \gamma)\right) ds dz$$

with

$$M_{a,\lambda,\gamma,\mathsf{P}} = (2\pi)^{-(1+\frac{n}{2})} \cdot \left[|\mathsf{P}| \cdot \prod_{k=1}^{n} \left(1 - \frac{2-a}{2\lambda}\gamma_k^2\right)\right]^{-\frac{1}{2}}$$

$$L = \mathrm{diag}(\tilde{\sigma})^{-1} \mathsf{P}^{-1} \mathrm{diag}(\tilde{\sigma})^{-1}\,.$$

This approach for the PDF enjoys the advantage of not having to numerically solve a n-dimensional integral involved with the inversion procedure of the characteristic function in equation (4.22). Instead, one is able to superimpose $f_{\bar{B}}(s;x)$, available in analytical

form with a weighting density function given by $f_S(s)$, whereas the latter one is without explicit representation. In case of the CDF in equation (4.24), even $G_{\bar{B}}(s;x)$ has no closed form solution but efficient algorithms contained in commercial software packages deliver precise numerical evaluations.

The determination of PDF and CDF was demonstrated for the restricted example of stdMVNTS distributions, although the same principle would apply for more general multivariate NTS distributions as well. Anticipating the application of multivariate NTS distributions in the context of portfolio analysis performed in chapter 5, one has to note that the handling of actual n-dimensional multivariate NTS distributions is totally avoided by only having to consider linear combinations for the intended purpose. Therefore, direct evaluation of neither PDF nor CDF of multivariate NTS distributions is necessary for this matter.

4.2.2 Linear Rescaling and MVNTS Distribution

In order to being no longer restricted to such multivariate distributions with standardized marginals only, one proceeds to one of several conceivable definitions of a general version of the MVNTS distribution with arbitrary first and second moment of their marginals. As has already been indicated, a definition of a general MVNTS distribution exhibiting useful properties, though not being overly intuitive, is obtained by linear rescaling of the above stdMVNTS distribution. An alternative suggestion for this definition is proposed and discussed subsequently.

A general *multivariate Normal Tempered Stable (MVNTS) distributed* random vector X is thus defined by a n-dimensional multivariate standard NTS distributed random vector $\tilde{X} \sim stdMVNTS(a, \lambda, \gamma, \mathsf{P})$ and the following simple linear transformation with coefficient vectors c and d,

$$X = \mathrm{diag}(c)\tilde{X} + d \tag{4.27}$$

where $d = (d_1, d_2, \ldots, d_n)^{\mathsf{T}} \in \mathbb{R}^n$, $c = (c_1, c_2, \ldots, c_n)^{\mathsf{T}}$, $c_k > 0, k = 1, 2, \ldots, n$. This distribution is further denoted by $X \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$.

Note that only a component-wise linear transformation is used in above equation (4.27) instead of a general one. Thus, one directly obtains the basic statistics of the the MVNTS distributed random vector X

$$E(X) = E(\tilde{X}) + d = d$$

$$Cov(X_j, X_k) = c_j c_k \left(\frac{2-a}{2\lambda} \gamma_j \gamma_k + \tilde{\sigma}_j \tilde{\sigma}_k \rho_{j,k} \right) \ , \ j,k = 1,2,\ldots,n$$

with familiar

$$\tilde{\sigma}_k = \sqrt{1 - \frac{2-a}{2\lambda}\gamma_k^2} \, , \, k = 1,2,\ldots,n$$

and compact matrix notation

$$\begin{aligned} Cov(X) &= \mathrm{diag}(c) Cov(\tilde{X}) \mathrm{diag}(c) \\ &= \mathrm{diag}(c) \left[\frac{2-a}{2\lambda} \gamma\gamma^{\mathsf{T}} + \mathrm{diag}(\tilde{\sigma}) \mathsf{P} \mathrm{diag}(\tilde{\sigma}) \right] \mathrm{diag}(c) \end{aligned}$$

$$Var(X) = (c_1^2, c_2^2, \ldots, c_n^2)^{\mathsf{T}} \, .$$

With the same argument of simple linear transformations, also skewness and excess kurtosis of the marginals associated with this MVNTS distribution could be determined. This would be achieved by the properties of these two particular univariate statistical measures with regard to one-dimensional linear transformations. At this point, only the main idea is illustrated while details of the computation and their results are omitted, however. Finally, it is trivial to recognize that a $stdMVNTS(a, \lambda, \gamma, \mathsf{P})$ distribution is equivalent to a $MVNTS(a, \lambda, \gamma, \mathsf{P}, \mathbf{1}, \mathbf{0})$ distribution with $\mathbf{1} = (1, \ldots, 1)^{\mathsf{T}}$ and $\mathbf{0} = (0, \ldots, 0)^{\mathsf{T}}$. This embedding of stdMVNTS in the set of MVNTS distributions will turn out to be useful in later applications.

Since every representation or parametrization, respectively, has its individual advantages

and shortcomings, the above definition via rescaling of standardized MVNTS is no exception, a complementing and more intuitive approach together with its inherent parametrization for the class of MVNTS should be presented. It is based on the idea of skipping the obscuring intermediate step via stdMVNTS. In doing so, the above random vector[16] $Y \equiv Y_1$ of a MVNTS process $Y_{(t)}$ with parameters $(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$ at time $t = 1$ is said to be $MNTS(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$ distributed, $Y \sim MNTS(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$. Associated moments, other statistical measures and the characteristic function immediately follow by inserting these parameters together with $t = 1$ into equations (4.13) – (4.15) and (4.10), respectively.

While the first principal definition and parametrization possesses advantages concerning the expression of moments as well as with regard to tractability for later parameter estimation and application in the context of portfolio analysis, the second one allows better interpretability of its incorporated building blocks in addition to improved presentability of the characteristic function. Furthermore, it will turn out to be more convenient for theoretical considerations.

In order to verify the equivalence of both classes, $X \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$ and $Y \sim MNTS(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$, we derive their parameter correspondence in the following. In other words, we give the MNTS parameter tuple $(a^*, \lambda^*, \gamma^*, \sigma^*, \mathsf{P}^*, \mu^*)$ of a random variable $X \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$. Comparing the characteristic function of X, obtained by the universal linear transformation property of characteristic functions

$$\phi_X(z) = \phi_{\mathrm{diag}(c)\tilde{X}+d}(z) = \phi_{\tilde{X}}(\mathrm{diag}(c)^{\mathsf{T}} z) \cdot \exp(id^{\mathsf{T}} z) = \exp(\Psi_X(z)), \; z \in \mathbb{R}^n$$

$$\Psi_X(z) = id^{\mathsf{T}} z + \frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i\gamma^{\mathsf{T}}\mathrm{diag}(c)z - \frac{1}{2}\left(\mathrm{diag}(c)z\right)^{\mathsf{T}} \mathcal{S}\mathrm{diag}(c)z)\right)^{\frac{a}{2}}\right] - i\gamma^{\mathsf{T}}\mathrm{diag}(c)z$$

$$\mathcal{S} = \mathrm{diag}\left(\left[\sqrt{1 - \frac{1}{\lambda}\left(1 - \frac{a}{2}\right)\gamma_k^2}\right]_{k=1,\ldots,n}\right) \mathsf{P}\mathrm{diag}\left(\left[\sqrt{1 - \frac{1}{\lambda}\left(1 - \frac{a}{2}\right)\gamma_k^2}\right]_{k=1,\ldots,n}\right),$$

[16] See introduction of section 4.2.

with the structure of the MNTS characteristic function[17], we get

$$\gamma^* = \mathrm{diag}(c)\gamma \tag{4.28}$$

$$\sigma^* = \mathrm{diag}(c)\tilde{\sigma} = \mathrm{diag}(c)\left[\sqrt{1 - \frac{1}{\lambda}\left(1 - \frac{a}{2}\right)\gamma_k^2}\right]_{k=1,\dots,n} \tag{4.29}$$

$$\mu^* = d\,, \tag{4.30}$$

through a consolidation of appropriate terms into new corresponding MNTS parameter values. The parameters (a^*, λ^*) of the former subordinator dynamics as well as the correlation matrix P^* of the employed Brownian motion however remain unaltered during this parameter conversion[18].

Its inverse[19] is embodied by the following relation

$$c = \left[\sqrt{\frac{2-a^*}{2\lambda^*}(\gamma_k^*)^2 + (\sigma_k^*)^2}\right]_{k=1,\dots,n}$$

$$\gamma = \left[\left(\frac{2-a^*}{2\lambda^*} + \frac{(\sigma_k^*)^2}{(\gamma_k^*)^2}\right)^{-1/2}\right]_{k=1,\dots,n} = \left[\gamma_k^* \cdot \sqrt{\frac{2\lambda^*}{(2-a^*)(\gamma_k^*)^2 + 2\lambda^*(\sigma_k^*)^2}}\right]^{\mathsf{T}}_{k=1,\dots,n}$$

with trivially $a = a^*$, $\lambda = \lambda^*$, $\mathsf{P} = \mathsf{P}^*$ and $d = \mu^*$. As both the parameter conversion and its inverse are bijective on their respective parameter spaces, the sets of existing MVNTS and MNTS distributions are congruent and differ only in their individual parametrization. This relation will prove essential on different occasions in subsequent chapters.

4.2.3 Linear Transformation

In this section, general linear transformations of all stdMVNTS, MVNTS and MNTS distributions are covered in full detail. The reason for this meticulous attention is that they constitute an essential tool for the operations performed in the context of parameter esti-

[17] The MNTS-CF is furthermore obtained by setting $t = 1$ in equation (4.10).
[18] MVNTS $\rightarrow$ MNTS
[19] That is MNTS $\rightarrow$ MVNTS. Details on the derivation can be found in appendix B.2.

mation and portfolio analysis in the forthcoming chapters. The effects of linear mappings are to be stated in terms of associated parameter values, from which e.g. moments and characteristic functions can be immediately resolved afterwards. In particular, as none of these three distributions offer the advantage of a closed form probability density function, the characteristic function has to serve as the only analytic representation available for performing the necessary calculations.

For all of the three considered distributions a general linear mapping $\mathbb{R}^n \to \mathbb{R}^m$ of a multivariate n-dimensional random variable $X \in \mathbb{R}^n$ is applied

$$V = AX + b\,,$$

specified by a $(m \times n)$-matrix A possessing full rank and a translation vector $b \in \mathbb{R}^m$.

In a first step a basic difficulty connected with the treatment of general linear transformations will be exemplified by means of MNTS distributions, where it can be observed most clearly. For this purpose, consider a m-dimensional random vector U created by $Y \sim MNTS(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$ and the above mapping

$$U = AY + b\,.$$

Its characteristic function is given by

$$\phi_U(z) = \phi_{AY+b}(z) = \exp(ib^\mathsf{T} z) \cdot \phi_Y(A^\mathsf{T} z) = \exp(\Psi_U(z))\,,\ z \in \mathbb{R}^m$$

$$\Psi_U(z) = ib^\mathsf{T} z + \frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i\gamma^\mathsf{T} A^\mathsf{T} z - \frac{1}{2}(A^\mathsf{T} z)^\mathsf{T}\mathrm{diag}(\sigma)\mathsf{P}\mathrm{diag}(\sigma)^\mathsf{T} A^\mathsf{T} z)\right)^{\frac{a}{2}}\right]$$

$$+ i(\mu - \gamma)^\mathsf{T} A^\mathsf{T} z$$

$$= \frac{2\lambda}{a}\left[1 - \left(1 - \frac{1}{\lambda}(i(A\gamma)^\mathsf{T} z - \frac{1}{2}(z^\mathsf{T} A\mathrm{diag}(\sigma)\mathsf{P}(A\mathrm{diag}(\sigma))^\mathsf{T} z))\right)^{\frac{a}{2}}\right]$$

$$+ i(A\mu + b - A\gamma)^{\mathsf{T}} z\,.$$

The calculations in the above equation result from applying the universal linear transformation formula for characteristic functions[20] to the specific context, which will likewise be done for the two remaining distributions. Finally, exploiting the present similarities of terms in corresponding expression (4.10) leads to the conclusion that U follows a $MNTS(a_U, \lambda_U, \gamma_U, \sigma_U, \mathsf{P}_U, \mu_U)$ distribution again. The extraction of

$$a_U = a\,,\ \lambda_U = \lambda\,,\ \gamma_U = A\gamma\,,\ \mu_U = A\mu + b$$

is straightforward, whereas the handling of the generated Brownian covariance structure

$$\Sigma_U^{(B)} = A\mathrm{diag}(\sigma)\mathsf{P}(A\mathrm{diag}(\sigma))^{\mathsf{T}}$$

proves to be more difficult. However, the only necessity is to accommodate this covariance matrix into the conventional parametrization for the MNTS distribution. This is achieved by decomposing $\Sigma_U^{(B)}$ into its associated vector of standard deviations σ_U and the separated correlation matrix P_U by the diagonal factorization

$$\Sigma_U^{(B)} = \mathrm{diag}(\sigma_U)\mathsf{P}_U\mathrm{diag}(\sigma_U)$$

$$\sigma_U^{(k)} = \Sigma_{U,(k,k)}^{(B)}$$

$$\mathsf{P}_U^{(i,j)} = \Sigma_{U,(i,j)}^{(B)} \cdot \left[\sigma_U^{(i)}\sigma_U^{(j)}\right]^{-1}\,.$$

Because of this unwieldy parameter representation, we restrict the further study of std-MVNTS and MVNTS to only simple linear transformations instead of considering the most general case. These simple linear transformations provide the central tool for performing linear standardizations in the following chapters, which is why they represent the subclass of the highest relevance with respect to our interests, nevertheless. Another central use of

[20] can be found e.g. in Rotar [Rota99, p. 361].

these simple linear combination to be anticipatorily mentioned at this point will be in the context of applying multivariate Normal Tempered Stable distributions for the innovation process in ARMA-GARCH time series models in chapter 6.

Before doing so, the basic findings concerning general linear transformations valid for all distributions of multivariate NTS type can be summarized as follows:

- Parameters (a, λ) associated with the CTS subordinator dynamics remain unaffected.
- The underlying implicit Brownian motion in the subordination is strictly linearly transformed by the mapping matrix A. This results in the respective mappings $A\gamma$ and $A\Sigma^{(B)}A^{\mathsf{T}}$ of its specifying parameters.
- The external drift component is subject to the entire affine linear transformation.

This principle means of influencing the single constituting components of multivariate NTS distributions is reflected in each of the now following cases of the restricted simple linear transformation.

For this purpose we furthermore consider a *simple linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^n$ of a n-dimensional random vector X

$$V = T(X) = \mathrm{diag}(\beta)X + \theta\,,$$

specified by a vector $\beta \in \mathbb{R}^n_{++}$ and a vector $\theta \in \mathbb{R}^n$. This kind of mapping is characterized by the peculiarity that each vector component of X is separately transformed by $v_i = \beta_i x_i + \theta_i, i = 1, \ldots, n$, which results in a high degree of tractability. This kind of transformation is meant by linear rescaling of X.

Starting with the MVNTS, we have that the random vector $V = T(X)$ created from a multivariate n-dimensional MVNTS distributed random vector $X \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$ possesses the characteristic function as follows

$$\phi_V(z) = \phi_X(\mathrm{diag}(\beta)z) \cdot \exp(i\theta^{\mathsf{T}}z) = \exp(\Psi_V(z))\,,\ z \in \mathbb{R}^n$$

$$= \exp\left[\frac{2\lambda}{a}\left(1-\left(1-\frac{1}{\lambda}\left(i(\mathrm{diag}(c)\gamma)^{\mathsf{T}}\mathrm{diag}(\beta)z-\frac{1}{2}(\mathrm{diag}(\beta)z)^{\mathsf{T}}\mathrm{diag}(c)\mathrm{diag}(\tilde{\sigma}_{(a,\lambda,\gamma)})\right.\right.\right.$$

$$\left.\left.\left.\cdot\,\mathsf{P}\mathrm{diag}(\tilde{\sigma}_{(a,\lambda,\gamma)})\mathrm{diag}(c)\mathrm{diag}(\beta)z\right)\right)^{\frac{a}{2}}\right)+i(d-\mathrm{diag}(c)\gamma)^{\mathsf{T}}\mathrm{diag}(\beta)z+i\theta^{\mathsf{T}}z\Bigg]$$

with

$$\tilde{\sigma}_{(a,\lambda,\gamma),k}=\sqrt{1-\frac{(2-a)}{2\lambda}\gamma_k^2}\,,\quad k=1,\ldots,n$$

which results in

$$\phi_V(z)=\exp\left[\frac{2\lambda}{a}\left(1-\left(1-\frac{1}{\lambda}\left(i(\mathrm{diag}(\beta)\mathrm{diag}(c)\gamma)^{\mathsf{T}}z-\frac{1}{2}z^{\mathsf{T}}\mathrm{diag}(\beta)\mathrm{diag}(c)\mathrm{diag}(\tilde{\sigma}_{(a,\lambda,\gamma)})\right.\right.\right.$$

$$\left.\left.\left.\cdot\,\mathsf{P}\mathrm{diag}(\beta)\mathrm{diag}(c)\mathrm{diag}(\tilde{\sigma}_{(a,\lambda,\gamma)})z\right)\right)^{\frac{a}{2}}\right)+i\left(\mathrm{diag}(\beta)d+\theta-(\mathrm{diag}(\beta)\mathrm{diag}(c)\gamma)\right)^{\mathsf{T}}z\Bigg]$$

When again exploiting the present structural similarities to the characteristic functions of MVNTS and MNTS, respectively, we find that V obeys a

- $MVNTS(a^*,\lambda^*,\gamma^*,\mathsf{P}^*,c^*,d^*)$ distribution with

$$a^*=a\,,\ \lambda^*=\lambda\,,\ \gamma^*=\gamma\,,\ \mathsf{P}^*=\mathsf{P} \tag{4.31}$$

$$c^*=\mathrm{diag}(\beta)c \tag{4.32}$$

$$d^*=\mathrm{diag}(\beta)d+\theta\,, \tag{4.33}$$

  or a

- $MNTS(a^*,\lambda^*,\gamma^{**},\sigma^*,\mathsf{P}^*,\mu^*)$ distribution with

$$\gamma^{**}=\mathrm{diag}(\beta)\mathrm{diag}(c)\gamma$$

$$\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)^\mathsf{T} = \mathrm{diag}(\beta)\mathrm{diag}(c)\tilde{\sigma}_{(a,\lambda,\gamma)} \text{ with } \sigma_k^* = \sqrt{1 - \frac{(2-a)}{2\lambda}\gamma_k^2 c_k \beta_k}$$

$$\mu^* = \mathrm{diag}(\beta)d + \theta = d^* \,,$$

where the latter representations could also be obtained by means of the parameter correspondence from section 4.2.2, p. 74 f. . Here, one notices in particular how the MVNTS distributions benefit from the extraction of their scales, which are already explicitly reflected in their parametrization. This, in turn, enables a rather compact description of the results.

When subsequently treating the stdMVNTS distribution, one can build on the above results and on the fact that by its specific construction a stdMNTS distribution is included in the general class of MVNTS distributions with $c = (1, \ldots, 1)^\mathsf{T} = \mathbf{1}$ and $d = (0, \ldots, 0)^\mathsf{T} = \mathbf{0}$. To elaborate on this in greater detail, a multivariate n-dimensional standard NTS distributed random vector $\tilde{X} \sim stdMVNTS(a, \lambda, \gamma, \mathsf{P}) \equiv MVNTS(a, \lambda, \gamma, \mathsf{P}, \mathbf{1}, \mathbf{0})$ together with a transformation $U = T(\tilde{X})$ are considered

$$U = T(\tilde{X}) = \mathrm{diag}(\beta)\tilde{X} + \theta \,.$$

By the above argument and the relations given in equations (4.32) and (4.33) it is straightforward to conclude that U is again MVNTS distributed $V \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c^*, d^*)$ with

$$c^* = \mathrm{diag}(\beta)\mathbf{1} = \beta$$

$$d^* = \mathrm{diag}(\beta)\mathbf{0} + \theta = \theta \,.$$

As a concluding remark, it should be stated that the class of multivariate NTS distributions is closed under linear transformations, which of course is independent of the choice of particular parametrization MVNTS or MNTS. This closedness in combination with the possibility of simple representation of the results, especially concerning linear rescalings,

enables the application of MVNTS distributions as innovation processes for of ARMA-GARCH time series models to be studied in chapter 6.

4.2.4 Linear Combinations

In this section, linear combinations of dimensions of multivariate Normal Tempered Stable distributions are presented. They constitute another specific subclass of general linear transformations covered in the section above and are a central manipulation tool for portfolio analysis to be treated in the forthcoming chapters. But unlike general linear transformations do, linear combinations in the context of MVNTS distributions possess a thoroughly desirable property which will provide convenient tractability. This fact can be attributed on the one hand to the simple structure of linear combinations, but on the other even more so to the particular properties of the MVNTS distributions due to its foresighted construction approach via Brownian subordination.

For a non-trivial vector $\omega \in \mathbb{R}^n$, one obtains a linear combination Z of dimensions of a n-dimensional MVNTS distributed random vector $X \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$ with weighting ω through

$$Z = \omega^{\mathsf{T}} X = \sum_{k=1}^{n} \omega_k X_k \,.$$

By the definition of MVNTS via Brownian subordination and subsequent linear rescaling presented in sections 4.2.1 and 4.2.2, it holds true that

$$\begin{aligned} Z &= \omega^{\mathsf{T}} \left(\mathrm{diag}(c)\tilde{X} + d \right) = \omega^{\mathsf{T}} \mathrm{diag}(c)\tilde{X} + \omega^{\mathsf{T}} d \\ &= \sum_{k=1}^{n} \omega_k d_k + \sum_{k=1}^{n} \omega_k c_k \tilde{X}_k \\ &= \omega^{\mathsf{T}} \mathrm{diag}(c) \left[(S-1)\gamma + \sqrt{S}\mathrm{diag}(\tilde{\sigma})\mathrm{chol}(\mathsf{P})W \right] + \omega^{\mathsf{T}} d \end{aligned}$$

$$= \sum_{k=1}^{n} \omega_k d_k + \sum_{k=1}^{n} \omega_k c_k \left(\gamma_k (S-1) + \tilde{\sigma}_k \sqrt{S} V_k \right)$$

with

$$\tilde{\sigma} = \left[\sqrt{1 - \frac{2-a}{2\lambda} \gamma_k^2} \right]_{k=1,\ldots,n}^{\mathsf{T}}$$

and

$$W \sim \mathcal{N}^n(0, I)\,,\ V \sim \mathcal{N}^n(0, \mathsf{P})$$

Hence,

$$Z = \omega^{\mathsf{T}} d + \omega^{\mathsf{T}} \mathrm{diag}(c) \gamma (S-1) + \mathrm{diag}(c) \mathrm{diag}(\tilde{\sigma}) \mathrm{chol}(\mathsf{P}) \sqrt{S} \omega^{\mathsf{T}} W \tag{4.34}$$

$$= \sum_{k=1}^{n} \omega_k d_k + \left(\sum_{k=1}^{n} \omega_k c_k \gamma_k \right) (S-1) + \sqrt{S} \left(\sum_{k=1}^{n} \omega_k c_k \tilde{\sigma}_k V_k \right)\,, \tag{4.35}$$

where $\tilde{X}$ denotes a corresponding multivariate $stdMVNTS(a, \lambda, \gamma, \mathsf{P})$ distributed random vector. The last expression in equation (4.34) and (4.35), respectively, are made up of a weighted sum of possibly correlated Normal distributions, which in turn, by the familiar property of Normal distributions, results in a univariate Normal distribution. With the conventional computation rules for Gaussian random variables, we obtain the the following expression

$$Z = m + g(S-1) + s\sqrt{S} U\,, \quad U \sim \mathcal{N}(0,1) \tag{4.36}$$

where

$$m = \omega^{\mathsf{T}} d = \sum_{k=1}^{n} \omega_k d_k \tag{4.37}$$

$$g = \omega^{\mathsf{T}} \mathrm{diag}(c) \gamma = \sum_{k=1}^{n} \omega_k c_k \gamma_k \tag{4.38}$$

$$s = \sqrt{\omega^{\mathsf{T}}\mathrm{diag}(c)\mathrm{diag}(\tilde{\sigma})\mathsf{P}\mathrm{diag}(\tilde{\sigma})\mathrm{diag}(c)\omega}\,,\ \mathsf{P} = (\rho_{j,k})_{j,k=1,\ldots,n}$$

$$= \sqrt{\sum_{j=1}^{n}\sum_{k=1}^{n}\omega_j\omega_k c_j c_k \tilde{\sigma}_j\tilde{\sigma}_k\rho_{j,k}} = \sqrt{\sum_{j=1}^{n}\sum_{k=1}^{n}\omega_j\omega_k c_j c_k \sqrt{1-\frac{(2-a)}{2\lambda}\gamma_j^2}\sqrt{1-\frac{(2-a)}{2\lambda}\gamma_k^2}\rho_{j,k}}\,, \tag{4.39}$$

which closely resembles the defining subordination scheme for univariate NTS distributions in equation (3.24), section 3.7. Therefore, the univariate random variable Z turns out to be distributed according to a univariate Normal Tempered Stable distribution with the above parameter values again, hence $Z \sim NTS(a, \lambda, g, s^2, m)$. Note that the above sum representations (4.39) attempt to clearly illustrate the composition of the emerging parameter values on the one hand, while on the other the more compact linear algebra notation is aimed at efficient implementation of the necessary calculation in a matrix-oriented programming environment like MATLAB®.

The reason for the existence of this useful property, is the fact that MVNTS distributions are constructed via Brownian subordination, which has already been stressed before. To be more specific, linear combinations are basically concerned with the summation of Normal distributions while the subordinator itself remains unaffected. In other words, it makes no difference whether linear combinations of a subordinated Brownian motion are considered or if the linear combination of the Brownian motion is subordinated directly, instead. At this point, it is worth mentioning another beneficial relation with regard to MVNTS parameter estimation to be carried out at a later stage. As the one-dimensional subordinator S is not influenced by the linear combination, the parameter values (a, λ) associated with its dynamics are directly passed from the multivariate MVNTS on to the univariate NTS. Hence, this operation preserves the pair (a, λ) of a MVNTS distribution which can furthermore be obtained from any non-trivial linear combination by applying a suitable NTS estimation procedure.

Chapter 5

Unconditional Models

This chapter is intended to provide a first field of application of the multivariate Normal Tempered Stable distribution within the scope of financial data modeling. In order to emphasize relevant further concepts yet to be presented, these mainly include parameter estimation as well as numerical approximations of probability density and cumulated distribution functions of MVNTS distributions, a deliberately naive and simple model framework is chosen at first. After these specific methods were introduced, the MVNTS distribution will be applied within the context of a successful state-of-the-art model for financial time series data in the next chapter.

In the remainder of this chapter, the above mentioned methods are utilized to fit the specified model to a set of empirical data taken from the *Dow Jones Industrial Average Index*. Furthermore, the goodness of fit of the estimated model will be assessed by a range of various powerful statistical standard tests designed for this purpose. In a last step, an optimal portfolio strategy based on the newly developed model is going to be be derived, whose performance is subsequently compared to a common widespread approach by means of an empirical backtest.

5.1 Model Specification

In contrast to the continuously compounded return rates usually employed in financial modeling, we utilize simply compounded rates

$$r_t = \frac{\Delta P_t}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{5.1}$$

instead, which should be subject of the subsequent modeling task. The reason lies in the advantages this choice of modeling variable entails for determining portfolio returns at a later stage of this chapter[1].

In the following, a comparatively simple model for the joint dynamics of asset prices is suggested. Actually, one assumes the simply compounded index and joint stock returns (5.1) in each 1-day-period to be serially independent and identically distributed (i.i.d.) random vectors R_t according to the multivariate Normal tempered stable distribution introduced in section 4.2.2. This formulation is further referred to as the *Stationary Model* of constant volatility and basically embodies a white noise process for the daily returns[2]

$$R_t \overset{i.i.d.}{\sim} MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)\,,\ t \in \mathbb{N}\,.$$

5.2 Empirical Data Set

The empirical study of a stationary model for asset returns is based on daily return data of 29 stocks included in the Dow Jones Industrial Average (DJIA) index in addition to the daily returns of the index itself. These time series of historical observations cover the time interval between September 25, 1997 and October 30, 2009. Hence, the entire data set consists of in total 3405 daily records for each of the 29 encompassed stocks. The DJIA component listing given in table 5.1 is that of October 25, 2010 and k denotes the position

[1] By the arguments of Brooks in [Broo08, p. 7].

[2] Note however, that this is in general not equivalent to a random walk for the prices, only aggregated log-returns would follow random walk processes.

k	Company Name	Ticker
1	3M Co.	MMM
2	Alcoa Inc.	AA
3	American Express Co.	AXP
4	AT&T Inc.	T
5	Bank of America Corp.	BAC
6	Boeing Co.	BA
7	Caterpillar Inc.	CAT
8	Chevron Corp.	CVX
9	Cisco Systems Inc.	CSCO
10	Coca-Cola Co.	KO
11	E.I. DuPont de Nemours & Co.	DD
12	Exxon Mobil Corp.	XOM
13	General Electric Co.	GE
14	Hewlett-Packard Co.	HPQ
15	Home Depot Inc.	HD

k	Company Name	Ticker
16	Intel Corp.	INTC
17	International Business Machines Corp.	IBM
18	Johnson & Johnson	JNJ
19	JPMorgan Chase & Co.	JPM
20	McDonald's Corp.	MCD
21	Merck & Co. Inc.	MRK
22	Microsoft Corp.	MSFT
23	Pfizer Inc.	PFE
24	Procter & Gamble Co.	PG
25	Travelers Cos. Inc.	TRV
26	United Technologies Corp.	UTX
27	Verizon Communications Inc.	VZ
28	Wal-Mart Stores Inc.	WMT
29	Walt Disney Co.	DIS

Table 5.1: Considered stocks with ticker symbols and index position.

in the data set[3]. One separate time series of the same length contains the associated DJIA index returns over the above sample period. The origin of this data is the OptionMetrics Ivy DataBase provided by Wharton Research Data Services.

5.3 Parameter Estimation

The basic tool for tackling the task of parameter estimation of the unconditional distribution $MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$, which in turn entirely specifies the stated model for the stock returns, will be a maximum likelihood (ML) approach. As a simultaneous estimation of all involved parameter values proves impossible by aspects of computational time consumed and robustness of resulting estimates, a decomposition and following sequential procedure for solving this problem is proposed here. Besides the multivariate model for the joint returns of the stocks, a univariate model for the index returns is separately considered in the first instance. By the particular construction approach it is then possible to easily incorporate the index returns as a further dimension into the multivariate NTS model for the 29 stocks afterwards.

[3]Of the 30 stocks currently comprised in the DJIA , Kraft Food Inc. (KFT) is left out of consideration because of incomplete return data history available to the author.

5.3.1 Sample Period

For implementation of the actual parameter estimation only historical daily return records from a 10-year-period, beginning January 2, 1999 and ending December 31, 2008, are considered. Hence, $N = 2514$ returns are contained in the sample period. It has to be taken into account, that the recent financial crisis, peaking in the collapse of U.S. investment bank Lehman Brothers on September 15, 2008, with its highly volatile market conditions, is included in the estimation period.

