Lévy Processes in Finance: The Change of Measure and Non-Linear Dependence

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von Jens Wannenwetsch
aus Langenau (Württemberg)

Tag der Promotion: 17. Juni 2005

Dekan:	Prof. Dr. Matthias Kräkel
Erstreferent:	Prof. Dr. Klaus Sandmann
Zweitreferent:	Prof. Dr. Klaus Schürger
Tag der mündlichen Prüfung:	17. Juni 2005

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Meinen Eltern

Danksagung

Mein Dank geht zuallererst an meinen Betreuer Prof. Klaus Sandmann für seine fachliche und organisatorische Unterstützung während der letzten Jahre, die wesentlich zum Gelingen dieser Dissertation beigetragen hat. Ebenso gilt ein besonderer Dank Prof. Klaus Schürger für die kritische Durchsicht dieser Arbeit und seine hilfreichen Anmerkungen. Für die Möglichkeit, ein interessantes und gewinnbringendes Jahr am Centre de Mathématiques Appliquées der Ecole Polytechnique in Palaiseau bei Paris zu verbringen, bin ich Dr. Rama Cont sehr zu Dank verpflichtet. Mit ihm und Dr. Peter Tankov hatte ich dort die Möglichkeit, mit zwei Experten aus dem Bereich der Anwendung von Sprungprozessen in der Finanzmathematik zu diskutieren und zu arbeiten. Die Anregung, mich mit Lévyprozessen zu befassen, verdanke ich Prof. Sergio Albeverio.

Erwähnen möchte ich an dieser Stelle auch meine Weggefährten aus der Betriebswirtschaftlichen Abteilung III und der Statistischen Abteilung der Universität Bonn Sven Balder, Michael Brandl, An Chen, Haishi Huang, Simon Jäger, Dr. Antje Mahayni, Anne Ruston, Patrick Schonefeld, Xia Su, Michael Suchanecki und Manuel Wittke. Sie haben mir mit fachlichen Kommentaren, aber auch mit leidenschaftlich geführten Diskussionen über Themen wie Fußball, Stromkästen und Schrödingerprozesse während zahlreicher Raucherpausen interessante Einsichten verschafft. Insbesondere möchte ich Sven Balder für das Korrekturlesen meiner Arbeit danken. Nicht unerwähnt bleiben sollen auch meine Kommilitonen von der Bonn Graduate School of Economics, von denen ich stellvertretend Rudolf Maier und Dr. Rolf Tisljar erwähnen möchte, mit denen ich in der ersten Zeit im Doktorandenprogramm ein Büro geteilt habe.

Des weiteren gilt mein Dank Prof. Georg Nöldeke und Prof. Urs Schweizer in ihrer Eigenschaft als Leiter der Bonner Graduiertenausbildung, in deren organisatorischem Rahmen meine Promotion erfolgt ist, sowie der Deutschen Forschungsgemeinschaft und dem Marie-Curie-Programm der Europäischen Kommission für deren finanzielle Unterstützung.

Der größte Dank aber gebührt meinen Eltern, die mich während all der Jahre in jeder erdenklichen Hinsicht gefördert und mir so meine Ausbildung erst ermöglicht haben. Einen wesentlichen Anteil hat auch Silja Linnemann, die mich in allen Phasen meiner Promotion verständnisvoll begleitet hat und für deren Unterstützung ich sehr dankbar bin.

Contents

Introduction

The emergence of Lévy processes in the finance literature is due to empirical observations finding that the distribution of equity returns is in general skewed and leptokurtotic. One of the first to account for this phenomenon was Mandelbrot (1963) long before the seminal Black-Scholes model came up. Merton was first to introduce explicitly jumps into an option pricing model, and through this contribution he opened up research on jump-diffusion processes in finance. But ever since option pricing has completely and thoroughly been changed by the contributions of Black and Scholes (1973) and Merton (1973), people thought about generalizing the model in many ways in order to weaken the strong assumptions behind the Black-Scholes model. The Black-Scholes model assumes normally distributed independent and stationary log returns and is thus a Lévy process. Probably the most natural way to extend the model is to consider a more realistic distribution of the increments maintaining all the other Black-Scholes assumptions. And indeed, a Lévy process on a Euclidean space is constructed by giving a distribution on this space with the only property of infinite divisibility, and the theory of stochastic processes provides the technical tools to construct a Lévy process with independent and stationary increments distributed according to the given distribution. Although a Wiener process is a Lévy process in mathematical terms, in this thesis we use the term 'Lévy process' when we talk about processes with a more flexible distribution of the increments than a Gaussian distribution provides.
A Lévy process is a superposition of three independent components: a non-stochastic drift process, a Wiener process and a jump process. The most interesting one is the latter which is entirely captured by the Lévy measure. The Lévy measure determines the jump intensity of every jump size and is the main distinguishing feature for a rough subdivision into three classes of the zoo of Lévy processes used in finance: The only source of randomness of the first and the second class is the jump structure, the Wiener process being scaled identically to zero. The first class possesses an infinite Lévy measure, which means that it has an infinite expected number of jumps in every finite time interval, and infinite variation, i.e. the jumps add up to infinity. Though this seems strange at first, one has to be aware that a simple Wiener process is also an infinite variation process. The second class has finite variation while the mass of its Lévy measure is infinite as well. Third, quite differently from these two pure jump processes, there are jump-diffusion processes where the main movement comes from a non-zero Wiener process whereas the jumps occur rarely, i.e. finitely

often in a finite time interval. Accordingly, the mass of the Lévy measure is finite. We want to argue now that this subdivision, though quite meaningful in describing statistical features of the processes, is not helpful when it comes to deciding which one is best suitable in a financial context. First of all, apart from this subdivision, there is the widely discussed question of whether an equity price process should be modelled by a jump or a continuous process. The basic problem is that this question cannot be answered directly because we want to model a discretely observed price movement by a continuous time stochastic process. There are a number of authors who have conceived statistical tests to settle the question. But when the issue is pricing, this question becomes irrelevant in a certain sense. As explained above, there is a probability measure, namely the distribution of the increments, which is hidden behind any Lévy process, whether it be continuous or not. And the task is to estimate the parameters of this one-dimensional distribution. The question if the Lévy process constructed out of it is continuous is a mathematical one: Given an infinitely divisible probability distribution, the correct statement is that the assigned canonical stochastic process is not per se continuous but has possibly a continuous modification. This is true for a Gaussian distribution, but for distributions with heavier tails we can in general obtain only a right-continuous modification with left limits. Hence, working with a (realistic) model with stationary and independent increments simply implies the occurence of jumps, and we do not have to be concerned about whether jumps are a statistically adequate description of an equity price process or not. And this is all the more true because we deal with the risk-neutral distribution which cannot directly be observed anyhow.

Hence, if we have accepted the jump structure, the same argument as above does away with the need to choose the special jump structure according to our subdivision as long as the underlying probability distribution of the chosen Lévy process can incorporate realistic equity return features. More precisely, we mean by this that it should be able to fit the observed first four standardized moments of equity returns, namely mean, volatility, skewness and kurtosis. The latter is the most widely used measure for heavy-tailedness and is determined by the Lévy measure.

Given a set of realistic infinitely divisible return distributions, the most important feature of a Lévy model choice is tractability as the application of certain pricing methods is concerned. And tractability - and not the question about the jump structure - is the reason why we use a pure jump infinite variation process, namely the normal inverse Gaussian Lévy process, in the first chapter of the thesis and the tempered stable process in the second chapter. The tempered stable Lévy process is pure jump and can also have infinite variation, but we use it only for the case where it has finite variation. The third and parts of the fourth chapter are built around Kou's model, which is a jump-diffusion model.

For different types of Lévy processes different pricing techniques have to be applied. Hence a second subdivision along different lines may be more helpful than the first one. Lévy processes can be defined by the distribution of its increments, by a time-changed Brownian motion or by the so-called Lévy triplet: In this thesis we work exclusively with the third definition. The Lévy triplet is given by a deterministic drift coefficient, the scaling of the standard Brownian motion and the Lévy measure. From the Lévy triplet we obtain immediately a representation of the characteristic function of the Lévy process by means of the well-known Lévy-Khinchin theorem. With that information one can price options of European type through Fourier inversion: The price of an option is then given by an inverse Fourier integral of the product of the Fourier transform of the contract to be priced and the characteristic function of the Lévy process. This integral can be numerically computed by the technique of the Fast Fourier transform. Whenever one has the factors of this product in a convenient closed form, pricing by Fourier inversion and Fast Fourier transform is strongly recommendable.

As already mentioned, the use of concrete Lévy processes is propagated in all four main chapters of this thesis, and we will use them to solve three different problems in this framework. In each chapter we choose the model which we believe to be the most suitable for each purpose. The pricing procedure depends much on the choice of the model, and in view of immediate applicability we stress concreteness versus abstractness. Each model is immediately applicable and suitable for pricing by Fourier inversion.

Up to now, we have tacitly assumed that we want to model the (continuously compounded) returns of an equity by a Lévy process. Translated into prices, this means that we work throughout the thesis with the exponential Lévy model, in which prices are defined by the exponential of a Lévy process. The interest rate process is given by a riskless savings account. This model has several advantages which will be elaborated on in due course.

Until this point we have had a discussion of how to model equity returns. Pricing is done by evaluating the discounted expected value of an option given the probability distribution of the equity at the maturity date of the contract. But arbitrage-free valuation means that the expectation is not taken under the statistical distribution of the assets but under a risk-neutral distribution. Risk-neutral means that seen under this measure the equity price process must be a martingale, that is a fair game with zero expected profits. Talking about *a* and not about *the* risk-neutral distribution highlights the feature which is by far the most important one in modelling asset returns by Lévy processes: Lévy markets are incomplete. This means in terms of pricing that the martingale measure is not unique and in terms of hedging that it is not possible to track entirely the price process of a derivative security by a hedging portfolio which consists of the underlying equity and a riskless bond. For brevity we say that a probability measure $\mathbf{P}$ is always the statistical one, and by $\mathbf{Q}$ we will understand a risk-neutral martingale measure. $\mathbf{P}$ can always be obtained by standard statistical estimation techniques from historical equity return data. This is different for the martingale measure $\mathbf{Q}$ because it is not directly observable. Basically, it is just a pricing rule. As such it is implicit in market option prices observed in option exchanges, i.e. the options market chooses in some way the risk-neutral martingale measure. In order to retrieve it, one has to solve an inverse problem, which is not at all a standard problem in a Lévy setting.

Here we are at the first topic of the thesis: the *change of measure*. Once being aware that there is quite a large number of possible martingale measures, one has to set up rules of how to select a specific one. This is all the more important when taking a look at the result of Eberlein and Jacod (1997) who prove that in a setting where the equity price process is given by an infinite variation process, the exponential Lévy model has the highest degree of incompleteness possible: Every option price in the no-arbitrage interval can be obtained through a corresponding change of measure. The nice thing about Lévy processes is that there is a simple parametrization of the change of measure, given the assumption that that the equity price process is again a Lévy process under $\mathbf{Q}$. The change of measure is then given by two objects: A real-valued variable changing the drift of the Wiener process and a non-stochastic positive function, henceforth called *measure change function*, which changes the intensity of the jumps. In the case where the equity price process is pure jump, all the content of the change of measure is contained in the measure change function. In this special case, the incompleteness issue becomes rather lucid: The measure change function is only required to be such that the price process is a martingale under $\mathbf{Q}$ which amounts to solving a one-dimensional equation in terms of an infinite-dimensional variable from a rather general function space.

There are about two different types of handling the problem as indicated by the upper two boxes of Fig. 1. The first one takes as given the process under the measure $\mathbf{P}$ and specifies a change of measure by giving a computationally simple parametric form of the measure change function and thereby reducing the problem to a finite-dimensional one. The most popular change of measure is the Esscher change of measure, which is fixed by determining the only parameter of an exponential measure change function such that it fulfils the martingale condition for the price process of the underlying. Therefore, we denote this class by the name *Esscher type martingale measures*. There is a wide range of justifications for these measures: They

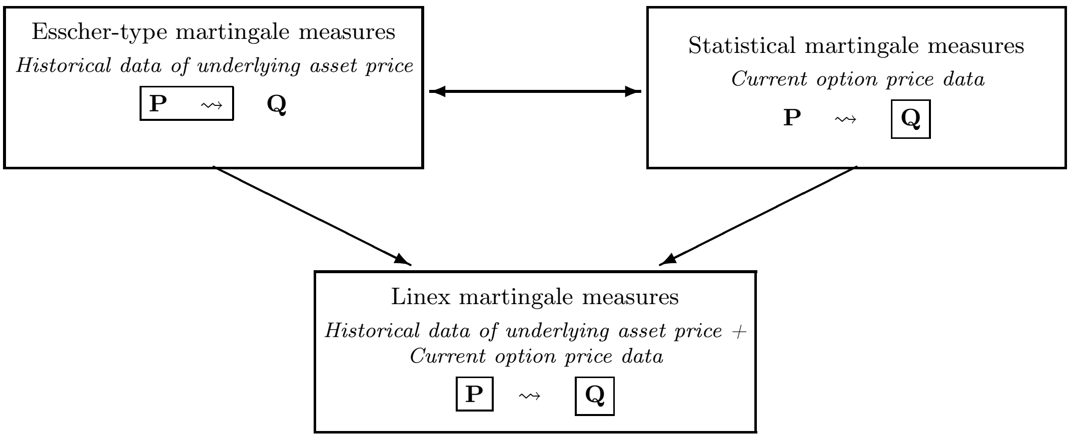


Fig. 1. Different concepts of the change of measure. The framed objects are the starting points of each method.

arise, for instance, as a result of a utility maximization problem of a representative investor or of a distance minimization theorem in the sense of information theory, or simply by tractability considerations.

Chapter 2 develops a class of Esscher-type martingale measures for normal inverse Gaussian Lévy processes which provides more degrees of freedom for modelling the risk-neutral distribution than the Esscher change of measure and which is therefore called the class of *flexible martingale measures.* Based on the contradictory facts that on the one hand all important measure change functions in the literature are monotone and on the other hand Carr et al. (2000a) observe that the option-implied measure change function should show some kind of symmetry with respect to the abscissa, the flexible measures are constructed such that they can incorporate both shapes. They share with the Esscher transform the property of being easily used for fast pricing by Fourier inversion.

Esscher-type martingale measures lack an essential property with regard to consistent option pricing. That is, this procedure implicitly assumes a specific risk-neutral distribution which may well differ from the one chosen by the market. The empirical literature (e.g. Bates (1991)) has dealt with the estimation of risk-neutral distributions for a long time. For instance, one conclusion was that the crash of 1987 significantly altered $\mathbf{Q}$ but not $\mathbf{P}$, for instance by charging a higher premium for out-of-the-money options as a traditional instrument for protection against downside risk. Hence it seems important to take these facts into consideration when designing a change of measure.

Coming back to Fig. 1 one way of handling this problem is the use of *statistical* martingale measures. This theory estimates the stock price process directly under the risk-neutral probability measure, i.e. by using current option price data from exchanges without worrying about the historical stock return distribution and any change of measure procedure. This is adequate as long as one is not in need of the statistical distribution as well. However, suppose we have the following situation: An option writer sells an option to an investor. This entails two tasks for the option writer: In the first place, he has to find an adequate, i.e. fair price for which he is ready to sell the option. Secondly, he might want to hedge the risk he incurs through the option. For the first task he needs the pricing rule, i.e. the risk-neutral distribution $\mathbf{Q}$ of the underlying asset. And hedging in a Lévy market setting could mean to follow a quadratic hedge to minimize the expected hedging error. But to be meaningful, this expectation should be taken under the statistical measure. Hence, this is an example of a situation where one actually needs both - the risk-neutral *and* the statistical distribution of the stock price process. The usual procedure to tackle this problem would be to estimate $\mathbf{P}$ as well as $\mathbf{Q}$ assuming the price process under both measures to be of the same class of Lévy processes. This entails the problem that these two measures need not necessarily be equivalent. On the contrary, for two widely used classes of Lévy processes Raible (2000) has shown that absolute equivalence, hence the no-arbitrage property of option prices, imposes rather strong constraints on the parameters, which leaves not much freedom of capturing both asset option price data sufficiently well. A further method would be to do the exact opposite of the Esscher-type approach: Estimating the stock price process under the risk-neutral law, and then performing a change of measure to obtain the process under the historical law. Assuming a Lévy process for the risk-neutral movement, one could be forced to consider changes of measures which do not any more preserve the Lévy property in order to obtain a good fit to historical data, e.g. to model some kind of dependence of the increments.

A possible solution to this problem is proposed in **Chapter 3** for the case of tempered stable Lévy processes. Given a specific parametric form of the considered Lévy process under the statistical distribution, it is not assumed that the class of the risk-neutral distribution is known. Instead we presuppose that only the risk-neutral second, third, and fourth moments are known. This is equivalent to being given the risk-neutral volatility, skewness, and kurtosis. In addition to the martingale condition this amounts to four restrictions. As in general the number of free parameters of the measure change function is bigger, we choose that function which minimizes the relative entropy with respect to the original measure, given those four constraints. Additionally, we can force the measure change function to have different monotonicity features and therefore resume the topic of Chapter 2 from a different angle. All in all, this amounts to solve a non-linear minimization problem, but with equality and inequality constraints which are all linear. The linearity is due to the specific form of the elements in the finite-dimensional space of measure change functions that we will choose. For reasons which become clear later on, this new change of measure will be called *linex change of measure*. Fig. 1 explains graphically the gap which this approach fits into.

Chapter 4 deals with a very different topic. However, it works with the technique of statistical martingale measures such that we cover all three types of change of measure in this thesis. This chapter focuses on basket option pricing in a multidimensional Lévy model with both linear and non-linear jump dependence of the components.
Basket options deal with the risk exposure of portfolios of risky assets. A European basket call resp. put option is just a plain-vanilla call resp. put option on a portfolio of assets instead on a single asset. While wanting to hedge against the risk involved with portfolios, basket options are generally cheaper than options on single assets, which in principle could also be used. The reason is that the prices of basket options are functions of the dependence of the assets. High negative dependence between two assets clearly reduces the risk of the portfolio and leads to lower prices for basket options on a portfolio of these two assets. Therefore there is a clear need to compute fair prices of basket options.
However, pricing is not straightforward, and up to now there is no closed-form solution even for relatively simple models for the price of the assets. The literature deals almost exclusively with the case where the underlying assets are modelled by the

multidimensional Black-Scholes model with correlated Wiener processes. The problem in this case is that there is no closed form of the distribution of the weighted sum of lognormally distributed random variables. In particular, it does not any more follow a lognormal distribution. Hence appropriate methods for approximate pricing even in this relatively simple model are required (see e.g. Gentle (1993), Huynh (1994), Blix (1998), Ju (2002), Deelstra et al. (2004) among many others).
Using the multi-dimensional Black-Scholes model implies normally distributed returns and linear correlation. Both implications are regularly rejected by financial data: As mentioned above, for one-dimensional returns one observes constantly that they are skewed and leptokurtotic and very far away from being appropriately described by a normal law (see e.g. Cont (2001)). As we have already mentioned, in hedging portfolios of assets it is cheaper to use basket options than options on the single assets contained in the portfolio because dependence risk is also priced. This means that dependence should be carefully modelled, and as there is evidence (e.g. Breymann et al. (2003)) that inter-asset dependence is sometimes not properly modelled just by linear correlation, it follows that other measures of dependence should be included in the model. The most important one is probably tail dependence. Tail dependence is an asymptotic concept which describes intuitively the probability of a big jump of one component, conditional on the event that another component has also a big jump. For linear correlation this probability is zero, but evidence suggests that non-zero tail dependence is indeed a property of financial data, given for example the downward comovement of equities when big macroeconomic crises hit the world economy.
All these stylized facts seem to suggest the use of a more sophisticated model that relaxes the assumptions inherent in the Black-Scholes world. Flamouris and Giamouridis (2004) seems to be the first reference to deal explicitly with the valuation (and hedging) of basket options in a multidimensional jump diffusion model for the underlying securities. A severe drawback with this approach is that the jumps are assumed to be independent. However, in a jump-diffusion model jumps are supposed to describe the impact of major macroeconomic events, hence jumps of course should be dependent to some extent.
As a first step to overcome these problems, in Chapter 4 we develop a tractable two-dimensional Lévy model based on Kou's one-dimensional jump-diffusion model whose dependence is modeled through a Lévy copula, a very recent approach which was introduced by Tankov (2003). Tractability means again that using this model we can very fast determine approximate values of arbitrage-free prices of options on a portfolio of two dependent assets.

The final **Chapter 5** deals with risk-minimizing hedging in an exponential Lévy model. Before solving for the hedge ratios in a multidimensional model, the essence of the problem is explained in a one-dimensional context.

Eventually, the very first **Chapter 1** provides an introduction into the mathematical tools which are used throughout the thesis in order to handle Lévy processes.

Chapter 1

Preliminaries

This chapter gives a short introduction into the tools which are related to the use of Lévy processes in option pricing theory and which are needed in the main part of this thesis. Beginning with the notion of a Poisson random measure we construct jump processes and discuss the stochastic analysis of Lévy processes. After the definition of an exponential Lévy model we review the three classes of Lévy processes which are used in the three main chapters of this thesis and show how the Fourier inversion pricing tool works in the case of Lévy processes. Finally, we give an introduction into the theory of Lévy copulas. This chapter is based mainly on four textbooks: Cont and Tankov (2004b), Jacod and Shiryaev (2003), Protter (1995) and Sato (1999).
As a general assumption we will always work with a filtered complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbf{P})$ with a finite time horizon $T \in [0,\infty)$ and a right-continuous filtration $(\mathcal{F}_t)_{0\leq t\leq T}$. We will also assume that $\mathcal{F}_0$ is the trivial σ-algebra $\{\emptyset, \Omega\}$. Unless otherwise mentioned, the expectation operator E is always with respect to $\mathbf{P}$.
Whenever we work in a multidimensional framework we will make use of a convenient matrix notation: The transposed of a matrix A is denoted by A'. A vector in $\mathbb{R}^n$ without the transposed sign is always understood as a column vector. For an arbitrary matrix Ξ we will denote its i-th row by the (column!) vector Ξ^i. This implies that for a vector ξ the symbol ξ^i stands for its i-th entry[1]. Correspondingly, Ξ^{ij} denotes the entry in the i-th row and j-th column of Ξ. With $|x|$ we denote an arbitrary norm of the finite-dimensional variable x, whether x be a vector or a matrix. For a real number x it denotes the standard norm $|x| = \sqrt{x^2}$. For the sake of a convenient notation we often omit the time parameter when we talk about stochastic process. Thus, the stochastic process $(X_t)_{0\leq t\leq T}$ will mostly be abbreviated by X. Moreover, for a complex number z we denote by $\Re(z)$ its real part and by $\Im(z)$ its imaginary part.

[1] However, in some parts of the thesis it is more convenient to use the subscript notation ξ_i, and we will freely choose the most adequate notation. In any case, there will be no danger of confusion.

1.1 Poisson random measures and jump processes

Let E be a Hausdorff space which is equipped with a σ-algebra $\mathcal{E}$. Typical choices for E in our context will be shown below.

Definition 1.1 (Radon measure). *Let $E \subset \mathbb{R}^n$. A measure ν on $(E, \mathcal{E})$ is a* Radon measure *if $\nu(B) < \infty$ for every compact set $B \in \mathcal{E}$.*

Given a Radon measure ν we show in this section that by way of a Poisson random measure one can construct a stochastic process which moves only by jumps. First we have to go from ν, which will be interpreted as an intensity measure, to the associated Poisson random measure μ:

Definition 1.2 (Poisson random measure and intensity measure). *Let $E \subset \mathbb{R}^n$ and ν be a positive Radon measure on $(E, \mathcal{E})$. A* Poisson random measure *on E with* intensity measure *ν is an integer-valued random measure:*

$$\begin{aligned} \mu : \Omega \times \mathcal{E} &\to \mathbb{N}_0 \\ (\omega, A) &\to \mu(\omega, A) \end{aligned}$$

such that

- *For almost all $\omega \in \Omega$, $\mu(\omega, \cdot)$ is an integer-valued Radon measure on E.*
- *For any compact $A \in \mathcal{E}$, $\mu(\cdot, A)$ is a Poisson random variable with parameter $\nu(A)$.*
- *For disjoint sets $A_1, \ldots, A_n \in \mathcal{E}$, the variables $\mu(\cdot, A_1), \ldots, \mu(\cdot, A_n)$ are independent.*

Proposition 1.3. *For any given Radon measure ν on $E \subset \mathbb{R}^n$, there exists a Poisson random measure μ on E with intensity ν.*

Proof. Cont and Tankov (2004b), p. 57. □

The step from a Poisson random measure towards the construction of a stochastic process is initiated by giving the space E a structure which is interpreted as the Cartesian product of time and jump size:

$$E = [0, T] \times (\mathbb{R}^n \setminus \{0\})$$

Denoting with δ_a the Dirac measure at the point a, any Poisson random measure on E has the representation

$$\mu = \sum_{i \geq 1} \delta_{(T_i, Y_i)}$$

as a counting measure associated to a randomly selected configuration of points $(T_i, Y_i) \in E$. For a more thorough interpretation of such a measure we define in terms of the filtration $\mathcal{F}$ and the sequence $(T_i, Y_i)_{i \geq 1}$:

Definition 1.4. *A Poisson measure* μ *is said to be* non-anticipating *if*

- $(T_i)_{i\geq 1}$ *are non-anticipating random times;*
- Y_i *is* $\mathcal{F}_{T_i}$*-measurable.*

This means that μ describes a jump structure in time and space in that it prescribes at the non-anticipating time T_i a jump whose size Y_i is not revealed until T_i. Stochastic processes are constructed through an integral of a real-valued measurable function $f : E \to \mathbb{R}$ with respect to a non-anticipating μ. Given a random measure μ on E, one starts as usual with the definition of an integral for a simple function $f = \sum_{j=1}^m c_j \mathbf{1}_{A_j}$ for $c_j \geq 0$ and disjoint sets $A_j \in \mathcal{E}, j = 1, \ldots, m$, and one defines

$$\int_0^T \int_{\mathbb{R}^n\setminus\{0\}} f(s,y)\mu(ds,dy) := \sum_{j=1}^m c_j\mu(A_j).$$

Then one uses the monotone convergence theorem to extend this definition first to positive integrands and then to an arbitrary measurable function[2]. Finally, given that for such a function f we have

$$\int_0^T \int_{\mathbb{R}^n\setminus\{0\}} |f(s,y)|\nu(ds,dy) < \infty, \tag{1.1.1}$$

we have defined the integral

$$\int_0^T \int_{\mathbb{R}^n\setminus\{0\}} f(s,y)\mu(ds,dy),$$

which is absolutely integrable due to (1.1.1), and we obtain

$$E\left[\int_0^T \int_{\mathbb{R}^n\setminus\{0\}} f(s,y)\mu(ds,dy)\right] = \int_0^T \int_{\mathbb{R}^n\setminus\{0\}} f(s,y)\nu(ds,dy).$$

by taking the expected value with respect to $\mathbf{P}$. In terms of the sequence $(T_i, Y_i)_{i\geq 1}$ associated to μ we have

$$\int_0^T \int_{\mathbb{R}^n\setminus\{0\}} f(s,y)\mu(ds,dy) = \sum_{\{i,T_i\in[0,T]\}} f(T_i, Y_i).$$

Introducing the useful abbreviation $f * \mu_T := \int_0^T \int_{\mathbb{R}^n\setminus\{0\}} f(s,y)\mu(ds,dy)$ and restricting the time domain of the integral from $[0,T]$ to $[0,t]$, $0 \leq t \leq T$, the random variables

$$X_t = f * \mu_t, \quad 0 \leq t \leq T,$$

[2]See Cont and Tankov (2004b), p.59.

define an adapted stochastic process on the given filtered probability space, which moves only by jumps of size $f(T_i, Y_i)$ at the random times T_i.
This section has shown that beginning with an intensity measure ν one can step by step define a pure jump stochastic processes from ν. In the special case where this process becomes a Lévy process the measure ν will turn out to be very closely related to the Lévy measure K to be defined in the next section.

1.2 Stochastic integrals

Before coming to the definition of a Lévy process we will give the definition of the stochastic integral of a function f with respect to a compensated Poisson random measure. The integral of $f * \mu_t$, which was defined in the previous section for a non-stochastic function f can be quite easily extended to predictable functions f:[3]
A function on $\Omega \times E$ is said to be *predictable* if it is measurable with respect to the σ-algebra on $\Omega \times E$ generated by all left-continuous adapted processes. The following definition of a stochastic integral is from Jacod and Shiryaev (2003), Definition II.1.27.

Definition 1.5. *a) We denote by $G(\mu)$ the set of all predictable real-valued functions f on $\Omega \times E$ such that the increasing process*

$$\left\{ \sum_{u \leq \cdot} [f(u, \Delta X_u) \mathbf{1}_{\Delta X_u \neq 0}]^2 \right\}^{1/2}$$

is locally integrable.

b) If $f \in G(\mu)$ we call stochastic integral of f with respect to $\mu - \nu$ and we denote by $f * (\mu - \nu)$ *any purely discontinuous local martingale X such that ΔX and f are indistinguishable.*

We add two useful definitions:

Definition 1.6. *Given a Poisson random measure μ, we denote by*

- *$\mathcal{J}^1(\mu)$ the set of all predictable functions f on $\Omega \times E$ such that $|f| * \nu$ is locally integrable and increasing.*
- *$\mathcal{J}^2(\mu)$ the set of all predictable functions f on $\Omega \times E$ such that $|f|^2 * \nu$ is locally integrable and increasing.*

The following proposition is a collection of useful results related to stochastic integrals with respect to a random measure. We say that a martingale X is *square-integrable* if $\sup_{t \in [0,T]} E[|X_t|^2] < \infty$. Moreover, given two stochastic processes X

[3] See Jacod and Shiryaev (2003), II.1.5.

and Y, the processes $[X, Y]$ and $\langle X, Y\rangle$ denote the quadratic covariation process[4] resp. the predictable quadratic covariation process[5] of X and Y. For the latter to be well-defined X and Y must be at least locally square-integrable martingales.

Proposition 1.7. *Let μ be a Poisson random measure.*

a) If $f \geq -1$ identically then $f \in G(\mu)$ if and only if the increasing process

$$\left(1-\sqrt{1+f}\right)^2 * \nu$$

is locally integrable.

b) If $f \in \mathcal{J}^1(\mu)$ then $f \in G(\mu)$ and

$$f * (\mu - \nu) = f * \mu - f * \nu.$$

c) If $f_1 \in G(\mu)$ and $f_2 \in \mathcal{J}^1(\mu)$ then

$$f_1 * (\mu - \nu) + f_2 * \mu = [f_1 + f_2] * (\mu - \nu) + f_2 * \nu.$$

d) If $f \in \mathcal{J}^2(\mu)$ then $f * (\mu - \nu)$ is a square-integrable martingale and

$$E\left[|f * (\mu - \nu)|^2\right] = E\left[|f|^2 * \nu\right].$$

e) If $f_1, f_2 \in \mathcal{J}^2(\mu)$ then

$$[f * (\mu - \nu), g * (\mu - \nu)] = (fg) * \mu$$

and

$$\langle f * (\mu - \nu), g * (\mu - \nu)\rangle = (fg) * \nu.$$

Proof. The items *a*) and *b*) are Theorem II.1.33d resp. Proposition II.1.28 in Jacod and Shiryaev (2003) whereas *c*) is a special case of Proposition 5.3 in Goll and Kallsen (2000). *d*) is Proposition 8.8 in Cont and Tankov (2004b), and the first part of *e*) is implied by Jacod and Shiryaev (2003), Theorem I.4.47, and Definition 1.5. The second part of *e*) follows by Theorem II.1.33a and the localization identity I.4.3 in Jacod and Shiryaev (2003). □

1.3 Lévy processes

Definition 1.8. *An adapted $\mathbb{R}^n$-valued stochastic process $X = (X_t)_{0 \leq t \leq T}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbf{P})$ is called a Lévy process if*

a) $X_0 = 0$ $\mathbf{P}$-a.s.;

[4] Jacod and Shiryaev (2003), I.4.45.
[5] Jacod and Shiryaev (2003), I.4.2

b) $X_t - X_s$ *is independent of* $\mathcal{F}_s$, $0 \leq s \leq t \leq T$ *(independent increments);*

c) $X_t - X_s \stackrel{d}{=} X_{t-s}$, $0 \leq s \leq t \leq T$ *(stationary increments);*

d) $\lim_{s\to t} X_s = X_t$ *in probability (stochastic continuity).*

In order to study the path properties of a Lévy process we need the notion of a càdlàg function: A function which is right-continuous and has left limits is said to be *càdlàg*.
Every Lévy process has a modification[6] which is càdlàg[7]. In the following we will always assume that we work with the càdlàg modification of a Lévy process. A stochastic process which is càdlàg has two important path properties: The total number of jumps is at most countable, and the number of jumps whose size is bigger (in absolute value) than any arbitrary $\epsilon > 0$ is finite[8]. A Lévy process X has a useful representation in terms of the characteristic function χ_t of X_t:

Proposition 1.9. *Let* $(X_t)_{0\leq t\leq T}$ *be a Lévy process on* $\mathbb{R}^n$*. There exists a continuous function* $\psi : \mathbb{R}^n \to \mathbb{C}$ *such that*

$$\chi_t(z) := E[e^{iz'X_t}] = e^{t\psi(z)}, \quad z \in \mathbb{R}^n.$$

Proof. Cont and Tankov (2004b), p.70. □

The investigation of the special structure of the involved *cumulant function* ψ is the purpose of the Lévy-Khinchin representation which will be introduced below. The following definition is taken from Sato (1999).

Definition 1.10. *A probability measure* m *on* $\mathbb{R}^n$ *is said to be* infinitely divisible *if, for any positive integer* p*, there is a probability measure* m_p *on* $\mathbb{R}^n$ *such that* m *is equal to the* p*-fold convolution of* m_p*. The latter is denoted by* m_p^{p*}*.*

A Lévy process can be constructed by means of a single measure which has to be infinitely divisible. Let P_ξ denote the law of the random variable ξ under $\mathbf{P}$.

Theorem 1.11. *If* m *is an infinitely divisible probability measure on* $\mathbb{R}^n$*, then for every* $t \in [0,T]$ m^{t*} *can be defined.* $(m^{t*})_{0\leq t\leq T}$ *is a convolution semigroup of probability measures on* $\mathbb{R}^n$ *from which a probability space* $(\Omega, \mathcal{F}, \mathbf{P})$ *and a Lévy process* $X = (X_t)_{0\leq t\leq T}$ *on* $(\Omega, \mathcal{F}, \mathbf{P})$ *can be constructed such that* $P_{X_1} = m = m^{1*}$*. This Lévy process is unique in law.*

Proof. Sato(1999), Lemma 7.9 and Theorem 7.10. □

In the following we will analyse the reverse direction of what was done in Section 1.1: Given a Lévy process X, we derive the jump measure μ^X and the Lévy measure K.

[6]A stochastic process X' is called a *modification* of a stochastic process X if $\mathbf{P}(\{X_t = Y_t\}) = 1$ for $0 \leq t \leq T$.
[7]See Sato (1999), Theorem 11.5.
[8]See Cont and Tankov (2004b), p.38.

Definition 1.12 (Lévy measure). *Let $(X_t)_{0\le t\le T}$, $T>1$, be a Lévy process on $\mathbb{R}^n$. The measure K on $\mathbb{R}^n$ defined by*

$$K(A)=E[\#\{t\in[0,1]:\Delta X_t\neq 0,\Delta X_t\in A\}],\quad A\in\mathcal{B}(\mathbb{R}^n),$$

is called the Lévy measure *of X.*

$K(A)$ thus gives the expected number of jumps per unit time with size in A. We assign now a random measure μ^X to a given Lévy process X on $\mathbb{R}^n$: For any measurable set $(\tau,A)\subset[0,\infty)\times\mathbb{R}^n\setminus\{0\}$ we define

$$\mu^X(\tau,A):=\#\{(t,X_t-X_{t-})\in(\tau\times A)\}. \tag{1.3.1}$$

If $\tau=[t_1,t_2]$ for $t_1<t_2$, this definition gives the random number of jumps of X between times t_1 and t_2 with sizes in the set A. The following proposition gives a fundamental characterization of the structure of a Lévy process. It provides a decomposition of a Lévy process in a deterministic drift part, a continuous diffusion part and a jump part. We use the following definition: Given a positive definite matrix c, we denote by $\sqrt{c}$ the square root of c, in the sense that $c=\sqrt{c}\sqrt{c}$. The matrix c is interpreted as a covariance matrix, so in addition to being positive definite it is also symmetric. Hence there exist an orthogonal matrix P and a diagonal matrix D with the (positive) eigenvalues as entries on the main diagonal such that

$$c=PDP'=(PD^{\frac{1}{2}}P')(PD^{\frac{1}{2}}P')'$$

where $D^{\frac{1}{2}}$ is generated from D by replacing the eigenvalues by their square roots. The matrix $\sqrt{c}:=PD^{\frac{1}{2}}P'$ is evidently symmetric, and we have the following relation between c and $\sqrt{c}$:

$$\sum_{k=1}^{n}(\sqrt{c})^{jk}(\sqrt{c})^{ik}=c^{ij},\quad i,j=1,\dots,n. \tag{1.3.2}$$

From the above decomposition of $\sqrt{c}$ it follows as well that the inverse c^{-1} of c is also symmetric.

Theorem 1.13 (Lévy-Itô decomposition). *Let $(X_t)_{0\le t\le T}$ be a Lévy process on $\mathbb{R}^n$ and K its Lévy measure. Then we have*

- *K is a Radon measure on $\mathbb{R}^n\setminus\{0\}$ and satisfies*

$$\int_{\mathbb{R}^n}\min(|x|^2,1)K(dx)<\infty. \tag{1.3.3}$$

- *The jump measure μ^X of X is a Poisson random measure on $[0,T]\times\mathbb{R}^n$ with intensity measure $dtK(dx)$;*

- *There exist a vector $b \in R^n$, an n-dimensional standard Brownian motion W and a symmetric non-negative definite matrix $c \in \mathbb{R}^{n\times n}$ such that*

$$X_t = bt + \sqrt{c}W_t + h(x) * (\mu^X - \nu)_t + (x - h(x)) * \mu_t^X \tag{1.3.4}$$

where $h : \mathbb{R}^n \to \mathbb{R}^n$ is given by $h(x) := x\mathbf{1}_{0\le|x|\le 1}$.

Proof. Cont and Tankov (2004b), p. 79. □

Theorem 1.13 provides a decomposition of X into a martingale and a predictable process. Hence a Lévy process is a special semimartingale[9], and the semimartingale characteristics[10] of X are given by $(bt, ct, dtK(dx))$. The integral $h(x) * (\mu^X - \nu)_t$ is the stochastic integral of h with respect to the compensated jump measure $(\mu^X - \nu)$ as given in Definition 1.5.
Adding formally all the integrals with respect to μ^X in (1.3.4) results in the new jump part $x * \mu_t^X$. Rather intuitively, this is nothing but the sum of the jumps of X. But for a general jump structure of a Lévy process this integral does not exist in the sense of our definition in Section 1.1. This is why it is necessary to truncate the small jumps by the *truncation function* h which ensures existence of the two jump-related integrals in (1.3.4). The choice of h made in Proposition 1.13 is the common one in the literature, but other measurable functions h with $h(x) = xg(x)$ and $g : \mathbb{R}^n \to \mathbb{R}$ can also be used as long as they are bounded and satisfy[11]

$$g(x) = 1 + o(|x|) \quad \text{for } |x| \to 0 \tag{1.3.5}$$

$$g(x) = O(1/|x|) \quad \text{for } |x| \to \infty. \tag{1.3.6}$$

However, it must be noted that the vector b depends on h whereas c and K are independent of the choice of h. From the Lévy-Itô decomposition it is not hard to obtain the Lévy-Khinchin representation which is a kind of extension of Proposition 1.9:

Theorem 1.14 (Lévy-Khinchin representation). *Let $(X_t)_{0\le t\le T}$ be a Lévy process on $\mathbb{R}^n$. Then there exist b, c, and K as in Proposition 1.13 such that the characteristic function χ_t of X_t has the representation*

$$\chi_t(z) = E[e^{iz'X_t}] = e^{t\psi(z)}$$

where

$$\psi(z) = ib'z - \frac{1}{2}z'cz + \int_{\mathbb{R}^n} (e^{iz'x} - 1 - iz'h(x))K(dx). \tag{1.3.7}$$

Proof. Sato (1999), Theorem 1.11 and Proposition 1.9. □

---

[9] See e.g. Jacod and Shiryaev (2003), p.43.
[10] Jacod and Shiryaev (2003), p.76
[11] Sato (1999), p.38.

For the function χ_1 we also write simply χ. Due to their fundamental role b, c, and K are summarized under the name *characteristic triplet* or *Lévy triplet*. Henceforward we define Lévy processes X by their Lévy triplet, and we write $X = (b, c, K)_{\mathbf{P}}$, thus keeping in mind that the Lévy triplet is always defined relative to the underlying probability measure. We thus fix a Lévy process which is unique in law.
From Theorem 1.14 the role of h resp. g becomes more lucid: For small $|x|$ we have because of (1.3.5)

$$|e^{iz'x} - 1 - iz'xg(x)| = O(|x|^2) \quad \text{for } |x| \to 0.$$

This together with the condition (1.3.3) on the Lévy measure K implies K-integrability at the origin of the integrand in the above formula for ψ. For large $|x|$ the integrand being bounded is a sufficient condition for integrability (with condition (1.3.3) in view), and we obtain indeed for a positive constant C via (1.3.6)

$$|e^{iz'x} - 1 - iz'xh(x)| \leq 2 + |z||x||g(x)| \leq 2 + C \quad \text{for } |x| \to \infty$$

The Lévy measure is the decisive variable with which important properties of the corresponding pure jump Lévy process $(0, 0, K)_{\mathbf{P}}$ can be deduced: As the paths of a Lévy process are almost surely càdlàg, the Lévy process has only a finite number of jumps with size bigger than any positive threshold value in every finite time interval. This observation corresponds to the property of the Lévy measure that $K(\mathbb{R}^n \setminus B_\varepsilon(0)) < \infty$ for any $\varepsilon > 0$. As for the small jumps, from condition (1.3.3) we see that by definition of a Lévy process we have $\int_{|x|\leq 1} |x|^2 K(dx) < \infty$. Hence the special case of infinite variation, i.e. $\int_{|x|\leq 1} |x| K(dx) = \infty$, is perfectly included in the definition of a Lévy process, which means that adding the jumps of such a pure jump Lévy process does not result in a random variable with finite expectation. But if we do have finite variation, we can choose a particularly easy truncation function, which does not have to satisfy condition (1.3.5), namely $h(x) \equiv 0$. Finite variation is given in the following two situations where the second one implies the first one:

- $\int_{|x|\leq 1} |x| K(dx) < \infty$. This means that the assigned Lévy process has jumps of *finite variation*, i.e. we can represent it as the sum of its jumps.
- $\int_{|x|\leq 1} K(dx) < \infty$. This resulting Lévy process has *finite activity*, i.e. in every finite time interval it jumps only finitely often.

Aside from these two special cases corresponding to the behaviour of K in the neighbourhood of zero there is one more special case which is worth mentioning. If we have $\int_{|x|\geq 1} |x| K(dx) < \infty$ then the assigned Lévy process X satisfies $E[|X_t|] < \infty$ for $0 \leq t \leq T$, following Lemma 1.17, and the choice $h(x) \equiv x$ is possible.
The following lemmas will be important in due course:

Theorem 1.15 (Doléans-Dade Exponential Formula). *Let $X = (b, c, K)_P$ be a real-valued Lévy process. The stochastic differential equation*

$$Z_t = 1 + \int_0^t Z_{s-}\, dX_s \tag{1.3.8}$$

has a (up to indistinguishability) unique solution which is of the form

$$Z_t = \mathcal{E}(X)_t = e^{X_t - \frac{c}{2}t} \prod_{s \leq t} (1 + \Delta X_s) e^{-\Delta X_s} \tag{1.3.9}$$

where $\Delta X_t := X_t - X_{t-}$, *and the product in (1.3.9) is absolutely convergent.*

Proof. Jacod and Shiryaev (2003), I.4.61. □

The product in (1.3.9) makes sense because X has only a countably infinite number of jumps due to its càdlàg property. The characteristic function $z \to \chi(z)$ of a probability measure always exists for every $z \in \mathbb{R}^n$ because e^{izx}, which is to be integrated, is bounded. For some purposes, though, it is useful to extend the domain of definition to some subsets of $\mathbb{C}^n$ which are not included in $\mathbb{R}^n$. But then the boundedness in question disappears, and for general $z \in \mathbb{C}^n$ we have no longer integrability of (1.3.7). The moment-generating function, defined by $t \to \chi(-it)$, deals with this problem. Regarding χ as a function with domain $\mathbb{C}^n$ (not necessarily finite everywhere) it cuts out a vertical straight line of the plane representation of $\mathbb{C}^n$. Clearly, the domain of existence of the moment-generating function is a very delicate issue. If it contains an open interval around zero, we have the following nice result:

Lemma 1.16. *Let Y be a random variable on $(\Omega, \mathcal{F}, \mathbf{P})$, whose moment generating function exists on some open interval on the real line containing* 0. *Then all moments exist, i.e.*

$$E[|Y|^\alpha] < \infty \qquad \forall \alpha > 0. \tag{1.3.10}$$

Proof. See Gut (1995), Theorem III.3.3. □

A further useful result is that the question of integrability of a Lévy process can be traced back to the integrability of the Lévy measure in the following way:

Lemma 1.17. *For the moments of a Lévy process $X = (b, c, K)_{\mathbf{P}}$ we have the following equivalences for $\alpha, \theta > 0$*

$$E[|X_t|^\alpha] < \infty \;\; \forall t \in [0, T] \qquad \Longleftrightarrow \qquad \int_{\{|x|>1\}} |x|^\alpha K(dx) < \infty. \tag{1.3.11}$$

and

$$E[e^{\theta|X_t|}] < \infty \;\; \forall t \in [0, T] \qquad \Longleftrightarrow \qquad \int_{\{|x|>1\}} e^{\theta|x|} K(dx) < \infty. \tag{1.3.12}$$

Proof. Sato(1999), Corollary 25.8. □

There is a simple corollary to Theorem 1.13.

Corollary 1.18. *The characteristic $b \in \mathbb{R}^n$ is linked with the expected value of a Lévy process X by*

$$E[X_1] = b + \int_{\mathbb{R}^n} (x - h(x))K(dx) \tag{1.3.13}$$

and $E[X_t] = tE[X_1]$.

Proof. Taking the expected value on both sides of (1.3.4) for $t = 1$ yields (1.3.13). The statement of the corollary is even valid for the case of non-existence of the first moment. In the light of Lemma 1.17 both sides of equation (1.3.13) are then infinity. □

Lemma 1.19. *Let X be an $\mathbb{R}^n$-valued Lévy process.*

a) If X is a local martingale, then it is a martingale.

b) If $e^X := (e^{X^1}, \dots, e^{X^n})'$ is a local martingale, then it is a martingale.

Proof. Following Cont and Tankov (2004b), Theorem 4.1, every linear transformation of a Lévy process is again a Lévy process. This means in particular that every $X^i, i = 1, \dots, n$, is a real-valued Lévy process, and we can apply Lemma 4.4 in Kallsen (2000). □

The following theorem develops a parametrization of all admissible changes of measure which preserve the property of independent and stationary increments. Certain objects, which depend on the underlying probability measure such as the expectation operator or the characteristic function, will henceforward have a corresponding index when ambiguity arises. We say a measure λ is absolutely equivalent with respect to another measure μ, which is defined on the same measurable space, if

$$\mu(A) = 0 \quad \Rightarrow \quad \lambda(A) = 0$$

for every measurable set A. We write then $\lambda \ll \mu$. They are said to be equivalent and we write $\lambda \sim \mu$ if $\lambda \ll \mu$ and $\mu \ll \lambda$.
We can identify a change of measure by a number β and a function y, i.e. by deterministic quantities, as specified in the following theorem:

Theorem 1.20. *Let $\mathbf{P}$ be a probability measure. Let X be a Lévy process on $\mathbb{R}^n$ with triplet $(b, c, K)_{\mathbf{P}}$. Then there is a probability measure $\mathbf{Q} \sim \mathbf{P}$ such that X is a $\mathbf{Q}$-Lévy process with triplet $(\tilde{b}, \tilde{c}, \tilde{K})_{\mathbf{Q}}$ if and only if there exist $\beta \in \mathbb{R}^n$ and a function y from $\operatorname{supp}(K) \subseteq \mathbb{R}^n$ into $\mathbb{R}_+$ satisfying*

$$\int_{\mathbb{R}^n} |h(x)(1 - y(x))|K(dx) < \infty \quad \textit{and} \quad \int_{\mathbb{R}^n} \left(1 - \sqrt{y(x)}\right)^2 K(dx) < \infty$$

and

$$\begin{aligned} \tilde{b} &= b + c\beta + \int_{\mathbb{R}^n} h(x)(y(x) - 1)\, K(dx) && (1.3.14) \\ \tilde{c} &= c \\ \frac{d\tilde{K}}{dK}(x) &= y(x). && (1.3.15) \end{aligned}$$

Proof. The necessary condition for equivalence of **P** and **Q** is from Jacod and Shiryaev (2003), IV.4.39c, and sufficiency was proved by Raible (2000) in the one-dimensional case. However, his proof extends easily to the case of a general dimension n. Moreover, the statement of the theorem is usually made for a function y with domain $\mathbb{R}^n$. However, y outside $\mathrm{supp}(K)$ does not have any impact on the theorem, thus this modification is trivial. □

The vector β changes the drift of the diffusion part of the Lévy process whereas y describes the jump structure under the new measure: It describes for every set $A \in \mathcal{B}(\mathbb{R}^n)$ of possible jump sizes of X the change of the jump intensity from $K(A)$ to $\int_A y(x)K(dx)$. For convenience we will refer to y as the *measure change function.* β and y are also called the *Girsanov quantities.* The density process Z which β and y describe has the following form:[12]

$$Z_t = \left.\frac{d\mathbf{Q}}{d\mathbf{P}}\right|_{\mathcal{F}_t} = \mathcal{E}(N_.)_t,\ t \in [0,T] \tag{1.3.16}$$

where

$$N = \beta W + (y(x)-1) * (\mu^X - \nu). \tag{1.3.17}$$

For Z to be a density process it must be a positive and uniformly integrable **P**-martingale with $Z_0 = 1$. The proof of this is implicit in Theorem 1.20 where we cited Raible (2000). He proved it by recurring to quite deep results from stochastic analysis. We think it worthwhile to give a more elementary proof.

Proposition 1.21. *If*

$$\int_{\mathbb{R}^n} \left(1 - \sqrt{y(x)}\right)^2 K(dx) < \infty \tag{1.3.18}$$

*then Z given in (1.3.16) is a uniformly integrable **P**-martingale.*

Proof. First of all, $y(x) - 1 \in G(\mu)$ due to (1.3.18) and Proposition 1.7a, thus N is a local **P**-martingale. It is positive because $y(x) \geq 0$ for all $x \in \mathbb{R}^n$ as can easily be seen from Theorem 1.15. Hence it is a supermartingale (see e.g. Harrison and Pliska (1981), 3.8), and following Jacod (1979), Lemma (7.10), it remains to show that $E^{\mathbf{P}}[Z_T] = E^{\mathbf{P}}[Z_0] = 1$.
N has independent and stationary increments, i.e. it is a Lévy process[13]. Following Kallsen (2000), Lemma 5.8, we can find a process $\hat{N}$ with $e^{\hat{N}} = \mathcal{E}(N)$ which is again a Lévy process. The proof of this result is essentially only an application of the Itô formula. As N is a local martingale, this is also true for $\mathcal{E}(N)$[14] and hence for $e^{\hat{N}}$. By Lemma 1.19 it is even a martingale, and we have $E^{\mathbf{P}}[Z_T] = E^{\mathbf{P}}[e^{\hat{N}_T}] = e^{\hat{N}_0} = 1$. □

This result is important for Chapters 2 and 3, in which we focus on the modelling of measure change functions that should eventually lead to an absolutely continuous change of measure.

[12] Jacod and Shiryaev (2003),III.5.19.
[13] See Sato (1999), Lemma 33.6.
[14] This can immediately be seen from 1.3.8.

1.4 The exponential Lévy model

In the context of the application of Lévy processes in option pricing theory many models work with the stochastic exponential, i.e. the stock price process S is given by

$$S_t = S_0\mathcal{E}(X)_t, 0 \le t \le T, \tag{1.4.1}$$

where X is a one-dimensional Lévy process. Despite its tractability from the point of view of stochastic analysis[15], it gives for most Lévy processes negative stock prices with positive probability. Furthermore, if the basic distribution of the Lévy model is determined by log return data, statistical estimation theory goes better together with the ordinary exponential than with the stochastic exponential model. The former is given by

$$S_t = S_0 e^{X_t}, 0 \le t \le T, \tag{1.4.2}$$

and the risk-free security is assumed to be

$$B_t = e^{rt}, 0 \le t \le T. \tag{1.4.3}$$

Note that for $X_t = (\mu - \sigma^2/2)t + \sigma W_t$ the geometric Brownian motion is obtained from Itô's lemma. Analogously, the multidimensional exponential Lévy model is given by

$$S_t = \mathcal{S}_0 e^{X_t}, 0 \le t \le T, \tag{1.4.4}$$

and (1.4.3) where $S := (S^1, \ldots, S^n)'$, $e^X := (e^{X^1}, \ldots, e^{X^n})'$ and $S_0^i > 0$ for $i \in \{1, \ldots, n\}$. Moreover we set $\mathcal{S} := \text{diag}(S^1, \ldots, S^n)$. As usual for this kind of models we state several further assumptions which we assume to be given throughout the thesis: We assume that trading takes place continuously without short-sale and borrowing and lending restrictions. Moreover, we assume the absence of transaction costs and taxes, and the assets are supposed to pay no dividends.

As the exponential Lévy model is incomplete the question arises how to price derivatives which are written on the security whose price is given by (1.4.2) resp. (1.4.4). Given a $\mathcal{F}_T$-measurable random variable H and assuming the existence of a probability measure $\mathbf{Q}$ under which the discounted asset price process $e^{-r\cdot}S$ is a martingale we propose

$$V_t := e^{-r(T-t)} E^{\mathbf{Q}}[H|\mathcal{F}_t] \tag{1.4.5}$$

as a reasonable price at time t for a contingent claim on H.
If H is attainable, (1.4.5) is its only sensible price, which moreover is unique over all martingale measures $\mathbf{Q}$.[16] For a non-attainable H we lose uniqueness, but (1.4.5) is still reasonable in the sense that it is not possible for an investor to generate an arbitrage strategy using the stock, the risk-free security *and* the derivative with the above price[17]. In other words, the extended market given by (1.4.2) resp. (1.4.4),

[15] One nice feature is that $\mathcal{E}(X)$ is a local martingale if X is a local martingale.
[16] See e.g. Björk (1992).
[17] See Keller (1997), Proposition 9.

(1.4.3) and (1.4.5) is arbitrage-free. For the following theorem see also Eberlein and Jacod (1997) for the one-dimensional case.

We introduce some more multidimensional notation. For $x = (x^1, \dots, x^n)$ we set

$$J : \mathbb{R}^n \ni x \to J(x) := (e^{x^1} - 1, \dots, e^{x^n} - 1)' \in \mathbb{R}^n.$$

and $\mathbf{1} := (1, \dots, 1)' \in \mathbb{R}^n$. Moreover, $\mathcal{C} := (c^{11}, \dots, c^{nn})'$ and $\tilde{\mathcal{C}} := (\tilde{c}^{11}, \dots, \tilde{c}^{nn})'$.