5.3.2 Index Model

At first, we turn to the one-dimensional NTS model for the DJIA index returns $R^{(ind)}$

$$R_t^{(ind)} \overset{i.i.d.}{\sim} NTS(a_{(ind)}, \lambda_{(ind)}, \gamma_{(ind)}, \sigma^2_{(ind)}, \mu_{(ind)}) \,,\ t \in \mathbb{N} \,,$$

and their time series of empirical realizations $\{r_t^{(ind)}\}$, $t = 1, \ldots, 2514$. Although this kind of parametrization was not explicitly introduced for the univariate NTS in chapter 3, the approach of removing the scales and modeling the standardized situation only, familiar from the multivariate setting, should be employed for the index data, also. This simplifies the subsequent estimation procedure on the one hand, on the other hand this decomposition will facilitate the conflation of index and stock returns in one joint model, which is required for later portfolio optimization in section 5.5.2. The empirical returns have arithmetic sample mean $\bar{r}^{(ind)} = 6.2227 \cdot 10^{-5}$ and sample standard deviation[4] $Std^*(r^{(ind)}) = \sqrt{Var^*(r^{(ind)})} = 0.0127$. Both values represent unbiased estimates for the expectation $E(R^{(ind)})$ and the standard deviation $Std(R^{(ind)}) = \sqrt{Var(R^{(ind)})}$ of the index returns $R^{(ind)}$, respectively, and moreover reflect the inherent scales of the data. As already indicated above, the principle of removing the scale and then modeling the emerging standardized distribution can be directly transferred from the multivariate case, although for the univariate case the NTS was not defined in the same way as the MVNTS

[4]Recall that the sample version of variance and standard deviation operate with a denominator of $(N-1)$ instead of N in order to ensure unbiasedness.

finally was. Of course it is obvious that the information contained in the removed scale has to be further carried along when reducing NTS to stdNTS.

Before actually performing the linear standardization, it is important to point out the following relation. One has to prove first, that the set of linearly standardized NTS distributions is congruent to the set of univariate stdNTS distributions presented in section 4.2. This verification involves two steps:

1.) Linearly standardized NTS distributions are entirely contained in the set of univariate stdNTS distributions.

2.) Univariate stdNTS distributions for their part constitute a subset of univariate and linearly standardized univariate NTS distributions.

The direction referred to by 1.) is proven by considering a random variable $X \sim NTS(a, \lambda, \gamma, \sigma^2, \mu)$ with expectation $E(X) = \mu$ and standard deviation[5]

$$Std(X) = \sqrt{Var(X)} = \sqrt{\sigma^2 + \frac{2-a}{2\lambda}\gamma^2}\,.$$

The corresponding linearly standardized random variable $\tilde{X} = \bar{b}X + \bar{d}$ is consequently attained by setting

$$\tilde{X} = \frac{X - E(X)}{Std(X)} = \frac{1}{Std(X)}X - \frac{E(X)}{Std(X)}$$

$$\bar{b} = \frac{1}{Std(X)} = \left(\sigma^2 + \frac{2-a}{2\lambda}\gamma^2\right)^{-1/2}$$

$$\bar{d} = -\frac{E(X)}{Std(X)} = -\mu\bar{b} = -\mu\left(\sigma^2 + \frac{2-a}{2\lambda}\gamma^2\right)^{-1/2}\,.$$

By the rules of univariate linear transformations developed in section 3.7.2 we have that $\tilde{X}$ is again NTS distributed with parameter vector $(\tilde{a}, \tilde{\lambda}, \tilde{\gamma}, \tilde{\sigma}^2, \tilde{\mu})$, i.e.

[5] See equation (3.21).

$\tilde{X} \sim NTS(\tilde{a}, \tilde{\lambda}, \tilde{\gamma}, \tilde{\sigma}^2, \tilde{\mu})$, and

$$\tilde{a} = a\,,\ \tilde{\lambda} = \lambda$$

$$\tilde{\gamma} = \bar{b}\gamma = \left(\sigma^2 + \frac{2-a}{2\lambda}\gamma^2\right)^{-1/2} = \left(\frac{\sigma^2}{\gamma^2} + \frac{2-a}{2\lambda}\right)^{-1/2}$$

$$\tilde{\sigma}^2 = \sigma^2\bar{b}^2 = \sigma^2\left(\sigma^2 + \frac{2-a}{2\lambda}\gamma^2\right)^{-1} = \left(1 + \frac{2-a}{2\lambda}\frac{\gamma^2}{\sigma^2}\right)^{-1} = \left(\frac{2\lambda\sigma^2 + (2-a)\gamma^2}{2\lambda\sigma^2}\right)^{-1} \tag{5.2}$$

$$\tilde{\mu} = \bar{b}\mu + \bar{d} = \bar{b}\mu - \mu\bar{b} = 0\,. \tag{5.3}$$

Furthermore, it remains to be shown that this NTS distribution matches the structure of a particular $stdNTS(a^*, \lambda^*, \gamma^*)$ distribution in $NTS(a^*, \lambda^*, \gamma^*, \left[\bar{\sigma}_{(a^*,\lambda^*,\gamma^*)}\right]^2, \bar{\mu})$ parametrization. In above equation (5.3), the condition

$$\bar{\mu} \stackrel{!}{=} 0 = \tilde{\mu}$$

is shown to be directly satisfied. The term $\bar{\sigma}_{(a^*,\lambda^*,\gamma^*)}$ implied by the stdNTS parameters

$$a^* = \tilde{a} = a\,,\ \lambda^* = \tilde{\lambda} = \lambda$$

$$\gamma^* = \tilde{\gamma} = \left(\frac{\sigma^2}{\gamma^2} + \frac{2-a}{2\lambda}\right)^{-1/2}$$

has to be equal to the expression in (5.2), which is finally proven by below calculations resulting in equation (5.4).

$$\left[\bar{\sigma}_{(a^*,\lambda^*,\gamma^*)}\right]^2 = 1 - \frac{2-a^*}{2\lambda^*}(\gamma^*)^2 = 1 - \frac{2-a^*}{2\lambda^*}\left(\frac{\sigma^2}{\gamma^2} + \frac{2-a}{2\lambda}\right)^{-1}$$

$$= 1 - \left(\frac{2\lambda\sigma^2}{(2-a)\gamma^2} + 1\right)^{-1} = 1 - \left(\frac{2\lambda\sigma^2 + (2-a)\gamma^2}{(2-a)\gamma^2}\right)^{-1}$$

$$= 1 - \frac{(2-a)\gamma^2}{2\lambda\sigma^2 + (2-a)\gamma^2} = \frac{2\lambda\sigma^2 + (2-a)\gamma^2 - (2-a)\gamma^2}{2\lambda\sigma^2 + (2-a)\gamma^2}$$

$$= \frac{2\lambda\sigma^2}{2\lambda\sigma^2 + (2-a)\gamma^2} \tag{5.4}$$

But as values of parameter γ for stdNTS distributions have to meet a certain restriction reflected by equation (3.28), its validity in the considered case is finally verified by

$$|\gamma^*| \overset{!}{<} \sqrt{\frac{2\lambda^*}{2-a^*}} \Leftrightarrow (\gamma^*)^2 \overset{!}{<} \frac{2\lambda^*}{2-a^*} = \left(\frac{2-a}{2\lambda}\right)^{-1}$$

$$\frac{2-a}{2\lambda} > 0 \Rightarrow (\gamma^*)^{-2} = \underbrace{\underbrace{\frac{\sigma^2}{\gamma^2}}_{>0} + \frac{2-a}{2\lambda}}_{>0} > \frac{2-a}{2\lambda}$$

$$\Rightarrow (\gamma^*)^2 = \left(\frac{\sigma^2}{\gamma^2} + \frac{2-a}{2\lambda}\right)^{-1} = \left[(\gamma^*)^{-2}\right]^{-1} < \left(\frac{2-a}{2\lambda}\right)^{-1}.$$

Thus, every linearly standardized NTS random variable X obeys a particular stdNTS distribution whose parameter values were explicitly derived above.

The part constituted by 2.) remains to be considered. An explicit verification like the one for the opposite direction carried out above would become more difficult here, wherefore an indirect approach is implemented instead. In doing so, one has to show that each arbitrarily rescaled stdNTS distribution is again contained in the class of NTS distributions. This statement is indeed true as every stdNTS distribution has an alternative representation in terms of NTS parameters and the class of NTS distributions is furthermore closed under the involved linear transformations[6]. Hence, by 1.) and 2.) the equivalence of the set of linearly standardized NTS distributions and the set of stdNTS distributions is finally verified. In conclusion, this explains why the assumption of stdNTS distributions for the linearly standardized index returns $\tilde{R}^{(ind)}$ is admissible when the initial $R^{(ind)}$ are assumed to be univariate NTS distributed. The likewise linearly standardized historical

[6]This closure has been pointed out in section 3.7.2.

index returns in question are created by

$$\tilde{r}^{(ind)} = \frac{r^{(ind)} - \bar{r}^{(ind)}}{Std^*(r^{(ind)})},$$

where the employed coefficients $\bar{r}^{(ind)}$ and $Std^*(r^{(ind)})$ are simply the unbiased estimates of the scale parameters $E(R^{(ind)})$ and $Std(R^{(ind)})$ obtained above.

The associated stdNTS parameters $(\tilde{a}_{(ind)}, \tilde{\lambda}_{(ind)}, \tilde{\gamma}_{(ind)})$ for $\tilde{R}^{(ind)}$ are estimated simultaneously by the maximum likelihood method, yielding point estimates and corresponding 95% confidence intervals (CI) given in table 5.2. The procedure relies on a heuristic search on the feasible region for parameter values, an approach which will be repeatedly applied in a similar form in the forthcoming sections. The required inversion formulas for the probability density function together with efficient techniques for their numerical evaluation[7] by Fast Fourier Transformations are comprehensively covered in the forthcoming section 5.3.3.1.

	$\hat{a}_{(ind)}$	$\hat{\lambda}_{(ind)}$	$\hat{\gamma}_{(ind)}$
estimate	1.1584	0.1812	-0.0296
lower CI boundary	0.8791	0.0521	-0.0765
upper CI boundary	1.4376	0.3104	0.0173

Table 5.2: Parameter estimates (stdNTS) for standardized index returns.

Supplementarily, it is noted that by using the above point estimates in combination with the estimates of mean and standard deviation the modeled $R_t^{(ind)}$ follow an estimated $NTS(\hat{a}, \hat{\lambda}, \hat{\eta}, \hat{\sigma}^2, \hat{\mu})$ distribution with

$$\hat{\sigma} = \bar{b} \cdot \tilde{\sigma}_{(\hat{a},\hat{\lambda},\hat{\gamma})} = Std(r^{(ind)}) \cdot \sqrt{1 - \frac{2-\hat{a}}{2\hat{\lambda}}\hat{\gamma}^2} = 0.01267$$

$$\hat{\eta} = \bar{b} \cdot \hat{\gamma}_{(ind)} = Std(r^{(ind)}) \cdot \hat{\gamma}_{(ind)} = -3.7592 \cdot 10^{-4}$$

$$\hat{\mu} = \bar{b} \cdot \tilde{\mu} + \bar{d} = \bar{d} = \bar{r}^{(ind)} = 6.2227 \cdot 10^{-5}.$$

[7] The numerical evaluation becomes necessary by the fact that closed-form solutions are not available as has been explained in section 4.2.1.3 previously.

This consequently results from the application of the calculation rules developed for linear transformations in section (3.7.2) and the constructive definition of the stdNTS distribution. The above parameter values indicate that the distribution is concentrated on a rather limited range around 0 which constitutes another reason for performing the required parameter estimations only after the observations have been previously standardized.

5.3.3 CF Inversion

5.3.3.1 Probability Density Function

The method of maximum likelihood parameter estimation requires evaluation of the NTS probability density function. As the PDF in case of the univariate NTS distribution cannot be expressed as an analytical function, so does the efficient point-wise numerical approximation of the PDF pose one of the main challenges and necessary preconditions for practical handling of tempered stable distributions. The same applies to cumulative distribution functions of the univariate NTS, which become relevant in the further steps of the analysis when quantiles and other risk measures have to be determined.

Treating the PDF $f_X(x)$ first, the following inversion relation[8] to the characteristic function $\phi_X(z)$ has to be solved,

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-izx)\phi_X(z)dz\,,\ \forall\, x \in \mathbb{R}\,. \tag{5.5}$$

The above complex integral embodies the familiar continuous Fourier transform so f_X corresponds to the representation of ϕ_X on the frequency domain $x \in \mathbb{R}$. For a single point $x \in \mathbb{R}$, the entire range of numerical integration techniques for the above integral in equation (5.5) would be available. All of these approaches, however, do not prove very suitable as only one single point x can be treated at the same time. To remedy this situation, Bailey and Swarztrauber [BaSw94] suggested the following solution, encompassing three particular steps:

[8] Similar to equations (4.22) and (4.25).

- The integration in equation (5.5) is discretized, the integrand is converted to a simple step function.
- The discretization grid $\{z_j\}_{j=0,1,\ldots,m}$ for the approximative evaluation on $z \in \mathbb{R}$ as well as a set of frequencies $\{x_k\}_{k=0,1,\ldots,m}$ are set in a way that the expression resembles a discrete Fourier transform of $\{\phi_X(z_j)\}$. It is central to note here that this discrete Fourier transform (DFT) involved is the same for every x_k.
- For the DFT an efficient FFT[9] algorithm is available which reduces the computational complexity from $\mathcal{O}(m^2)$ to $\mathcal{O}(m \log m)$.

This approach is going to be presented here in full detail as it will provide the basic structure for the numerical inversion of the characteristic function in order to obtain the CDF. Following Bailey and Swarztrauber [BaSw94], the complex exponent function in the analytic inversion term[10] is discretized to a simple step function

$$\int_{-\infty}^{\infty} \exp(-ixz)\phi(z)dz \approx \sum_{j=0}^{m-1} \exp(-ix_k z_j)\phi(z_j)\Delta_z =: I_k \tag{5.6}$$

on a grid $\{x_k\}$ at first. The discretization is performed on a grid $\{z_j\} = \{s_z + j\Delta_z,\, j = 0, 1, \ldots, m\}$ and a set of corresponding evaluation frequencies $\{x_k\} = \{s_x + \Delta_x,\, k = 0, 1, \ldots, m\}$. These two sets are each specified by a respective start value s_z, s_x and step size Δ_z, Δ_x for m steps. Incorporating this grid specification into equation (5.6) leads to

$$\begin{aligned} I_k &= \sum_{j=0}^{m-1} \exp(-i(s_x + k\Delta_x)(s_z + j\Delta_z))\phi(s_z + j\Delta_z)\Delta_z \\ &= \Delta_z \sum_{j=0}^{m-1} \exp(-i(s_x s_z + s_x\Delta_z j + \Delta_x s_z k + \Delta_x\Delta_z kj))\phi(s_z + j\Delta_z) \\ &= \Delta_z \exp(-is_x s_z) \exp(-i\Delta_x s_z k) \sum_{j=0}^{m-1} \exp(-i\Delta_x\Delta_z kj) \exp(-is_x\Delta_z j)\phi(s_z + j\Delta_z)\,. \end{aligned} \tag{5.7}$$

[9] abbrv.: Fast Fourier Transformation
[10] see equation (5.5)

Recall that the discrete Fourier transform (DFT) of a complex vector $p \in \mathbb{C}^m$ is defined as the likewise complex vector $D \in \mathbb{C}^m$ with

$$D_k = \sum_{j=0}^{m-1} e^{-2\pi \mathrm{i} \cdot \frac{jk}{m}} \cdot p_j \,, \; k = 0, 1, \ldots, m-1 \,. \tag{5.8}$$

Provided $m \in 2^{\mathbb{N}}$, i.e. m is a power of two, the Fast Fourier Transformation algorithm as a means of efficient computation of the DFT D can be applied. On this account it is desired to arrange the grids in a way that the expression in equation (5.7) adopts the structure of the DFT in equation (5.8) while meeting the prerequisites for the FFT. This is basically achieved by setting

$$\Delta_x \Delta_z = \frac{2\pi}{m} \Rightarrow \Delta_x = \frac{2\pi}{m\Delta_z} \,, \tag{5.9}$$

when assuming that Δ_z and $m \in 2^{\mathbb{N}}$ have arbitrary but fixed values, yielding

$$\begin{aligned} I_k &= \Delta_z \exp(-is_x s_z) \exp(-i\frac{2\pi}{m\Delta_z} s_z k) \sum_{j=0}^{m-1} \exp(-i2\pi \frac{kj}{m}) \exp(-is_x \Delta_z j) \phi(s_z + j\Delta_z) \\ &= \Delta_z \exp(-is_x s_z) \exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z} k) D_k \left[[\exp(-is_x \Delta_z j) \phi(s_z + j\Delta_z)]_{j=0,1,\ldots,m-1} \right] . \end{aligned}$$

Note that the above relation (5.9) enables a representation of the summation as a DFT of a slightly modified vector of characteristic function values, evaluated on the z-grid. Finally, a reasonable but more or less arbitrary choice for the starting points of the grids

$$s_z = -\Delta_z \cdot \frac{m}{2} = -\frac{m\Delta_z}{2} \tag{5.10}$$

$$s_x = -\Delta_x \cdot \frac{m}{2} = -\frac{\pi}{\Delta_z} \tag{5.11}$$

position them almost centered around the origo. This completes the grid specification and leads to

$$I_k = \Delta_z \exp(-i\pi \frac{s_z}{\Delta_z}) \exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z} k) D_k \left[[\exp(-i\pi j) \phi(s_z + j\Delta_z)]_{j=0,1,\ldots,m-1} \right]$$

$$= \Delta_z \exp(-i\pi\frac{m}{2}) \exp(-i\pi k) D_k \left[\left[(-1)^j \phi(\Delta_z(j - \frac{m}{2})) \right]_{j=0,1,\ldots,m-1} \right]$$

$$= \Delta_z \exp(-i\pi k) D_k \left[\left[(-1)^j \phi(\Delta_z(j - \frac{m}{2})) \right]_{j=0,1,\ldots,m-1} \right] , \; k = 0, 1, \ldots, m-1 \, .$$

By applying the above conventions for grid construction one finds that with a fixed m controlling the general level of approximation precision or computation expenditure, respectively, the step size of the z-grid is the only remaining free determination variable for both grids

$$\Delta_z \;\; , \;\; s_z = -\Delta_z \cdot \frac{m}{2} = -\frac{m\Delta_z}{2} \quad (z\text{-grid})$$

$$\Delta_x = \frac{2\pi}{m\Delta_z} \;\; , \;\; s_x = -\Delta_x \cdot \frac{m}{2} = -\frac{\pi}{\Delta_z} \quad (x\text{-grid}) \, .$$

These implied parameters together with Δ_z are responsible for how large approximation errors can grow[11] on the one hand and how dense and wide-ranging the returned approximate values of $f_X(x)$ are on the other.

5.3.3.2 Cumulative Distribution Function

For the univariate CDF, various alternative inversion formulas of the characteristic function are present to this date. One developed by Kim et al. is

$$F_X(x) = \frac{e^{x\rho}}{\pi} \Re \left(\int_0^\infty e^{-ixz} \frac{\phi_X(z + i\rho)}{c(\rho - iz)} dz \right) , \; x \in \mathbb{R} \, , \tag{5.12}$$

[11] The precision of this technique is mainly determined by a combination of either the range $m\Delta_z$, which should therefore be as large as possible, and the step size s_z, which should be as small as possible for reasonable results. Because with fixed m this cannot be achieved simultaneously, finding a good value for Δ_z is a question of adequately balancing these two aspects.

which is presented in [KRBF09] together with an appropriate FFT discretization. The variable c admits

$$c = \begin{cases} 1 & \exists\ \rho > 0 : |\phi_X(z)| < \infty\ \forall \{z \in \mathbb{C} : \Im(z) = \rho\} \\ -1 & \exists\ \rho < 0 : |\phi_X(z)| < \infty\ \forall \{z \in \mathbb{C} : \Im(z) = \rho\} \end{cases} .$$

The last condition basically boils down to the question whether the moment generating $M_X(u)$ of X possesses a non-degenerate domain G from which a constant ρ can be chosen. At the same time, this already illustrates the central shortcoming of this approach. While the characteristic function of X is generally defined on the whole real line $\mathbb{R}$, this cannot be assumed for $M_X(u)$. Therefore, before applying the above method a feasible value $\rho \in G$ or G itself, respectively, has to be determined.

In order to remedy this weakness and to eliminate the obstructive threshold for application, an alternative inversion approach is presented here, which will be furthermore conditioned for efficient tractability with FFT tools. As X is a continuous random variable, the inversion formula of Gil-Pelaez [GilP51]

$$F_X(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\Im[e^{-ixz}\phi_X(z)]}{z}\, dz = \frac{1}{2} - \frac{1}{\pi} \cdot \Im \left[\int_0^\infty \frac{e^{-ixz}\phi_X(z)}{z}\, dz \right]$$

valid for continuity points applies for all $x \in \mathbb{R}$ in the considered case. For the numerical approximation

$$\int_0^\infty \frac{\exp(-ixz)\phi(z)}{z} dz \approx \sum_{j=0}^{m-1} \exp(-ix_k z_j) \frac{\phi(z_j)}{z_j} \Delta_z =: I_k$$

one has to pay attention to the fact that the integrand now exhibits a singularity around 0. In the same manner as previously in the Bailey and Swarztrauber approach, a z- and x-grid is specified by the grid coefficients $m, s_z, \Delta_z, s_x, \Delta_x$

$$I_k = \Delta_z \exp(-is_x s_z) \exp(-i\Delta_x s_z k) \sum_{j=0}^{m-1} \exp(-i\Delta_x \Delta_z kj) \exp(-is_x \Delta_z j) \frac{\phi(s_z + j\Delta_z)}{(s_z + j\Delta_z)} . \tag{5.13}$$

Again, the necessary condition for a conversion of the sum into a discrete Fourier transform

is $\Delta_x = \frac{2\pi}{m\Delta_z}$ for a pre-specified fixed $m \in 2^{\mathbb{N}}$. By insertion into equation (5.13), this yields

$$I_k = \Delta_z \exp(-is_x s_z) \exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z}k) \sum_{j=0}^{m-1} \exp(-i2\pi\frac{kj}{m}) \exp(-is_x\Delta_z j) \frac{\phi(s_z + j\Delta_z)}{(s_z + j\Delta_z)}$$

$$= \Delta_z \exp(-is_x s_z) \exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z}k) D_k \left[\left[\exp(-is_x\Delta_z j) \frac{\phi(s_z + j\Delta_z)}{(s_z + j\Delta_z)} \right]_{j=0,1,\ldots,m-1} \right] \tag{5.14}$$

Without any up-front information concerning the location of the distribution of X it seems appropriate to position the x-grid around the origo, which is done by a proper choice of s_x. Although s_x could be set as in equation (5.11), it has proven to deliver rather imprecise results for the Gil-Pelaez formula near $x = 0$ however. An improved alternative is the construction of an absolute symmetric x-grid where by the even number of steps m no x_k coincides with 0.

$$s_x = -\Delta_x \left(\frac{m}{2} - \frac{1}{2} \right) = -\frac{2\pi}{m\Delta_z} \left(\frac{m-1}{2} \right) = -\frac{\pi(m-1)}{m\Delta_z}\,.$$

After the x-grid is entirely specified by the above coefficients, only s_z remains for determination. As a singularity is present in $z = 0$, $s_z = 0$ fails to be a valid choice. When Δ_z is reasonably small, approximations of the first order of the integrand become sufficiently precise, so $s_z = \frac{\Delta_z}{2}$ is as close as possible to the true result, ceteris paribus. Inserting s_x and s_z into equation (5.14) leads to

$$I_k = \Delta_z \exp(i\frac{\pi(m-1)s_z}{m\Delta_z}) \exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z}k) D_k \left[\left[\exp(i\frac{\pi(m-1)}{m\Delta_z}\Delta_z j) \frac{\phi(s_z + j\Delta_z)}{(s_z + j\Delta_z)} \right]_{j=0,1,\ldots,m-1} \right]$$

$$= \Delta_z \exp(i\pi\frac{(m-1)s_z}{m\Delta_z}) \exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z}k) D_k \left[\left[\exp(i\frac{\pi(m-1)}{m}) \frac{\phi(s_z + j\Delta_z)}{(s_z + j\Delta_z)} \right]_{j=0,1,\ldots,m-1} \right]$$

$$= \Delta_z \exp(i\pi\frac{(m-1)}{2m\Delta_z}) \exp(-i\frac{\pi}{m}k) D_k \left[\left[\exp(i\frac{\pi(m-1)}{m}) \frac{\phi(\Delta_z(j+\frac{1}{2}))}{\Delta_z(j+\frac{1}{2})} \right]_{j=0,1,\ldots,m-1} \right].$$

This expression provides a possibility for an efficient approximation of the CDF $F_X(x)$ of X at a given set of points x, when using the FFT method for the required DFT D_k. Aggregating the above steps finally gives

$$F_X\left(\frac{\pi(2k-m+1)}{m\Delta_z}\right) = \frac{1}{2} - \frac{1}{\pi}\cdot\Im\left(\Delta_z \exp(i\pi\frac{(m-1)s_z}{m\Delta_z})\exp(-i\frac{2\pi}{m}\frac{s_z}{\Delta_z}k)\right.$$
$$\left.\cdot\, D_k\left[\left[\exp(i\frac{\pi(m-1)}{m})\frac{\phi(s_z+j\Delta_z)}{(s_z+j\Delta_z)}\right]_{j=0,1,\ldots,m-1}\right]\right),$$
$$k = 0, 1, \ldots, m-1\,.$$

In order to give an impression of the possible precision of the method, the implied grid coefficients for fixed m and a chosen Δ_z according to the above conventions are summarized below

$$\Delta_z \;,\; s_z = \frac{\Delta_z}{2} \quad (z\text{-grid})$$
$$\Delta_x = \frac{2\pi}{m\Delta_z} \;,\; s_x = -\Delta_x\cdot\left(\frac{m}{2}-\frac{1}{2}\right) = -\frac{\pi(m-1)}{m\Delta_z} \quad (x\text{-grid})\,.$$

One has to repeatedly stress the basic improvement of this adaptation of the Gil-Pelaez inversion theorem to a DFT scheme in comparison to the one in equation (5.12) here. It lies in the fact that this approach does not rely on an arbitrary choice of a value ρ coming from an existing domain G of $M_X(u)$, which is not easy to ensure. The main application of CDFs is related to quantiles, alternative risk measures and statistical goodness of fit testing, which is why this approximation scheme is of exceptional relevance in the forthcoming sections. Finally, one has to mention another alternative approach for symmetric α-stable distributions in this context suggested by Zieliński [Ziel01], which is likewise based on the Gil-Pelaez inversion formula.

5.3.4 Stock Model

When considering the MVNTS distribution for joint stock returns R with parameter tuple $(a, \lambda, \gamma, \mathsf{P}, c, d)$, it is again made use of the fact that MVNTS results from a simply rescaled stdMVNTS. This particular feature, besides the properties of linear combinations of MVNTS distributions, enables the following decomposition of the estimation procedure. Such a decomposition proves to be necessary as a simultaneous estimation of the entire parameter set is not tractable in any way. The high dimensionality together with the given fact that only numerical approximations of univariate PDFs are available would render this direct approach rather non-robust. Therefore, the decomposition basically aims at desegregating the problem into single one-dimensional estimations on standardized data in order to reduce the number of variables to be simultaneously estimated as far as possible.

For improved clarity the decomposition and subsequent estimation of the $MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$ distribution assumed for the multivariate random vector R is presented in the following four separate steps, where $N = 2514$ denotes the number of daily samples recorded for $n = 29$ stocks for this purpose.

1.) Subordinator dynamics:

The parameters (a, λ) specifying the originally underlying subordinator dynamics in the MVNTS distribution offer the advantage of being maintained under both linear combinations and simple linear transformations. By this argument the univariate random variable Q

$$Q = \frac{(\omega^{\mathsf{T}} R) - E(\omega^{\mathsf{T}} R)}{\sqrt{Var(\omega^{\mathsf{T}} R)}},$$

where ω denotes a non-trivial vector in $\mathbb{R}^n$, is $stdNTS(a, \lambda, g)$ distributed. The linear combination with ω converts the initial MVNTS assumption in a univariate NTS assumption and, furthermore, the subsequent linear standardization implies the stdNTS assumption for Q, where the last transition is justified by the arguments given in the above section 5.3.2. Afterwards, the values in the parameter set are estimated on the

correspondingly transformed empirical data

$$q = \frac{(\omega^{\mathsf{T}} r) - \omega^{\mathsf{T}} \bar{r}}{Std(\omega^{\mathsf{T}} r)} \, .$$

One has to note, however, that the specific choice of ω will slightly influence the parameter estimates $(\hat{a}, \hat{\lambda}, \hat{g})$[12]. For this reason and in order to ensure that the index return model can be incorporated later on into this model for joint stock returns, it would be appropriate to employ the index returns $R^{(ind)} = \omega_{DJIA}^{\mathsf{T}} R$, which are obtained by the DJIA weighting ω_{DJIA}[13]. This implies that $(\hat{a}, \hat{\lambda})$ for the joint stock model are equal to the ones estimated on the standardized index return observations given in table 5.2. Another natural choice for ω would be an equally weighted vector of total mass one. In the end, both results only differ by a very small extent.

2.) Multivariate linear standardization:

After (a, λ) was estimated as described above, the estimate of (c, d) is determined in this step. Fortunately the parameters (c, d) are directly linked to the inherent scales of the MVNTS distribution or rather the expectation and standard deviations of their marginal distributions, respectively. Hence, $(\hat{c}, \hat{d})$ can be simply determined by the following familiar unbiased estimators

$$\hat{d} = \bar{r} = \left(\bar{r}^{(1)}, \bar{r}^{(2)}, \ldots, \bar{r}^{(29)} \right)^{\mathsf{T}}$$

$$\hat{c} = \sqrt{Var(r)} = \frac{1}{\sqrt{N-1}} \left(\sqrt{\sum_{t=1}^{N} (r_t^{(1)} - \bar{r}^{(1)})^2}, \ldots, \sqrt{\sum_{t=1}^{N} (r_t^{(29)} - \bar{r}^{(29)})^2} \right)^{\mathsf{T}} ,$$

with resulting numerical values given in table 5.3.

[12] This is rather a practical issue concerning the appropriateness of the distributional assumption and the randomness in the sample than a theoretical one.

[13] Although the weights in the DJIA vary slightly over time this approach still seems admissible here. An exemplary weighting ω_{DJIA} can be found in appendix C.1.

k	$\hat{d}_k$	$\hat{c}_k$
1	$4.146 \cdot 10^{-4}$	0.0164
2	$2.542 \cdot 10^{-4}$	0.0275
3	$1.461 \cdot 10^{-4}$	0.0245
4	$1.176 \cdot 10^{-4}$	0.0207
5	$2.146 \cdot 10^{-4}$	0.0262
6	$3.962 \cdot 10^{-4}$	0.0212
7	$5.950 \cdot 10^{-4}$	0.0217
8	$5.183 \cdot 10^{-4}$	0.0180
9	$3.285 \cdot 10^{-4}$	0.0308
10	$5.467 \cdot 10^{-5}$	0.0162
11	$2.017 \cdot 10^{-5}$	0.0193
12	$5.565 \cdot 10^{-4}$	0.0179
13	$1.124 \cdot 10^{-5}$	0.0203
14	$5.439 \cdot 10^{-4}$	0.0277
15	$1.101 \cdot 10^{-4}$	0.0239

k	$\hat{d}_k$	$\hat{c}_k$
16	$1.962 \cdot 10^{-4}$	0.0295
17	$2.059 \cdot 10^{-4}$	0.0203
18	$3.239 \cdot 10^{-4}$	0.0147
19	$3.368 \cdot 10^{-4}$	0.0270
20	$4.231 \cdot 10^{-4}$	0.0182
21	$1.881 \cdot 10^{-7}$	0.0203
22	$9.823 \cdot 10^{-5}$	0.0228
23	$-5.574 \cdot 10^{-5}$	0.0190
24	$3.415 \cdot 10^{-4}$	0.0164
25	$4.659 \cdot 10^{-4}$	0.0227
26	$5.352 \cdot 10^{-4}$	0.0200
27	$1.622 \cdot 10^{-4}$	0.0196
28	$3.475 \cdot 10^{-4}$	0.0191
29	$1.924 \cdot 10^{-4}$	0.0226

Table 5.3: Estimates $(\hat{c}, \hat{d})$ of empirical stock returns.

One further removes these inherent scales of a MVNTS distribution by a simple linear transformation of the random vector, where the necessary coefficients can be directly obtained from (c, d) or their corresponding estimates, respectively, available at this point. Based on the initial assumption of $R_t \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$ with already determined estimates $(\hat{a}, \hat{\lambda}, \hat{c}, \hat{d})$ this leads to the assumption of $\tilde{R} \sim stdMVNTS((\hat{a}, \hat{\lambda}, \gamma, \mathsf{P}) \equiv MVNTS((\hat{a}, \hat{\lambda}, \gamma, \mathsf{P}, \mathbf{1}, \mathbf{0})$ for

$$\tilde{R} = \mathrm{diag}(\hat{c}^{-1})(R - \hat{d})\,,\ \hat{c}^{-1} = \left(\hat{c}_1^{-1}, \ldots, \hat{c}_n^{-1}\right)^{\mathsf{T}}$$

$$\tilde{r}_t = \mathrm{diag}(\hat{c}^{-1})(r_t - \hat{d})\,,\ t = 1, \ldots, N\,.$$

Note that we do not have to deal with the explicit equivalence of the linearly standardized MVNTS and the stdMVNTS here as was necessary for the univariate index returns in section 5.3.2. The reason for this is that such a linear standardization of any MVNTS distribution just sets its parameter values (c, d) to $(\mathbf{1}, \mathbf{0})$ which directly corresponds to a stdMVNTS distribution in MVNTS parametrization. This is straightforward by the construction of MVNTS via rescaled stdMVNTS and does not require

further calculations.

3.) Estimation of γ:

Unaffected by the still unknown P, each single dimension $\tilde{R}_k$, $k = 1, \ldots, n$, or marginal distribution of the multivariate stdMVNTS distributed $\tilde{R}$ obeys a univariate $stdNTS(a, \lambda, \gamma_k)$ distribution, which is again due to the Brownian subordination[14]. By this relation, the single dimensions γ_k of γ can be estimated by the maximum likelihood method on the corresponding dimensions of standardized empirical stock return data $\tilde{r}^{(k)}$ under fixed parameter values $(\hat{a}, \hat{\lambda})$. The results together with their 95% confidence intervals (CI) are given in the table 5.4.

k	$\hat{\gamma}_k$	lower CI bound.	upper CI bound.
1	0.0578	0.0131	0.1025
2	0.0357	-0.0105	0.0819
3	0.0170	-0.0279	0.0619
4	0.0193	-0.0251	0.0637
5	0.0147	-0.0354	0.0648
6	0.0022	-0.0431	0.0475
7	0.0185	-0.0264	0.0633
8	-0.0254	-0.0728	0.0219
9	0.0088	-0.0366	0.0543
10	0.0310	-0.0146	0.0767
11	0.0697	0.0249	0.1145
12	-0.0314	-0.0784	0.0155
13	0.0333	-0.0129	0.0794
14	0.0341	-0.0111	0.0793
15	0.0588	0.0137	0.1039

k	$\hat{\gamma}_k$	lower CI bound.	upper CI bound.
16	0.0162	-0.0279	0.0604
17	0.0153	-0.0309	0.0615
18	0.0512	0.0057	0.0966
19	0.0676	0.0219	0.1132
20	0.0268	-0.0181	0.0718
21	-0.0269	-0.0741	0.0204
22	0.0427	-0.0027	0.0880
23	0.0255	-0.0193	0.0703
24	0.0017	-0.0480	0.0513
25	0.0678	0.0209	0.1148
26	0.0148	-0.0313	0.0609
27	0.0411	-0.0038	0.0860
28	0.0637	0.0193	0.1081
29	0.0395	-0.0066	0.0855

Table 5.4: Estimate for γ (stdMVNTS) of standardized stock returns.

4.) Calibration of P:

As the realizations of the subordinator $S_{(t)}$ and the Brownian motion $B_{(s)}$ cannot be observed themselves, there is no possibility for a direct estimation of the correlation matrix P of the Brownian motion $B_{(s)}$ at hand. Rather, one has to calibrate P under the given estimates of the remaining parameter values such that the implied correlation in $\tilde{R}$ matches the correlation of the standardized empirical stock return data in $\tilde{r}$. Recall from equation (4.21) that the correlation matrix of a MVNTS distributed random

[14] See section 4.2.1.

vector R, which is equal to the covariance matrix of the associated linearly standardized stdMVNTS random vector $\tilde{R}$ is

$$Corr(R) = Cov(\tilde{R}) = \text{diag}(\tilde{\sigma}_{(a,\lambda,\gamma)})\mathsf{P}\text{diag}(\tilde{\sigma}_{(a,\lambda,\gamma)}) + \frac{2-a}{2\lambda}\gamma\gamma^{\mathsf{T}}$$

$$\tilde{\sigma}_{(a,\lambda,\gamma)} = \left[\sqrt{1 - \frac{2-a}{2\lambda}\gamma_k^2}\right]_{k=1,2,\ldots,n}^{\mathsf{T}}.$$

In order to match the theoretical correlation with its empirical counterpart while considering the current estimates $(\hat{a}, \hat{\lambda}, \hat{\gamma}, \hat{c}, \hat{d})$ the following relation has to hold

$$Corr(r) = Cov(\tilde{r}) \stackrel{!}{=} \text{diag}(\tilde{\sigma}_{(\hat{a},\hat{\lambda},\hat{\gamma})})\hat{\mathsf{P}}\text{diag}(\tilde{\sigma}_{(\hat{a},\hat{\lambda},\hat{\gamma})}) + \frac{2-\hat{a}}{2\hat{\lambda}}\hat{\gamma}\hat{\gamma}^{\mathsf{T}}$$

from which the appropriately calibrated $\hat{\mathsf{P}}$ is finally resolved

$$\Rightarrow \hat{\mathsf{P}} = \text{diag}\left(\tilde{\sigma}^{-1}_{(\hat{a},\hat{\lambda},\hat{\gamma})}\right)\left[Cov(\tilde{r}) - \frac{2-\hat{a}}{2\hat{\lambda}}\hat{\gamma}\hat{\gamma}^{\mathsf{T}}\right]\text{diag}\left(\tilde{\sigma}^{-1}_{(\hat{a},\hat{\lambda},\hat{\gamma})}\right) \tag{5.15}$$

$$\tilde{\sigma}^{-1}_{(\hat{a},\hat{\lambda},\hat{\gamma})} = \left[\left(1 - \frac{2-a}{2\lambda}\gamma_k^2\right)^{-1/2}\right]_{k=1,2,\ldots,n}^{\mathsf{T}}.$$

The numerical evaluation of the above expression on the underlying empirical data set leads to the calibrated Brownian correlation matrix in table 5.5.

After having performed these steps, we have arrived at being able to entirely specify the estimated MVNTS parameters for the joint unconditional distribution of stock returns as well as the stdMVNTS parameters for their standardized counterpart. Moreover, this is synonymous to the complete specification of the stated simple stationary model defined in section 5.1. After assessing the adequacy of its obtained statistical fit in the next section this model will be put to a practical application in the field of portfolio analysis and optimization, where it is made use of its risk prediction capabilities.

$$\hat{\mathrm{P}} = \begin{pmatrix}
1 & 0.455 & 0.451 & 0.316 & 0.403 & 0.394 & 0.49 & 0.394 & 0.294 & 0.333 & 0.544 & 0.442 & 0.47 & 0.268 & 0.391 & 0.315 & 0.308 & 0.338 & 0.399 & 0.281 & 0.286 & 0.287 & 0.333 & 0.346 & 0.411 & 0.526 & 0.301 & 0.363 & 0.352 \\
0.455 & 1 & 0.447 & 0.344 & 0.374 & 0.414 & 0.539 & 0.508 & 0.311 & 0.259 & 0.574 & 0.493 & 0.463 & 0.302 & 0.361 & 0.33 & 0.332 & 0.246 & 0.389 & 0.297 & 0.282 & 0.347 & 0.321 & 0.234 & 0.376 & 0.472 & 0.348 & 0.276 & 0.414 \\
0.451 & 0.447 & 1 & 0.443 & 0.674 & 0.403 & 0.464 & 0.396 & 0.43 & 0.338 & 0.49 & 0.406 & 0.638 & 0.379 & 0.513 & 0.428 & 0.432 & 0.361 & 0.678 & 0.33 & 0.359 & 0.436 & 0.405 & 0.306 & 0.533 & 0.495 & 0.414 & 0.451 & 0.496 \\
0.316 & 0.344 & 0.443 & 1 & 0.377 & 0.294 & 0.313 & 0.339 & 0.293 & 0.309 & 0.364 & 0.398 & 0.385 & 0.254 & 0.32 & 0.283 & 0.322 & 0.329 & 0.401 & 0.268 & 0.326 & 0.328 & 0.333 & 0.3 & 0.332 & 0.326 & 0.68 & 0.342 & 0.339 \\
0.403 & 0.374 & 0.674 & 0.377 & 1 & 0.346 & 0.42 & 0.351 & 0.357 & 0.264 & 0.449 & 0.349 & 0.558 & 0.291 & 0.46 & 0.35 & 0.374 & 0.289 & 0.755 & 0.287 & 0.301 & 0.353 & 0.357 & 0.283 & 0.516 & 0.434 & 0.362 & 0.37 & 0.369 \\
0.394 & 0.414 & 0.403 & 0.294 & 0.346 & 1 & 0.384 & 0.334 & 0.285 & 0.274 & 0.423 & 0.364 & 0.409 & 0.29 & 0.357 & 0.311 & 0.273 & 0.283 & 0.366 & 0.281 & 0.279 & 0.322 & 0.3 & 0.244 & 0.362 & 0.537 & 0.292 & 0.323 & 0.364 \\
0.49 & 0.539 & 0.464 & 0.313 & 0.42 & 0.384 & 1 & 0.411 & 0.304 & 0.278 & 0.56 & 0.415 & 0.467 & 0.302 & 0.412 & 0.335 & 0.336 & 0.287 & 0.433 & 0.28 & 0.267 & 0.345 & 0.296 & 0.262 & 0.387 & 0.517 & 0.29 & 0.354 & 0.372 \\
0.394 & 0.508 & 0.396 & 0.339 & 0.351 & 0.334 & 0.411 & 1 & 0.212 & 0.296 & 0.433 & 0.827 & 0.349 & 0.237 & 0.283 & 0.246 & 0.258 & 0.342 & 0.331 & 0.254 & 0.334 & 0.326 & 0.361 & 0.241 & 0.382 & 0.398 & 0.337 & 0.268 & 0.347 \\
0.294 & 0.311 & 0.43 & 0.293 & 0.357 & 0.285 & 0.304 & 0.212 & 1 & 0.168 & 0.311 & 0.23 & 0.474 & 0.517 & 0.363 & 0.636 & 0.516 & 0.171 & 0.436 & 0.247 & 0.178 & 0.566 & 0.239 & 0.149 & 0.305 & 0.356 & 0.315 & 0.308 & 0.393 \\
0.333 & 0.259 & 0.338 & 0.309 & 0.264 & 0.274 & 0.278 & 0.296 & 0.168 & 1 & 0.319 & 0.352 & 0.306 & 0.162 & 0.289 & 0.192 & 0.178 & 0.369 & 0.253 & 0.273 & 0.307 & 0.25 & 0.323 & 0.395 & 0.268 & 0.311 & 0.302 & 0.308 & 0.249 \\
0.544 & 0.574 & 0.49 & 0.364 & 0.449 & 0.423 & 0.56 & 0.433 & 0.311 & 0.319 & 1 & 0.462 & 0.5 & 0.298 & 0.411 & 0.325 & 0.329 & 0.317 & 0.45 & 0.326 & 0.322 & 0.317 & 0.343 & 0.357 & 0.42 & 0.527 & 0.342 & 0.379 & 0.399 \\
0.442 & 0.493 & 0.406 & 0.398 & 0.349 & 0.364 & 0.415 & 0.827 & 0.23 & 0.352 & 0.462 & 1 & 0.384 & 0.253 & 0.314 & 0.278 & 0.301 & 0.379 & 0.343 & 0.272 & 0.355 & 0.353 & 0.397 & 0.281 & 0.381 & 0.423 & 0.378 & 0.316 & 0.359 \\
0.47 & 0.463 & 0.638 & 0.385 & 0.558 & 0.409 & 0.467 & 0.349 & 0.474 & 0.306 & 0.5 & 0.384 & 1 & 0.383 & 0.5 & 0.442 & 0.469 & 0.37 & 0.617 & 0.329 & 0.358 & 0.464 & 0.421 & 0.337 & 0.467 & 0.527 & 0.411 & 0.453 & 0.485 \\
0.268 & 0.302 & 0.379 & 0.254 & 0.291 & 0.29 & 0.302 & 0.237 & 0.517 & 0.162 & 0.298 & 0.253 & 0.383 & 1 & 0.318 & 0.517 & 0.493 & 0.146 & 0.373 & 0.223 & 0.198 & 0.455 & 0.21 & 0.126 & 0.252 & 0.352 & 0.27 & 0.262 & 0.386 \\
0.391 & 0.361 & 0.513 & 0.32 & 0.46 & 0.357 & 0.412 & 0.283 & 0.363 & 0.289 & 0.411 & 0.314 & 0.5 & 0.318 & 1 & 0.362 & 0.357 & 0.272 & 0.499 & 0.343 & 0.289 & 0.376 & 0.322 & 0.267 & 0.375 & 0.43 & 0.323 & 0.591 & 0.405 \\
0.315 & 0.33 & 0.428 & 0.283 & 0.35 & 0.311 & 0.335 & 0.246 & 0.636 & 0.192 & 0.325 & 0.278 & 0.442 & 0.517 & 0.362 & 1 & 0.514 & 0.202 & 0.427 & 0.221 & 0.23 & 0.598 & 0.249 & 0.18 & 0.315 & 0.355 & 0.315 & 0.312 & 0.38 \\
0.308 & 0.332 & 0.432 & 0.322 & 0.374 & 0.273 & 0.336 & 0.258 & 0.516 & 0.178 & 0.329 & 0.301 & 0.469 & 0.493 & 0.357 & 0.514 & 1 & 0.233 & 0.431 & 0.239 & 0.227 & 0.479 & 0.259 & 0.172 & 0.319 & 0.36 & 0.32 & 0.316 & 0.397 \\
0.338 & 0.246 & 0.361 & 0.329 & 0.289 & 0.283 & 0.287 & 0.342 & 0.171 & 0.369 & 0.317 & 0.379 & 0.37 & 0.146 & 0.272 & 0.202 & 0.233 & 1 & 0.265 & 0.269 & 0.499 & 0.273 & 0.526 & 0.405 & 0.292 & 0.312 & 0.299 & 0.328 & 0.283 \\
0.399 & 0.389 & 0.678 & 0.401 & 0.755 & 0.366 & 0.433 & 0.331 & 0.436 & 0.253 & 0.45 & 0.343 & 0.617 & 0.373 & 0.499 & 0.427 & 0.431 & 0.265 & 1 & 0.285 & 0.328 & 0.426 & 0.342 & 0.241 & 0.498 & 0.46 & 0.407 & 0.403 & 0.435 \\
0.281 & 0.297 & 0.33 & 0.268 & 0.287 & 0.281 & 0.28 & 0.254 & 0.247 & 0.273 & 0.326 & 0.272 & 0.329 & 0.223 & 0.343 & 0.221 & 0.239 & 0.269 & 0.285 & 1 & 0.26 & 0.225 & 0.274 & 0.293 & 0.296 & 0.321 & 0.253 & 0.331 & 0.289 \\
0.286 & 0.282 & 0.359 & 0.326 & 0.301 & 0.279 & 0.267 & 0.334 & 0.178 & 0.307 & 0.322 & 0.355 & 0.358 & 0.198 & 0.289 & 0.23 & 0.227 & 0.499 & 0.328 & 0.26 & 1 & 0.264 & 0.551 & 0.331 & 0.32 & 0.301 & 0.319 & 0.306 & 0.281 \\
0.287 & 0.347 & 0.436 & 0.328 & 0.353 & 0.322 & 0.345 & 0.326 & 0.566 & 0.25 & 0.317 & 0.353 & 0.464 & 0.455 & 0.376 & 0.598 & 0.479 & 0.273 & 0.426 & 0.225 & 0.264 & 1 & 0.312 & 0.182 & 0.342 & 0.388 & 0.348 & 0.354 & 0.403 \\
0.333 & 0.321 & 0.405 & 0.333 & 0.357 & 0.322 & 0.345 & 0.326 & 0.239 & 0.323 & 0.343 & 0.397 & 0.421 & 0.21 & 0.322 & 0.249 & 0.259 & 0.526 & 0.342 & 0.274 & 0.551 & 0.312 & 1 & 0.341 & 0.321 & 0.356 & 0.331 & 0.366 & 0.319 \\
0.346 & 0.234 & 0.306 & 0.3 & 0.283 & 0.244 & 0.262 & 0.241 & 0.149 & 0.395 & 0.357 & 0.281 & 0.337 & 0.126 & 0.267 & 0.18 & 0.172 & 0.405 & 0.241 & 0.293 & 0.331 & 0.182 & 0.341 & 1 & 0.272 & 0.333 & 0.273 & 0.321 & 0.209 \\
0.411 & 0.376 & 0.533 & 0.332 & 0.516 & 0.362 & 0.387 & 0.382 & 0.305 & 0.268 & 0.42 & 0.381 & 0.467 & 0.252 & 0.375 & 0.315 & 0.319 & 0.292 & 0.498 & 0.296 & 0.32 & 0.342 & 0.321 & 0.272 & 1 & 0.437 & 0.322 & 0.344 & 0.349 \\
0.526 & 0.472 & 0.495 & 0.326 & 0.434 & 0.537 & 0.517 & 0.398 & 0.356 & 0.311 & 0.527 & 0.423 & 0.527 & 0.352 & 0.43 & 0.355 & 0.36 & 0.312 & 0.46 & 0.321 & 0.301 & 0.388 & 0.356 & 0.333 & 0.437 & 1 & 0.338 & 0.388 & 0.423 \\
0.301 & 0.348 & 0.414 & 0.68 & 0.362 & 0.292 & 0.29 & 0.337 & 0.315 & 0.302 & 0.342 & 0.378 & 0.411 & 0.27 & 0.323 & 0.315 & 0.32 & 0.299 & 0.407 & 0.253 & 0.319 & 0.348 & 0.331 & 0.273 & 0.322 & 0.338 & 1 & 0.325 & 0.357 \\
0.363 & 0.276 & 0.451 & 0.342 & 0.37 & 0.323 & 0.354 & 0.268 & 0.308 & 0.308 & 0.379 & 0.316 & 0.453 & 0.262 & 0.591 & 0.312 & 0.316 & 0.328 & 0.403 & 0.331 & 0.306 & 0.354 & 0.366 & 0.321 & 0.344 & 0.388 & 0.325 & 1 & 0.331 \\
0.352 & 0.414 & 0.496 & 0.339 & 0.369 & 0.364 & 0.372 & 0.347 & 0.393 & 0.249 & 0.399 & 0.359 & 0.485 & 0.386 & 0.405 & 0.38 & 0.397 & 0.283 & 0.435 & 0.289 & 0.281 & 0.403 & 0.319 & 0.209 & 0.349 & 0.423 & 0.357 & 0.331 & 1
\end{pmatrix} \tag{5.16}$$

Table 5.5: Calibrated $\hat{\mathrm{P}}$ (stdMVNTS) of standardized stock returns.

5.4 Goodness of Fit Tests

For assessing the goodness of fit (GoF) of the estimated distributions associated with the above simple models for both index and joint stock returns a number of different methods will be applied in this section. Basic theory on formal statistical goodness of fit tests are provided in many standard textbooks on the topic. Moreover, an additional overview is given in appendix D. Since in case of the fitted index return model we have to deal with only univariate data, even graphical measures besides the conventional statistical test procedures are become available. For the multivariate model of joint stock returns however, only formal goodness-of-fit measures such as the Kolmogorov-Smirnov (KS) distance and Anderson-Darling (AD) statistics are applicable. One has to be well aware of the fact that these standard tests are only able to consider the marginal distributions of the fitted MVNTS, whereas the adequacy of the modeled dependencies structure between single dimensions of the multivariate distribution cannot be validated. The difficulty here basically lies in the comparably high number of considered model dimensions. Otherwise, the task of assessing the multivariate dependence structure could be addressed by the entire set of χ^2 tests on the joint distribution of pairwise projections (R_j, R_k), $j = 1, \ldots, n$, $k > j$. However, this approach likewise does not prove to be very tractable either, as $\frac{n(n-1)}{2} = 406$ existing pairs have to be considered. An opportunity of at least gaining a particular and informal revision is to perform these 1-dimensional GoF tests for linear combinations of MVNTS components with arbitrary but suitable weights. There, the resulting theoretical distribution implied by the estimated MVNTS is tested against the empirical counterpart of the linear combination.

5.4.1 Index Model

Starting with the index model and a visual examination first, figure 5.1 displays a graph of the standardized empirical index return kernel density. This density is furthermore compared to the probability density functions of the fitted stdNTS[15] vs. a standard Normal

[15] Its associated parameter values were given in table 5.2.

distribution. The plot clearly demonstrates how the NTS, in contrast to the standard Normal distribution, is able to follow the kernel densities peaked form with heavier weights in the border areas of the distribution support. Skewness is not perceptibly present in this data set of historical index returns. Hence, although the standard Normal distribution in contrast to the more flexible stdNTS is not able to account for asymmetry, this particular aspect is not responsible for its inferior fitting ability, which is rather due to the aforementioned inferior adaptation to peakedness and tail features.

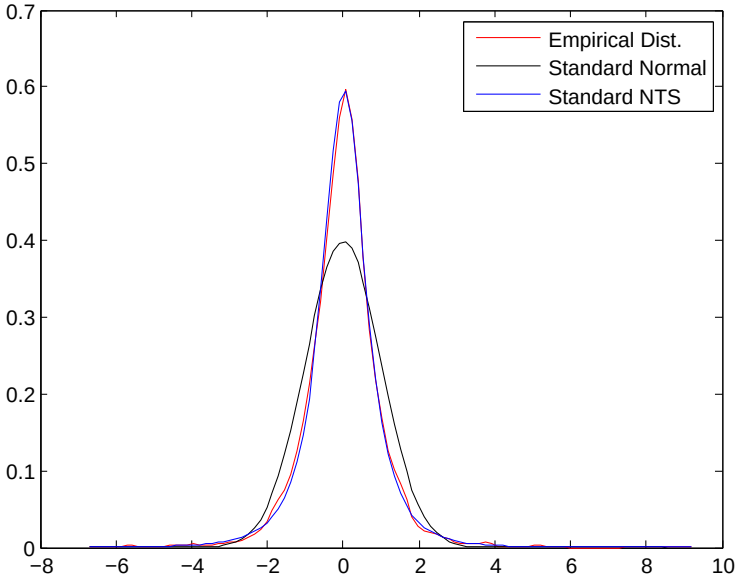
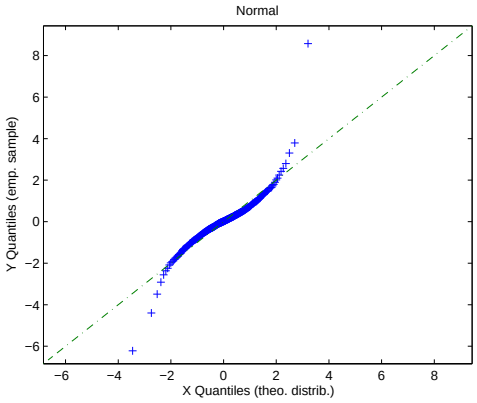


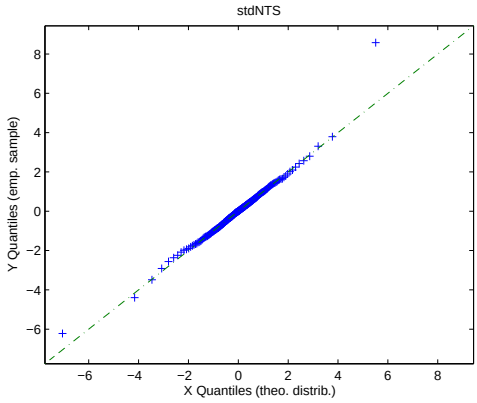
Figure 5.1: Kernel densities of standardized empirical DJIA index returns, compared to fitted NTS and Normal densities.

The second means of graphical inspection are the two separate quantile-quantile plots shown in figure 5.2. In accordance to the density graphs, these plots confirm the NTS ability to better fit the empirical data than the Normal distribution, which is true both in the peaked center and the particularly pronounced tails of the distribution. The Q-Q plots furthermore allow for a more detailed inspection of the tails than the kernel densities in regions of low observation frequencies.

Although these graphical approaches already strongly suggest superiority of the NTS with



(a) Quantile-quantile plot of standardized empirical returns vs. standard Normal distribution assumption.



(b) Quantile-quantile plot of standardized empirical returns vs. standard NTS distribution assumption.

Figure 5.2: Quantile-quantile plots of standardized empirical returns.

regard to its adaptability over the Normal distribution, this finding is further substantiated by the following results of conducted formal tests, them having the advantage of an unambiguous interpretation of the output[16]. The NTS distribution fitted to the index return distributions generate the test statistics presented in table 5.6 together with their corresponding p-values.

	KS	AD	AD^2
statistics	0.0133	0.1051	0.3957
p-value	0.7618	0.0769	0.8530

Table 5.6: Results Goodness-of-fit tests of NTS for empirical index returns.

While the Kolmogorov-Smirnov test (KS) pays more attention to the center of the distribution, both the conventional Anderson-Darling test (AD) as well as the quadratic Anderson-Darling test (AD^2) are designed for detecting deviations in the tails. Thus, from the presented p-values one is able to conclude that the NTS constitutes a very convincing model for both the center and especially the tails of the distribution.

[16]Which is present in form of p-values.

5.4.2 Joint Stock Model

The same tests can furthermore be carried out for each of the 29 stocks and their marginal distributions separately, for which the following results in table 5.7 are achieved. They confirm with only few exceptions a rather good fit of the NTS for the marginal distributions of joint stock returns. One is able to observe that in over 75% of the cases the stock generates sufficiently small values of KS, AD and AD^2, respectively, whose associated p-values are given in table 5.8. As the distributions of these statistics, which are necessary for converting the obtained statistics into corresponding p-values, depend on the sample size $N = 2514$ and cannot be stated in closed form, one has to resort to a simulative approach [MaMa04][17].

k	KS	AD	AD^2
1	0.0177	0.0502	0.6589
2	0.0158	0.0650	0.8388
3	0.0226	0.0487	1.3246
4	0.0164	0.0650	0.9362
5	0.0712	0.1525	22.7916
6	0.0269	0.0669	3.6305
7	0.0297	0.0737	3.9486
8	0.0308	0.1718	4.0562
9	0.0149	0.0594	0.5124
10	0.0156	0.0400	0.5913
11	0.0141	0.0498	0.7722
12	0.0285	0.0790	2.6815
13	0.0203	0.0560	0.6578
14	0.0123	0.0390	0.4716
15	0.0130	0.2639	0.6362

k	KS	AD	AD^2
16	0.0190	0.0507	1.7242
17	0.0142	0.0522	0.3491
18	0.0155	0.1499	0.5271
19	0.0374	0.0799	4.9703
20	0.0321	0.0772	4.2806
21	0.0119	0.2241	0.5309
22	0.0208	0.0584	0.9170
23	0.0232	0.0508	1.8334
24	0.0306	2.1532	4.4743
25	0.0265	0.0878	1.5562
26	0.0130	0.4183	0.7881
27	0.0172	0.0438	1.2699
28	0.0185	0.0445	1.1662
29	0.0157	0.0456	0.4414

Table 5.7: Goodness-of-fit statistics DJIA stocks (MVNTS).

5.5 Portfolio Analysis and Optimization

Before addressing the task of static portfolio optimization in the estimated stationary model, risk and performance of a portfolio with a specific natural composition is investigated in further detail. Besides the consideration of diversification effects under a ho-

[17]The asymptotic distributions ($N \to \infty$) for both AD and AD^2 statistics were already stated in the original publication of Anderson and Darling [AnDa52].

k	KS	AD	AD^2
1	0.4067	0.5797	0.5940
2	0.5504	0.2669	0.4538
3	0.1520	0.6241	0.2246
4	0.5044	0.2669	0.3926
5	0.0000	0.0351	0.0000
6	0.0522	0.2438	0.0133
7	0.0229	0.1824	0.0093
8	0.0165	0.0274	0.0082
9	0.6284	0.3557	0.7341
10	0.5669	0.8730	0.6562
11	0.6911	0.5914	0.5014
12	0.0330	0.1506	0.0399
13	0.2501	0.4271	0.5949
14	0.8378	0.8954	0.7758
15	0.7847	0.0114	0.6142

k	KS	AD	AD^2
16	0.3221	0.5653	0.1310
17	0.6856	0.5232	0.8971
18	0.5799	0.0363	0.7193
19	0.0017	0.1462	0.0030
20	0.0110	0.1602	0.0064
21	0.8629	0.0160	0.7154
22	0.2250	0.3753	0.4039
23	0.1314	0.5624	0.1137
24	0.0175	0.0002	0.0052
25	0.0577	0.1156	0.1637
26	0.7858	0.0046	0.4897
27	0.4428	0.7717	0.2424
28	0.3495	0.7511	0.2808
29	0.5629	0.7180	0.8068

Table 5.8: Goodness-of-fit p-values DJIA stocks (MVNTS).

mogeneous mixture of single stocks, the adequacy of dependence modeling of MVNTS distributions for empirical return data will also be assessed. This supplements the GoF tests for the marginal stock return distributions presented in the previous section.