Theorem 1.22. *Let $X = (b, c, K)_{\mathbf{P}}$ be a Lévy process. Let $\beta \in \mathbb{R}^n$ and $y : \mathbb{R}^n \to \mathbb{R}_+$ be the Girsanov quantities which satisfy:*

$$\int_{\mathbb{R}^n} \left(1 - \sqrt{y(x)}\right)^2 K(dx) < \infty \tag{1.4.6}$$

and

$$\int_{|x|\geq 1} (e^{x^i} - 1) y(x) K(dx) < \infty \quad \textit{for} \quad i = 1, \dots, n. \tag{1.4.7}$$

Then the discounted asset price process $\hat{S} := (e^{-rt}\mathcal{S}_0 e^{X_t})_{0\leq t\leq T}$ is a $\mathbf{Q}$-martingale if and only if

$$b - r\mathbf{1} + c\beta + \frac{1}{2}\mathcal{C} + \int_{\mathbb{R}^n} (J(x)y(x) - h(x))K(dx) = 0 \in \mathbb{R}^n. \tag{1.4.8}$$

In particular, we have for $n = 1$

$$b - r + c\left(\beta + \frac{1}{2}\right) + \int_{-\infty}^{\infty} [(e^x - 1)y(x) - h(x)]\, K(dx) = 0. \tag{1.4.9}$$

Proof. For simplicity, we assume that $h(x) = x\mathbf{1}_{|x|\leq 1}$. We see that

$$\begin{aligned} |h^i(x)(y(x) - 1)| &= \left| h^i(x) \left\{ y(x) - 1 - \left(\sqrt{y(x)} - 1\right)^2 + \left(\sqrt{y(x)} - 1\right)^2 \right\} \right| \\ &\leq 2|h^i(x)||\sqrt{y(x)} - 1| + \left(\sqrt{y(x)} - 1\right)^2 \\ &\leq |h^i(x)|^2 + 2\left(\sqrt{y(x)} - 1\right)^2 \end{aligned}$$

where the first inequality follows essentially from straightforward calculations and $|h^i(x)| \leq 1$, and for the second the second binomial formula are used. From this we have that

$$\int_{\mathbb{R}^n} |h^i(x)(y(x) - 1)| K(dx) \leq \int_{B_1(0)} \left(x^i\right)^2 K(dx) + 2\int_{\mathbb{R}^n} \left(\sqrt{y(x)} - 1\right)^2 K(dx) < \infty.$$

Both integrals are finite because of (1.3.3) and (1.4.6). This justifies the application of Theorem 1.20. We have under $\mathbf{Q}$ through an application of Itô's lemma[18] to the

[18] For a version of Itô's Lemma where the function to be differentiated is vector-valued, see e.g. Arnold (1973), p.103, for the case of a diffusion process X.

discounted asset price $\hat{S}_t = e^{-rt}\mathcal{S}_0 e^{X_t}$

$$\begin{aligned}\hat{S}_t &= \hat{S}_0 - \int_0^t \hat{\mathcal{S}}_{u-} r\mathbf{1} du + \int_0^t \hat{\mathcal{S}}_{u-} dX_u + \frac{1}{2}\int_0^t \hat{\mathcal{S}}_{u-}\tilde{\mathcal{C}} du + [\hat{\mathcal{S}}_-(J(x)-x)] * \mu_t^X \\ &= \hat{S}_0 + \int_0^t \hat{\mathcal{S}}_{u-}(\tilde{b} - r\mathbf{1}) du + \frac{1}{2}\int_0^t \hat{\mathcal{S}}_{u-}\tilde{\mathcal{C}} du + \int_0^t \hat{\mathcal{S}}_{u-}\sqrt{\tilde{c}} d\tilde{W}_u \\ &\quad + [\hat{\mathcal{S}}_- h(x)] * (\mu^X - \tilde{\nu})_t + [\hat{\mathcal{S}}_-(J(x) - h(x))] * \mu_t^X. \end{aligned} \tag{1.4.10}$$

Condition (1.4.6) ensures that $\tilde{K}$ with $\tilde{K}(A) = \int_A y(x)K(dx)$, $A \in \mathcal{B}(\mathbb{R}^n)$, is again a Lévy measure. Therefore the fact that $e^{x^i} - 1 - x^i = O(|x|^2)$ for $|x| \to 0$ implies in connection with (1.4.7) that we have

$$\int_{\mathbb{R}^n} (e^{x^i} - 1 - h^i(x))\tilde{K}(dx) < \infty.$$

This means that we can apply Proposition 1.7c to S, and we obtain from (1.4.10)

$$\begin{aligned}\hat{S}_t &= S_0 + \int_0^t \hat{\mathcal{S}}_{u-}(\tilde{b} - r\mathbf{1}) du + \frac{1}{2}\int_0^t \hat{\mathcal{S}}_{u-}\tilde{\mathcal{C}} du + \int_0^t \hat{\mathcal{S}}_{u-}\sqrt{\tilde{c}} d\tilde{W}_u \\ &\quad + [\hat{\mathcal{S}}_- J(x)] * (\mu^X - \tilde{\nu})_t + [\hat{\mathcal{S}}_-(J(x) - h(x))] * \tilde{\nu}_t.\end{aligned}$$

Using Theorem 1.20 and substituting for the Lévy triplet under $\mathbf{Q}$ its drift is zero if and only if

$$\begin{aligned}& b - r\mathbf{1} + c\beta + \int_{\mathbb{R}^n} h(x)(y(x) - 1)K(dx) + \frac{1}{2}\mathcal{C} + \int_{\mathbb{R}^n} (J(x) - h(x))y(x)K(dx) \\ = \;& b - r\mathbf{1} + c\beta + \frac{1}{2}\mathcal{C} + \int_{\mathbb{R}^n} (J(x)y(x) - h(x))K(dx) = 0.\end{aligned}$$

This entails that the discounted asset price process is a local martingale. Following Lemma 1.19 it is even a martingale. □

1.5 Examples of Lévy processes

This section presents three classes of one-dimensional Lévy processes, which will be the workhorse for the Chapters 2, 3 and 4.

1.5.1 Normal inverse Gaussian (NIG) Lévy processes

Generalized hyperbolic (GH) distributions are defined by their five-parameter density function f_{GH} with respect to the Lebesgue measure, which is equal to

$$f_{GH}(x) := C\ (\alpha\sqrt{\delta^2 + (x-\mu)^2})^{\lambda - 1/2} e^{\beta(x-\mu)} K_{\lambda - 1/2}(\alpha\sqrt{\delta^2 + (x-\mu)^2}) \tag{1.5.1}$$

where

$$C := \frac{(\delta\sqrt{\alpha^2 - \beta^2})^\lambda}{\sqrt{2\pi}\alpha^{2\lambda - 1}\delta^{2\lambda} K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})}.$$

Its parameters are given by $\lambda \in \mathbb{R}$, $\alpha > 0$, $-\alpha < \beta < \alpha$, $\delta > 0$ and the location parameter $\mu \in \mathbb{R}$. K_ν denotes the modified Bessel function of the third kind[19] of the order ν.
In view of the exponential Lévy model we impose the additional requirement

$$1 < \alpha - \beta \tag{1.5.2}$$

which will become clear later on. GH distributions are infinitely divisible and thus can be used as a starting point to define a Lévy process. Such a Lévy process does not have a diffusion component, hence we have $c = 0$ in terms of the Lévy-Khinchin representation in Theorem (1.14).
The four-parameter normal inverse Gaussian (NIG) distribution corresponds to a generalized hyperbolic distribution with $\lambda = -1/2$. In this case f_{GH} boils down to

$$f_{NIG}(x) = \frac{\delta\alpha}{\pi} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)} \frac{K_1(\alpha\sqrt{\delta^2+(x-\mu)^2})}{\sqrt{\delta^2+(x-\mu)^2}}.$$

The characteristic function of the time-t-member of the convolution semigroup generated by a NIG distribution has the computationally simple representation

$$\chi_t(z) = \chi_t(z;\alpha,\beta,\delta,\mu) = e^{i\mu t z} \frac{\exp(t\delta\sqrt{\alpha^2-\beta^2})}{\exp(t\delta\sqrt{\alpha^2-(\beta+iz)^2})}$$

or in terms of the cumulant

$$\log \chi_t(z) = t\psi(z) = t\left(i\mu z + \delta\sqrt{\alpha^2-\beta^2} - \delta\sqrt{\alpha^2-(\beta+iz)^2}\right).$$

We see immediately from this representation of χ_t that

$$\chi_t(z;\alpha,\beta,\delta,\mu) = \chi(z;\alpha,\beta,t\delta,t\mu)$$

which means that unlike in the general case of a GH distribution the convolution semigroup generated by a NIG distribution is completely contained in the class of NIG distributions. This is of course a very convenient feature which allows to price options consistently in the NIG framework across all maturity dates.
The moment generating function φ is for an NIG distributed random variable X

$$\varphi(u;\alpha,\beta,\delta,\mu) := E[e^{uX}] = e^{\mu u} \frac{\exp(\delta\sqrt{\alpha^2-\beta^2})}{\exp(\delta\sqrt{\alpha^2-(\beta+u)^2})} \tag{1.5.3}$$

which exists for $|\beta+u| < \alpha$. The Lévy measure of a NIG distribution has full support $\mathbb{R}$ and is equal to

$$K(dx) = \frac{\delta\alpha}{\pi} e^{\beta x}|x|^{-1}K_1(\alpha|x|)\,dx, \tag{1.5.4}$$

For further information on NIG Lévy processes we refer to Barndorff-Nielsen (1997). For statistical parameter estimation by the maximum likelihood method Appendix A contains the log-likelihood function of the NIG distribution as well as a different parametrization, which will be useful for the rescaling technique that will be introduced in Chapter 2.

[19] See e.g. Abramowitz and Stegun (1972).

1.5.2 Tempered Stable Lévy Processes

A *tempered stable Lévy process* X has a Lévy measure whose density k is defined by

$$k(x) := \begin{cases} k_+(x) := c\,|x|^{-1-\nu}e^{-\lambda_+|x|} & , \quad x > 0 \\ 0 & , \quad x = 0 \\ k_-(x) := c\,|x|^{-1-\nu}e^{-\lambda_-|x|} & , \quad x < 0. \end{cases} \tag{1.5.5}$$

where $c > 0$, $\lambda_+, \lambda_- > 1$, and $\nu < 1, \nu \neq 0$. In the literature[20] the definition is usually more general: c is allowed to be different for the positive and the negative branch of k, and ν is only required to be smaller than 2. Our more restrictive condition ensures that

$$\int_{|x|\leq 1} |x|K(dx) < \infty, \tag{1.5.6}$$

which means that X is a finite variation process. This assumptions is needed in Section 3, but apart from this it can be empirically justified by arguments in Carr et al. (2000a). If we have even $\nu < 0$ it is a compound Poisson process with a finite number of expected jumps on each bounded interval.
The property of a finite variation makes it possible to choose the truncation function $h(x) \equiv 0$. If so, then the drift parameter b equals $\mu - \int_{-\infty}^{\infty} xK(dx)$ with $\mu \in \mathbb{R}$ where $\mu := E[X_1]$. The expectation of X_1 exists because we have $\int_{|x|>1} |x|K(dx) < \infty$, from which the existence follows by Lemma 1.17.
The tempered stable Lévy process has the advantage of having quite a simple Lévy measure which is easy to work with. For large $|x|$ the parameters λ_+ and λ_- govern the question of K-integrability of certain functions whereas ν determines the behaviour of K and hence the integrability question for small $|x|$.

1.5.3 Jump-diffusion processes and Kou's model

This section introduces jump-diffusion processes as an important subclass of the set of Lévy processes. More exactly, they correspond to Lévy processes with a finite Lévy measure. The results and definitions in this section are taken from Cont and Tankov (2004b).

Definition 1.23. *A compound Poisson process with intensity $\lambda > 0$ and jump size distribution f is a stochastic process X defined as*

$$X_t = \sum_{i=1}^{N_t} Y_i$$

where the jump sizes Y_i are independently and identically distributed with distribution f, and N_t is a Poisson process with intensity λ, independent from Y_i, $i \geq 1$.

As in (1.3.1) to a compound Poisson process X_t one can assign a random measure μ^X on $[0,T] \times \mathbb{R}^n$ which contains the jump structure of X. The link to the Lévy measure K is given by the following proposition:

[20] See e.g. Cont and Tankov (2004b).

Proposition 1.24. *Let X be a compound Poisson process with intensity λ and jump size distribution f. Its jump measure μ^X is a Poisson random measure on $\mathbb{R}^n \times [0, \infty)$ with intensity measure*

$$\nu(dt, dx) = K(dx)dt = \lambda f(dx)dt$$

Proof. Cont and Tankov (2004b), p. 75. □

This means that the Lévy measure is finite with mass λ and can therefore be represented as the intensity λ of the underlying Poisson process N_t times the jump size probability distribution $f(dx)$. This is why every compound Poisson process can be written as the sum of its jumps by using the random measure $\mu^X(dt, dx)$ of Proposition 1.24:

$$X_t = \sum_{s \in [0,t]} \Delta X_s \mathbf{1}_{\Delta X_t \neq 0} = \int_{[0,t] \times \mathbb{R}^n} x \mu^X(ds, dx) = x * \mu_t^X.$$

Hence X is not only a finite variation process but also has finite activity, which means that the above sum has almost surely only a finite number of terms.
A jump-diffusion process is the sum of a scaled Wiener process with drift and a compound Poisson process. Such processes are used to describe the random fluctuation of stock price processes where the diffusion part is responsible for modelling the small random fluctuations of the price process whereas the jumps represent extreme events that occur rarely. The latter are responsible for introducing skewness and kurtosis in the return distribution. Put differently, jump-diffusion processes are precisely the Lévy processes with a non-zero diffusion part and a finite non-zero Lévy measure.

Kou (2002) and Kou (1999) have introduced a model for one-dimensional equity price movement, henceforward called *Kou's model*, which possesses some nice features. They assume the stock price process to be given by

$$\frac{dS_t}{S_{t-}} = bdt + \sqrt{c}dW_t + d\left(\sum_{i=1}^{N_t}(V_i - 1)\right) \tag{1.5.7}$$

and $S_0 > 0$ where the random variables $\log V_i$ are distributed with density function

$$f(x) = p\lambda_+ e^{-\lambda_+ x}\mathbf{1}_{\{x>0\}} + (1-p)\lambda_- e^{-\lambda_- |x|}\mathbf{1}_{\{x<0\}}$$

for $0 \leq p \leq 1$, $\lambda_+, \lambda_- > 0$, and N_t is a Poisson process with intensity λ. According to Theorem 1.15 the solution is the stochastic exponential of the term on the right hand side of (1.5.7), i.e.

$$\begin{aligned} S_t &= S_0 e^{(b-\frac{1}{2}c)t + \sqrt{c}W_t} \prod_{i=1}^{N_t} V_i \\ &= S_0 \exp\left((b - \frac{1}{2}c)t + \sqrt{c}W_t + \sum_{i=1}^{N_t} \log V_i\right). \end{aligned} \tag{1.5.8}$$

It follows that Kou's model can be moulded into the form of an exponential Lévy model. Skipping the variance correction of the deterministic part of the argument of the exponential function in (1.5.8), we obtain a slightly different version of Kou's model: Defining the Lévy process $X = (b, c, K)_{\mathbf{P}}$, i.e.

$$X_t = bt + \sqrt{c}W_t + \sum_{s \le t} \Delta X_s,$$

where $b \in \mathbb{R}, c > 0$ and $K(dx) = \lambda f(x)dx$, the equity price model is given by $S_t = S_0 e^{X_t}$. Hence we obtain as the characteristic function of X_t

$$E[e^{izX_t}] = e^{t\psi(z)}$$

where

$$\begin{aligned} \psi(z) &= ibz - \frac{c}{2}z^2 + \int_{-\infty}^{\infty} (e^{izx} - 1)K(dx) \\ &= ibz - \frac{c}{2}z^2 + i\lambda z \left(\frac{p}{\lambda_+ - iz} - \frac{1-p}{\lambda_- + iz} \right) \end{aligned}$$

which exists for $-\lambda_+ < \Im(z) < \lambda_-$. This is the model which will be referred to when we are concerned about Kou's model in Chapter 4.
The jump size distribution is particularly suited for Monte-Carlo simulations. Its distribution function is equal to

$$F(x) = \int_{-\infty}^{x} f(\xi)d\xi = \begin{cases} (1-p)e^{\lambda_- x}, & , \quad x < 0 \\ 1 - pe^{-\lambda_+ x}, & , \quad x \ge 0 \end{cases} \tag{1.5.9}$$

and can be analytically inverted to yield

$$F^{-1}(y) = \begin{cases} -\frac{1}{\lambda_-} \log \frac{1-p}{y}, & , \quad y < 1-p \\ \frac{1}{\lambda_+} \log \frac{p}{1-y}, & , \quad y \ge 1-p \end{cases} \tag{1.5.10}$$

for $y \in [0, 1]$.

1.6 Fourier inversion and Fast Fourier transform

Pricing of path-independent options in the exponential Lévy model can take advantage of the technique of Fourier inversion and Fast Fourier transform. Given a one-dimensional asset price process S and a Lévy process X such that $S_t = S_0 e^{X_t}$ as well as a payoff function $w(S_T)$ there are essentially two prerequisites for a practical application of this method: Both the characteristic function χ of X under a martingale measure and the two-sided Laplace transform[21] φ_v of the modified payoff function $v(x) := w(e^{-x})$ must be available in an analytically tractable form, and

[21] See Doetsch (1950).

they must fulfil $\chi(iR) < \infty$ and $\varphi_v(R) < \infty$ for a constant $R \in \mathbb{R}$. The algorithm which we present here is essentially the one given in Raible (2000), p.61 ff.

First, we evaluate the price of a European call option with strike price $K = 1$ to which the case of a general K can be traced back to. For $K = 1$ we have $v(x) = (e^{-x}-1)^+$ and $\varphi_v(z) = \frac{1}{z(z+1)}$, where the latter equality is true for $\Re(z) < -1$. Then the price $V(\zeta)$ at time 0 of the option as a function of the negative log forward price $\zeta := -\log(e^{rT}S_0)$ of the underlying is given by[22]

$$V(\zeta) = \frac{e^{\zeta R - rT}}{2\pi} \int_{-\infty}^{\infty} e^{iu\zeta} \underbrace{\varphi_v(R+iu)\chi^{(r)}(iR-u)}_{=:g(u)} du, \tag{1.6.1}$$

where the function $\chi^{(r)}(z) := e^{-iurT}\chi(z)$ is the modified characteristic function, which takes into account slightly different definitions of the exponential Lévy model in Raible (2000) and this thesis[23].

The Fourier transformation in (1.6.1) can be calculated by approximation via the discrete Fourier transform and evaluation of the latter by means of the Fast Fourier Transform algorithm (FFT). Hence we have for a sufficiently large $N \in \mathbb{N}$, which should be a power of 2, and a small step size $\Delta u > 0$ the integral approximation[24]

$$V(\zeta) \approx \frac{e^{\zeta R - rT}}{\pi} \Delta u \Re \underbrace{\left(\sum_{n=0}^{N-1} e^{in\Delta u \zeta} g_n \right)}_{=:G(\zeta)} \tag{1.6.2}$$

with $g_0 := g(0)/2$ and $g_n := g(n\Delta u)$ for n $\in \{1, \ldots, N-1\}$. The imaginary part of the Fourier transformation in (1.6.1) drops out because $g(-u) = \overline{g(u)}$ for all u for which $g(u)$ is defined.

The FFT computes quite efficiently the sum $G(\zeta)$ on the right hand side of (1.6.2) for an equally spaced number of values $\zeta_k = k\Delta\zeta$ simultaneously. The step size $\Delta\zeta$ induced by the choice of N and Δu is given by $\Delta\zeta = \frac{2\pi}{N\Delta u}$. The index k runs over all $k \in \{0, \ldots, N-1\}$ and thus covers a complete period of the periodic function $k \to e^{2\pi ink/N}$. This and the observation that the FFT gives the best approximation for $k \approx 0$ and $k \approx N-1$ leads to the formula

$$V(\zeta_k) \approx \frac{e^{\zeta_k R - rT}}{\pi} \Delta u \Re(G(\zeta_k)) := p_k, \tag{1.6.3}$$

this time for $k \in \{-N/2, \ldots, N/2\}$ where the region of best approximation is around $k = 0$.

For a general strike price $K > 0$ the pricing function $V^K(\zeta)$ of an option can be

[22]Raible (2000), Theorem 3.2

[23]See Raible (2000), Appendix B.2

[24]Raible(2000), p.70.

traced back to $V^1 \equiv V$ via[25]

$$V^K(\zeta) = K\, V(\zeta + \log K). \tag{1.6.4}$$

Finally, the last task to obtain the price at $t = 0$ of a European call option with maturity T and strike price K whose underlying is worth S_0 in $t = 0$ is to evaluate (1.6.3) for a suitable k. Given (1.6.4) we have

$$k\Delta\zeta = \zeta + \log K = \log \frac{K}{e^{rT} S_0}. \tag{1.6.5}$$

The last term is the logarithm of the ratio of the strike price and the forward price of the underlying. Clearly, its interesting values lie around zero, and we are in the region of best approximation of V. (1.6.5) yields directly[26]

$$k^* = \frac{1}{\Delta\zeta} \log \frac{K}{e^{rT} S_0} \quad \text{and} \quad k^u = \lceil k^* \rceil \quad \text{and} \quad k^d = \lfloor k^* \rfloor .$$

The approximate option price as a function of the negative log forward price $\zeta = -\log(e^{rT} S_0)$ of the underlying is then calculated by linear interpolation, and we have

$$V^K(\zeta) \approx p_{k^d} + \frac{k^* - k^d}{k^u - k^d}(p_{k^u} - p_{k^d}). \tag{1.6.6}$$

The accuracy of the above algorithm can be checked by comparing its results with the ones from the Black-Scholes formula for the case where the Lévy process is actually a Wiener process with drift. Fig. 1.1 depicts the resulting absolute pricing error in a typical parameter setting as a function of the moneyness, defined as the K/S_0, and time to maturity in days. We see that in this example the FFT algorithm computes option prices with an error of less than $3 * 10^{-6}$, thus yielding quite a satisfactory approximation.

1.7 Lévy Copulas

A (probabilistic) copula is a real-valued function of two or more variables which is used to link one-dimensional probability laws in order to construct a multivariate probability distribution function on $\mathbb{R}^n$ with prescribed marginals. Plainly spoken, the new probability distribution function is given by plugging the marginal distribution functions into the copula. For an in-depth treatment of copulas we refer to Nelsen (1999).

The very same principle can be used in order to build a multidimensional Lévy measure out of one-dimensional Lévy measures, where the copulas in this case are called *Lévy copulas*. The following two paragraphs summarize the results in Tankov (2003) and Cont and Tankov (2004b) which will be of importance in Chapter 4.

[25] Raible (2000), Lemma 3.3.

[26] $\lceil x \rceil := \min\{n \in \mathbb{Z} | n \geq x\}$, $\lfloor x \rfloor := \max\{n \in \mathbb{Z} | n \leq x\}$.

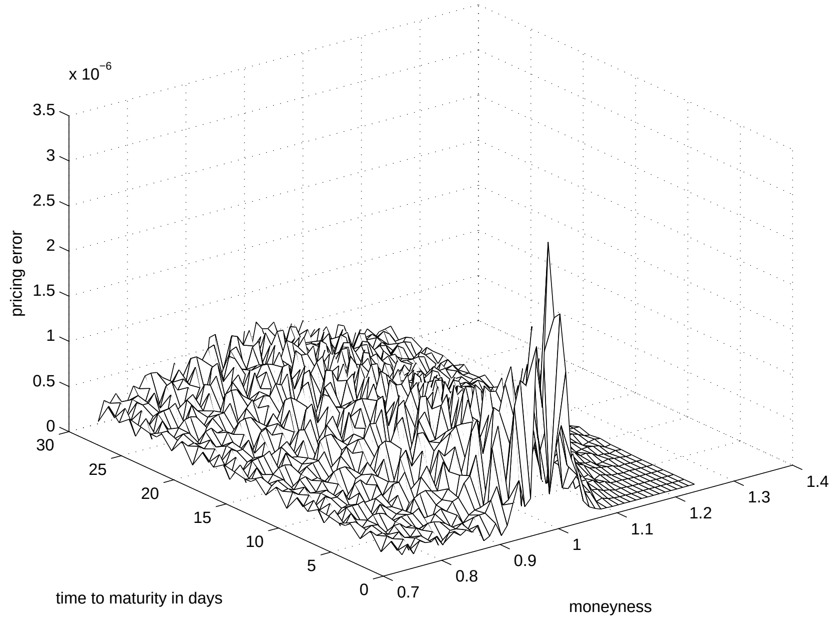


Fig. 1.1. Pricing error of FFT method for $S_0 = 40$, $\sigma = 0.026$ and $r = 4\%$ p.a. with parameters $N = 2^{18} = 262144$, $R = -25$ and $\Delta u = 0.12$ as a function of the moneyness ($= K/S_0$) and time to maturity.

The most striking difference between the two concepts lies in the domain of definition. While a copula is defined on $[0,1]^n$, i.e. the Cartesian product of the ranges of the one-dimensional distribution functions, the domain of definition of a Lévy copula must be somewhat different. This comes from the observation that the role of the distribution function in the theory of probabilistic copulas is now played by the tail integral which is not well-defined everywhere on the real axis: Lévy measures in general are not integrable in a neighbourhood of zero, so in the first place one works only with tail integrals of Lévy measures on the positive real line, i.e. one is up to this point only concerned about the dependence structure of positive jumps. But even with this simplification the non-integrability forces a Lévy copula to be defined on $[0,\infty]^n$. Later on this concept will be extended to the modelling of the dependence of positive *and* negative jumps.

1.7.1 Positive Lévy copulas

Tail integrals and Lévy measure

A Lévy copula is a continuous n-place real function C with certain properties to be made precise below. The domain and the range of C are denoted by Dom$C$ and Ran$C$. In the following we suppose that Dom$C$ is given by $S^1 \times S^2 \times \cdots \times S^n$ where each S^k has a smallest and a greatest element $\underline{a}^k = \min S^k$ and $\bar{b}^k = \max S^k, k =$

$1, \dots n$. An *n-box* $B = [a, b]$ is defined as the set $[a^1, b^1] \times [a^2, b^2] \times \dots \times [a^n, b^n]$ for $a = (a^1, \dots, a^n)$ and $b = (b^1, \dots, b^n)$. We begin this section with a number of definitions which will eventually pin down the notion of a Lévy copula. The symbol $\bar{\mathbb{R}}$ denotes the extended real line $\mathbb{R} \cup \{\infty\} \cup \{-\infty\}$.

Definition 1.25. *Let $S^1, S^2, \dots, S^n$ be nonempty subsets of $\bar{\mathbb{R}}$, and let C be an n-place real function such that* $\mathrm{Dom}C = S^1 \times S^2 \times \cdots \times S^n$. *Let $B = [a, b]$ be an n-box all of whose vertices are in* $\mathrm{Dom}C$. *Then the C-*volume *of B is given by*

$$V_C(B) = \sum \mathrm{sgn}(u) C(u)$$

where the sum is taken over all vertices u of B, and sgn*(u) is given by*

$$\mathrm{sgn}(u) = \begin{cases} 1, & , \text{ if } u^k = a^k \text{ for an even number of } k\text{'s} \\ -1, & , \text{ if } u^k = a^k \text{ for an odd number of } k\text{'s.} \end{cases} \tag{1.7.1}$$

Definition 1.26. *An n-place real-valued function C*

- *is n-increasing if $V_C(B) \geq 0$ for all n-boxes B whose vertices lie in* $\mathrm{Dom}C$.
- *is grounded if $C(u) = 0$ for all $u \in \mathrm{Dom}C$ such that $u^k = \underline{a}^k$ for at least one $k \in \{1, \dots, n\}$.*
- *has (one-dimensional) margins*

$$C_k(u) := C(\bar{b}^1, \dots, \bar{b}^{k-1}, u, \bar{b}^{k+1}, \dots, \bar{b}^n), \ u \in S^k.$$

  with $\mathrm{Dom}C_k = S^k$.

Before coming to the definition of a Lévy copula, we have to deal with the substitute of the distribution function in the world of probabilistic copulas: the tail integral and its connection to the Lévy measure.

Definition 1.27. *An n-dimensional tail integral is a function $U : [0, \infty]^n \to [0, \infty]$ such that*

a) $(-1)^n U$ *is a n-increasing function;*

b) U *is equal to zero if one of its arguments is equal to* ∞*;*

c) U *is finite everywhere except at zero and* $U(0, \dots, 0) = \infty$.

As already noted every n-dimensional Lévy measure K can be assigned a tail integral U^K by the following definition:

Definition 1.28. *Let K be a Lévy measure on $[0, \infty)^n \setminus \{0\}$. Its tail integral U^K is a function $[0, \infty]^n \to [0, \infty]$ such that*

a) U^K *is equal to zero if one of its arguments is equal to* ∞*;*

b) $U^K(0,\ldots,0)=\infty$*;*

c) For $(x^1,\ldots,x^n)\in[0,\infty)^n\setminus\{0\}$*.*

$$U^K(x^1,\ldots,x^n)=K([x^1,\infty)\times\ldots\times[x^n,\infty)).$$

It is important to note some properties of the above definition: Apart from its value at the origin, U^K is finite on $[0,\infty]$. This follows immediately from the fact that every Lévy measure integrates the constant in a domain which is bounded away from zero[27]. In addition we have that for every right-open left-closed interval I we obtain the identity

$$K(I)=(-1)^n V_{U^K}(I),$$

which connects the Lévy measure and the volume assigned to the tail integral U^K. Finally U^K integrates $|x|^2$ near 0 just because K does by definition. Note that U^K is a tail integral in the sense of Definition 1.28. Conversely, starting from a tail integral U one can recover the Lévy measure hidden behind it. Given a tail integral U one can recover the Lévy measure

Proposition 1.29. *Let U be a n-dimensional tail integral, left-continuous in each variable, which integrates $|x|^2$ in a neighbourhood of zero. Then there exists a unique Lévy measure K on $\mathcal{B}([0,\infty)^n\setminus\{0\})$ such that U is the tail integral of K.*

Proof. Tankov (2003), Corollary 3.1. □

Sklar's theorem and convergence

With the help of Definition 1.26 we can now define the notion of a Lévy copula as well as state a version of Sklar's theorem for Lévy measures:

Definition 1.30 (Lévy copula). *An n-dimensional* Lévy copula *is an n-increasing grounded function $C:[0,\infty]^n\to[0,\infty]$ with margins C_k, $k=1,\ldots,n$, which satisfy $C_k(u)=u$ for all u in $[0,\infty]$.*

Theorem 1.31 (Sklar's theorem for Lévy processes). *Let K be a Lévy measure on $[0,\infty)^n\setminus\{0\}$ with tail integral U^K and marginal Lévy measures $K_1,\ldots,K_n$. There exists a Lévy copula C such that*

$$U^K(x^1,\ldots,x^n)=C(U^{K_1}(x^1),\ldots,U^{K_n}(x^n)), \tag{1.7.2}$$

where $U^{K_1},\ldots,U^{K_n}$ are tail integrals of $K_1,\ldots,K_n$. If the marginal Lévy measures $K_1,\ldots,K_n$ are infinite and have no atoms, this copula is unique.
Conversely, if C is a Lévy copula and $K_1,\ldots,K_n$ are Lévy measures on $(0,\infty)$ with tail integrals $U^{K_1},\ldots,U^{K_n}$ then (1.7.2) defines a tail integral of a Lévy measure on $[0,\infty)^n\setminus\{0\}$.

[27] i.e. due to its càdlàg property every Lévy process has almost surely a finite number of jumps bigger than some arbitrary positive bound.

Proof. Tankov (2003), Theorem 3.1. □

For Lévy measures constructed from Lévy copulas weak convergence can be traced back to the Lévy copula, as the following proposition shows:

Proposition 1.32. *Let $(C^i)_{i\in\mathbb{N}}$ and C be Lévy copulas and $U_1, \dots U_n$ marginal tail integrals of Lévy measures. Set*

$$U^i(x^1, \dots, x^n) = C^i(U_1(x^1), \dots, U_n(x^n)),\ i \in \mathbb{N}$$

and

$$U(x^1, \dots, x^n) = C(U_1(x^1), \dots, U_n(x^n)),$$

and let K^i resp. K be the Lévy measures associated with U^i resp. U by Proposition 1.29. Then for $i \to \infty$ the Lévy measures K^i converge weakly to K, i.e.

$$\lim_{i\to\infty} \int_{\mathbb{R}^n} f(x)K^i(dx) = \int_{\mathbb{R}^n} f(x)K(dx)$$

for all bounded continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$, if and only if C^i converges pointwise to C.

Proof. This is a special case of Theorem 3.3 in Barndorff-Nielsen and Lindner (2004). □

Finally we can define the *support* of a Lévy copula in exactly the same way as it is defined for probabilistic copulas[28]: The *support* of a Lévy copula C is the complement of the union of all open sets A in $[0, \infty]^n$ such that $V_C(A) = 0$. The same applies to the definitions of absolute continuity and singularity of Lévy copulas.

Specific Lévy copulas

The two most important Lévy copulas are the independence and the complete dependence copula. A family of Lévy copulas that comprises both of them is called *comprehensive.*

Definition 1.33. *A subset S of $\bar{\mathbb{R}}^n$ is called* nondecreasing *if for every two vectors $v, u \in S$ either $v^k \le u^k\ \forall k$ or $v^k \ge u^k\ \forall k$. S is called* increasing *if for every two vectors $v, u \in S$, $v \ne u$ either $v^k < u^k\ \forall k$ or $v^k > u^k\ \forall k$.*

Definition 1.34. *Let $X = (X^1, \dots, X^n)$ be a Lévy process with only positive jumps. Its jumps are said to be* completely dependent *if there exists an increasing subset S of $\mathbb{R}^n_+$ such that $K(\bar{\mathbb{R}}^n_+ \setminus S) = 0$.*

Proposition 1.35 (Independence and complete dependence).
Let $X = (X^1, \dots, X^n)$ be a pure jump Lévy process with only positive jumps.

[28] See e.g. Nelsen (1999).

- *Its components are independent if and only if its Lévy copula (or one of them if there are many) has the form*

$$C_{\perp}(u^1,\dots,u^n) = u^1\mathbf{1}_{\{u^2=\infty,\dots,u^n=\infty\}} + \dots + u^n\mathbf{1}_{\{u^1=\infty,\dots,u^{n-1}=\infty\}}.$$

  If the marginal Lévy measures of X are infinite and have no atoms, then $C_{\perp}$ is the unique Lévy copula of X.

- *Let X be supported by a non-decreasing set S. Then a possible Lévy copula of X is the complete dependence Lévy copula given by*

$$C_{||}(u^1,\dots,u^n) = \min(x^1,\dots,x^n).$$

  Conversely, if the Lévy copula of X is given by $C_{||}$ then the Lévy measure of X is supported by a non-decreasing set. If, in addition, the tail integrals of components of X are continuous, then the jumps of X are completely dependent.

Proof. Tankov (2003), Propositions 4.2 and 4.3. □

The complete dependence Lévy copula and the complete dependence probabilistic copula are formally the same (with different domains, though) whereas the corresponding copulas linking independent components look different. This has the plain consequence that in order to construct a comprehensive family Lévy copulas one cannot make use of a redefinition of a family of probabilistic copulas, and one is forced to pursue a very different path. This will be seen in Chapter 4.

1.7.2 General Lévy copulas

For the rest of this section we confine ourselves to the case $n = 2$. Lévy copulas on $\mathbb{R}^2$ that deal with the dependence structure of both positive and negative jumps are called *general Lévy copulas* or simply Lévy copulas.

Definition 1.36. $F(x,y) : [-\infty,\infty]^2 \to [-\infty,\infty]$ *is a (general) Lévy copula if it has the following three properties:*

a) F *is 2-increasing.*

b) $F(0,x) = F(x,0) = 0 \quad \forall x \in \bar{\mathbb{R}}$.

c) $F(x,\infty) - F(x,-\infty) = F(\infty,x) - F(-\infty,x) = x$.

Tail integrals for Lévy measures on $\mathbb{R}^n$ are somewhat more cumbersomely to define because one has to work around the possible singularity at zero. We deal with the cases $n = 1$ and $n = 2$.

Definition 1.37. *Let K be a Lévy measure. This measure has two tail integrals, $U^+ : [0,\infty] \to [0,\infty]$ for the positive part and $U^- : [-\infty,0] \to [-\infty,0]$ for the negative part, defined as follows:*

$$\begin{aligned} U^+(x) &= K([x,\infty)) \quad \text{for } x \in (0,\infty),\ U^+(0)=\infty,\ U^+(\infty)=0, \\ U^-(x) &= -K([-\infty,x)) \quad \text{for } x \in (-\infty,0),\ U^-(0)=-\infty,\ U^+(-\infty)=0. \end{aligned}$$

Let K be a Lévy measure on $\mathbb{R}^2$ with marginal tail integrals U_1^+, U_1^-, U_2^+ and U_2^-. This measure has four tail integrals U^{++}, U^{+-}, U^{-+} and U^{--}, where each tail integral is defined on its respective quadrant, including the coordinate axes, as follows:

$$\begin{array}{rll} U^{++}(x,y) = & K([x,\infty)\times[y,\infty)), & \text{if } x\in(0,\infty) \text{ and } y\in(0,\infty),\\ U^{+-}(x,y) = & -K([x,\infty)\times[-\infty,y)), & \text{if } x\in(0,\infty) \text{ and } y\in(-\infty,0),\\ U^{-+}(x,y) = & -K([-\infty,x)\times[y,\infty)), & \text{if } x\in(-\infty,0) \text{ and } y\in(0,\infty),\\ U^{--}(x,y) = & K([-\infty,x)\times[-\infty,y)), & \text{if } x\in(-\infty,0) \text{ and } y\in(-\infty,0). \end{array}$$

If x or y is equal to $+\infty$ or $-\infty$, the corresponding tail integral is zero and if x or y is equal to zero, the tail integrals satisfy the following conditions:

$$\begin{aligned} U^{++}(x,0) - U^{+-}(x,0) &= U_1^+(x)\\ U^{-+}(x,0) - U^{--}(x,0) &= U_1^-(x)\\ U^{++}(0,y) - U^{-+}(0,y) &= U_2^+(y)\\ U^{+-}(0,y) - U^{--}(0,y) &= U_2^-(y). \end{aligned}$$

Based on the theory of positive Lévy copulas, given a two-dimensional Lévy measure K, we can write down the four tail integrals for every quadrant of the plane. This amounts to nothing but Sklar's theorem for Lévy processes with positive and negative jumps:

Theorem 1.38. *Let K be a Lévy measure on $\mathbb{R}^2$ with marginal tail integrals U_1^+, U_1^-, U_2^+ and U_2^-. There exists a Lévy copula C such that U^{++}, U^{+-}, U^{-+} and U^{--} are tail integrals of K where*

$$\begin{array}{rll} U^{++}(x,y) = & C(U_1^+(x),U_2^+(y)), & \text{if } x\geq 0 \text{ and } y\geq 0,\\ U^{+-}(x,y) = & C(U_1^+(x),U_2^-(y)), & \text{if } x\geq 0 \text{ and } y\leq 0,\\ U^{-+}(x,y) = & C(U_1^-(x),U_2^+(y)), & \text{if } x\leq 0 \text{ and } y\geq 0,\\ U^{--}(x,y) = & C(U_1^-(x),U_2^-(y)), & \text{if } x\leq 0 \text{ and } y\leq 0. \end{array}$$

If the marginal tail integrals are absolutely continuous, and K does not charge the coordinate axes, the Lévy copula is unique.
Conversely, if C is a Lévy copula and U_1^+, U_1^-, U_2^+ and U_2^- are tail integrals of one-dimensional Lévy measures then the above formulas define a set of tail integrals of a Lévy measure.

Proof. Cont and Tankov (2004b), Theorem 5.7. □

One of the simplest non-trivial Lévy copulas is the following one:

Example 1.39. If C^+ and C^- are positive Lévy copulas, the following function C on $\mathbb{R}^2$ clearly satisfies Definition 1.37.

$$C(u,v) := C^+(|u|,|v|)\mathbf{1}_{\{u\geq 0, v\geq 0\}} + C^-(|u|,|v|)\mathbf{1}_{\{u\leq 0, v\leq 0\}} \tag{1.7.3}$$

But contrary to the general case it prescribes a zero jump intensity for jumps which, for the two components of a Lévy process, go in reverse directions.

Chapter 2

A class of tractable martingale measures

2.1 Introduction

In this section we concentrate entirely on one-dimensional Lévy processes. We assume to be given a normal inverse Gaussian (NIG) Lévy process X on the given probability space $(\Omega, \mathcal{F}, \mathbf{P})$. This class of Lévy processes has been introduced in Section 1.5.1. Given X, the equity price process is given by

$$S_t = S_0 e^{X_t}, 0 \leq t \leq T,$$

and the risk-free security evolves according to

$$B_t = e^{rt}, 0 \leq t \leq T.$$

Theorems 1.20 and 1.22 give the foundation of a closer look at possible types of measure changes and reveal that the stochastic problem of choosing a change of measure can be translated into a deterministic one if we assume that X is a Lévy process under both $\mathbf{P}$ and the risk-neutral measure $\mathbf{Q}$ to be determined. It suffices to select a positive measurable function y on the real line with one or more parameters that satisfies conditions (1.4.6) and (1.4.7) and is able to fit its parameters such that the martingale conditon (2.2.6) is also met.
We now give a survey of commonly used martingale measure for the exponential Lévy model of Section 1.4, and we stick to the y-centred point of view that we have developed in Chapter 1. The involved measure change functions each have one unknown parameter which is pinned down by the martingale condition in Theorem 1.22. In each case the letter θ is used for this parameter. In order to simplify the representation we confine ourselves to pure jump Lévy processes.
The measure $\mathbf{Q}$ which is obtained from $\mathbf{P}$ via the *Esscher* transform is given via the transformation

$$\frac{d\mathbf{Q}}{d\mathbf{P}} = \frac{\exp(\theta X_T)}{E^{\mathbf{P}}[\exp(\theta X_T)]} = e^{\theta X_T - T\psi(-i\theta)}$$

where ψ is the cumulant function of X_1. The corresponding measure change function is given by[1]

$$y(x) = e^{\theta x}.$$

Keller (1997), Section 1.4.3., provides a justification for the Esscher transform in terms of economic theory. It arises as the change of measure which is induced by a general equilibrium representative agent model with a constant-relative-risk-aversion utility function with $-\theta$ as coefficient of relative risk aversion, where θ is the Esscher parameter. A somewhat stronger case for the Esscher transform is made by the argument that it is the simplest change of measure in a certain sense: Starting from the understanding that y must be strictly positive on the real axis one sets $y(x) = e^{f(x)}$ for an arbitrary differentiable function f and obtains the Esscher transform as the first-order Taylor approximation of f under the assumption of $y(0) = 1$. This assumption must be fulfilled for infinite Lévy measures due to (1.4.6). This vindication had been put forward by Madan and Milne (1991) even before the name 'Esscher transform' appeared. The Esscher measure is a very comfortable change of measure from the computational point of view: There is a large class of Lévy processes including the normal inverse Gaussian Lévy process, which are closed under an Esscher change of measure, i.e. starting from a certain class of Lévy processes under the measure $\mathbf{P}$, the Lévy process under the new measure $\mathbf{Q}$, which is obtained through the Esscher transform, is of the same class.

The *minimal martingale measure*, introduced in Föllmer and Schweizer (1991), is of some interest because it supports a hedging strategy with a minimal exposure to hedging error in a certain sense. We will clarify this notion in Chapter 4. Unfortunately the minimal martingal measure defines a signed measure in many cases depending on the type of Lévy process chosen and on its parameters. The prices calculated by using the minimal martingale measures thus introduce arbitrage possibilites. For a constant θ the corresponding measure change process for an exponential Lévy model is given by[2]

$$Z_t = \mathcal{E}\left(-\int_0^{\cdot} \frac{\theta}{S_{u-}} dM_u\right)_t$$

where $M := [S_-(e^x - 1)] * (\mu^X - \nu^{\mathbf{P}})$ is the martingal part of S. Comparing coefficients with representation (2.2.5) gives

$$y(x) = 1 - \theta(e^x - 1),$$

which is clearly not identically bigger than zero.

The *minimal entropy martingale measure* has been investigated e.g. by Chan (1999), Miyahara (1999) and Miyahara (2001). It is defined as the measure $\mathbf{Q}$ which minimizes the relative entropy $\int_\Omega \frac{d\mathbf{Q}}{d\mathbf{P}} \log \frac{d\mathbf{Q}}{d\mathbf{P}}\, d\mathbf{P}$ of $\mathbf{Q}$ with respect to $\mathbf{P}$. Both Chan (1999) and Cont and Tankov (2004b) show that in this case

$$y(x) = e^{\theta(e^x - 1)}.$$

[1] See Keller (1997), Lemma 21.

[2] See e.g. Schweizer (1995).

It is worthwhile to mention that for the stochastic exponential stock price model (1.4.1) the minimal entropy and the Esscher measure coincide[3].
There are some further measures in the context of utility maximization which yield e.g.

$$y(x) = e^{-p\theta(e^x-1)}, \tag{2.1.1}$$

which is derived from maximizing the CARA[4] utility function $u(x) = 1 - \frac{1}{p}e^{-px}$ in a consumption-based representative agent model[5]. However, this change of measure shows a sort of trade-off between θ and the risk-aversion parameter p of the utility function: The measure change function (2.1.1) does not change if p is modified because - assuming a unique solution of the martingale condition for fixed p - θ changes accordingly; shortly, θ and p cannot be identified through the martingale condition.

What all these measures have in common is a sort of asymmetric behaviour: They are monotone, i.e. they either rise or fall on the entire real line. By a separate estimation of $\mathbf{Q}$ through the technique of statistical martingale measures and $\mathbf{P}$ Carr et al. (2000a) investigate the shape of the measure change function and find a somewhat different shape. Their functions y are minimal at zero and show a U-shape. The steepness on both sides is interpreted as a kind of risk aversion assigned to positive and negative jumps. For brevity we say that monotone functions have an *asymmetric* shape, while the U-shape is *symmetric*. Carr et al. (2000a) also give a theoretical founding of this specific shape: They use that a Lévy measure can be computed as the limit of the $t-$member of the corresponding convolution semigroup divided by t as t goes to zero[6]. Based on the fact that in a representative agent model there is no demand for options in contrast to a heterogeneous agent model, they derive the asymmetric shape for the former model and the symmetric one for the latter. Although this reasoning is intuitive rather than being derived from a rigorous theoretical model, it serves as a starting point in order to construct symmetric measure change functions in the framework of an absolutely continuous change of measure.
With all this preparatory work we can adumbrate the line of thought in the following key part of this chapter:

- We examine to what extent the shape of y changes the pricing behaviour of the correspondingly defined martingale measure. This means that we will give a parametric class of martingale measures which incorporate both shapes covered in the motivation above, and we will discuss the results.

- Independently from the first item, this new class of measures will provide an alternative to traditional pricing with the Esscher measure. The idea is similar

[3]See Esche (2004).
[4]CARA=constant absolute risk aversion.
[5]See Kallsen (2000). For some further measures derived from distance-minimization criteria see Goll and Rüschendorf (2001)
[6]See Barndorff-Nielsen (2000).

to the one of statistical martingale measures: There are several Esscher-type martingale measures, and the exchange-traded option prices can decide which one fits best the current market situation. This decision can be made for example by a comparison of the implied volatility smile of these models with the empirical volatility smile. The advantage of this procedure over the use of statistical martingale measures is that, based on an estimation of $\mathbf{P}$, arbitrage-free prices via an absolutely continuous change of measure will be obtained.
A further advantage is more intricate. Statistical martingale measures have an infinite number of degrees of freedom, i.e. they depend very much on the subtleties of market prices, especially on mispricing because of low liquidity, and on the chosen distance measure for minimization. On the contrary, it could be more meaningful to have only a finite (and very low) number of measures from which one measure is chosen according to the general market trend. This specific measure yields prices which should be more robust to sporadic mispricing tendencies in the options market.

2.2 Martingale measures for NIG Lévy processes

2.2.1 Change of measure

Let X be an NIG Lévy process with Lévy measure

$$K(dx) = \frac{\delta\alpha}{\pi} e^{\beta x} |x|^{-1} K_1(\alpha|x|)dx \tag{2.2.1}$$

with the parameters and the Bessel function specified in Section 1.5.1. The moment generating function (1.5.3) exists in an open interval containing zero, hence due to Lemma 1.16 all moments of an NIG distribution exist, especially the first one, which by differentiation of the moment generating function is equal to

$$\mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}}. \tag{2.2.2}$$

Due to Lemma 1.17 applied for $\alpha = 1$ and the discussion after Theorem 1.14 we can choose the truncation function $h(x) \equiv x$. Taking into account Corollary 1.18 we can from now on work with the Lévy triplet $(b, 0, K)_P$ where the Lévy measure is given by (1.5.4) and

$$b = E^{\mathbf{P}}[X_1] = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}}. \tag{2.2.3}$$

Moreover, it is a pure jump Lévy process such that the diffusion component c is equal to zero. All told, this leads to the cumulant function

$$\psi(z) = ibz + \int_{-\infty}^{\infty} (e^{izx} - 1 - izx) K(dx) \tag{2.2.4}$$

of the NIG process $X_t = bt + x * (\mu^X - \nu^{\mathbf{P}})_t$.
As the process moves only by jumps, the measure change process Z is completely given by the measure change function y and is equal to

$$Z = \mathcal{E}\left((y(x) - 1) * (\mu^X - \nu^{\mathbf{P}})\right). \tag{2.2.5}$$

The martingale condition (1.4.9) looks as follows:

$$b - r + \int_{-\infty}^{\infty} \{(e^x - 1)y(x) - x\} K(dx) = 0. \tag{2.2.6}$$

Owing to the special role of the parameter β in the NIG density function, which accounts for the asymmetry of the Lévy measure, we sometimes denote the Lévy measure by K^β instead of K.

In order to apply the FFT method of Section 1.6 it is necessary to have a simple analytic form of the characteristic function of the random variable X_1 under the risk-neutral measure. For the Esscher change of measure the characteristic function is again an NIG Lévy process with different parameter β. For other cases the Lévy-Khinchin formula in Theorem 1.14 is available. The convenient applicability of the last approach, though theoretically always possible, depends for the most part on the selected change of measure. This is because of the integral in the Lévy-Khinchin representation, for which numerical evaluation is difficult due to the fact that, first, this is an indefinite integral over $\mathbb{R}$ and, second, the Lévy measure has a singularity at zero. This is all the more important because the characteristic function, thus the integral, has to be evaluated many times during the computation of the Fast Fourier transform. All this means that a simple form of the risk-neutral characteristic function is required. Using Theorem 1.20, equation (2.2.4) gives the cumulant function under $\mathbf{Q}$

$$\begin{aligned}
\psi^{\mathbf{Q}}(z) &= \\
&= ib'z + \int_{\mathbb{R}} (e^{izx} - 1 - izx) K'(dx) \\
&= ibz + iz \int_{\mathbb{R}} x(y(x) - 1)\, K(dx) + \int_{\mathbb{R}} (e^{izx} - 1 - izx) y(x) K(dx).
\end{aligned} \tag{2.2.7}$$

The task of this section is to carefully construct the change of measure in such a way that the integral can be calculated analytically. Hence we now define a set of probability measures by giving the function y that fully determines the change of measure starting from an underlying probability measure $\mathbf{P}$.

Definition 2.1. *A measure change function y is called* flexible *if for*

- $\Gamma = (\gamma_1, \ldots, \gamma_n, \bar{\gamma}_1, \ldots, \bar{\gamma}_m)'$, $\gamma_i, \bar{\gamma}_i \geq 2$ *and*
- $\Theta = (\theta_1, \ldots, \theta_m, \bar{\theta}_1, \ldots, \bar{\theta}_m)'$, $-\alpha - \beta < \theta_i < \alpha - \beta - 1$, $-\alpha + \beta + 1 < \bar{\theta}_j < \alpha + \beta$,

$i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}$ we have

$$y(x) - 1 := \begin{cases} \sum_{i=1}^{n} x^{\gamma_i} e^{\theta_i x} & , \quad x \geq 0 \\ \sum_{j=1}^{m} |x|^{\bar{\gamma}_j} e^{\bar{\theta}_j |x|} & , \quad x < 0. \end{cases} \tag{2.2.8}$$

A measure $\mathbf{Q}$ *which is constructed from a measure* $\mathbf{P}$ *by the measure change process (2.2.5) with a flexible measure change function* y *is called* flexible *measure.*

In order to simplify the notation, whenever we talk about the parameters γ resp. θ we actually mean $\gamma_i, \bar{\gamma}_j$ resp. $\theta_i, \bar{\theta}_j$ for $i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}$. For further reference we define the functions y_1 and y_2 according to $y(x) - 1 \equiv y_1(x)\mathbf{1}_{\{x\geq 0\}} + y_2(x)\mathbf{1}_{\{x<0\}}$.

It is worthwhile to make some remarks. The term x^γ that appears in y prevents the measure change function $y \equiv 1$, which corresponds to the trival change of measure, from being embedded in the class of flexible measure change functions. This means that we implicitly assume that under the statistical martingale measure $\mathbf{P}$ the Lévy process X is not a martingale.

Nevertheless, the factor x^γ is important because it allows for an analytical evaluation of the integrals with respect to K which appear in the Lévy-Khinchin representation. The measure $x^\gamma K(dx)$ is finite, and this is precisely the reason why some compound integrals that appear in the computations below can be written as a sum of simpler integrals.

A last remark must be made as to how the term x^γ influences the shape of y. For large $|x|$ the exponential part of a flexible measure change function is relevant for integrability for $|x| \to \infty$. On the contrary, for small $|x|$ the polynomial part is essential. Its local behaviour at zero is $y(x) - 1 = O(x^\gamma)$ for $\gamma \geq 2$ whereas for the Esscher measure we have $y(x) - 1 = O(x)$. The following lemma is needed in the sequel:

Lemma 2.2. *For $a, b \geq 0$ we have*

$$(1 - \sqrt{1+a+b})^2 \leq 2(1 - \sqrt{1+a})^2 + 2(1 - \sqrt{1+b})^2. \tag{2.2.9}$$

Proof. The statement is equivalent to

$$f(a,b) := 6 + a + b + 2\sqrt{1+a+b} - 4\sqrt{1+a} - 4\sqrt{1+b} \ \geq 0$$

which by substituting $x := \sqrt{a+1}$ and $y := \sqrt{b+1}$ becomes

$$\left(\sqrt{x^2+y^2-1} + 1\right)^2 - 4(x+y-1) \geq 0.$$

For the statement to be proved it is sufficient to show that the function $h(y) := \sqrt{x^2+y^2-1} + 1 - 2\sqrt{x+y-1}$ is non-negative for $y \geq x \geq 1$ with fixed x. We have $h(1) = |x| + 1 - 2\sqrt{x} = (\sqrt{x}-1)^2 \geq 0$ as $x \geq 1$, and for $y \geq x \geq 1$ we have by trivial transformations that

$$h'(y) = \frac{y}{\sqrt{x^2+y^2-1}} - \frac{1}{\sqrt{x+y-1}} \geq 0$$

if and only if

$$y^2(x+y) - 2y^2 - x^2 + 1 \geq 0.$$

But this is true because

$$y^2(x+y) - 2y^2 - x^2 + 1 \geq 2y^2x - 2y^2 - (x^2-1) \geq (2y^2 - (x+1))(x-1) \geq 0$$

for $y \geq x \geq 1$. □

Proposition 2.3. *Every flexible measure change function y determines an absolutely continuous change of measure, i.e. it is positive and satisfies*

$$\int_{-\infty}^{\infty} (\sqrt{y(x)} - 1)^2 K(dx) < \infty. \tag{2.2.10}$$

In addition, we have

$$\int_{1}^{\infty} (e^x - 1) y(x) K(dx) < \infty. \tag{2.2.11}$$

Proof. For the first condition it suffices to consider the case $n = m = 1$ thanks to Lemma 2.2 and a simple induction argument over n. The index 1 of the parameters will be dropped in the proof. We have

$$\int_0^1 \left(\sqrt{1 + x^{\gamma} e^{\theta x}} - 1\right)^2 K(dx) \leq \int_0^1 \left(\sqrt{1 + x^2 e^{\theta x}} - 1\right)^2 K(dx) < \infty \tag{2.2.12}$$

because of $\left(\sqrt{1 + x^2 e^{\theta x}} - 1\right)^2 = \frac{x^2}{2} + O(x^3)$ and (1.3.3). By the very special form of K we have

$$e^{\theta x} K(dx) = e^{\theta x} K^{\beta}(dx) = \frac{\delta \alpha}{\pi} e^{\theta x} e^{\beta x} |x|^{-1} K_1(\alpha |x|)\, dx = K^{\beta+\theta}(dx).$$

Going back to the definition of an NIG distribution $K^{\beta+\theta}(dx)$ is again a Lévy measure of an NIG distribution (with asymmetry parameter $\beta+\theta$ instead of β) if $-\alpha < \beta+\theta < \alpha$. But this is exactly what is assumed in Definition 2.1. Hence

$$\int_1^{\infty} \left(\sqrt{1 + x^{\gamma} e^{\theta x}} - 1\right)^2 K(dx) < \int_1^{\infty} x^{\gamma} e^{\theta x} K(dx) = \int_1^{\infty} x^{\gamma} K^{\beta+\theta}(dx) < \infty. \tag{2.2.13}$$

The first inequality is due to $\sqrt{1+a} - 1 < \sqrt{a}$ for $a > 0$, and existence results from the fact that all moments of an NIG distribution exist, which implies via Lemma 1.17 that the integral in (2.2.13) exists.
Equations (2.2.12) and (2.2.13) imply (2.2.10) for the positive part of the real axis. The negative one follows analogously.
As for condition (2.2.11) we write

$$\int_1^{\infty} (e^x - 1)(1 + \sum_{i=1}^{n} x^{\gamma_i} e^{\theta_i x}) K(dx) \leq \sum_{i=1}^{n} \int_1^{\infty} (e^x - 1)(1 + x^{\gamma_i} e^{\theta_i x}) K(dx)$$

because of $c(1+\sum_{i=1}^n a_i) \le c(n+\sum_{i=1}^n a_i) = \sum_{i=1}^n c(1+a_i)$ for a constant $c>0$, $a_i>0, i\in\{1,\dots,n\}$. Furthermore

$$\begin{aligned}&\int_1^\infty (e^x-1)(1+x^\gamma e^{\theta x})K(dx)\\ &= \int_1^\infty e^x K(dx) - K\{[1,\infty)\} + \int_1^\infty x^\gamma K^{\beta+\theta+1}(dx) - \int_1^\infty x^\gamma K^{\beta+\theta}(dx) < \infty.\end{aligned}$$

The existence of the first integral follows from condition (1.5.2), and $K\{[1,\infty)\}$ is finite because it is a Lévy measure, i.e. (1.3.3) is fulfilled. The last two integrals are finite using the same argument as in the first part of the proof if we take for granted the admissible range of θ. □

For the main statement of the section the Bessel function K_1 ocurring in the NIG Lévy measure (2.2.1) has to be represented in a convenient form. From Watson (1966), p.182, eq. (8), we have for $x>0$

$$K_1(x) = \frac{1}{2}\int_0^\infty u^{-2} e^{-\frac{1}{2}x(u+\frac{1}{u})} du.$$

Substitution $u \to \frac{1}{u}$ yields the simpler form[7]

$$K_1(x) = \frac{1}{2}\int_0^\infty e^{-\frac{1}{2}x(u+\frac{1}{u})} du. \tag{2.2.14}$$

In addition we need Euler's integral representation of the Gamma function[8] $x \to \Gamma(x)$

$$\Gamma(\gamma) = \int_0^\infty t^{\gamma-1} e^{-t} dt,\ \gamma > 0. \tag{2.2.15}$$

From (2.2.7) we have

$$\psi^{\mathbf{Q}}(z) = ibz + iz\int_{\mathbb{R}} x(y(x)-1)\,K(dx) + \int_{\mathbb{R}} (e^{izx}-1-izx)y(x)K(dx)$$

[7]For a compact account of Bessel functions see Cont and Tankov (2004b), Appendix A. A simple substitution transforms representation (2.2.14) into the Sommerfeld integral representation of the order one of the modified Bessel function of the second kind.