5.5.1 Equal Weights Portfolio

When simply selecting an equally weighted portfolio (EWPF) composed of the 29 stocks, the resulting portfolio return can be determined by a linear combination with weight vector ω and

$$\omega = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)^{\mathsf{T}} = \left(\frac{1}{29}, \ldots, \frac{1}{29}\right)^{\mathsf{T}}$$

$$\omega^{\mathsf{T}}\mathbf{1} = \sum_{j=1}^{n} \omega_j = 1\,. \tag{5.17}$$

As was indicated at the beginning of this chapter, the return representation by simply compounded rates has the advantage that portfolio returns can be determined by

$$R_{ewpf} = \omega^{\mathsf{T}} R\,, \quad r_t^{(ewpf)} = \omega^{\mathsf{T}} r_t\,,\, t = 1, \ldots, N \tag{5.18}$$

with ω reflecting the portfolio weights. Moreover, ω has to satisfy the conditions of an affine combination, i.e. equation (5.17) has to hold true. Affine combinations embody a restriction of general linear combinations in that the sum of involved weights equals one. This corresponds to the situation that exactly the amount of reference capital or basis for the return calculation is indeed invested without any upward or downward deviations. Equation (5.18) covers both the series of realized EWPF returns $r_t^{(ewpf)}$ generated from the historical stock return data set contained in r, as well as the distribution assumption of i.i.d. EWPF returns reflected by the random variable R_{ewpf}. The latter one is implied by the estimated model in section 5.3.4 and the calculation rules valid for general linear combinations developed in section 4.2.4. In other words, from the now present model assumption of $R \sim MVNTS(\hat{a}, \hat{\lambda}, \hat{\gamma}, \hat{\mathsf{P}}, \hat{c}, \hat{d})$ follows that R_{ewpf} is a univariate random variable distributed according to $NTS(a_{(ewpf)}, \lambda_{(ewpf)}, g_{(ewpf)}, s_{(ewpf)}, m_{(ewpf)})$ with

$$(a_{(ewpf)}, \lambda_{(ewpf)}) = (\hat{a}, \hat{\lambda}) = (1.1584, 0.1812)$$

$$g_{(ewpf)} = \omega^{\mathsf{T}} \mathrm{diag}(\hat{c}) \hat{\gamma} = 5.9466 \cdot 10^{-4}$$

$$s_{(ewpf)} = \sqrt{\omega^{\mathsf{T}} \mathrm{diag}(\hat{c}) \mathrm{diag}\left(\sqrt{1 - \frac{2-\hat{a}}{2\hat{\lambda}}\hat{\gamma}^2}\right) \hat{\mathsf{P}} \mathrm{diag}\left(\sqrt{1 - \frac{2-\hat{a}}{2\hat{\lambda}}\hat{\gamma}^2}\right) \mathrm{diag}(\hat{c}) \omega}$$

$$= 0.0133$$

$$m_{(ewpf)} = \omega^{\mathsf{T}} \hat{d} = 2.7085 \cdot 10^{-4}\,.$$

Note however that the vector of parameter values for R_{ewpf} are not obtained by direct estimation on the series $r_t^{(ewpf)}$ but arise from the MVNTS distribution for joint stock returns. Hence, the uncertainties in $(a_{(ewpf)}, \lambda_{(ewpf)}, g_{(ewpf)}, s_{(ewpf)}, m_{(ewpf)})$ are subject to the randomness of the sampling as well as to possible misspecification of the MVNTS model.

As the portfolio return distribution comes from a homogeneously diversified linear com-

bination of MVNTS dimensions, its shape will be considerably influenced by present dependencies between the single stocks. For this reason, a good fit of the implied theoretical distribution of R_{ewpf} to the corresponding empirical data $r^{(ewpf)}$ would not only confirm a good fit of the marginal distributions of the single stocks contained in this stationary model, but would even more substantially serve as an indicator for an adequately captured dependence structure. Of course it must be clearly stated that this consistency check, as being based on one single weighting only, is thus not able to assess the modeled dependencies in a systematical way or to full extent.

Diagram 5.3 illustrates, as was already displayed for the index returns before, the kernel density of the historical EWPF returns. This kernel density is moreover compared to the theoretical probability densities of the assumed NTS distribution of R_{ewpf} and an implied Normal distribution. The parameter values of the latter were determined by the unbiased estimates of means and covariance matrix of R, which entirely describe a multivariate Normal distribution for the joint stock returns. These estimates were further processed by the familiar summation rules for correlated Gaussian random variables. The NTS assumption achieves an apparently precise fit in the center as well as at the tails of the distribution, although a slight difference in the skews must be noticed around the distribution center in 0. This nevertheless promising impression is further substantiated by the associated quantile-quantile plots in figure 5.4 which constitute a second possibility of visual validation. There, the NTS assumption does not cause any remarkable or systematic deviation from the diagonal, which would exceed the randomness of the sample. On the other hand, as was already demonstrated for the case of the index returns, the Normal assumption fails to capture both the peakedness (see fig. 5.3) and the heavy-tailedness (see fig. 5.4(a)) of the empirical EWPF return distribution. However, this result is presumably less due to an insufficient representation of the dependence structure but it is rather suspected that, as was previously indicated by the example of index returns, the Normal assumption is already inappropriate for a characterization of the marginal return distributions.

Now, the findings of the graphical diagnosis above are supplemented by the three familiar GoF statistics. But in contrast to the interpretation of these graphical results, one has

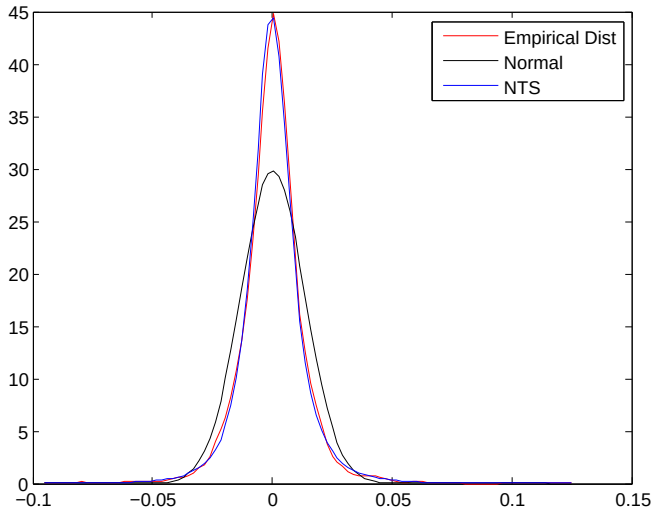
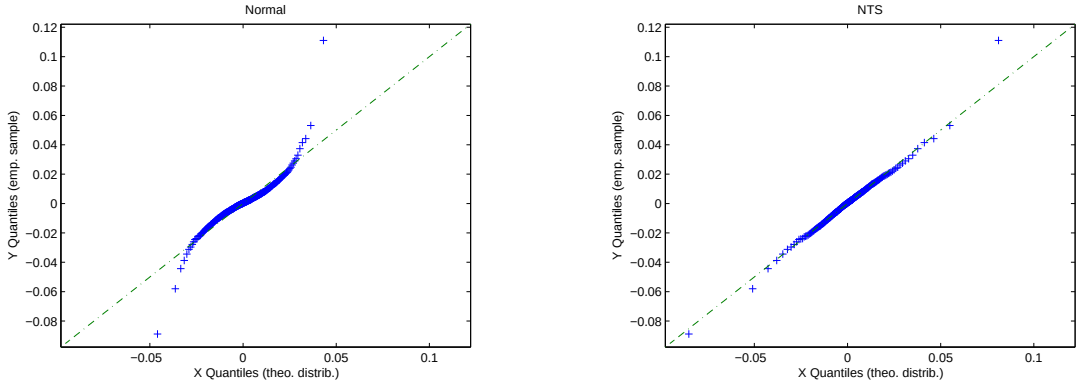


Figure 5.3: Kernel density of EWPF returns in comparison with NTS and Normal distribution.



(a) Quantile-quantile plot of empirical EWPF returns vs. Normal distribution assumption.

(b) Quantile-quantile plot of empirical EWPF returns vs. NTS distribution assumption.

Figure 5.4: Quantile-quantile plots for empirical EWPF returns based on Normal and NTS distributions.

to admit that none of the KS, AD and AD^2 statistics together with their respective p-values given in table 5.9 are able to fully support the MVNTS assumption, whereas a strong rejection of the assumption can not be justified at a reasonable significance level either. In particular, the poor KS statistic is due to the aforementioned slight offset of the densities in the center of the distribution, as was illustrated in figure 5.3. It is still an open question whether the deviations are primarily caused by deficits in represention of either the marginal distributions or the dependence structure.

	KS	AD	AD^2
statistics	0.0255	0.0613	1.3018
p-value	0.0750	0.3218	0.2318

Table 5.9: GoF results for EWPF.

To highlight the capabilities of the estimated model with regard to risk quantification and risk prediction, sequences of the widespread risk measures Value-at-Risk (VaR) and Average Value-at-Risk (AVaR)[18] for the EWPF are given in figure 5.5. These two risk measures are numerically evaluated for respective significance levels α[19] ranging between 0.1 and 5%. When comparing these sequences to their empirical equivalents for both risk measures[20], the MVNTS-based prediction clearly outperforms its Gaussian-based counterparts. This is especially true for low significance levels α in case of VaR as well as AVaR in general. Although this determination and comparison of risk measures is not declared as such, it is another confirming test, in the broader sense, of model adequacy concerning the tail quantiles and probabilities in particular.

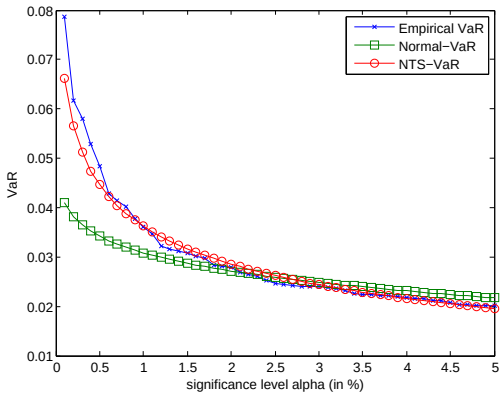
5.5.2 Portfolio Optimization

This study of risk assessment and quantification is continued with the transition from the rather naive EWPF to portfolios with optimized weights. In a first step, this optimization

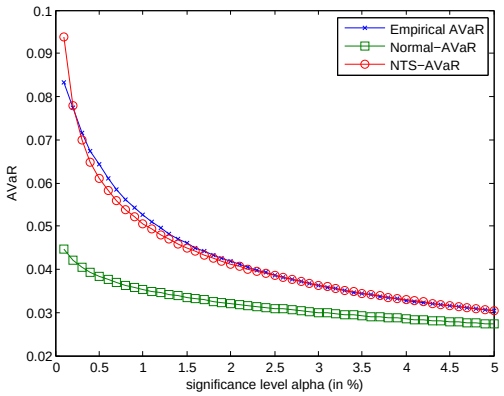
[18]For a definition of VaR and AVaR see for example [RaSF08, p. 182] and [RaSF08, p. 208], respectively, just to name a few. The latter one is also known in the literature as *Expected Shortfall* or *Conditional Value-at-Risk*.

[19]In alternative definitions, the above risk measures are specified by the confidence level $p = 1 - \alpha$ in the established literature.

[20]Their empirical definitions and calculation methods can be found in [RMFF$^+$07, p. 312] and [RaSF08, p. 214]



(a) VaR sequences for different significance levels α.



(b) AVaR sequences for different significance levels α.

Figure 5.5: Comparison of VaR and AVaR sequences for the EWPF.

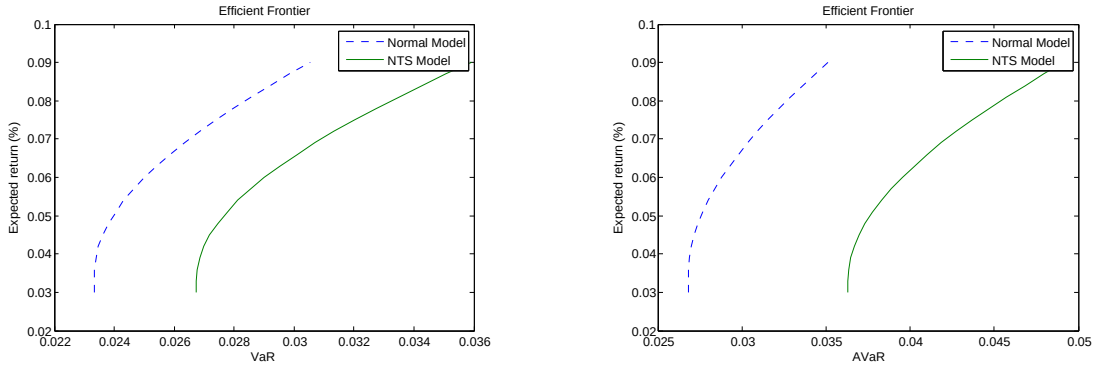
is carried out with respect to VaR and AVaR subject to several levels of expected return[21]. The following two diagrams contained in figure 5.6 show the resulting efficient frontiers from this optimization for the employed risk measure VaR and AVaR, respectively. They are generated under either the Normal or the Normal Tempered Stable model[22] for joint stock returns.

Although the efficient frontier does not imply a unique allocation decision yet, it represents the set of dominating[23] portfolio compositions available in the modeled stock market. Both diagrams demonstrate how the risk present in the market is underestimated under the Gaussian assumption. The MVNTS model in contrast has proven its superior ability to fit the empirical data under several criteria in section 5.4, thus generating more realistic representations of market risk. They surpass those obtained under the Gaussian model on average by around 30%. Therefore, using the Gaussian model instead of the MVNTS would even lead to two pitfalls at the same time. On the one hand, potential for beneficial diversification can not be fully exploited, as not being adequately modeled. Moreover, the risk

[21] See appendix C.2 for the precise formulation of the optimization problem.

[22] The explicit mention of the fact that the model refers to multivariate distributions should be occasionally dropped from now on for the sake of straightforward description. Hence, NTS becomes an abbreviation for Normal Tempered Stable models, univariate as well as multivariate ones, in general, including the multivariate case of the former MVNTS.

[23] As long as a usually risk-averse investor is assumed, who prefers higher returns over lower returns and lower risks over higher risks, with regard to the employed risk measures.



(a) Efficient frontiers with respect to VaR.

(b) Efficient frontiers with respect to AVaR.

Figure 5.6: Efficient frontiers for risk measure significance level $\alpha = 0.01$ under Gaussian vs. MVNTS assumption.

contained in the thereby constructed sub-optimal portfolio is additionally underestimated.

When moving on to a global portfolio optimization, it is necessary to aggregate the associated risk and return into a one-dimensional objective function. This task is accomplished by three different performance ratios to be considered in this study, given in table 5.10 [24]. The Sharpe ratio is solely defined in terms of distribution moments, i.e. no explicit distributional assumption has to be made while only estimates for means and covariances are required. But the Sharpe ratio, being a performance measure developed in the 1960ies, suffers from the shortcomings of its utilized risk measure, the standard deviation. For this reason, the two alternative performance measures VaR ratio and STAR ratio should be preferred for portfolio optimization and later backtesting. Their risk measures incorporate a stronger emphasis on downside risk, asymmetry and heavy tailed distributions. While AVaR is even contained in the class of coherent risk measures[25], VaR is not. In the following, their significance levels are set to $\alpha = 0.01$ again, in order to focus on the more extreme risks in the course of optimization.

For the forthcoming optimization, the following aspects have to be noted:

1.) Instead of conventional portfolio returns $R_p = \omega^\top R$ considered so far, one gives way

[24] Their definitions follow Rachev et al. [RaSF08].

[25] Coherence is a desired property of risk measures, see Artzner et al. [ADEH99] as well as Acerbi and Tasche [AcTa02] for further details.

	definition	numerator	denominator
Sharpe ratio	$\frac{E(R_p)-E(R_b)}{\sqrt{Var(R_p-R_b)}}$	mean of pf. excess return	standard deviation of portfolio excess return
VaR_α ratio	$\frac{E(R_p)-E(R_b)}{VaR_\alpha(R_p-R_b)}$	mean of pf. excess return	VaR of pf. excess return, significance level α
STAR_α ratio	$\frac{E(R_p)-E(R_b)}{AVaR_\alpha(R_p-R_b)}$	mean of pf. excess return	AVaR of pf. excess return, significance level α

Table 5.10: Comparison of employed performance ratios in the portfolio optimization.

to excess returns $(R_p - R_b)$. They become the basis for both the numerator and the individual denominator in the respective instances of the performance ratio.

2.) The DJIA index serves as benchmark return R_b. In order to be able to determine the distribution of the difference $R_p - R_b$, one has to include R_b as an additional dimension into the distribution of joint returns. This is required both for the case of the Gaussian and the MVNTS assumption[26].

3.) For the estimation of the correlation structure, it appears essential to base it on a 1-year period of most recent return data, while the other parameters are still based on the 10-year period.

In order to keep a sufficiently large time interval for purposes of later out-of-sample back-testing available, the 10-years sample period used for model estimation is relocated to the beginning of the data set, i.e. now ranging from October 1, 1997 to September 30, 2007. This excludes most events which happened during the financial crisis from the sample period. The results of the portfolio optimization[27], carried out in a heuristic fashion, are given in table 5.11.

Furthermore, the corresponding optimal weights ω^* can be found in table 5.12. These portfolio weights will be referred to as static strategies under respective optimization criteria, i.e. theoretical performance ratios, and model assumptions for R in the further context.

[26]This subsequently possible incorporation is facilitated by the consistent choice of $(\hat{a}, \hat{\lambda})$ for stocks and the index in section 5.3.4, which then solely requires a re-estimation of joint correlations between each stock and the index.

[27]The corresponding optimization problem can be found in appendix C.3.

		ratio	reward	risk measure
	Sharpe ratio maximizing PF	0.1293	0.0287 %	0.2222 %
Normal Model	VaR ratio maximizing PF	0.0589	0.0287 %	0.4882 %
	STAR ratio maximizing PF	0.0510	0.0287 %	0.5635 %
NTS Model	VaR ratio maximizing PF	0.0919	0.0254 %	0.2766 %
	STAR ratio maximizing PF	0.0783	0.0254 %	0.3246 %

Table 5.11: Output of portfolio optimization under various ratio criteria.

		Normal Model		NTS Model	
k	Sharpe ratio maximizing portfolio	VaR ratio maximizing portfolio	STAR ratio maximizing portfolio	VaR ratio maximizing portfolio	STAR ratio maximizing portfolio
	ω^*_{Sharpe}	$\omega^*_{VaR,Normal}$	$\omega^*_{STAR,Normal}$	$\omega^*_{VaR,NTS}$	$\omega^*_{STAR,NTS}$
1	0.0799	0.0799	0.0799	0.0814	0.0811
2	0.0373	0.0373	0.0373	0.0417	0.0417
3	0.0579	0.0579	0.0579	0.0526	0.0534
4	0.0310	0.0310	0.0310	0.0260	0.0254
5	0.0303	0.0303	0.0303	0.0179	0.0181
6	0.0362	0.0362	0.0362	0.0337	0.0340
7	0.0599	0.0599	0.0599	0.0514	0.0513
8	0.0201	0.0201	0.0201	0.0260	0.0257
9	0.0191	0.0191	0.0191	0.0048	0.0046
10	0.0053	0.0053	0.0053	0.0220	0.0220
11	0.0074	0.0074	0.0074	0.0289	0.0291
12	0.0685	0.0685	0.0685	0.0383	0.0382
13	0.0277	0.0277	0.0277	0.0330	0.0335
14	0.0356	0.0356	0.0356	0.0361	0.0368
15	0.0265	0.0265	0.0265	0.0278	0.0272
16	0.0138	0.0138	0.0138	0.0189	0.0183
17	0.0698	0.0698	0.0698	0.0717	0.0714
18	0.0660	0.0660	0.0660	0.0698	0.0701
19	0.0264	0.0264	0.0264	0.0366	0.0365
20	0.0295	0.0295	0.0295	0.0292	0.0293
21	0.0350	0.0350	0.0350	0.0273	0.0271
22	0.0324	0.0324	0.0324	0.0375	0.0376
23	-0.0103	-0.0103	-0.0103	-0.0022	-0.0017
24	0.0632	0.0632	0.0632	0.0526	0.0517
25	0.0056	0.0056	0.0056	0.0152	0.0160
26	0.0704	0.0704	0.0704	0.0541	0.0534
27	-0.0004	-0.0004	-0.0004	0.0061	0.0058
28	0.0388	0.0388	0.0388	0.0397	0.0406
29	0.0173	0.0173	0.0173	0.0219	0.0219
Sum	1	1	1	1	1

Table 5.12: Optimal portfolio weights ω^* for different optimization criteria.

5.5.3 In-Sample and Out-of-Sample Tests

Now, beyond the perspective of formal GoF tests and supporting graphical assessments performed in sections 5.4 and 5.5.1, two of the above optimal static strategies are compared against each other with regard to their performance on empirical return data. For a justified selection of two significant candidates from the five available strategies, one has to consider the following. The Sharpe ratio of a particular allocation is independent of the actual choice of either the Normal or the MVNTS assumption in the theoretical context. This is because the mean vector as well as the covariance matrix obtained from the respective estimation procedure are identical under both assumptions and coincide with their empirical counterparts. Moreover, as the multivariate Normal distribution is entirely specified by the mean vector and the covariance matrix, the VaR and STAR ratio, although not identical, deliver the same optimal strategy ω^* under this assumption, however. This is confirmed numerically in table 5.12. Hence, ω^*_{Sharpe} serves as the representative of conventional portfolio construction. From the strategies under MVNTS assumption, $\omega^*_{STAR,NTS}$ is chosen, as its associated risk measure considers the shape of distribution beyond VaR, whereas the two strategies do not differ substantially anyway.

The in-sample test compares a number of different indicators reflecting rewards and risks of hypothetical excess return realizations obtained from strategies ω^*_{Sharpe} and $\omega^*_{STAR,NTS}$, denoted with r_p^{Sharpe} and $r_p^{STAR,NTS}$, respectively. The sample period under consideration is October 1, 1997 to September 30, 2007, which consists of 2514 return samples. The same is done in the out-of-sample test, with the only difference that the sample period changes to October 1, 2007 through September 30, 2009 with 515 return samples contained. Besides the empirical performance ratios, various other figures of both risk and reward are represented in tables 5.13 and 5.14. The empirical benchmark return is still taken from the DJIA index on respective corresponding time intervals.

On the reward side, sample mean and maximum are included. Cumulated returns[28] apply for a static investment according to two different allocation weights. The VaR_α and AVaR_α for significance level $\alpha = 0.01$ of the returns instead of the losses represent the upper

[28] Returns again refer to excess returns $r_p - r_b$ in this section.

		in-sample test	
		$r_p^{Sharpe} - r_b$	$r_p^{STAR,NTS} - r_b$
Rewards	sample mean	0.0287 %	0.0254 %
	maximum return	1.1290 %	1.1389 %
	cumulative return	104.63 %	88.35 %
	empirical VaR returns (1 %)	0.6673 %	0.6301 %
	empirical AVaR returns (1 %)	0.7899 %	0.7664 %
Risks (Uncertainty)	sample standard deviation	0.2222 %	0.2081 %
	maximum return − minimum return	2.4089 %	2.1617 %
	empirical VaR returns (1 %) + empirical VaR (1 %)	1.2094 %	1.0999 %
Risks (Loss)	Maximum Loss	1.1899 %	1.228 %
	empirical VaR (1 %)	0.6047 %	0.5500 %
	empirical AVaR (1 %)	0.7448 %	0.6797 %
Performance Measure	empirical Sharpe ratio	0.1293	0.1221
	empirical VaR ratio	0.0475	0.0462
	empirical STAR ratio	0.0386	0.0374
	empirical Rachev ratio	1.0609	1.1277

Table 5.13: In-sample test results of empirical excess returns.

		out-of-sample test	
		$r_p^{Sharpe} - r_b$	$r_p^{STAR,NTS} - r_b$
Rewards	sample mean	0.0468 %	0.0491 %
	maximum return	2.4010 %	2.3958 %
	cumulative return	26.25 %	27.76 %
	empirical VaR returns (1 %)	1.1122 %	1.0573 %
	empirical AVaR returns (1 %)	1.8032 %	1.6559 %
Risks (Uncertainty)	sample standard deviation	0.3622 %	0.3438 %
	maximum return − minimum return	3.7523 %	3.6663 %
	empirical VaR returns (1 %) + empirical VaR (1 %)	2.0166 %	1.8943 %
Risks (Loss)	Maximum Loss	1.3512 %	1.2705 %
	empirical VaR (1 %)	0.9044 %	0.8370 %
	empirical AVaR (1 %)	1.1720 %	1.1227 %
Performance Measure	empirical Sharpe ratio	0.1292	0.1428
	empirical VaR ratio	0.0518	0.0589
	empirical STAR ratio	0.0399	0.0437
	empirical Rachev ratio	1.5386	1.4749

Table 5.14: Out-of-sample test results of empirical excess returns.

$(1-\alpha)$-quantile and the conditional expectation of returns exceeding the $(1-\alpha)$-quantile, respectively[29]. While for the in-sample test, the Sharpe-optimal strategy ω^*_{Sharpe} shows advantages considering all five reward figures, the NTS STAR-optimal strategy $\omega^*_{STAR,NTS}$ shows at least a higher mean and cumulative return in the out-of-sample test. For clarity of presentation, the first strategy will further be denoted by A, the latter by B, respectively.

Subsequently, the reward potential is opposed by the risk and uncertainties one has to bear in exchange. On the side of uncertainties, meaning the possibilities for random return fluctuation in both negative and positive direction, we find that the sample standard deviation, the range of realized excess returns as well as the distance between α-quantile and $(1-\alpha)$-quantile is consistently lower for strategy B than for A. This result is true for both in-sample- and out-of-sample perspective. When turning to actual risks, i.e. the potential for negative returns only, the statement remains the same. There, for maximum loss[30] as well as for the conventional VaR and AVaR, strategy B generates lower risk figures than A throughout all risk figures for both in- and out-of-sample tests. So far, although strategy A offers advantages on the reward side, one clearly has to accept a higher level of risk in return.

Higher risk is not unfavorable a priori, however. Performance ratios aggregate these rewards and risks into one single measure. There, the results are well distinguished between the in-sample and out-of-sample case. While the covered in-sample period leaves out the peak of the global financial crisis, these events are included in the out-of-sample period. For Sharpe, VaR and STAR ratio, strategy A dominates in the in-sample test while strategy B does so in the out-of-sample test. Among the performance ratios, only Rachev ratio[31] indicates an opposite result in both cases.

Therefore, it seems that an optimal allocation decision based on the first two distributional moments maximized with respect to the Sharpe ratio (strategy A) is able to deliver superior results when applied to periods of moderate and constant market volatilities over time.

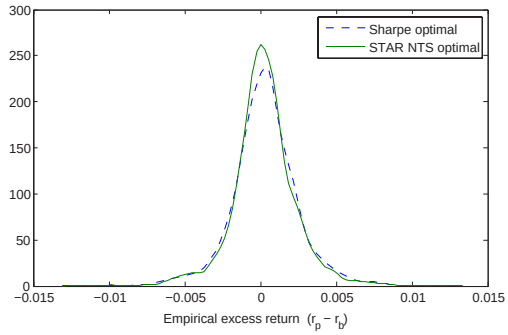
[29] For this reason, these two measures can be determined by calculating the conventional VaR and AVaR for returns with opposite sign.

[30] or minimal return

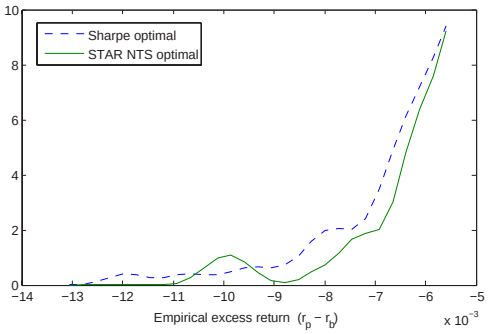
[31] For a definition and first applications see Biglova et al. [BORS04].

Strategy B instead, based on the MVNTS distribution as an assumption and optimized with respect to a performance measure focused on extreme negative tail events, shows significant advantages in generalizing the information of the sample period to an out-of-sample period.

Finally, diagrams 5.7 and 5.8 try to reinforce the above conclusions about the derived risk and reward quantities by visual features of the empirical portfolio excess return. In contrast to the different scalar performance measures which, by design, are only able to represent one isolated feature of the excess return distribution, whereas these aspects are displayed simultaneously in the kernel density plots. For the in-sample backtest one observes a higher peak of the excess return distribution generated by strategy B (STAR NTS-optimal allocation weights) compared to strategy A (Sharpe-optimal allocation weights). In addition, the distribution of strategy A is slightly skewed to the left while the distribution of strategy B is almost symmetric. In figure 5.7(b), which only shows the distribution of the left tail from the 1%-quantile downward, one notices further that strategy A has more pronounced left tails and therefore carrying a higher associated risk of facing severe losses. In the overall picture 5.8(a) of the out-of-sample backtest, the distributions of A and B closely resemble each other in symmetry and peakedness, the slight skewness of distribution A from the in-sample backtest has vanished. But also in the out-of-sample test, comprising market returns with substantially increased volatilities, excess returns of A show fatter tails and higher risks, as was already true for the in-sample test. While AVaR only measures the first moment of the negative's tail distribution, the whole shape of the negative tail becomes apparent in diagram 5.8(b), which is indispensable for a comprehensive risk assessment.

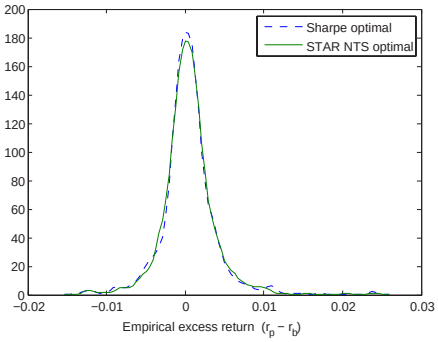


(a) Kernel densities of empirical excess returns for ω^*_{Sharpe} and $\omega^*_{STAR,MVNTS}$

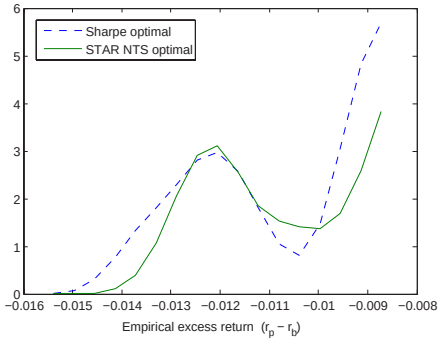


(b) Magnification of kernel densities below 1%-quantile

Figure 5.7: Kernel densities of the in-sample test.



(a) Kernel densities of empirical excess returns for ω^*_{Sharpe} and $\omega^*_{STAR,MVNTS}$



(b) Magnification of kernel densities below 1%-quantile

Figure 5.8: Kernel densities of the out-of-sample test.

Chapter 6

ARMA-GARCH Models

6.1 Basic Theory and Definitions

While the former model in chapter 5 is based on the rather simple and unrealistic assumption of i.i.d. returns, this assumption is replaced by an advanced time series model for conditional mean and variance. In this context, the classical ARMA model constitutes a linear time series model for the conditional mean with constant conditional volatility. ARMA models can be furthermore combined with ARCH and GARCH components, which explicitly take volatility clustering into account. The innovative idea behind ARMA-GARCH models is to model the linear autocorrelation of the squared process variables. By this approach, the model requires only one single source of risk for the standardized innovations, which simultaneously drives the return process and the current level of volatility. This offers the advantage of parsimonious parametrization combined with the availability of efficient estimation methods.

For the scope to be considered here, an ARMA(1,1)-GARCH(1,1) model appears to be sufficiently flexible while maintaining a considerable level of analytical tractability with regard to parameter estimation. To be more precise, the ARMA(1,1)-GARCH(1,1) model

for the joint daily asset returns $R_t = \left(R_t^{(1)}, \ldots, R_t^{(n)}\right)$ at time t is formulated as follows

$$R_{t+1}^{(k)} = C_k + b_k R_t^{(k)} + \eta_k \sigma_{k,t} \xi_{k,t} + \sigma_{k,t+1} \xi_{k,t+1} \tag{6.1}$$

$$\sigma_{t+1,k}^2 = \alpha_k + \theta_k \sigma_{k,t}^2 \xi_{k,t}^2 + \beta_k \sigma_{k,t}^2 \,, \quad k = 1, \ldots, n \tag{6.2}$$

Note that the return of each single asset is driven by a merely univariate ARMA(1,1)-GARCH(1,1) model. Therefore, the conditional mean and variance of the day-ahead return $R_{t+1}^{(k)}$ of one particular asset component k solely depends on its own history of realizations up to time t. The only source of dependence between $R_t^{(k)}$ for different k is the dependence structure of the vector of standardized innovations

$$\xi_t = (\xi_{1,t}, \ldots, \xi_{n,t})^{\mathsf{T}} \overset{i.i.d.}{\sim} stdMVNTS(a, \lambda, \gamma, \mathsf{P}) \,.$$

For a comprehensive survey of GARCH models with non-Gaussian innovations in the multivariate case and in the context of financial modeling in particular, see Aas et al. [AaHHD06] and the references therein.

6.2 Empirical Estimation and GoF Tests

6.2.1 Index Returns

In a preparing step, the ARMA-GARCH model is applied to the one-dimensional DJIA index returns, i.e. $n = 1$ in equations (6.1) and (6.2). The subsequent estimation of model coefficients and distribution parameters is based on the sample period from October 20, 1998 to September 26, 2008. This period encompasses a 10-year return history with $N = 2500$ daily records in total. The reason for choosing September 26, 2008 as the end date of the sample period, is the market crash that occurred on the next trading day, September 29, 2008. On that particular day the DJIA index lost 6.98% in value, triggered by the collapse of U.S. investment bank Lehman Brothers, which has happened two weeks before.

Tables 6.1 and 6.2 show the ARMA and GARCH coefficients, respectively, obtained by efficient ML estimation based on the assumption of a t-distributed random variable for the innovations ξ_t. This assumption makes standard estimation procedures available while heavy tails are sufficiently taken into account.

$\hat{C}_{ind}$	$\hat{b}_{ind}$	$\hat{\eta}_{ind}$
$1.1936 \cdot 10^{-4}$	0.7425	-0.7826

Table 6.1: Estimates of ARMA coefficients for index returns under t-assumption.

$\hat{\alpha}_{ind}$	$\hat{\theta}_{ind}$	$\hat{\beta}_{ind}$
$5.1670 \cdot 10^{-7}$	0.0634	0.9350

Table 6.2: Estimates of GARCH coefficients for index returns under t-assumption.

With the estimated ARMA-GARCH coefficients and the corresponding sequence of conditional volatilities $\{\hat{\sigma}_{ind,t}\}_{t=1,\dots,N}$, one is moreover able to generate the time series of estimated standardized residuals $\{\hat{\xi}_{ind,t}\}_{t=1,\dots,N}$. By the assumptions implied in equations (6.1) and (6.2), these standardized residuals should be free of autocorrelation. Therefore, the univariate stdMVNTS distribution parameters of ξ_{ind} can be estimated by using the methods described in section 5.3.2. The result of this estimation is contained in table 6.3.

$\hat{a}_{ind}$	$\hat{\lambda}_{ind}$	$\hat{\gamma}_{ind}$
0.0100	3.2007	-0.1886

Table 6.3: Parameter estimates (stdNTS) for standardized index residuals.

Moreover, the results of the GoF test of the estimated stdNTS distribution with respect to the standardized index residuals can be found in table 6.4. In this table, the realized values of KS, AD and AD^2 together with their corresponding p-values are presented. They can be compared to the opposing GoF results obtained for a Gaussian ARMA-GARCH with normally distributed innovations in table 6.5. Note that the Gaussian ARMA-GARCH model has different coefficient estimates from the ones presented in tables 6.1 and 6.2. This is due to the transition to the alternative assumption of a normal distribution for the innovations. Although none of the three evaluated statistics can seriously be rated a clear

indication for the adequacy of the stdNTS assumption for the process innovations, only the *AD* test is able to formally reject this hypothesis at a 1% significance level, however. In contrast, all three GoF tests clearly reject the Gaussian hypothesis at this significance level. With regard to the index residuals of an ARMA-GARCH model, the stdNTS distribution seems nevertheless superior to the normal distribution. The remarkably high value $\hat{\lambda}_{ind}$ of approximately 3.2 of the tempering coefficient λ in combination with a rather small value of $\hat{a}_{ind}$ close to zero leads to the conclusion that the tails of the considered distribution are located between those of an NTS and the Gaussian alternative.

statistics	value	p-value
KS	0.0285	0.0335
AD	0.3146	0.0080
AD^2	2.5957	0.0442

Table 6.4: GoF results for standardized index residuals (stdNTS).

statistics	value	p-value
KS	0.0379	0.0015
AD	62.9629	0.0000
AD^2	4.3241	0.0061

Table 6.5: GoF results for standardized index residuals (Normal).

6.2.2 Stock Returns

The promising results obtained for the index returns in favor of normal tempered stable distributions should now be further substantiated. The model for the joint stock returns is composed of the univariate ARMA-GARCH time series models for each single stock according to equations (6.1) and (6.2). Similar to the treatment of index returns, two different model versions are considered. The model termed *MVNTS ARMA-GARCH* employs the assumption of univariate t-distributions for heavy-tailed process innovations to efficiently determine estimates of the process coefficients. These estimates are subsequently used for generating the time series of standardized residuals $\{\hat{\xi}_{k,t}\}, k = 1, \ldots, n\,,\, t = 1, \ldots, N$. The standardized residuals of the components k are further aggregated into the n-dimensional time series $\{\hat{\xi}_t\}, t = 1, \ldots, N$. Due to the previous filtering of the empirical return ob-

servations, autocorrelation of first and second order $\{\hat{\xi}_t\}$ should have been eliminated by this procedure. This enables the assumption of $\{\hat{\xi}_t\}$ being independently sampled from identical $stdMVNTS(a, \lambda, \gamma, \mathsf{P})$ distributions. However, the second model termed *Normal ARMA-GARCH* basically involves similar procedural steps. In contrast to the MVNTS ARMA-GARCH, it entirely relies on the Gaussian assumption for the distribution of innovations. This is both true for the estimation of the ARMA-GARCH coefficients as well as for the innovation distribution itself.

The estimates of the MVNTS ARMA-GARCH coefficients are given in table 6.6. The estimation is performed for the 29 selected DJIA stocks contained in table 5.1. It is furthermore based on the same observation period as employed in section 6.2.1. Although the MVNTS ARMA-GARCH model will be compared to the Normal ARMA-GARCH model in the course of GoF tests, explicit estimation results for the latter are omitted here. This is justified by the focus on interpretable results only. Finally, one has to note that the estimated coefficients differ substantially when comparing different stocks. This, in turn, indicates that each stock possesses its individual characteristics with regard to return dynamics.

The parameter estimation of the corresponding stdMVNTS distribution is carried out by performing the procedure developed in section 5.3.4. As was explained, the parameters (a, λ) related to the subordinator dynamics are simultaneously estimated, based on linear combinations of the joint residuals $\hat{\xi}_t = \left(\hat{\xi}_{1,t}, \ldots, \hat{\xi}_{29,t}\right)^{\mathsf{T}}$ with a weighting ω. The most reliable and meaningful outcome is generated when using an equally weighted portfolio, i.e.

$$\omega = \omega_{ewpf} = \left(\frac{1}{29}, \ldots, \frac{1}{29}\right)^{\mathsf{T}} . \tag{6.3}$$

This choice leads to the estimates

$$\hat{a} = 0.01$$

$$\hat{\lambda} = 2.6925 .$$

Having these common fixed values for every dimension $k = 1, \ldots, 29$, one is able to se-

k	$\hat{C}_k$	$\hat{b}_k$	$\hat{\eta}_k$	$\hat{\alpha}_k$	$\hat{\theta}_k$	$\hat{\beta}_k$
1	$5.9275 \cdot 10^{-5}$	0.8072	-0.8332	$2.0000 \cdot 10^{-7}$	0.0224	0.9771
2	$4.8087 \cdot 10^{-4}$	-0.6776	0.7125	$1.8860 \cdot 10^{-6}$	0.0368	0.9611
3	$1.0424 \cdot 10^{-4}$	0.7871	-0.8431	$7.4933 \cdot 10^{-7}$	0.0558	0.9442
4	$7.8242 \cdot 10^{-4}$	-0.9325	0.9259	$5.4533 \cdot 10^{-7}$	0.0387	0.9608
5	$4.6341 \cdot 10^{-5}$	0.9170	-0.9372	$1.3748 \cdot 10^{-6}$	0.0882	0.9118
6	$8.4114 \cdot 10^{-4}$	-0.1499	0.1001	$3.4631 \cdot 10^{-6}$	0.0476	0.9440
7	$1.2915 \cdot 10^{-3}$	-0.6808	0.7062	$9.9813 \cdot 10^{-7}$	0.0159	0.9816
8	$3.6463 \cdot 10^{-4}$	0.5662	-0.6236	$4.6748 \cdot 10^{-6}$	0.0671	0.9139
9	$2.2367 \cdot 10^{-4}$	0.6591	-0.6988	$4.3463 \cdot 10^{-7}$	0.0244	0.9751
10	$2.5916 \cdot 10^{-4}$	-0.6091	0.6319	$2.3288 \cdot 10^{-7}$	0.0322	0.9674
11	$2.5215 \cdot 10^{-4}$	-0.5519	0.5615	$9.3181 \cdot 10^{-7}$	0.0356	0.9618
12	$1.5038 \cdot 10^{-3}$	-0.9560	0.9480	$3.1502 \cdot 10^{-6}$	0.0573	0.9294
13	$3.6220 \cdot 10^{-5}$	0.8588	-0.8859	$2.8945 \cdot 10^{-7}$	0.0317	0.9683
14	$1.8906 \cdot 10^{-4}$	0.7415	-0.7853	$6.6811 \cdot 10^{-7}$	0.0306	0.9693
15	$1.0384 \cdot 10^{-4}$	-0.9580	0.9693	$1.5445 \cdot 10^{-6}$	0.0418	0.9557
16	$1.4604 \cdot 10^{-4}$	0.6982	-0.7167	$9.7592 \cdot 10^{-7}$	0.0278	0.9708
17	$4.3450 \cdot 10^{-4}$	-0.3269	0.2926	$5.6013 \cdot 10^{-7}$	0.0308	0.9678
18	$3.2311 \cdot 10^{-4}$	-0.4373	0.4699	$5.8986 \cdot 10^{-7}$	0.0547	0.9443
19	$7.7352 \cdot 10^{-5}$	0.8244	-0.8458	$1.0110 \cdot 10^{-6}$	0.0592	0.9408
20	$9.8250 \cdot 10^{-5}$	0.8353	-0.8554	$1.6811 \cdot 10^{-6}$	0.0323	0.9620
21	$2.7484 \cdot 10^{-4}$	0.0154	0.0069	$5.3944 \cdot 10^{-6}$	0.0474	0.9371
22	$6.4453 \cdot 10^{-5}$	-0.1247	0.0771	$5.7350 \cdot 10^{-7}$	0.0431	0.9569
23	$-1.6279 \cdot 10^{-4}$	-0.5710	0.6154	$2.8998 \cdot 10^{-6}$	0.0632	0.9298
24	$2.1816 \cdot 10^{-4}$	0.5807	-0.6493	$9.1479 \cdot 10^{-7}$	0.0436	0.9516
25	$1.9488 \cdot 10^{-4}$	0.1880	-0.2210	$4.6307 \cdot 10^{-6}$	0.0863	0.9075
26	$1.5361 \cdot 10^{-4}$	0.7831	-0.8328	$1.2956 \cdot 10^{-6}$	0.0486	0.9480
27	$2.3897 \cdot 10^{-4}$	-0.8822	0.8526	$9.8393 \cdot 10^{-7}$	0.0542	0.9447
28	$4.8395 \cdot 10^{-6}$	0.9350	-0.9592	$2.5855 \cdot 10^{-7}$	0.0277	0.9719
29	$4.3704 \cdot 10^{-5}$	0.8718	-0.8996	$8.2264 \cdot 10^{-7}$	0.0245	0.9728

Table 6.6: Estimated coefficients for the MVNTS ARMA-GARCH stock model.

quentially estimate the single components γ_k of γ. The current estimates are given in table 6.7, together with their corresponding 95% confidence intervals.

k	$\hat{\gamma}_k$	lower c.i. bound	upper c.i. bound
1	0.0799	-0.0364	0.1961
2	0.1569	0.0204	0.2935
3	0.1493	0.0158	0.2828
4	0.1121	-0.0240	0.2483
5	-0.0306	-0.0839	0.0226
6	0.0541	-0.0831	0.1912
7	0.0620	-0.0670	0.1910
8	-0.1331	-0.2883	0.0220
9	0.0946	-0.0460	0.2352
10	0.1553	0.0427	0.2679
11	0.2217	0.0906	0.3528
12	-0.1668	-0.3122	-0.0213
13	0.1547	0.0194	0.2901
14	0.0832	-0.0367	0.2031
15	0.1258	0.0983	0.1533

k	$\hat{\gamma}_k$	lower c.i. bound	upper c.i. bound
16	-0.0377	-0.1833	0.1079
17	-0.0207	-0.1438	0.1024
18	0.1434	0.0108	0.2761
19	0.1440	0.0154	0.2725
20	0.1038	-0.0331	0.2407
21	-0.0001	-0.0002	0.0001
22	0.0477	0.0476	0.0477
23	-0.0314	-0.1648	0.1021
24	0.0208	-0.0649	0.1065
25	0.1020	-0.0074	0.2114
26	0.0872	-0.0219	0.1963
27	0.0945	-0.0494	0.2385
28	0.1977	0.0520	0.3433
29	0.1156	0.0119	0.2193

Table 6.7: Estimate $\hat{\gamma}$ of the MVNTS ARMA-GARCH stock model.

In a final step, together with the previous estimates $(\hat{a}, \hat{\lambda}, \hat{\gamma})$ the calibration of $\hat{\mathsf{P}}$ can be carried out according to equation (5.15). Similar to the stationary model in chapter 5, the calibration of $\hat{\mathsf{P}}$ completes the specification of the entire MVNTS ARMA-GARCH model for the joint stock return dynamics R_t. Table 6.8 contains the calibrated correlation matrix $\hat{\mathsf{P}}$ of the underlying multivariate Brownian motion.

Statistical GoF tests constitute a difficult task when more than two dimensions are involved[1]. The most problematic aspect of multidimensional tests such as the χ^2 test, is the non-robustness of their results. This moreover impedes their capability of clearly distinguishing distributions with a sufficiently adequate model fit from those possessing a poor model fit, in particular with regard to the distribution tails. Therefore, one is restricted to one-dimensional GoF tests of the marginal distributions, while the adequacy of the dependence structure cannot be formally assessed in this context. Tables 6.9 and 6.10 contain the three familiar univariate GoF statistics for the 29 marginal distributions of

[1]For detailed explanations see sections 5.4.2.

$$\hat{\mathrm{P}} = \begin{pmatrix}
1 & 0.373 & 0.384 & 0.241 & 0.374 & 0.324 & 0.431 & 0.305 & 0.281 & 0.302 & 0.493 & 0.353 & 0.43 & 0.254 & 0.327 & 0.296 & 0.291 & 0.267 & 0.36 & 0.238 & 0.181 & 0.254 & 0.29 & 0.295 & 0.337 & 0.45 & 0.254 & 0.328 & 0.305 \\
0.373 & 1 & 0.352 & 0.253 & 0.325 & 0.32 & 0.445 & 0.385 & 0.262 & 0.172 & 0.48 & 0.398 & 0.37 & 0.227 & 0.294 & 0.292 & 0.262 & 0.134 & 0.362 & 0.246 & 0.16 & 0.269 & 0.253 & 0.186 & 0.28 & 0.402 & 0.258 & 0.237 & 0.301 \\
0.384 & 0.352 & 1 & 0.379 & 0.603 & 0.34 & 0.407 & 0.293 & 0.402 & 0.319 & 0.431 & 0.338 & 0.566 & 0.335 & 0.46 & 0.398 & 0.399 & 0.299 & 0.627 & 0.312 & 0.253 & 0.387 & 0.338 & 0.3 & 0.445 & 0.441 & 0.384 & 0.434 & 0.438 \\
0.241 & 0.253 & 0.379 & 1 & 0.342 & 0.229 & 0.24 & 0.222 & 0.282 & 0.274 & 0.285 & 0.288 & 0.351 & 0.218 & 0.293 & 0.259 & 0.292 & 0.27 & 0.365 & 0.22 & 0.236 & 0.285 & 0.289 & 0.293 & 0.276 & 0.272 & 0.645 & 0.311 & 0.276 \\
0.374 & 0.325 & 0.603 & 0.342 & 1 & 0.302 & 0.384 & 0.268 & 0.34 & 0.296 & 0.411 & 0.306 & 0.506 & 0.286 & 0.436 & 0.338 & 0.343 & 0.296 & 0.683 & 0.273 & 0.236 & 0.324 & 0.356 & 0.29 & 0.438 & 0.413 & 0.353 & 0.402 & 0.351 \\
0.324 & 0.32 & 0.34 & 0.229 & 0.302 & 1 & 0.325 & 0.236 & 0.241 & 0.22 & 0.352 & 0.273 & 0.356 & 0.233 & 0.294 & 0.269 & 0.242 & 0.201 & 0.328 & 0.229 & 0.185 & 0.249 & 0.241 & 0.209 & 0.295 & 0.479 & 0.243 & 0.275 & 0.288 \\
0.431 & 0.445 & 0.407 & 0.24 & 0.384 & 0.325 & 1 & 0.318 & 0.265 & 0.22 & 0.47 & 0.345 & 0.414 & 0.253 & 0.359 & 0.308 & 0.302 & 0.209 & 0.395 & 0.235 & 0.157 & 0.291 & 0.246 & 0.214 & 0.337 & 0.464 & 0.238 & 0.322 & 0.309 \\
0.305 & 0.385 & 0.293 & 0.222 & 0.268 & 0.236 & 0.318 & 1 & 0.166 & 0.206 & 0.332 & 0.775 & 0.272 & 0.179 & 0.181 & 0.182 & 0.215 & 0.262 & 0.183 & 0.191 & 0.211 & 0.254 & 0.18 & 0.242 & 0.31 & 0.227 & 0.193 & 0.228 \\
0.281 & 0.262 & 0.402 & 0.282 & 0.34 & 0.241 & 0.265 & 0.166 & 1 & 0.203 & 0.29 & 0.198 & 0.443 & 0.458 & 0.341 & 0.566 & 0.479 & 0.157 & 0.412 & 0.242 & 0.159 & 0.487 & 0.277 & 0.187 & 0.281 & 0.319 & 0.301 & 0.317 & 0.349 \\
0.302 & 0.172 & 0.319 & 0.274 & 0.296 & 0.22 & 0.22 & 0.206 & 0.203 & 1 & 0.271 & 0.293 & 0.336 & 0.169 & 0.275 & 0.224 & 0.223 & 0.322 & 0.279 & 0.236 & 0.245 & 0.257 & 0.299 & 0.386 & 0.252 & 0.27 & 0.28 & 0.287 & 0.224 \\
0.493 & 0.48 & 0.431 & 0.285 & 0.411 & 0.352 & 0.47 & 0.332 & 0.29 & 0.271 & 1 & 0.376 & 0.435 & 0.256 & 0.355 & 0.308 & 0.308 & 0.246 & 0.414 & 0.276 & 0.214 & 0.287 & 0.312 & 0.279 & 0.361 & 0.457 & 0.287 & 0.352 & 0.336 \\
0.353 & 0.398 & 0.338 & 0.288 & 0.306 & 0.273 & 0.345 & 0.775 & 0.198 & 0.293 & 0.376 & 1 & 0.335 & 0.199 & 0.244 & 0.22 & 0.25 & 0.267 & 0.304 & 0.204 & 0.223 & 0.253 & 0.299 & 0.241 & 0.281 & 0.347 & 0.285 & 0.256 & 0.269 \\
0.43 & 0.37 & 0.566 & 0.351 & 0.506 & 0.356 & 0.414 & 0.272 & 0.443 & 0.336 & 0.435 & 0.335 & 1 & 0.342 & 0.462 & 0.441 & 0.434 & 0.363 & 0.53 & 0.306 & 0.273 & 0.427 & 0.408 & 0.36 & 0.417 & 0.466 & 0.365 & 0.439 & 0.428 \\
0.254 & 0.227 & 0.335 & 0.218 & 0.286 & 0.233 & 0.253 & 0.179 & 0.458 & 0.169 & 0.256 & 0.199 & 0.342 & 1 & 0.276 & 0.441 & 0.432 & 0.107 & 0.35 & 0.215 & 0.151 & 0.372 & 0.203 & 0.145 & 0.224 & 0.293 & 0.252 & 0.256 & 0.32 \\
0.327 & 0.294 & 0.46 & 0.293 & 0.436 & 0.294 & 0.359 & 0.181 & 0.341 & 0.275 & 0.355 & 0.244 & 0.462 & 0.276 & 1 & 0.348 & 0.351 & 0.255 & 0.466 & 0.316 & 0.211 & 0.343 & 0.304 & 0.263 & 0.333 & 0.367 & 0.284 & 0.563 & 0.348 \\
0.296 & 0.292 & 0.398 & 0.259 & 0.338 & 0.269 & 0.308 & 0.182 & 0.566 & 0.224 & 0.308 & 0.22 & 0.41 & 0.441 & 0.348 & 1 & 0.465 & 0.181 & 0.41 & 0.223 & 0.182 & 0.501 & 0.254 & 0.199 & 0.287 & 0.323 & 0.284 & 0.325 & 0.334 \\
0.291 & 0.262 & 0.399 & 0.292 & 0.343 & 0.242 & 0.302 & 0.198 & 0.479 & 0.223 & 0.308 & 0.25 & 0.434 & 0.432 & 0.351 & 0.465 & 1 & 0.219 & 0.382 & 0.246 & 0.164 & 0.433 & 0.245 & 0.195 & 0.29 & 0.32 & 0.297 & 0.331 & 0.355 \\
0.267 & 0.134 & 0.299 & 0.27 & 0.296 & 0.201 & 0.209 & 0.215 & 0.157 & 0.322 & 0.246 & 0.267 & 0.363 & 0.107 & 0.255 & 0.181 & 0.219 & 1 & 0.253 & 0.22 & 0.388 & 0.24 & 0.462 & 0.358 & 0.241 & 0.221 & 0.252 & 0.296 & 0.204 \\
0.36 & 0.362 & 0.627 & 0.365 & 0.683 & 0.328 & 0.395 & 0.262 & 0.412 & 0.279 & 0.414 & 0.304 & 0.53 & 0.35 & 0.466 & 0.41 & 0.382 & 0.253 & 1 & 0.287 & 0.248 & 0.38 & 0.333 & 0.264 & 0.455 & 0.439 & 0.38 & 0.424 & 0.407 \\
0.238 & 0.246 & 0.312 & 0.22 & 0.273 & 0.229 & 0.235 & 0.183 & 0.242 & 0.236 & 0.276 & 0.204 & 0.306 & 0.215 & 0.316 & 0.223 & 0.246 & 0.22 & 0.287 & 1 & 0.181 & 0.219 & 0.244 & 0.271 & 0.254 & 0.275 & 0.212 & 0.285 & 0.246 \\
0.181 & 0.16 & 0.253 & 0.236 & 0.236 & 0.185 & 0.157 & 0.191 & 0.159 & 0.245 & 0.214 & 0.223 & 0.273 & 0.151 & 0.211 & 0.182 & 0.164 & 0.388 & 0.248 & 0.181 & 1 & 0.199 & 0.453 & 0.259 & 0.2 & 0.2 & 0.239 & 0.235 & 0.19 \\
0.254 & 0.269 & 0.387 & 0.285 & 0.324 & 0.249 & 0.291 & 0.211 & 0.487 & 0.257 & 0.287 & 0.253 & 0.427 & 0.372 & 0.343 & 0.501 & 0.433 & 0.24 & 0.38 & 0.219 & 0.199 & 1 & 0.273 & 0.221 & 0.287 & 0.334 & 0.323 & 0.345 & 0.329 \\
0.29 & 0.253 & 0.338 & 0.289 & 0.356 & 0.241 & 0.246 & 0.254 & 0.277 & 0.299 & 0.312 & 0.299 & 0.408 & 0.203 & 0.304 & 0.254 & 0.245 & 0.462 & 0.333 & 0.244 & 0.453 & 0.273 & 1 & 0.313 & 0.264 & 0.289 & 0.295 & 0.355 & 0.277 \\
0.295 & 0.186 & 0.3 & 0.293 & 0.29 & 0.209 & 0.214 & 0.18 & 0.187 & 0.386 & 0.279 & 0.241 & 0.36 & 0.145 & 0.263 & 0.199 & 0.195 & 0.358 & 0.264 & 0.271 & 0.259 & 0.221 & 0.313 & 1 & 0.261 & 0.287 & 0.28 & 0.293 & 0.209 \\
0.337 & 0.28 & 0.445 & 0.276 & 0.438 & 0.295 & 0.337 & 0.242 & 0.281 & 0.252 & 0.361 & 0.281 & 0.417 & 0.224 & 0.333 & 0.287 & 0.29 & 0.241 & 0.455 & 0.254 & 0.2 & 0.287 & 0.264 & 0.261 & 1 & 0.372 & 0.293 & 0.345 & 0.296 \\
0.45 & 0.402 & 0.441 & 0.272 & 0.413 & 0.479 & 0.464 & 0.31 & 0.319 & 0.27 & 0.457 & 0.347 & 0.466 & 0.293 & 0.367 & 0.323 & 0.32 & 0.221 & 0.439 & 0.275 & 0.2 & 0.334 & 0.289 & 0.287 & 0.372 & 1 & 0.294 & 0.339 & 0.368 \\
0.254 & 0.258 & 0.384 & 0.645 & 0.353 & 0.243 & 0.238 & 0.227 & 0.301 & 0.28 & 0.287 & 0.285 & 0.365 & 0.252 & 0.284 & 0.284 & 0.297 & 0.252 & 0.38 & 0.212 & 0.239 & 0.323 & 0.295 & 0.28 & 0.293 & 0.294 & 1 & 0.313 & 0.308 \\
0.328 & 0.237 & 0.434 & 0.311 & 0.402 & 0.275 & 0.322 & 0.193 & 0.317 & 0.287 & 0.352 & 0.256 & 0.439 & 0.256 & 0.563 & 0.325 & 0.331 & 0.296 & 0.424 & 0.285 & 0.235 & 0.345 & 0.355 & 0.293 & 0.345 & 0.339 & 0.313 & 1 & 0.314 \\
0.305 & 0.301 & 0.438 & 0.276 & 0.351 & 0.288 & 0.309 & 0.228 & 0.349 & 0.224 & 0.336 & 0.269 & 0.428 & 0.32 & 0.348 & 0.334 & 0.355 & 0.204 & 0.407 & 0.246 & 0.19 & 0.329 & 0.277 & 0.209 & 0.296 & 0.368 & 0.308 & 0.314 & 1
\end{pmatrix} \tag{6.4}$$

Table 6.8: Calibrated $\hat{\mathrm{P}}$ of the MVNTS ARMA-GARCH stock model.

the MVNTS ARMA-GARCH and the Normal ARMA-GARCH model, respectively[2].

k	KS	p-value KS	AD	p-value AD	AD^2	p-value AD^2
1	0.0220	0.1761	3.6931	0.0001	2.2950	0.0636
2	0.0106	0.9372	0.0757	0.1689	0.4101	0.8386
3	0.0103	0.9517	0.5806	0.0023	0.2104	0.9873
4	0.0120	0.8629	0.0992	0.0873	0.2596	0.9650
5	0.0138	0.7223	332.2412	0.00000	0.6390	0.6117
6	0.0125	0.8243	0.6922	0.0017	0.6655	0.5882
7	0.0118	0.8716	2.7883	0.0001	0.4132	0.8355
8	0.0298	0.0234	0.0744	0.1776	2.2398	0.0681
9	0.0137	0.7333	2.0703	0.0002	0.8501	0.4462
10	0.0148	0.6384	6.8903	0.0000	0.7453	0.5220
11	0.0137	0.7312	0.4811	0.0034	0.6238	0.6256
12	0.0240	0.1098	0.0565	0.4158	1.3874	0.2059
13	0.0104	0.9466	0.4899	0.0033	0.4773	0.7700
14	0.0203	0.2501	8.4180	0.0000	2.3844	0.0571
15	0.0121	0.8501	49.2358	0.0000	0.5655	0.6811
16	0.0176	0.4156	1.4522	0.0004	0.9664	0.3754
17	0.0168	0.4792	5.3703	0.0000	1.2180	0.2609
18	0.0150	0.6189	0.5546	0.0026	0.6922	0.5652
19	0.0124	0.8354	0.0776	0.1578	0.3172	0.9246
20	0.0131	0.7800	0.1652	0.0297	0.7150	0.5462
21	0.0266	0.0565	8.3511	0.0000	4.8909	0.0033
22	0.0206	0.2345	1279.1789	0.0000	1.4880	0.1794
23	0.0154	0.5872	2.8938	0.0001	0.6472	0.6044
24	0.0160	0.5393	372.7506	0.0000	1.1690	0.2797
25	0.0173	0.4415	9.0923	0.0000	1.2094	0.2641
26	0.0094	0.9782	48854.8877	0.0000	0.5521	0.6943
27	0.0149	0.6319	0.0591	0.3619	0.5951	0.6526
28	0.0237	0.1173	0.0552	0.4453	1.2846	0.2374
29	0.0157	0.5661	8.7223	0.0000	0.8085	0.4749

Table 6.9: GoF tests for residuals of the MVNTS ARMA-GARCH stock model.