[8]See e.g. Abramowitz and Stegun (1972).

$$
\begin{aligned}
&= ibz + iz\int_0^\infty xy_1(x)\,K(dx) + iz\int_{-\infty}^0 xy_2(x)\,K(dx) \\
&\quad + \int_0^\infty (e^{izx} - 1 - izx)y_1(x)K(dx) + \int_{-\infty}^0 (e^{izx} - 1 - izx)y_2(x)K(dx) \\
&\quad + \psi^{\mathbf{P}}(z) - ibz \\
&= \psi^{\mathbf{P}}(z) + iz\int_0^\infty xy_1(x)\,K(dx) + iz\int_{-\infty}^0 xy_2(x)\,K(dx) \\
&\quad + \int_0^\infty (e^{izx} - 1 - izx)y_1(x)K(dx) + \int_{-\infty}^0 (e^{izx} - 1 - izx)y_2(x)K(dx) \\
&= \psi^{\mathbf{P}}(z) + \int_0^\infty (e^{izx} - 1)y_1(x)K(dx) + \int_{-\infty}^0 (e^{izx} - 1)y_2(x)K(dx) \\
&= \psi^{\mathbf{P}}(z) + \int_0^\infty e^{izx}y_1(x)K(dx) + \int_{-\infty}^0 e^{izx}y_2(x)K(dx) \\
&\quad - \int_0^\infty y_1(x)K(dx) - \int_{-\infty}^0 y_2(x)K(dx). \qquad (2.2.16)
\end{aligned}
$$

The last but one equality is justified by the fact that the measures $y_i(x)K(dx)$ for $i = 1, 2$ integrate the function $x \to x$ due to the factor x^γ. For a convenient representation of the results we set

$$
\begin{aligned}
\mathcal{I}_1(z, \Gamma, \Theta) &= \int_0^\infty e^{izx}y_1(x)K(dx), \\
\mathcal{I}_2(z, \bar{\Gamma}, \bar{\Theta}) &= \int_{-\infty}^0 e^{izx}y_2(x)K(dx),
\end{aligned}
$$

where we suppress the dependence of $\mathcal{I}_1$ and $\mathcal{I}_2$ of the parameters of X. The most important thing is to evaluate efficiently the first two integrals in (2.2.16) because in our algorithm they must be evaluated for many values of z. This will be done in the next lemma.

Lemma 2.4. *For the above integrals we have:*

$$
\mathcal{I}_1(z, \Gamma, \Theta) = \frac{\delta\alpha}{2\pi}\sum_{k=1}^n \left(\Gamma(\gamma_k)\int_0^\infty \frac{u^{\gamma_k}}{\left(\frac{1}{2}\alpha u^2 - c_k(z)u + \frac{\alpha}{2}\right)^{\gamma_k}}du\right) \qquad (2.2.17)
$$

and

$$
\mathcal{I}_2(z, \bar{\Gamma}, \bar{\Theta}) = \frac{\delta\alpha}{2\pi}\sum_{j=1}^m \left(\Gamma(\bar{\gamma}_j)\int_0^\infty \frac{u^{\bar{\gamma}_j}}{\left(\frac{1}{2}\alpha u^2 - \bar{c}_j(z)u + \frac{\alpha}{2}\right)^{\bar{\gamma}_j}}du\right) \qquad (2.2.18)
$$

where $c_k(z) := \theta_k + \beta + iz$ *and* $\bar{c}_j(z) := \bar{\theta}_j - \beta - iz$ *for* $k \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$. *For the special case* $\gamma_k = \bar{\gamma}_j = 2$ *for* $k \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$ *this leaves us with*

$$
\mathcal{I}_1(z, 2, \Theta) =
$$

$$= \sum_{k=1}^{n} \frac{\delta}{2\pi(\alpha^2 - c_k(z)^2)} \left\{ 2c_k(z) + \frac{\alpha^2}{(\alpha^2 - c_k(z)^2)^{1/2}} \left[\pi - 2\arctan\left(\frac{-c_k(z)}{\sqrt{\alpha^2 - c_k(z)^2}} \right) \right] \right\} \tag{2.2.19}$$

and

$$\mathcal{I}_2(z, 2, \bar{\Theta}) =$$

$$= \sum_{j=1}^{m} \frac{\delta}{2\pi(\alpha^2 - \bar{c}_j(z)^2)} \left\{ 2\bar{c}_j(z) + \frac{\alpha^2}{(\alpha^2 - \bar{c}_j(z)^2)^{1/2}} \left[\pi - 2\arctan\left(\frac{-\bar{c}_j(z)}{\sqrt{\alpha^2 - \bar{c}_j(z)^2}} \right) \right] \right\} \tag{2.2.20}$$

Proof. As before we will do this for the first integral and for $n = 1$, and the index 1 will again be dropped.

$$\begin{aligned} \int_0^\infty e^{izx} y_1(x) K(dx) &= \int_0^\infty e^{izx} x^\gamma e^{\theta x} K(dx) \\ &= \frac{\delta\alpha}{\pi} \int_0^\infty \exp\left[(iz + \theta + \beta)x\right] x^{\gamma - 1} K_1(\alpha x) dx \\ &= \frac{\delta\alpha}{2\pi} \int_{x=0}^\infty \exp\left[(iz + \theta + \beta)x\right] x^{\gamma-1} \int_{u=0}^\infty \exp\left[-\frac{1}{2}\alpha x(u + \frac{1}{u}) \right] du\, dx. \end{aligned}$$

Here the integral representation (2.2.14) is used. The Fubini theorem, substitution $y = \left(\frac{1}{2}\alpha(u + \frac{1}{u}) - \theta - \beta - iz\right) x$ and Euler's integral representation (2.2.15) for the Gamma function yield

$$\begin{aligned} &\int_0^\infty e^{izx} y_1(x) K(dx) \\ &= \frac{\delta\alpha}{2\pi} \int_{u=0}^\infty \int_{x=0}^\infty \exp\left[-\left(\frac{1}{2}\alpha(u + \frac{1}{u}) - \theta - \beta - iz \right) x \right] x^{\gamma-1} dx\, du \quad (2.2.21) \\ &= \frac{\delta\alpha}{2\pi} \int_{u=0}^\infty \left(\frac{1}{2}\alpha(u + \frac{1}{u}) - \theta - \beta - iz \right)^{-\gamma} \int_{y=0}^\infty e^{-y} y^{\gamma-1} dy\, du \\ &= \frac{\delta\alpha}{2\pi} \Gamma(\gamma) \int_{u=0}^\infty \left(\frac{1}{2}\alpha(u + \frac{1}{u}) - \theta - \beta - iz \right)^{-\gamma} du \\ &= \frac{\delta\alpha}{2\pi} \Gamma(\gamma) \int_{u=0}^\infty \frac{u^\gamma}{\left(\frac{1}{2}\alpha u^2 - (\theta + \beta + iz)u + \frac{\alpha}{2}\right)^\gamma} du \quad (2.2.22) \end{aligned}$$

The computation of the second equality, which recovers the gamma function, actually involves a complex contour integral and can be justified as follows: If we set $v(u) := \frac{1}{2}\alpha(u + \frac{1}{u}) - \theta - \beta$ for $u > 0$, the inner integral in (2.2.21) becomes

$$\lim_{t \to \infty} \int_{\mathcal{J}(t)} y^{\gamma-1} e^{-y} dy$$

for fixed $u > 0$ after substitution where $\mathcal{J}(t) := \{(v(u) - iz)s | 0 \le s \le t\}$. Using the Cauchy theorem for integrals over closed contours we obtain

$$\int_{\mathcal{J}(t)} y^{\gamma-1} e^{-y} dy = \int_{\mathcal{J}_1(t)} y^{\gamma-1} e^{-y} dy + \int_{\mathcal{J}_2(t)} y^{\gamma-1} e^{-y} dy$$

with $\mathcal{J}_1(t) := \{v(u)s|0 \leq s \leq t\}$ and $\mathcal{J}_2(t) := \{-izs|0 \leq s \leq t\}$. Due to $u + \frac{1}{u} \geq 2$ for $u > 0$ and Definition 2.1 we have $v(u) \geq \alpha - \theta - \beta > 1$, and hence the first integral on the right-hand side converges to the gamma function evaluated at γ as $t \to \infty$. The second one disappears for $t \to \infty$ because $v(u) > 0$ and

$$\left| \int_{\mathcal{J}_2(t)} y^{\gamma-1} e^{-y} dy \right| \leq v(u) e^{-v(u)t} \int_0^{|z|} |v(u)t - is|^{\gamma-1} ds$$

where the integral on the right-hand side, depending on the value of γ, displays an at most polynomial increase in t.

For $\gamma = 2$ with Bronstein et al. (1993), 19.5.1.2, formula 48, setting $c(z) := \theta + \beta + iz$ we have

$$\begin{aligned}
\int_0^\infty e^{izx} y_1(x) K(dx) \Big|_{\gamma=2} &= \lim_{M\to\infty} \int_0^M e^{izx} y_1(x) K(dx) \Big|_{\gamma=2} = \\
&= \frac{\delta\alpha}{2\pi} \Gamma(2) \lim_{M\to\infty} \left\{ \left[\frac{2(2c(z)^2 - \alpha^2)u - 2\alpha c(z)}{\alpha(\alpha^2 - c(z)^2)(\alpha u^2 - 2c(z)u + \alpha)} \right.\right. \\
&\quad \left.\left. + \frac{2\alpha}{(\alpha^2 - c(z)^2)^{3/2}} \arctan\left(\frac{\alpha u - c(z)}{\sqrt{\alpha^2 - c(z)^2}} \right) \right]_{u=0}^{u=M} \right\}.
\end{aligned}$$

The limit of the complex arctan-function for a fixed imaginary part and the real part of the argument tending to plus infinity is $\pi/2$. Therefore

$$\begin{aligned}
&= \frac{\delta\alpha}{2\pi} \lim_{M\to\infty} \left\{ \frac{2(2c(z)^2 - \alpha^2)M - 2\alpha c(z)}{\alpha(\alpha^2 - c(z)^2)(\alpha M^2 - 2c(z)M + \alpha)} \right. \\
&\quad + \frac{2\alpha}{(\alpha^2 - c(z)^2)^{3/2}} \arctan\left(\frac{\alpha M - c(z)}{\sqrt{\alpha^2 - c(z)^2}} \right) \\
&\quad \left. + \frac{2\alpha c(z)}{\alpha^2(\alpha^2 - c(z)^2)} - \frac{2\alpha}{(\alpha^2 - c(z)^2)^{3/2}} \arctan\left(\frac{-c(z)}{\sqrt{\alpha^2 - c(z)^2}} \right) \right\} \\
&= \frac{\delta\alpha}{2\pi} \left\{ \frac{\pi\alpha}{(\alpha^2 - c(z)^2)^{3/2}} + \frac{2\alpha c(z)}{\alpha^2(\alpha^2 - c(z)^2)} \right. \\
&\quad \left. - \frac{2\alpha}{(\alpha^2 - c(z)^2)^{3/2}} \arctan\left(\frac{-c(z)}{\sqrt{\alpha^2 - c(z)^2}} \right) \right\} \\
&= \frac{\delta}{2\pi(\alpha^2 - c(z)^2)} \left\{ 2c(z) + \frac{\alpha^2}{(\alpha^2 - c(z)^2)^{1/2}} \left[\pi - 2\arctan\left(\frac{-c(z)}{\sqrt{\alpha^2 - c(z)^2}} \right) \right] \right\}.
\end{aligned}$$

Note that for the second integral over the negative part of the real line we just have to change the signs of β and z in order to apply the above calculations. □

For general γ solving the integrals in (2.2.22) requires the Appell hypergeometric function. For tractability reasons we will choose the special case where all γ-parameters are equal to 2. This will also be our choice for the analysis in the next section.

Now we have again a look at Theorem 1.22 and the martingale condition (2.2.6):

$$b - r + \int_{-\infty}^{\infty} ((e^x - 1)y(x) - x)K(dx) = 0. \tag{2.2.23}$$

Lemma 2.5. *Equation* (2.2.23) *is equivalent to*

$$\mathcal{I}_1(-i, \Gamma, \Theta) + \mathcal{I}_2(-i, \bar{\Gamma}, \bar{\Theta}) - \mathcal{I}_1(0, \Gamma, \Theta) - \mathcal{I}_2(0, \bar{\Gamma}, \bar{\Theta}) = r - \log(\varphi(1)). \tag{2.2.24}$$

Proof. The integral in (2.2.23) is

$$\begin{aligned}
&\int_{-\infty}^{\infty} \{(e^x - 1)y(x) - x)K(dx)\} \\
&= \int_{-\infty}^{\infty} \{(e^x - 1)(1 + y_1(x)I_{x>0} + y_2(x)I_{x<0}) - x\} K(dx) \\
&= \int_0^{\infty} (e^x - 1)y_1(x)K(dx) + \int_{-\infty}^0 (e^x - 1)y_2(x)K(dx) + \\
&\quad \int_{-\infty}^{\infty} (e^x - 1 - x)K(dx) \\
&= \int_0^{\infty} (e^x - 1)y_1(x)K(dx) + \int_{-\infty}^0 (e^x - 1)y_2(x)K(dx) + \log(\varphi(1)) - b
\end{aligned}$$

because $\psi^{\mathbf{P}}(-i) = \log(\varphi(1)) = b + \int_{-\infty}^{\infty}(e^x - 1 - x)K(dx)$. Note that condition (1.5.2) ensures that $\varphi(1) > 0$ is finite and hence the domain of $\psi^{\mathbf{P}}$ can be extended such that $\psi^{\mathbf{P}}(-i)$ is well-defined.
The statement follows from equation (2.2.23) and splitting the integrals, which is possible due to the finiteness of all appearing integrals. ☐

Finally we have obtained an analytic expression of the cumulant function $\psi^{\mathbf{Q}}$:

Proposition 2.6. *The cumulant function $\psi^{\mathbf{Q}}$ of X under the flexible change of measure with parameters Θ, Γ, $\bar{\Theta}$ and $\bar{\Gamma}$ has the representation*

$$\psi^{\mathbf{Q}}(z) = \psi^{\mathbf{P}}(z) + \mathcal{I}_1(z, \Gamma, \Theta) + \mathcal{I}_2(z, \bar{\Gamma}, \bar{\Theta}) - \mathcal{I}_1(0, \Gamma, \Theta) + \mathcal{I}_2(0, \bar{\Gamma}, \bar{\Theta})$$

where $\psi^{\mathbf{P}}$ has the form (1.5.3), and $\mathcal{I}_1$ and $\mathcal{I}_2$ are given by (2.2.17) and (2.2.18). ☐

2.3 Option pricing with flexible measures

In this section we focus entirely on the easiest types of flexible measures by setting $n = m = 1$ and $\gamma_1 = \bar{\gamma}_1 = 2$ in Definition 2.1, where again the index 1 will be dropped henceforth. We explore if they have the capability of significantly changing the pricing performance of the NIG-Lévy model in connection with the Esscher transform. This model will be referred to as the NIG-Esscher model.
The procedure in this section is as follows: First, the solution of the martingale condition (2.2.24) is computed in terms of the two remaining parameters θ and $\bar{\theta}$. By fixing γ and $\bar{\gamma}$ to be equal to 2 we lose control over the behaviour of y in the region close to 0. Nevertheless, through the remaining single degree of freedom, which we have through choosing θ and $\bar{\theta}$ under the constraint (2.2.24), we have enough flexibility so as to obtain different shapes of the measure change function y, while still retaining an acceptable degree of computational complexity. Then the FFT algorithm from Section 1.6 will be employed for the computation of a whole range of European option prices for strikes around the current price of the underlying. We confine ourselves to call options in the sequel.

2.3.1 Differences among flexible measures

The first natural question to ask is about which of the many flexible martingale measures provided by the solution of the martingale condition (2.2.24) should be chosen. For this the zeros of the function[9]

$$(-\alpha-\beta, \alpha-\beta-1) \times (-\alpha+\beta+1, \alpha+\beta) \ni (\theta_1, \theta_2) \to$$
$$F(\theta_1,\theta_2) = \mathcal{I}_1(-i,2,\theta_1) + \mathcal{I}_2(-i,2,\theta_2) - \mathcal{I}_1(0,2,\theta_1) - \mathcal{I}_2(0,2,\theta_2) - r + \log(\varphi(1))$$

have to be examined. This occurs exemplarily for a set of parameters[10], which are detailed in Fig. 2.1, first keeping the interest rates at zero and then raising them to the annual rate of 4%. Fig. 2.1 shows the graphs of F on the top and its zeros at the bottom. The different colouring of the bottom pictures indicates the form of the measure change function y. We select three showcase points T1, T2, T3 where $\theta_1 \in \{-20, 0, 10\}$. T1 corresponds to an asymmetric measure while T3 determines a symmetric measure. T2 is also symmetric, though its increase on the positive real axis is of polynomial order because the exponential term vanishes. A positive interest rate has the effect of lowering the graphs of F in Fig. 2.1. The more the interest rate increases the more downwards moves the line of zeros. For very large values of the interest rate the graph of the line of zeros looks like the left graph in Fig. 2.1 reflected at the diagonal of the rectangular, i.e. the θ_1-part of all martingale measures becomes positive and we have only the symmetric shape. However, for large values of θ_1 the pricing algorithm becomes unstable and produces prices which

[9]For convenience we set $\theta_2 := \bar{\theta}_1$ such that we work with θ_1 and θ_2 instead of θ and $\bar{\theta}$.
[10]These are the Volkswagen parameters used in the next but one section.

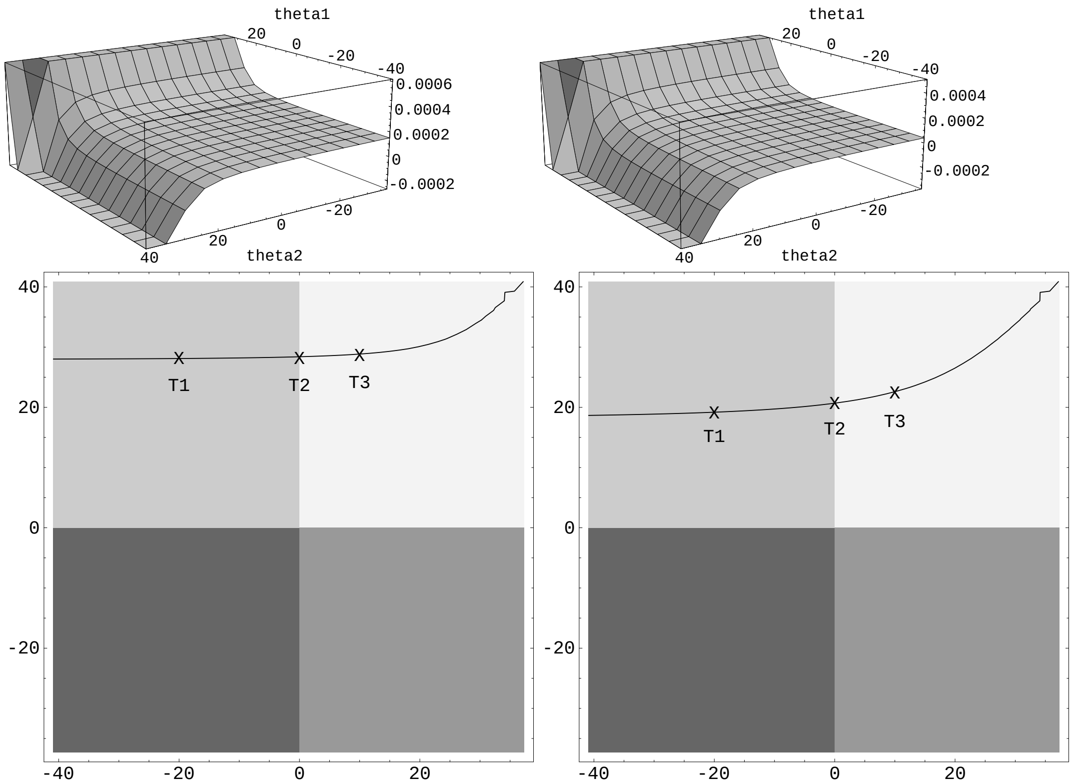


Fig. 2.1. Visualization of the martingale condition for $\alpha = 40$, $\beta = 1.3$, $\delta = 0.03$, $\mu = -0.0012$, $S_0 = 1$, $T = 1$ day and an annualized interest rate of 0 % for the left-hand side and 4 % for the right hand side. The pictures at the bottom each provide a top view of the two-dimensional manifolds right above them. θ_1 runs in both cases along the horizontal axis. The line connects all the point which give rise to zeros of the function F. The three selected points are the ones chosen for the analysis in the main text.

are not arbitrage-free. This is also the reason why T3 was not chosen to be bigger than $\theta_1 = 10$ in the first component.

The pricing behaviour of T1, T2, T3 is examined in Fig. 2.2. First we see the different forms of the involved measure change functions discussed above. The values of θ_1 and θ_2 enforce a considerable steepness of the functions such that the scaling of the abscissa has to be adjusted carefully in order to discern their shape. The graphs in the middle and at the bottom clearly show two things: First, the difference among the various measures is tiny, almost not discernible in the pictures in the middle. Second, zooming in yields that there is a certain structure among the measures going with the order structure of the parameters θ_1 - the bigger θ_1 is the bigger is the option price. This results in the following conclusion: Fine tuning of the measure to obtain reasonable prices (e.g. prices of exchange traded options) is possible because we can - at least in this and some other examples - obtain a higher resp. lower price by

increasing resp. decreasing θ_1. However, the difference is very small and in some cases even negligible.
This very last statement leads us to consider one representative out of the sample of all flexible martingale measure and call it *the* flexible measure. Our choice is $\theta_1 = 0$, which will prove to be appropriate in the subsequent sections when assessing the performance of the flexible measure relative to the Esscher-Lévy case and Black-Scholes prices.
As to the question of the shape of the measure change function we see that its effect on option pricing can be neglected in our framework.

2.3.2 Options sensitivities

Now that we have chosen a canonical flexible measure with $\theta_1 = 0$ we can compare the NIG-flexible model, the NIG-Esscher model and the Black-Scholes model with respect to their response to changes in the marginal distributions of the underlying Lévy process for the stock price movement. These changes are best embodied by mean, variance, skewness and kurtosis. As it is, these figures are not very intuitively reflected in the NIG parameters $(\alpha, \beta, \delta, \mu)$. The solution is to reparametrize the NIG distribution in order to obtain parameters (ξ, χ, σ, m) which better suit this purpose. They are derived in Appendix A.2. ξ ranging from 0 to 1 is a measure of kurtosis (the bigger ξ the higher the kurtosis), χ represents skewness, σ and μ are standard deviation resp. mean. The time to maturity is $T = 1$ day, and the current price of the underlying is normalized to 1 as well. As they are the most important factors, the impact of kurtosis and volatility is presented in Figures 2.3 and 2.4.
The analysis is carried out in terms of the difference between Lévy prices and Black-Scholes prices. The qualitative structure of the differences between the flexible prices and the Black-Scholes prices on the one hand and between the Esscher prices and the Black-Scholes prices on the other hand is the same. This is the first important observation to make. At-the-money options are cheaper and away-from-the-money options are more expensive in both Lévy models compared to Black-Scholes prices whereas for options which are either far in-the-money or far out-of-the-money the difference vanishes. The latter observation is more a consistency check of the pricing method than a structural observation because in every sensible model prices have to converge to 0 resp. to S_0 if they are very far out-of-the-money resp. in-the-money.
In spite of this qualitative similarity the quantitative differences are sizable and grow further with increasing kurtosis. The differences between Black-Scholes and Lévy-flexible prices are also remarkable: For instance, in the high kurtosis case ($\xi = 0.8$) the ATM option is roughly 10 % more expensive in the Black-Scholes case than in the Lévy case with the flexible change of measure. The case of volatility reveals an even more marked distance. Note that the chosen parameters have all values which are quite reasonable for empirical data.
The differences between flexible and Esscher prices is always positive and biggest for options very slightly in-the-money, monotonically decreasing towards both sides. As for the well-known problem of underpriced out-of-the-money options in the Black-

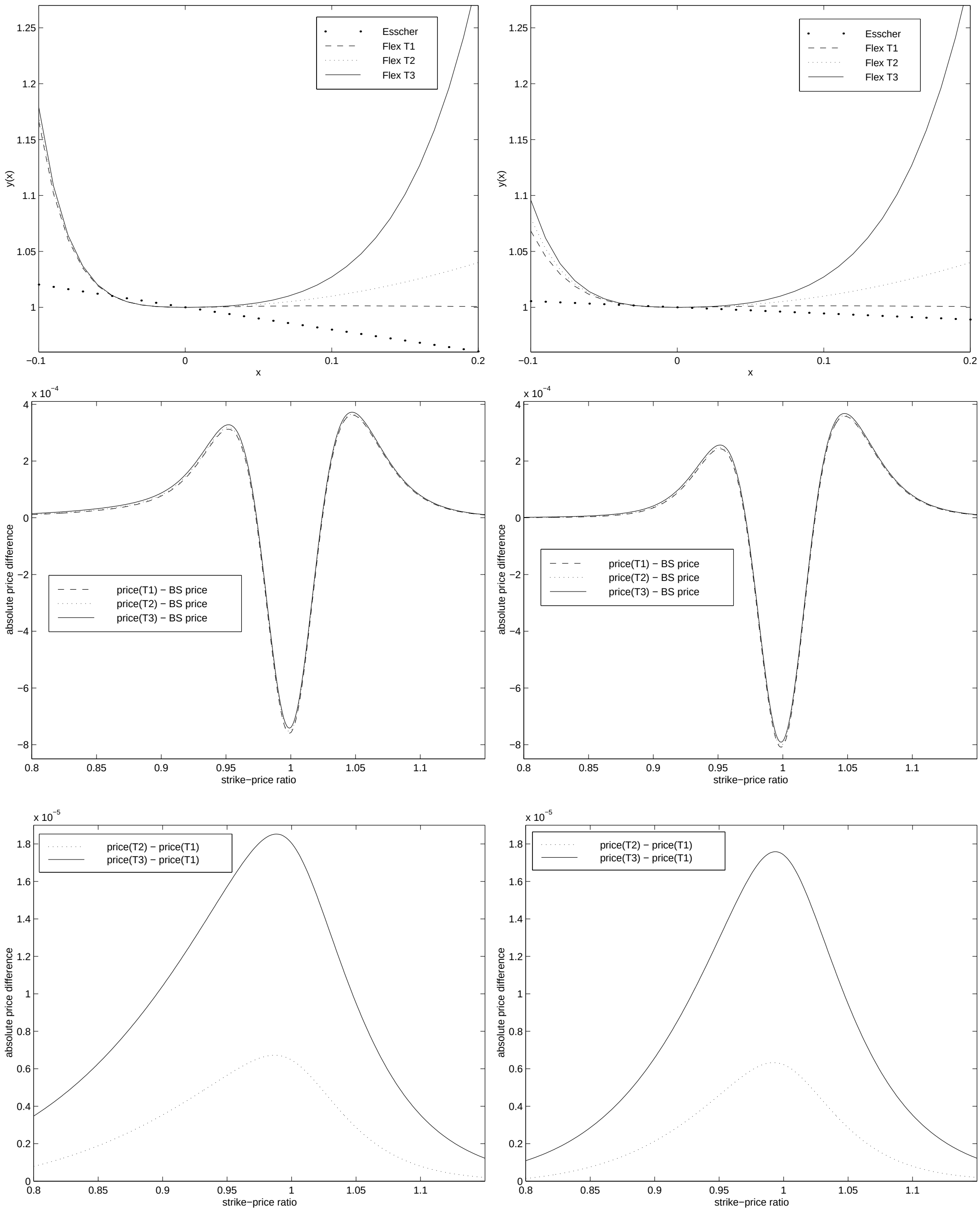


Fig. 2.2. Pricing performance of different flexible measures for an annual interest rate of 0% (left column) and 4% (right column). The top pictures graph the measure change functions y for T1, T2, T3 as well as the corresponding Esscher measure change function. In the middle the distance of these prices relative to Black-Scholes prices is given as a function of the strike-price ratio. At the bottom the difference of T2 and T3 versus T1 is depicted.

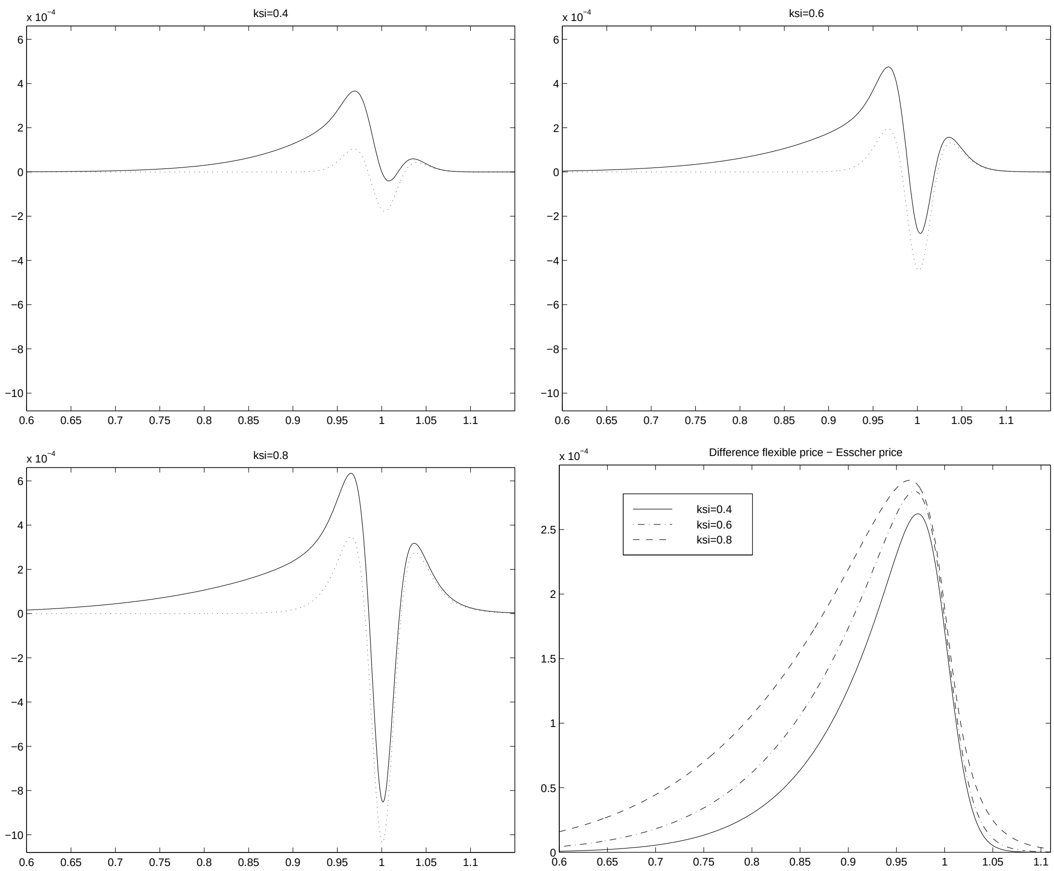


Fig. 2.3. The impact of changes of the kurtosis parameter ξ on the difference between flexible price and Esscher price for $\chi = -0.02$, $\sigma = 0.02$, $m = 0.0002$ and $\xi \in \{0.4, 0.6, 0.8\}$. The first three pictures show the differences of both models with respect to the Black-Scholes model (the solid line for the flexible measure and the dotted line for the Esscher measure), the last one illustrates the differences within the two classes of prices, each for the three values of ξ.

Scholes world flexible measures tend to provide a more pronounced correction to option prices than Esscher prices do. An analogous examination of the effects of skewness and mean lead to similar graphs like above. The impact of the mean is in the same quantitative order as the one of kurtosis whereas skewness plays a minor role.

2.3.3 A practical analysis of pricing performance

This section serves to demonstrate the pricing performance of the NIG-flexible model compared with the NIG-Esscher model and the Black-Scholes model by means of empirical data.

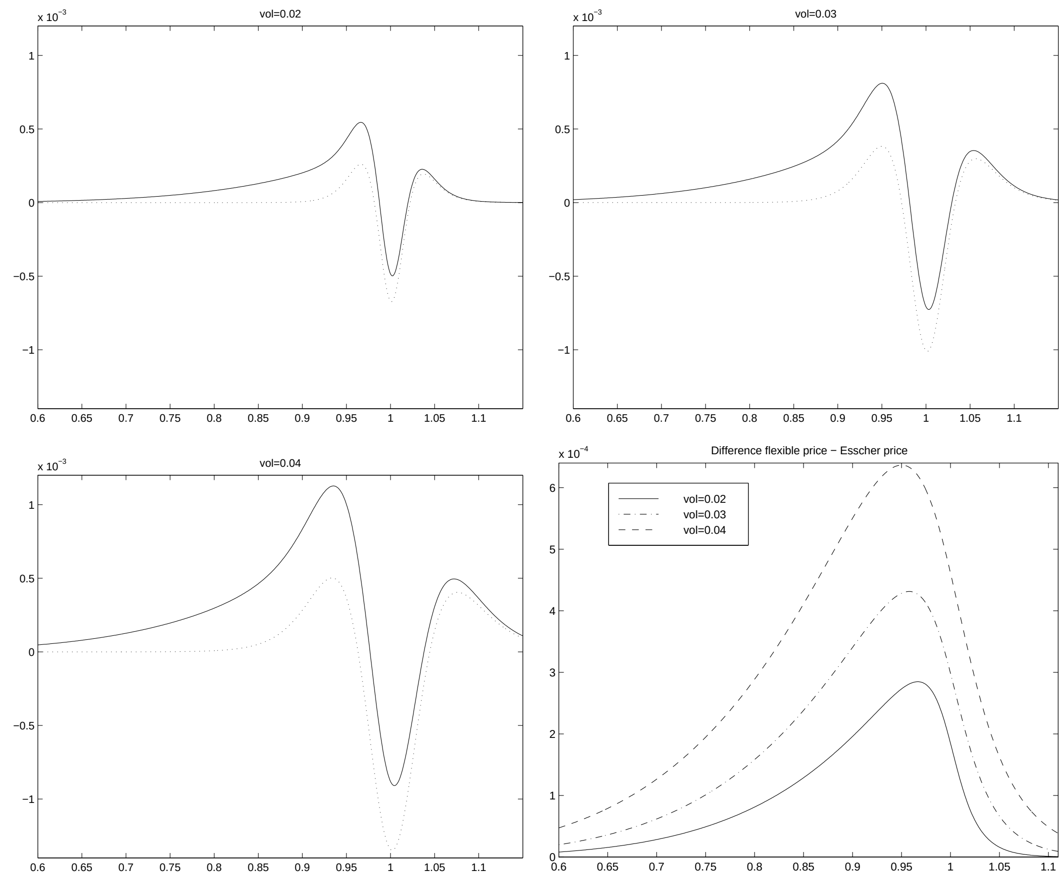


Fig. 2.4. The impact of changes of the volatility σ on the difference between flexible price and Esscher price for $\xi = 0.7$, $\chi = -0.02$, $m = 0.0002$ and $\sigma \in \{0.02, 0.03, 0.04\}$. The structure of the graphs is like in Fig. 2.3.

For each model we analyse call option prices on four German stocks - Daimler-Chrysler (DCX), Deutsche Bank (DBK), SAP, and Volkswagen (VOW) - from different sectors which are traded on the electronic trading platform Xetra. For the maximum likelihood estimation of the Lévy parameters we use daily data from 1/1/1998 until 24/3/2003 in the case of Deutsche Bank, SAP, and Volkswagen. Daimler-Chrysler stock price data are available since 27/10/98. This adds up to 1321 resp. 1114 observations. Dividends are taken into account through an appropriate one-off increase of the returns whenever dividends are paid[11]. In Germany this occurs only once a year.

[11] See Eberlein and Keller(1995).

	mean	std	skew	kurt	Jarque-Bera
DCX	-0.145	0.391	0.046	1.17	63.75
DBK	-0.055	0.430	-0.027	2.45	330.30
SAP	-0.0225	0.625	0.299	6.02	520.23
VOW	-0.0625	0.417	-0.030	1.95	209.32

The table on the right shows some basic statistical figures of the empirical distribution of stock log return data: mean, standard deviation, skewness, kurtosis and the test statistic for the Jarque-Bera test for goodness-of-fit to a normal distribution. The scaling sensitive variables mean and standard deviation are annualized using the convention of 250 trading days per year. It can be seen that the hypothesis of normal distribution is clearly rejected for every stock (99 % quantile of chi-squared distribution: 9.21).[12]
The parameters are estimated by a steepest descent maximum likelihood method. For this we make use of the corresponding procedure in the C programme 'hyp' of Blæsild and Sørensen (1992) which was written for generalized hyperbolic distributions. The log-likelihood function for the NIG case is given in Section A.1 in the appendix. However, these parameters still have to be changed in order to use them for pricing.

The role of the volatility in the sensitivity analysis of the last section suggests a careful handling of this issue. Of course, the parameter estimates from the maximum likelihood procedure determine something like a global variance over more than five years. However, the effect of volatility clustering must be somehow taken into consideration given that we do not explicitly model this phenomenon. Hence we estimate a 30 day historical volatility and rescale the parameters such that they yield this very value of volatility, retaining the values of skewness and kurtosis. This procedure is described in Prause (1999), p.38. Basic to this approach is the view that skewness and kurtosis define roughly the shape of the distribution which remains relatively stable over time. The rescaled parameters are given in Table 2.1 together with the annualized historical 30 day volatility *hist30* and the Esscher measure change parameter θ.

	α	β	δ	μ	hist30	θ
DCX	41.8274	3.6960	0.0462	-0.0036	0.528	-0.874
DBK	28.7281	-0.5901	0.0413	0.0004	0.599	-0.132
SAP	34.5103	1.6752	0.0363	-0.0022	0.514	-0.002
VOW	29.9364	0.9825	0.0372	-0.0012	0.558	-0.484

Tab. 2.1. Parameter estimates of fitted NIG distribution *after* rescaling.

[12] It is interesting to note that the values of the kurtosis are remarkably lower nowadays than they were ten years ago. This could be a result of an increasing trading frequency (due to electronic platforms, growing importance of stocks among private investors) which could potentially lead to less heavy tails via the central limit theorem.

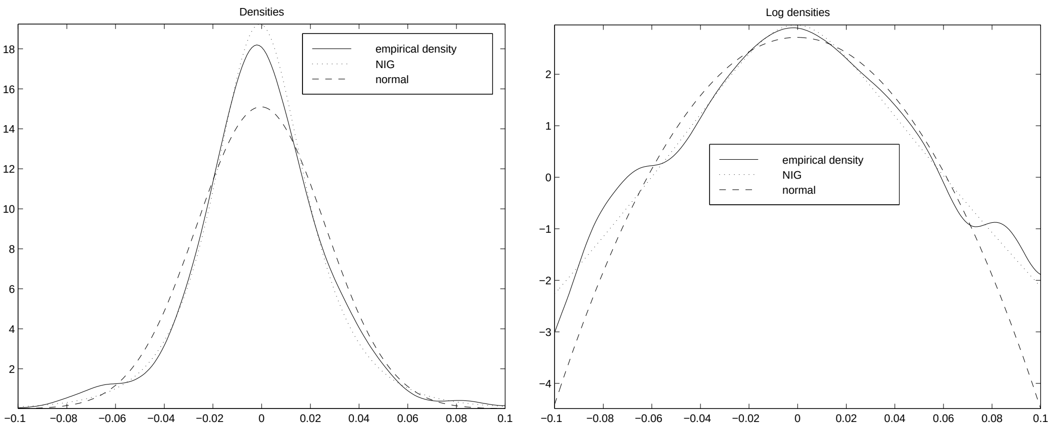


Fig. 2.5. Densities and log densities of empirical density function, fitted NIG and fitted normal distribution for the Volkswagen data.

Fig. 2.5 exemplifies the good fit of the NIG distribution compared to the normal distribution with Volkswagen data. The empirical density function is represented using a Gaussian kernel estimator. Compared to the normal distribution the NIG distribution has more mass around the origin and in the tails. Note that for evaluating the fit in the tails the log density representation on the right hand side of Fig. 2.5 is much more suitable.

Fig. 2.6 depicts for each of the Lévy models the distance to Black-Scholes prices for Daimler-Chrysler and Deutsche Bank, and time to maturity $T = 1, 5, 10$ days. In the case of Daimler-Chrysler it can be seen that the different behaviour of flexible and Esscher prices is quite striking. Whereas the Esscher price differences level off over time, the difference between flexible and BS prices becomes much more pronounced and the shape changes from an almost point-symmetric graph to a bell-shaped curve. This can be of importance if the view is held that the Esscher price corrections of the BS price are too small. Comparing the scaling of the DCX and DBK figures one detects that the differences in the latter case are quite modest. A similar difference can be recognized in Fig. 2.7 which shows these differences for VOW and SAP across the dimensions strike-price ratio and time to maturity. Flexible prices tend to become higher for Volkswagen with time to maturity growing both in comparison with Esscher and Black-Scholes prices whereas for the SAP stocks the differences flatten over time.
An interesting observation is that those stocks with big differences between Esscher and flexible measure, i.e. DCX and VOW, are those whose Esscher parameter θ is close to zero, i.e. the processes under the statistical probability measure are already close to a martingale. But this means possibly that a change of measure does not bring about much, whether it be of the Esscher or the flexible type, and this could be a reason why prices under these two martingales measures do not differ much.

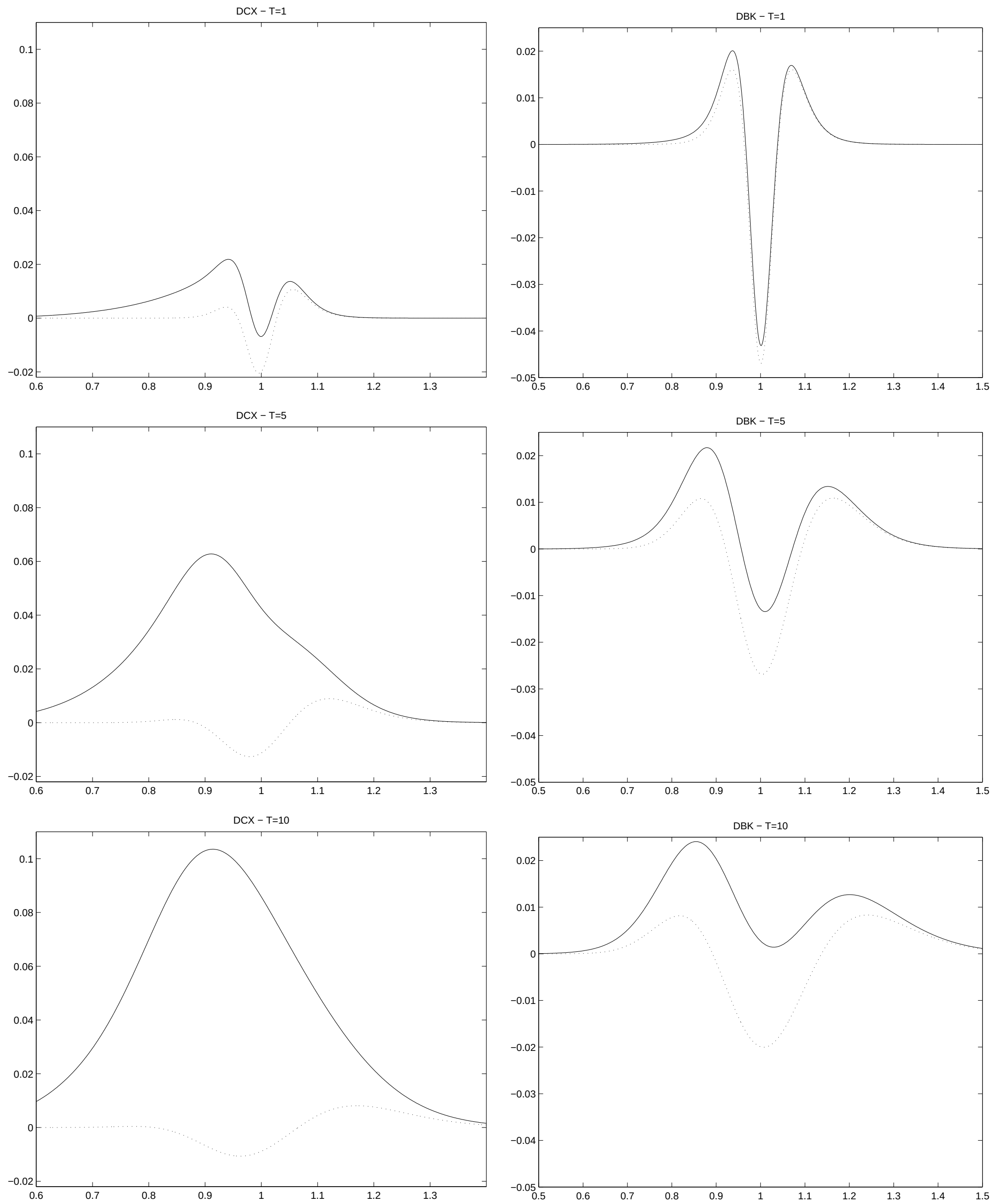


Fig. 2.6. Absolute price differences between the NIG-flexible model and Black-Scholes (solid line) and between the NIG-Esscher model and Black-Scholes (dotted line). From top to bottom time to maturity goes from $T = 1$ day via $T = 5$ days to $T = 10$ days. The left hand side shows Daimler-Chrysler, the right hand side Deutsche Bank

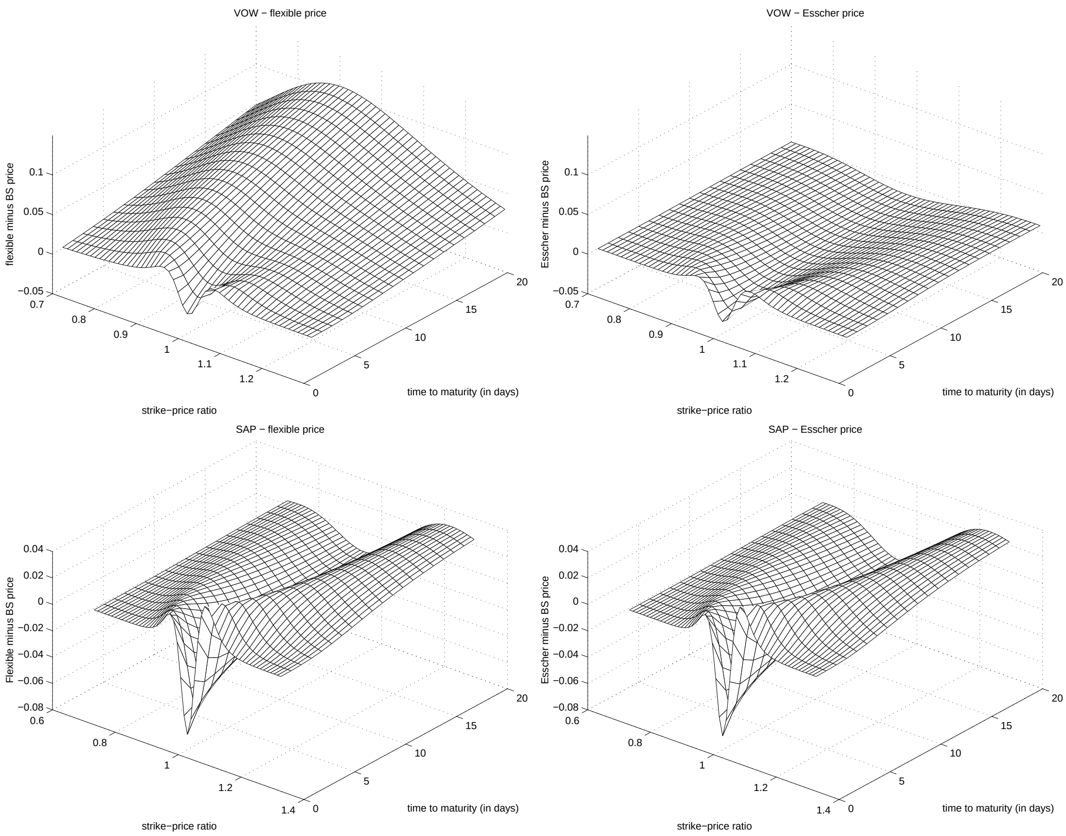


Fig. 2.7. Flexible-Black-Scholes (left) and Esscher-Black-Scholes (right) differences for SAP (top) and Volkswagen (bottom). Time to maturity ranges from 1 to 20 days.

Of course, this argument depends intuitively on the assumption that the change of measure looks for the most direct way to obtaining a martingale. This should be the case for the Esscher measure, but in the case of the flexible measure this is less clear. However, as we chose T2 in Section 2.3.1 with $\theta_1 = 0$, this assumption might be satisfied.
On the other hand, DBK and SAP require a 'sizable' change of measure because they are far from being a martingale, and hence the repsective prices differ significantly. For results regarding the issue of time consistency of models based on Lévy processes see Eberlein and Özkan (2003).

2.4 Concluding remarks

This section presents an approach of pricing derivatives on securities whose prices are modeled as exponential Lévy processes. Throughout the presentation we focus on the special case of normal inverse Gaussian Lévy processes. As the model is in-

complete a specific change of measure must be chosen in order to close the model. We introduce the parametrized class of flexible martingale measures which are conveniently handled when calculating option prices via FFT methods - a property which so far seems to be shared only by the Esscher measure.
We see that for a certain model flexible prices for different parameters of the measure change function change only slightly so that we chose one specific measure as a credible representative of all the flexible measure. A comparison of the pricing performance of Black-Scholes prices, Esscher prices and flexible prices yields an interesting observation: The correction of Black-Scholes prices provided by Esscher prices is qualitatively similar to the difference of Black-Scholes and flexible prices. However, the deviation is more pronounced in the latter case, so it possibly provides a better model in cases where the Esscher measure corrects Black-Scholes prices too modestly.
The basis of flexible measures is the understanding that utility maximization in a restrictive representative agent model and distance minimization are not very convincing as a basis of choosing a specific change of measure. On the contrary in this chapter the entirely practical view is held that it is enough to specify a measure and let observed derivative prices decide on whether the proposed change of measure does or does not make sense.
An additional result is that the chapter provides a small hint at the question about the connection of Lévy processes and Lévy prices in a world where the incompleteness appears to be of a hopeless degree. We chose the flexible measure whose measure change function apparently deviates substantially from the one of the Esscher change of measure, and nevertheless we obtain prices which exhibit a strong resemblance.

Chapter 3

Moment-matching of the change of measure

3.1 Introduction

The framework in this chapter is given again by a one-dimensional exponential Lévy model of the form (1.4.2) and (1.4.3). We assume a time horizon of $T = 1$ which is also the maturity date of the contracts which will be considered here. The class of *tempered stable Lévy processes*, introduced in Section 1.5.2, lends itself in a very convenient way to the problem to be solved in this chapter. $X = (b, 0, K)_{\mathbf{P}}$ denotes a tempered stable Lévy process with its characteristic function $\chi_{\mathbf{P}}$ equal to

$$\chi_{\mathbf{P}}(z) = E^{\mathbf{P}}[e^{izX_t}] = e^{t\psi_{\mathbf{P}}(z)}$$

with the cumulant function

$$\psi_{\mathbf{P}}(z) = ibz + \int_{-\infty}^{\infty} (e^{izx} - 1)K(dx).$$

It is important to recall that we can choose the truncation function $h(x) \equiv 0$. The cumulant function has an infinite series representation[1] $\psi_{\mathbf{P}}(z) = \sum_{n=0}^{\infty} \kappa_n (iz)^n/n!$ in terms of the cumulants κ_n. From this it follows directly that

$$\kappa_n = \frac{1}{i^n}\, \psi_{\mathbf{P}}^{(n)}(z)\Big|_{z=0}, \tag{3.1.1}$$

and the cumulants are directly related to the moments of X under $\mathbf{P}$. Changing to another probability measure $\mathbf{Q}$, which is absolutely equivalent to $\mathbf{P}$ and which preserves the Lévy property, changes the Lévy triplet to $(b', 0, K')_{\mathbf{Q}}$ where[2]

$$b' = b \quad \text{and} \quad \frac{dK'}{dK}(x) = y(x)$$

[1] See Abramowitz and Stegun (1972), 26.1.12

[2] See Theorem 1.20.

such that the cumulant function under $\mathbf{Q}$ is equal to

$$\psi_{\mathbf{Q}}(z) = ibz + \int_{-\infty}^{\infty} (e^{izx} - 1) y(x) K(dx). \tag{3.1.2}$$

The *relative entropy* of $\mathbf{Q}$ to $\mathbf{P}$

$$I_t(\mathbf{Q}, \mathbf{P}) := E^{\mathbf{P}} \left[\left. \frac{d\mathbf{Q}}{d\mathbf{P}} \right|_{\mathcal{F}_t} \log \left(\left. \frac{d\mathbf{Q}}{d\mathbf{P}} \right|_{\mathcal{F}_t} \right) \right] \tag{3.1.3}$$

which is defined if $\mathbf{Q}|_{\mathcal{F}_t} \ll \mathbf{P}|_{\mathcal{F}_t}$ for all $t \in [0, T]$. We set $I(\mathbf{Q}, \mathbf{P}) := I_1(\mathbf{Q}, \mathbf{P})$. Following Theorem 1.20 giving a function y with the properties specified therein determines a change of measure which, in addition, pins down a martingale measure if the martingale condition

$$b - r + \int_{-\infty}^{\infty} (e^x - 1) y(x) K(dx) = 0. \tag{3.1.4}$$

holds. Usually, as for instance in the case of the Esscher transform, one chooses a simple functional form for y with one free parameter and then tries to solve (3.1.4) for it.

The method introduced in this chapter tries to find a more adequate answer to this problem. Given the statistical distribution of asset prices and some moments of the risk-neutral distribution, it gives a procedure of how to obtain a pricing rule in the form of the measure change function y. This function determines completely the change of measure in our setting where the asset price process follows a Lévy process both under the statistical and the risk-neutral probability measure. In this chapter we break the infinite-dimensional problem (3.1.4) into a problem with a finite number of degrees of freedom, i.e. we construct a finite-dimensional space $\mathcal{Y}$ of admissible measure change functions y. Though finite, the number of parameters is in general quite large, contrary to the usual approach. We use these degrees of freedom in order to equate not only (3.1.4) but also the empirical second, third, and fourth moment of the risk-neutral distribution. As this problem is generically still underdetermined we seek that y among the ones fulfilling the constraints which is closest to the original measure $\mathbf{P}$ as measured by the relative entropy of $\mathbf{Q}$ with respect to $\mathbf{P}$. Briefly, this means that we are concerned with the minimization problem

$$\min_{y \in \mathcal{Y}} I(\mathbf{Q}^y, \mathbf{P}) \tag{3.1.5}$$

such that

- The martingale condition (3.1.4) is satisfied.
- The second, third and fourth moment of the risk-neutral distribution of X_1 equal the respective empirical moments.
- Certain regularity conditions hold.

The following tasks remain to be done: Firstly, in Section 3.2 the set $\mathcal{Y}$ is designed in such a way that the equality constraints are linear and the corresponding coefficients can be calculated analytically. The final formulation of the central minimization problem does not appear until the end of Section 3.3 which is preceded by an analytical computation of all coefficients and functions which are needed in this respect. Section 3.4 reviews the linex pricing procedure while Section 3.5 gives a method of how to obtain the required risk-neutral moments. Section 3.6 illustrates the results with artificial data, and finally we conclude in Section 3.7.
For the practical use of this model it may be quite unrealistic to assume the risk-neutral moments to be known. Nevertheless, these moments are directly related to characteristics of market participants such as risk aversion. Hence the model can be used in order to show the impact of such characteristics on the structure of option prices.

3.2 Parametrization of the measure change function

We are given k_- negative real numbers $x_1, \dots, x_{k_-}$ and k_+ positive real numbers $x_{k_-+1}, \dots, x_{k_-+k_+}$ where $x_1 < \dots < x_{k_-}$ and $x_{k_-+1} < \dots < x_{k_-+k_+}$. Together with the origin they provide a partition of the real line. The unbounded intervals $(-\infty, x_1]$ and $[x_{k_-+k_+}, \infty)$ are called the *boundary* part of the real line. The *origin* part is given by $[x_{k_-}, 0)$ and $[0, x_{k_-+1})$, and, finally, the *inner* part consists of all the rest.
For the distances between these points two different settings will be used:

- *Equidistant* spacing: The points are equidistant such that for $d_+ := (x_{k_-+k_+} - x_{k_-+1})/(k_+ - 1)$ and $d_- := (x_{k_-} - x_1)/(k_- - 1)$ we have

$$x_i = x_{k_-+1} + (i - (k_- + 1))d_+ \ , \ i \in \{k_- + 1, \dots, k_- + k_+\}$$

and

$$x_i = x_1 + (i-1)d_- \ , \ i \in \{1, \dots, k_-\}.$$

- *Geometric* spacing: The distance increases by a factor σ_+ resp. $1/\sigma_-$ as the points tend to become bigger in absolute value. Formally, this means that for $d_{\sigma_+} := (x_{k_-+k_+} - x_{k_-+1})(1-\sigma^+)/(1-\sigma_+^{k_+-1})$ and $d_{\sigma_-} := (x_{k_-} - x_1)(1 - \sigma_-)/(1-\sigma_-^{k_--1})$ we have

$$x_i = x_{k_-+1} + \sigma_+^{i-(k_-+1)} d_{\sigma_+} \ , \ i \in \{k_- + 1, \dots, k_- + k_+\}$$

and

$$x_i = x_1 + \sigma_-^{i-1} d_{\sigma_-} \ , \ i \in \{1, \dots, k_-\}.$$

The constant d_{σ_-}, for instance, comes about as follows: Given

$$x_2 = x_1 + d, \quad x_3 = x_2 + \sigma_- d, \quad \dots \quad x_{k_-} = x_{k_--1} + (\sigma_-)^{k_--2} d$$

summation over all intervals yields

$$x_{k_-} - x_1 = \sum_{i=1}^{k_- -1} (x_{i+1} - x_i) = d(1 + \sigma_- + \sigma_-^2 + \ldots + \sigma^{k_- -2}) = d\frac{1 - \sigma_-^{k_- -1}}{1 - \sigma_-}.$$

The only solution is $d = d_{\sigma_-}$. We must require $\sigma_+ > 1$ and $0 < \sigma_- < 1$ in order to obtain the desired effect of an increase of the length of the intervals towards minus and plus infinity. This effect turns out to be quite useful in obtaining solutions for the optimization problem in cases where the search algorithm fails to find a solution for equidistant spacing.

The task is now to define a set $\mathcal{Y}$ of measure change functions, which lead to a tractable option pricing formula. In order to separate integrability problems every function $y \in \mathcal{Y}$ is split into three parts according to the given partition:

$$y(x) = y^I(x) + y^O(x) + y^{B-}(x) + y^{B+}(x)$$

where y^O defines the part around the origin, y^{B-} resp. y^{B+} stand for the behaviour of y for values which are great in absolute value, and y^I, which covers the inner part, fills the gap between the two. All three functions are positive on the real line but with different support.
For the inner part we take a piecewise linear function whose kinks are determined by the given partition of the real line. Defining the index set by $K = \{1, \ldots, k_- - 1, k_- + 1, \ldots, k_- + k_+ - 1\}$ we define

$$y^I(x) := \sum_{i \in K} \left[y_i + (x - x_i) \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right] \mathbf{1}_{[x_i, x_{i+1})}(x)$$

such that $y^I(x_i) = y_i$ for positive numbers y_i where $i \in K$.

For the origin we take the indicator function of the inner part, i.e.

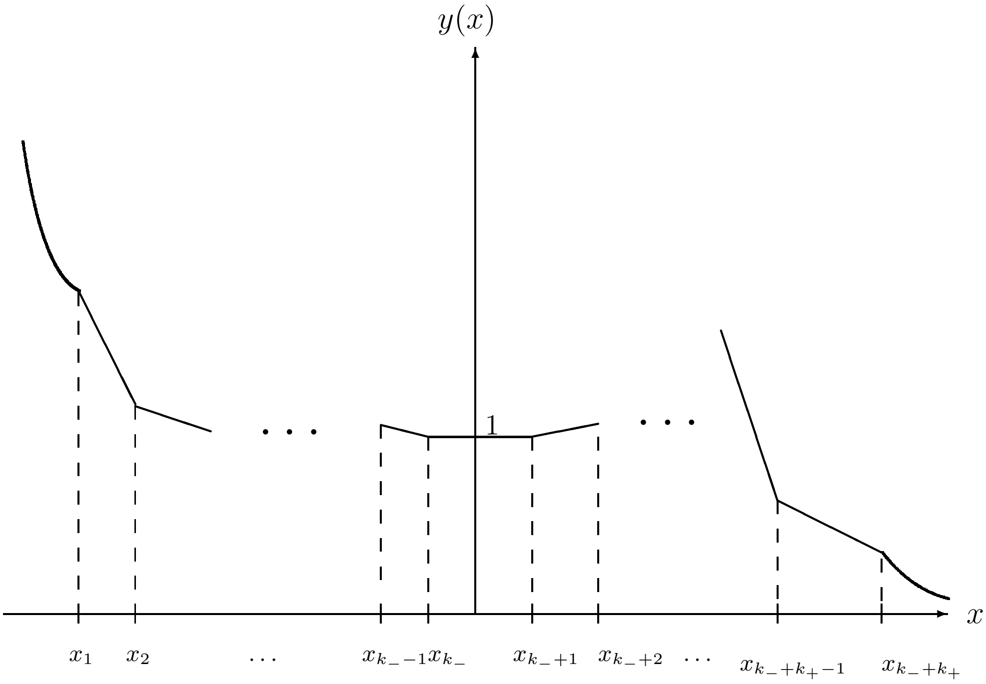
$$y^O(x) = \mathbf{1}_{[x_{k_-}, x_{k_-+1})}(x).$$

Note that (1.4.6) enforces $y(0) = 1$ for a general Lévy measure. For the boundary part we assume an exponential behaviour with a fixed 'exogeneous' coefficient in the exponent. We consider two types of qualitative boundary behaviour: *rising* or *falling*, corresponding to different monotonicity behaviour of y, i.e. whether y is symmetric or asymmetric. Depending on the sign of the coefficient the function either rises to infinity or approaches zero. Thus we resume the corresponding discussion in Chapter 2 about the shape of the measure change function. We have

$$y^{B+}(x) = y_{k_-+k_+} e^{\theta_+(x - x_{k_-+k_+})} \mathbf{1}_{[x_{k_-+k_+}, \infty)}(x) \tag{3.2.1}$$

resp.

$$y^{B-}(x) = y_1 e^{\theta_-(x_1 - x)} \mathbf{1}_{(-\infty, x_1)}(x) \tag{3.2.2}$$



where θ_+ and θ_- are positive constants in the rising case and negative constants in the falling case[3]. In any case they are required to satisfy $\theta_+ < \lambda_+ - 1$ and $\theta_- > -\lambda_-$[4]. We now go on to show that for $y \in \mathcal{Y}$ the conditions of (1.22) are satisfied.

Definition 3.1. *A function $y \in \mathcal{Y}$ is called* linex measure change function *due to its partly linear, partly exponential shape. The measure $\mathbf{Q}^y$ obtained from $\mathbf{P}$ by a measure change function of the form (3.3.3) is called linex measure.*

We make a remark concerning notation: Given fixed coefficients θ_+ and θ_- of the boundary part of a linex measure change function, there is a one-to-one correspondence between linex measure change functions $x \to y(x)$ and positive vectors $y \in \mathbb{R}^{k_- + k_+}$ with $y_{k_-} = y_{k_- + 1} = 1$ we will freely use the notation y for both objects. Sometimes we will even consider y as vector in $\mathbb{R}^{k_- + k_+ - 2}$ in order to get rid of the two components equal to 1. Which object is meant in each case will be obvious from the context.

Proposition 3.2. *Every linex measure change function y satisfies conditions (1.4.6) and (1.4.7).*

---

[3]The required linearity of the constraints of the minimization problem makes it impossible to endogenize θ_+ and θ_-.

[4]The conditions for θ_+ and θ_- are such that the functions A_+, A_-, B_+, B_- to be defined in Section 3.3 are well-defined.

Proof. For the positive boundary integral of (1.4.6) we have

$$\int_{x_{k_-+k_+}}^{\infty} (1-\sqrt{e^{\theta_+ x}})^2 x^{-1-\nu} e^{-\lambda_+ x} dx \leq$$
$$\int_{\min(x_{k_-+k_+},1)}^{1} (1-\sqrt{e^{\theta_+ x}})^2 x^{-1-\nu} e^{-\lambda_+ x} dx + \int_{\max(x_{k_-+k_+},1)}^{\infty} e^{\theta_+ x} x^{-1-\nu} e^{-\lambda_+ x} dx.$$

The inequality is valid because $0 \leq (1-\sqrt{e^{\theta_+ x}})^2 = (1-2\sqrt{e^{\theta_+ x}}) + e^{\theta_+ x} \leq e^{\theta_+ x}$ for $x > 0$. The first integral poses no problem because it is either zero or finite because the integration of the continuous integrand is over a compact interval. The finiteness of the second integral can be shown with the same reasoning as above given the restrictions on θ_+ and θ_-. The negative boundary works the same way. Hence the above integral is finite.
The integral (1.4.6) in the neighbourhood of the origin is zero, and the middle part means again integrating a continuous function over a compact interval.
For the central part of y we have to check the integrability of a constant, the middle part requires consideration of a function $y = mx+b$ - m and b are arbitrary constants in $\mathbb{R}$ - whereas for the boundary integral to be finite the exponential function must be integrated.
For (1.4.7) we have for the positive boundary integrals

$$\int_{x_{k_-+k_+}}^{\infty} (e^x - 1) e^{\theta_+ x} x^{-1-\nu} e^{-\lambda_+ x} dx < \infty. \tag{3.2.3}$$

This is because the integrand $(e^x-1)e^{\theta_+ x} x^{-1-\nu} e^{-\lambda_+ x} = (e^{(\theta_+-\lambda_++1)x} - e^{(\theta_+-\lambda_+)x}) x^{-1-\nu}$ satisfies

$$\lim_{x\to\infty} \frac{(e^{(\theta_+-\lambda_++1)x} - e^{(\theta_+-\lambda_+)x}) x^{-1-\nu}}{1/x^2} = \lim_{x\to\infty} (e^{(\theta_+-\lambda_++1)x} - e^{(\theta_+-\lambda_+)x}) x^{1-\nu} = 0$$

due to $\theta_+ < \lambda_+ - 1 < \lambda_+$. Hence, the integrand in (3.2.3) is bounded by $1/x^2$ for large x with $\int_a^\infty 1/x^2 dx$ being finite for all $a > 0$. Hence all this results in the existence of the integrals in (1.4.6) and (1.4.7). □

3.3 Formulation of the optimization problem

3.3.1 Technical part

First of all we recall that the integral

$$\int_0^\infty x^{-1+v} e^{-x} dx \tag{3.3.1}$$

converges if and only if $v > 0$. If so, then it is defined as the value of the gamma function Γ at v.

For an analytic expression of the relevant integrals we can make use of the *incomplete* gamma function[5]

$$\Gamma[v,z] = \int_z^\infty x^{-1+v} e^{-x} dx, \quad v > 0, z > 0. \tag{3.3.2}$$

In order to make the notation more comprehensible, it is useful to give simple expressions of the functions A_+, A_-, B_+, and B_- in the four variables a, b, λ, and ν. A_+ and A_- are used for the formulation of the martingale condition:

$$\begin{aligned} A_+[a,b,\theta,p] &:= c\int_a^b (e^x-1)x^p e^{\theta x} x^{-1-\nu} e^{-\lambda_+ x} dx \\ A_-[a,b,\theta,p] &:= c\int_a^b (e^x-1)x^p e^{\theta x} |x|^{-1-\nu} e^{-\lambda_-|x|} dx. \end{aligned}$$

Furthermore the functions B_+ and B_- are helpful for the objective function and the moment conditions:

$$\begin{aligned} B_+[a,b,\theta,p] &:= c\int_a^b x^p e^{\theta x} x^{-1-\nu} e^{-\lambda_+ x} dx \\ B_-[a,b,\theta,p] &:= c\int_a^b x^p e^{\theta x} |x|^{-1-\nu} e^{-\lambda_-|x|} dx. \end{aligned}$$

The range of parameters as well as simple analytic expressions in terms of the incomplete gamma function can be found in the appendix. The superscripts '+' resp. '-' refer to integration along parts of the positive resp. negative real axis such that $b > a > 0$ resp. $0 > b > a$.

Objective function

Given a measure change function y, i.e. a measurable function $\mathbb{R} \to \mathbb{R}_+$, the measure change process is given by

$$Z_t = \left.\frac{d\mathbf{Q}}{d\mathbf{P}}\right|_{\mathcal{F}_t} = \mathcal{E}(N)_t \tag{3.3.3}$$

where $N_t = (y(x)-1) * (\mu^X - \nu^{\mathbf{P}})_t$ is the jump-type stochastic integral process of the time-independent random function $y(x) - 1$. This means that y determines entirely the new probability measure $\mathbf{Q}$, and we can express the relative entropy (3.1.3) of $\mathbf{Q}$ with respect to $\mathbf{P}$ in terms of the measure change function by the following lemma:

Lemma 3.3. *The relative entropy process has the following representation in terms of the measure change function y*

$$I_t(\mathbf{Q},\mathbf{P}) = t\int_{-\infty}^\infty (y(x)\log y(x) - (y(x)-1))K(dx) < \infty. \tag{3.3.4}$$

[5] See Abramowitz and Stegun (1972), p.260, 6.5.1 and 6.5.2.

Proof. See Cont and Tankov (2004a). □

The existence of the integral (3.3.4) for a linex measure change function y follows from the following discussion. Note that the integrand is always nonnegative[6], and it is zero if and only if $y(x) \equiv 1\ K(dx) - a.s.$ Unfortunately this integral cannot easily be calculated for the inner part. By defining

$$f(x) := \left(y^I(x)\log(y^I(x)) - (y^I(x)-1)\right)k(x) \tag{3.3.5}$$

and $f_i := f(x_i)$ for $i \in K$ where $y^I(x_i) = y_i$, we can approximate the integral in the inner part by the trapezoidal rule.

$$\begin{aligned}
&\int_{x_{k_-+1}}^{x_{k_-+k_+}} (y^I(x)\log y^I(x) - (y^I(x)-1))k_+(dx) \\
=& \sum_{i=k_-+1}^{k_-+k_+-1} \int_{x_i}^{x_{i+1}} f(x)dx \\
\approx& \frac{1}{2}\sum_{i=k_-+1}^{k_-+k_+-1} (x_{i+1}-x_i)(f_{i+1}+f_i) \\
=& \frac{1}{2}\left((x_{k_-+2}-x_{k_-+1})f_{k_-+1} + \sum_{i=k_-+2}^{k_-+k_+-1}(x_{i+1}-x_{i-1})f_i\right) \\
&+\frac{1}{2}(x_{k_-+k_+}-x_{k_-+k_+-1})f_{k_-+k_+} \\
=& \frac{1}{2}\left(\sum_{i=k_-+2}^{k_-+k_+-1}(x_{i+1}-x_{i-1})f_i + (x_{k_-+k_+}-x_{k_-+k_+-1})f_{k_-+k_+}\right).
\end{aligned}$$

The last equality is a sort of telescope sum[7]. It holds because $f_{k_-+1} = 0$ due to $y_{k_-+1} = 1$. Likewise we have $f_{k_-} = 0$ for the following term for the integration on

[6] See Abramowitz and Stegun (1972), p.68, 4.1.33.

[7] For the negative integral this can be seen by

$$\begin{aligned}
&\sum_{i=1}^{k_--1}(x_{i+1}-x_i)(f_{i+1}+f_i) \\
=& (x_2-x_1)(f_2+f_1) + (x_3-x_2)(f_3+f_2) + (x_4-x_3)(f_4+f_3) + \ldots + \\
& (x_{k_--1}-x_{k_--2})(f_{k_--1}+f_{k_--2}) + (x_{k_-}-x_{k_--1})(f_{k_-}+f_{k_--1}) \\
=& (x_2-x_1)f_1 + (x_3-x_1)f_2 + (x_4-x_2)f_3 + \ldots + (x_{k_-}-x_{k_--2})f_{k_--1} + (x_{k_-}-x_{k_--1})f_{k_-}
\end{aligned}$$

the negative real line:

$$\begin{aligned}
&\int_{x_1}^{x_{k_-}} (y^I(x)\log y^I(x) - (y^I(x)-1))k_-(dx)\\
&\approx \frac{1}{2}\left((x_2-x_1)f_1 + \sum_{i=2}^{k_--1}(x_{i+1}-x_{i-1})f_i\right).
\end{aligned}$$

Especially easy is the origin part where due to the constancy of y the relative entropy is zero. The boundary part can again exactly be calculated. Purely formal calculation for the negative part yields

$$\begin{aligned}
&\int_{-\infty}^{x_1} (y^{B_-}(x)\log y^{B_-}(x) - (y^{B_-}(x)-1))k_-(x)dx\\
= &\int_{-\infty}^{x_1}\left(y_1 e^{\theta_-(x_1-x)}\log\left(y_1 e^{\theta_-(x_1-x)}\right) - (y_1 e^{\theta_-(x_1-x)}-1)\right)k_-(x)dx\\
= &\int_{-\infty}^{x_1}\left((\log y_1 - 1)y_1 e^{\theta_-(x_1-x)} + \theta_-(x_1-x)y_1 e^{\theta_-(x_1-x)} + 1\right)k_-(x)dx\\
= &(\log y_1 - 1 + \theta_- x_1)y_1 e^{\theta_- x_1}\int_{-\infty}^{x_1} e^{-\theta_- x}k_-(x)dx - \theta_- y_1 e^{\theta_- x_1}\int_{-\infty}^{x_1} x e^{-\theta_- x}k_-(x)dx\\
&+\int_{-\infty}^{x_1} k_-(x)dx\\
= &(\log y_1 - 1 + \theta_- x_1)y_1 e^{\theta_- x_1}B_-(-\infty, x_1, -\theta_-, 0) - \theta_- y_1 e^{\theta_- x_1}B_-(-\infty, x_1, -\theta_-, 1)\\
&+B_-(-\infty, x_1, 0, 0)\\
=: &f^-(y_1)
\end{aligned}$$

As all these terms in the last line are finite we can read the set of equations backwards and confirm the finiteness of the relative entropy associated with y. Likewise we obtain for the positive part

$$\begin{aligned}
&\int_{x_{k_-+k_+}}^{\infty} (y^{B_+}(x)\log y^{B_+}(x) - (y^{B_+}(x)-1))k_+(x)dx\\
= &(\log y_{k_-+k_+} - 1 - \theta_+ x_{k_-+k_+})y_{k_-+k_+}e^{-\theta_+ x_{k_-+k_+}}B_+(x_{k_-+k_+}, \infty, \theta_+, 0)\\
&+\theta_+ y_{k_-+k_+}e^{-\theta_+ x_{k_-+k_+}}B_+(x_{k_-+k_+}, \infty, \theta_+, 1) + B_+(x_{k_-+k_+}, \infty, 0, 0)\\
=: &f^+(y_{k_-+k_+})
\end{aligned}$$

Hence we can write down the approximate relative entropy $\tilde{\mathcal{I}}(y) = \tilde{\mathcal{I}}(\mathbf{Q}, \mathbf{P})$ in terms of the y_i. Via definition of

$$\tilde{f}_i := \tilde{f}(y_i) := (y_i \log y_i - (y_i - 1))$$

and $K' = K - \{1, k_-, k_- + 1, k_- + k_+\}$ we obtain

$$\begin{aligned}
\tilde{\mathcal{I}}(y) = &\frac{1}{2}[(x_2-x_1)\tilde{f}(y_1)k(x_1) + \sum_{i\in K'}(x_{i+1}-x_{i-1})\tilde{f}(y_i)k(x_i)\\
&+(x_{k_-+k_+} - x_{k_-+k_+-1})\tilde{f}(y_{k_-+k_+})k(x_{k_-+k_+})] + f^-(y_1) + f^+(y_{k_-+k_+}).
\end{aligned}$$

Martingale condition

The aim is to get the martingale condition in an analytic form, i.e. to calculate

$$\int_{-\infty}^{\infty} (e^x - 1)y(x)K(dx) = r - b. \tag{3.3.6}$$

To start with, the measure change function y for the inner part can be written as

$$y^I(x) = \sum_{i \in K} [(y_i - x_i h_i(y_{i+1} - y_i)) + h_i(y_{i+1} - y_i)x] \mathbf{1}_{[x_i, x_{i+1})}(x) \tag{3.3.7}$$

using $h_i := 1/(x_{i+1} - x_i)$, and we obtain

$$\begin{aligned}
&\int_{x_{k_-+1}}^{x_{k_-+k_+}} (e^x - 1)y^I(x)k_+(x)dx \\
&= c \sum_{i=k_-+1}^{k_-+k_+-1} \int_{x_i}^{x_{i+1}} (e^x - 1)[(y_i - x_i h_i(y_{i+1} - y_i)) + h_i(y_{i+1} - y_i)x] x^{-1-\nu} e^{-\lambda_+ x} dx \\
&= \sum_{i=k_-+1}^{k_-+k_+-1} (A_+[x_i, x_{i+1}, 0, 0](y_i - x_i h_i(y_{i+1} - y_i)) + A_+[x_i, x_{i+1}, 0, 1]h_i(y_{i+1} - y_i)) \\
&= \sum_{i=k_-+1}^{k_-+k_+-1} \{(A_+[x_i, x_{i+1}, 0, 0](1 + x_i h_i) - A_+[x_i, x_{i+1}, 0, 1]h_i)\, y_i \\
&\quad + (A_+[x_i, x_{i+1}, 0, 1] - A_+[x_i, x_{i+1}, 0, 0]x_i)\, h_i y_{i+1}\}
\end{aligned}$$

and likewise

$$\begin{aligned}
&\int_{x_1}^{x_{k_-}} (e^x - 1)y^I(x)k_-(x)dx \\
&\quad = \sum_{i=1}^{k_- -1} \{(A_-[x_i, x_{i+1}, 0, 0](1 + x_i h_i) - A_-[x_i, x_{i+1}, 0, 1]h_i)\, y_i \\
&\quad\quad + (A_-[x_i, x_{i+1}, 0, 1] - A_-[x_i, x_{i+1}, 0, 0]x_i)\, h_i y_{i+1}\}
\end{aligned}$$

For the origin part we have

$$\begin{aligned}
\int_{x_{k_-}}^{0} (e^x - 1)y^O(x)k_-(x)dx \quad &+ \quad \int_0^{x_{k_-+1}} (e^x - 1)y^O(x)k_+(x)dx \\
&= \quad A_-[x_{k_-}, 0, 0, 0] + A_+[0, x_{k_-+1}, 0, 0].
\end{aligned}$$

Finally, the boundary region is equal to

$$\begin{aligned}
&\int_{-\infty}^{x_1}(e^x-1)y^{B-}(x)k_-(x)dx + \int_{x_{k_-+k_+}}^{\infty}(e^x-1)y^{B+}(x)k_+(x)dx \\
= \;& c\int_{-\infty}^{x_1}(e^x-1)y_1e^{\theta_-(x_1-x)}|x|^{-1-\nu}e^{-\lambda_-|x|}dx \\
&+c\int_{x_{k_-+k_+}}^{\infty}(e^x-1)y_{k_-+k_+}e^{\theta_+(x-x_{k_-+k_+})}x^{-1-\nu}e^{-\lambda_+x}dx \\
= \;& A_-[-\infty,x_1,-\theta_-,0]e^{\theta_-x_1}y_1 + A_+[x_{k_-+k_+},\infty,\theta_+,0]e^{-\theta_+x_{k_-+k_+}}y_{k_-+k_+}.
\end{aligned}$$

Using all this information we can define the vector $\bar{\mu} \in \mathbb{R}^{k_-+k_+-2}$ of coefficients of the values y_i:

$$\bar{\mu}_i := \begin{cases}
(A_-[x_1,x_2,0,0](1+x_1h_1) - A_-[x_1,x_2,0,1]h_1) + \\ A_-[-\infty,x_1,-\theta_-,0]e^{\theta_-x_1} & i=1 \\[2ex]
(A_-[x_i,x_{i+1},0,0](1+x_ih_i) - A_-[x_i,x_{i+1},0,1]h_i) + \\ (A_-[x_{i-1},x_i,0,1] - A_-[x_{i-1},x_i,0,0]x_{i-1})\, h_{i-1} & i=2,\dots,k_--1 \\[2ex]
(A_+[x_i,x_{i+1},0,0](1+x_ih_i) - A_+[x_i,x_{i+1},0,1]h_i) + \\ (A_+[x_{i-1},x_i,0,1] - A_+[x_{i-1},x_i,0,0]x_{i-1})\, h_{i-1} & i=k_-+2,\dots, \\ & \quad k_-+k_+-1 \\[2ex]
\left\{A_+[x_{k_-+k_+-1},x_{k_-+k_+},0,1] - \right. \\ \left. A_+[x_{k_-+k_+-1},x_{k_-+k_+},0,0]x_{k_-+k_+-1}\right\} h_{k_-+k_+-1} + \\ A_+[x_{k_-+k_+},\infty,\theta_+,0]e^{-\theta_+x_{k_-+k_+}} & i=k_-+k_+.
\end{cases}$$

Observe that in this representation we have already accounted for the condition $y_{k_-} = y_{k_-+1} = 1$. Hence $\bar{\mu}$ has dimension k_-+k_+-2 and not k_-+k_+. We define the real constant $\bar{\beta}$ by:

$$\begin{aligned}
\bar{\beta} \;=\;& r - b - (A_-[x_{k_--1},x_{k_-},0,1] - A_-[x_{k_--1},x_{k_-},0,0]x_{k_--1})h_{k_--1} \\
&-A_+[x_{k_-+1},x_{k_-+2},0,0](1+x_{k_-+1}h_{k_-+1}) + A_+[x_{k_-+1},x_{k_-+2},0,1]h_{k_-+1} \\
&-A_-[x_{k_-},0,0,0] - A_+[0,x_{k_-+1},0,0].
\end{aligned}$$

Then the drift parameter b in the Lévy triplet is given by[8]

$$\begin{aligned}
b \;&=\; \mu - \int_{-\infty}^{\infty} xK(dx) \\
&=\; \mu + c\int_0^{\infty} x^{-1+(1-\nu)}e^{-\lambda_-x}dx - c\int_0^{\infty} x^{-1+(1-\nu)}e^{-\lambda_+x}dx \\
&=\; \mu + c\,(\lambda_-^{\nu-1} - \lambda_+^{\nu-1})\Gamma(1-\nu).
\end{aligned}$$

[8]For the parameter $\mu = E^{\mathbf{P}}[X_1]$ see Section 1.5.2.

Moment conditions

The objective is to compute the risk-neutral moments in a preferably simple way in terms of the measure change function y resp. the values $(y_j)_{j=1,\dots,k_-+k_+}$. The starting point is provided by the standardized central moments of the distribution of X_1, which correspond to the log returns over one period of the considered asset price process. These moments, namely mean value *mean*, volatility *vol*, skewness *skew*, and kurtosis *kurt* are the ones which can intuitively be grasped and compared to the respective values of other assets. Given the mean and standard deviation of a distribution they are defined according to

$$skew = \frac{\mu_3}{vol^3} \quad \text{and} \quad kurt = \frac{\mu_4}{vol^4}$$

where μ_n, $n = 1, \dots, 4$ denote the central moments. Put differently we have

$$\mu_1 = mean, \quad \mu_2 = vol^2, \quad \mu_3 = skew * vol^3, \quad \mu_4 = kurt * vol^4. \tag{3.3.8}$$

It is well-known that the n-th moment is obtained by n-fold differentiation of the characteristic function $z \to E^{\mathbf{Q}}[e^{izX_1}]$ evaluated at $z = 0$ whereas the cumulants of order n correspond to the n-th derivative of the cumulant function at $z = 0$ according to (3.1.1). Putting this together we can represent the cumulants in terms of the moments:[9]

$$\kappa_1 = \mu_1, \quad \kappa_2 = \mu_2, \quad \kappa_3 = \mu_3, \quad \kappa_4 = \mu_4 - 3\mu_2^2. \tag{3.3.9}$$

The cumulants are now the quantities whose theoretical values can be relatively comfortably computed in terms of $(y_j)_{j=1,\dots,k_-+k_+}$ by the formula (3.1.1). The following inequality holds:

$$\int_{-\infty}^{\infty} |x|^k y(x)k(x)dx < \infty \quad \forall\, k \in \mathbb{N}. \tag{3.3.10}$$

Given the parameter restrictions the only thing which is to be checked is the finiteness around zero. But this is clear due to $k-\nu > 0$ for $k \geq 1$ which is necessary and - given the admissible range of parameters - sufficient for the convergence of the integrals that possess the structure of the gamma function and that will be calculated in the sequel. Due to Lemma 1.17 equation (3.3.10) entails the existence of all moments of integer order of the risk-neutral probability measure. We observe that X is no longer a tempered stable Lévy process under $\mathbf{Q}$.

Lemma 3.4.

$$\psi_{\mathbf{Q}}^{(k)}(0) = i^k \int_{-\infty}^{\infty} x^k y(x)k(x)dx \quad \forall\, k \in \mathbb{N} \setminus \{1\}. \tag{3.3.11}$$

Proof. We have to differentiate the cumulant function of the risk-neutral probability measure $\mathbf{Q}$, i.e.

$$\psi_{\mathbf{Q}}(z) = ib'z + \int_{-\infty}^{\infty} (e^{izx} - 1)K'(dx).$$

[9] See e.g. Abramowitz and Stegun (1972), 26.1.13.

The k-th derivative of the integrand is $(ix)^k e^{izx}$ and we have $|(ix)^k e^{izx}| = |x|^k$ for any non-trivial interval containing zero. As seen above $|x|^k$ is K'-integrable. Hence according to the differentiation lemma (see Bauer (1992), Lemma 16.2) we can interchange differentiation and integration. For $z = 0$ the result is obtained. □

Formula (3.1.1) and Lemma 3.4 provide us with analytical expressions for the cumulants of the distribution of X_1 under the risk-neutral probability measure $\mathbf{Q}$. Examining the integral $\int_{-\infty}^{\infty} x^k y(x)k(x)dx$ more closely we obtain for the positive inner part the expression

$$\begin{aligned}
&\int_{x_{k_-+1}}^{x_{k_-+k_+}} x^k y^I(x)k_+(x)dx \\
&= c\sum_{i=k_-+1}^{k_-+k_+-1}\int_{x_i}^{x_{i+1}} x^k[(y_i - x_ih_i(y_{i+1}-y_i)) + h_i(y_{i+1}-y_i)x]x^{-1-\nu}e^{-\lambda_+ x}dx \\
&= \sum_{i=k_-+1}^{k_-+k_+-1}(B_+[x_i,x_{i+1},0,k](y_i - x_ih_i(y_{i+1}-y_i)) + B_+[x_i,x_{i+1},0,k+1]h_i(y_{i+1}-y_i)) \\
&= \sum_{i=k_-+1}^{k_-+k_+-1}\{(B_+[x_i,x_{i+1},0,k](1+x_ih_i) - B_+[x_i,x_{i+1},0,k+1]h_i)\,y_i \\
&\quad + (B_+[x_i,x_{i+1},0,k+1] - B_+[x_i,x_{i+1},0,k]x_i)\,h_iy_{i+1}\}\,,
\end{aligned}$$

and for the negative one

$$\begin{aligned}
&\int_{x_1}^{x_{k_-}} x^k y^I(x)k_-(x)dx \\
&= \sum_{i=1}^{k_--1}\{(B_-[x_i,x_{i+1},0,k](1+x_ih_i) - B_-[x_i,x_{i+1},0,k+1]h_i)\,y_i \\
&\quad + (B_-[x_i,x_{i+1},0,k+1] - B_-[x_i,x_{i+1},0,k]x_i)\,h_iy_{i+1}\}\,.
\end{aligned}$$

For the neighbourhood around the origin and for the boundary part we obtain

$$\begin{aligned}
\int_{x_{k_-}}^{0} x^k y^O(x)k_-(x)dx \quad &+ \quad \int_0^{x_{k_-+1}} x^k y^O(x)k_+(x)dx \\
&= \quad B_-[x_{k_-},0,0,k] + B_+[0,x_{k_-+1},0,k]
\end{aligned}$$

and

$$\begin{aligned}
&\int_{-\infty}^{x_1} x^k y^{B-}(x) k_-(x)dx + \int_{x_{k_-+k_+}}^{\infty} x^k y^{B+}(x) k_+(x)dx \\
&= c\int_{-\infty}^{x_1} x^k y_1 e^{\theta_-(x_1-x)}|x|^{-1-\nu} e^{-\lambda_-|x|}dx \\
&\quad + c\int_{x_{k_-+k_+}}^{\infty} x^k y_{k_-+k_+} e^{\theta_+(x-x_{k_-+k_+})} x^{-1-\nu} e^{-\lambda_+ x}dx \\
&= B_-[-\infty, x_1, -\theta_-, k] e^{\theta_- x_1} y_1 + B_+[x_{k_-+k_+}, \infty, \theta_+, k] e^{-\theta_+ x_{k_-+k_+}} y_{k_-+k_+}.
\end{aligned}$$

Analogously to the case of the martingale condition the vectors $\mu^{(n)} \in \mathbb{R}^{k_-+k_+-2}$, $n = 2, 3, 4$, of coefficients of the values y_i are the following ones:

$$\bar{\mu}_i^{(n)} := \begin{cases}
(B_-[x_1, x_2, 0, n](1 + x_1 h_1) - B_-[x_1, x_2, 0, n+1]h_1) & \\
+B_-[-\infty, x_1, -\theta_-, n]e^{\theta_- x_1} & i = 1 \\
& \\
(B_-[x_i, x_{i+1}, 0, n](1 + x_i h_i) - B_-[x_i, x_{i+1}, 0, n+1]h_i) & \\
+ (B_-[x_{i-1}, x_i, 0, n+1] - B_-[x_{i-1}, x_i, 0, n]x_{i-1})\, h_{i-1} & i = 2, \ldots, k_- - 1 \\
& \\
(B_+[x_i, x_{i+1}, 0, n](1 + x_i h_i) - B_+[x_i, x_{i+1}, 0, n+1]h_i) & \\
+ (B_+[x_{i-1}, x_i, 0, n+1] - B_+[x_{i-1}, x_i, 0, n]x_{i-1})\, h_{i-1} & i = k_- + 2, \ldots, \\
& \quad k_- + k_+ - 1 \\
& \\
\{B_+[x_{k_-+k_+-1}, x_{k_-+k_+}, 0, n+1] - & \\
B_+[x_{k_-+k_+-1}, x_{k_-+k_+}, 0, n]x_{k_-+k_+-1}\}\, h_{k_-+k_+-1} & \\
+B_+[x_{k_-+k_+}, \infty, \theta_+, n]e^{-\theta_+ x_{k_-+k_+}} & i = k_- + k_+
\end{cases}$$

The real constants $\bar{\beta}^{(n)}$ are equal to

$$\begin{aligned}
\bar{\beta}^{(n)} &= \kappa_n - (B_-[x_{k_--1}, x_{k_-}, 0, n+1] - B_-[x_{k_--1}, x_{k_-}, 0, n]x_{k_--1})h_{k_--1} \\
&\quad - B_+[x_{k_-+1}, x_{k_-+2}, 0, n](1 + x_{k_-+1}h_{k_-+1}) + B_+[x_{k_-+1}, x_{k_-+2}, 0, n+1]h_{k_-+1} \\
&\quad - B_-[x_{k_-}, 0, 0, n] - B_+[0, x_{k_-+1}, 0, n].
\end{aligned}$$

3.3.2 The optimization problem and regularity conditions

As mentioned, in the previous calculations we have already accounted for the condition $y_{k_-} = y_{k_-+1} = 1$ which comes from our assumption $y(0) = 1$. That is why the vector-valued decision variable y has dimension $k_- + k_+ - 2$. To indicate this, we use the index j instead of i.
For a concise representation of the martingale and moment constraints of the minimization problem set $\bar{\mu} = (\bar{\mu}_j)_{j=1,\ldots,k_-+k_+-2}$, $\bar{\mu}^{(n)} = (\bar{\mu}_j^{(n)})_{j=1,\ldots,k_-+k_+-2}$ for $n = 2, 3, 4$, and $M = [\bar{\mu}^{(2)}, \bar{\mu}^{(3)}, \bar{\mu}^{(4)}, \bar{\mu}]' \in \mathbb{R}^{4\times(k_-+k_+-2)}$. Moreover, we define the vector $\beta = [\bar{\beta}^{(2)}, \bar{\beta}^{(3)}, \bar{\beta}^{(4)}, \bar{\beta}]' \in \mathbb{R}^4$.

Some additional constraints have to be imposed: First, all y_j must be positive by definition of an absolutely continuous equivalent change of measure. For numerical reasons - the relative entropy is only defined for strictly positive measure change functions - we impose the stronger condition that all components of y are bounded below by some $\epsilon > 0$. Secondly, we must have in mind that after accounting for the martingale condition and the moment conditions it remains still a considerable number of degrees of freedom depending on the number of sampling points $k_- + k_+$. In order to prevent an all too irregular shape of the measure change function we have therefore the freedom to impose some regularity conditions on the shape. We try to impose the *symmetric* shape, which means that we require y to rise as a function of the absolute value of x. However, it turns out that in Section 3.6 for numerical reasons we must admit an exception from this condition in the neighbourhood of zero. Hence, the condition for symmetry reads like this:

$$y_j \geq y_{j+1} \quad \forall\, j = 1, \ldots, k_- - 2 \quad \text{and} \quad y_j \leq y_{j+1} \quad \forall\, j = k_-, \ldots k_- + k_+ - 3.$$

Assuming $\{e_1, \ldots, e_{k_-+k_+-2}\}$ to be the standard basis in $\mathbb{R}^{k_-+k_+-2}$, this means that for

$$\begin{aligned} \phi_j &:= -e_j + e_{j+1} \in \mathbb{R}^{k_-+k_+-2}, \quad j = 1, \ldots k_- - 2; \\ \phi_j &:= e_j - e_{j+1} \in \mathbb{R}^{k_-+k_+-2}, \quad j = k_-, \ldots k_- + k_+ - 3 \end{aligned}$$

we have

$$\Phi\, y \leq 0 \in \mathbb{R}^{k_-+k_+-2}$$

where $\Phi = [\phi_1, \ldots, \phi_{k_-+k_+-4}]' \in \mathbb{R}^{(k_-+k_+-4)\times(k_-+k_+-2)}$ and $y = (y_j)_{j=1,\ldots,k_-+k_+-2}$. Finally, the following minimization problem is set up:

$$\begin{aligned} &\min_{y \in \mathcal{Y}} \tilde{\mathcal{I}}(y) \\ s.t. \quad &My = \beta \\ &\Phi y \leq 0 \\ y_1, \ldots, y_{k_-+k_+-2} &\geq \epsilon \end{aligned} \tag{3.3.12}$$

This means that we look for a positive vector $y = (y_j)_{j=1,\ldots,k_-+k_+-2}$ which minimizes the relative entropy with respect to the original probabaility measure $\mathbf{P}$ under the equality constraints of the martingale and moment conditions (represented by $My = \beta$) and inequality constraint of the regularity condition (represented by $\Phi y \leq 0$). The advantage of using linex measures is that the equality constraints are linear. The only non-linearity in this optimization problem is the objective function.

Remark. The intention of the introduction of linex measures is to make the problem (3.1.5) easier to solve. However, one can also try to solve it for a general measure change function y instead of requiring $y \in \mathcal{Y}$. If one is ready to skip the monotonicity constraints, an application of standard optimization theory shows that the optimal measure change function y^* exhibits the following structure:

$$y^*(x) = e^{\lambda_1(e^x-1)+\lambda_2 x^2+\lambda_3 x^3+\lambda_4 x^4}.$$

The parameters $\lambda_1, \ldots, \lambda_4$ must be such that the martingale equation and the moment constraints are fulfilled. Moreover integrability requires at least λ_1 and λ_4 to be non-positive. Given existence, getting them in an explicit form is a rather difficult task because it involves a non-linear system of four equations which are given by integrals over the entire real line. Linex measures allow to simplify this problem and to make it accessible for option pricing by Fourier inversion.

3.4 Pricing with linex measures

In general, when working on pricing derivative securities in the framework of the mainstream arbitrage theory, one has to solve the following problem: Given the stochastic process of the underlying security under the statistical probability measure **P** one has to find a probability measure **Q** under which the very same process becomes a martingale. **Q** is therefore called martingale measure or risk-neutral measure. This means that there are two distributions, each equipped with its own characteristics such as mean, variance, skewness and kurtosis. Arbitrage-free valuation seen from this point of view amounts to making statements about the interrelation between the two. The key to this problem is provided by the techniques of absolutely continuous changes of measure, i.e. Girsanov's theorems.

The most prominent model is of course the Black-Scholes model, for which this problem is immensely simplified. It assumes a Gaussian distribution for the returns of the underlying with the two parameters mean and volatility, i.e. standard deviation. The standard Girsanov theorem in the Black-Scholes world of geometric Brownian motion now says that there is only one martingale measure which leads inevitably to another normal distribution with (a priori different) risk-neutral mean and volatility. For pricing we need only the parameters of the risk-neutral distribution. One statement of Girsanov's theorem in this case is that the risk-neutral mean is equal to the risk-free interest rate, so we do not have to bother about estimation of the mean. For the risk-neutral volatility we look for the market price of an option (or even of several options) which is in some way similar to the one we want to price. Solving for that volatility parameter which makes the Black-Scholes formula reproduce the observed market price, we obtain the risk-neutral volatility, which the Black-Scholes theory calls implied volatility. Strangely enough, this is not the only way of finding a volatility estimate: Another way is to use the historical volatility estimated from time series data. In general, this is wrong because historical volatility corresponds to the statistical volatility under the measure **P**. However, this is permitted by a second statement of Girsanov's theorem saying that the statistical is equal to the risk-neutral volatility. Hence, the Black-Scholes theory justifies the use of both, and

it is left to the economic considerations of the user to choose one of them.
Yet this is different for models based on Lévy processes, partly because the martingale measure is not unique. For pricing based on Esscher-type martingale measures one estimates carefully the statistical distribution of historical return data, an approach which lies at the heart of the theory of Lévy processes because the phenomenon of skewed and leptokurtotic returns relates first and foremost to statistical returns. Fixing then some martingale measure according to whatever criterion in fact gives arbitrage-free prices. Nevertheless together with the martingale measure these models implicitly pin down all the risk-neutral moments and hence the shape of the volatility smile, which may not correspond to the one obtained by analysing market prices of exchange traded options. All told, caring too much about returns under $\mathbf{P}$ while returns under $\mathbf{Q}$ are actually important is not quite appropriate for pricing derivatives. Seen from this angle the use of Esscher-type martingale measures is even a step behind Black-Scholes pricing combined with implied volatility.
This is now where the linex measures set in. We specify both the statistical and the risk-neutral distribution, the latter only via the second, third and fourth moment, and then try to find a change of measure - based on the sound foundations of the theory of absolutely continuous changes of measure - that succeeds in transforming one distribution into the other. The measure change function, which completely describes the change of measure in the world of purely discontinuous stochastic processes with independent and stationary increments, should be easy to handle, and this is exactly what the shape of the linex measure change function, partly affine linear, partly exponential, is supposed to do. A whole branch of the empirical literature, see e.g. Corrado and Su (1997), deals with the estimation of risk-neutral distributions resp. its moments.
From this angle linex measures provide a natural extension of Black-Scholes implied volatility pricing with the difference that in a Lévy world higher moments and the incompleteness problem require to use more information in the form of both statistical and risk-neutral parameters and a theory of connecting them.

3.5 Recovery of risk-neutral moments

One of the advantages of the linex change of measure is that for its application we just have to have the second, the third and the fourth risk-neutral moments instead of the entire distribution of returns. In theory we can recover these moments exactly. Suppose we want to replicate the payoff $f(F)$ where f is a twice differentiable payoff function and F is the forward price of the underlying security. Then we have from Carr et al. (2000b) the following formula which results from a Taylor expansion with arbitrary expansion point κ:

$$f(F) = f(\kappa) + f'(\kappa)(F-\kappa) + \int_{\kappa_+}^{\infty} f''(K)(F-K)^+ dK + \int_0^{\kappa_-} f''(K)(K-F)^+ dK.$$

where $\kappa_+ \geq \kappa$ and $\kappa_- \leq \kappa$. This equation implies that every payoff can exactly be replicated by static positions in a bond, a forward contract and a continuum of plain-

vanilla call and put options for which a market price is required to exist. Needless to mention that this does not help for practical purposes. Instead we are forced to approximate the occuring integrals by making use of exchange-traded options of different (and usually very few) strikes.
The trick for solving our moment recovery problem is now to see that the definitions of risk-neutral standard deviation, skewness and kurtosis can be interpreted as the risk-neutral expectations of certain contracts, from here on referred to as the moment contracts. Hence the above recipe can be used for this purpose. But under some circumstances we can enormously simplify the problem by finding that there are sometimes traded contracts whose payoff functions are themselves quite similar to those of the moment contracts. For this reason we take a closer look at currency options markets.
It occurs that in markets for currency options there are usually three exchange-traded contracts for several currencies and maturities which represent each bets on volatility, skewness, and kurtosis[10]: the *at-the-money straddle*, the *25-delta risk reversal*, and the *25-delta strangle*. The at-the-money straddle is a combination of two long positions in a standard call and put options which are struck at-the-money. The other two positions each involve a call and a put options which are both out-of-the-money such that their deltas are equal to 0.25. Whereas the risk reversal requires a short position in the put and a long position in the call, the strangle consists of two long positions. This means that the strangle can be seen as a straddle which is pulled apart symmetrically to its strike price. These three contracts give approximations of the moments we are looking for.
In order to make this plausible, we make the simplifying assumptions that the riskless interest rate is equal to zero and that the asset returns $X_t = \log(S_T/S_0)$ in our exponential Lévy model have zero expectation under any flexible martingale measure $\mathbf{Q}$. Maturity dates are always set to $t = T$. Beginning with the straddle with strike $K = S_0$, its price C_{vol} is given by

$$C_{vol} = E^{\mathbf{Q}}[(S_T - S_0)^+] + E^{\mathbf{Q}}[(S_0 - S_T)^+] = E^{\mathbf{Q}}[|S_T - S_0|]. \tag{3.5.1}$$

Under our assumptions the risk-neutral variance of the returns X can be obtained by valuing the contract paying $(\log(S_T/S_0))^2$ at maturity. A Taylor expansion around S_0 leads to

$$\left\{\log\left(\frac{S_T}{S_0}\right)\right\}^2 = \frac{1}{S_0^2}|S - S_0|^2 + O(|S_T - S_0|^3 \quad \text{for} \quad S_T \to S_0.$$

Taking this approximation for granted, we obtain

$$S_0\left[E^{\mathbf{Q}}\left\{\log\left(\frac{S_T}{S_0}\right)\right\}^2\right]^{1/2} \approx \left[E^{\mathbf{Q}}|S_T - S_0|^2\right]^{1/2} \tag{3.5.2}$$

[10]This remark is due to Peter Carr. See also Gereben (2002).

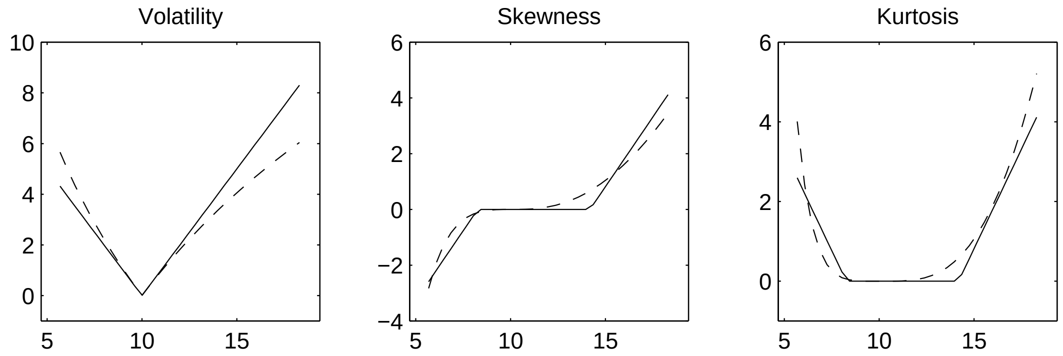


Fig. 3.1. Comparison of payoffs at maturity of moment contracts and exchange-traded combinations where the former are depicted by solid and the latter by dashed lines. From left to right: Approximation of S_0 times volatility by the straddle, of skewness by the risk reversal and of the kurtosis by the strangle. Data of example: $S_0 = 10$, $\sigma_{BS} = 0.4$ and $T = 1$

where the term in square brackets on the left hand side of (3.5.2) is equal to the risk-neutral volatility. Moreover by Jensen's inequality and (3.5.1) we have

$$\left[E^{\mathbf{Q}}|S_T - S_0|^2\right]^{1/2} \leq E^{\mathbf{Q}}|S_T - S_0| = C_{vol}. \tag{3.5.3}$$

Now our line of reasoning becomes less formal: Given that the approximation (3.5.2) is not that bad and the inequality (3.5.3) is rather tight, putting together (3.5.2) and (3.5.3) imply that the payoff of the at-the-money straddle is approximately equal to S_0 times the risk-neutral volatility. But this means that an approximation of the latter can be directly inferred from the market price of the straddle.

The first picture in Figure 3.1 shows the payoff function of the straddle and the one of the left hand side of (3.5.2) for $S_0 = 10$ and $T = 1$. It suggests that the approximation is not that bad.

The standard deviation obtained through this procedure has been determined without assuming any model for the behaviour of the asset prices. Contrary to this procedure, it is market practice to use the implied Black-Scholes volatility σ_{BS} of the straddle as the risk-neutral volatility. This proxy will be used for the scaling of the third and fourth moment to obtain risk-neutral skewness and kurtosis.

The contracts for obtaining the exact values of skewness and kurtosis of the risk-neutral return distribution are

$$\frac{1}{\sigma_{BS}^3} E^{\mathbf{Q}} \left\{ \log\left(\frac{S_T}{S_0}\right) \right\}^3 \quad \text{resp.} \quad \frac{1}{\sigma_{BS}^4} E^{\mathbf{Q}} \left\{ \log\left(\frac{S_T}{S_0}\right) \right\}^4. \tag{3.5.4}$$

We add the assumption of $\sigma_{BS} = 0.4$ to the above example and can once more see from the second and third picture of Figure 3.1 that the risk reversal and strangle should have prices pretty similar to those of the moment contracts in (3.5.4), i.e. to

skewness and kurtosis.
As a matter of course, these conclusions do only make sense if the prices of those three contracts are backed by a sufficient market volume, i.e. if their prices can reasonably be regarded as true market valuations.
It goes without saying that as the above contracts are simple combinations of call and put options, this technique is extendable to all options markets where call and put options are frequently traded. The only problem could be to have the suitable strikes available to synthesize these combinations.

3.6 Implementation and examples

In this section we will illustrate the application of the tool of linex measures by means of an example. We choose a tempered stable Lévy process with parameters $c = 0.3$, $\lambda_- = 3.75$, $\lambda_+ = 4.0$ and $\nu = 0.6$ yielding[11] $vol^{\mathbf{P}} = 0.28$, $skew^{\mathbf{P}} = -0.10$ and $kurt^{\mathbf{P}} = 5.81$ for the statistical standardized central moments. The expected value of the drift is set to a tiny 0.0002 % per year. The number of sampling points on each side of the origin is taken to be eight, and the exponential parts grow with the coefficient $\theta_- = \theta_+ = 1$. The support of the inner part of the measure change function is $[-6.0 \ldots -10^{-8}] \cup [10^{-8} \ldots 6.0]$, which is partitioned by a geometric spacing with parameters $\sigma_- = 0.7$ and $\sigma_+ = 1.3$.
We solve the minimization problem (3.3.12) for four different parameter constellations which differ only in the risk-neutral standardized central moments $vol^{\mathbf{Q}}$, $skew^{\mathbf{Q}}$ and $kurt^{\mathbf{Q}}$. The first case is a benchmark case. The second, third and fourth constellations each change only one moment in order $skew^{\mathbf{Q}}$, $kurt^{\mathbf{Q}}$ and $vol^{\mathbf{Q}}$. The change in $vol^{\mathbf{Q}}$ is a special case, which remains to be seen in the following sections. The left column of Fig. 3.2 gives all the necessary information.

3.6.1 Measure change functions

Theoretically, the problem (3.3.12) is a standard optimization problem. However, there are a number of practical issues to be tackled in order to implement it on the computer. The algorithm to solve (3.3.12) involves a line search at each optimization step, which makes it quite possible to find a vector y with one or more negative components. This means that occasionally the algorithm evaluates the relative entropy for a measure change function which becomes negative somewhere and thus gives a complex-valued result. This of course renders the obtained results useless respectively causes the algorithm to diverge. The constituting function of the relative entropy is $y' \to y' \log(y') - (y' - 1)$. The solution to this problem is to extend this function continuously in way that does not change its monotonicity structure. This is done by

$$y' \to \begin{cases} y' \log(y') - (y' - 1) & , \quad y' > 0 \\ -(y' - 1) & , \quad y' \leq 0. \end{cases} \tag{3.6.1}$$

[11] All considered volatlites are annualized.

Continuity, which is required for an application of the optimization algorithm follows from $\lim_{y' \to 0+} y' \log(y') = 0$.
The number of sampling points of the measure change functions in the example is a compromise between having enough degrees of freedom and a relatively fast computation. The number of variables to be determined must be bigger than the number of equality constraints, i.e.

$$k_+ + k_- - 2 \geq 4 \tag{3.6.2}$$

in order to obtain a solution. The appropriate type of spacing turns out to be the geometric one because in this and in most other examples it proves to be more flexible than the equidistant spacing in finding a feasible linex measure change function, i.e. one that satisfies all the restrictions of (3.3.12).
The determination of the optimal measure change function in the sense of (3.3.12) proceeds as follows: First, assuming (3.6.2), we try to find a feasible solution by solving all 4×4 subsystems of $My = \beta$ that are not numerically singular and yield a positive solution. A positive solution y^* of these smaller problems is in any case a solution of the original problem provided that the vector y^* is extended by zeros at those entries which correspond to eliminated columns. The mean of all these solutions is again a solution and is taken as the initial value for the optimization algorithm. This procedure results in most relevant cases in a positive solution of $My = \beta$. We call this solution the *unconstrained* linex measure.
Beginning with the found initial value we start the optimization algorithm and find four measure change functions y_1, y_2, y_3 and y_4 for each of the four cases we are investigating. They are drawn in the right column of Fig. 3.2.
Unfortunately, we did not succeed in imposing a condition on the shape of the measure change function for the given parameter configurations. More exactly, a condition of this type causes the algorithm to diverge. The four resulting measure change functions y_1, y_2, y_3, and y_4 are depicted in the right column of Fig. 3.2.