First of all it can be observed that the KS test does not reject the stdNTS hypothesis for the marginal standardized innovations at the 1% significance level for any of the incorporated stocks. Quite contrary, it rather delivers remarkably high p-values in most of the cases. This result is further substantiated by the output of the squared AD test, where the null hypothesis is only rejected for stock $k = 21$ at the given significance level. Most other p-values range between 0.5 up to 0.98, however. The stock $k = 21$ seems to be exceptional for the following reason. Under the alternative Normal ARMA-GARCH model

[2]The symbol ∞ indicates the occurrence of a numerical overflow during the calculation, due to extremely high values. The corresponding p-value is practically zero.

k	KS	p-value KS	AD	p-value AD	AD^2	p-value AD^2
1	0.0499	0.0000	$8.3390 \cdot 10^{3}$	0.0000	13.7287	0.0000
2	0.0355	0.0035	1.7528	0.0003	5.2295	0.0000
3	0.0431	0.0001	$2.3091 \cdot 10^{2}$	0.0000	5.8110	0.0000
4	0.0351	0.0041	3.9956	0.0001	4.4903	0.0051
5	0.0364	0.0025	$1.2148 \cdot 10^{11}$	0.0000	6.1758	0.0000
6	0.0365	0.0025	$1.9087 \cdot 10^{3}$	0.0000	6.2636	0.0000
7	0.0399	0.0006	$1.8699 \cdot 10^{5}$	0.0000	7.1843	0.0000
8	0.0217	0.1868	$2.4727 \cdot 10^{-1}$	0.0131	1.4414	0.1912
9	0.0386	0.0011	$3.7963 \cdot 10^{4}$	0.0000	6.3495	0.0000
10	0.0500	0.0000	$1.3099 \cdot 10^{6}$	0.0000	9.5867	0.0000
11	0.0407	0.0005	$4.6706 \cdot 10^{1}$	0.0000	7.2222	0.0000
12	0.0221	0.1729	$4.1480 \cdot 10^{-1}$	0.0046	2.8307	0.0334
13	0.0344	0.0051	$1.8085 \cdot 10^{2}$	0.0000	4.9505	0.0031
14	0.0554	0.0000	∞	0.0000	∞	0.0000
15	0.0381	0.0013	$1.4554 \cdot 10^{10}$	0.0000	6.9869	0.0000
16	0.0307	0.0174	$1.0433 \cdot 10^{4}$	0.0000	4.9986	0.0029
17	0.0478	0.0000	$2.1225 \cdot 10^{5}$	0.0000	10.3437	0.0000
18	0.0474	0.0000	$1.5220 \cdot 10^{2}$	0.0000	8.5022	0.0000
19	0.0421	0.0002	2.5201	0.0001	5.7134	0.0000
20	0.0416	0.0003	4.4951	0.0000	6.7040	0.0000
21	0.0690	0.0000	$1.3151 \cdot 10^{23}$	0.0000	25.9682	0.0000
22	0.0512	0.0000	$2.8644 \cdot 10^{11}$	0.0000	13.3420	0.0000
23	0.0352	0.0040	$9.4186 \cdot 10^{4}$	0.0000	6.9954	0.0000
24	0.0482	0.0000	$5.2377 \cdot 10^{14}$	0.0000	12.7904	0.0000
25	0.0518	0.0000	$4.1476 \cdot 10^{6}$	0.0000	11.3507	0.0000
26	0.0471	0.0000	$2.1642 \cdot 10^{20}$	0.0000	7.8455	0.0000
27	0.0286	0.0332	1.0800	0.0007	3.8150	0.0108
28	0.0266	0.0567	$2.8377 \cdot 10^{-1}$	0.0099	4.0093	0.0087
29	0.0471	0.0000	∞	0.0000	∞	0.0000

Table 6.10: GoF tests for residuals of the Normal ARMA-GARCH stock model.

its associated p-values for all three GoF tests likewise are virtually zero up to at least four significant digits. Therefore, the characteristics of this stock are not ascertainable by neither the MVNTS nor the Normal model. The supremum-type AD test, however, rejects the above marginal null hypothesis of stdNTS for almost each of the stocks. This raises reasonable doubt concerning the overall appropriateness of this type of AD test for the situation under study. One reason for this apparently overly strict behavior in the formal sense might be explained by its membership in the class of supremum-type tests. Hence, the AD statistics sensitively reacts to outliers inevitably contained in data sets of large size. Nevertheless, AD seems to detect one distinct feature of the distributional fit which does not match the obtained standardized residuals. Again by definition of AD,

this feature is supposedly related to the weak extreme tails of the empirical frequency distribution. This suspected assumption is due to the fact that the ARMA-GARCH model now accounts for a considerable proportion of unconditional volatility and excess kurtosis. Recall that in the stationary model without ARMA-GARCH effect presented in chapter 5, the MVNTS distribution proved excellent fitting capabilities confirmed by all three GoF tests. Following these arguments, it may seem reasonable to not assign any major relevance to the case of supremum-type AD test for the time being. Instead, one relies on the significant verification of the marginal MVNTS fit indicated by KS and AD^2, although this aspect should further on be kept in mind. Even more so, the poor AD results appears negligible in the first instance since additional backtests performed in section 6.3.2 formally prove the MVNTS ARMA-GARCH to be a valid model for the DJIA stock return series in an out-of-sample period of high market volatility.

Finally considering the Normal ARMA-GARCH model, not the slightest evidence of agreement with empirical data can be found. With only a small number of stocks excepted, all three GoF tests simultaneously reject the Gaussian hypothesis for corresponding standardized residuals at a given significance level of 1%. This provides further support for the acknowledged fact that ARMA-GARCH time series models are not able to capture the entire amount of excess kurtosis and skewness when employing normal distributions for their standardized innovations.

Since the KS test covers the center of the distributions and the squared AD test assigns significant emphasis on the tails, the stdMVNTS proves to be a reasonable distributional assumption in the context of the ARMA-GARCH model. So far, this is at least true when only marginal distributions are considered. An alternative comparison of Normal and MVNTS ARMA-GARCH from a risk measurement perspective is presented in section 6.3.1, which moreover incorporates an implicit assessment the dependence structure.

Before turning to the analysis of equally-weighted portfolios and the application of the multivariate ARMA-GARCH time series models with regard to risk prediction, the day-ahead forecasts for VaR and AVaR on September 26, 2008 are illustrated in table 6.11. Although various empirical backtests are left to the forthcoming section 6.3.2, the sig-

nificant differences between the predictions under the MVNTS ARMA-GARCH and the Normal ARMA-GARCH model become clearly apparent in this comparison.

	VaR 1%		AVaR 1%	
k	Normal	MVNTS	Normal	MVNTS
1	0.0387	0.0378	0.0446	0.0457
2	0.0702	0.0750	0.0804	0.0900
3	0.1212	0.1221	0.1391	0.1462
4	0.0510	0.0548	0.0584	0.0661
5	0.1975	0.2333	0.2255	0.2810
6	0.0569	0.0605	0.0653	0.0728
7	0.0500	0.0545	0.0575	0.0659
8	0.0543	0.0613	0.0622	0.0743
9	0.0623	0.0625	0.0713	0.0751
10	0.0361	0.0380	0.0414	0.0457
11	0.0522	0.0534	0.0599	0.0640
12	0.0527	0.0593	0.0602	0.0719
13	0.0782	0.0776	0.0896	0.0934
14	0.0465	0.0534	0.0533	0.0643
15	0.0722	0.0766	0.0829	0.0923

	VaR 1%		AVaR 1%	
k	Normal	MVNTS	Normal	MVNTS
16	0.0570	0.0593	0.0652	0.0716
17	0.0463	0.0453	0.0532	0.0548
18	0.0273	0.0278	0.0313	0.0333
19	0.1821	0.1895	0.2089	0.2269
20	0.0389	0.0409	0.0446	0.0494
21	0.0452	0.0522	0.0517	0.0630
22	0.0547	0.0584	0.0627	0.0701
23	0.0557	0.0600	0.0636	0.0722
24	0.0319	0.0348	0.0366	0.0420
25	0.1097	0.1197	0.1258	0.1442
26	0.0528	0.0521	0.0610	0.0630
27	0.0532	0.0579	0.0610	0.0696
28	0.0404	0.0422	0.0462	0.0505
29	0.0524	0.0471	0.0601	0.0566

Table 6.11: Comparison of day-ahead VaR and AVaR forecasts for each of the stocks on September 26, 2008.

6.3 Portfolio Analysis

6.3.1 Risk Prediction

In order to adapt the model to currently available information in the best possible manner, one considers an ARMA-GARCH model with continuous re-estimation. This model variant offers improved flexibility as conditional mean and variance are not solely influenced by past realizations of the standardized innovations, which is sampled from a distribution with fixed parameter values. In fact, the coefficients of the ARMA-GARCH structure, processing the information contained in past realizations, are adjusted to newly available information every day. This has the advantage of not only being able to process current realizations in a fixed structure of equations reflecting the general characteristics of observed price dynamics, but to make this structure itself subject to constant change. Besides the ARMA-GARCH coefficients, this adaptation also affects the parameters of the distribution of innovations.

Formally, the ARMA coefficients $C_{k,t}$, $b_{k,t}$, $\eta_{k,t}$ as well as the GARCH coefficients $\alpha_{k,t}$, $\theta_{k,t}$ and $\beta_{k,t}$ in equations (6.1) and (6.2) are supplemented by an additional time index t to indicate that their estimation is based on a sample period $[t-\tau, t]$ of length τ. The same is true for the parameter set $(a_t, \lambda_t, \gamma_t, \mathsf{P}_t)$ of the stdMVNTS distribution of joint standardized innovations. Note that, in contrast to $\xi_{t,k}$ and $\sigma^2_{k,t}$, this time-dependency is not implied by the dynamic structure of equations (6.1) and (6.2) but rather an exogenously imposed assumption concerning the model inputs.

As was pointed out in section 4.2.2, the MVNTS distribution possesses a particular advantage with regard to its parametrization. According to equations (6.1) and (6.2), the conditional mean of the return R_{t+1} under the information $\mathcal{F}_t$ up to time t is[3]

$$\begin{aligned} E(R_{t+1}|\mathcal{F}_t) &= \left[E\left(C_{k,t} + b_{k,t}R_t^{(k)} + \eta_{k,t}\sigma_{k,t}\xi_{k,t} + \sigma_{k,t+1}\xi_{k,t+1}|\mathcal{F}_t\right)\right]^{\mathsf{T}}_{k=1,\ldots,n} \\ &= \left[C_{k,t} + b_{k,t}R_t^{(k)} + \eta_{k,t}\sigma_{k,t}\xi_{k,t} + \sigma_{k,t+1}\cdot E\left(\xi_{k,t+1}|\mathcal{F}_t\right)\right]^{\mathsf{T}}_{k=1,\ldots,n} \\ &= \left[C_{k,t} + b_{k,t}R_t^{(k)} + \eta_{k,t}\sigma_{k,t}\xi_{k,t}\right]^{\mathsf{T}}_{k=1,\ldots,n}. \end{aligned}$$

The corresponding vector of conditional variances is

$$\begin{aligned} Var(R_{t+1}|\mathcal{F}_t) &= \left[Var\left(C_{k,t} + b_{k,t}R_t^{(k)} + \eta_{k,t}\sigma_{k,t}\xi_{k,t} + \sigma_{k,t+1}\xi_{k,t+1}|\mathcal{F}_t\right)\right]^{\mathsf{T}}_{k=1,\ldots,n} \\ &= \left[\sigma^2_{k,t+1}\cdot Var\left(\xi_{k,t+1}|\mathcal{F}_t\right)\right]^{\mathsf{T}}_{k=1,\ldots,n} \\ &= \left[\sigma^2_{k,t+1}\right]^{\mathsf{T}}_{k=1,\ldots,n}. \end{aligned}$$

Therefore, equation (6.1) can be interpreted as a simple rescaling of the random vector of standardized innovations $\xi_{t+1} = (\xi_{1,t+1}, \ldots, \xi_{n,t+1})^{\mathsf{T}} \sim stdMVNTS(a_t, \lambda_t, \gamma_t, \mathsf{P}_t)$, leading to the conclusion that the day-ahead joint stock return is MVNTS distributed according

[3]Note that $\sigma_{t+1,k}$ is $\mathcal{F}_t$-measurable.

to

$$R_{t+1}|\mathcal{F}_t \sim MVNTS\left(a_t, \lambda_t, \gamma_t, \mathsf{P}_t, [\sigma_{k,t+1}]_{k=1,\dots,n}^{\mathsf{T}}, \left[C_{k,t} + b_{k,t}R_t^{(k)} + \eta_{k,t}\sigma_{k,t}\xi_{k,t}\right]_{k=1,\dots,n}^{\mathsf{T}}\right)$$

This is basically due to the fact that the parameters (c, d) of a MVNTS distribution are directly linked to the mean and standard deviations or the coefficient vectors of a simple linear transformation, respectively.

In the following, an equally-weighted portfolio (EWPF) comprising the $n = 29$ stocks is considered. The associated portfolio return R_{t+1}^{ewpf} is generated by the linear combination of dependent stock returns $R_{k,t+1}$

$$R_{t+1}^{ewpf} = \omega_{ewpf}^{\mathsf{T}} R_{t+1}\,,$$

using the weight vector ω_{ewpf} from equation (6.3). When applying the beneficial property of MVNTS distributions concerning linear combinations[4], one is able to conclude from equations (4.37) – (4.39) that $R_{t+1}^{ewpf} \sim NTS(a_t, \lambda_t, g_t, s_t^2, m_t)$ where

$$g_t = \omega_{ewpf}^{\mathsf{T}}\mathrm{diag}(\sigma_{t+1})\gamma_t$$

$$m_t = \omega_{ewpf}^{\mathsf{T}}\left[C_{k,t} + b_{k,t}R_t^{(k)} + \eta_{k,t}\sigma_{k,t}\xi_{k,t}\right]_{k=1,\dots,n}^{\mathsf{T}}$$

$$s_t = \left[\omega_{ewpf}^{\mathsf{T}}\mathrm{diag}(\sigma_{t+1})\mathrm{diag}(\mathcal{S}_t)\mathsf{P}_t\mathrm{diag}(\mathcal{S}_t)\mathrm{diag}(\sigma_{t+1})\omega_{ewpf}\right]^{\frac{1}{2}}$$

and

$$\mathcal{S}_t = \left[\sqrt{1 - \frac{2-a}{2\lambda}\gamma_{k,t}^2}\right]_{k=1,\dots,n}^{\mathsf{T}}$$

$$\sigma_{t+1} = [\alpha_{k,t} + \theta_{k,t}\sigma_{k,t}\xi_{k,t} + \beta_{k,t}\sigma_{k,t}]_{k=1,\dots,n}^{\mathsf{T}}\,.$$

[4]See section 4.2.4.

Figure 6.1 illustrates the development of the day-ahead predictions of VaR and AVaR of the EWPF return R_{t+1}^{ewpf} for a significance level of 1%. In this figure, the respective predictions are obtained from the time-adaptive versions of MVNTS ARMA-GARCH and Normal ARMA-GARCH and their dynamic parameter estimates. Both models use a historical sample period of 10 years, which enables predictions in the time interval between September 1, 2007 and September 25, 2009 on the available data set.

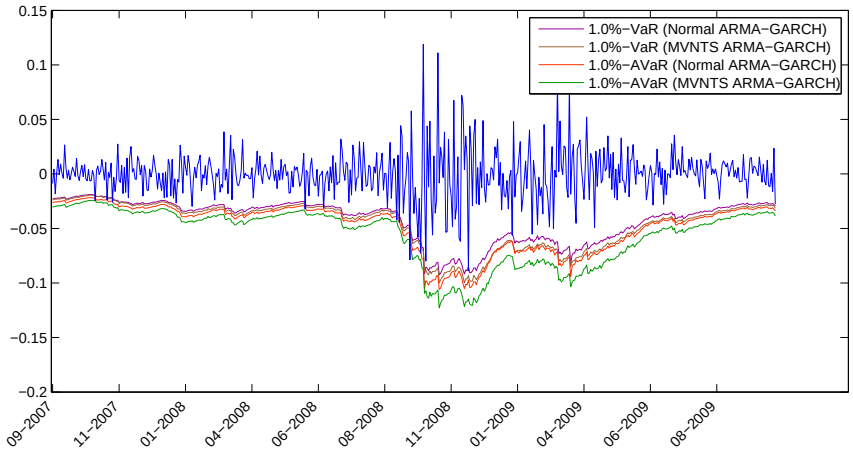


Figure 6.1: Sequence of EWPF VaR_α and AVaR_α predictions of MVNTS ARMA-GARCH and Normal ARMA-GARCH, $\alpha = 0.01$.

With the entire sequence of estimated ARMA-GARCH model parameters for $t = 1, \ldots, N$, a further analysis of the market crash on September 29, 2008 is carried out. On that day, the EWPF would have generated a loss of 7.87 %. The daily probabilities for an event with equal or even greater losses under the MVNTS ARMA-GARCH model and the Normal ARMA-GARCH model with time adapted parameters and coefficients are compiled in table 6.12. Given that standardized innovations are independent over time, the respective probabilities are converted into expected time intervals between crashes of this magnitude. While the frequency of roughly 10 years implied by the MVNTS ARMA-GARCH model appears to be in conformity with observable reality, the corresponding statement of the Normal ARMA-GARCH can reasonably be doubted.

	Event probability	Exp. waiting time
Normal ARMA-GARCH	0.0048 %	82.2743 years
MVNTS ARMA-GARCH	0.0539 %	7.3688 years

Table 6.12: Daily market crash probabilities and corresponding waiting times.

	1%-VaR	1%-AVaR	0.5%-VaR	0.5%-AVaR
Normal ARMA-GARCH	0.0469	0.0537	0.0519	0.0583
MVNTS ARMA-GARCH	0.0496	0.0597	0.0567	0.066

Table 6.13: EWPF VaR_α and AVaR_α predictions of MVNTS ARMA-GARCH and Normal ARMA-GARCH, $\alpha = 0.01$, $\alpha = 0.005$.

6.3.2 ARMA-GARCH Backtests

Various formal backtests concerning the adequateness of an estimated ARMA-GARCH model and its distributional assumption for the implied standardized residuals are based on the number of quantile violations. For each time step contained in a backtesting period, the realized value of the standardized residual is compared to upper and lower α-quantiles of the assumed distribution of standardized innovation. A violation occurs when the residual is not within the interval between upper and lower quantile. A model is rejected when violations occur too often or too rarely in the course of time, based on certain formal criteria. On average, 2α violations should occur as a reference value when the model is correct. A violation on the downside can be interpreted as a realized return falling below the negative VaR for significance level α. Two such tests are the Christoffersen test[5] and the Berkowitz test[6], which both implement the basic approach of likelihood ratio tests. The results of their different variants are presented in tables 6.14 and 6.15. Specifically, MVNTS ARMA-GARCH models with dynamic parameter values are employed to generate distribution forecasts of EWPF returns for the time horizon of one day. These dynamic forecasts are compared to the EWPF return realizations in the 2-year backtesting period from September 18, 2007 until September 16, 2009. The individual test variants are in

[5]See Christoffersen in [Chri98].
[6]See Berkowitz in [Berk01].

case of the Christoffersen test

- unconditional coverage (LR_{uc})
- test for serial independence (LR_{ind1})
- test for conditional coverage (LR_{cc})

and in the case of the Berkowitz test

- lower tail test (LR_{tail})
- test for serial independence (LR_{ind2}) .

All likelihood ratios (LR) can be easily converted into corresponding p-values, since they are asymptotically χ^2-distributed. The coverage and the tail tests are based on quantiles for $\alpha = 0.01$ and $\alpha = 0.005$, as indicated in the respective results table.

	Christofferson tests		
	LR_{uc} (p-value)	LR_{ind1} (p-value)	LR_{cc} (p-value)
Normal ARMA-GARCH	5.3222 (0.0211)	0.4919 (0.4831)	5.8141 (0.0546)
MVNTS ARMA-GARCH	2.5482 (0.1104)	0.3280 (0.5669)	2.8762 (0.2374)

	Berkowitz tests	
	LR_{tail} (p-value)	LR_{ind2} (p-value)
Normal ARMA-GARCH	7.1044 (0.0287)	7.0075 (0.0081)
MVNTS ARMA-GARCH	2.7303 (0.2553)	6.0850 (0.0136)

Table 6.14: Christoffersen and Berkowitz backtests for $\alpha = 0.01$.

Before turning to the interpretation of the results, it has to be emphasized at this point that these tests do not only concern the ARMA-GARCH structure and the marginal distribution but explicitly incorporate the dependence structure of the stdMVNTS distribution assumption for joint innovations, since EWPF returns or linear combinations are considered.

	Christofferson tests		
	LR_{uc} (p-value)	LR_{ind1} (p-value)	LR_{cc} (p-value)
Normal ARMA-GARCH	10.0375 (0.0015)	0.3280 (0.5669)	10.3654 (0.0056)
MVNTS ARMA-GARCH	0.0866 (0.7686)	0.0360 (0.8495)	0.1226 (0.9406)

	Berkowitz tests	
	LR_{tail} (p-value)	LR_{ind2} (p-value)
Normal ARMA-GARCH	9.9465 (0.0069)	7.0075 (0.0081)
MVNTS ARMA-GARCH	0.0980 (0.9522)	6.0850 (0.0136)

Table 6.15: Christoffersen and Berkowitz backtests for $\alpha = 0.005$.

For $\alpha = 0.01$, Normal and MVNTS ARMA-GARCH models show the same performance in the Christoffersen tests. None of them can be rejected for significance levels below 10%. In the Berkowitz tail test under the same conditions, the Normal ARMA-GARCH model is rejected at the significance level of 1%, while the MVNTS ARMA-GARCH model is not. When α is set to a value of 0.005, the tests assign increased relevance to more extreme events. The Christoffersen tests for conditional and unconditional coverage as well as the Berkowitz tail test reject the Normal ARMA-GARCH model at significance level of 1%. The MVNTS ARMA-GARCH model passes[7] these three tests with resulting p-values ranging between 0.06 and 0.16. Therefore, the shortcomings of Normal ARMA-GARCH models become clearly evident when extreme risks are to be quantified. For this reason Normal ARMA-GARCH models do not seem to be suitable for implementing sustainable risk management principles. A MVNTS ARMA-GARCH model should be the preferred choice for performing these tasks, instead.

The assumption of independence is not doubted in the Christoffersen test, whose results are naturally not affected by different values of α. The Berkowitz test rejects the hypothesis of independent draws at a significance level of 1% for the Normal ARMA-GARCH model while this is not possible for the MVNTS ARMA-GARCH model. However, one has to consider the substantial differences between the Christoffersen and the Berkowitz test concerning their results with regard to the independence hypothesis.

[7] In the sense that they cannot be rejected at the respective significance levels.

6.4 Literature Review

At this point, a comprehensive summary of similar multivariate Brownian subordination approaches present in the literature to this date should be given.

Øigård et al. [ØHHG05] recently defined a simple multivariate extension of the NIG model based on Brownian subordination by maintaining the univariate IG subordinator and only replace the original univariate Brownian motion by its multivariate counterpart. Moreover, these authors present an efficient expectation maximization (EM) algorithm for the maximum likelihood (ML) estimation of the parameter values based on a method developed by Karlis [Karl02]. Aas et al. [AaHHD06] apply the multivariate NIG (MNIG) to financial econometrics involving GARCH time series models. With their estimated model they are able to show the superiority of this multivariate distribution assumption over the Gaussian alternative in the context of optimal portfolios and in-sample performance evaluations.

One has to remark however, that MNIG was not one of the first multivariate Brownian subordination-based models, however. Previously, a series of *multivariate skewed-t* distributions in different specifications were introduced by Azzalini and Capitanio [AzCa03], Bauwens and Laurent [BaLa05] and Dokov et al. [DoSR08], where the latter article emphasizes on the practical computation of the common risk measures Value-at-Risk and Average Value-at-Risk and the construction of the distribution via Brownian subordination by means of employing an Inverse Gamma subordination process. In the work of the former authors it has to be primarily criticized that their versions of the multivariate skewed-t distribution fail when data exhibits asymmetric tails, a feature which can be captured by the MNIG instead[8]. This shortcoming, however, was remedied by the further development and sophistication achieved by Dokov et al. Moreover, while the multivariate skewed-t distribution possesses polynomial decay of the distribution tails, this decay is of exponential type in the case of the MNIG distribution, a feature which appears to be more consistent with empirical observations of asset returns.

The principal advantage of the MNIG and the multivariate skewed-t model lies in the fact

[8] see Aas et al. [AaHHD06].

that both are embedded in the more general class of multivariate generalized hyperbolic distributions discovered by Barndorff-Nielsen [Barn77], which is due to utilizing a specific subordinator from the generalized inverse Gaussian class. In particular, as a result, both MNIG and the skewed-t approach allow for closed form probability density functions.

One of the latest attempts in this particular direction is made by Eberlein and Madan [EbMa10]. In their setting of Brownian subordination, they establish a multivariate Brownian subordination with the distinguishing innovation that each dimension is driven by a separate independent Gamma process for the virtual time flow. By this construction a rather advanced and flexible multivariate extension of the familiar Variance Gamma model is obtained, which is denoted by MVVG. The drawback one has to accept by the choice of separate independent subordinators, is that the characteristic function as the central means of representation of many non-Gaussian processes is not available in closed form. This considerably complicates the theoretical handling as well as the practical application of the model. Moreover, one cannot ensure that the resulting distributions maintain the property of infinite divisibility. One of their most noteworthy findings lies in the universal fact that even in their flexible setting the achieved correlation structure of the distribution is rather limited. In other words, one would not be able to calibrate the MVVG to a more extreme correlation structure, where a Gaussian model is able to incorporate any given covariance matrix. However, for capturing realistic empirical correlation patterns, the capabilities of the MVVG seem to be sufficient in this respect. In an application to portfolio optimization the MVVG nevertheless delivers superior performance results compared to the Gaussian alternative as it is able to better capture the present dependence structures of higher order which go beyond simple correlation.

Chapter 7

Conclusion and Outlook

7.1 Conclusion

Adequate statistical modeling of multidimensional asset price processes constitutes one of the principal tasks in several areas of modern quantitative finance. One of the most critical aspects are asymmetry and features of heavy-tailedness of return distributions. Moreover, volatility clustering is observed on the time scale. These phenomena are widely acknowledged as stylized facts of financial market data.

Basic concepts and definitions are briefly reviewed in chapter 2. This review encompasses the theory of multivariate Lévy processes, the time-change of Lévy processes as well as a compact overview with respect to univariate tempered stable processes.

In order to prepare a transition to multivariate extensions in chapter 4, univariate normal tempered stable (NTS) models are thoroughly treated in chapter 3. The presentation is oriented towards a consistent formulation of the associated processes and distributions within the Brownian subordination perspective. This includes a detailed derivation of various process and distribution characteristics. It is achieved by the application of general calculation methods for Lévy processes to the specific setting of a Brownian subordination with a CTS process for the virtual time. The focus is on a clear and transparent description of the procedure. This should ease the application of the methods when turning to the

more complex case of $n > 2$ dimensions. A supplementing result is the correspondence in terms of parameters between the two existing definitions of the NTS model. This particular relation proves only to be present in the one-dimensional case, however.

Chapter 4 introduces the multivariate normal tempered stable (MVNTS) process. Its definition is likewise based on the construction technique of Brownian subordination in a multivariate setting. Subsequently, the Lévy triplet as the fundamental means of process specification, is derived. This is accompanied by associated multivariate characteristic functions and distributional moments. Based on the evaluation of the process after a unit time interval, the corresponding MVNTS distribution is defined. Particular care was taken to design a parametrization that implies a reasonable degree of tractability with respect to practical application. In this context, it proves beneficial to consider a class of distributions with zero mean and unit variances first, before proceeding to a version with arbitrary moments, which is created by a simple linear rescaling. Moreover, properties of various kinds of transformations such as linear mappings and linear combinations are derived. These two aforementioned operations prove to be particulary relevant in the further empirical analysis. In addition to the advantages offered by the specific parametrization, the general construction approach of Brownian subordination enables the circumvention of many complex calculations involved with the derivation of joint moments, characteristic functions as well as probability density and cumulative distribution functions.

The empirical assessment of model performance and adaptability is subdivided into several individual parts. In order to implement a first application of the MVNTS distribution, a simple i.i.d. model for the unconditional joint daily returns of the DJIA index and the stocks contained therein is stated in chapter 5. Subsequently, a full estimation procedure for the entire parameter set of the MVNTS distribution is developed. It is shown that there exists a particular possibility for decomposing the estimation problem into several sequential steps. Each step is only concerned with a certain parameter subset, where not more than three parameters have to be estimated simultaneously by means of maximum likelihood. This decomposition ensures improved stability and makes the method feasible even for a higher number of dimensions. It can basically be attributed to the specific choice

of the definition and associated parametrization. By close connection to the underlying Brownian motion, calibration of the dependence structure can be easily achieved.

With this procedure at hand, an unconditional model of the DJIA index returns and the joint stock returns are estimated. The estimation is based on an empirical 10-years sampling interval from 1999 through 2008. Assessing the goodness of fit of marginal distributions, the model generates sufficiently high p-values in the application of the KS and both conventional and squared AD tests. Since KS and AD tests assign different weights to the center and the tails of the distribution, the statistical fit is evaluated on the entire support. These formal test results are further substantiated by graphical inspections, where an alternative Gaussian model proves to be insufficient. A result of the same quality was obtained when putting the modeled n-dimensional MVNTS dependence structure to a test, which was particulary realized by creating an equally-weighted portfolio (EWPF). However, one has to note a slight offset in the center of the distributions, indicating an imperfectly captured dependence structure.

Furthermore, the chapter presents a new numerical inversion scheme of the characteristic function for generating approximations of the CDF. This approach is based on the Gil-Pelaez inversion theorem. Compared to the conventional CDF inversion, this method does not rely on the existence of exponential moments and proves to be universally applicable.

This advantage in MVNTS fitting precision can be exploited to construct optimal portfolios in a static context. Besides the comparison of efficient frontiers VaR and AVaR, various alternative performance ratios are placed in opposition to each other. They result in rather different portfolio weights. Two representative static strategies were subsequently exposed to an in-sample as well as an out-of-sample evaluation. The in-sample period comprises a pre-crisis time interval from October 1997 to September 2007, while the out-of-sample period from October 2007 through September 2009 contains market conditions with extremely high volatilities. In both periods, the Sharpe ratio-optimal strategy carries higher risks while on the other hand offering a higher reward potential compared to an opposing STARR-optimal strategy derived under MVNTS assumption. When aggregating risk and reward into a single performance ratio, the conventional ratios consistently favor the

Sharpe ratio-optimized strategy in the in-sample analysis, while in the more decisive out-of-sample analysis the STARR-optimal strategy is consistently preferred.

The study of portfolios involves linear combinations of asset returns. It is shown that linear combinations of MVNTS dimensions result in a univariate NTS distributed random variable. Moreover, by the specifically designed estimation procedure based on the properties of the parametrization, it is possible to merge two specified MVNTS models into a joint model. The latter property proves to be essential for employing the DJIA index returns as the benchmark in the course of portfolio optimization.