3.6.2 Pricing

For a demonstration of the technique of linex measures we have chosen a tempered stable Lévy process because its characteristic function under the linex measure can be analytically computed. The only challenge is to provide a fast algorithm for the computation of the incomplete gamma function for complex arguments to which the cumulant function (3.1.2) is reduced. Given such a representation of the cumulant function we can then effciently price options by the Fourier inversion technique summarized in Section 1.6.
For $\nu < 0$ we have

$$\Gamma(-\nu, z) = \frac{1}{\nu}(z^{-\nu} e^{-z} - \Gamma(1-\nu, z)). \tag{3.6.3}$$

This follows from putting together three formulae from Abramowitz and Stegun (1972): Combining the recurrence formula 6.5.22 for the incomplete gamma function

Lévy process:
c=0.3 λ_-=3.75 λ_+=4 ν=0.6
Statistical parameters:
Vol=0.28 Skew=-0.10 Kurt=5.81
Risk-neutral parameters:
Vol=0.28 Skew=-0.10 Kurt=4
***Linex* specification of y(x):**
k_-=8 k_+=8 θ_-=1 θ_+=1
Inner part: [-6 ... -10^{-8}] $\cup$ [10^{-8} ... 6]
Spacing: geometric σ_-=0.7 σ_+=1.3

Lévy process:
c=0.3 λ_-=3.75 λ_+=4 ν=0.6
Statistical parameters:
Vol=0.28 Skew=-0.10 Kurt=5.81
Risk-neutral parameters:
Vol=0.28 Skew=-0.6 Kurt=4
***Linex* specification of y(x):**
k_-=8 k_+=8 θ_-=1 θ_+=1
Inner part: [-6 ... -10^{-8}] $\cup$ [10^{-8} ... 6]
Spacing: geometric σ_-=0.7 σ_+=1.3

Lévy process:
c=0.3 λ_-=3.75 λ_+=4 ν=0.6
Statistical parameters:
Vol=0.28 Skew=-0.10 Kurt=5.81
Risk-neutral parameters:
Vol=0.28 Skew=-0.10 Kurt=6
***Linex* specification of y(x):**
k_-=8 k_+=8 θ_-=1 θ_+=1
Inner part: [-6 ... -10^{-8}] $\cup$ [10^{-8} ... 6]
Spacing: geometric σ_-=0.7 σ_+=1.3

Lévy process:
c=0.3 λ_-=3.75 λ_+=4 ν=0.6
Statistical parameters:
Vol=0.28 Skew=-0.10 Kurt=5.81
Risk-neutral parameters:
Vol=0.40 Skew=-0.10 Kurt=4
***Linex* specification of y(x):**
k_-=8 k_+=8 θ_-=1 θ_+=1
Inner part: [-6 ... -10^{-8}] $\cup$ [10^{-8} ... 6]
Spacing: geometric σ_-=0.7 σ_+=1.3

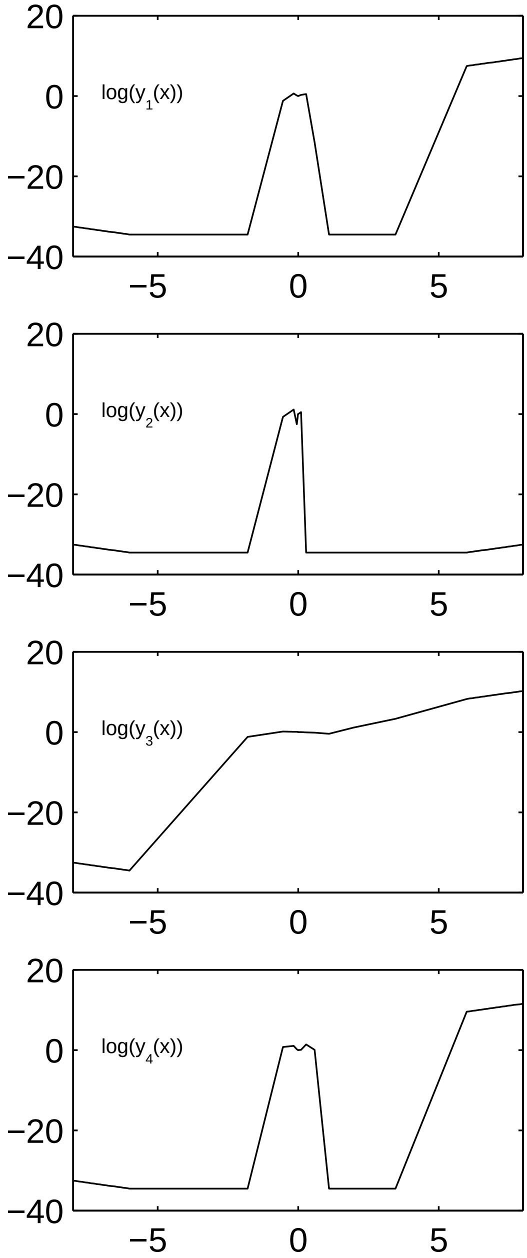


Fig. 3.2. Parameter configurations and implied measure change functions for the four cases: benchmark case, high negative risk-neutral skewness, high risk-neutral kurtosis, high risk-neutral volatility (from top to bottom).

with 6.5.3 and 6.1.17 for the ordinary gamma function we derive (3.6.3). This equality reduces the computational costs by implying that it is enough to evaluate the incomplete gamma function only for the first coefficent $1-\nu$ although the algorithm requires both $\Gamma(-\nu, z)$ and $\Gamma(1-\nu, z)$ to be evaluated for many different values of z. Due to the special form of the linex measure change functions the really time-expensive part of the pricing procedure is the computations of the coefficients of M in (3.3.12). But they depend only on the statistical parameters. Given these parameters we can very fast calculate option prices for different settings of risk-neutral parameters.
We compare the pricing performance of the linex measures with both the Esscher change of measure and Black-Scholes prices. The Esscher transform comes along with the measure change function $y(x) = e^{\theta x}$ where $\theta \in \mathbb{R}$ is the unique solution of the martingale equation

$$b - r + \int_{-\infty}^{\infty} (e^x - 1)e^{\theta x} K(dx) = 0$$

which amounts to solve

$$b - r + A_-[-\infty, 0, \theta, 0] + A_+[0, \infty, \theta, 0] = 0 \tag{3.6.4}$$

for θ.

In order to show the inadequacy of the Esscher technique as regards the implied risk-neutral moments, equalities (3.3.8) and (3.3.9) are used the other way round: The cumulants κ_j^{es} for $j = 2, 3, 4$ are easily calculated from the risk-neutral cumulant function such that (3.3.8) and (3.3.9) yield the standardized central moments under $\mathbf{Q}^{es}$, namely

$$vol^{es} = \sqrt{\kappa_2^{es}}, \quad skew^{es} = \frac{\kappa_3^{es}}{(\kappa_2^{es})^{3/2}}, \quad kurt^{es} = 3 + \frac{\kappa_4^{es}}{(\kappa_2^{es})^2}. \tag{3.6.5}$$

For the example we price standard European call options on an underlying asset with current price $S_0 = 20$, time to maturity $T = 1$ year, and a risk-free interest rate of 2% p.a. for strike prices around S_0[12].
Fig. 3.3 shows the results. The left column compares prices obtained with both the linex and the Esscher change of measure with Black-Scholes prices by drawing their difference whereas the right column shows the difference between linex and Esscher measure. The rows show the four cases described in section 3.6.1. The abscissa shows the moneyness K/S_0[13]. Hence out-of-the-money (OTM) options are on the right hand side of 1.
For the volatility parameter in the Black-Scholes formula $vol^{\mathbf{Q}}$ is used, which should

[12] It is understood that the choice of the time to maturity is just for illustrational purposes. Significant deviations from Gaussianity can be observed for daily or intra-day data; hence one would not use the present model for pricing options with as much as one year to maturity.

[13] Recall that in the literature there are multiple definitions of moneyness.

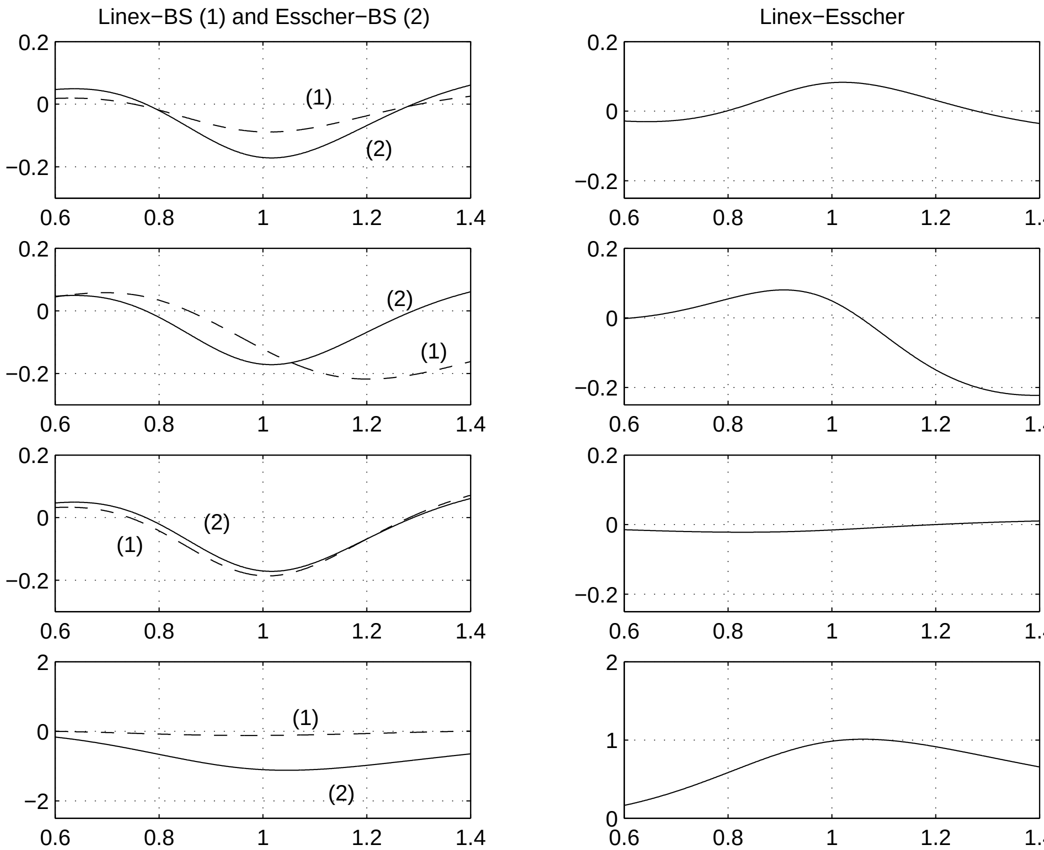


Fig. 3.3. Comparison of all four cases in terms of pricing behaviour: For the option described in the text the left column shows the pricing performance of both the linex measure and the Esscher measure compared to Black-Scholes prices. The right column depicts the price differences between linex and Esscher pricing.

come close to some implicit volatility estimate. By this procedure we try to imitate pricing by implicit rather than historical volatility. It would be quite unfair for the Black-Scholes formula to use the latter one when comparing it to a measure which is adjusted to risk-neutral parameters.

The Esscher corrections of Black-Scholes prices have the form which is typical for Lévy prices. According to a general agreement in the literature the risk-neutral distribution has the same qualitative features as the statistical one - less mass for values around the mean and more mass in the tails compared to the Gaussian distribution - it can be seen from the pricing formula that Lévy prices are lower for options around the at-the-money option and higher for those which are far out-of-the-money. The reason for the latter observation is that in a heavy tail environment OTM options have a greater chance to become ITM than in the Gaussian case.

Linex prices have the same characteristics concerning the direction of Black-Scholes

price corrections. In the benchmark case linex prices, which are by their definition closer to the market, correct Esscher prices in the direction of Black-Scholes prices. Hence Esscher prices deviate too much from Black-Scholes prices for both the over-pricing and the underpricing case.
The case of high risk-neutral skewness gives a highly asymmetric picture. This phenomenon corresponds to a situation where investors have a very pronounced degree of risk-aversion and therefore put more mass on losses in their subjective expectations of future asset prices. In such a situation the traditional instruments for protection against downside risk, OTM put options, have a greater price as compared to the benchmark case. The counterpart of this observation is that OTM call options are relatively cheap. And the latter statement can be verified in Fig. 3.3 where the prices of call options are drawn.

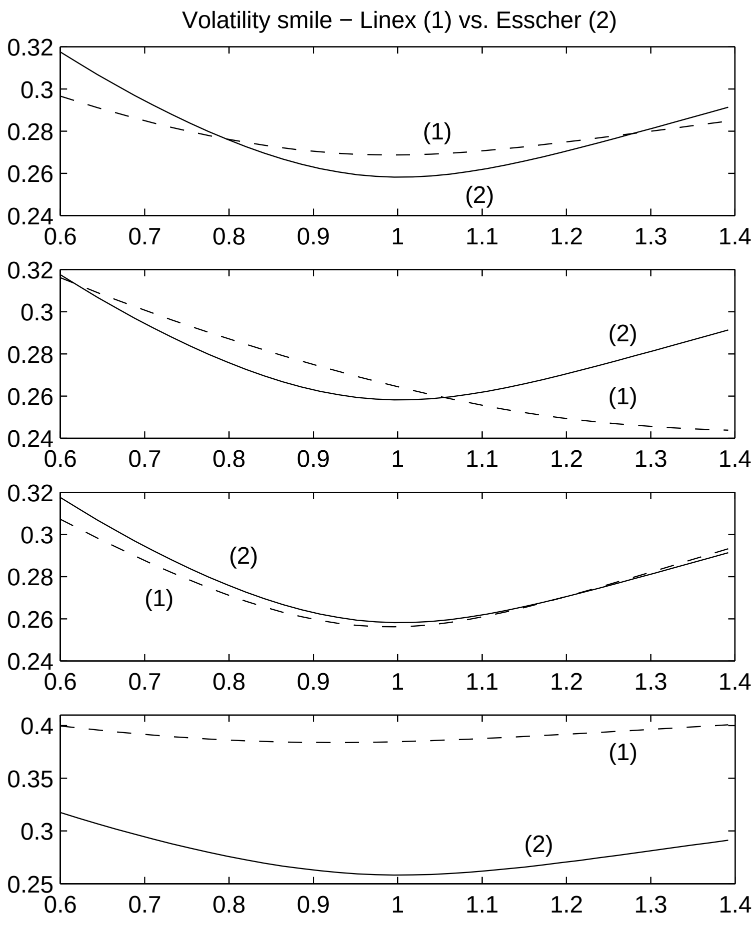


Fig. 3.4. Volatility smile of *linex* prices (dashed lines) each compared with the Esscher-implied volatility smile (solid line), which is identical for all cases. Note the different scaling of the ordinate of the lowermost picture.

Forcing up the risk-neutral kurtosis gives a picture which comes quite close to the Esscher structure. This is not very surprising because the risk-neutral parameters enforced by the Esscher change of measure are not far away from the ones of the linex measures, i.e. $vol^{\mathbf{Q}} = 0.28$ vs. $vol^{es} = 0.28$, $skew^{\mathbf{Q}} = -0.1$ vs. $skew^{es} = -0.3$, and $kurt^{\mathbf{Q}} = 6$ vs. $kurt^{es} = 5.91$. The skewness gap, which makes up the only significant difference, can be seen from the third picture in Fig. 3.4 by the more pronounced curvature of the Esscher volatility smile. All in all, this gives an indication that the risk-neutral distribution is quite satisfactorily captured by the second, the third, and the fourth moment.

A completely different picture is obtained in the latter case where statistical and risk-neutral volatility are very different. As the Esscher measure does not take into account risk-neutral volatility, it ignores this difference. On the contrary, Black-Scholes prices do if implied volatility is used. Hence in this case they should be preferred to Lévy-Esscher prices. Typically, in empirical studies (e.g. Corrado and Su (1997)) the estimated risk-neutral skewness is highly negative - often even more negative than in our high skewness case - , contradicting our Esscher-implied values, and the kurtosis is not as big as in our Esscher case.

Fig. 3.4 is a very interesting conclusion from the pricing behaviour of the considered models. It depicts the volatility smile which is implied by the different measures. In our parameter configuration, the Esscher measure implies a very pronounced curvature of the smile, and the smile is almost symmetric. Contrary to this, a high skewness translates into a very asymmetric shape of the volatility smile, which is then called volatility skew. But this is exactly what has been increasingly observed since the crash of 1987. Hence, linex pricing in the scenario of the third case is much closer to reality than pricing with the Esscher measure.

Again the case of a high risk-neutral volatility compared to the statistical one shows how wrong Esscher pricing becomes in this case.

More interesting in practical terms are the relative instead of the absolute differences. Again, for all four cases, the solid lines in Fig. 3.5 show the relative percentage differences between implied volatilities of Linex and Esscher pricing. And once again we see that for the range of strike prices considered we have differences up to 20% in implied volatility for the first three pictures and an even bigger discrepancy for the difference in the case of differing risk-neutral volatilities.

The dashed lines in Fig. 3.5 examine a totally different aspect of the linex change of measure. In Section 3.6.1 we have said what we understand by a unconstrained linex measure: It fulfills all the equality restrictions of a linex measure, but it only served as the initial value of the relative entropy minimization. The dashed lines show the percentage differences of implied volatilities between the linex measures and its corresponding starting values. Hence they are expected to reveal some information about to what extent the three assumed moments pin down the price of an option. If they covered all the information content in option prices the dashed curves should be straight lines identical to zero. A large deviation would mean that we should add higher moments in order to better capture the risk-neutral distribution. Fig. 3.5 reveals that the deviation is quite small with the exception of the high kurtosis case.

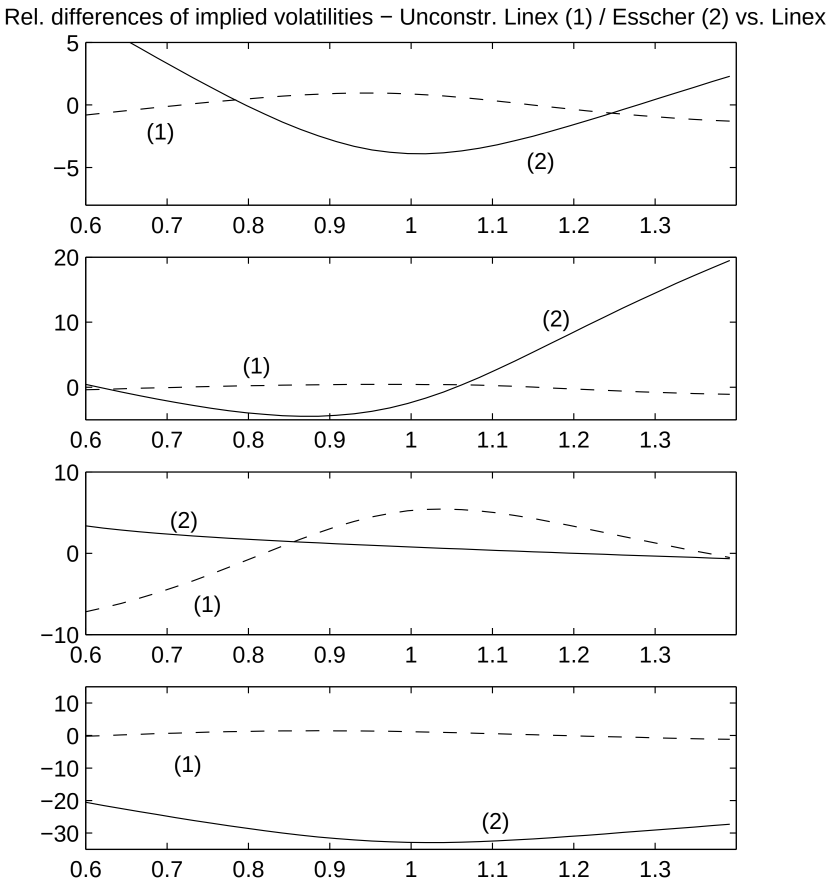


Fig. 3.5. Volatility smile differences in % for the cases of Fig. 3.4.

It is in particular small for the high skewness case, which, as we said, is the most likely scenario for present-day options markets.

3.7 Concluding remarks

This chapter has shown on the one hand that Lévy models are not always closer to reality than the classical Black-Scholes setting. The reason is that the non-uniqueness of the martingale measure gives rise to many methods of selection which may be consistent with economic theory and lead to a fast and efficient pricing algorithm but which are not in line with observed exchange-traded option prices, provided that a sufficiently liquid market for the latter exists. Pricing with the Esscher transform is one example of this procedure: Esscher pricing only depends on statistical parameters and implies a certain structure of the risk-neutral distribution which in general is inconsistent with the risk-neutral distribution observed in the market. In the above example, compared with the magnitude usually found by empirical studies, the implied risk-neutral skewness following the Esscher approach is too low in absolute value. Moreover, the implied kurtosis is far too high.

On the other hand it has been shown that this deficiency can be mended by modelling explicitly the measure change process so as to equate the theoretical risk-neutral moments to the empirical ones estimated from exchange-traded option price data. Hence we use statistical as well as risk-neutral moments to price options, contrary to Esscher pricing (only statistical parameters) and pricing by statistical martingale measures (only risk-neutral parameters). The approach of linex measures comes up to these requirements by offering a flexible form of the measure change function, partly linear, partly exponential, in such a way that finding a feasible solution is essentially a linear, thus simple problem. Based on the analysis of an example we see that linex measures have the potential to explain the volatility smile within a very simple Lévy model. Through this approach we succeed in connecting the curvature resp. the asymmetry of the volatility smile to the risk-neutral moments kurtosis resp. skewness.

The entire analysis is performed with tempered stable Lévy processes offering both a reasonable fit to stock return data and a simple structure of the Lévy measure, which is essential to get analytical results.

Chapter 4

Non-linear dependence and option pricing

4.1 Introduction

In this chapter an approach of valuing basket options is developed where the underlying assets are modelled by a two-dimensional jump-diffusion process with skewed and leptokurtotic returns and a more general dependence structure than the usual linear dependence in Gaussian models. The approach is based on the theory of Lévy copulas, which originates from Tankov (2003) and which is outlined in Section 1.7. Parts of this chapter are based on Wannenwetsch (2004).

The procedure is arranged in four steps. In Section 4.2 a comprehensive one-parameter family of positive Lévy copulas is introduced. The aim of this construction is to represent all possible degrees of tail dependence by a suitable choice of the parameter. At the same time the tractability argument enforces it to be as simple as possible in the sense that it should be possible to integrate the resulting Lévy measure analytically in order to obtain an analytical expression of its characteristic function via the Lévy-Khinchin theorem. This opens up the possibility of pricing by Fourier inversion and Fast Fourier transform. All the analysis is limited to the two-dimensional case. This restriction is unfortunate but the general case does not seem to be a trivial extension. However, the main focus of this section is on highlighting the effect of non-linear dependence on option pricing via a tractable model, and for this purpose two dimensions are sufficient. Moreover, many interesting cases of practical relevance can be traced back to the two-dimensional case.
With the help of the Lévy measure obtained this way Section 4.3 constructs a two-dimensional version of Kou's model, described in Section 1.5.3, which is given in terms of its characteristic triplet and an analytical expression of the characteristic function.
Section 4.4 gives a method of valuing basket options on a portfolio of two assets which are modelled according to the new model. A strong point of the model is that all

moments of the two-dimensional return process can very easily be exactly calculated without having to resort to numerical methods. By a suitable transformation of the payoff function of a basket option it is possible to reduce the multidimensional pricing problem to a one-dimensional one which is solved by Fourier inversion.
The last step in Section 4.5 evaluates the model as well as its pricing performance for a set of artificial data and tries to quantify the pricing error which traditional models make by just using linear correlation as a dependence measure.

4.2 A comprehensive family of bivariate Lévy copulas

This section introduces a family of positive Lévy copulas in two dimensions. The integral of the Lévy measure, which is constructed out of two marginal distributions and a member of this family, can be comfortably decomposed into a sum of simpler integrals with respect to the marginal Lévy measures. Finally, the questions of tail dependence and simulation of the jump size distribution are dealt with.

4.2.1 Definition

Proposition 4.1. *For $\theta \geq 1$ the following functions $C_\theta : [0,\infty]^2 \to [0,\infty]$ with*

$$C_\theta(u,v) := \frac{1}{\theta+1}(\min(\theta u, v) + \min(u, \theta v)) \tag{4.2.1}$$

form a comprehensive family of Lévy copulas. The complete dependence copula is attained by $\theta = 1$, and $\theta \to \infty$ yields the independence copula.

Proof. It can easily be seen that C_θ is grounded and has the right margins. To see that C_θ is 2-increasing consider a 2-box $[u_1, u_2] \times [v_1, v_2]$, $u_1 < u_2$, $v_1 < v_2$ and

$$\begin{aligned} & V_{C_\theta}([u_1,u_2]\times[v_1,v_2]) \\ = \; & \frac{1}{\theta+1}\left\{V_{C_{||}}([\theta u_1, \theta u_2]\times[v_1,v_2]) + V_{C_{||}}([u_1,u_2]\times[\theta v_1, \theta v_2])\right\} \geq 0, \end{aligned}$$

which follows from the fact that $C_{||}$ is a Lévy copula and thus 2-increasing. The other statements are obvious. □

We can construct new Lévy copulas by building convex combinations of some copulas C_θ. This can be done for discrete and continuous integrals which we denote differently because they have rather different properties.

Definition 4.2. *Let*

$$C_{w_d,\Theta}(u,v) := \sum_{\theta\in\Theta} w_d(\theta) C_\theta(u,v) \tag{4.2.2}$$

where Θ is a countable subset of $[1,\infty]$ and $w_d : \Theta \to [0,1]$ is a weighting scheme with $\sum_{\theta\in\Theta} w_d(\theta) = 1$. Moreover, let

$$C_{w_c}(u,v) := \int_1^\infty w_c(\theta) C_\theta(u,v) d\theta \tag{4.2.3}$$

where $w_c : [1, \infty) \to \mathbb{R}_+$ *is continuous and such that* $\int_1^\infty w_c(\theta) d\theta = 1$.

Definition 4.2 assumes of course that w_c and w_d are such that the corresponding integrals exist.

Proposition 4.3. *For* Θ, w_c *and* w_d *satisfying the assumptions of Definition 4.2,* $C_{w_d,\Theta}$ *is a singular Lévy copula with support*

$$\bigcup_{\theta \in \Theta} \{(u, v) \in [0, \infty]^2 | \theta u = v \text{ or } u = \theta v\}$$

whereas C_{w_c} *is absolutely continuous with full support.*

Proof. The first statement is obvious if one recalls that the complete dependence copula is singular with support $\{(u, v) \in [0, \infty]^2 | u = v\}$[1] and that a countable union of sets of measure zero has again measure zero. For the copula C_w we have

$$\begin{aligned} C_{w_c}(u, v) &= \int_1^\infty w_c(\theta) \left(\frac{1}{\theta + 1} \min(\theta u, v) + \frac{1}{\theta + 1} \min(u, \theta v) \right) d\theta \\ &= u \int_1^{\max\{1, v/u\}} \frac{\theta w_c(\theta)}{\theta + 1} d\theta + v \int_{\max\{1, v/u\}}^\infty \frac{w_c(\theta)}{\theta + 1} d\theta \\ &\quad + v \int_1^{\max\{1, u/v\}} \frac{\theta w_c(\theta)}{\theta + 1} d\theta + u \int_{\max\{1, u/v\}}^\infty \frac{w_c(\theta)}{\theta + 1} d\theta. \end{aligned}$$

For $u > v$ this equals

$$v \int_1^\infty \frac{w_c(\theta)}{\theta + 1} d\theta + v \int_1^{u/v} \frac{\theta w_c(\theta)}{\theta + 1} d\theta + u \int_{u/v}^\infty \frac{w_c(\theta)}{\theta + 1} d\theta$$

where it can immediately be seen that $C_{w_c}(u, v)$ is differentiable in the cone $A_1 := \{(u, v) \in (0, \infty) \times (0, \infty) | u > v\}$. By symmetry we have differentiability in $A_1 \cup A_2$ where $A_2 := \{(u, v) \in (0, \infty) \times (0, \infty) | u < v\}$. We find that

$$\frac{\partial^2}{\partial u \partial v} C_{w_c}(u, v) = \begin{cases} \frac{u}{v} \frac{w_c(u/v)}{u+v} & , \text{ for } u > v \\ \frac{v}{u} \frac{w_c(v/u)}{u+v} & , \text{ for } u < v \end{cases} \tag{4.2.4}$$

and for an arbitrary element $(\bar{u}, \bar{u})$ of the diagonal we have by the continuity of w

$$\lim_{n \to \infty} \left(\frac{u_{1,n}}{v_{1,n}} \frac{w_c(u_{1,n}/v_{1,n})}{u_{1,n} + v_{1,n}} \right) = \frac{w_c(1)}{2\bar{u}} = \lim_{n \to \infty} \left(\frac{u_{2,n}}{v_{2,n}} \frac{w_c(u_{2,n}/v_{2,n})}{u_{2,n} + v_{2,n}} \right)$$

for sequences $(u_{1,n}, v_{1,n})_{n \in \mathbb{N}}$ in A_1 and $(u_{2,n}, v_{2,n})_{n \in \mathbb{N}}$ in A_2 both converging to $(\bar{u}, \bar{u})$. □

[1] See e.g. Embrechts et al. (2001), p.8.

The Lévy copulas given above are defined in a way to construct tractable bivariate models for the joint movement of two dependent financial assets: Together with some specific Lévy measures on $\mathbb{R}$ they result in a two-dimensional Lévy measure via Theorem 1.31 that is easy enough to integrate analytically the exponential function. But with Theorem 1.14 in mind, this means that one can build bivariate Lévy processes with an analytical representation of their characteristic functions.
$C_{w_d,\Theta}$ and C_{w_c} provide a more realistic structure of the jump copula than C_θ does, in the sense that they their support is bigger than the one of C_θ, and there is no economic reason why one should a priori limit possible jump sizes.
A final remark concerns the copula C_{w_c}. As w_c is a continuous non-negative function, the finiteness of $\int_1^\infty w(\theta)d\theta$ implies that $w(\theta) \to 0$ for $\theta \to \infty$, which in turn means that the independence copula cannot be attained by any choice of w. A possible remedy would be to build a convex combination of C_w and the independence copula.
In the following, we will deal only with the basic Lévy copula C_θ. This is because C_θ itself seems to provide quite a realistic dependence structure of equity returns, and, secondly, because $C_{w_d,\Theta}$ is simply a linear (possibly infinite) combination of Lévy copulas of the type C_θ, hence extensions are possible and trivial.

4.2.2 Integration of the Lévy measure

Throughout this chapter we will maintain the following assumption:

Assumption 4.4. *For every one-dimensional tail integral U we have*

$$U(0+) := \lim_{t \to 0+} U(t) < \infty.$$

Two remarks must be made: First, Assumption 4.4 means that only integrable Lévy measures, which lead to finite activity Lévy processes, are considered. Second, Definition 1.27 requires that $U(0) = \infty$, hence we always have a discontinuity at zero of all tail integrals.
Given two tail integrals U_1 and U_2 of one-dimensional Lévy measures K_1 and K_2, which satisfy Assumption 4.4, we can define a two-dimensional tail integral with the help of the copulas introduced in the preceding section. As will be argued in Section 4.5 the very simple copula C_θ already suffices to produce interesting and realistic dependence structures. Hence we will restrict ourselves to this special copula in the following. The step from this tail integral, denoted by U_θ, to the corresponding Lévy measure, called K_θ, is done by Proposition 1.29. We have

$$U_\theta(x, y) = C_\theta(U_1(x), U_2(y)). \tag{4.2.5}$$

To ease notation we assume without loss of generality

Assumption 4.5.

$$\frac{U_2(0+)}{U_1(0+)} \geq 1.$$

As $\theta \geq 1$ this implies

$$\theta \geq \frac{U_1(0+)}{U_2(0+)}. \tag{4.2.6}$$

The support of C_θ, which was derived in Proposition 4.3, implies a very special structure of the support of K_θ. It consists of parts of the axes, representing jumps of only one component, and an off-axis part, where both components jump at the same time. As for this point, the following proposition distinguishes two cases depending on the size of θ:

- $\theta > \frac{U_2(0+)}{U_1(0+)}$. Define the two functions $\sigma_a, \sigma_b : \mathbb{R}_+ \to \mathbb{R}_+$ describing the off-axis support of K_θ by

$$\theta U_1(\sigma_a(y)) = U_2(y+) \quad \text{and} \quad U_1(x+) = \theta U_2(\sigma_b(x)).$$

- $\theta \leq \frac{U_2(0+)}{U_1(0+)}$. In this case let $\varsigma_a, \varsigma_b : \mathbb{R}_+ \to \mathbb{R}_+$ be the functions implicitly given by

$$\theta U_1(x+) = U_2(\varsigma_a(x)) \quad \text{and} \quad U_1(x+) = \theta U_2(\varsigma_b(x)).$$

The support of the Lévy measure K_θ is the union of the sets

$$\mathcal{X} := \{(x,y) \in \mathbb{R}_+^2 | y = 0\}, \qquad \mathcal{Y} := \{(x,y) \in \mathbb{R}_+^2 | x = 0\}$$
$$\mathcal{D}_a := \{(x,y) \in \mathbb{R}_+^2 | \theta U_1(x+) = U_2(y+)\}, \qquad \mathcal{D}_b := \{(x,y) \in \mathbb{R}_+^2 | U_1(x+) = \theta U_2(y+)\}.$$

For $\theta = 1$ we have $\mathcal{D}_a = \mathcal{D}_b$. If $\theta > 1$ we have for $(\bar{x}, y^a) \in \mathcal{D}_a$ and $(\bar{x}, y^b) \in \mathcal{D}_b$ the relation $U_2(y^a) = \theta U_1(\bar{x}) = \theta^2 U_2(y^b) > U_2(y^b)$, which means that $y^a < y^b$, i.e. $\mathcal{D}_a$ is completely below $\mathcal{D}_b$, and there is no point of intersection. Both cases are depicted in Fig. 4.1 for the special case where the graphs of $\mathcal{D}_a$ and $\mathcal{D}_b$ are straight lines. The integration with respect to K_θ proceeds as follows:

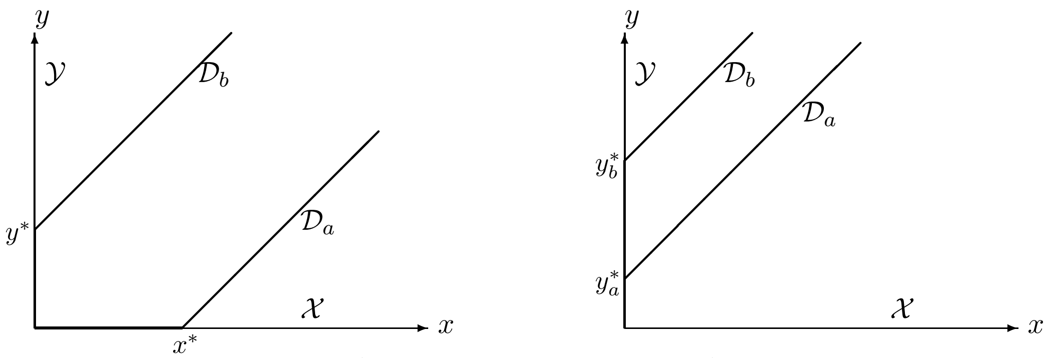


Fig. 4.1. The supports (=union of $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{D}_a$ and $\mathcal{D}_b$) of the two-dimensional Lévy measure K_θ for the two cases $\theta > U_2(0+)/U_1(0+)$ (left) and $\theta \leq U_2(0+)/U_1(0+)$ (right) which are considered in Proposition 4.6.

Proposition 4.6. *Let $f : \mathbb{R}_+^2 \to \mathbb{C}$ be a continuous function, and let Assumption 4.4 be satisfied. If $\theta > \frac{U_2(0+)}{U_1(0+)}$ then*

$$\int_0^\infty \int_0^\infty f(x,y)K_\theta(dx,dy) = \frac{\theta}{\theta+1}\int_0^{x^*} f(x,0)K_1(dx) + \frac{\theta}{\theta+1}\int_0^{y^*} f(0,y)K_2(dy)$$
$$\frac{1}{\theta+1}\int_0^\infty f(x,\sigma_b(x))K_1(dx) + \frac{1}{\theta+1}\int_0^\infty f(\sigma_a(y),y)K_2(dy), \tag{4.2.7}$$

where $x^* = \sigma_a(0)$ and $y^* = \sigma_b(0)$, and if $\theta \leq \frac{U_2(0+)}{U_1(0+)}$ then

$$\int_0^\infty \int_0^\infty f(x,y)K_\theta(dx,dy) = \int_0^{y_a^*} f(0,y)K_2(dy) + \frac{\theta}{\theta+1}\int_{y_a^*}^{y_b^*} f(0,y)K_2(dy)$$
$$+\frac{\theta}{\theta+1}\int_0^\infty f(x,\varsigma_a(x))K_1(dx) + \frac{1}{\theta+1}\int_0^\infty f(x,\varsigma_b(x))K_1(dx)$$

with $y_a^* = \varsigma_a(0)$ and $y_b^* = \varsigma_b(0)$.

Proof. Appendix C.1. □

4.2.3 Simulation of the jump size distribution

As the preceding calculation shows, the finiteness of the Lévy measures K_1 and K_2 is passed on to the finiteness of K_θ. Hence it defines the jump structure of a bivariate compound Poisson process. This section describes how to simulate dependent positive jumps which arise from K_θ. This is essential for the simulation of the jump-diffusion process to be defined in the next section under the name Kou's two-dimensional model.

We are given two one-dimensional Lévy measures with mass on $\mathbb{R}_+$ with tail integrals U_1 and U_2. Coupling them with the Lévy copula C_θ results in the two-dimensional tail integral U_θ given by (4.2.5). The resulting Lévy measure has mass

$$\lambda := K_\theta(\mathbb{R}_+^2) = U_1(0+) + U_2(0+) - C_\theta(U_1(0+), U_2(0+)).$$

This means that the jump size distribution Π^j is given by its tail integral

$$\Pi^j(X \geq x, Y \geq y) = \frac{1}{\lambda}U_\theta(x,y), \tag{4.2.8}$$

where X resp. Y are the random jump sizes of the first respectively the second component. We start by computing the conditional distributions $\Pi^j(X > x|Y = y')$, which splits up into a total of three cases depending on the value of $y' \geq 0$. We do all this only for the first case considered in Proposition 4.6.

Given that $0 \leq \underline{y} \leq y' \leq \bar{y}$, we obtain by definition of a conditional probability

$$\begin{aligned} \Pi^j(X \geq x | \underline{y} \leq Y \leq \bar{y}) &= \frac{C_\theta(U_1(x), U_2(\underline{y})) - C_\theta(U_1(x), U_2(\bar{y}))}{C_\theta(U_1(0), U_2(\underline{y})) - C_\theta(U_1(0), U_2(\bar{y}))} \\ &= \frac{C_\theta(U_1(x), U_2(\underline{y})) - C_\theta(U_1(x), U_2(\bar{y}))}{U_2(\underline{y}) - U_2(\bar{y})}. \end{aligned} \tag{4.2.9}$$

We observe that for $\underline{y} \to y'$, $\bar{y} \to y'$ and $y' > 0$ the tail integral of the conditional jump size distribution can only take three different values: $\Pi^j(X \geq x | \underline{y} \leq y' \leq \bar{y}) = 1$ or $\frac{1}{\theta+1}$ or 0.
Coming back to the announced three cases we assume first that $y' > y^*$ where y^* stems from the proof of Proposition 4.6. Let (x_1', y') and (x_2', y') be the points on $\mathcal{D}_a$ and D_b corresponding to y', i.e.

$$\theta U_1(x_1') = U_2(y') \quad \text{and} \quad U_1(x_1') = \theta U_2(y').$$

Then we obtain from (4.2.9) that $\Pi^j(X \geq x | Y = y')$ is the tail integral of a Bernoulli distributed random variable where

$$\Pi^j(X = x_1' | Y = y') = \frac{\theta}{\theta + 1} \quad \text{and} \quad \Pi^j(X = x_2' | Y = y') = \frac{1}{\theta + 1}.$$

In a corresponding manner, the second case $y' \leq y^*$ results in

$$\Pi^j(X = 0 | Y = y') = \frac{\theta}{\theta + 1} \quad \text{and} \quad \Pi^j(X = x_2' | Y = y') = \frac{1}{\theta + 1}.$$

Finally, tackling the third case, we assume $y' = 0$, i.e. Y does not jump. This case is different because conditional on the event $\{Y = 0\}$, X has a continuous distribution:

$$\begin{aligned} \Pi^j(X \geq x | 0 \leq Y \leq \epsilon) &= \frac{C_\theta(U_1(x), U_2(0)) - C_\theta(U_1(x), U_2(\epsilon))}{\lim_{\xi \to 0}(C_\theta(U_1(\xi), U_2(0)) - C_\theta(U_1(\xi), U_2(\epsilon)))} \\ &= \frac{U_1(x) - \frac{1}{\theta+1}(U_2(\epsilon) + U_1(x))}{\lim_{\xi \to 0}\{U_1(\xi) - \frac{1}{\theta+1}[U_2(\epsilon) + U_1(\xi)]\}} \end{aligned}$$

such that for $\epsilon \to 0$ we obtain

$$\Pi^j(X \geq x | Y = 0) = \frac{\theta U_1(x) - U_2(0+)}{\theta U_1(0+) - U_2(0+)}.$$

Given a random variable generator for the distribution of Y with the tail integral $U_2(y)/U_2(0+)$, we are thus able to simulate the two-dimensional random variable (X, Y). Kou's model can easily be simulated by the standard inversion method because the inverse of the jump size distribution function can be obtained explicitly and is equal to (1.5.10).

4.2.4 Tail dependence of a Lévy copula

Tail dependence is an asymptotic concept of dependence for two random variables X and Y which assesses how likely it is to have two big realizations of X and Y at the same time, where big is meant in terms of the absolute value.

Definition 4.7. *Let Π^j be a probability measure which at the same time is the two-dimensional jump measure of a jump-diffusion Lévy process and which is constructed according to (1.7.2) from a Lévy copula C and two tail integrals U_1 and U_2. Then Π^j has a coefficient of tail dependence λ_C*

$$\lambda_C := \lim_{u \to 0+} \frac{C(u,u)}{u},$$

provided that this limit exists.

It is interesting to observe that tail dependence is a copula property, independent of the marginal tail integrals resp. Lévy measures. Definition 4.7 is inspired by a similar definition for probabilistic copulas[2]. The difference with this definition is due to the fact that in our case we work with tail integrals instead of distribution functions. The intuitive justification of the definition is as follows: Let (X, Y) be random jump sizes with probability distribution Π^j, and consider for $u \le \min\{\lambda_1, \lambda_2\}$ (with λ_1 and λ_2 denoting the masses of the Lévy measures of X and Y) the conditional probability

$$\Pi^j(Y > U_2^{-1}(u) | X > U_1^{-1}(u)). \tag{4.2.10}$$

Plainly speaking, this is the conditional probability of a big value of Y given that X is big. The bigger u the more one approaches the extreme values of the jump size distribution. Using the representation (4.2.8) of Π^j and the definition of a conditional probability, Π^j is equal to

$$\frac{\Pi^j(X > U_1^{-1}(u), Y > U_2^{-1}(u))}{\Pi^j(X > U_1^{-1}(u))} = \frac{C(u,u)}{u}.$$

This interpretation makes clear that we have always $\lambda_C \in [0,1]$. Having once more a look at equation (4.2.10), one observes that u cannot be interpreted as a quantile, which is the usual thing to do in the analogous definition for probabilistic copulas. The reason is that U_1 and U_2 are not tail integrals of a probability distribution but of a Lévy measure. Despite of this lack of interpretation, the coefficient of tail dependence of C_θ is in line with our intuition. We obtain

$$\lambda_{C_\theta} = \frac{2}{\theta + 1},$$

which means that for the complete dependence case $\theta = 1$ we have maximal tail dependence $\lambda_{C_\theta} = 1$ whereas $\lambda_{C_\theta} = 0$ for two independent Lévy processes ($\theta \to \infty$). Hence C_θ can incorporate all degrees of tail dependence in the sense of our Definition 4.7.

[2]See Embrechts et al. (2001), p. 15.

4.3 Kou's model in two dimensions

This section develops a model which can be seen as an extension to two dimensions of Kou's model, which was sketched in Section 1.5.3. Kou's model in conjunction with the Lévy copula C_θ results in a two-dimensional jump-diffusion process which can incorporate some realistic properties while it is still tractable in the sense that it allows to obtain an analytical expression of the characteristic function. This especially entails the ability to exactly calculate all moments of this process.

4.3.1 Kou's Lévy measure in two dimensions

Lévy measure for positive Lévy copulas

As a first step towards the extension of Kou's model to two dimensions we deal with the case of coupling two subordinators of the Kou type via the copula C_θ. The tail integrals U_1 and U_2 are then given by

$$U_1(x) = \begin{cases} \infty, & x = 0 \\ p_1\lambda_1 e^{-\lambda_{1+}x}, & 0 < x < \infty \\ 0, & x = \infty \end{cases} \quad \text{and} \quad U_2(y) = \begin{cases} \infty, & y = 0 \\ p_2\lambda_2 e^{-\lambda_{2+}y}, & 0 < y < \infty \\ 0, & y = \infty \end{cases}$$

where $0 \leq p_1, p_2 \leq 1$, $\lambda_1, \lambda_2, \lambda_{1+}, \lambda_{2+} > 0$. Then the tail integral $U_\theta(x, y) = C_\theta(U_1(x), U_2(y))$ of the resulting Lévy measure K_θ is given by

$$U_\theta(x,y) = \begin{cases} \frac{1}{\theta+1}\left(\min\{\theta p_1\lambda_1 e^{-\lambda_{1+}x}, p_2\lambda_2 e^{-\lambda_{2+}y}\}\right. \\ \left. +\min\{p_1\lambda_1 e^{-\lambda_{1+}x}, \theta p_2\lambda_2 e^{-\lambda_{2+}y}\}\right), & x > 0, y > 0 \\ p_1\lambda_1 e^{-\lambda_{1+}x}, & x > 0, y = 0 \\ p_2\lambda_2 e^{-\lambda_{2+}y}, & x = 0, y > 0 \\ \infty, & x = 0, y = 0. \end{cases} \tag{4.3.1}$$

The aim of this subsection is to compute the characteristic function of the two-dimensional version of Kou's model, which essentially amounts to applying Proposition 4.6 to the function $f(x, y) \equiv e^{izx+iz'y}$, i.e. to compute

$$\int_{x=0}^{\infty}\int_{y=0}^{\infty} e^{izx+iz'y} U_\theta(dx, dy).$$

We still maintain Assumption 4.5, which in this case takes the form

$$\theta \geq \frac{p_1\lambda_1}{p_2\lambda_2}.$$

The sets $\mathcal{D}_a$ and $\mathcal{D}_b$ from the previous section are

$$\begin{aligned} \mathcal{D}_a &= \{(x,y) \in \mathbb{R}_+^2 | \theta p_1\lambda_1 e^{-\lambda_{1+}x} = p_2\lambda_2 e^{-\lambda_{2+}y}\} \quad \text{and} \\ \mathcal{D}_b &= \{(x,y) \in \mathbb{R}_+^2 | p_1\lambda_1 e^{-\lambda_{1+}x} = \theta p_2\lambda_2 e^{-\lambda_{2+}y}\}, \end{aligned}$$

from which we have that

$$\begin{aligned}
\sigma_a(y) &= \frac{\lambda_{2+}}{\lambda_{1+}} y + \frac{1}{\lambda_{1+}} \log\left(\theta \frac{p_1\lambda_1}{p_2\lambda_2}\right), \\
\sigma_b(x) = \varsigma_b(x) &= \frac{\lambda_{1+}}{\lambda_{2+}} x + \frac{1}{\lambda_{2+}} \log\left(\theta \frac{p_2\lambda_2}{p_1\lambda_1}\right), \\
\varsigma_a(x) &= \frac{\lambda_{1+}}{\lambda_{2+}} x + \frac{1}{\lambda_{2+}} \log\left(\frac{1}{\theta} \frac{p_2\lambda_2}{p_1\lambda_1}\right),
\end{aligned}$$

which are all positive at zero for the corresponding parameter configurations, thus determining the points $x^*, y^*, y_a^*, y_b^* \geq 0$.
In the light of Proposition 4.6 we obtain for the parameter configuration $\theta > \frac{p_2\lambda_2}{p_1\lambda_1}$ the integral over $\mathcal{D}_a$ as

$$\begin{aligned}
&\Phi_a^\sigma(\theta; z, z') := \int_0^\infty f(\sigma_a(y), y) K_2(dy) \\
&= \int_0^\infty \exp\left\{ iz \left[\frac{\lambda_{2+}}{\lambda_{1+}} y + \frac{1}{\lambda_{1+}} \log\left(\theta \frac{p_1\lambda_1}{p_2\lambda_2}\right)\right] + iz'y \right\} p_2\lambda_2\lambda_{2+} e^{-\lambda_{2+}y} dy \\
&= p_2\lambda_2\lambda_{2+} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\theta \frac{p_1\lambda_1}{p_2\lambda_2}\right)\right] \int_0^\infty \exp\left\{ -\left[\lambda_{2+} - iz\frac{\lambda_{2+}}{\lambda_{1+}} - iz'\right] y \right\} dy \\
&= \frac{p_2\lambda_2\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{2+} - iz'\lambda_{1+}} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\theta \frac{p_1\lambda_1}{p_2\lambda_2}\right)\right]. \qquad (4.3.2)
\end{aligned}$$

By analogy with (4.3.2), we obtain for the integration on $\mathcal{D}_b$ by interchanging indices as well as z and z'

$$\begin{aligned}
\Phi_b^\sigma(\theta; z, z') &:= \int_0^\infty f(x, \sigma_b(x)) K_1(dx) \\
&= \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\theta \frac{p_2\lambda_2}{p_1\lambda_1}\right)\right]. \qquad (4.3.3)
\end{aligned}$$

The other case $\theta \leq \frac{p_2\lambda_2}{p_1\lambda_1}$ is done similarly, and one has

$$\begin{aligned}
\Phi_a^\varsigma(\theta; z, z') &:= \int_0^\infty f(x, \varsigma_a(x)) K_1(dx) \\
&= \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\frac{1}{\theta} \frac{p_2\lambda_2}{p_1\lambda_1}\right)\right]
\end{aligned}$$

and

$$\Phi_b^\varsigma(\theta; z, z') := \int_0^\infty f(x, \varsigma_b(x)) K_1(dx) = \Phi_b^\sigma(\theta; z, z').$$

From the arguments of the above exponential functions to be integrated - e.g. in the line above (4.3.2) - one verifies that these computations can be done for $(z, z') \in$

$\mathcal{M}_{12}^{+} := \{(\zeta, \zeta') \in \mathbb{C}^2 | \lambda_{1+}\lambda_{2+} + \lambda_{2+}\Im(\zeta) + \lambda_{1+}\Im(\zeta') > 0\}$.
Hence, by Proposition 4.6 we have for $\theta > \frac{p_2\lambda_2}{p_1\lambda_1}$ and $f(x,y) \equiv e^{izx+iz'y}$

$$\int_0^\infty \int_0^\infty e^{izx+iz'y} K_\theta(dx,dy) = \frac{1}{\theta+1} \left\{ \Phi_b^\sigma(\theta; z, z') + \Phi_a^\sigma(\theta; z, z') \right. \\ \left. + \frac{\theta p_1 \lambda_1 \lambda_{1+}}{\lambda_{1+} - iz} [1 - e^{-(\lambda_{1+} - iz)x^*}] + \frac{\theta p_2 \lambda_2 \lambda_{2+}}{\lambda_{2+} - iz'} [1 - e^{-(\lambda_{2+} - iz')y^*}] \right\}, \tag{4.3.4}$$

and for $\theta \leq \frac{p_2\lambda_2}{p_1\lambda_1}$

$$\int_0^\infty \int_0^\infty e^{izx+iz'y} K_\theta(dx,dy) = \frac{1}{\theta+1} \left\{ \theta \Phi_a^\varsigma(\theta; z, z') + \Phi_b^\varsigma(\theta; z, z') \right. \\ \left. \frac{\theta p_2 \lambda_2 \lambda_{2+}}{\lambda_{2+} - iz'} [e^{-(\lambda_{2+} - iz')y_a^*} - e^{-(\lambda_{2+} - iz')y_b^*}] \right\} + \frac{p_2 \lambda_2 \lambda_{2+}}{\lambda_{2+} - iz'} [1 - e^{-(\lambda_{2+} - iz')y_a^*}]. \tag{4.3.5}$$

The domain of existence for the one-dimensional integrals are $\mathcal{M}_1^{+} := \{(\zeta, \zeta') \in \mathbb{C}^2 | \Im(\zeta) > -\lambda_{1+}\}$ and $\mathcal{M}_2^{+} := \{(\zeta, \zeta') \in \mathbb{C}^2 | \Im(\zeta') > -\lambda_{2+}\}$. Hence the above calculations can be carried out for[3] $(z, z') \in \mathcal{M}_1^{+} \cap \mathcal{M}_2^{+} \cap \mathcal{M}_{12}^{+}$.

The first two terms in (4.3.4) and (4.3.4) are the ones that determine the dependence structure of the Lévy process. For independent jumps, i.e. $\theta = \infty$, they must cleary vanish. Indeed: The mixed term of the first case is

$$\frac{\Phi_b^\sigma(\theta; z, z') + \Phi_a^\sigma(\theta; z, z')}{\theta+1} \tag{4.3.6}$$

$$= \frac{1}{\theta+1} \left\{ \frac{p_2\lambda_2\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{2+} - iz'\lambda_{1+}} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\theta \frac{p_1\lambda_1}{p_2\lambda_2} \right) \right] \right. \\ \left. + \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\theta \frac{p_2\lambda_2}{p_1\lambda_1} \right) \right] \right\}$$
$$= \frac{p_2\lambda_2\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{2+} - iz'\lambda_{1+}} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\theta \frac{p_1\lambda_1}{p_2\lambda_2} \right) - \log(\theta+1) \right] \\ + \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\theta \frac{p_2\lambda_2}{p_1\lambda_1} \right) - \log(\theta+1) \right]. \tag{4.3.7}$$

The real part of the argument of the first exponential reduces to

$$\log\left[\left(\frac{p_1\lambda_2}{p_2\lambda_2} \right)^{\frac{-\Im(z)}{\lambda_{1+}}} \frac{\theta^{\frac{-\Im(z)}{\lambda_{1+}}}}{\theta+1} \right]. \tag{4.3.8}$$

As $\frac{-\Im(z)}{\lambda_{1+}} < 1$ due to $z \in \mathcal{M}_1$ for $\theta \to \infty$ we have convergence of the argument of the logarithm to zero, i.e. the first term in (4.3.7) vanishes. The second exponential

[3] Regardless of the value of θ, $\mathcal{M}_{12}^{+}$ is the domain of existence of the mixed term. The reason is that for $\theta < \frac{p_2\lambda_2}{p_1\lambda_1}$ equation (4.3.3) can be derived as (4.3.2) with the same integrability condition but with a smaller domain of integration.

is treated similarly. As for the integrals over $\mathcal{X}$ and $\mathcal{Y}$, x^* and y^* go to infinity for $\theta \to \infty$. These considerations yield

$$\begin{aligned}\lim_{\theta\to\infty}\int_0^\infty\int_0^\infty e^{izx+iz'y}K_\theta(dx,dy) &= \frac{p_1\lambda_1\lambda_{1+}}{\lambda_{1+}-iz}+\frac{p_2\lambda_2\lambda_{2+}}{\lambda_{2+}-iz'}\\ &= \int_0^\infty\int_0^\infty e^{izx+iz'y}K_\perp(dx,dy),\end{aligned}$$

i.e. the sum of the two integrals of e^{ixz} and $e^{iyz'}$ with respect to the marginal Lévy measures. And this is equal to the integral of f over the two-dimensional Lévy measure $K_\perp$ with margins K_1 and K_2 linked by the independence copula $C_\perp$.
This result could have been established more elegantly but less intuitively by noting that $C_\theta(u,v) \to C_\perp(u,v)\ \forall(u,v) \in \bar{\mathbb{R}}^2$ for $\theta \to \infty$ and applying Proposition 1.32 on weak convergence of Lévy measures.