The second part of the empirical study addresses the application of the MVNTS distribution in the context of an advanced multivariate ARMA-GARCH time series model for DJIA returns. Chapter 6 compares the MVNTS assumption to the multivariate Gaussian assumption for the distribution of ARMA-GARCH innovations. Statistical goodness of fit tests indicate a clear superiority of the MVNTS assumption over the alternative model with Gaussian innovations. This becomes more evident when comparing the 1-day predictions of VaR and AVaR of continuously updated ARMA-GARCH models over a time interval of 2 years. One has to stress the fact that these results were obtained during the crisis period between 2007 and 2009, where Gaussian ARMA-GARCH innovations are not able to adequately capture the dynamics of volatile market conditions. Although the differences are not as significant as in the unconditional framework, the subsequent Christoffersen and Berkowitz tests for various criteria tend to favor the MVNTS assumption.

In summary, two goals were accomplished in the present thesis. First, one defined a versatile stochastic model in an arbitrary number of dimensions. A comprehensive description of its representations and properties together with a toolbox for basic manipulations is provided. The particular construction approach in combination with an expediently designed parametrization results in a remarkably tractable model. This enables efficient estimation method as well as eased implementation of portfolio optimization. On the other hand, this tractability is available in spite of a considerable level of flexibility and fitting capability. The superior adaptability becomes evident in particular during volatile market periods.

Moreover, the MVNTS model is one of the first multivariate extensions which genuinely implement the common principle of tempered stable models. Many of its highly desirable properties would vanish when merely defining a multivariate distribution from tempered stable marginal distributions, combined by means of a copula. In addition, it was demonstrated that previous approaches, like the multivariate versions of variance gamma and normal inverse Gaussian processes, are embedded in the MVNTS model.

Finally, the MVNTS model might possess the potential of gaining a respectable level of acceptance. The reason for this is that the widespread assumption of Brownian motions does not have to be completely disregarded. Only the idea of a linearly and deterministically evolving trading time has to be replaced by a stochastic process of virtual trading time, while the Brownian motion can be basically maintained.

7.2 Outlook

This section concludes the thesis by identifying areas of further research.

For situations where NTS distributions might turn out to be inadequate, one can still make use of the natural dependence structure inherently contained in the MVNTS distribution. For this purpose, the associated MVNTS copula can be isolated. The parametrization of this copula is closely related to the parametrization of a stdMVNTS. Since the MVNTS is generated by a mixture model of elliptical distributions, the MVNTS copula is capable of representing asymmetric dependence structures. This property proves relevant, as in financial market data one can observe that the dependence of lower tails is significantly stronger than the dependence of upper tails[1]. Therefore, a model using the MVNTS copula is more likely to be universally valid for market up-swings as well as market downturns.

An improved flexibility of the MVNTS dependence structure could be achieved when appointing an individual but mutually independent CTS subordinator to each dimension of the multivariate Brownian motion. However, this would entail the vanishing of the considerable level of tractability that accompanied the original model definition. As has

[1] See McNeil et al. [McFE05].

been pointed out, both the joint multivariate distribution as well as the distribution of linear combinations do not possess a closed-form characteristic function. This in turn makes it hard to tell whether this subordination generates Lévy processes or infinitely divisible distributions at all.

Bibliography

[AaHHD06] Aas, K., Hobæk Haff, I., and Dimakos, X. K. Risk estimation using the multivariate normal inverse Gaussian distribution. *Journal of Risk 8*, 2 (2006), 39–60.

[AcTa02] Acerbi, C., and Tasche, D. Expected shortfall: a natural coherent alternative to value at risk. *Economic Notes 31*, 2 (2002), 379–388.

[Ande03] Anderson, T. W. *An introduction to multivariate statistical analysis*, 3. ed. Wiley series in probability and statistics. Wiley, Hoboken, NJ, 2003.

[AnDa52] Anderson, T. W., and Darling, D. A. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics 23*, 2 (June 1952), 193–212.

[Appl05] Applebaum, D. *Lévy processes and stochastic calculus*, 2. ed., vol. 93 of *Cambridge studies in advanced mathematics*. Cambridge University Press, Cambridge, 2005.

[ADEH99] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. Coherent measures of risk. *Mathematical Finance 9* (1999), 203–228.

[AzCa03] Azzalini, A., and Capitanio, A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65*, 2 (May 2003), 367–389.

[Bach00] Bachelier, L. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure 17*, 3 (1900), 21–86.

[BaSw94] Bailey, D. H., and Swarztrauber, P. N. A fast method for the numerical evaluation of continuous Fourier and Laplace transforms. *SIAM Journal on Scientific Computing 15*, 5 (September 1994), 1105–1110.

[Barn77] Barndorff-Nielsen, O. E. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society A 353* (1977), 401–419.

[Barn97] Barndorff-Nielsen, O. E. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*, 1 (March 1997), 1–13.

[Barn98] Barndorff-Nielsen, O. E. Processes of normal inverse Gaussian type. *Finance and Stochastics 2*, 1 (1998), 41–68.

[BNLe01] Barndorff-Nielsen, O. E., and Levendorskiĭ, S. Z. Feller processes of normal inverse Gaussian type. *Quantitative Finance 1*, 3 (2001), 318–331.

[BNMR01] Barndorff-Nielsen, O. E., Mikosch, T., and Resnick, S. I., Eds. *Lévy processes: theory and applications*. Birkhäuser, Boston, 2001.

[BNSh02] Barndorff-Nielsen, O. E., and Shephard, N. Normal modified stable processes. *Theory of Probability and Mathematical Statistics 65* (2002), 1–20.

[BaLa05] Bauwens, L., and Laurent, S. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics 23*, 3 (July 2005), 346–354.

[Bell68] Bell, W. W. *Special functions for scientists and engineers*. Van Nostrand, London, 1968.

[Berk01] Berkowitz, J. Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics 19*, 4 (October 2001), 465–474.

[Bert05] Bertoin, J. *Lévy processes*, 1. paperback ed., 4. print. ed., vol. 121 of *Cambridge tracts in mathematics*. Cambridge University Press, Cambridge, 2005.

[BRKF10a] Bianchi, M. L., Rachev, S. T., Kim, Y. S., and Fabozzi, F. J. Tempered infinitely divisible distributions and processes. *Theory of Probability and its Applications 55*, 1 (2010), 59–86.

[BRKF10b] Bianchi, M. L., Rachev, S. T., Kim, Y. S., and Fabozzi, F. J. Tempered stable distributions and processes in finance: numerical analysis. In *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, M. Corazza and C. Pizzi, Eds. Springer, Milan, 2010, pp. 33–42.

[BORS04] Biglova, A., Ortobelli, S., Rachev, S. T., and Stoyanov, S. V. Different approaches to risk estimation in portfolio theory. *The Journal of Portfolio Management 31*, 1 (Fall 2004), 103–112.

[BiBr99] Bilodeau, M., and Brenner, D. *Theory of multivariate statistics*. Springer texts in statistics. Springer, New York, 1999.

[BlSc73] Black, F., and Scholes, M. The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 3 (1973), 637–654.

[BüTr94] Büning, H., and Trenkler, G. *Nichtparametrische statistische Methoden*, 2., ext. and rev. ed. de Gruyter, Berlin, 1994.

[Boll86] Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 3 (1986), 307–327.

[BoLe00] Boyarchenko, S., and Levendorskiĭ, S. Option pricing for truncated Lévy processes. *International Journal for Theory and Applications in Finance 3* (2000), 549–552.

[Broo08] Brooks, C. *Introductory econometrics for finance*, 2. ed. Cambridge University Press, 2008.

[CGMY02] Carr, P., Geman, H., Madan, D. B., and Yor, M. The fine structure of asset returns: an empirical investigation. *Journal of Business 75*, 2 (2002), 305–322.

[CaMa98] Carr, P. P., and Madan, D. B. Option valuation using the fast Fourier transform. *Journal of Computational Finance 2* (1998), 61–73.

[ChRF07] Chernobai, A. S., Rachev, S. T., and Fabozzi, F. J. *Operational risk: a guide to Basel II capital requirements, models, and analysis*. The Frank J. Fabozzi series, Wiley finance. Wiley, Hoboken, N.J., 2007.

[ChLV05] Cherubini, U., Luciano, E., and Vecchiato, W. *Copula methods in finance*, repr. ed. Wiley finance series. Wiley, Chichester, 2005.

[Chri98] Christoffersen, P. F. Evaluating interval forecasts. *International Economic Review 39* (1998), 841–862.

[Clar73] Clark, P. K. A subordinated stochastic process model with finite variation for speculative prices. *Econometrica 41*, 1 (January 1973), 135–155.

[CoTa04] Cont, R., and Tankov, P. *Financial modelling with jump processes*. Chapman & Hall/CRC financial mathematics series. Chapman & Hall/CRC, Boca Raton, Fla., 2004.

[CMTT90] Crawford Moss, L. L. C., Taylor, M. S., and Tingey, H. B. Quantiles of the Anderson-Darling statistic. *Communications in Statistics - Simulation and Computation 19*, 3 (1990), 1007–1014.

[DeMc05] Demarta, S., and McNeil, A. J. The t-copula and related copulas. *International Statistical Review 73*, 1 (2005), 111–129.

[DoSR08] Dokov, S., Stoyanov, S. V., and Rachev, S. T. Computing VaR and AVaR of skewed-t distribution. *Journal of Applied Functional Analysis 3* (2008), 189–209.

[Eber01] Eberlein, E. Applications of generalized hyperbolic Lévy motion to finance. In *Lévy Processes – Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick, Eds. Birkhäuser, Boston, 2001, pp. 319–336.

[EbKP98] Eberlein, E., Keller, U., and Prause, K. New insights into smile, mispricing and value at risk: The hyperbolic model. *Journal of Business 71* (1998), 371–405.

[EbMa10] Eberlein, E., and Madan, D. B. On correlating Lévy processes. *Journal of Risk 13*, 1 (2010), 3–16.

[EbRa01] Eberlein, E., and Raible, S. Some analytic facts on the generalized hyperbolic model. In *Proceedings of the 3rd European Meeting of Mathematicians* (2001), vol. 202.

[EmLM03] Embrechts, P., Lindskog, F., and McNeil, A. Modeling dependence with copulas and applications to risk management. In *Handbook of Heavy-Tailed Distributions in Finance*, S. T. Rachev, Ed., Handbooks in Finance: Book 1. Elsevier/North Holland, 2003, pp. 329–384.

[Engl82] Engle, R. F. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica 50* (1982), 987–1008.

[Fama63] Fama, E. F. Mandelbrot and the stable Paretian hypothesis. *The Journal of Business 36*, 4 (October 1963), 420–429.

[Fama65] Fama, E. F. The behavior of stock market prices. *Journal of Business 38* (1965), 34–105.

[Fama70] Fama, E. F. Efficient capital markets: A review of theory and empirical work. *Journal of Finance 25*, 2 (1970), 383–417.

[Fell68] Feller, W. *An introduction to probability theory and its applications*, 3. ed., rev. print. ed., vol. 1 of *Wiley series in probability and mathematical statistics*. Wiley, New York, NY, 1968.

[Fell71] Feller, W. *An introduction to probability theory and its applications*, 2. ed., repr. ed., vol. 2 of *Wiley series in probability and mathematical statistics*. Wiley, New York, NY, 1971.

[GeMY01] Geman, H., Madan, D. B., and Yor, M. Asset prices are Brownian motions: Only in business time. In *Quantitative Analysis in Financial Markets*, M. Avellaneda, Ed. World Scientific: River Edge, New Jersey, 2001, pp. 103–146.

[GilP51] Gil-Pelaez, J. Note on the inversion theorem. *Biometrika 38* (1951), 481–482.

[Grei11] Greiner, S. P. *Ben Graham Was a Quant: Raising the IQ of the Intelligent Investor*, 1. ed. Wiley Finance. Wiley, April 2011.

[HaØi01] Hansen, A., and Øigård, T. A. The normal inverse Gaussian distribution: A versatile model for heavy-tailed stochastic processes. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (May 2001), vol. 6, pp. 3985–3988.

[HeKl02] Henze, N., and Klar, B. A treatment of multivariate skewness, kurtosis and related statistics. *Journal of Multivariate Analysis 83* (2002), 141–165.

[KaSh00] Karatzas, I., and Shreve, S. E. *Brownian motion and stochastic calculus*, 2. ed., corr. 6. print. ed., vol. 113 of *Graduate texts in mathematics*. Springer, New York, 2000.

[Karl02] Karlis, D. An EM type algorithm for maximum likelihood estimation for the normal inverse Gaussian distribution. *Statistics & Probability Letters 57* (2002), 43–52.

[Kim05] Kim, Y. S. *The Modified Tempered Stable Process with Application to Finance*. PhD thesis, Sogang University, Seoul, Department of Mathematics, 2005.

[KRBF08a] Kim, Y. S., Rachev, S. T., Bianchi, M. L., and Fabozzi, F. J. Financial market models with Lévy processes and time-varying volatility. *Journal of Banking and Finance 32*, 7 (2008), 1363–1378.

[KRBF08b] Kim, Y. S., Rachev, S. T., Bianchi, M. L., and Fabozzi, F. J. A new tempered stable distribution and its application to finance. In *Risk Assessment: Decisions in Banking and Finance*, G. Bol, S. T. Rachev, and R. Würth, Eds., Contributions to Economics. Physica-Verlag, Springer, Heidelberg, 2008, pp. 77–109.

[KRBF09] Kim, Y. S., Rachev, S. T., Bianchi, M. L., and Fabozzi, F. J. Computing VaR and AVaR in infinitely divisible distributions. *Yale ICF Working Paper Series 09-07* (March 2009), 1–37.

[KRBF10] Kim, Y. S., Rachev, S. T., Bianchi, M. L., and Fabozzi, F. J. Tempered stable and tempered infinitely divisible GARCH models. *Journal of Banking & Finance 34*, 9 (September 2010), 2096–2109.

[KRBM+11] Kim, Y. S., Rachev, S. T., Bianchi, M. L., Mitov, I., and Fabozzi, F. J. Time series analysis of financial market meltdowns. *Journal of Banking & Finance 35*, 8 (August 2011), 1879–1891.

[KRCB08] Kim, Y. S., Rachev, S. T., Chung, D. M., and Bianchi, M. L. A modified tempered stable distribution with volatility clustering. In *New Developments in Financial Modelling*, J. O. Soares, J. P. Pina, and M. Catalao-Lopes, Eds. Cambridge Scholars Publishing, 2008, pp. 344–365.

[KRCB09] Kim, Y. S., Rachev, S. T., Chung, M. D., and Bianchi, M. L. The modified tempered stable distribution, GARCH models and option pricing. *Probability and Mathematical Statistics 29*, 1 (2009), 91–117.

[Kopo95] Koponen, I. Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. *Physical Review E 52*, 1 (July 1995), 1197–1199.

[Kypr06] Kyprianou, A. E. *Introductory lectures on fluctuations of Lévy processes with applications*. Universitext. Springer, Berlin, 2006.

[Lé48] Lévy, P. *Processus stochastiques et mouvement brownien*, vol. 6 of *Monographies des probabilités*. Gauthier-Villars, Paris, 1948.

[Lewi61] Lewis, P. A. W. Distribution of the Anderson-Darling statistic. *The Annals of Mathematical Statistics 32*, 4 (December 1961), 1118–1124.

[Luka70] Lukacs, E. *Characteristic functions*, 2. ed., rev. and enlarged ed. Griffin books of cognate interest. Griffin, London, 1970.

[MaCC98] Madan, D. B., Carr, P. P., and Chang, E. C. The variance gamma process and option pricing. *European Finance Review 2*, 1 (1998), 79–105.

[MaSe87] Madan, D. B., and Seneta, E. Simulation of estimates using the empirical characteristic function. *International Statistical Review 55*, 2 (August 1987), 153–161.

[MaSe90] Madan, D. B., and Seneta, E. The variance gamma (v.g.) model for share market returns. *The Journal of Business 63*, 4 (October 1990), 511–524.

[Mand63] Mandelbrot, B. The variation of certain speculative prices. *Journal of Business 36*, 4 (October 1963), 394–419.

[Mand67] Mandelbrot, B. The variation of some other speculative prices. *Journal of Business 40*, 4 (October 1967), 393–413.

[MaTa67] Mandelbrot, B., and Taylor, H. On the distribution of stock price differences. *Operations Research 15*, 6 (November - December 1967), 1057–1062.

[Mark52] Markowitz, H. M. Portfolio selection. *The Journal of Finance 7*, 1 (March 1952), 77–91.

[MaMa04] Marsaglia, G., and Marsaglia, J. Evaluating the Anderson-Darling distribution. *Journal of Statistical Software 9*, 2 (2004), 1–5.

[McCu87] McCullagh, P. *Tensor methods in statistics*. Monographs on statistics and applied probability. Chapman & Hall, London, 1987.

[McFE05] McNeil, A. J., Frey, R., and Embrechts, P. *Quantitative risk management: concepts, techniques and tools*. Princeton series in finance. Princeton University Press, Princeton, NJ, 2005.

[Mert73] Merton, R. C. Theory of rational option pricing. *Bell Journal of Economics and Management Science 4*, 1 (Spring 1973), 141–183.

[MöPe10] Mörters, P., and Peres, Y. *Brownian motion*, 1. publ. ed., vol. 30 of *Cambridge series in statistical and probabilistic mathematics*. Cambridge University Press, Cambridge, 2010.

[Nels06] Nelsen, R. B. *An introduction to copulas*, 2. ed. Springer series in statistics. Springer, New York, 2006.

[Nola12] Nolan, J. P. *Stable Distributions - Models for Heavy Tailed Data*. Birkhäuser, Boston, 2012. In progress, Chapter 1-3 online at academic2.american.edu/∼jpnolan.

[ØHHG05] Øigård, T. A., Hanssen, A., Hansen, R. E., and Godtliebsen, F. EM-estimation and modeling of heavy-tailed processes with the multivariate normal inverse Gaussian distribution. *Signal Processing 85* (2005), 1655–1673.

[Økse07] Øksendal, B. K. *Stochastic differential equations: an introduction with applications*, 6. ed., corr. 4. print. ed. Universitext. Springer, Berlin, 2007.

[RKBF11] Rachev, S. T., Kim, Y. S., Bianchi, M. L., and Fabozzi, F. J. *Financial Models with Lévy Processes and Volatility Clustering*, 1. ed. The Frank J. Fabozzi Series, Wiley Finance. John Wiley & Sons, 2011.

[RaMF05] Rachev, S. T., Menn, C., and Fabozzi, F. J. *Fat-tailed and skewed asset return distributions: implications for risk management, portfolio selection, and option pricing.* The Frank J. Fabozzi series. Wiley, Hoboken, NJ, 2005.

[RaMi00] Rachev, S. T., and Mittnik, S. *Stable Paretian models in finance.* Series in financial economics and quantitative analysis. Wiley, Chichester, 2000.

[RMFF+07] Rachev, S. T., Mittnik, S., Fabozzi, F. J., Focardi, S. M., and Jašić. *Financial econometrics: from basics to advanced modeling techniques.* The Frank J. Fabozzi series, Wiley Finance. John Wiley & Sons, Inc., Hoboken, NJ, 2007.

[RaSF08] Rachev, S. T., Stoyanov, S. V., and Fabozzi, F. J. *Advanced stochastic models, risk assessment, and portfolio optimization: the ideal risk, uncertainty, and performance measures.* The Frank J. Fabozzi series, Wiley Finance. John Wiley & Sons, Inc., Hoboken, NJ, 2008.

[Rosi07] Rosiński, J. Tempering stable processes. *Stochastic Processes and their Applications 117*, 6 (June 2007), 677–707.

[Ross76] Ross, S. The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*, 3 (December 1976), 341–360.

[Rota99] Rotar, V. *Probability Theory.* World Scientific Publishing, 1999.

[Rydb97] Rydberg, T. The normal inverse Gaussian Lévy process: Simulation and approximation. *Communications in Statistics. Stochastic Models 13*, 4 (1997), 887–910. ISSN 0882-0287.

[Rydb99] Rydberg, T. H. Generalized hyperbolic diffusion processes with applications in finance. *Mathematical Finance 9*, 2 (April 1999), 183–201.

[SSØH01] Salberg, A., Swami, A., Øigård, T. A., and Hansen, A. The normal inverse Gaussian distribution as a model for MUI. In *35th Asilomar Conference on Signals, Systems and Computers* (2001), vol. 1, pp. 143–147.

[SaTa00] Samorodnitsky, G., and Taqqu, M. S. *Stable non-Gaussian random processes: stochastic models with infinite variance*, 1. repr. ed. Stochastic modeling. Chapman & Hall/CRC, Boca Raton, Fla., 2000.

[Sato99] Sato, K.-I. *Lévy processes and infinitely divisible distributions*, vol. 68 of *Cambridge studies in advanced mathematics*. Cambridge University Press, Cambridge, 1999.

[Scho03] Schoutens, W. *Lévy processes in finance: pricing financial derivatives*. Wiley series in probability and statistics. Wiley, Chichester, 2003.

[ScCa09] Schoutens, W., and Cariboni, J. *Lévy processes in credit risk*. Wiley finance series. Wiley, Chichester, 2009.

[Shar64] Sharpe, W. F. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance 19*, 3 (1964), 425–442.

[Shep91a] Shephard, N. G. From characteristic function to distribution function: a simple framework for the theory. *Economic Theory 7* (1991), 519–529.

[Shep91b] Shephard, N. G. Numerical integration rules for multivariate inversions. *Journal of Statistical Computation and Simulation 39*, 1 (1991), 37–46.

[UpSh79] Upton, D. E., and Shannon, D. S. The stable Paretian distribution, subordinated stochastic processes, and asymptotic lognormality: An empirical investigation. *The Journal of Finance 34*, 4 (September 1979), 1031–1039.

[Ziel01] Zieliński, R. High-accuracy evaluation of the cumulative distribution function of α-stable symmetric distributions. *Journal of Mathematical Sciences 105*, 6 (2001), 2630–2632.

[Zolo86] Zolotarev, V. M. *One-dimensional stable distributions*, vol. 65 of *Translations of mathematical monographs*. American Mathematical Society, Providence, RI, 1986.

Appendix A

Lévy Processes

A.1 Definitions

General Lévy processes in contrast to the ubiquitous Brownian motions are considered to be the more advanced types in the area of continuous-time stochastic processes. Although their representation sometimes appear rather involved and discouraging, every single instance contained in this class is constructed by implementing the same simple principle, which namely is the serial independence and identical distribution of increments.

Following the presentation in Sato [Sato99, p. 3 f.], a possibly multivariate Lévy process $X_{(t),t>0}$ on $\mathbb{R}^n$ is defined by the conditions

- $X_0 = 0$ P-almost surely.
- Independent increments: For every strictly increasing sequence $0 \leq t_0 < t_1 < \ldots < t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.
- Stationary increments: The law of $X_{t+h} - X_t$ does not depend on t and is the same as the law of X_h. This is also known as *temporal homogeneity*.
- Stochastic continuity: $\forall \varepsilon > 0, \lim_{h \to 0} P(|X_{t+h} - X_t| \geq \varepsilon) = 0$
  Stochastic continuity should not be mistaken, however, with the continuity of process paths.

- Although not imposed by every author, a common additional condition is that trajectories are *càdlàg*[1]. It can be shown that every Lévy process has a unique modification which then possesses the càdlàg property. In other words, there is $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$ such that, for every $\omega \in \Omega_0$, $X_t(\omega)$ is right-continuous in $t \geq 0$ and has left limits in $t > 0$.

As this kind of definition via conditions required to be met is not very illustrative at all, one resorted to a more constructive perspective by means of the yet to follow Lévy-Itō decomposition. With the help of this decomposition one is able to explicitly state the entire set of Lévy processes according to the above conditions.

An extensive treatment of the basic theory on Lévy processes this section draws upon, can be found in the textbooks of Sato [Sato99], Bertoin [Bert05] or Applebaum [Appl05]. Books which already have an orientation towards applications of various kinds are the ones recently written by Schoutens [Scho03], Schoutens and Cariboni [ScCa09] as well as the book from Kyprianou [Kypr06].

Only the aspects most relevant to the course of this work are explicitly recapitulated at this point. First, it should be emphasized that a general multivariate Lévy process $X_{(t),t>0}$ is a superimposition of three independent components:

- Diffusion component: Arithmetic Brownian motion without trend and a covariance matrix A controlling the diffusion behavior.
- External trend: Their trajectories are simply $\mathbb{R}^n$-valued linear functions of time t which are specified by a vector $m \in \mathbb{R}^n$.
- Jump component: The most central but complex item in the aggregation. Its jump behavior is determined by a random measure[2] ν on $\mathbb{R}^n \backslash \{0\}$ of Radon type which satisfies the two conditions

$$\int_{|x|\leq 1} |x|^2 \nu(dx) < \infty \quad , \qquad \int_{|x|\geq 1} \nu(dx) < \infty \, . \tag{A.1}$$

[1] French: continue à droite, limitée à gauche; right-continuous with left limits

[2] For details on how random measures of Poisson and general type determine a pure jump process, see Applebaum [Appl05, p. 103].

Therefore, each Lévy process can be uniquely represented by its associated Lévy triplet (A, ν, m) where each item in the triplet specifies one of the above components.

Conditions similar to those in equation (A.1) characterize different types of involved jump components.

1.) finite variation: $\int_{|x|\leq 1} |x|\nu(dx) < \infty$

2.) finite activity: $\int_{|x|\leq 1} \nu(dx) < \infty$

3.) existing expectation: $\int_{\mathbb{R}^n} |x|\nu(dx) < \infty$

The characteristic function $\phi_{X_t}(z)$ of the random process vector X_t of a multivariate Lévy process $X_{(t),t>0}$ with Lévy triplet (A, ν, m) at time t is determined by the *Lévy-Khintchine formula*[3]

$$\begin{aligned}\phi_{X_t}(z) &= E(\exp(iz^\mathsf{T} X_t)) \\ &= \exp\left[itm^\mathsf{T} z - \frac{t}{2} z^\mathsf{T}\Sigma z + t\int_{\mathbb{R}^n\setminus\{0\}} \left(\exp(iz^\mathsf{T} x) - 1 - iz^\mathsf{T} \underbrace{x\mathbb{1}_{|x|<1}}_{=:\vartheta(x)}\right)\nu(dx)\right].\end{aligned} \tag{A.2}$$

Due to its generation by a Lévy process with accordingly defined increments, the distribution of X_t is moreover *infinitely divisible* (ID). This constitutes a universal relation as every Lévy process is linked to one particular ID distribution and vice versa. Note that for general Lévy jump measures fulfilling the condition in equation (A.1) but possessing infinite variation, the integral in the Lévy-Khintchine formula in equation (A.2) would not exist without the *compensation term* $\vartheta(x) = x\mathbb{1}_{|x|<1}$. Its meaning becomes clear when considering the *Lévy-Itō decomposition*[4]. The central result is a formal decomposition rather similar to the qualitative motivation in enumeration on p. 162. Accordingly, a Lévy

[3] See e.g. Bertoin [Bert05, p. 13 f.].

[4] See, for instance, Applebaum [Appl05, p. 112 ff.] and Kyprianou [Kypr06, p. 33 ff.].

process $X_{(t)}$ with triplet (A, ν, m) can be divided into the following three independent sub-processes: $X_{(t)} = X^{(1)}_{(t)} + X^{(2)}_{(t)} + X^{(3)}_{(t)}$

- $X^{(1)}_{(t)}$: Arithmetic Brownian motion with drift m and covariance structure of the diffusion parameter contained in matrix A.
- $X^{(2)}_{(t)}$: Compound Poisson process determined by the measure ν on the domain of 'large' jumps $\{x \in \mathbb{R}^n : |x| \geq 1\}$.
- $X^{(3)}_{(t)}$: Square-integrable pure jump martingale with an almost surely countable number of jumps on each finite time interval which are of magnitude less than 1 termed 'small' jumps. This martingale is generated by the assumption of ν on domain $\{x \in \mathbb{R}^n : |x| < 1\}$ is driving a compensated[5] jump process.

Finally, the compensation term in eq. (A.2) is a direct result of compensated nature of $X^{(3)}_{(t)}$. The following arguments play a crucial role for different ways of representing particular types of Lévy processes. First note that the expression in the Lévy-Khintchine formula and the associated Lévy-Itō decomposition exist for any kind of Lévy processes whose measure ν has to satisfy the existence conditions in (A.1). If, however, the Lévy process meets one of the additional conditions below, one can switch to a representation of simpler fashion compared to the general case:

i) Bounded variation: $\int_{|x|\leq 1} |x|\nu(dx) < \infty$

Compensation is not essential for the process to exist, not even for small jumps of magnitude less than one. Since the compensation of small jumps[6] driven by ν would by its regularity merely create an existing deterministic linear trend with bounded gradient. This linear trend could also be separated from the process of small jumps and incorporated in the previous trend component. Consequently, when ν implies finite variation, the process composed by the triplet $(A, \nu, m^{(1)})$ under conventional compensation of small jumps is equivalent to a process with $(A, \nu, m^{(0)})$, where ν

[5] Compensation of Poisson and general random measures is defined and explained e.g. in Cont and Tankov [CoTa04, p. 59 f.].

[6] This is termed as the usual compensation convention as it is universally applicable.

controls an uncompensated jump process encompassing the entire spectrum of jump sizes. This jump process exists without any required compensation. Note the fact that ultimately $m^{(0)}$ denotes the case where ν is treated without compensation. The following equivalence to the usual and universally admissible convention characterized by $m^{(1)}$ and the compensation of small jumps holds

$$\vartheta^{(0)}(x) = 0 \quad \text{(degenerated compensation function)}$$

$$m^{(0)} = m^{(1)} - \int_{|x|<1} iz^{\mathsf{T}} x \nu(dx)\,,$$

where the letter equation reflects the necessary external trend when it is switched to the convention of uncompensated jumps.

ii) Finite first moment: $E(X_t) < \infty \; \forall\, t > 0 \Leftrightarrow \int_{\mathbb{R}^n \setminus \{0\}} |x| \nu(dx) < \infty$
By the characteristics of ν it would even be admissible to compensate the entire jump spectrum, corresponding to the compensation function for the Lévy-Khintchine formula

$$\vartheta^{(2)}(x) = x\,.$$

This convention also leads to an existing internal compensation trend with bounded gradient vector while the external trend denoted by $m^{(2)}$ again indicates the particular type of compensation convention. Hence, by similar arguments as in i.), a random measure ν satisfying above condition generates two equivalent Lévy processes for $(A, \nu, m^{(1)})$ and $(A, \nu, m^{(2)})$, if and only if

$$m^{(2)} = m^{(1)} + \int_{|x|\geq 1} iz^{\mathsf{T}} x \nu(dx)\,.$$

This case of $|x|$-integrable ν constitutes a convenient subcase as the admissible $m^{(2)}$ is directly reflecting the entire drift present in the process. The jump component became a martingale by the applied full compensation, so $m^{(2)}$ is solely responsible for the evolution of the expectation $E(X_t)$ of $X_{(t)}$ over time t.

By these explanations it was illustrated how different compensation and representation conventions can be equivalently converted into each other when certain existence and regularity requirements are met, and, in consequence, how these conversions may lead to favorable representations of the situation at hand. In particular $m^{(2)}$ has the advantage of containing the complete information about the mean, whereas $m^{(0)}$ with $\vartheta(x) = 0$ is very often able to contribute an expression of simple analytical tractability in the Lévy-Khinchine formula (A.2).

A.2 Stable Processes

The class of α-stable processes entirely comprise the set of self-similar Lévy processes [7]. Besides these, only Brownian motions possess the self-similarity property for Lévy processes with corresponding similarity index of $H = 1/2$ in particular. The α-stable process $X_{(t),t\geq 0}$ in the univariate case is characterized by its purely non-Gaussian character reflected in the full Lévy triplet $(0, \nu_X, m_X^{(1)})$ with

$$\nu_X(x) = \frac{C_+}{x^{1+\alpha}}\mathbb{1}_{x>0} + \frac{C_-}{|x|^{1+\alpha}}\mathbb{1}_{x<0}$$

and $\alpha \in (0,2)$, $C_+, C_- > 0$, exhibiting polynomial decay[8]. Processing the Lévy triplet and the Lévy triplet in particular with the Lévy-Khintchine formula from equation (A.2) results in the characteristic function of the *α-stable random variable*

$$\phi_{X_t}(z) = \exp\left(izt m + t\int_{\mathbb{R}\setminus\{0\}} \left(\exp(-izx) - 1 - izx\mathbb{1}_{|x|<1}\right)\left[\frac{C_+}{x^{1+\alpha}}\mathbb{1}_{x>0} + \frac{C_-}{|x|^{1+\alpha}}\mathbb{1}_{x<0}\right] dx \right) \tag{A.3}$$

$$= \begin{cases} \exp\left(i\mu z t - t|\sigma z|^\alpha\left(1 - i\beta\,\mathrm{sgn}(z)\tan\frac{\pi\alpha}{2}\right)\right)\ ,\ \alpha \neq 1 \\ \exp\left(i\mu z t - t\sigma|z|\left(1 + i\beta\frac{2}{\pi}\mathrm{sgn}(z)\ln|z|\right)\right)\ ,\ \alpha = 1 \end{cases} \tag{A.4}$$

[7] For a proof see Samorodnitsky and Taqqu [SaTa00, p. 351 f.].
[8] See for example Rachev et al. [KRBF10].

which can be found in Rachev and Mittnik [RaMi00]. Note that it proves to be very useful to switch from the original parametrization in terms of (α, C_+, C_-, m) to the one employed in equation (A.4) using $(\alpha, \beta, \sigma, \mu)$, $\beta \in (-1, 1)$ for the level of asymmetry, $\sigma > 0$ for the scale and $\mu \in \mathbb{R}$ for the location of the distribution. The correspondence of both parameterizations is derived in Sato [Sato99, p. 80 ff.] for the general multivariate case and is furthermore illustrated by Nolan [Nola12, ch. 3]. The distinction of the cases $\alpha \neq 1$ and $\alpha = 1$ is finally caused by a series representation necessary for the integration in equation (A.3), where the α-stable Lévy density is in square brackets. This fact is stressed here since one encounters the same structures and implications when later turning to CTS and NTS.

A.3 Characteristic Function of the CTS Subordinator

This section is intended to give an outline of the integration techniques involved with the transition from a CTS-Lévy density to its characteristic function by means of the Lévy-Khintchine formula. The actual benefit of this demonstration will turn out to be twofold:

- The analytic solution of the integral relies on a series representation of the exponential function and a subsequent aggregation of the outcome in a binomial series. This procedure will be reused for both the univariate as well as the multivariate NTS process.
- The results for the CTS are required to determine the corresponding input of the CTS subordinator for insertion into the subordination formulas.

The following calculations are guided by the directions given in Kim [Kim05, p. 18 ff.] and Cont and Tankov [CoTa04, p. 121 f.]. Since a CTS Lévy density function $\nu_{CTS}(x)$ with parameters $(\alpha, C_+, C_-, \lambda_+, \lambda_-, m)$ satisfies the condition

$$\int_{|x|\geq 1} |x| \nu_{CTS}(x) dx < \infty \,,$$

full compensation of the jump component, linked to the comprehension of m as external drift coefficient $m^{(2)}$, is admissible. This requires the solutions of integrals of the form

$$I = \int_0^\infty (\exp(izx) - 1 - izx)\nu_{CTS}(x) = \int_0^\infty (\exp(izx) - 1 - izx)\frac{\exp(-\lambda x)}{x^{1+\alpha}}dx$$

in the Lévy-Khintchine formula. The Taylor series expansion of the first exponential function yields

$$I = \int_0^\infty \left(\sum_{n=0}^\infty \frac{(izx)^n}{n!} - 1 - izx\right)(\exp(-\lambda x)x^{-\alpha-1})dx\,.$$

After eliminating the first two series terms

$$I = \int_0^\infty \left(\sum_{n=2}^\infty \frac{(izx)^n}{n!}\right)(\exp(-\lambda x)x^{-\alpha-1})dx\,,$$

one is able to interchange summation und integration under some weak regularity condition, which is satisfied here.

$$I = \sum_{n=2}^\infty \frac{(iz)^n}{n!}\int_0^\infty \exp -\lambda x x^{n-\alpha-1}dx$$

With the substitution $\lambda x = t \Rightarrow x = \frac{1}{\lambda}t\,,\ dt = \lambda dx \Rightarrow dx = \frac{1}{\lambda}dt$ and the integral representation of the Gamma function

$$\Gamma(1-\alpha) = \int_0^\infty \frac{1}{x^\alpha}\exp(-x)dx = \int_0^\infty x^{-\alpha}\exp(-x)dx$$

one ultimately obtains

$$
\begin{aligned}
I &= \sum_{n=2}^{\infty} \frac{(iz)^n}{n!} \int_0^{\infty} \exp -t(\frac{1}{\lambda}t)^{n-\alpha-1}\frac{1}{\lambda}dt \\
&= \sum_{n=2}^{\infty} \frac{(iz)^n}{n!}\lambda^{\alpha-n} \int_0^{\infty} t^{n-\alpha-1}\exp(-t)dt \\
&= \sum_{n=2}^{\infty} \frac{(iz)^n}{n!}\lambda^{\alpha-n}\Gamma(n-\alpha)\,,
\end{aligned}
$$

where by the functional equation $x\cdot\Gamma(x) = \Gamma(1+x)$ for the Gamma function the following power series becomes clearly recognizable

$$
\begin{aligned}
I = \lambda^{\alpha}\Gamma(2-\alpha)\Bigg[&\frac{1}{2!}\left(\frac{iz}{\lambda}\right)^2 + \frac{(2-\alpha)}{3!}\left(\frac{iz}{\lambda}\right)^3 + \frac{(2-\alpha)(3-\alpha)}{4!}\left(\frac{iz}{\lambda}\right)^4 \\
&+ \frac{(2-\alpha)(3-\alpha)(4-\alpha)}{5!}\left(\frac{iz}{\lambda}\right)^5 + \frac{(2-\alpha)(3-\alpha)(4-\alpha)(5-\alpha)}{6!}\left(\frac{iz}{\lambda}\right)^6 + \ \ldots\ \Bigg]
\end{aligned}
$$

For $\alpha \in \mathbb{R}\backslash\{0,1\}$, the bracketed expression resembles the power series

$$
(1+x)^{\alpha} = \sum_{n=0}^{\infty}\binom{\alpha}{n}x^n\,,\ \alpha, x \in \mathbb{C}\backslash\mathbb{N}_0 \quad \text{(binomial series)}
$$

Hence, after performing the transformation step

$$
\begin{aligned}
I = \lambda^{\alpha}\frac{\Gamma(2-\alpha)}{\alpha(\alpha-1)}\Bigg[&\frac{1}{2!}\left(-\frac{iz}{\lambda}\right)^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}\left(-\frac{iz}{\lambda}\right)^3 + \frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}{4!}\left(-\frac{iz}{\lambda}\right)^4 \\
&+ \frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)}{5!}\left(-\frac{iz}{\lambda}\right)^5 + \ \ldots\ \Bigg] \qquad \text{(A.5)}
\end{aligned}
$$

a binomial series from the third term onwards appears

$$I = \lambda^\alpha \Gamma(-\alpha) \left[\left(1 - \frac{iz}{\lambda}\right)^\alpha - 1 - \alpha \left(-\frac{iz}{\lambda}\right) \right] = \Gamma(-\alpha) \left[(\lambda - iz)^\alpha - \lambda^\alpha + iz\alpha\lambda^{\alpha-1} \right] .$$

As indicated above, the two particular cases $\alpha = 0$ and $\alpha = 1$ cannot be captured with the binomial series and must be considered separately from the expression in equation (A.5) on:

- $\alpha = 1$:

$$I = \lambda \sum_{n=2}^{\infty} \frac{1}{n!} \left(\frac{iz}{\lambda}\right)^n \Gamma(n-1) = \lambda \sum_{n=2}^{\infty} \frac{1}{n!} \left(\frac{iz}{\lambda}\right)^n (n-2)!$$

$$= \lambda \sum_{n=2}^{\infty} \left(\frac{iz}{\lambda}\right)^n \frac{(n-2)!}{n!} = \lambda \sum_{n=2}^{\infty} \left(\frac{iz}{\lambda}\right)^n \frac{1}{n(n-1)}$$

  For the sake of clarity and compact display, the simple substitution $u = \frac{iz}{\lambda}$ with its associated derivatives of the sum term is considered.

$$f(u) := \sum_{n=2}^{\infty} u^n \frac{1}{n(n-1)}$$

$$f'(u) = \sum_{n=2}^{\infty} u^{n-1} \frac{1}{n-1}$$

$$f''(u) = \sum_{n=2}^{\infty} u^{n-2} = \sum_{n=0}^{\infty} u^n = \frac{1}{1-u}$$

  Note that the convergence radius of the last series is not of relevance. Integration yields

$$f'(x) = \int f''(x)dx = \int \frac{1}{1-x} dx = -\ln(1-x) + C_1 .$$

The substitution $y = 1 - u$ once again enables the reversal of the chain rule:

$$f(x) = \int f'(x)dx = \int -\ln(1-x) + C_1 dx = (1-x)\ln(1-x) - (1-x) + C_1 x + C_2$$

which draws on the familiar relation

$$\int \ln(y)dy = y\ln(y) - y\,.$$

This solution of the differential equation of second order is adapted to its boundary conditions by the technique of coefficient comparison:

$$f'(x=0) = 0 \stackrel{!}{=} C_1 \Rightarrow C_1 = 0$$

$$f(x=0) = 0 \stackrel{!}{=} C_2 - 1 \Rightarrow C_2 = 1\,.$$

This results in

$$f(x) = (1-x)\ln(1-x) + x$$

which finally yields

$$I = \lambda f\left(\frac{iz}{\lambda}\right) = (\lambda - iz)\ln\left(1 - \frac{iz}{\lambda}\right) + iz\,.$$

Although the involved calculations are not absolutely identical, this gives an impression why the expression of the α-stable characteristic function is different for the particular case of $\alpha = 1$ only.

- $\alpha = 0$: The required calculations in this case exhibit a rather similar structure and suitable operations as for the case $\alpha = 1$ above. Therefore, the explicit derivation is omitted here as $\alpha = 0$ is not included in the parameter range of the CTS anyway. Nevertheless, one should note that the value of $\alpha = 0$ is characteristic for the Gamma process, which constitutes a central instance of Lévy processes. For example, it serves as the subordinator for the Brownian motion in the construction of the *Variance Gamma (VG)* process.

Appendix B

Brownian Subordination

B.1 Univariate Subordination

The detailed proof of equation (3.17) follows the outline of Cont and Tankov given in [CoTa04, p. 118].

$$\nu_X(x) = \int_0^\infty \underbrace{\frac{1}{\sqrt{2\pi}\sigma\sqrt{s}} \exp\left(-\frac{(x-\mu s)^2}{2\sigma^2 s}\right)}_{\text{time-dependent normal density}} \cdot \underbrace{\frac{\lambda^{1-\frac{a}{2}}}{\Gamma(1-\frac{a}{2})} \frac{\exp(-\lambda s)}{s^{\frac{a}{2}+1}}}_{\text{Lévy density of subordinator}} \; ds$$

$$= \frac{\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})} \int_0^\infty \exp\left(-\frac{(x-\mu s)^2}{2\sigma^2 s} - \lambda s\right) s^{-(\frac{3}{2}+\frac{a}{2})} ds$$

$$= \frac{\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})} \int_0^\infty \exp\left(\frac{x^2 + 2x\mu s - \mu^2\sigma^2}{2\sigma^2 s} - \lambda s\right) s^{-(\frac{3}{2}+\frac{a}{2})} ds$$

$$= \frac{\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})} \exp\left(\frac{x\mu}{\sigma^2}\right) \int_0^\infty \exp\left(-\left(\frac{\mu^2}{2\sigma^2} + \lambda\right) s - \frac{x^2}{2\sigma^2} \cdot \frac{1}{s}\right) s^{-(\frac{a}{2}+\frac{3}{2})} ds \,.$$

Integral representation of the modified Bessel function of the second kind of order p:

$$K_p(y) = 2^{-(p+1)} y^p \int_0^\infty \exp(-t - \frac{y^2}{4t}) t^{-(p+1)} dt \Leftrightarrow \int_0^\infty \exp(-t - \frac{y^2}{4t}) t^{-(p+1)} dt = 2^{p+1} y^{-p} K_p(y)$$

Inserting the suitable substitution

$$t = \left(\frac{\mu^2}{2\sigma^2}+\lambda\right)s \Rightarrow s = \frac{t}{\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)} \Rightarrow \frac{1}{s} = \frac{\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)}{t}$$

$$dt = \left(\frac{\mu^2}{2\sigma^2}+\lambda\right)ds \Rightarrow ds = \frac{dt}{\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)}$$

yields

$$\begin{aligned}\nu_X(x) &= \int_0^\infty \exp\left(-t - \frac{x^2}{2\sigma^2}\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)\frac{1}{t}\right)\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)^{\frac{a}{2}+\frac{3}{2}} t^{-(\frac{a}{2}+\frac{3}{2})}\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)^{-1} dt \\ &\quad \cdot \frac{\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})}\exp\left(\frac{x\mu}{\sigma^2}\right) \\ &= \int_0^\infty \exp\left(-t - \frac{x^2\mu^2+2\sigma^2x^2\lambda}{4\sigma^4}\frac{1}{t}\right) t^{-(\frac{a}{2}+\frac{3}{2})} dt \\ &\quad \cdot \frac{\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})}\exp\left(\frac{x\mu}{\sigma^2}\right)\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)^{\frac{a+1}{2}}.\end{aligned}$$

Resolving the resemblancies in the above expressions

$$p+1 = \frac{a}{2}+\frac{3}{2} \Rightarrow p = \frac{a+1}{2}$$

$$y^2 = \frac{x^2\mu^2+2\sigma^2x^2\lambda}{\sigma^4} \Rightarrow y = \frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}$$

subsequently leads to

$$\begin{aligned}\nu_X(x) &= \frac{\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})}\exp\left(\frac{x\mu}{\sigma^2}\right)\left(\frac{\mu^2}{2\sigma^2}+\lambda\right)^{\frac{a+1}{2}} 2^{\frac{a}{2}+\frac{3}{2}}\left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right)^{-\frac{a+1}{2}} \\ &\quad \cdot K_{\frac{a+1}{2}}\left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right)\end{aligned}$$

$$= \frac{2\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})} \exp\left(\frac{x\mu}{\sigma^2}\right) \left(\frac{\mu^2+2\sigma^2\lambda}{\sigma^2}\right)^{\frac{a+1}{2}} \left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right)^{-\frac{a+1}{2}}$$

$$\cdot K_{\frac{a+1}{2}}\left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right)$$

$$= \frac{2\lambda^{1-\frac{a}{2}}}{\sqrt{2\pi}\sigma\Gamma(1-\frac{a}{2})} \exp\left(\frac{x\mu}{\sigma^2}\right) \left(\frac{\sqrt{\mu^2+2\sigma^2\lambda}}{|x|}\right)^{\frac{a+1}{2}} K_{\frac{a+1}{2}}\left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right).$$

Aggregating the remaining terms gives the final result

$$\nu_X(x) = \frac{\sqrt{2}\lambda^{1-\frac{a}{2}}\left(\mu^2+2\sigma^2\lambda\right)^{\frac{a+1}{4}}}{\sqrt{\pi}\sigma\Gamma(1-\frac{a}{2})} \exp\left(\frac{x\mu}{\sigma^2}\right) \frac{K_{\frac{a+1}{2}}\left(\frac{|x|\sqrt{\mu^2+2\sigma^2\lambda}}{\sigma^2}\right)}{|x|^{\frac{a+1}{2}}}.$$

Barndorff-Nielsen and Shephard are discussing a similar approach in [BNSh02] including an explicit solution on p. 9, whereas the original publication of Barndorff-Nielsen and Levendorskiĭ [BNLe01, p. 7 ff.] for univariate NTS processes only contains the characteristic exponent of the process law instead of presenting its entire Lévy triplet. Nevertheless, Barndorff-Nielsen and Levendorskiĭ demonstrate another more involved determination procedure for this characteristic exponent than the simple subordination formula in equations (3.13) and (3.15).

B.2 Inverted Parameter Conversion of MNTS Class

In this section it will be explicitly shown that every particular instance of the MNTS distribution with parameters $(a, \lambda, \gamma, \sigma, \mathsf{P}, \mu)$ has an corresponding formulation in terms of the MVNTS distribution with parameter tuple $(a^*, \lambda^*, \gamma^*, \mathsf{P}^*, c^*, d^*)$. This is achieved by inverting the parameter correspondence for the opposite direction given in section 4.2.2.

Recall from equations (4.28) – (4.30) that

$$a = a^*,\ \lambda = \lambda^*,\ \mathsf{P} = \mathsf{P}^*,\ \mu = d^*$$

$$\gamma \;=\; [c_k^* \cdot \gamma_k^*]_{k=1,\ldots,n}^{\mathsf{T}} \tag{B.1}$$

$$\sigma \;=\; \left[c^* \cdot \sqrt{1 - \frac{1}{\lambda^*}\left(1 - \frac{a}{2}\right)(\gamma_k^*)^2}\right]_{k=1,\ldots,n}^{\mathsf{T}} . \tag{B.2}$$

Inverting the relations in the first row is trivial and directly yields

$$a^* = a\,,\ \lambda^* = \lambda\,,\ \mathsf{P}^* = \mathsf{P}\,,\ d^* = \mu\,.$$

The remaining parameter values have to be simultaneously resolved from the other two rows in the above equations (B.1) and (B.2) which can be achieved by performing the following two steps:

- As c^* is directly linked to the vector of standard deviations inherent in the MNTS, one can immediately conclude that

$$c^* \;=\; \left[\sqrt{\frac{2-a}{2\lambda}\gamma_k^2 + \sigma_k^2}\right]_{k=1,\ldots,n}^{\mathsf{T}} = \left[\sqrt{\frac{(2-a)\gamma_k^2 + 2\lambda\sigma_k^2}{2\lambda}}\right]_{k=1,\ldots,n}^{\mathsf{T}} ,$$

  where the configuration of the last expression is useful for the subsequent verification of the validity condition on the MVNTS parameter set.

- Moreover, by the relation of the distinct values and meanings of γ in the MVNTS and the MNTS distribution via c, one is able to deduce

$$\gamma^* \;=\; \left[\gamma_k \cdot (c_k^*)^{-1}\right]_{k=1,\ldots,n}^{\mathsf{T}} = \left[\gamma_k \cdot \sqrt{\frac{2\lambda}{(2-a)\gamma_k^2 + 2\lambda\sigma_k^2}}\right]_{k=1,\ldots,n}^{\mathsf{T}} .$$

  In order to verify the validity of γ^* created by the above equation it is sufficient to check

$$(\gamma^*)^2 \;=\; \frac{2\lambda\gamma^2}{(2-a)\gamma^2 + 2\lambda\sigma^2} = \left(\frac{(2-a)\gamma^2 + 2\lambda\sigma^2}{2\lambda\gamma^2}\right)^{-1}$$

$$= \left(\frac{(2-a)}{2\lambda} + \frac{\sigma^2}{\gamma^2}\right)^{-1} \leq \left(\frac{(2-a)}{2\lambda}\right)^{-1} .$$

Appendix C

Optimization Problems

C.1 Dow Jones Industrial Average Weights

Table C.1 presents the component weighting of stocks contained in the Dow Jones Industrial Average (DJIA) index as of October 25, 2010.

Ticker symbol	weight in %
MMM	6.163
AA	0.873
AXP	2.664
T	1.923
BAC	0.757
BA	4.858
CAT	5.359
CVX	5.754
CSCO	1.601
KO	4.136
DD	3.234
XOM	4.488
GE	1.089
HPQ	2.907
HD	2.129

Ticker symbol	weight in %
INTC	1.347
IBM	9.481
JNJ	4.338
JPM	2.513
KFT	2.201
MCD	5.336
MRK	2.537
MSFT	1.708
PFE	1.195
PG	4.308
TRV	3.769
UTX	5.085
VZ	2.193
WMT	3.658
DIS	2.404

Table C.1: Component weighting of the DJIA index on October 25, 2010

C.2 Efficient Frontier

Efficient frontiers concerning $VaR_\alpha(R_p)$ and $AVaR_\alpha(R_p)$ of a portfolio return R_p under a fixed level of α is obtained by solving the respective optimization problems for a series of minimal expected returns $\mu_k^{(b)}$, $k = 1, 2, \ldots, m$.

$$\min_{\omega \in \mathbb{R}^n} VaR_\alpha(R_p) \qquad \min_{\omega \in \mathbb{R}^n} AVaR_\alpha(R_p)$$

$$s.t. \qquad s.t.$$

$$R_p = \omega^\mathsf{T} R = \sum_{j=1}^{n} \omega_j R_j \qquad R_p = \omega^\mathsf{T} R = \sum_{j=1}^{n} \omega_j R_j$$

$$E(R_p) \geq \mu_k^{(b)} \qquad E(R_p) \geq \mu_k^{(b)}$$

$$\sum_{j=1}^{n} \omega_j = 1 \qquad \sum_{j=1}^{n} \omega_j = 1$$

$$\omega_j \geq -C\,,\ j = 1, \ldots, n \qquad \omega_j \geq -C\,,\ j = 1, \ldots, n$$

The constraint in the last row of each optimization problem represents a particular short selling restriction. None of the involved stocks are allowed to be shortened by an amount of more than $C \cdot 100\%$ of the total portfolio capital invested. For the optimizations carried out in the text, this restriction has not been active however.

Furthermore, there are two distinct assumptions for the distribution of joint stock returns R:

1.) Normal model: $R = (R_1, \ldots, R_n)^\mathsf{T} \sim \mathcal{N}^n(\mu, \Sigma)$

2.) Normal Tempered Stable model: $R = (R_1, \ldots, R_n)^\mathsf{T} \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$

C.3 Global Stationary Portfolio Optimization

A unique optimal static allocation is derived by solving the following optimization problem.

$$\max_{\omega \in \mathbb{R}^n} \frac{E(R_p - R_b)}{\Upsilon(R_p - R_b)}$$

$$s.t.$$

$$R_p = \omega^{\mathsf{T}} R = \sum_{j=1}^{n} \omega_j R_j$$

$$\sum_{j=1}^{n} \omega_j = 1$$

$$\omega_j \geq -C$$

The optimization problem is varied by alternative choices of the operator $\Upsilon(R)$ in the denominator of the objective function, which is nothing but the performance ratio. The different choices of $\Upsilon(R)$ comprise

- $\Upsilon(R) = Std(R) = \sqrt{Var(R)} \Rightarrow$ Sharpe-Ratio
- $\Upsilon(R) = VaR_\alpha(R) \Rightarrow$ VaR-Ratio
- $\Upsilon(R) = AVaR_\alpha(R) \Rightarrow$ STAR-Ratio

Each of the operators can be combined with the two model assumptions for the joint distribution of stock and index returns $(R^{\mathsf{T}}, R_b) = (R^{\mathsf{T}}, R^{(ind)})$

1.) Normal model: $(R^{\mathsf{T}}, R^{(ind)}) = (R_1, \ldots, R_{n+1}) \sim \mathcal{N}^n(\mu, \Sigma)$

2.) Normal Tempered Stable model:

$$(R^{\mathsf{T}}, R^{(ind)}) = (R_1, \ldots, R_{n+1}) \sim MVNTS(a, \lambda, \gamma, \mathsf{P}, c, d)$$

As was the case for the efficient frontiers no short-selling restriction ($C = \infty$) has been activated here.

Appendix D

Goodness of Fit Tests

After having estimated values for the parameters of a specified model, there are several techniques available for assessing the goodness of model fit onto the empirical data. This model fit diagnosis should also assist in identifying and locating possible misspecifications and deficiencies with regard to flexibility and adaptability. For addressing this task, graphical techniques for a visual inspection are recommended, such as kernel density plots and quantile-quantile plots which allow for a particular detection of potential weaknesses.

The methods for objectively testing statistical hypotheses concerning the nature of the data generating process the empirical observations were sampled from, presented in the following sections have in common that they, without exception, rely on the distance between the theoretical cumulative distribution function (CDF) and the empirical distribution function (EDF). Therefore, they are usually termed as *empirical distribution function-based tests*, where $F_n(x)$ denotes the EDF for sample size n and $F_0(x)$ indicates the theoretical equivalent of the null hypothesis. Their presentation is following Büning and Trenkler [BüTr94, p. 69 ff., p. 83 f.] and Chernobai et al. [ChRF07, p. 207 ff.]. Note however that this is solely done for the case of univariate distributions.

D.1 Kolmogorov-Smirnov Test

The statistics of the Kolmogorov-Smirnov (KS) test measures the largest vertical distance between $F_n(x)$ and $F_0(x)$. In other words, KS is represented mathematically as

$$KS = \sqrt{n} \sup_{x \in \mathbb{R}} |F_0(x) - F_n(x)| \; . \tag{D.1}$$

Although its distribution is independent of F_0 as long as F_0 is assumed to be continuous, it cannot be stated in closed form. Therefore, critical values for the test statistics as well as corresponding p–values have to be obtained via Monte Carlo simulation or alternative numerical approximation approaches whose results can be found in statistical tables as well as contained in commercial standard software packages. With the sample order statistics $x_{(j)}$ and consequentially $z_{(j)} = F_0(x_{(j)})$ the KS statistics can be directly obtained by

$$KS = \sqrt{n} \max \left(\sup_{j=1,\ldots,n} \left(\frac{j}{n} - z_{(j)} \right), \sup_{j=1,\ldots,n} \left(z_{(j)} - \frac{j-1}{n} \right) \right) \tag{D.2}$$

from the empirical data set $\{x_j\}_{j=1,\ldots,n}$.

D.2 Anderson-Darling Test

Like the KS test, Anderson-Darling (AD) tests are an alternative goodness of fit (GoF) test of supremum type. Their test statistic is formulated as

$$AD = \sqrt{n} \sup_{x \in \mathbb{R}} \left| \frac{F_0(x) - F_n(x)}{\sqrt{F_0(x)(1 - F_0(x))}} \right| , \tag{D.3}$$

with computing formula

$$AD = \sqrt{n} \max \left(\sup_{j=1,\ldots,n} \left(\frac{\frac{j}{n} - z_{(j)}}{\sqrt{z_{(j)}(1 - z_{(j)})}} \right), \sup_{j=1,\ldots,n} \left(\frac{z_{(j)} - \frac{j-1}{n}}{\sqrt{z_{(j)}(1 - z_{(j)})}} \right) \right) . \tag{D.4}$$

Another central test of supremum type in the literature is the Kuiper test whose explicit

presentation is omitted here. In comparison to the KS test, the AD test is employing an additional weighting function for the distances in order to emphasize deviations in the tails. This makes the AD test in particular suitable for testing hypotheses of heavy-tailedness, whereas the KS statistics, by its equal weighting, is foremost able to capture vertical deviations in the center of the distribution where discrepancies between EDF and CDF are most likely to have the largest spread.

D.3 Cramér-von Mises Test

With respect to the above supremum type GoF tests the general class of Cramér-von Mises (CM) tests implement a contrasting principle. They do not only measure a single supremum of the vertical distance between EDF and CDF but consider a weighted average of all squared vertical distances on $\mathbb{R}$

$$W_H^2 = n \int_{-\infty}^{\infty} (F_n(x) - F_0(x))^2 H(F_0(x)) dF_0(x)$$

containing a specific weighting function $H(t) \geq 0$.

Besides the entire class of CM tests, a particular Cramér-von Mises test exists, defined by the choice of $H(t) = 1$. In this case, the test statistic W^2 can be explicitly computed from empirical data by

$$W^2 = \frac{n}{3} + \frac{1}{n} \sum_{j=1}^{n} (1 - 2j) z_{(j)} + \sum_{j=1}^{n} z_{(j)}^2 \,.$$

The general Cramér-von Mises class of GoF tests forms the basis for another variant of Anderson-Darling tests described in the following section.

D.4 Squared Anderson-Darling Test

The choice of $H(t) = \frac{1}{t(t-1)}$, which aims at assigning significant weight to the tails, results in the *Squared Anderson-Darling* (AD^2) *test* with statistics

$$AD^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(x)(1 - F_0(x))} dF_0(x)$$

and its computational equivalent on empirical data

$$AD^2 = -n + \frac{1}{n} \sum_{j=1}^{n} (1 - 2j) \ln z_{(j)} - \frac{1}{n} \sum_{j=1}^{n} (1 + 2(n - j)) \ln(1 - z_{(j)}) \,.$$

The distributions of both AD and AD^2 only depend on the sample size n, where already Anderson and Darling in their original research [AnDa52] were able to derive their asymptotical distributions for $n \to \infty$. Crawford Moss et al. [CMTT90], based on the work of Lewis [Lewi61], provide tables for small sample sizes. Additionally, one has to point out the simulative approach of Marsaglia and Marsaglia [MaMa04], who kindly provide their software tool for computation, which processes arbitrary values of n. As a concluding remark, one has to note that by their specific design towards the screening of the tail properties, both Anderson-Darling tests belong to the most powerful tests for rejecting Normality in the presence of heavy-tailed data.

D.5 Critical Remarks

In practical applications as well as in many commercial software packages sightly modified definitions of KS and AD are widespread. This modification basically contains the omission of the leading factor $\sqrt{n}$ in equations (D.1) – (D.4), respectively. For reasons of consistency and comparability this convention is used throughout this study in chapters 5 and 6.

Nevertheless, a substantial drawback of this modification has to be mentioned at this point. The laws of given expressions in equation (D.2) and (D.4) converge for $n \to \infty$ to

a non-trivial distribution, respectively. For sufficiently large n, this limiting distribution can then be used for accurate approximations of p-values for KS and AD statistics. This important property vanishes for the modified versions without factor $\sqrt{n}$, however. Their laws converge to a meaningless Dirac measure in point 0 instead.