4.3.2 General Lévy copulas and Fourier transformation

Up to now we have established all our results for positive Lévy copulas, i.e. for copulas linking dependence of jumps which go in one direction. Any sensible model for financial markets must allow for positive and negative jumps and thus account for dependence modelling in every four quadrants of $\mathbb{R}^2$: Dependence of positive jumps of one component with positive jumps of the second, dependence of positive jumps of one component with negative jumps of the second, and so on. Thus we extend the copula C_θ to cover the case of coupling Lévy measures with support on the entire $\mathbb{R}$. The Lévy copula

$$C_{\theta_+,\theta_-}(u,v) := C_{\theta^+}(|u|,|v|)\mathbf{1}_{\{u\geq 0,v\geq 0\}} + C_{\theta^-}(|u|,|v|)\mathbf{1}_{\{u\leq 0,v\leq 0\}} \tag{4.3.9}$$

was already mentioned in Example 1.39. Accordingly, we refer to Section 1.7.2 for a mathematical treatment of general Lévy copulas. This is apparently a very simple structure because we put no mass on the simultaneous occurrence of a positive jump of one Lévy process and a negative one of the other. Consequently, we only model positive dependence, which seems reasonable for modelling the price process of many equities via a jump diffusion process. This could be, for instance, because the jumps are held responsible for macroeconomic risk factors which are likely to move the prices equities in the same direction, either both up or both down. In any case, the general case of modelling all four quadrants of the two-dimensional Lévy measure would be quite easy to obtain but would render the exposition of the model more cumbersome.
What is still possible with the restricted model is to model separately the dependence of positive and negative jumps via the parameters θ_+ and θ_-. This is reasonable if one thinks e.g. that the risks of downward jumps are more dependent than those of upward jumps.
The tail integrals of the two considered marginal Lévy measures are subdivided into

two tail integrals corresponding to the positive and the negative real axis, which results in a function decreasing on both sides of zero. They are given by

$$U_1^+(x) = \begin{cases} \infty, & x = 0 \\ p_1\lambda_1 e^{-\lambda_{1+}x}, & x > 0 \\ 0, & x = \infty \end{cases} \quad \text{and} \quad U_1^-(x) = \begin{cases} 0, & x = -\infty \\ -(1-p_1)\lambda_1 e^{-\lambda_{1-}|x|}, & x < 0 \\ -\infty, & x = 0 \end{cases} \tag{4.3.10}$$

respectively

$$U_2^+(y) = \begin{cases} \infty, & y = 0 \\ p_2\lambda_2 e^{-\lambda_{2+}y}, & y > 0 \\ 0, & y = \infty \end{cases} \quad \text{and} \quad U_2^-(y) = \begin{cases} 0, & y = -\infty \\ -(1-p_2)\lambda_2 e^{-\lambda_{2-}|y|}, & y < 0 \\ -\infty, & y = 0. \end{cases} \tag{4.3.11}$$

Analogously to the previous section the tail integral of the two-dimensional Lévy measure $K_{\theta+,\theta-}$ can be defined for $x, y \leq 0$, and we obtain for $x, y \neq 0$

$$U_{\theta+,\theta-}(x,y) = C_{\theta+,\theta-}(U_1(x), U_2(y)) \tag{4.3.12}$$

where[4]

$$U_1(x) = U_1^+(x)\mathbf{1}_{\{x>0\}} + U_1^-(x)\mathbf{1}_{\{x<0\}} \quad \text{and} \quad U_2(y) = U_2^+(y)\mathbf{1}_{\{y>0\}} + U_2^-(y)\mathbf{1}_{\{y<0\}}. \tag{4.3.13}$$

To compute the integral

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{izx+iz'y} K_{\theta_+,\theta_-}(dx,dy) \\ = &\int_0^{\infty}\int_0^{\infty} e^{izx+iz'y} K_{\theta_+}(dx,dy) + \int_{-\infty}^0\int_{-\infty}^0 e^{izx+iz'y} K_{\theta_-}(dx,dy) \end{aligned} \tag{4.3.14}$$

three cases must be distinguished for each of the integrals in (4.3.14). This is done in Appendix C.2.
We have already seen that the first integral in (4.3.14) exists in $\mathcal{M}_1^+ \cap \mathcal{M}_2^+ \cap \mathcal{M}_{12}^+$. It turns out that both integrals are finite for $(z, z') \in \mathcal{M} = \mathcal{M}_1 \cap \mathcal{M}_2 \cap \mathcal{M}_{12}$ where $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_{12}$ being given by

$$\begin{aligned} \mathcal{M}_1 &= \{(\zeta,\zeta') \in \mathbb{C}^2 | \lambda_{1-} > \Im(\zeta) > -\lambda_{1+}\}, \\ \mathcal{M}_2 &= \{(\zeta,\zeta') \in \mathbb{C}^2 | \lambda_{2-} > \Im(\zeta') > -\lambda_{2+}\}, \quad \text{and} \\ \mathcal{M}_{12} &= \{(\zeta,\zeta') \in \mathbb{C}^2 | \lambda_{1+}\lambda_{2+} + \lambda_{2+}\Im(\zeta) + \lambda_{1+}\Im(\zeta') > 0 \quad \wedge \\ &\qquad \lambda_{1-}\lambda_{2-} - \lambda_{2-}\Im(\zeta) - \lambda_{1-}\Im(\zeta') > 0\}. \end{aligned}$$

Later it will only be important that a neighbourhood of $(0,0)$ is in $\mathcal{M}$ which is, of course, true if all λ's in the above formulae are positive. Finally, we have the following result:

[4] Observe that this representation is valid because the four occurring integrals are defined such that they have the same sign as their arguments.

Proposition 4.8. *The function* $\psi_d : \mathbb{C}^2 \supset \mathcal{M} \to \mathbb{C}$ *with*

$$\psi_d(z, z') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (e^{izx+iz'y} - 1) K_{\theta^+,\theta^-}(dx, dy)$$

has the analytic representation

$$\psi_d(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{izx+iz'y} K_{\theta^+,\theta^-}(dx, dy) - \lambda_1 - \lambda_2 + A_{\theta_+,\theta_-} \tag{4.3.15}$$

where the integral in (4.3.15) is given in the appendix, and

$$A_{\theta_+,\theta_-} = C_{\theta^+}(p_1\lambda_1, p_2\lambda_2) + C_{\theta^-}((1-p_1)\lambda_1, (1-p_2)\lambda_2)$$

is a constant that tends towards zero for $(\theta^+, \theta^-) \to (\infty, \infty)$.

Proof. Due to the finiteness of K_{θ^+,θ^-}, ψ_d can be written as the difference of two integrals, the last one being the mass of K_{θ^+,θ^-}. But the mass is equal to the intensity of the two-dimensional Lévy measure, which is equal to $\lambda_1 + \lambda_2 - A_{\theta^+,\theta^-}$. □

The analytical form of the characteristic function obtained by the preceding proposition can now be used to calculate exactly all the mixed moments of the Lévy process described by this characteristic function. This task requires the calculation of all the pure und mixed derivatives of the cumulant function (4.3.15). These derivatives follow the pattern of the derivatives of the function ϕ that will be defined in the lemma below. For any differentiable two-place function ϕ we use the notation

$$\phi^{s,t}(z, z') := \frac{\partial^s \partial^t}{\partial z^s \partial z'^t} \phi(z, z'). \tag{4.3.16}$$

Lemma 4.9. *Let*

$$\phi(z, z') = \frac{\alpha}{\beta - i\gamma z - i\delta z'} e^{i\epsilon z}$$

and $\alpha, \beta, \gamma, \delta, \epsilon \in \mathbb{R}$. *Then its mixed partial derivatives for* $n, m \in \mathbb{N}_0$ *are given by*

$$\phi^{n,m}(z, z') = \alpha \delta^m e^{i\epsilon z} \sum_{j=0}^{n} \binom{n}{j} (j+m)! \frac{i^{n+m} \gamma^j \epsilon^{n-j}}{(\beta - i\gamma z - i\delta z')^{j+m+1}}. \tag{4.3.17}$$

Proof. For $m = 0$ (4.3.17) becomes

$$\phi^{n,0}(z, z') = \alpha e^{i\epsilon z} \sum_{j=0}^{n} \frac{n!}{(n-j)!} \frac{i^n \gamma^j \epsilon^{n-j}}{(\beta - i\gamma z - i\delta z')^{j+1}} \tag{4.3.18}$$

which can be proved by induction over n. For $n = 0$ the claim is true. Furthermore

$$\begin{aligned}
&\frac{\partial}{\partial z}\phi^{n,0}(z,z') \\
= \;& \alpha i\epsilon e^{i\epsilon z}\sum_{j=0}^{n}\frac{n!}{(n-j)!}\frac{i^n\gamma^j\epsilon^{n-j}}{(\beta-i\gamma z-i\delta z')^{j+1}} + \alpha e^{i\epsilon z}\sum_{j=0}^{n}\frac{n!}{(n-j)!}\frac{i^{n+1}\gamma^{j+1}\epsilon^{n-j}(j+1)}{(\beta-i\gamma z-i\delta z')^{j+2}} \\
= \;& \alpha e^{i\epsilon z}\left[\frac{i^{n+1}\epsilon^{n+1}}{\beta-i\gamma z-i\delta z'} + \sum_{j=1}^{n}\frac{n!}{(n-j)!}\frac{i^{n+1}\gamma^j\epsilon^{n-j+1}}{(\beta-i\gamma z-i\delta z')^{j+1}}\right. \\
&\left. + \frac{i^{n+1}\gamma^{n+1}(n+1)!}{(\beta-i\gamma z-i\delta z')^{n+2}} + \sum_{j=0}^{n-1}\frac{n!}{(n-j)!}\frac{i^{n+1}\gamma^{j+1}\epsilon^{n-j}(j+1)}{(\beta-i\gamma z-i\delta z')^{j+2}}\right] \\
= \;& \alpha e^{i\epsilon z}\left[\frac{i^{n+1}\epsilon^{n+1}}{\beta-i\gamma z-i\delta z'} + \sum_{j=1}^{n}\left(\frac{n!}{(n-j)!} + j\frac{n!}{(n-j+1)!}\right)\frac{i^{n+1}\gamma^j\epsilon^{n-j+1}}{(\beta-i\gamma z-i\delta z')^{j+1}}\right. \\
&\left. + \frac{i^{n+1}\gamma^{n+1}(n+1)}{(\beta-i\gamma z-i\delta z')^{n+2}}\right] \\
= \;& \phi^{n+1,0}(z,z'),
\end{aligned}$$

observing that $\frac{n!}{(n-j)!} + j\frac{n!}{(n-j+1)!} = \frac{(n+1)!}{(n-j+1)!}$. The claim follows easily by taking the m-th derivative of (4.3.18) with respect to z' and Fubini's theorem. □

4.3.3 Kou's model in two dimensions

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Suppose we are given

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2 \quad \text{and} \quad c = \begin{pmatrix} \sigma_1^2 & \varrho_{12}\sigma_1\sigma_2 \\ \varrho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

with $\sigma_1, \sigma_2 > 0$ and $-1 < \varrho_{12} < 1$. We want c to be non-singular, hence we have to exclude the case where we have perfectly positive and perfectly negative correlation. Furthermore we are given two one-dimensional Lévy measures K_1 and K_2 of Kou's double exponential form, a copula C_{θ_+,θ_-} with parameters $\theta_+, \theta_- \in [1,\infty]$ which result in the two-dimensional Lévy measure K_{θ_+,θ_-} in (4.3.12). Now define the two-dimensional Lévy process

$$(X_t, Y_t)_{0\le t\le T} := (b, c, K_{\theta_+,\theta_-})_{\mathbf{P}}.$$

This definition implies the existence of a two-dimensional Wiener process W with uncorrelated components W_1 and W_2 and of a random measure μ on $[0,T]\times\mathbb{R}^2$ which is constructed from K_{θ_+,θ_-} according to Section 1.1. As a truncation function $h(x) \equiv 0$ is chosen. Hence the two jump processes $\tilde{X}$ and $\tilde{Y}$ are given by

$$\tilde{X}_t := \int_0^t\int_{\mathbb{R}^2} x\mu(ds,dx,dy) \quad \text{and} \quad \tilde{Y}_t := \int_0^t\int_{\mathbb{R}^2} y\mu(ds,dx,dy). \tag{4.3.19}$$

Eventually, we obtain the following representation of the process (X,Y):

$$(X_t, Y_t)' = bt + \sqrt{c}W_t + (\tilde{X}_t, \tilde{Y}_t)'. \tag{4.3.20}$$

This is a two-dimensional jump-diffusion process with the diffusion part given by the volatilities σ_1 and σ_2 and the correlation coefficient ϱ_{12} of the diffusion part. We equip the given probability space with the filtration

$$\mathcal{F}_t = \sigma(W_s^1, W_s^2, \tilde{X}_s, \tilde{Y}_s, 0 \le s \le t),$$

which is the minimum filtration with respect to which all four processes are adapted. The financial market is then given by two risky assets

$$S_t^X = S_0^X \exp(X_t) \quad \text{and} \quad S_t^Y = S_0^Y \exp(Y_t) \tag{4.3.21}$$

and a riskless savings account

$$B_t = \exp(rt) \tag{4.3.22}$$

for a risk-free interest rate $r > 0$. This model is one possible two-dimensional version of Kou's model for one-dimensional asset price dynamics, and this is now the place where the KTD model is explicitly defined:

Definition 4.10. *The model (4.3.21) and (4.3.22) is called* Kou's two-dimensional (KTD) model*, and the stochastic process* (X,Y) *is called the KTD jump-diffusion Lévy process.*

Before we analyse the KTD process, a remark on the probability measures must be made. We will make use of the KTD model in conjunction with the technique of statistical martingale measures, i.e. we assume that X and Y are martingales under the given probability measure $\mathbf{P}$. As was explained in detail in the first chapters of this thesis, this approach has meanwhile become the usual thing to do in financial applications of Lévy processes, see for instance the books by Cont and Tankov (2004b) and Schoutens (2003). Thanks to the copula approach the tasks of adjusting the parameters such that both marginal processes are martingales are independent from one another.
Here it is appropriate to introduce a useful definition: In case we want to highlight the dependence of the characteristic function $\chi(z,z')$ or the cumulant function $\psi(z,z')$ on the first parameters b_1 and b_2 of the Lévy triplet we write $\chi(z,z';b_1,b_2)$ and $\psi(z,z';b_1,b_2)$.
Recalling that the cumulant functions of the marginals are easily obtained from $\psi(z,z')$ by setting z resp. z' to zero we obtain in the case of X from (4.3.21) and the martingale condition $E[e^{-rt}S_t^X] = S_0^X$ (remember that $X_0 = 0$ by definition of a Lévy process)

$$\begin{aligned} e^{rt} = E[e^{X_t}] = e^{t\psi_t(1/i,0;\,b_1,0)} &= \exp\left(t[b_1 + \frac{1}{2}\sigma_1^2 + \int_{-\infty}^{\infty}(e^x - 1)K_1(dx)]\right) \\ &= e^{tb_1}\, e^{t\psi(1/i,0;\,0,0)}. \end{aligned}$$

The parameter b_2 is obtained in the same way, and thus $b = (b_1, b_2)$ is already determined via the martingale requirement by

$$b_1 = r - \psi(1/i, 0; 0, 0) \quad \text{and} \quad b_2 = r - \psi(0, 1/i; 0, 0). \tag{4.3.23}$$

Together with (4.3.23) Corollary 1.18 implies that the means $(\mu_1, \mu_2) := (E[X_1], E[Y_1])$ of X_1 and Y_1 are equal to

$$(\mu_1, \mu_2) = \Bigg(r - \psi(1/i, 0; 0, 0) + \lambda_1 \left(\frac{p_1}{\lambda_{1+}} - \frac{1-p_1}{\lambda_{1-}} \right), \\ r - \psi(0, 1/i; 0, 0) + \lambda_2 \left(\frac{p_2}{\lambda_{2+}} - \frac{1-p_2}{\lambda_{2-}} \right) \Bigg), \tag{4.3.24}$$

and $E[X_t] = tE[X_1]$ and $E[Y_t] = tE[Y_1]$.
The following proposition follows immediately from the Lévy-Khinchin representation in Theorem 1.14 and Proposition 4.8 and summarizes what we have done so far in this section.

Proposition 4.11. *The characteristic function $\chi_t(z, z')$ of the two-dimensional Lévy process (X, Y) exists for $(z, z') \in \mathcal{M}$ and is given by*

$$\chi_t(z, z') = E[e^{izX_t + iz'Y_t}] = e^{t[\psi_c(z,z') + \psi_d(z,z')]},$$

where

$$\psi_c(z) = ib_1 z + ib_2 z' - \frac{1}{2}(\sigma_1^2 z^2 + 2\varrho_{12}\sigma_1\sigma_2 zz' + \sigma_2^2 z'^2),$$

and ψ_d is provided by (4.3.15). □

The KTD process is a two-dimensional jump-diffusion process. The key to the non-linear dependence is the use of Lévy copulas, which can of course be applied to all kinds of Lévy processes. Nevertheless, they are tailor-made for the specific features of Lévy processes with infinite Lévy measures. As for our jump-diffusion case we deal with a finite Lévy measure, hence we could in principle also use probabilistic copulas. The reason why we do not is chiefly based on the structure of the independence copula: In the case of two-dimensional copulas, the probabilistic independence copula is $\Pi^2(u, v) = uv$ whereas the Lévy copula representing independence is given by $C_\perp(u, v) = u\mathbf{1}_{\{v=\infty\}} + v\mathbf{1}_{\{u=\infty\}}$. $C_\perp$ cannot easily be adapted such that - restricted to a suitable domain - it becomes a probabilistic copula, hence it is a real problem to model independent Lévy processes by a probabilistic copula linking two one-dimensional Lévy measures. In particular, this is a problem if one wants to construct a comprehensive family of copulas, which is what we have done by using Lévy copulas.
An alternative procedure to Lévy copulas would be the following: Suppose the compound Poisson processes X and Y are used to model dependent equity price processes. Modelling dependence must deal with two channels through with X and Y are linked: the jump times and the jump sizes. Jump sizes can be modelled by a

probabilistic copula, but the independence copula will not lead to independent X and Y unless it is guaranteed that they do not have common jump times. The idea of a Lévy copula is however a natural approach for Lévy processes because it specifies the two channels at the same time, and modelling of complete dependence and independence is quite easy with this approach.

Later we will often refer to the *pure diffusion model*. We will understand by this notion an exponential Lévy model with a bond price process of the form (4.3.21) and (4.3.22) where (X, Y) is a two-dimensional Gaussian process consisting of a drift and a scaled Wiener process. In this model returns are only linked by linear dependence.

4.3.4 Decompositions of the KTD process

With the help of Proposition 4.11 we can deduce how the variances and the covariance of X_T and Y_T decompose into a diffusion part and a jump part. This will turn out to be useful in Section 4.5. We have for the joint characteristic function $\chi_T(z, z')$ of X_T and Y_T

$$\chi_T^{1,1}(0,0) = T\psi^{1,1}(0,0) + T^2\psi^{1,0}(0,0)\psi^{0,1}(0,0) = -E[X_T Y_T].$$

The means are $\mu_1 T = \frac{T}{i}\psi^{1,0}(0,0)$ and $\mu_2 T = \frac{T}{i}\psi^{0,1}(0,0)$, and hence the covariance of X_T and Y_T is given by

$$\operatorname{cov}(X_T, Y_T) = E[X_T Y_T] - \mu_1\mu_2 T^2 = -T\psi^{1,1}(0,0).$$

Writing down $\psi(0,0)$ more explicitly in terms of its decomposition into a diffusion and a jump contribution results in

$$\operatorname{cov}(X_T, Y_T) = T[c_{12} - \psi_d^{1,1}(0,0)]. \tag{4.3.25}$$

Analogously, the variances $\operatorname{var}(X_T)$ and $\operatorname{var}(Y_T)$ can be written as

$$\begin{aligned}\operatorname{var}(X_T) &= T[c_1 - \psi_d^{2,0}(0,0)]\\ \operatorname{var}(Y_T) &= T[c_2 - \psi_d^{0,2}(0,0)].\end{aligned}$$

The special feature of the KTD process is the kind of dependence of its two components: The two diffusion components are linearly dependent (if $\varrho_{12} \neq 0$), whereas the jump components possibly show some kind of tail dependence, which is based on the special construction of the Lévy copula (4.3.9). The law of the KTD process is uniquely characterized by its characteristic function at an arbitrary t, say $t = 1$. This is because two random variables with the same characteristic function have the same law and because a Lévy process is already characterized by the law of an arbitrary increment.
Now we have a closer look at the jump parts $(\tilde{X}, \tilde{Y})$ of (X, Y) which were defined

in (4.3.19). Each of the two marginals of these jump processes disintegrates further into two more jump processes:

$$\begin{aligned}\tilde{X}_t &= \tilde{X}_t^{\perp} + \tilde{X}_t^{||},\\ \tilde{Y}_t &= \tilde{Y}_t^{\perp} + \tilde{Y}_t^{||}.\end{aligned}$$

The processes $\tilde{X}_t^{\perp}$ and $\tilde{Y}_t^{\perp}$ are independent, whereas $\tilde{X}_t^{||}$ and $\tilde{Y}_t^{||}$ jump at the same time as prescribed by the copula C_{θ_+,θ_-}. In terms of the discussion of the Lévy measure K_θ in Section 4.2.2 and its extension to K_{θ_+,θ_-} in (4.3.9), this means that we divide its support into a part on the axes and another one outside the axes.
Both one-dimensional marginal Lévy measures K_1 and K_2 are of Kou's double exponential form with intensities

$$\lim_{x\to 0+} U_1^+(x) + \lim_{x\to 0-} |U_1^-(x)| = p_1\lambda_1 + (1-p_1)\lambda_1 = \lambda_1$$

resp. λ_2. By integrating the two-dimensional Lévy measure we obtain the intensity λ of the process $(\tilde{X}, \tilde{Y})$, where[5]

$$\begin{aligned}\lambda &= \lambda_1 + \lambda_2 - \lim_{x,y\to 0+} C_{\theta^+}(U_1^+(x), U_2^+(y)) - \lim_{x,y\to 0-} C_{\theta^-}(|U_1^-(x)|, |U_2^-(y)|)\\ &= \lambda_1 + \lambda_2 - C_{\theta^+}(p_1\lambda_1, p_2\lambda_2) - C_{\theta^-}((1-p_1)\lambda_1, (1-p_2)\lambda_2).\end{aligned}$$

This is exactly the sum of the intensity of $(\tilde{X}_t^{||}, \tilde{Y}_t^{||})$, i.e. of

$$C_{\theta^+}(p_1\lambda_1, p_2\lambda_2) + C_{\theta^-}((1-p_1)\lambda_1, (1-p_2)\lambda_2)$$

and of the intensities of the two independent jump processes $\tilde{X}_t^{\perp}$ and $\tilde{Y}_t^{\perp}$

$$\lambda_1 - C_{\theta^+}(p_1\lambda_1, p_2\lambda_2) - C_{\theta^-}((1-p_1)\lambda_1, (1-p_2)\lambda_2)$$

and

$$\lambda_2 - C_{\theta^+}(p_1\lambda_1, p_2\lambda_2) - C_{\theta^-}((1-p_1)\lambda_1, (1-p_2)\lambda_2).$$

These results can be used for simulating the KTD process. The Lévy-Khinchin representation shows that the cumulant function is the sum of the cumulant functions of a drift, a scaled Wiener process, a positive jump part and a negative jump part. But this means that they are independent and consequently can be simulated separately, and the overall jump-diffusion process arises from the superposition of these four parts. The simulation of a Wiener process on a fixed time grid is performed by simulating normally distributed increments. The jump size distribution of positive and negative jumps is simulated according to Section 4.2.3 where the numbers of jump times in the interval $[0, T]$ are Poisson distributed with intensities

$$T[p_1\lambda_1 + p_2\lambda_2 - C_{\theta^+}(p_1\lambda_1, p_2\lambda_2)]$$

for the positive jumps and

$$T[(1-p_1)\lambda_1 + (1-p_2)\lambda_2 - C_{\theta^-}((1-p_1)\lambda_1, (1-p_2)\lambda_2)]$$

for the negative jumps.

[5] This results from adding the masses of the margins and noting that by doing this one has integrated the off-axes area of $\mathbb{R}^2$ twice; subsequently it has to be subtracted once.

4.4 Approximate basket option pricing

The purpose of this section is to use the KTD model to evaluate options on two assets. One of the simplest classes of such options one can imagine are European basket (or index) options: Given weights ω_1 and ω_2, a strike price $K > 0$, and a maturity date $T > 0$, the payoff C_T of a basket call option is

$$C_T = (\omega_1 S_T^X + \omega_2 S_T^Y - K)^+,$$

i.e. it is simply an option on a portfolio of assets instead of single assets. This can more conveniently be written as

$$(S_0(\bar{\omega}_1 e^{X_T} + \bar{\omega}_2 e^{Y_T}) - K)^+ \tag{4.4.1}$$

where $S_0 := \omega_1 S_0^X + \omega_2 S_0^Y$, and $\bar{\omega}_1 := \omega_1 S_0^X / S_0$, $\bar{\omega}_2 := \omega_2 S_0^Y / S_0$. The new weights $\bar{\omega}_1$ and $\bar{\omega}_2$ are non-negative and satisfy $\bar{\omega}_1 + \bar{\omega}_2 = 1$.
We will use two methods for pricing options: The first is based on the technique of Fourier inversion and Fast Fourier transform, which is outlined in Section 1.6, and will be the main focus of this section. The goodness of fit will be evaluated by a Monte-Carlo simulation: Once given the results on the simulation of the KTD jump size distribution in Section 4.2.3, the implementation of this method is straightforward. However, as a plain Monte-Carlo method turns out to be not very accurate, we apply a variance reduction scheme[6] which uses the technique of control variates. The estimator $\hat{C}_0$ of the option price C_0 at $t = 0$ is equal to

$$\hat{C}_0 = e^{-rT} \frac{1}{N} \sum_{j=1}^{N} \{C_T(\omega_j) - \xi(\Gamma(\omega_j) - E[\Gamma])\} \tag{4.4.2}$$

for $\xi \in \mathbb{R}$, a number N of replications of a simulation and a control variate Γ which should be as highly correlated as possible with C_T. The variable ω_j indicates the state of the world for the j-th simulation trial in the sample. The number ξ is chosen so as to minimize the standard deviation of (4.4.2). The resulting ξ_{opt} is the quotient of the covariance of C_T and Γ and the variance of Γ which must be estimated by

$$\xi_{opt} = \frac{\sum_{j=1}^{N} (C_T(\omega_j) - \bar{C}_T)(\Gamma(\omega_j) - \bar{\Gamma})}{\sum_{j=1}^{N} (\Gamma(\omega_j) - \bar{\Gamma})^2}$$

with $\bar{C}_T$ and $\bar{\Gamma}$ denoting the sample means of C_T and Γ.
All in all, it remains to choose a control variate with, first, high correlation with C_T and, second, a conveniently computable expectation value, which enters in (4.4.2). A common choice[7] is the weighted sum of two single-assets options on each of the

[6] See Glasserman (2004) for a general reference and Cont and Tankov (2004b) in the context of basket options.
[7] See among others Laurence and Wang (2001) and Cont and Tankov (2004b), p. 375.

two underlyings. By the convexity of the maximum operator and setting $\tilde{K} := K/S_0$ we have for a path ω_j

$$\begin{aligned} C_0(\omega_j) &= e^{-rT}\left(S_0(\bar{\omega}_1 e^{X_T(\omega_j)} + \bar{\omega}_2 e^{Y_T(\omega_j)}) - K\right)^+ \\ &= e^{-rT} S_0\left(\bar{\omega}_1(e^{X_T(\omega_j)} - \tilde{K}) + \bar{\omega}_2(e^{Y_T(\omega_j)} - \tilde{K})\right)^+ \\ &\leq e^{-rT} S_0\left\{\bar{\omega}_1(e^{X_T(\omega_j)} - \tilde{K})^+ + \bar{\omega}_2(e^{Y_T(\omega_j)} - \tilde{K})^+\right\} \end{aligned}$$

The right hand side is therefore an upper bound of the option price for every path ω_j, but nevertheless the random variables corresponding to the left and right hand side are positively correlated.

In the rest of this section we will address the problem of finding an approximate pricing method in the KTD model that is faster than a Monte-Carlo simulation.

4.4.1 Pricing methodology

For the one-dimensional exponential Lévy model there is a very convenient method to obtain very fast and accurately the prices of a variety of European options: One represents the price of the option as a Fourier integral of the product of the characteristic function of the Lévy process and the Fourier transform of the underlying asset. The method based on Raible (2000), which was presented in Section 1.6, works instead with the Laplace transform, but essentially it amounts to the same thing as the method of Lewis (2001) based on the Fourier transform. The obtained integral can then be approximated via the Fast Fourier transform.
There is a variety of contracts which can be priced with this technique[8]. While many common payoff functions do not have a classical Fourier transform (in the sense that a Fourier transform is understood as a function with a real argument), the trick of Lewis (2001) is to extend the definition of the Fourier transform to the complex plane, thus possibly gaining integrability for certain arguments outside the real axis. The most obvious example is the payoff function $w(x) = (e^x - K)^+$ of a European call option with strike $K > 0$ whose Fourier transform $z \to \hat{w}(z)$ exists only for $\Im(z) > 1$.
We now turn to the KTD model given by (4.3.21) and (4.3.22). A priori, there is no reason why the above described method should not be extended to cover the case of payoff functions that depend on more than one assets. The standard but non-trivial contract in this case is a European call option on a basket of two assets with payoff function $w(x, y) = (w_1 e^x + w_2 e^y - K)^+$ for some positive variables ω_1, ω_2, and K. However, its Fourier transform

$$\hat{w}(u, v) = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} e^{iux+iuy}(w_1 e^x + w_2 e^y - K)^+ dy dx,$$

[8] See again Lewis (2001) and Raible (2000).

does not exist for any $(u,v) \in \mathbb{C}^2$ in the sense of absolute convergence: It suffices to see that

$$\begin{aligned} &\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} |e^{iux+iuy}(w_1e^x + w_2e^y - K)^+|dydx \\ = &\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} e^{-\Im(u)x-\Im(v)y}(w_1e^x + w_2e^y - K)^+dydx \\ \geq &\int_{x=-\infty}^{\infty}\int_{y=\log(K/w_2)}^{\infty} e^{-\Im(u)x-\Im(v)y}(w_1e^x + w_2e^y - K)dydx \qquad (4.4.3) \end{aligned}$$

where the latter integral clearly is not finite. The reason why one has integrability in the one-dimensional case is that the corresponding integrand in (4.4.3) is forced down to zero for $x \to \infty$ by its first factor through an appropriate choice of $\Im(u)$ and equals zero for negative x which are sufficiently large in absolute value because $K > 0$. This argument ceases to hold in the multidimensional case. Although the same kind of analysis for the put option reveals that it is indeed integrable, the resulting expression cannot easily be integrated to obtain an analytic form.
All this means that even the simplest payoff functions cannot be explicitly priced by Fourier inversion in several dimensions. The main idea for a possible solution to this problem is quite common in the basket option literature: One tries to reduce the multidimensional problem to a one-dimensional one. The following notation prepares this technique:

We recall the representation (4.4.1) of the payoff function, which is written in view of the later application of the Fast Fourier transform. This kind of approximation is best for K near S_0[9], and this is exactly the domain we are interested in for economic reasons. Consider the random variables Z_T and Z'_T:

$$Z'_T := e^{Z_T} := \bar{\omega}_1 e^{X_T} + \bar{\omega}_2 e^{Y_T} \qquad (4.4.4)$$

i.e. $Z_T := \log[\bar{\omega}_1 e^{X_T} + \bar{\omega}_2 e^{Y_T}]$. If we had the characteristic function of Z_T, we could proceed directly with the methods of Lewis (2001) or Raible (2000), i.e. one could price basket options with the tools of one-dimensional Fourier inversion and Fast Fourier transform. Unfortunately, getting a computationally tractable form of this function seems to be impossible, hence we have to be satisfied with a numerical approximation.
The literature on basket option pricing in markets which are only driven by a Wiener process in principle has to face the same problem: A sum of lognormally distributed variables is not any more of the lognormal type. Thus in this case one usually resorts to Gram-Charlier or Edgeworth series expansions[10] of density functions where usually the approximation is done using the first four moments of the true distribution. These approximations become the worse the more skewed and heavy-tailed the

[9]See Raible (2000).
[10]See e.g. Stuart and Ord (1987).

underlying distribution is[11]. But as just these phenomena are to be expected in a Lévy environment where the single assets have skewed and heavy-tailed distribution of returns, this approach is not convenient. This is still true if one includes more than four moments, which leads to unstable results due to the well-known problem of the occurrence of high-order polynomials. Alternatively, one could resort to a saddlepoint expansion which in general gives quite good approximations both in the centre as well as in the tails[12]. But saddlepoint approximations require additional information of the model which the KTD model cannot provide in a tractable way.
A second attempt could be to determine the characteristic function of Z'_T, thus avoiding the inconvenient logarithm in the formulation of Z_T. The calculation of moments of Z'_T then involves expressions of the form $E[e^{jX_T+kY_T}]$ for j,k becoming as big as the order of the moment which is to be calculated. Clearly, this would mean to evaluate the characteristic function χ_T of (X_T, Y_T) at $(u,v) = (-ji, -ki)$ where $i = \sqrt{-1}$. Taking a look at its domain of existence $\mathcal{M}$ given in Section 4.3.2, it turns out that these moments are less likely to exist the more heavy-tailed the distributions are. But the feature of heavy tails is directly driven by the size of the parameters λ_{1+}, λ_{1-}, λ_{2+} and λ_{2-}. Taking a leap forward to the example in Section 4.5, Table 4.1 shows that the fourth moment of Z'_T does not exist because $-4 < -\lambda_{2+} = -3.7$. This is the reason why we attach importance to the necessity of working with Z_T instead of Z'_T. In fact, the moments of Z_T exist:

Lemma 4.12. *For (X_T, Y_T) as defined above we have*

$$E\left[|\log(\bar{\omega_1}e^{X_T} + \bar{\omega_2}e^{Y_T}|^p\right] < \infty, \quad p \geq 1.$$

Proof. First of all note that for $a, b \in \mathbb{R}$ and $p \geq 1$

$$(a+b)^p \leq [|a|+|b|]^p \leq [2\max(|a|,|b|)]^p = 2^p \max(|a|^p, |b|^p) \leq 2^p(|a|^p + |b|^p). \quad (4.4.5)$$

Then it follows that

$$\begin{aligned}
& E\left[\left|\log\left\{\bar{\omega_1}e^{X_T} + \bar{\omega_2}e^{Y_T}\right\}\right|^p\right] \\
\leq \quad & E\left[\left|\log\left\{\max(\bar{\omega_1},\bar{\omega_2})\left(e^{|X_T|} + e^{|Y_T|}\right)\right\}\right|^p\right] \\
\leq \quad & E\left[\left|\log\left\{2\max(\bar{\omega_1},\bar{\omega_2})e^{|X_T|+|Y_T|}\right\}\right|^p\right] \\
\leq \quad & 2^p\left\{|\log\{2\max(\bar{\omega_1},\bar{\omega_2})\}|^p + E\left[(|X_T| + |Y_T|)^p\right]\right\} \\
\leq \quad & 2^p\left|\log\{2\max(\bar{\omega_1},\bar{\omega_2})\}\right|^p + 4^p\left\{E[(|X_T|^p] + E[|Y_T|^p]\right\},
\end{aligned}$$

where in the last two steps (4.4.5) was applied. Looking at the domain of existence $\mathcal{M}$ of $\chi(z,z')$ it is clear that the marginal moment generating functions of X_T and Y_T exist in a neighbourhood of zero. Then by Lemma 1.16 all moments of X_T and Y_T exist, and the above expression is finite for any $p \geq 1$. □

[11] See Stuart and Ord (1987), p.228/9. Jensen (1995), p.2, provides an example where the Edgeworth expansion completely fails in the tails of the distribution to be approximated.

[12] Jensen (1995).

The three step pricing procedure for basket option pricing in the KTD model will be the following:

a) The cumulants of the two-dimensional distribution (X_T, Y_T) will be calculated by means of the analytic form of its cumulant function and its derivatives. These cumulants will be transformed to obtain the moments. Furthermore the derivatives of the functions $(x, y) \rightarrow (\log(\bar{\omega}_1 e^x + \bar{\omega}_2 e^y))^p$ are computed for $p = 1, \ldots, 4$.

b) Via two-dimensional Taylor expansions for $p = 1, \ldots, 4$ we obtain the first four moments of Z_T in a recursive form.

c) The first four moments of Z_T fix a unique distribution of the normal inverse Gaussian (NIG) form, which appears to be tailor-made for this purpose. Knowing that the characteristic function of the NIG distribution has an easy analytic representation, one can then use the pricing procedure of one-dimensional Fourier inversion.

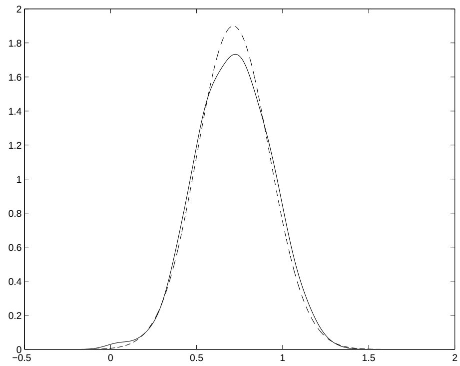


Fig. 4.2. Approximate density of the distribution of $\log(e^X + e^Y)$ (solid line) and a fitted normal distribution (dashed line).

The third step requires some explanation with the help of Figure 4.2. The graph shows the approximation of the density function of the distribution of $\log(e^X + e^Y)$ with normally distributed and positively correlated normal random variables[13] X and Y which is obtained by a Monte-Carlo simulation and a kernel-smoothed density estimator. It can be observed that the shape of this density function is not very different from the one of equity returns. But for the latter there are lots of parametric models which allow to describe very precisely the distribution of returns and to characterize the main sources of risk. Of course, this argument would not go through for a density function with, for instance, two peaks and hence a risk structure totally different from the one of returns, which therefore could not any more be sufficiently described by the first four moments. But in our case we can go by what Jarrow and Rudd (1982), p. 349, wrote it in the context of one-dimensional option pricing where the density function of the distribution of the underlying asset is not known explicitly, 'Intuition suggests that for practical purposes the first four moments of the underlying distribution should capture the majority of its influence as it affects option pricing.'

[13] X and Y are jointly normally distributed with means zero, standard deviations 0.2 resp. 0.3 and correlation 0.5.

The same way as in Chapter 4.2 throughout this chapter we will make use of the notation (4.3.16) for abbreviating partial derivatives of a function. The following three subsections each deal with one step of the pricing procedure sketched above.

4.4.2 Derivatives

A simple form of the characteristic function χ of (X_T, Y_T) is given in Proposition 4.11, and to obtain the moments of Z_T one has to calculate its derivatives. The procedure starts by computing the derivatives of the corresponding cumulant functions, which are easily computed using Lemma 4.9 for the jump part. For the diffusion we have

$$\begin{aligned} \psi_c^{1,0}(z,z') = ib_1 - c_{11}z - c_{12}z', &\qquad \psi_c^{0,1}(z,z') = ib_2 - c_{22}z' - c_{12}z \\ \psi_c^{2,0}(z,z') \equiv -c_{11}, &\qquad \psi_c^{0,2}(z,z') \equiv -c_{22} \\ \psi_c^{1,1}(z,z') \equiv -c_{12}, \quad \text{and} &\qquad \psi_c^{s,t}(z,z') \equiv 0 \ \text{ for } \ s+t \geq 3 \end{aligned}$$

The second task in the first step of the pricing procedure is to transform the obtained cumulants into central moments. The means μ_1 and μ_2 of X_1 and Y_1 were calculated in (4.3.24). By using the Lévy-Khinchin representation for the characteristic function $\chi_t(z,z')$ of (X_t, Y_t) we have

$$\begin{aligned} \bar{\chi}_t(z,z';b_1,b_2) &= E[e^{iz(X_t-\mu_1 t)+iz'(Y_t-\mu_2 t)}] \\ &= e^{-iz\mu_1 t - iz'\mu_2 t}E[e^{izX_t+iz'Y_t}] = \chi_t(z,z';b_1-\mu_1,b_2-\mu_2). \end{aligned}$$

where $\bar{\chi}_t(z,z') = \bar{\chi}_t(z,z';b_1,b_2)$ is the characteristic function of the vector $(X_t - \mu_1 t, Y_t - \mu_2 t)$. The corresponding cumulant function is denoted by $\bar{\psi}_t(z,z')$.

Given the cumulants, i.e. the appropriately scaled derivatives of the cumulant function

$$\kappa^{n,m} = \frac{1}{i^{n+m}} \frac{\partial^n \partial^m}{\partial z^n \partial z'^m} \bar{\psi}_T(z,z') \bigg|_{z=z'=0},$$

and recalling that $\chi(z,z') = e^{T\psi(z,z')}$ we can calculate the central moments

$$M^{n,m} := E[(X_T - \mu_1 T)^n (Y_T - \mu_2 T)^m] \tag{4.4.6}$$

of (X_T, Y_T) for $n \geq 1$ and $m \geq 0$ from the recursions by

$$\begin{aligned} M^{n,m} &= \frac{1}{i^{n+m}} \bar{\chi}_T^{n,m}(0,0) = \frac{1}{i^{n+m}} \frac{\partial^n \partial^m}{\partial z^n \partial z'^m} \bar{\chi}_T(z,z') \bigg|_{z=z'=0} \\ &= \frac{1}{i^{n+m}} \frac{\partial^{n-1}\partial^m}{\partial z^{n-1}\partial z'^m} T\bar{\psi}_T^{1,0}(z,z')\bar{\chi}_T(z,z') \bigg|_{z=z'=0} \\ &= \frac{T}{i^{n+m}} \sum_{j=0}^{n-1}\sum_{k=0}^{m} \binom{n-1}{j}\binom{m}{k} \bar{\psi}_T^{j+1,k}(z,z')\bar{\chi}_T^{n-1-j,m-k}(z,z') \bigg|_{z=z'=0} \\ &= T\sum_{j=0}^{n-1}\sum_{k=0}^{m} \binom{n-1}{j}\binom{m}{k} \kappa^{j+1,k} M^{n-1-j,m-k}, \end{aligned}$$

given all the moments $M^{0,m}$ for $m \geq 1$. The latter can recursively be obtained from $M^{0,0} = 1$ by

$$M^{0,m} = T \sum_{k=0}^{m-1} \binom{m-1}{k} \kappa^{0,k+1} M^{0,m-1-k}.$$

Before proceeding to the Taylor approximations, we are still in need of the partial derivatives of $\log(\sum_{i=1}^{2} \bar{\omega}_i e^{x_i})$. The rest of this section is devoted to the their time-saving calculation, which goes along the lines of the method used to obtain the moments $M^{n,m}$, i.e. we will obtain them in a recursive form. We provide a lemma which uses the following definitions: Let

$$F(x_1, x_2) := \log(\sum_{i=1}^{2} \bar{\omega}_i e^{x_i})$$

and

$$f_1(x_1, x_2) := \frac{\partial}{\partial x_1} F(x_1, x_2) = \frac{\bar{\omega}_1 e^{x_1}}{\sum_{i=1}^{2} \bar{\omega}_i e^{x_i}}$$

as well as

$$f_2(x_1, x_2) := \frac{\partial}{\partial x_2} F(x_1, x_2) = \frac{\bar{\omega}_2 e^{x_2}}{\sum_{i=1}^{2} \bar{\omega}_i e^{x_i}}.$$

As we have to consider the derivatives of the powers of F we also introduce

$$F^{s,t,m}(x_1, x_2) := \frac{\partial^s}{\partial x_1^s} \frac{\partial^t}{\partial x_2^t} [F(x_1, x_2)]^m, \; m \in \mathbb{N}_0, \tag{4.4.7}$$

for all $s \leq p$ and $t \leq q$, where clearly $F^{s,t,1} \equiv F^{s,t}$, and we distinguish with an abuse of notation the two classes of functions $F^{s,t,m}$ and $F^{s,t}$ by their numbers of superscripts.

Lemma 4.13. *Let $0 \leq s \leq p$ and $0 \leq t \leq q$.*
Step 1: *The 'pure' derivatives of F are given by*

$$F^{0,0} = \log(\sum_{i=1}^{n} \bar{\omega}_i e^{x_i}) \quad \textit{and} \quad F^{1,0} = f_1 \quad \textit{and} \quad F^{0,1} = f_2$$

and the recursive equations[14]

$$F^{s,0} = f_1^{s-1,0} = f_1^{s-2,0}(1 - f_1) - \sum_{j=0}^{s-3} \binom{s-2}{j} f_1^{j,0} f_1^{s-2-j,0} \tag{4.4.8}$$

and

$$F^{0,t} = f_2^{0,t-1} = f_2^{0,t-2}(1 - f_2) - \sum_{j=0}^{t-3} \binom{t-2}{j} f_2^{0,j} f_2^{0,t-2-j} \tag{4.4.9}$$

[14]The usual sum convention $\sum_n^m := 0$ for $n > m$ applies here.

for $s, t \geq 2$.
Step 2: *Given the derivatives of Step 1 the cross derivatives are, again by recursion,*

$$F^{s,1} = f_2^{s,0} = -\sum_{j=0}^{s-1} \binom{s-1}{j} f_2^{j,0} f_1^{s-1-j,0}$$

for $s \geq 1$ *and* $t = 1$. *For the general case* $s \geq 1$ *and* $t \geq 2$ *we have*

$$F^{s,t} = f_2^{s,t-1} = -\sum_{j=0}^{s-1}\sum_{k=0}^{t-2} \binom{s-1}{j}\binom{t-1}{k} f_2^{j,k} f_2^{s-j,t-2-k} - \sum_{j=0}^{s-1}\binom{s-1}{j} f_2^{j,t-1} f_1^{s-1-j,0}. \tag{4.4.10}$$

Step 3: *Given all the partial derivatives of* F *from Steps 1 and 2 we have*

$$F^{0,0,m} = [F(x_1, x_2)]^m.$$

If $s \geq 1$ *and* $t = 0$, *this yields by recursion*

$$F^{s,0,m} = m\sum_{j=0}^{s-1}\binom{s-1}{j} F^{j,0,m-1} F^{s-j,0}. \tag{4.4.11}$$

For $s \geq 0$ *and* $t \geq 1$ *we have the recursive relation*

$$F^{s,t,m} = m\sum_{j=0}^{t-1}\sum_{k=0}^{s}\binom{t-1}{j}\binom{s}{k} F^{k,j,m-1} F^{s-k,t-k}. \tag{4.4.12}$$

Proof. Appendix C.3. □

4.4.3 Taylor approximation

The intention of this section is to compute the moments of Z_T:

$$E[Z_T^p] = E[(\log[\bar{\omega}_1 e^{X_T} + \bar{\omega}_2 e^{Y_T}])^p] \quad \text{for} \quad p = 1, 2, 3, 4.$$

We define the joint law of X_T and Y_T by

$$\bar{P}_T(B) := \mathbf{P}((X_T(\omega), Y_T(\omega)) \in B), \quad B \in \mathcal{B}^2([0, T]).$$

For a real-valued function f with domain $D \subset \mathbb{R}^d$ the Taylor formula around some point $x \in \mathbb{R}^d$ is given by

$$f(x+h) = \sum_{|\alpha| \leq n} \frac{1}{\alpha!}\partial^\alpha f(x) h^\alpha + R$$

where

$$R = \sum_{|\alpha| = n+1} \frac{1}{\alpha!}\partial^\alpha f(x + \theta h) h^\alpha$$

and the multiindex conventions $|\alpha| = \alpha_1 + \ldots + \alpha_d$, $\alpha! = \alpha_1! \cdot \ldots \cdot \alpha_d!$, $x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_d^{\alpha_d}$, and $\partial^\alpha = \partial^{|\alpha|}/\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_d} x_n$ for a multiindex $\alpha \in \mathbb{N}_0^d$.
The expansion point of the Taylor expansion is $(\mu_1 T, \mu_2 T) = (E[X_T], E[Y_T])$ from (4.3.24). Therefore considering $B_{\epsilon_1,\epsilon_2}(\mu_1 T, \mu_2 T)$ to be the neighbourhood of the point $(\mu_1 T, \mu_2 T)$ where the Taylor approximation is 'good'[15] we obtain for $(x, y) \in B_{\epsilon_1,\epsilon_2}(\mu_1 T, \mu_2 T)$

$$(\log[\omega_1 e^x + \omega_2 e^y])^p = \sum_{j+k \le n} \frac{1}{j!k!} F^{j,k,p}(\mu_1 T, \mu_2 T)(x - \mu_1 T)^j (y - \mu_2 T)^k + R_n$$

where[16] $R_n = o(|(x, y) - (\mu_1 T, \mu_2 T)|^n)$ for $(x, y) \to (\mu_1 T, \mu_2 T)$, and $|\cdot|$ is any norm in $\mathbb{R}^2$. Then we can integrate $(\log[\bar{\omega}_1 e^x + \bar{\omega}_2 e^y])^p$ with respect to x and y, and we have

$$\begin{aligned} E[(\log[\bar{\omega}_1 e^{X_T} + \bar{\omega}_2 e^{Y_T}])^p] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\log[\bar{\omega}_1 e^x + \bar{\omega}_2 e^y])^p \bar{P}_T(dx, dy) \\ &= \sum_{j+k \le n} \frac{1}{j!k!} F^{j,k,p}(\mu_1 T, \mu_2 T) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_1 T)^j (y - \mu_2 T)^k \bar{P}_T(dx, dy) + E[R_n] \\ &= \sum_{j+k \le n} \frac{1}{j!k!} F^{j,k,p}(\mu_1 T, \mu_2 T) M^{j,k} + E[R_n]. \end{aligned}$$

where the moments $M^{j,k}$ and the derivatives $F^{j,k,p}$ were computed in (4.4.6) respectively Lemma 4.13.
Unfortunately a good estimate of the remainder $E[R_n]$ is difficult to obtain. This is partly due to the fact that the derivatives $F^{j,k,p}$ are not given explicitly but merely in a recursive form. Therefore we restrict ourselves to showing numerically that for the applications in Chapter 4.5 we indeed have a very good approximation for low n leaving out $E[R_n]$.

4.4.4 Fitting an NIG distribution

The last step is to fit the parameters of a suitable distribution to the obtained four moments of Z_T. This can be done by a normal inverse Gaussian distribution. Its characteristic function χ_{NIG} exists for $z \in \{z \in \mathbb{C} | \Re(z) \in \mathbb{R}, -\alpha - \beta < \Im(z) < \alpha - \beta\}$ and is equal to

$$\chi_{NIG}(z) = e^{\psi_{NIG}(z)} \tag{4.4.13}$$

where

$$\psi_{NIG}(z) = i\mu z + \delta\sqrt{\alpha^2 - \beta^2} - \delta\sqrt{\alpha^2 - (\beta + iz)^2}. \tag{4.4.14}$$

[15] At this point we do not make precise this notion since we only discuss heuristics.
[16] See e.g. Königsberger (1997).

The cumulants have the following representation[17]:

$$\kappa_1 = \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}} + \mu, \quad \kappa_2 = \frac{\delta\alpha^2}{(\sqrt{\alpha^2-\beta^2})^3}$$
$$\kappa_3 = \frac{3\delta\alpha^2\beta}{(\sqrt{\alpha^2-\beta^2})^5}, \quad \kappa_4 = \frac{3\delta\alpha^2(\alpha^2+4\beta^2)}{(\sqrt{\alpha^2-\beta^2})^7}. \tag{4.4.15}$$

Given the first four (finite) moments $\mu'_j = E[Z^j], j = 1, \ldots, 4$, of a random variable Z the task is to invert (4.4.15) to obtain an approximation of its distribution within the class of NIG distributions. After transforming the moments to cumulants, a non-linear inverse problem must be solved for which a priori neither existence nor uniqueness is guaranteed. However, the class of NIG distributions seems to be tailor-made for this purpose because it turns out that, given the parameter restrictions, a unique solution exists.

The idea is to break up the problem by choosing a different parametrization than the one in terms of α, β, δ, and μ, which corresponds more closely to the interpretable parameters mean, volatility, skewness, and kurtosis. Most importantly, we introduce two parameters χ resp. ξ, which are scale and location invariant and represent measures for the skewness resp. the heaviness of the tails of the distribution of Z. The other parameters are the location parameter m and the scale parameter σ, which have an immediate interpretation in terms of mean and standard deviation. To cut a long story short, we change the parametrizaton from $(\alpha, \beta, \delta, \mu)$ to (ξ, χ, σ, m). This trick has already been used in Chapter 2, and the procedure is sketched in Appendix A.2.

The first step to apply this result is to transform the moments $\mu'_j, j = 1, \ldots 4$, into the central moments $\mu_j, j = 1, \ldots, 4$, by

$$\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= \mu'_2 - \mu'^2_1 \\ \mu_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1 \\ \mu_4 &= \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1. \end{aligned} \tag{4.4.16}$$

As we have

$$\mu_2 = \kappa_2, \quad \mu_3 = \kappa_3, \quad \mu_4 = \kappa_4 + 3\kappa_2^2$$

skewness $\bar{\mu}_3$ and (excess) kurtosis $\bar{\mu}_4$ are given by

$$\bar{\mu}_3 = \frac{\kappa_3}{(\kappa_2)^{3/2}} \quad \text{and} \quad \bar{\mu}_4 = \frac{\kappa_4}{(\kappa_2)^2}.$$

Using in addition the parameters ζ and ϱ with

$$\zeta = \delta\sqrt{\alpha^2-\beta^2} \quad \text{and} \quad \varrho = \beta/\alpha \tag{4.4.17}$$

[17] See Rydberg (1997).

we have from (4.4.15) that

$$\bar{\mu}_3 = 3\frac{\varrho}{\zeta^{1/2}} = 3\frac{\chi}{(1-\xi^2)^{1/2}} \quad \text{and} \quad \bar{\mu}_4 = 3\frac{1+4\varrho^2}{\zeta} = 3\frac{\xi^2+4\chi^2}{1-\xi^2}.$$

Solving for ξ and χ in terms of $\bar{\mu}_3$ and $\bar{\mu}_4$ yields

$$\xi = \left(\frac{\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2}{\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2 + 3}\right)^{1/2} \quad \text{and} \quad \chi = \frac{\bar{\mu}_3}{3}\left(\frac{3}{\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2 + 3}\right)^{1/2},$$

provided that $\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3 > 0$. In order not to lose the overview the result is summarized in the following proposition:

Proposition 4.14. *Let $\mu'_j = E[Z^j]$ for $j = 1, \dots, 4$ be the moments of a random variable Z. The central moments $\mu_j, j = 1, \dots, 4$, are then given by (4.4.16), and skewness $\bar{\mu}_3$ and kurtosis $\bar{\mu}_3$ by*

$$\bar{\mu}_3 = \frac{\mu_3}{\mu_2^{3/2}} \quad \textit{and} \quad \bar{\mu}_4 = \frac{\mu_4}{\mu_2^2} - 3.$$

Then if

$$\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2 > 0 \tag{4.4.18}$$

the unique NIG approximation to the distribution of Z in terms of matching the first four moments is given by

$$\begin{aligned} \xi &= \left(\frac{\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2}{\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2 + 3}\right)^{1/2} \\ \chi &= \frac{\bar{\mu}_3}{3}\left(\frac{3}{\bar{\mu}_4 - \frac{4}{3}\bar{\mu}_3^2 + 3}\right)^{1/2} \\ \sigma &= \mu_2^{1/2} \\ m &= \mu_1. \end{aligned}$$

□

As for condition (4.4.18), it should be added that it is not very restrictive. If Z has a symmetric distribution under P, i.e. if $\bar{\mu}_3 = 0$, then every distribution with heavier tails than the Gaussian distribution will satisfy (4.4.18). For the historical distribution of stock returns typical values for the skewness range from -0.5 to $+0.5$, whereas the excess kurtosis is usually considerably bigger than 1. For this range (4.4.18) is satisfied. For the distribution of the logarithm of the sum of exponential Lévy processes under the risk-neutral measure condition (4.4.18) is also likely to be not very restrictive.

One general remark concerning the NIG distribution should be made: Pricing options by means of the above procedure has an important advantage. There are at least three pricing methods which can easily be applied. First, as we already saw, Monte-Carlo pricing is easily feasible. Second, given that we accept the moment approximation by the NIG distribution described in this chapter we can price derivatives by numerical integration because the NIG distribution possesses an analytical density function and is closed under convolution. Third, of course, one can apply the technique of Fourier inversion introduced in this chapter.
The advantage of Fourier inversion is that it is by far the fastest way to compute prices if one realizes that it is based on an efficient numerical evaluation of one integral for many strike prices. In the realm of Lévy processes where analytical formula are almost completely absent, it should therefore be applied whenever possible. However, the other methods are suitable to check programming errors as well as the goodness of the moment approximation in Section 4.4.3, which is clearly a precondition for an application of the Fourier inversion method. This will be done in the next section.

4.5 Applications

The goal of this section is to apply the KTD model to concrete examples of data configuration in order to shed light on the model in three different aspects: We want to highlight the statistical properties of the model, especially concerning its modelling of the dependence between asset returns; we will be concerned about the goodness of approximation of the prices obtained by the Fourier inversion method introduced in the last paragraph; and we will compare the KTD prices with the ones for the standard model of multidimensional diffusion and linear dependence.

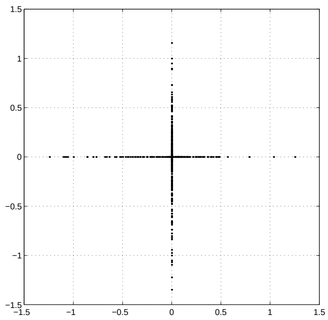
4.5.1 Return distributions in Kou's model

In this section we work with Kou's model in two dimensions where the marginal laws follow the one-dimensional Kou's model. The time horizon is $T = 1$. The parameters, which are maintained throughout the section, are given by the values shown in Table 4.1. The index j is 1 for X_1 and 2 for Y_1.

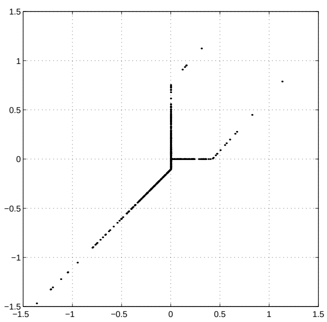
	σ_j	p_j	λ_j	λ_{j+}	λ_{j-}	vol_j	$skew_j$	$kurt_j$
Law of X_1	0.21	0.45	0.25	4.1	3.6	0.281	-0.357	7.714
Law of Y_1	0.3	0.5	0.4	3.7	3.6	0.387	-0.035	5.406

Tab. 4.1. Parameter values for the marginals.

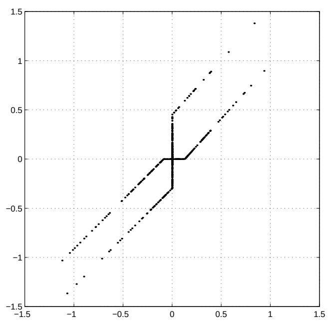
As these data are chosen to model a risk-neutral distribution, X and Y must be martingales. This is achieved by adjusting the drifts b_1 and b_2 given the complete specification of all other variables in the model. The part of linear dependence is modelled by a correlation of 0.5. X_1 has a high negative skewness as well as a high kurtosis. This means that its implied volatility smile is both very curved and



(a) $\theta_+ = \theta_- = \infty$



(b) $\theta_+ = 10$, $\theta_- = 1$



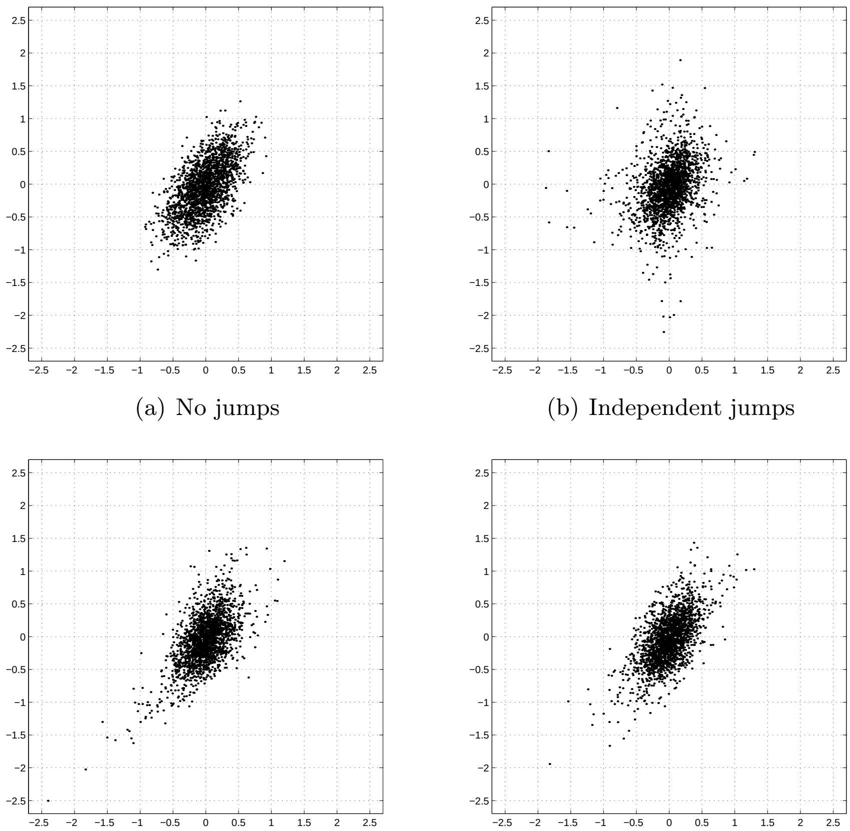
(c) $\theta_+ = 3$, $\theta_- = 2$

Fig. 4.3. Monte-Carlo simulation of two-dimensional jump measure, 500 simulations.

markedly asymmetric. By contrast, Y_1 implies a less curved and almost symmetric volatility smile.

Fig. 4.3 depicts the results of a simulation of the jump size distribution of the two-dimensional stochastic process (X, Y). In the first line jumps are independent ($\theta_+ = \theta_- = \infty$) whereas the lower two pictures show the simulation of dependent jumps: Complete dependence of the negative jumps ($\theta_- = 1$) of the left picture and a smaller degree of dependence in all the other cases.
As already mentioned in Section 4.2.4, where tail dependence for a Lévy copula was defined, it is rather difficult to say something about the relationship between tail dependence of the (two-dimensional) jump size distribution and the probabilistic tail dependence idea for the return distributions. This would be important because the latter is the one which is in principle observable and from which the tail dependence coefficient of the Lévy copula should be deduced.

The scatter plots in Fig. 4.4 reveal that - at least in our example - there should be a positive functional relationship between the two measures of tail dependence. The cases considered in Fig. 4.4 are the same as before except for the first one which is obtained from a two-dimensional normal distribution. In order to exclude scale



(a) No jumps (b) Independent jumps

(c) $\theta_+ = 10$, $\theta_- = 1$ (d) $\theta_+ = 3$, $\theta_- = 2$

Fig. 4.4. Monte-Carlo simulation of two-dimensional returns, 2000 simulations.

effects the latter has the same variances and covariance as the plot with independent jumps.

Fig. 4.5 displays an empirical scatter plot of two stock index returns: the German DAX and the Dutch AEX. By comparison with Fig. 4.4 we see that it has a similar structure as the case (d). Judging just by the look of the scatter plots with due care and attention, the index plot fits best with one of the two lowermost pictures in Fig. 4.4, and the pure diffusion and the independent jump cases (a) and (b) can be dismissed as not very realistic. This is an indication that a pronounced tail dependence is likely to be present in empirical data. However, we must have in mind that this comparison is not without problems because Fig. 4.5 shows a scatter plot from a statistical probability distribution whereas for option pricing we always have to be concerned about the risk-neutral distribution.

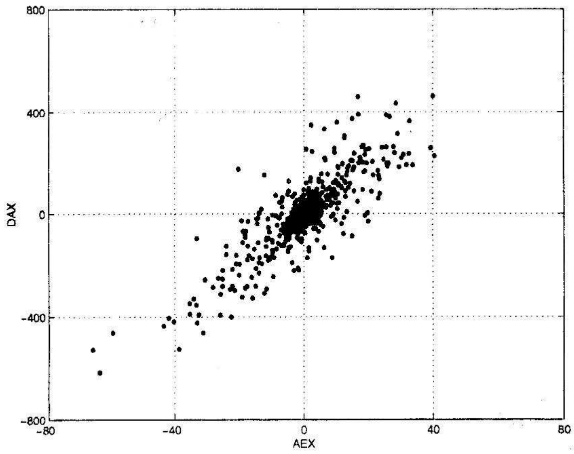


Fig. 4.5. Scatter plot of weakly returns of DAX and AEX (taken from Cont and Tankov (2004b), p. 166).

4.5.2 Goodness of approximation of the FFT pricing method

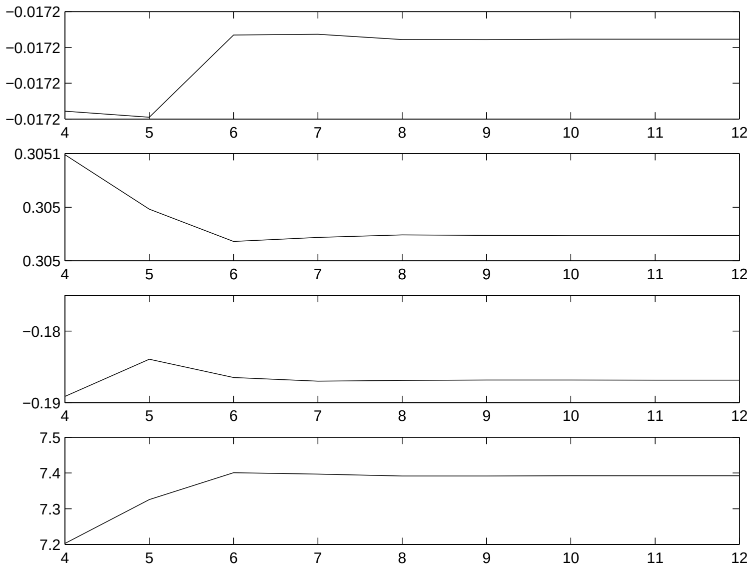
The key object of the Fourier method of Section 4.4 is the random variable Z_T, which was defined in Section 4.4.1 to be

$$Z_T = \log[\bar{\omega}_1 e^{X_T} + \bar{\omega}_2 e^{Y_T}]$$

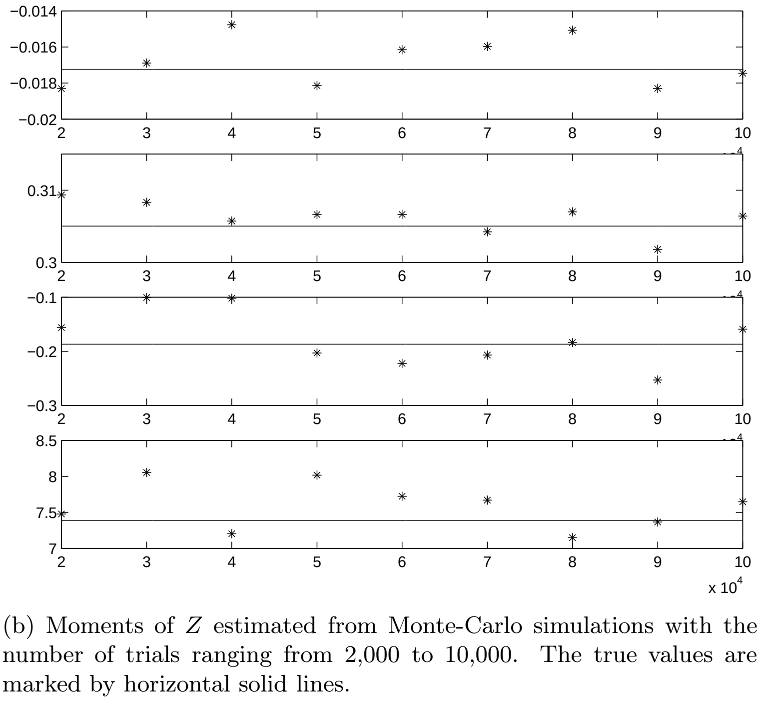
where (X, Y) is a KTD jump-diffusion process. Fig. 4.6(a) shows the convergence behaviour of the moment approximations of Z_T as a function of the order of the Taylor approximation for a European call option on a basket of two equities with marginal laws as in Table 4.1 and weights $\omega_1 = \omega_2 = 1$. The current prices of the equities are $S_0^X = S_0^Y = 10$, the correlation coefficient is 0.5, and the jumps are linked by the parameters $\theta_+ = 3$ and $\theta_- = 2$ of the Lévy copula C_{θ_+,θ_-}. The order of the Taylor approximation on the abscissa goes from 4 to 12. It can be deduced from Fig. 4.6(a) that a reasonable goodness of approximation is already obtained for orders of 7 or 8, which allows for a very fast computation. In contrast to this observation, Fig. 4.6(b) reveals that the estimation of the same moments from simulated data is quite unsatisfactory because the perceived convergence is very slow if visible at all. The results are displayed as a function of the number of drawings in a plain Monte-Carlo simulation which ranges from 20000 to 100000. The results become the worse the higher the moment is, a fact which has long been known in the literature[18].

We now proceed to compare the approximate prices obtained by the Fourier inversion technique with Monte-Carlo prices and, quite importantly, also with prices obtained in the pure diffusion model[19]. The latter model is important in the con-

[18]See e.g. Cont (2001).
[19]This model was defined at the end of Section 4.3.3.



(a) Approximate moments of Z as a function of the order of the Taylor approximation.

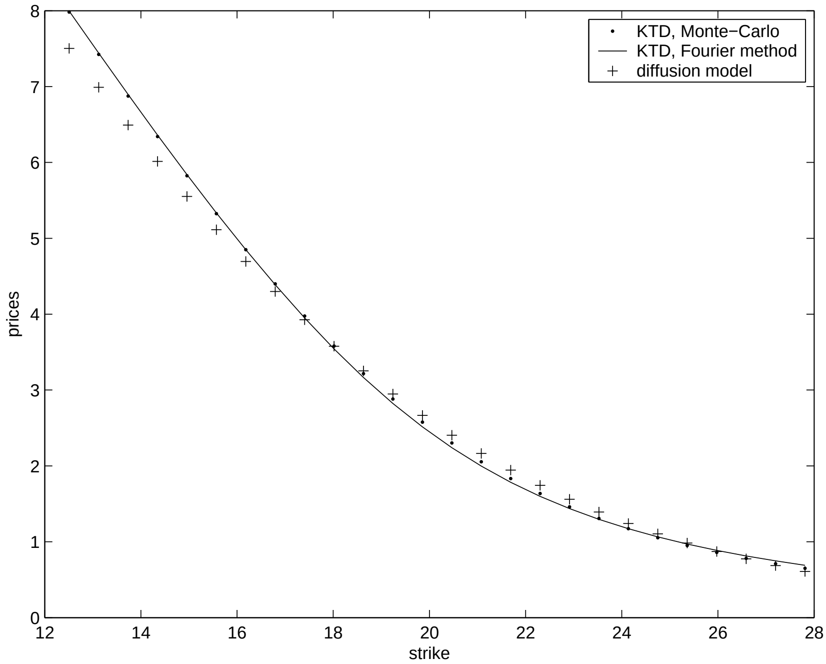


(b) Moments of Z estimated from Monte-Carlo simulations with the number of trials ranging from 2,000 to 10,000. The true values are marked by horizontal solid lines.

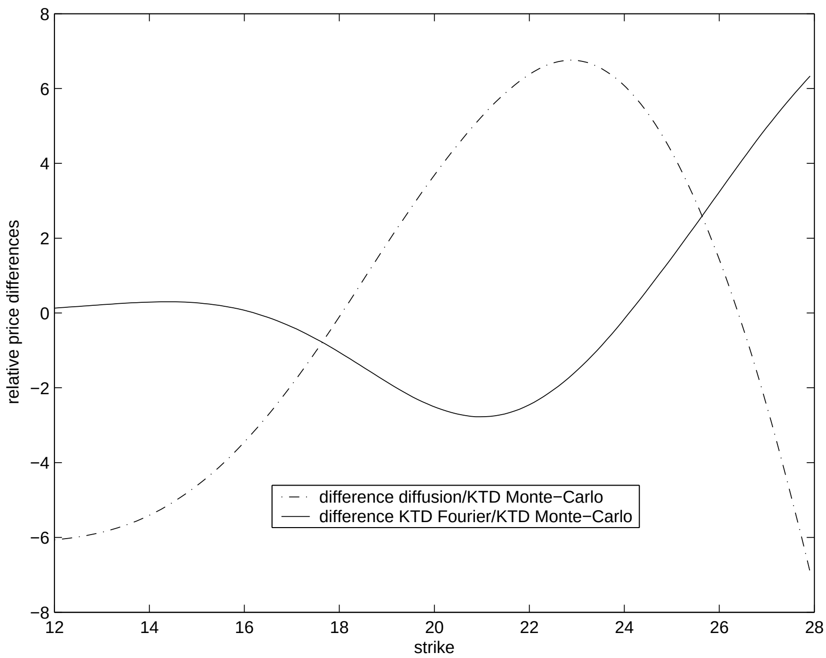
Fig. 4.6. Moments of Z: Both figures display in their four parts the moments on the horizontal axis: mean, standard deviation, skewness and kurtosis (from top to bottom).

text of this section because it is a benchmark relative to which pricing errors of the Fourier method in the KTD model must be assessed. For instance, a relative pricing error of 1% can be regarded as negligible if the pure diffusion price deviates from the true KTD price by 5% whereas for a 1% deviation the pricing error is quite big.
Prices in the pure diffusion model are computed with a very sufficient accurateness by a Monte-Carlo method with variance reduction, which is outlined in Appendix C.4 following Pellizzari (1998). As we have seen that a plain Monte-Carlo simulation in the case of Lévy processes usually converges very slowly, variance reduction is an even more pressing issue for simulating the KTD model, and we use control variates according to the method sketched at the beginning of Section 4.4. One observes that this method has a much better speed of convergence, and with 300,000 drawings the prices can be regarded as quite precise in the sense that different simulations do not visibly alter Fig. 4.7.
Fig. 4.7 compares different prices for strikes with moneyness ranging from 0.6 to 1.4, i.e. symmetrically around the at-the-money strike $K = S_0 = \omega_1 S_0^X + \omega_2 S_0^Y$. In Fig. 4.7(a) the solid line represents Fourier prices in the KTD model, and for a discrete set of strikes the dots stand for KTD Monte-Carlo prices, while the crosses are prices computed in the pure diffusion model. The differences can be seen more clearly in Fig. 4.7(b) which displays the relative price differences (in %) of KTD Fourier prices and diffusion prices in relation to the KTD Monte-Carlo prices. The latter are assumed to be accurate enough. For ITM options Fourier prices are quite precise, but become less accurate for ATM options with a maximum error of a bit more than 2%. But if we compare these differences to the intra-model differences reflected by the dash-dotted line, the picture shows that it is still much better to use the approximate KTD prices than prices in the diffusion model. For options far out-of-the-money the approximation becomes worse.

These results lead to more profound problems. The Fourier inversion method to determine approximate prices of basket options is accurate up to the first four moments of the distribution of the random variable Z_T from Section 4.4.1. This means that the price inaccuracies are due to the effects of the fifth and higher moments. But taking account of the fact that the estimation of the marginal return distributions proceeds in most cases using a parametric family of distributions with no more than four parameters and hence an exactness up to more than four moments is usually not within reach, it becomes clear that the higher moments of Z_T cannot be controlled. Consequently, the approximate Fourier prices can be seen as equally well justifiable as the seemingly exact Monte-Carlo prices. And for most strikes KTD Fourier prices are much closer to KTD Monte-Carlo prices than to diffusion prices. These facts should be taken into consideration when evaluating the goodness of the approximate pricing method.



(a) Option prices



(b) Option price differences in %

Fig. 4.7. Comparison of call option prices in the KTD model obtained by the FFT method and Monte-Carlo simulation, the latter both for the KTD and the common diffusion model. The number of simulations in both models is 300000. S_0 from (4.4.1) equals 20.

4.5.3 Pricing with the KTD model

Once again coming back to Fig. 4.7(b) and the dash-dotted line we see that the KTD model corrects diffusion prices qualitatively in the same way as in the better known case of standard (one-dimensional) European call options (see Chapter 3 or Chapter 4): Diffusion prices are too high for options around the ATM region and too small far away-from-the-money options[20].

Fig. 4.8 deals with a small experiment, which helps to assess the effect of tail dependence on basket option prices. The covariance decomposition (4.3.25), which was

$$\mathrm{cov}(X_T, Y_T) = T[c_{12} - \psi_d^{1,1}(0,0)],$$

shows how the total covariance of X_T and Y_T is decomposed into the covariance of the two involved Wiener processes and a part provided by the jump dependence. In a thought experiment we can change the parameters θ_+ and θ_- of C_{θ_+,θ_-}, which implies that $\psi_d^{1,1}(0,0)$ is modified. The decomposition result allows to adjust c_{12} such that $\mathrm{cov}(X_T, Y_T)$ stays the same. This means that while altering the jump dependence structure the only measure of dependence of a modeller who works only with the pure jump model and hence with linear dependence does not change. The result is that he cannot distinguish all these situations although they correspond to rather different jump structures. Fig. 4.8 depicts the impact of such a change of θ_+ and θ_- on the computed prices. The largest price difference is equal to 0.17, which is attained by the difference between completely dependent and independent jumps, and this appears to be quite sizable.

Another interesting insight can be gained from Fig. 4.9 which shows in two graphs the first four moments of Z_T as a function of $1/\theta_+ = 1/\theta_-$. This inverse parametrization is chosen because the case of independence ('$1/\theta_+ = 1/\theta_- = 0$') is more easily represented. The difference between both figures is that contrary to 4.9(a), in the second Fig. 4.9(b) the covariance of (X_T, Y_T) is held constant over all values of $1/\theta_+$ and $1/\theta_+$. Interestingly, these graphs show that the distribution of Z_T can be positively skewed although the marginals are both negatively skewed. This effect emerges even in Fig. 4.9(b), and it indicates once again that skewness, kurtosis and non-linear dependence in the data lead to unforeseeable properties of the random variable Z_T, which can be better captured by calibrating four rather than two moments.

4.6 Concluding remarks

This work belongs to a field of research starting with Tankov (2003) who proposed a method of how to construct multidimensional Lévy processes. Only very few contributions (e.g. Prause (1999), Rydberg (1997)) deal with this kind of problem. The innovation of Tankov (2003) is to bring the statistical copula idea of separating marginals and dependence into the world of Lévy processes and to provide a general

[20] Fig. 4.7(b) must be turned upside down to compare it to the corresponding figures in the previous chapters because the differences in those chapters are defined the other way round.

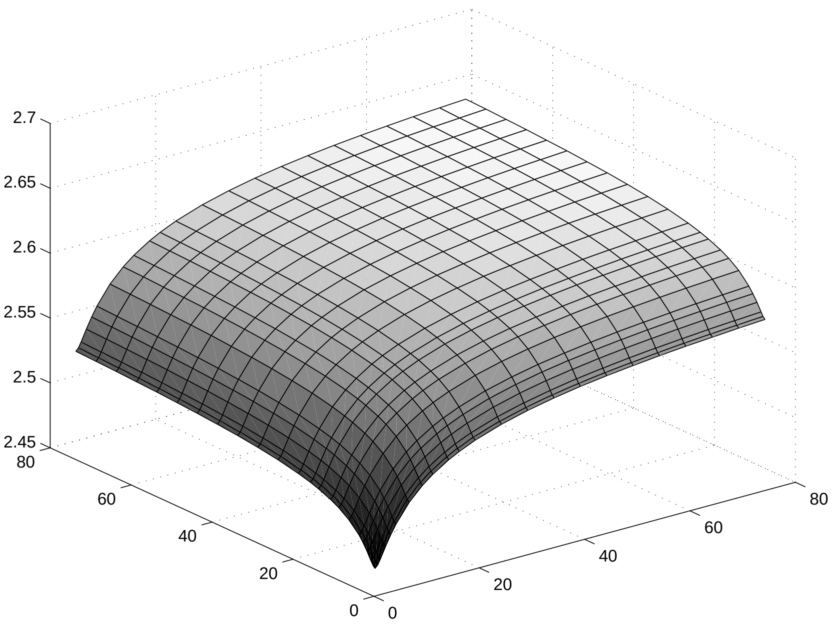
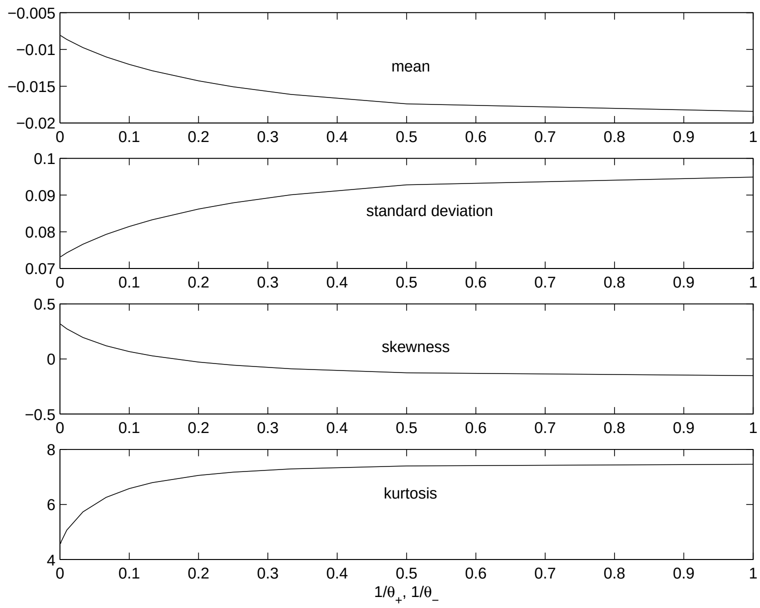
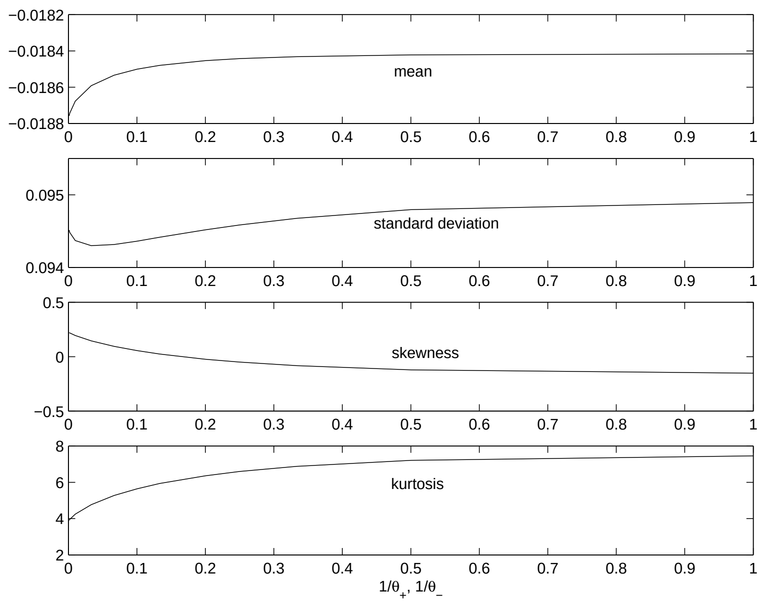


Fig. 4.8. Prices for a European ITM call coption ($K = S_0$) for different values of θ_+ and θ_- while at the same time the covariance is kept constant. The maximum difference is 0.17.

framework of constructing multidimensional Lévy processes rather than to generalize a special distribution to the multidimensional case. But the real challenge of the idea of a Lévy copula seems to be the requirement of tractability. And it is the tractability property which is the leitmotif of this contribution. We deliberately stayed as concrete as possible in order to find a very special model which can incorporate very realistic features (leptokurtotic and skewed marginals, tail dependent jumps) and at the same time remain tractable.
However, there are still some problems with this approach concerning the model itself as well as the proposed pricing methodology. As for the model techniques the question of how to estimate the two parameters has not been tackled, which clearly remains one of the most urgent things to do. More important than the technical side of the estimation procedure is availability of data. As we have to estimate the parameters of the risk-neutral as opposed to the statistical distribution there is a need for market prices of basket options in order to capture the risk-neutral dependence structure. It seems, though, that basket options are not very often traded at exchanges. An alternative approach would be to find economic reasons why it could perhaps be sufficient to take the statistical dependence structure. Of course, it would also be interesting to study the interrelation between the risk-neutral and the statistical dependence.
The pricing methodology in the KTD model immediately calls for an extension of the model to more than two dimensions because realistic basket options are mostly



(a) Moments of Z_T as a function of $1/\theta_+$ and $1/\theta_-$



(b) Like (a), but with constant covariance

Fig. 4.9. Parameters: σ_1=0.21, σ_2=0.3, p_1=0.45, p_2=0.5, λ_1=0.25, λ_2=0.4, λ_{1+}=4.1, λ_{2+}=3.7, λ_{1-}=3.6, λ_{2-}=3.6. Moments of Y as dependent of θ_+ and θ_-, which both are driven from 1 to 20. Note that the marginals both have negative skewness.

claims on whole indices rather than on just two assets. The first thing would be to extend the copula to more than two dimensions, which does not appear to be a straightforward task. Secondly, computations would become much less tractable, and it is not clear whether the feature of an analytical characteristic function would be preserved.

Finally, we could consider different (preferably infinitely often) differentiable payoff functions. If we define the strike price to be an asset, Eberlein and Papapantoleon (2004) provide an exact method (in the sense of being directly priced by Fourier inversion) to value a variety of exotic options depending on two assets. An interesting example of such an option is the exchange option with payoff function $p(x, y) = (x - y, 0)^+$. But as soon as this option is coupled with a strike price $K > 0$, i.e. $p(x, y) = (x - y, K)^+$, we have three assets just like in the case of the two-dimensional basket option, and the method of Eberlein and Papapantoleon (2004) cannot be applied any longer. It seems that such an option can be priced by a suitable modification of our approximation method.

Chapter 5

Risk-minimizing hedging

5.1 Introduction

Models of financial markets with assets driven by Lévy processes are in general incomplete. This means that not every contingent claim can be hedged completely, and hence one is forced to think about hedging strategies which cover the risk as well as possible in a certain sense. The hedging strategies that we consider here are assumed to use only the underlying asset and a riskless bond. In a complete market the hedging problem of a final payoff is solved by investing in a portfolio and pursuing a self-financing trading strategy which produces the same final payoff. A possible solution would be to maintain the latter payoff requirement while foregoing the self-financing assumption. Such a strategy can be found in a trivial way[1] and is not interesting at all. The idea of Föllmer and Sondermann (1986) is to replace the notion of a self-financing strategy by the weaker one of a mean self-financing strategy (see below). This leads to a non-trivial optimization problem, and one comes to the notion of a risk-minimizing hedging strategy, which we define in the sense of Schweizer (1991). The objective of this section is to compute a risk-minimizing strategy for options that depend on more than one asset in the framework of an exponential Lévy model.
This chapter has two main sections. The main result will not appear until Section 5.3 where we give an explicit presentation of a risk-minimizing hedging strategy and, correspondingly, of the minimal martingale measure in a multidimensional exponential Lévy world where the derivative is allowed to be written on more than one asset. This general result is of a very technical nature, and therefore the main economic ingredient might well be overlooked. Because of this, we decided to outline this procedure in Section 5.2 for the case where the derivative in question depends merely on one asset in a way that makes the structure of the general result more lucid. In doing this, we perform mere formal computations without bothering about whether the calculated objects exist or not. In fact, all of the objects in this first section will exist under suitable assumptions, but this will not be seen until Section 5.3.

[1] See Schweizer (2001).

There is a second reason why the introductory Section 5.2 was included. Risk-minimizing hedging can be given a twofold interpretation, depending on whether the tracking error is to be minimized under a (however obtained) risk-neutral or under the statistical probability measure. It will turn out that formally both approaches amount to the same result, but their interpretations are quite different: The first one is easier to obtain but has an economic flaw whereas the second one is mathematically more challenging but more meaningful in economic terms. Section 5.2 deals with this problem, discusses the results and classifies the present literature accordingly. The general result in Section 5.3 deals only with the second approach. The very first thing to do is to define the notions that we are going to talk about.

Definition 5.1. *Let M, N be square-integrable martingales. M and N are said to be orthogonal if $\langle M, N\rangle = 0$.*

This is not the common way of defining the notion of orthogonal martingales. Following the usual definition, two local martingales are then said to be orthogonal if their product is again a local martingale. However, for square-integrable martingales both definitions are equivalent.
Throughout this section we assume the asset price process to be given by a special semimartingale S with Doob-Meyer decomposition

$$S = S_0 + M + A$$

where M is a square-integrable martingale with $M_0 = 0$ and A a predictable process whose components have finite variation.

Definition 5.2. *Let Θ denote the space of all predictable $\mathbb{R}^n$-valued processes ϑ with*

$$E\left[\int_0^T \vartheta'_u d\langle M\rangle_u \vartheta_u + \left(\int_0^T |\vartheta'_u dA_u|\right)^2\right] < \infty \tag{5.1.1}$$

where $\langle M\rangle := (\langle M^i, M^j\rangle)_{i,j=1,\dots,n}$.
A portfolio strategy is a pair (C, ϑ) where $C > 0$ and $\vartheta \in \Theta$. Its associated value process is given by

$$V_t^\vartheta = C + \int_0^t \vartheta'_u dS_u.$$

We borrow from Choulli et al. (1998) the definition of a Föllmer-Schweizer decomposition:

Definition 5.3. *Given a semimartingale S, we say that a square-integrable $\mathcal{F}_T$-measurable random variable H admits a Föllmer-Schweizer decomposition if it can be written as*

$$H = H_0 + \int_0^T \vartheta'_u dS_u + \Gamma_T,$$

where H_0 is an $\mathcal{F}_0$-measurable random variable, $\vartheta \in \Theta$, and Γ is a square-integrable martingale starting at zero which is orthogonal to all components of M.

Central to finding a risk-minimizing hedging strategy is a martingale measure with special properties, the so-called *minimal martingale measure*, which is constructed for stochastic processes satisfying the structure condition from Schweizer (1994).

Definition 5.4. *Let S be a special semimartingale with Doob-Meyer decomposition $S = S_0 + M + A$, where*

$$A^i \ll \langle M^i \rangle, i = 1, \ldots, n$$

with predictable density υ^i. Moreover, let

$$\eta_t^i := \upsilon_t^i \Sigma_t^{ii} \quad \text{for} \quad i = 1, \ldots, d \tag{5.1.2}$$

and

$$\Sigma_t^{ij} := \frac{d\langle M^i, M^j \rangle_t}{dB_t} \quad \text{for} \quad i, j = 1, \ldots, d.$$

B is understood as any fixed increasing predictable càdlàg process starting at zero with $\langle M^i \rangle \ll B$. If there exists a predictable process $\hat{\lambda}$ with

$$\Sigma_t \hat{\lambda}_t = \eta_t, \quad \mathbf{P} - a.s. \text{ for } t \in [0, T], \tag{5.1.3}$$

and for the process $\hat{K}$, which is called the mean-variance trade-off process*, we have*

$$\hat{K}_t := \left\langle \int_0^\cdot \hat{\lambda}_u' dM_u \right\rangle_t = \int_0^t \hat{\lambda}_u' dA_u < \infty \quad \mathbf{P} - a.s. \text{ for all } t \in [0, T],$$

then S is said to satisfy the structure condition (SC).

As we will see later, for the case where S is an n-dimensional exponential Lévy process we can choose B with $B_t = t$, $t \in [0, T]$, such that (5.1.2) considerably simplifies to

$$\eta_t^i = \frac{dA_t^i}{dt} \quad \text{for} \quad i = 1, \ldots, d. \tag{5.1.4}$$

We give the definition of a *minimal martingale measure* in Föllmer and Schweizer (1991):

Definition 5.5. *A martingale measure $\hat{\mathbf{P}} \sim \mathbf{P}$ will be called* minimal *if any square-integrable $\mathbf{P}$-martingale Γ which is orthogonal to M under $\mathbf{P}$ remains a martingale under $\hat{\mathbf{P}}$.*[2]

The minimal martingale measure is in general a signed measure. In the following two sections we will give conditions under which it is indeed a probability measure.

Let a contingent claim H be given with assumptions made in Definition 5.3. The price of this claim at time t is denoted by $V(t, S_t)$ as a function of time and the price

[2] Föllmer and Schweizer (1991) assume in addition that $\hat{\mathbf{P}} = \mathbf{P}$ on $\mathcal{F}_0$. This condition is redundant in our case because we have assumed that $\mathcal{F}_0$ is the trivial σ-algebra.

of the underlying assets S. The financial notion 'risk-minimizing' from the beginning of the section translates henceforward into a mathematical definition. For this we still need to introduce the cost process Γ of a strategy ϑ.

$$\Gamma_t := V(t, S_t) - \int_0^t \vartheta'_u dS_u.$$

The following definition is based on Schweizer (2001).

Definition 5.6. *Let H be a contingent claim. A portfolio strategy (C, ϑ) with $C = V(0, S_0)$, $\vartheta \in \Theta$ and $V_T^\vartheta = H$ $\mathbf{P}-a.s.$ is called risk-minimizing for H if ϑ is such that Γ is a $\mathbf{P}$-martingale which is orthogonal to M.*

From the definition of the cost process we see that $\Gamma_0 = V(0, S_0)$. This means that with a strategy (C, ϑ) that satisfies only the martingale property of Definition 5.6 the cost process oscillates around the starting point $V(0, S_0)$. Such a strategy is called *mean-self-financing* because Γ is not identically equal to $V(0, S_0)$ as in the case of a self-financing strategy but still equal to $V(0, S_0)$ in expectation. A risk-minimizing strategy takes into account that Lévy markets are in general incomplete, and a full elimination of the risk of a derivative by continuous trading in the underlying and the risk-neutral asset is not possible. Hence it contents oneself with minimizing the hedging error in a certain sense: The strategy is such that the hedging error Γ is orthogonal (in the sense of Definition 5.1) to the price process of the underlying, i.e. it is the best possible way to cover the risk inherent in the terminal payoff H.

5.2 The one-dimensional case

In this section we use the one-dimensional case in order to assess the problems involved with risk-minimization in a jump process environment and to calculate the optimal hedge ratio for an exponential Lévy model. We assume throughout this section that $r = 0$. In the end, we will show that this assumption can be easily relaxed for a constant interest rate process. The hedge ratio can be interpreted in two different ways, and the two points of view will be discussed. We recall that the calculations in this subsection are completely of a formal nature and do not deal with existence questions. We take as given two probability measures $\mathbf{R}$ and $\tilde{\mathbf{R}}$ where at least the latter is a martingale measure.

Let the price process of a risky asset under the probability measure $\mathbf{R}$ be $S_t = S_0 e^{X_t}$ with $S_0 > 0$ and

$$X_t = bt + \sqrt{c} W_t + h(x) * (\mu^X - \nu)_t + (x - h(x)) * \mu_t^X \tag{5.2.1}$$

The Doob-Meyer decomposition into a martingale M and a predictable process A, obtained by an application of Itô's formula, reads

$$\begin{aligned} S_t &= S_0 + \sqrt{c} \int_0^t S_{u-} dW_u + (S_-(e^x - 1)) * (\mu^X - \nu)_t \\ &\quad + \left(b + \frac{c}{2} + \int_{\mathbb{R}} (e^x - 1 - h(x)) K(dx) \right) \int_0^t S_{u-} du. \end{aligned} \tag{5.2.2}$$

Given the martingale measure $\tilde{\mathbf{R}}$ we can give the pricing functional in the form of the risk-neutral pricing formula

$$V(t,S) := E^{\tilde{\mathbf{R}}}[w(S_T)|\mathcal{F}_t] = E^{\tilde{\mathbf{R}}}[w(S_T)|S_t = S] \tag{5.2.3}$$

where the last equality follows from the Markov property of the equity price process S. Applying Itô's formula to $V(\cdot,S)$ yields

$$\begin{aligned} V(t,S_t) &= V(0,S_0) + \int_0^t \frac{\partial V}{\partial S}(u,S_{u-})\sqrt{\tilde{c}}S_{u-}d\tilde{W}_u \\ &\quad +[V(u,S_{u-}e^x) - V(u,S_{u-})] * (\mu^X - \tilde{\nu})_t \end{aligned} \tag{5.2.4}$$

where we have used the price process S under the risk-neutral measure $\tilde{\mathbf{R}}$. The drift terms actually showing up in (5.2.4) must add up to zero[3] because $V(\cdot,S)$ is an $\tilde{\mathbf{R}}$-martingale. We set

$$\phi_t^c := \frac{\partial V}{\partial S}(t,S_{t-}) \quad \text{and} \quad \phi_t^d(x) := [V(t,S_{t-}e^x) - V(t,S_{t-})]. \tag{5.2.5}$$

Passing on to the actual hedging problem the cost process of the hedging strategy ϑ is defined by

$$\Gamma_t := V(t,S_t) - \int_0^t \vartheta_u dS_u.$$

The following assumption is worth emphasizing:

Assumption 5.7. *Γ is an $\mathbf{R}$-martingale.*

Therefore

$$\Gamma_t = \int_0^t (\phi_u^c - \vartheta_u)\sqrt{c}S_{u-}dW_u + \Big[\phi^d - \vartheta S_-(e^x - 1)\Big] * (\mu^X - \nu)_t \tag{5.2.6}$$

where again the drift terms sum up to zero.
The hedging strategy ϑ must be chosen such that the two $\mathbf{R}$-martingales Γ and M are orthogonal, i.e.

$$\langle \Gamma, M\rangle \equiv 0. \tag{5.2.7}$$

[3]This condition gives a partial integro-differential equation which is the jump analogy to the Black-Scholes partial differential equation. Explicitly, the PIDE is as follows

$$\frac{\partial V}{\partial t}(t,S) + \frac{\tilde{c}}{2}\frac{\partial^2 V}{\partial S^2}(t,S)S^2 + \int_{\mathbb{R}}[V(t,Se^x) - V(t,S) - \frac{\partial V}{\partial S}(t,S)S(e^x-1)]\tilde{K}(dx) = 0$$

and $V(T,S) = w(S_T)$. Note that the coefficients are the ones under some arbitrary risk-neutral measure $\tilde{\mathbf{R}}$, which is where the incompleteness problem emerges. Cont and Tankov (2004b) give a method of how to solve this equation (Cont and Tankov (2004b), equation (12.7), p.383) with numerical methods.

We have

$$\begin{aligned}\langle \Gamma, M\rangle_t &= \int_0^t (\phi^c_u - \vartheta_u) c S^2_{u-} du \\ &+ \int_0^t \int_{-\infty}^{\infty} \left[\left(\phi^d_u(x) - \vartheta S_{u-}(e^x - 1) \right) S_{u-}(e^x - 1) \right] K(dx) du\end{aligned}$$

hence we have to solve

$$(\phi^c - \vartheta) c S^2_- + \int_{-\infty}^{\infty} \left(\phi^d(x) - \vartheta S_-(e^x - 1) \right) S_-(e^x - 1) K(dx) = 0$$

which yields the predictable hedging strategy $\vartheta = \hat{\vartheta}$

$$\hat{\vartheta} = \frac{c\phi^c S_- + \int_{\mathbb{R}} (e^x - 1)\phi^d(x) K(dx)}{S_- \left(c + \int_{\mathbb{R}} (e^x - 1)^2 K(dx) \right)}. \tag{5.2.8}$$

It is intuitive to see how the results change if we set $r > 0$ in our setting. Because of

$$\bar{S}_t := e^{-rt} S_t = S_0 \exp\left[(b - r)t + \sqrt{c} W_t + h(x) * (\mu^X - \nu)_t + (x - h(x)) * \mu^X_t \right]$$

it suffices to replace b by $b - r$ in order to incorporate non-zero (but constant) interest rates according to the bond model (4.3.22). Going once again through the derivation of the hedging strategy (5.2.8), one observes that its form is not influenced. However, it should be self-evident that ϕ^c and ϕ^d look differently for different interest rates.

We now proceed to the economic interpretation of the abstract probability measures $\mathbf{R}$ and $\tilde{\mathbf{R}}$ in terms of the well-known probability measures: $\mathbf{P}$ is the statistical probability measure, and $\mathbf{Q}$ is a martingale measure, which is typically estimated directly from option price data, and $\hat{\mathbf{P}}$ is the minimal martingale measure in the sense of Schweizer (2001). One distinguishes between two ways of interpretation:

First case: $\mathbf{R} = \tilde{\mathbf{R}} = \mathbf{Q}$.

This case is conceptionally the simpler one: One starts directly with the specification of the asset price process S under a risk-neutral measure without any relation to the statistical martingale measure $\mathbf{P}$ and uses $\mathbf{Q}$ for the definition of V in (5.2.4). Assumption 5.7 is trivially satisfied because Γ is the difference between two $\mathbf{Q}$-martingales. Hence the orthogonality condition (5.2.7) is formulated in terms of the risk-neutral measure $\mathbf{Q}$ which means that the parameters showing up in the solution (5.2.8) are the risk-neutral ones, and the risk is minimized under $\mathbf{Q}$. But this is of course disputable: One would actually want to minimize the hedging error under $\mathbf{P}$ because, as a matter of fact, the costs of mishedging occur under the real-world measure $\mathbf{P}$ and not under $\mathbf{Q}$, which is just a pricing rule and not a statistical description of observed data[4]. Formula (5.2.8) in this context has been computed in Cont and Tankov (2004b), equation (10.39).

[4] See Cont and Tankov (2004b), p.339.

Second case: $\mathbf{R} = \mathbf{P}$, $\tilde{\mathbf{R}} = \hat{\mathbf{P}}$.

This approach resumes that criticism. Here the starting point is indeed the statistical measure $\mathbf{P}$ which is assumed not to be a martingale measure. Hence Assumption 5.7 is not fulfilled per se, and we need a very special change of measure. But this is exactly what $\hat{\mathbf{P}}$ does: Apart from being a martingale measure for S, it defines the process V via (5.2.3) such that Γ becomes a $\mathbf{P}$-martingale. The measure change process from $\mathbf{P}$ to $\hat{\mathbf{P}}$ is completely determined by the Girsanov quantities $\hat{\beta}$ and $\hat{y}$ (see Theorem 1.20) which are given by

$$\hat{\beta} = -\frac{\sqrt{c}\left(b + \frac{c}{2} + \int_{\mathbb{R}}(e^{x'} - 1 - h(x')K(dx')\right)}{c + \int_{\mathbb{R}}(e^{x'} - 1)^2 K(dx')}$$

and

$$\hat{y}(x) = 1 - \frac{b + \frac{c}{2} + \int_{\mathbb{R}}(e^{x'} - 1 - h(x')K(dx')}{c + \int_{\mathbb{R}}(e^{x'} - 1)^2 K(dx')}(e^x - 1). \tag{5.2.9}$$

At this point we are concerned with a well-known drawback of the minimal martingale measure and hence of risk-minimizing hedging: For many interesting asset price models the minimal martingale measure is signed. It follows that prices obtained by discounted expectation of a payoff function with respect to this measure are not arbitrage-free. This is typically the case for Lévy processes with unbounded jumps. But many Lévy processes used in finance have Lévy measures K with $\mathrm{supp}(K) = \mathbb{R}$. Here emerges a further advantage of the exponential Lévy model as opposed to the stochastic exponential Lévy model. In the latter case the measure change function of the minimal martingale measure is $y(x) = 1 - \theta x$ for a fixed constant θ. Assuming other conditions to be given, y being identically bigger than zero is equivalent to y describing an absolutely continuous change of measure. But y for the stochastic exponential Lévy model is affine-linear, hence positivity can only be achieved by restricting the size of the jumps, i.e. the support of the Lévy measure.
For the measure change function (5.2.9) this is different. We define the constants α and μ by

$$\begin{aligned} \alpha &:= b + \frac{c}{2} + \int_{\mathbb{R}} \left(e^{x'} - 1 - h(x')\right) K(dx'), \\ \mu &:= c + \int_{\mathbb{R}} \left(e^{x'} - 1\right)^2 K(dx'). \end{aligned}$$

and we have

$$\begin{aligned} \hat{y}(x) > 0 \quad \forall\, x \in \mathbb{R} &\iff \left(1 + \frac{\alpha}{\mu}\right) - \frac{\alpha}{\mu} e^x > 0 \\ &\iff -1 \leq \frac{\alpha}{\mu} \leq 0. \end{aligned} \tag{5.2.10}$$

This means that, given (5.2.10), in an exponential Lévy model with a Lévy measure even satisfying $\mathrm{supp}(K) = \mathbb{R}$ the minimal martingale measure is a real probability measure[5]. Writing (5.2.2) in the usual shorthand notation for stochastic integro-differential equations

$$\frac{dS_t}{S_{t-}} = \sqrt{c}dW_t + \int_{\mathbb{R}} (e^x - 1)(\mu^X - \nu)(dt, dx) + \left(b + \frac{c}{2} + \int_{\mathbb{R}} (e^x - 1 - h(x))K(dx) \right) dt$$

and forgetting about mathematical rigour, we can consider αdt and μdt as the mean and the variance of the returns of S over an infinitesimal time increment dt. As $\mu > 0$ a necessary condition for (5.2.10) is that α is negative. Adding that for $r > 0$ b in α must be replaced by $b - r$ this does not seem very unrealistic.

If (5.2.10) is not satisfied we have two alternatives: Either one restricts the sizes of the jumps, i.e. the model must be adjusted to the properties of the measure, or we accept that we have a signed measure. Though theoretically not meaningful the hedging strategy $\hat{\vartheta}$ and the measure change function y are still well-defined, and it would be a matter of hedging error comparison in a simulation to decide whether this approach is meaningful in a practical sense.

Clearly, the parameters in $\hat{\vartheta}$ in (5.2.8) are now the ones under $\mathbf{P}$, and the computation of V can be done in the usual way because S is again a Lévy process under $\hat{\mathbf{P}}$. This is obvious from the observation that $\hat{\beta}$ and $\hat{y}$ are non-stochastic and time-independent (see Theorem 1.20). Using a limit argument from a discrete-time argument Černý (2004a) and Černý (2004b) derive (5.2.8) with this interpretation as well as Colwell and Elliott (1993) for a more general model comprising the exponential Lévy model.

5.3 The multidimensional case

5.3.1 Hedging an option depending on n assets

Now we describe how an option on a general number n of assets can be hedged within the concept of risk-minimization. As the same ambiguity concerning the interpretation of the ocurring probability measures arises we focus here straight away on the second and more difficult case when the equity price process S is given under the statistical probability measure $\mathbf{P}$. In doing this, we clarify some technical issues which we postponed to this section.

Let $S^i = S_0^i e^{X^i}$ with $S_0^i > 0, i = 1, \dots, n$, describe the dynamics of the n risky assets under the statistical probability measure $\mathbf{P}$ where

$$X_t = (X_t^1, \dots, X_t^n)' = bt + \bar{c}W_t + h(x) * (\mu^X - \nu)_t + (x - h(x)) * \mu_t^X \qquad (5.3.1)$$

with $x = (x^1, \dots x^n)'$, $b = (b^1, \dots, b^n)'$ and $\bar{c} := \sqrt{c}$ denoting the square root[6] of $c = (c^{ij})_{i,j=1,\dots,n}$. W is a n-dimensional Wiener process with independent components and μ^X a jump measure on $[0, T] \times \mathbb{R}^n$. If we denote by $h^i(x)$ the i-th component of the vector $h(x)$ the components of X are given by

$$X_t^i = b^i t + (\bar{c}^i)' W_t + h^i(x) * (\mu^X - \nu)_t + (x^i - h^i(x)) * \mu_t^X, \quad i = 1, \dots, n.$$

[5] A similar result was obtained by Zhang (1994) in the framework of a jump-diffusion model.

[6] See the discussion in Section 1.3.

Applying Itô's lemma one obtains for $i = 1, \dots, n$

$$\begin{aligned}
S_t^i &= S_0^i + \int_0^t S_{u-}^i dX_u^i + \frac{1}{2}\sum_{k=1}^n (\bar{c}^{ik})^2 \int_0^t S_{u-}^i du + (S_-^i(e^{x^i} - 1 - x^i)) * \mu_t^X \\
&= S_0^i + b^i \int_0^t S_{u-}^i du + \int_0^t S_{u-}^i (\bar{c}^i)' dW_u + (S_-^i h^i(x)) * (\mu^X - \nu)_t \\
&\quad + (S_-^i[x^i - h^i(x)]) * \mu_t^X + \frac{c^{ii}}{2}\int_0^t S_{u-}^i du + (S_-^i(e^{x^i} - 1 - x^i)) * \mu_t^X \\
&= S_0^i + \int_0^t S_{u-}^i (\bar{c}^i)' dW_u + (S_-^i(e^{x^i} - 1)) * (\mu^X - \nu)_t \\
&\quad + \left(b^i + \frac{c^{ii}}{2} + \int_{\mathbb{R}^n} (e^{x^i} - 1 - h^i(x)) K(dx)\right) \int_0^t S_{u-}^i du. \qquad (5.3.2)
\end{aligned}$$

Hence we have the Doob-Meyer decompositions $S_t^i = S_0^i + M_t^i + A_t^i$ of S^i where

$$M_t^i := \int_0^t S_{u-}^i (\bar{c}^i)' dW_u + (S_-^i(e^{x^i} - 1)) * (\mu^X - \nu)_t, \qquad (5.3.3)$$

and

$$A_t^i := \left(b^i + \frac{c^{ii}}{2} + \int_{\mathbb{R}^n} (e^{x^i} - 1 - h^i(dx)) K(dx)\right) \int_0^t S_{u-}^i du. \qquad (5.3.4)$$

Moreover, given an $\mathcal{F}_T$-measurable payoff H with payoff function w, i.e. $H = w(S_T^1, \dots, S_T^n)$, we define the value process $V(\cdot, S)$ of H as a risk-neutral expectation with respect to $\hat{\mathbf{P}}$:

$$\begin{aligned}
V(t, S^1, \dots, S^n) &:= E^{\hat{\mathbf{P}}}[w(S_T^1, \dots, S_T^n) | \mathcal{F}_t] \\
&= E^{\hat{\mathbf{P}}}[w(S_T^1, \dots, S_T^n) | S_t^1 = S^1, \dots, S_t^n = S^n]. \qquad (5.3.5)
\end{aligned}$$

V is a uniformly integrable $\hat{\mathbf{P}}$-martingale by definition, hence the drift terms cancel after an application of Itô's lemma, and, given that $\hat{W}$ is a standard Wiener process under $\hat{P}$, we obtain

$$V(t, S_t) = V(0, S_0) + \int_0^t (\phi_u^c)' S_{u-} \bar{c}\, d\hat{W}_u + \phi^d(x) * (\mu^X - \nu^{\hat{\mathbf{P}}})_t \qquad (5.3.6)$$

by defining the predictable processes

$$\phi_t^c := \left(\frac{\partial V}{\partial S^1}(t, S_{t-}), \dots, \frac{\partial V}{\partial S^n}(t, S_{t-})\right)' \qquad (5.3.7)$$

and

$$\phi_t^d(x) = \left[V(t, S_{t-}^1 e^{x^1}, \dots, S_{t-}^n e^{x^n}) - V(t, S_{t-}^1, \dots, S_{t-}^n)\right]. \qquad (5.3.8)$$

For this to be feasible we assume that $V(t,S)$ is differentiable in t and twice differentiable in S.[7]

We recall the definitions $J(x) := (e^{x^1}-1,\dots,e^{x^n}-1)$ and $\mathcal{S} := \mathrm{diag}(S^1,\dots,S^n)$ that we have made in Chapter 1.
We are now ready to define the constant matrix $\mu \in \mathbb{R}^{n\times n}$, the constant vectors $\alpha \in \mathbb{R}^n$ and the predictable $\mathbb{R}^n$-valued process d by

$$\mu \;:=\; c + \int_{\mathbb{R}^n} J(x)J'(x)K(dx), \tag{5.3.9}$$

$$\alpha \;:=\; \left(b^i + \frac{c^{ii}}{2} + \int_{\mathbb{R}^n}(e^{x^i}-1-h^i(x))K(dx)\right)_{i=1,\dots,n}, \tag{5.3.10}$$

$$d_t \;:=\; c\mathcal{S}_{t-}\phi_t^c + \int_{\mathbb{R}^n}\phi_t^d(x)J(x)K(dx). \tag{5.3.11}$$

By means of the representations (5.3.3) and (5.3.4) we obtain then explicit representations of the processes $\langle M^i, M^j\rangle$ and A^i for $i,j = 1,\dots,n$. They are given by[8]

$$\begin{aligned} d\langle M^i, M^j\rangle_t &= \mu^{ij}S^i_{t-}S^j_{t-}dt \\ dA^i_t &= \alpha^i S^i_{t-}dt. \end{aligned}$$

Definitions (5.3.9), (5.3.10) and (5.3.11) allow us to give M and A compactly in matrix notation:

$$M_t \;=\; \int_0^t \mathcal{S}_{u-}\bar{c}dW_u + (\mathcal{S}_- J(x)) * (\mu^X - \nu^{\mathbf{P}})_t, \tag{5.3.12}$$

$$A_t \;=\; \int_0^t \mathcal{S}_{u-}\alpha du. \tag{5.3.13}$$

The main theorem can now be stated in terms of the previously defined constants. We repeat that by $|\cdot|$ we refer both to a vector norm and to the matrix norm which is generated by the former. We have in particular that $|Ax| \le |A||x|$ for $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$ for every $m,n \in \mathbb{N}$ [9].

Theorem 5.8. *Let $S = S_0e^X$ with $S^i_0 > 0$, $i = 1,\dots,n$, be the n-dimensional exponential Lévy process (5.3.1) with $X = (b,c,K)_{\mathbf{P}}$ and the Lévy-Itô decomposition $S = S_0 + M + A$ with M and A given by (5.3.12) and (5.3.13), and let*

$$\int_{|x|\ge 1} e^{2|x|}K(dx) < \infty. \tag{5.3.14}$$

[7] For $n = 1$ there are convenient conditions in terms of the diffusion component c and the Lévy measure K for this assumption to hold. This is the case, for instance, for a jump-diffusion process with a non-zero diffusion component (See Cont and Tankov (2004b), p.385 and Proposition 3.12.). Unfortunately there does not seem to be a similar result for $n > 1$.

[8] We have used (1.3.2).

[9] See e.g. Barnett (1990).

Let the matrix μ, which is given in (5.3.9), be non-singular $\mathbf{P}-a.s$, and furthermore suppose that

$$1-\alpha'\mu^{-1}J(x)>0 \quad \forall\, x\in \operatorname{supp}(K), \tag{5.3.15}$$

where $\alpha\in\mathbb{R}^n$ is given in (5.3.10).
The predictable processes $\Sigma=(\Sigma^{ij})_{i,j=1,\ldots,n}\in\mathbb{R}^{n\times n}$ and $\eta=(\eta^i)_{i=1,\ldots,n}\in\mathbb{R}^n$ are defined by

$$\Sigma_t^{ij}=\Sigma_t^{ji}:=\mu^{ij}S_-^iS_-^j=(\mathcal{S}_-\mu\mathcal{S}_-)^{ij}$$

and

$$\eta_t^i:=\alpha^iS_-^i=(\mathcal{S}_-\alpha)^i.$$

a) Then there exists a predictable process $\hat{\lambda}$ that satisfies

$$\Sigma\hat{\lambda}=\eta \quad \mathbf{P}-a.s., \tag{5.3.16}$$

and the martingale measure $\hat{\mathbf{P}}$ is given by the measure change process $\hat{Z}$, which is a square integrable $\mathbf{P}$-martingale:

$$\hat{Z}_t:=\left.\frac{d\hat{\mathbf{P}}}{d\mathbf{P}}\right|_{\mathcal{F}_t}=\mathcal{E}\left(-\int_0^{\cdot}\hat{\lambda}'dM\right)_t,\quad t\in[0,T]. \tag{5.3.17}$$

It is such that S is again a Lévy process under $\hat{\mathbf{P}}$.

b) Let H be an $\mathcal{F}_T$-measurable random variable with $E^{\mathbf{P}}[H^2]<\infty$, which is of the form $H=w(S_T)$ for a positive measurable function $w:\mathbb{R}^n\to\mathbb{R}_+$ that satisfies for $L>0$ the Lipschitz condition

$$|w(x)-w(y)|\le L|x-y|,\quad x,y\in\mathbb{R}^n.$$

Then the process $V(\cdot,S)$ in (5.3.5) is well-defined. If we assume that $V(t,S)$ is differentiable, once in t and twice in S, then via the use of definitions (5.3.7), (5.3.8) and (5.3.11) we obtain the process $\delta=(\delta^i)_{i=1,\ldots,n}\in\mathbb{R}^n$ by

$$\delta_t^i:=d^iS_-^i=(\mathcal{S}_-d)^i.$$

Then there is a predictable solution $\hat{\vartheta}\in\Theta$ to the linear problem

$$\Sigma\hat{\vartheta}=\delta \quad \mathbf{P}-a.s. \tag{5.3.18}$$

such that the process

$$\Gamma_t=V(t,S_t)-\int_0^t\hat{\vartheta}_u'dS_u,$$

is a square-integrable real-valued $\mathbf{P}$-martingale which is orthogonal to all components of the n-dimensional vector process M.

Proof. a) First of all, we state that condition (5.3.14) entails that M is square-integrable martingale. The non-singularity of μ implies the one of Σ because $\mathcal{S}$ is non-singular as $S_0^i > 0\ \forall i = 1, \ldots, n$. Moreover, A^i is absolutely continuous with respect to $\langle M^i \rangle$. This observation and the fact that $\langle M^i \rangle$ is absolutely continuous with respect to the nonstochastic process $B_t = t$ entail the validity of (5.1.4), and the problem (5.1.3) can indeed be reduced to (5.3.16). We have then the measure change process as the stochastic exponential of the process

$$-\int_0^t \hat{\lambda}' dM = -\int_0^t \hat{\lambda}' \mathcal{S} \sqrt{c} dW - \left[\hat{\lambda}' \mathcal{S} J(x)\right] * (\mu^X - \nu^{\mathbf{P}}),$$

hence

$$\hat{\beta}' = -\hat{\lambda}' \mathcal{S} \sqrt{c} \quad \text{and} \quad \hat{y}(x) - 1 = -\hat{\lambda}' \mathcal{S} J(x). \tag{5.3.19}$$

The problem $\Sigma \hat{\lambda} = \eta$ becomes $\mathcal{S} \mu \mathcal{S} \hat{\lambda} = \mathcal{S} \alpha$, and we have a unique $\hat{\lambda} = \mathcal{S}^{-1} \mu^{-1} \alpha$ by invertibility of $\mathcal{S}$ and μ. On the one hand this implies that the mean-variance trade-off process $\hat{K}$ becomes

$$\hat{K}_t = \int_0^t \hat{\lambda}_u' dA_u = \int_0^t \alpha \mu^{-1} \mathcal{S}^{-1} \mathcal{S} \alpha dt = \alpha \mu^{-1} \alpha\, t,$$

which is deterministic and hence finite on $[0, T]$. Hence the structure condition (SC) from Definition 5.4 is fulfilled.

On the other hand, we have the Girsanov quantities

$$\hat{\beta} = -\hat{\lambda}' \mathcal{S} \sqrt{c} = -\alpha' \mu^{-1} \sqrt{c} \quad \text{and} \quad y(x) - 1 = -\hat{\lambda}' \mathcal{S} J(x) = -\alpha' \mu^{-1} J(x) \tag{5.3.20}$$

which are neither random nor depend on time. This implies that S is a Lévy process under $\hat{\mathbf{P}}$. A short computation shows that condition (1.3.18) in Proposition 1.21 is fulfilled, therefore $\hat{Z}$ is a uniformly integrable martingale. As it is easily seen, the form of the Girsanov quantities in conjunction with (5.3.14) shows the $\mathbf{P}$-square-integrability of $\int \hat{\lambda} dM$, hence the $\mathbf{P}$-square-integrability of $\hat{Z}$ via the structure condition (SC) according to Choulli et al. (1998), Proposition 3.7. By (5.3.15) the measure change function y is positive on the support of K, from which it follows that $\hat{\mathbf{P}}$ is indeed a probability measure.

The probability measure $\hat{\mathbf{P}}$ is a martingale measure: Changing the measure in (5.3.2) leads to a drift part which is the time integral of the process

$$\begin{aligned}
& \mathcal{S} \sqrt{c} \hat{\beta} + \mathcal{S} \int_{\mathbb{R}^n} J(x)(y(x) - 1) K(dx) - \mathcal{S} \alpha \\
= \quad & \mathcal{S} c \mathcal{S} \hat{\lambda} + \int_{\mathbb{R}^n} \mathcal{S} J(x) J(x)' \mathcal{S} \hat{\lambda} K(dx) - \mathcal{S} \alpha \\
= \quad & \left(\mathcal{S} c \mathcal{S} + \int_{\mathbb{R}^n} \mathcal{S} J(x) J(x)' S K(dx) \right) \Sigma^{-1} \mathcal{S} \alpha - \mathcal{S} \alpha = 0
\end{aligned}$$

because the term in brackets in the previous line equals Σ. Hence the drift term is zero, and by Lemma (1.19) S is a true $\hat{\mathbf{P}}$-martingale.

b) By Schwarz's inequality $E^{\mathbf{P}}[H^2] < \infty$ and the $\mathbf{P}$-square-integrability of $\hat{Z}$ imply that $V(\cdot, S)$ is well-defined. The solution to problem (5.3.18) is again unique, and we can write $\hat{\vartheta} = \Sigma^{-1}\delta$.

The next thing to show is that its solution $\hat{\vartheta}$ is in Θ, i.e.

$$E^{\mathbf{P}}\left[\int_0^T \vartheta_u' d\langle M\rangle_u \vartheta_u + \left(\int_0^T |\vartheta_u' dA_u|\right)^2\right] < \infty.$$

We begin with the first integral. Using $\hat{\vartheta} = \mathcal{S}^{-1}\mu^{-1}d$ and $\Sigma = \mathcal{S}\mu\mathcal{S}$ we obtain

$$\int_0^t \hat{\vartheta}_u' d\langle M\rangle_u \hat{\vartheta}_u = \int_0^t d_u'\mu^{-1}d_u du \tag{5.3.21}$$

and subsequently

$$\int_0^t |\hat{\vartheta}_u' d\langle M\rangle_u \hat{\vartheta}_u| \leq |\mu^{-1}| \int_0^t |d_u|^2 du,$$

hence it is sufficient to show that $E^{\mathbf{P}}[|d|^2] < \infty$. We have because of Lipschitz continuity

$$\begin{aligned}
&|V(t, S_{t-}^1 + \Delta, \ldots, S_{t-}^n) - V(t, S_{t-}^1, \ldots, S_{t-}^n)| \\
&\leq \; E^{\hat{\mathbf{P}}}\Big[|w((S_{t-}^1 + \Delta)e^{X_{T-t}^i}, S_{t-}^2 e^{X_{T-t}^2}, \ldots, S_{t-}^n e^{X_{T-t}^n}) \\
&\qquad -w(S_{t-}^1 e^{X_{T-t}^1}, S_{t-}^2 e^{X_{T-t}^2}, \ldots, S_{t-}^n e^{X_{T-t}^n})|\Big|\mathcal{F}_t\Big] \\
&\leq \; L|\Delta| E^{\hat{\mathbf{P}}}[e^{X_{T-t}^1}|\mathcal{F}_t] = L|\Delta| \quad \mathbf{P} - a.s.
\end{aligned}$$

because $E^{\hat{\mathbf{P}}}[e^{X_{T-t}^1}|\mathcal{F}_t] = E^{\hat{\mathbf{P}}}[e^{X_T^1 - X_t^1}|\mathcal{F}_t]$ due to the property of independent and stationary increments of X and the $\hat{\mathbf{P}}$-martingale property of $e^{X_t^1}$, which implies that the partial derivative of V with respect to S^1 is $\mathbf{P} - a.s.$ bounded by the constant L. Repeating this analysis for the other partial derivatives we see that for a positive constant C_1 we have due to the equivalence of all norms on $\mathbb{R}^n$

$$|\phi^c| \leq C_1|\phi^c|_\infty \leq C_1 L =: C_c \quad \mathbf{P} - a.s. \tag{5.3.22}$$

where $|\cdot|_\infty$ is the maximum norm $|x|_\infty := \max\{x^1, \ldots, x^n\}$. For a vector $x = (x^1, \ldots, x^n)$ we define $\mathbf{e}^x = \mathrm{diag}(e^{x^1}, \ldots, e^{x^n}) \in \mathbb{R}^{n\times n}$. Continuing with the treatment of ϕ^d, we obtain for $\phi_t^d(x) = V(t, S_- e^x) - V(t, S_-)$

$$\begin{aligned}
&|V(t, S_{t-}^1 e^{x^1}, \ldots, S_{t-}^n e^{x^n}) - V(t, S_{t-}^1, \ldots, S_{t-}^n| \\
&\leq \; E^{\hat{\mathbf{P}}}\left[|w(S_{t-}^1 e^{x^1} e^{X_{T-t}^1}, \ldots, S_{t-}^n e^{x^n} e^{X_{T-t}^n}) - w(S_{t-}^1 e^{X_{T-t}^1}, \ldots, S_{t-}^n e^{X_{T-t}^n})|\Big|\mathcal{F}_t\right] \\
&\leq \; L E^{\hat{\mathbf{P}}}\left[|\mathcal{S}_{t-}\mathbf{e}^{X_{T-t}} J(x)|\big|\mathcal{F}_t\right] \leq L|\mathcal{S}_{t-}| E^{\hat{\mathbf{P}}}\left[|\mathbf{e}^{X_{T-t}}|\big|\mathcal{F}_t\right] |J(x)|.
\end{aligned} \tag{5.3.23}$$

By linearity and monotonicity of the conditional expectation we have for the maximum norm $|\cdot|_\infty$

$$\begin{aligned} E^{\hat{\mathbf{P}}}[|\mathbf{e}^{X_{T-t}}|_\infty|\mathcal{F}_t] &= E^{\hat{\mathbf{P}}}[\max\{e^{X^1_{T-t}},\dots,e^{X^n_{T-t}}\}|\mathcal{F}_t] \\ &\le \sum_{i=1}^n E^{\hat{\mathbf{P}}}[e^{X^i_{T-t}}|\mathcal{F}_t] \le n \quad \mathbf{P}-a.s. \end{aligned}$$

because all e^{X^i} are positive. Again due to the norm equivalence on $\mathbb{R}^{n\times n}$ we then have from (5.3.23)

$$|\phi^d_t(x)| \le nC_d|\mathcal{S}_{t-}|\,|J(x)| \tag{5.3.24}$$

for a constant $C_d > 0$.
From (5.3.22), (5.3.24) and $(a+b)^2 \le 2(a^2+b^2)$ for real a, b we obtain

$$\begin{aligned} |d|^2 &\le 2|c|^2|S_{t-}|^2|\phi^c|^2 + 2\left(\int_{\mathbb{R}^n}|\phi^d_t(x)||J(x)|K(dx)\right)^2 \\ &\le 2|\mathcal{S}_t|^2\left\{|c|^2C_c^2 + n^2C_d^2\left(\int_{\mathbb{R}^n}|J(x)|^2K(dx)\right)\right\} \end{aligned}$$

The integral with respect to K in this expression exists because of (5.3.14). This term is $\mathbf{P}$-integrable, which again entails $\mathbf{P}$-integrability of (5.3.21). Likewise we see that using Schwarz's inequality

$$\begin{aligned} E^{\mathbf{P}}\left[\left(\int_0^t \hat{\vartheta}'_u dA_u\right)^2\right] &= E^{\mathbf{P}}\left[\left(\int_0^t d'_u\mu^{-1}\alpha du\right)^2\right] \\ &\le |\mu^{-1}\alpha|^2 t E^{\mathbf{P}}[\int_0^t |d_u|^2 du] < \infty. \end{aligned}$$

All this implies that $\hat{\vartheta} \in \Theta$.

We show now that Γ is a $\mathbf{P}$-square-integrable martingale. First it is a local martingale: $V(\cdot, S)$ has the representation (5.3.6) under the measure $\hat{\mathbf{P}}$, and changing the measure back to $\mathbf{P}$ results in the drift of the process Γ under $\mathbf{P}$ which is equal to

$$-(\phi^c)'S\sqrt{c}\hat{\beta} - \int_{\mathbb{R}^n}\phi^d_t(y(x)-1)\ K(dx) - \vartheta' S\alpha.$$

Plugging in $\hat{\beta}$ and $\hat{y}$ from (5.3.19) and afterwards $\hat{\lambda} = \Sigma^{-1}S\alpha$ and $\vartheta = \Sigma^{-1}\mathcal{S}d$ yields

$$\begin{aligned} &(\phi^c)'\mathcal{S}c\mathcal{S}\hat{\lambda} + \int_{\mathbb{R}^n}\phi^d_t J'(x)\mathcal{S}\hat{\lambda}K(dx) - \vartheta'\mathcal{S}\alpha \\ =\ &\left(c\mathcal{S}\phi^c + \int_{\mathbb{R}^n}\phi^d_t J(x)K(dx) - d\right)'\mathcal{S}\Sigma^{-1}\mathcal{S}\alpha = 0 \end{aligned}$$

in view of the definition of d in (5.3.11). Here we have used that the matrices $\mathcal{S}$, c and Σ^{-1} are symmetric. The symmetry of Σ^{-1} follows immediately from the

considerations on p.15.
Γ is even a square-integrable martingale: As it is a local martingale, the drift is zero, and we have

$$\Gamma_t = V(0,S_0) + \int_0^t (\phi^c - \hat{\vartheta})' \mathcal{S}_{u-} \bar{c} dW_u + [\phi^d(x) - \hat{\vartheta}' \mathcal{S}_- J(x)] * (\mu^X - \nu). \quad (5.3.25)$$

We must show that for $u \in [0,T]$ we have $E^{\mathbf{P}}[\Gamma_u^2] < \infty$. That Γ is a true square-integrable martingale follows then from Protter (1995), Theorem I.47, by

$$E^{\mathbf{P}}\left[\sup_{u\in[0,T]} |\Gamma_u|\right] \leq \left(E^{\mathbf{P}}\left[\sup_{u\in[0,T]} |\Gamma_u|^2\right]\right)^{1/2} \leq 2\left(\sup_{u\in[0,T]} E^{\mathbf{P}}[\Gamma_u^2]\right)^{1/2},$$

where we have used first Jensens's and then Doob's inequality[10]. The square-integrability of both integrals in (5.3.25) can be seen from the isometry formulas[11] So we compute, using $\hat{\vartheta} = \mathcal{S}^{-1}\mu^{-1}d$

$$\begin{aligned} & E^{\mathbf{P}}\left[\int_0^t \left((\phi_u^c - \hat{\vartheta}_u)' \mathcal{S}_{u-} \bar{c}\right)^2 du\right] \\ \leq \ & 2|c| E^{\mathbf{P}}\left[\int_0^t \left\{|\phi_u^c|^2 |\mathcal{S}_{u-}|^2\right\} du\right] + 2|\mu^{-1}|^2 |c| E^{\mathbf{P}}\left[\int_0^t |d_u|^2 du\right]. \end{aligned}$$

This is finite given the $\mathbf{P}-a.s.$ boundedness of $|\phi^c|$ and the $\mathbf{P}$-square-integrability of $|d|$. Likewise we obtain for the second integral considering condition (5.3.14)

$$\begin{aligned} & E^{\mathbf{P}}\left[\int_0^t \int_{\mathbb{R}^n} \left(\phi_u^d(x) - \hat{\vartheta}_u' \mathcal{S}_{u-} J(x)\right)^2 K(dx)du\right] \\ \leq \ & 2E^{\mathbf{P}}\left[\int_0^t \int_{\mathbb{R}^n} |\phi_u^d(x)|^2 K(dx)du\right] + 2|\mu^{-1}|^2 E^{\mathbf{P}}\left[\int_0^t \int_{\mathbb{R}^n} |d_u|^2 |J(x)|^2 K(dx)du\right]. \end{aligned}$$

The final statement to verify is the orthogonality of Γ and $M^i, i = 1, \ldots, n$. In view of the representations of (5.3.3) and (5.3.25) an application of Proposition 1.7e yields

$$\begin{aligned} & \langle \Gamma, M^i \rangle_t \\ = \ & \int_0^t (\phi_u^c - \hat{\vartheta}_u)' \mathcal{S}_{u-} c^i S_{u-}^i du + \int_0^t \int_{\mathbb{R}^n} (\phi_u^d(x) - \hat{\vartheta}_u' \mathcal{S}_{u-} J(x)) S_{u-}^i (e^{x^i} - 1) K(dx)du \\ = \ & \int_0^t (\phi_u^c)' \mathcal{S}_{u-} c^i S_{u-}^i du + \int_0^t \int_{\mathbb{R}^n} \phi_u^d(x) S_{u-}^i (e^{x^i} - 1) K(dx)du \\ & - \int_0^t \hat{\vartheta}_u' \mathcal{S}_{u-} S_{u-}^i \left\{c^i + \int_{\mathbb{R}^n} J(x)(e^{x^i} - 1) K(dx)\right\} du, \end{aligned}$$

[10]See e.g. Protter (1995), p.12.
[11]See Revuz and Yor (2001), Theorem IV.2.2, for the first, diffusion-related and Proposition 1.7d of this thesis for the second, jump-related integral.

where the term in brackets in the last line is equal to μ^i. A look at (5.3.11) reveals that the expression in the last but one line is equal to $\int_0^t d_u^i S_u^i du$. Therefore, plugging in $\hat{\vartheta}' = d'\mu^{-1}\mathcal{S}^{-1}$ we obtain

$$\langle \Gamma, M^i \rangle_t = \int_0^t d_u^i S_u^i du - \int_0^t \hat{\vartheta}' \mathcal{S} S^i \mu^i du = \int_0^t d_u^i S_u^i du - \int_0^t d' \mu^{-1} \mu^i S^i du = 0$$

because $\mu^{-1}\mu^i \in \mathbb{R}^n$ is a vector with 1 at the i-th place and 0 otherwise. Thus Γ is orthogonal to M^i for $i = 1, \dots, n$. □

The economic interpretation of what we have done in Theorem 5.8 is provided by the following corollary whose most important statement is that we have found a risk-minimizing hedging strategy $\hat{\vartheta}$ in a rather explicit form. The process Γ is interpreted as the cost process of the hedging strategy.

Corollary 5.9. *Given all the objects and conditions in Theorem 5.8, the measure $\hat{\mathbf{P}}$ is a minimal martingale measure, and $(\hat{C}, \hat{\vartheta})$ with*

$$\hat{C} := E^{\hat{\mathbf{P}}}[H] \tag{5.3.26}$$

is the unique risk-minimizing hedging strategy for the contingent claim H in the sense of Definition 5.6.

Proof. In Theorem 5.8 we have constructed the cost process Γ such that:

$$V(T, S_T) = H = \int_0^T \hat{\vartheta}_u' dS_u + \Gamma_T = \Gamma_0 + \int_0^T \hat{\vartheta}_u' dS_u + (\Gamma_T - \Gamma_0) \tag{5.3.27}$$

and $\Gamma - \Gamma_0$ is a real-valued square-integrable $\mathbf{P}$-martingale starting at zero which is orthogonal to all components of the martingale part of S. As S satisfies the structure condition (SC), the assumptions of Choulli et al. (1998), Proposition 3.7, are fulfilled, and Choulli et al. (1998), Theorem 5.5, guarantees the uniqueness (and existence) of the Föllmer-Schweizer decomposition of H. This shows that it must be given by (5.3.27) and that $\hat{\vartheta}$ is unique.
In order to verify that $\hat{\mathbf{P}}$ is a minimal martingale measure we must show that the $\mathbf{P}$-martingale Γ is a $\hat{\mathbf{P}}$-martingale as well. Moreover, for the calculation of $\hat{C}$ we prove that $G_t(\hat{\vartheta}) := \int_0^t \hat{\vartheta}_u' dS_u$ is a $\hat{\mathbf{P}}$-martingale.
$G(\hat{\vartheta})$ is a local $\hat{\mathbf{P}}$ martingale, i.e. $G(\hat{\vartheta})\hat{Z}$ is a local $\mathbf{P}$ martingale[12], because S is a $\hat{\mathbf{P}}$-martingale, and it is square-integrable for every $\vartheta \in \Theta$ by definition of Θ. The process $\Gamma\hat{Z}$ is a local $\mathbf{P}$-martingale as well[13] because Γ is orthogonal to $\hat{Z}$ due to $\langle \Gamma, M \rangle = 0$. Following Doob's inequality $\sup_{0 \le t \le T} |\Gamma_t|$ and $\sup_{0 \le t \le T} |G_t(\hat{\vartheta})|$ are square-integrable with respect to $\mathbf{P}$. Due to Schwarz' inequality and Protter (1995),

[12] See Jacod and Shiryaev (2003), III.3.8.
[13] See the remark after Definition 5.1.

Theorem I.47,

$$\begin{aligned} E^P\left[\sup_{0\le t\le T}|\Gamma_t\hat{Z}_t|\right] &\le E^P\left[\sup_{0\le t\le T}|\Gamma_t|\sup_{0\le t\le T}|\hat{Z}_t|\right] \\ &\le \left(E^P\left[\sup_{0\le t\le T}|\Gamma_t|\right]^2\right)^{\frac{1}{2}}\left(E^P\left[\sup_{0\le t\le T}|\hat{Z}_t|\right]^2\right)^{\frac{1}{2}} < \infty \end{aligned}$$

$\Gamma\hat{Z}$ is a $\mathbf{P}$-martingale, hence Γ is a $\hat{\mathbf{P}}$-martingale. Likewise, one proves that $G(\hat{\vartheta})$ is a $\hat{\mathbf{P}}$-martingale. Then taking the expectation with respect to $\hat{\mathbf{P}}$ of (5.3.27) yields

$$\hat{C} = \Gamma_0 = V(0, S_0) = E^{\hat{\mathbf{P}}}[V(T, S_T)] = E^{\hat{\mathbf{P}}}[H].$$

Hence $(\hat{C}, \hat{\vartheta})$ is a risk-minimizing portfolio strategy. □

5.3.2 Hedging in the KTD model

For more explicit results we go back to the case $n = 2$ of the KTD model of Section 4.3.3. In this case we can even give a simple sufficient condition for the non-singularity of μ, namely $c^{11}, c^{22} > 0$ and $\varrho \neq 0$ where ϱ is the covariance matrix of the Gaussian part of S. In fact, we have for the determinant D of μ

$$\begin{aligned} D := \det(\mu) &= \mu^{11}\mu^{22} - (\mu^{12})^2 \\ &= \left(c^{11} + \int_{\mathbb{R}^2}(e^{x^1}-1)^2K(dx)\right)\left(c^{22} + \int_{\mathbb{R}^2}(e^{x^2}-1)^2K(dx)\right) \\ &\quad - \left(c^{12} + \int_{\mathbb{R}^2}(e^{x^1}-1)(e^{x^2}-1)K(dx)\right)^2 \\ &= C_1 + C_2 + C_3 \end{aligned}$$

where

$$\begin{aligned} C_1 &:= c^{11}c^{22} - \varrho^2 c^{11}c^{22} \\ C_2 &:= \int_{\mathbb{R}^2}(e^{x^1}-1)^2K(dx)\int_{\mathbb{R}^2}(e^{x^2}-1)^2K(dx) - \left(\int_{\mathbb{R}^2}(e^{x^1}-1)(e^{x^2}-1)K(dx)\right)^2 \\ C_3 &:= c^{11}\int_{\mathbb{R}^2}(e^{x^2}-1)^2K(dx) + c^{22}\int_{\mathbb{R}^2}(e^{x^1}-1)^2K(dx) \\ &\quad -2\varrho\sqrt{c^{11}}\sqrt{c^{22}}\int_{\mathbb{R}^2}(e^{x^1}-1)(e^{x^2}-1)K(dx). \end{aligned}$$

By assumption we have $C_1 > 0$, and $C_2 \geq 0$ due to Schwarz' inequality. Finally we have $C_3 \geq 0$ as a consequence of the second binomial inequality. Hence we have

$D > 0$ and thus the non-singularity of Σ. This means that for the KTD model the structure condition (SC) is always fulfilled.
The solution to (5.3.16) which becomes $S\mu S\hat{\lambda} = S\alpha$ is then

$$\hat{\lambda}_1 = \frac{1}{S^1_- D}\left(\mu^{22}\alpha^1 - \mu^{12}\alpha^2\right) \quad \text{and} \quad \hat{\lambda}_2 = \frac{1}{S^2_- D}\left(\mu^{11}\alpha^2 - \mu^{12}\alpha^1\right).$$

From representation (5.3.17) we obtain the measure change process

$$\mathcal{E}\left(\hat{\beta}^1 W^1_\cdot + \hat{\beta}^2 W^2_\cdot + [\hat{y}(x_1,x_2) - 1] * (\mu^X - \nu)_\cdot\right)_t$$

with

$$\begin{aligned}
\hat{\beta}^1 &= D^{-1}\left\{\bar{c}_1\left(\mu^{22}\alpha^1 - \mu^{12}\alpha^2\right) + \bar{c}_{12}\left(\mu^{11}\alpha^2 - \mu^{12}\alpha^1\right)\right\}, \\
\hat{\beta}^2 &= D^{-1}\left\{\bar{c}_2\left(\mu^{11}\alpha^2 - \mu^{12}\alpha^1\right) + \bar{c}_{12}\left(\mu^{22}\alpha^1 - \mu^{12}\alpha^2\right)\right\}
\end{aligned}$$

and

$$\hat{y}(x_1,x_2) = 1 + D^{-1}\left\{\left(\mu^{22}\alpha^1 - \mu^{12}\alpha^2\right)(e^{x_1} - 1) + \left(\mu^{11}\alpha^2 - \mu^{12}\alpha^1\right)(e^{x_2} - 1)\right\}$$

Quite similarly, we obtain the optimal hedge ratio $\hat{\vartheta} = (\hat{\vartheta}_1, \hat{\vartheta}_2)$

$$\hat{\vartheta}_1 = \frac{1}{S^1_- D}\left(\mu^{22}d^1 - \mu^{12}d^2\right) \quad \text{and} \quad \hat{\vartheta}_2 = \frac{1}{S^2_- D}\left(\mu^{11}d^2 - \mu^{12}d^1\right).$$

Now we can stop to reflect a bit on the economic implications of a risk-minimizing hedging-strategy. Going briefly back to the one-dimensional case, we note that the optimal hedging strategy was given by

$$\hat{\vartheta}_t = \frac{c\phi^c S_{t-} + \int_{\mathbb{R}}(e^x - 1)\phi^d(x)K(dx)}{S_{t-}\left(c + \int_{\mathbb{R}}(e^x - 1)^2 K(dx)\right)}.$$

The strategy $\hat{\vartheta}$ prescribes some sort of weighted average between hedging the diffusion part and the jump part of S. Setting the Lévy measure $K \equiv 0$ results in the common delta-hedging approach. But this approach would be suboptimal in the sense of risk minimization if it were maintained in the case with jumps[14].
In the two-dimensional problem the off-diagonal elements Σ^{12} of Σ are equal to

$$\left(c_{12} + \int_{\mathbb{R}^2}(e^{x_1} - 1)(e^{x_2} - 1)K(dx_1, dx_2)\right) S^1 S^2. \tag{5.3.28}$$

This is zero if S^1 and S^2 are independent processes: Then $\varrho = 0$, hence $c_{12} = 0$, and the integral in (5.3.28) is also zero. This is because according to Proposition 4.6, which says how to integrate a KTD Lévy measure, this integral is reduced to the sum

[14]See also Cont and Tankov (2004b), p. 337.

of two one-dimensional integrals with respect to the two marginal Lévy measures. But with $f(x_1, x_2) = (e^{x_1} - 1)(e^{x_2} - 1)$ we have $f(x_1, 0) = f(0, x_2) = 0$.
However, these conditions are not necessary: We might just as well imagine a very special situation where the two terms in (5.3.28) sum up to zero because both dependence structures go in different directions. In this case we would have a dependent two-dimensional process but the hedge ratios would not depend on one another. The reason for this strange behaviour is that risk minimization is a quadratic criterion which just deals with the second moments, i.e. with variances and covariances. Therefore (5.3.28), which is some kind of covariance, can stipulate an 'independent' hedging scheme when in fact S^1 and S^2 are dependent due to some higher moment dependence.

5.4 Concluding remarks

There are two remarks which should be added to the above considerations: Beginning with the one-dimensional case of Section 5.2, the optimal hedge parameter $\hat{\vartheta}$ of (5.2.8) is not yet explicit enough to obtain a numerical value. More precisely, the quantities ϕ^c and $\phi^d(x)$ in (5.2.5) involve the price functional $V(\cdot, S)$ of the option in the form of its first derivative and the integral of $V(\cdot, Se^x) - V(\cdot, S)$ with respect to the Lévy measure. The solution lies again in the field of Fourier inversion: As $V(\cdot, S)$ can be represented as a Fourier integral[15] the hedge parameter $\hat{\vartheta}$ can also be written as a Fourier integral if the technical conditions allow for an interchange of the order of the necessary integrations. Looking a bit closer at this Fourier integral, the integrand has an analytic form in terms of the cumulant function of the used Lévy process. This fact stresses again the importance of having nicely computable expression for the characteristic functions of risk-neutral distributions. In the case where we start with the statistical probability measure, the martingale preserving property of the minimal martingale measure is of great benefit in this respect.
The multidimensional case is more difficult. Calculating (5.3.18) by the same method as in the one-dimensional case amounts to finding a Fourier transform representation of the price process $V(\cdot, S)$ of the option. But here we are back at Chapter 4, Section 4.4.1, where we argued that such a representation rarely exists. Using our moment matching approach in the framework of the KTD model in this chapter is not feasible because our variable Z_T which is to be approximated contains the current asset prices in quite a complicated way; so differentiation is difficult. An additional reason is that our method approximates the price, but does not necessarily provide a good estimation of its derivative.
The second remark concerns extensions of risk-minimizing hedging. As this concept says how to hedge an option using only the underlying variable and the riskless bond, it is meant to improve on the usual delta hedging approach. A number of authors[16] have established connections between the gamma of an option and the hedging error

[15] See Section 1.6.
[16] See Černý (2004b), paragraph 12.6.

in a risk-minimizing strategy. It would surely be a very interesting approach to add options to the portfolio and to formulate a modified delta-gamma hedging approach, where delta is now given by a risk-minimizing strategy instead of the Black-Scholes delta.

Final remarks

As the main part of the option pricing literature since the 1970s this thesis is concerned with an extension of the method launched by Black and Scholes (1973) and Merton (1973). In doing this, it should have conveyed a threefold basic message:
First, arguing for the use of Lévy processes in finance means essentially to say that describing an evidently non-Gaussian distribution by means of its first four moments is preferable to adjusting only for the first two. Of course, this is a trivial but nevertheless important statistical statement. But in the same way as the fact that chopping down a tree is better done with a hatchet than a knife does not lead people to use the latter to this end, this statement should not be used to dismiss Lévy processes as irrelevant. On the contrary, it should be a strong case for the further use of Lévy processes in a field where seemingly small differences in results have a big impact.
Second, as we observe consistently in science, the advantages of models incorporating new empirical findings must be weighed up with the experience that they are almost surely less tractable than older models. But tractability matters in particular in option pricing theory where the speed of the applicable pricing procedures is of overwhelming importance. Providing tractable Lévy models was exactly the main focus of this thesis where we defined implicitly tractability by the requirement of being easily used for fast option pricing by Fourier techniques. Tractability is a feature of a concrete model rather than of a class of models. That is why most objects are defined in view of how well they go together to result in a tractable model: the NIG Lévy process with the flexible change of measure in Chapter 2, the tempered stable Lévy process with the linex measure change functions in Chapter 3 as well as Kou's model with the Lévy copula K_{θ_+,θ_-} in Chapter 4.
The third message is a quite pragmatic view of the incompleteness issue. Lévy markets are generically incomplete, and for a large class of Lévy processes adding options to the market does not complete the market unless the number of options is infinite. Hence we cannot assign unique prices to options. However, an incomplete Lévy model provides a large number of degrees of freedom to be fitted to the prices of a finite number of options, and by means of a well-posed calibration procedure we are able to obtain a single risk-neutral distribution of the returns, although the market remains of course incomplete in theory. This degree of freedom argument is important to account for the phenomenon that risk-neutral distributions change over time without a corresponding change of the statistical distribution.

Appendix A

A class of tractable martingale measures

A.1 Log-likelihood function

The log-likelihood function L of n independent observations x_i from a NIG distributed sample is

$$\begin{aligned} L = \; & n \log \frac{\delta\alpha}{\pi} + n(\delta\sqrt{\alpha^2-\beta^2} - \beta\mu) + \beta \sum_{i=1}^{n} x_i \\ & + \sum_{i=1}^{n} \left(\log K_1(\alpha\sqrt{\delta^2 + (x_i-\mu)^2}) - \frac{1}{2}\log(\delta^2 + (x_i-\mu)^2) \right). \end{aligned}$$

K_ν denotes the modified Bessel function of the third kind of the order ν, and with the convenient abbreviation $R(\cdot) \equiv \frac{K_0(\cdot)}{K_1(\cdot)}$ we get

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \frac{n\delta\alpha}{\sqrt{\alpha^2-\beta^2}} - \sum_{i=1}^{n} \sqrt{\delta^2+(x_i-\mu)^2} R(\alpha\sqrt{\delta^2+(x_i-\mu)^2}) \\ \frac{\partial L}{\partial \beta} &= \sum_{i=1}^{n} x_i - n\left(\mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}\right) \\ \frac{\partial L}{\partial \delta} &= \frac{n}{\delta} + n\sqrt{\alpha^2-\beta^2} - \sum_{i=1}^{n} \left(\frac{2\delta}{\delta^2+(x_i-\mu)^2} + \frac{\alpha\delta R(\alpha\sqrt{\delta^2+(x_i-\mu)^2})}{\sqrt{\delta^2+(x_i-\mu)^2}} \right) \\ \frac{\partial L}{\partial \mu} &= -n\beta + \sum_{i=1}^{n} \frac{x_i-\mu}{\sqrt{\delta^2+(x_i-\mu)^2}} \\ & \quad \times \left(\frac{2}{\sqrt{\delta^2+(x_i-\mu)^2}} + \alpha R(\alpha\sqrt{\delta^2+(x_i-\mu)^2}) \right). \end{aligned}$$

A.2 Reparametrization of the NIG distribution

A more convenient parametrization can be given in terms of more intuitive parameters, namely mean m, standard deviation σ and two parameters χ and ξ representing skewness and kurtosis. Given an NIG distribution with parameters $(\alpha, \beta, \delta, \mu)$ we define[1]

$$\zeta := \delta\sqrt{\alpha^2 - \beta^2} \quad \text{and} \quad \varrho := \beta/\alpha \tag{A.2.1}$$

as well as

$$\xi := (1+\zeta)^{-1/2} \quad \text{and} \quad \chi := \xi\varrho.$$

The important thing about these parameters is that they are scale and location invariant. The standard deviation σ is given by[2]

$$\sigma^2 = \frac{\delta\alpha^2}{(\sqrt{\alpha^2 - \beta^2})^3} \tag{A.2.2}$$

Using (A.2.2) and the definition of ζ in (A.2.1) we have (after eliminating β) two (non-linear) equations in α and δ, where the solution is unique given the sign restrictions on both variables. Finally we have for the parameter transformation $(\xi, \chi, \sigma, m) \to (\alpha, \beta, \delta, \mu)$

$$\begin{aligned}
\alpha &= \frac{\sqrt{\zeta}}{\sigma(1-\varrho^2)} \\
\beta &= \frac{\sqrt{\zeta}\varrho}{\sigma(1-\varrho^2)} \\
\delta &= \sigma\sqrt{\zeta(1-\varrho^2)} \\
\mu &= m - \sigma\varrho\sqrt{\zeta}
\end{aligned} \tag{A.2.3}$$

where $\zeta = \xi^{-2} - 1$ and $\varrho = \chi/\xi$.
Vice versa we obtain for $(\alpha, \beta, \delta, \mu) \to (\xi, \chi, \sigma, m)$

$$\xi = (1 + \delta\sqrt{\alpha^2 - \beta^2})^{-1/2} \tag{A.2.4}$$

$$\chi = \frac{\beta}{\alpha}(1 + \delta\sqrt{\alpha^2 - \beta^2})^{-1/2} \tag{A.2.5}$$

$$\sigma = \sqrt{\frac{\delta\alpha^2}{(\sqrt{\alpha^2 - \beta^2})^3}} \tag{A.2.6}$$

$$m = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}}. \tag{A.2.7}$$

[1] See Prause(1999).
[2] See Rydberg (1997).

Appendix B

Moment-matching of the change of measure

B.1 The functions A_+, A_-, B_+, and B_-

Depending on the range of ν we have in each case two different representations. For $b > a > 0$, $\theta < \lambda_+ - 1$ and $p - \nu < 0$ with $p \in \mathbb{N} \cup \{0\}$ we have

$$
\begin{aligned}
A_+[a, b, \theta, p] &= \int_a^b (e^x - 1) x^p e^{\theta x} x^{-1-\nu} e^{-\lambda_+ x} dx \\
&= \int_a^b x^{-1+(p-\nu)} e^{-(\lambda_+ - 1 - \theta)x} dx - \int_a^b x^{-1+(p-\nu)} e^{-(\lambda_+ - \theta)x} dx \\
&= (\lambda_+ - 1 - \theta)^{\nu - p} \int_{(\lambda_+ - 1 - \theta)a}^{(\lambda_+ - 1 - \theta)b} x^{-1+(p-\nu)} e^{-x} dx \\
&\quad -(\lambda_+ - \theta)^{\nu - p} \int_{(\lambda_+ - \theta)a}^{(\lambda_+ - \theta)b} x^{-1+(p-\nu)} e^{-x} dx \\
&= (\lambda_+ - 1 - \theta)^{\nu - p} (\Gamma[p - \nu, (\lambda_+ - 1 - \theta)a] - \Gamma[p - \nu, (\lambda_+ - 1 - \theta)b]) \\
&\quad -(\lambda_+ - \theta)^{\nu - p} (\Gamma[p - \nu, (\lambda_+ - \theta)a] - \Gamma[p - \nu, (\lambda_+ - \theta)b]).
\end{aligned}
$$

Given the range of ν the only possibility to violate the inequality $p - \nu < 0$ is the case $p = 0$ and $0 < \nu < 1$ In this case partial integration gives

$$
\begin{aligned}
A_+[a, b, \theta, 0] &= \int_a^b (e^x - 1) e^{\theta x} x^{-1-\nu} e^{-\lambda_+ x} dx \\
&= \int_a^b (e^{-(\lambda_+ - 1 - \theta)x} - e^{-(\lambda_+ - \theta)x}) x^{-1-\nu} dx \\
&= \frac{-1}{\nu} (e^{-(\lambda_+ - 1 - \theta)x} - e^{-(\lambda_+ - \theta)x}) x^{-\nu} \Big|_a^b \\
&\quad -\frac{1}{\nu} \int_a^b \left((\lambda_+ - 1 - \theta) e^{-(\lambda_+ - 1 - \theta)x} - (\lambda_+ - \theta) e^{-(\lambda_+ - \theta)x} \right) x^{-\nu} dx
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{-1}{\nu}\left((e^{-(\lambda_+-1-\theta)b} - e^{-(\lambda_+-\theta)b})b^{-\nu} - (e^{-(\lambda_+-1-\theta)a} - e^{-(\lambda_+-\theta)a})a^{-\nu}\right) \\
&\quad -\frac{(\lambda_+-1-\theta)^\nu}{\nu}\int_{(\lambda_+-1-\theta)a}^{(\lambda_+-1-\theta)b} x^{-1+(1-\nu)}e^{-x}dx \\
&\quad +\frac{(\lambda_+-\theta)^\nu}{\nu}\int_{(\lambda_+-\theta)a}^{(\lambda_+-\theta)b} x^{-1+(1-\nu)}e^{-x}dx \\
&= \frac{-1}{\nu}\left[(e^{-(\lambda_+-1-\theta)b} - e^{-(\lambda_+-\theta)b})b^{-\nu} - (e^{-(\lambda_+-1-\theta)a} - e^{-(\lambda_+-\theta)a})a^{-\nu}\right] \\
&\quad -\frac{(\lambda_+-1-\theta)^\nu}{\nu}(\Gamma[1-\nu,(\lambda_+-1-\theta)a] - \Gamma[1-\nu,(\lambda_+-1-\theta)b)] \\
&\quad +\frac{(\lambda_+-\theta)^\nu}{\nu}(\Gamma[1-\nu,(\lambda_+-\theta)a] - \Gamma[1-\nu,(\lambda_+-\theta)b]).
\end{aligned}
$$

Integration over subintervals of the negative real line is captured by the function A_-. For $0 > b > a$, $\theta > -\lambda_- - 1$ and $p - \nu < 0$ with $p \in \mathbb{N} \cup \{0\}$ we can make use of the previous computations and obtain

$$
\begin{aligned}
A_-[a,b,\theta,p] &= \int_a^b (e^x-1)x^p e^{\theta x}|x|^{-1-\nu}e^{-\lambda_-|x|}dx \\
&= (-1)^p\int_{-b}^{-a}(e^{-x}-1)x^p e^{-\theta x}x^{-1-\nu}e^{-\lambda_- x}dx \\
&= (-1)^p\int_{-b}^{-a}(e^{-(\lambda_-+1+\theta)x} - e^{-(\lambda_-+\theta)x})x^{-1+(p-\nu)}dx \\
&= (-1)^p(\lambda_-+1+\theta)^{\nu-p}(\Gamma[p-\nu,-(\lambda_-+1+\theta)b] - \Gamma[p-\nu,-(\lambda_-+1+\theta)a]) \\
&\quad -(-1)^p(\lambda_-+\theta)^{\nu-p}(\Gamma[p-\nu,-(\lambda_-+\theta)b] - \Gamma[p-\nu,-(\lambda_-+\theta)a]).
\end{aligned}
$$

Similarly, the case $p = 0$ and $0 < \nu < 1$ is covered by

$$
\begin{aligned}
A_-[a,b,\theta,0] &= \int_a^b (e^x-1)e^{\theta x}|x|^{-1-\nu}e^{-\lambda_-|x|}dx \\
&= \int_{-b}^{-a}(e^{-x}-1)e^{-\theta x}x^{-1-\nu}e^{-\lambda_- x}dx \\
&= \frac{-1}{\nu}\left((e^{(\lambda_-+1+\theta)a} - e^{(\lambda_-+\theta)a})(-a)^{-\nu} - (e^{(\lambda_-+1+\theta)b} - e^{(\lambda_-+\theta)b})(-b)^{-\nu}\right) \\
&\quad -\frac{(\lambda_-+1+\theta)^\nu}{\nu}(\Gamma[1-\nu,-(\lambda_-+1+\theta)b] - \Gamma[1-\nu,-(\lambda_-+1+\theta)a)] \\
&\quad +\frac{(\lambda_-+\theta)^\nu}{\nu}(\Gamma[1-\nu,-(\lambda_-+\theta)b] - \Gamma[1-\nu,-(\lambda_-+\theta)a]).
\end{aligned}
$$

For $b > a > 0$, $\theta < \lambda_+$ and $\nu < 0$ the functions B_- and B_+ are given by

$$\begin{aligned} B_+[a,b,\theta,p] &= \int_a^b x^p e^{\theta x} x^{-1-\nu} e^{-\lambda_+ x} dx \\ &= (\lambda_+ - \theta)^{\nu - p} \int_{(\lambda_+-\theta)a}^{(\lambda_+-\theta)b} x^{-1+(p-\nu)} e^{-x} dx \\ &= (\lambda_+ - \theta)^{\nu-p}(\Gamma[p-\nu,(\lambda_+-\theta)a] - \Gamma[p-\nu,(\lambda_+-\theta)b]), \end{aligned}$$

and for $0 < \nu < 1$ this expression results in

$$\begin{aligned} B_+[a,b,\theta,0] &= \int_a^b x^{-1-\nu} e^{-(\lambda_+-\theta)x} dx \\ &= \frac{-1}{\nu} x^{-\nu} e^{-(\lambda_+-\theta)x}\Big|_a^b - \frac{\lambda_+-\theta}{\nu}\int_a^b x^{-\nu} e^{-(\lambda_+-\theta)x} dx \\ &= \frac{1}{\nu}(a^{-\nu}e^{-(\lambda_+-\theta)a} - b^{-\nu}e^{-(\lambda_+-\theta)b}) - \frac{(\lambda_+-\theta)^{\nu}}{\nu}\int_{(\lambda_+-\theta)a}^{(\lambda_+-\theta)b} x^{-1+(1-\nu)} e^{-x} dx \\ &= \frac{1}{\nu}(a^{-\nu}e^{-(\lambda_+-\theta)a} - b^{-\nu}e^{-(\lambda_+-\theta)b}) \\ &\quad -\frac{(\lambda_+-\theta)^{\nu}}{\nu}(\Gamma[1-\nu,(\lambda_+-\theta)a] - \Gamma[1-\nu,(\lambda_+-\theta)b]) \end{aligned}$$

The negative branch of the real axis is covered by

$$\begin{aligned} B_-[a,b,\theta,p] &= \int_a^b x^p e^{\theta x}|x|^{-1-\nu} e^{-\lambda_-|x|} dx = (-1)^p \int_{-b}^{-a} x^{-1+(p-\nu)} e^{-(\lambda_-+\theta)x} dx \\ &= (-1)^p(\lambda_-+\theta)^{\nu-p}(\Gamma[p-\nu,-(\lambda_-+\theta)b] - \Gamma[p-\nu,-(\lambda_-+\theta)a]) \end{aligned}$$

for $0 > b > a$, $\theta > -\lambda_-$ and $\nu < 0$. The exceptionel case is

$$\begin{aligned} B_-[a,b,\theta,0] &= \int_{-b}^{-a} x^{-1-\nu} e^{-(\lambda_-+\theta)x} dx \\ &= \frac{-1}{\nu} x^{-\nu} e^{-(\lambda_-+\theta)x}\Big|_{-b}^{-a} - \frac{\lambda_-+\theta}{\nu}\int_{-b}^{-a} x^{-\nu} e^{-(\lambda_-+\theta)x} dx \\ &= \frac{1}{\nu}(e^{-(\lambda_--\theta)a}a^{-\nu} - e^{-(\lambda_--\theta)b}b^{-\nu}) \\ &\quad -\frac{(\lambda_-+\theta)^{\nu}}{\nu}\int_{-(\lambda_-+\theta)b}^{-(\lambda_-+\theta)a} x^{-1+(1-\nu)} e^{-x} dx \\ &= \frac{1}{\nu}(e^{-(\lambda_--\theta)a}a^{-\nu} - e^{-(\lambda_--\theta)b}b^{-\nu}) \\ &\quad -\frac{(\lambda_-+\theta)^{\nu}}{\nu}(\Gamma[1-\nu,-(\lambda_-+\theta)b] - \Gamma[1-\nu,-(\lambda_-+\theta)a]). \end{aligned}$$

The variables a and b can become zero of infinity. In these cases we freely insert the symbols for zero and infinity into the functions $A_+, A_-, B_+, and B_-$ where we actually mean the respective limits. The values for the limits are easily obtained by noting that the incomplete gamma functions turn either to zero or to the ordinary gamma function and by $\lim_{x\to 0(\infty)}(e^{cx} - e^{c'x})/x^{\nu} = 0$ for $c > 0, c' \geq 0, c \neq c'$ and $0 < \nu < 1$, which can easily be derived by l'Hôpital's rule.

Appendix C

Non-linear dependence and option pricing

C.1 Proof of Proposition 4.6

Assume $\theta > U_2(0+)/U_1(0+)$ such that also $\theta > U_1(0+)/U_2(0+)$. Define x^*, y^* by $x^* = \sigma_a(0)$ and $y^* = \sigma_b(0)$, i.e.

$$\theta U_1(x^*) = U_2(0+) \quad \text{and} \quad \theta U_2(y^*) = U_1(0+)$$

where, given the assumptions, x^* and y^* are both unique and positive. Take $\bar{x} > x^*$ and $\bar{y} > y^*$ and choose $m, q \in \mathbb{N}$ sufficiently large such that the implicitly defined variables $\underline{x} > 0$ and $\underline{y} > 0$

$$\theta U_1(x^* + \frac{1}{m}(\bar{x} - x^*)) = U_2(\underline{y}) \quad \text{and} \quad \theta U_2(y^* + \frac{1}{q}(\bar{y} - y^*)) = U_1(\underline{x}) \tag{C.1.1}$$

are such that[1]

$$\theta > U_1(x)/U_2(y_1)\ \forall x > 0 \quad \text{and} \quad \theta > U_2(y)/U_1(x_1)\ \forall y > 0. \tag{C.1.2}$$

This is possible because

$$\theta > \frac{U_1(0+)}{U_2(0+)} > \frac{U_1(x)}{U_2(0+)} \quad \text{and} \quad \theta > \frac{U_2(0+)}{U_1(0+)} > \frac{U_2(y)}{U_1(0+)}$$

for arbitrary $x, y > 0$ and due to the continuity of U_1 and U_2. Now consider for arbitrary $n, p \in \mathbb{N}$ the points

$$x_i = \begin{cases} \underline{x} + \frac{i}{n}(x^* - \underline{x}), & , \quad i = 0, \ldots, n \\ x^* + \frac{i-n}{m}(\bar{x} - x^*), & , \quad i = n+1, \ldots, n+m \end{cases}$$

[1] Geometrically the conditions (C.1.2) mean that in approximating the integral on $\mathcal{D}_a$, we do not have to be concerned about $\mathcal{D}_b$ because $\mathcal{D}_b$ and the squares considered for the integration on $\mathcal{D}_a$ are disjoint. A corresponding statement applies to the integration on $\mathcal{D}_b$.

and

$$y_j = \begin{cases} \underline{y} + \frac{j}{p}(y^* - \underline{y}), & , \quad j = 0, \ldots, p \\ y^* + \frac{j-p}{q}(\bar{y} - y^*), & , \quad j = p+1, \ldots, p+q, \end{cases}$$

which partition the intervals $[\underline{x}, \bar{x}]$ and $[\underline{y}, \bar{y}]$ on both axes. Equations (C.1.1) and (C.1.2) imply that[2]

$$\left.\begin{array}{ll} \theta U_1(x) > U_2(\underline{y}) \ , & x < x_{n+1} \\ \theta U_1(x) < U_2(\underline{y}) \ , & x > x_{n+1} \end{array}\right\} \text{ and } \left\{\begin{array}{ll} \theta U_2(y) > U_1(\underline{x}) \ , & y < y_{p+1} \\ \theta U_2(y) < U_1(\underline{x}) \ , & y > y_{p+1} \end{array}\right. \tag{C.1.3}$$

as well as

$$\theta U_2(\underline{y}) > U_1(x) \ \forall x > 0 \quad \text{and} \quad \theta U_1(\underline{x}) > U_2(y) \ \forall y > 0. \tag{C.1.4}$$

Now define $\tilde{\mathcal{X}}_i = [x_i, x_{i+1}] \times [0, \underline{y}]$, $\tilde{\mathcal{X}} = \bigcup_{i=0}^{n+m} \tilde{\mathcal{X}}_i$, $\tilde{\mathcal{Y}}_j = [0, \underline{x}] \times [y_j, y_{j+1}]$ and $\tilde{\mathcal{Y}} = \bigcup_{j=0}^{p+q} \tilde{\mathcal{Y}}_j$. Choose arbitrary points $(\xi_i, \xi) \in \tilde{\mathcal{X}}_i$ and $(\eta, \eta_j) \in \tilde{\mathcal{Y}}_j$ and consider the sums

$$\begin{aligned} & \frac{1}{\theta+1} \sum_{i=0}^{n+m} f(\xi_i, \xi)[U_\theta(x_{i+1}, \underline{y}) - U_\theta(x_{i+1}, 0) - U_\theta(x_i, \underline{y}) + U_\theta(x_i, 0)] \\ = \ & \frac{1}{\theta+1} \sum_{i=0}^{n+m} f(\xi_i, \xi) \left[\min(\theta U_1(x_{i+1}), U_2(\underline{y})) + \min(U_1(x_{i+1}), \theta U_2(\underline{y}))\right. \\ & - \min(\theta U_1(x_i), U_2(\underline{y})) - \min(U_1(x_i), \theta U_2(\underline{y})) \\ & \left. -\theta U_1(x_{i+1}) - U_1(x_{i+1}) + \theta U_1(x_i) + U_1(x_i)\right] \\ = \ & \frac{-\theta}{\theta+1} \sum_{i=0}^{n} f(\xi_i, \xi)(U_1(x_{i+1}) - U_1(x_i)), \end{aligned}$$

where (C.1.3) and (C.1.4) have been used for the latter equality. Note that for $x > x^*$ all terms cancel out which is why the sum performs the integration only from $\underline{x}$ to x^*. Taking into account the continuity of f, this expression consequently converges for $n, m \to \infty$ towards

$$\frac{-\theta}{\theta+1} \int_{\underline{x}}^{x^*} f(x, 0) U_1(dx).$$

The same line of reasoning applies to the integration over $\tilde{\mathcal{Y}}$ yielding

$$\frac{-\theta}{\theta+1} \int_{\underline{y}}^{y^*} f(0, y) U_2(dy).$$

If we then take into account that $\underline{x} > 0$ and $\underline{y} > 0$ could easily have been chosen smaller they can be substituted by zero in the limit. Second, the mass of the axes above x^* and y^* does not play a role, hence $\bar{x}$ and $\bar{y}$ can be replaced by infinity. Finally, we have obtained the integrals over $\mathcal{X}$ and $\mathcal{Y}$ from zero to infinity.

[2] Note that $x_{n+1} = x^* + (\bar{x} - x^*)/m$ and $y_{p+1} = y^* + (\bar{y} - y^*)/q$ are the 'threshold' values from (C.1.1).

Now we are concerned with the integration over $\tilde{\mathcal{D}}_a = \bigcup_{j=0}^{p+q} \tilde{\mathcal{D}}_a^j$ and $\tilde{\mathcal{D}}_b = \bigcup_{i=0}^{n+m} \tilde{\mathcal{D}}_j^2$ where $\tilde{\mathcal{D}}_a^j = [x'_j, y_j] \times [x'_{j+1}, y_{j+1}]$ and $\tilde{\mathcal{D}}_b^i = [x_i, y'_i] \times [x_{i+1}, y'_{i+1}]$. The points x'_j and y'_i lie on $\mathcal{D}_a$ and $\mathcal{D}_b$, i.e. $x'_j = \sigma_a(y_j)$ and $y'_i = \sigma_b(x_i)$. The fact that there is no point of intersection between $\mathcal{D}_a$ and $\mathcal{D}_b$ allows to obtain for $(\xi_j, \eta_j) \in \tilde{\mathcal{D}}_a^j$

$$
\begin{aligned}
&\frac{1}{\theta+1}\sum_{j=0}^{p+q} f(\xi_j,\eta_j)[U_\theta(x'_{j+1},y_{j+1}) - U_\theta(x'_{j+1},y_j) - U_\theta(x'_j,y_{j+1}) + U_\theta(x'_j,y_j)] \\
= \; &\frac{1}{\theta+1}\sum_{j=0}^{p+q} f(\xi_j,\eta_j)\left[\min(\theta U_1(x'_{j+1}),U_2(y_{j+1})) + \min(U_1(x'_{j+1}),\theta U_2(y_{j+1}))\right. \\
&- \min(\theta U_1(x'_j),U_2(y_{j+1})) - \min(U_1(x'_j),\theta U_2(y_{j+1})) \\
&- \min(\theta U_1(x'_{j+1}),U_2(y_j)) - \min(U_1(x'_{j+1}),\theta U_2(y_j)) \\
&\left.+ \min(\theta U_1(x'_j),U_2(y_j)) + \min(U_1(x'_j),\theta U_2(y_j))\right]
\end{aligned}
$$

$$
\overset{p,q\to\infty}{\to} \frac{-1}{\theta+1}\int_{\underline{y}}^{\bar{y}} f(\sigma_a(y),y)dU_2(y).
$$

For $(\xi_i, \eta_i) \in \tilde{\mathcal{D}}_b^i$ we have the corresponding result

$$
\frac{-1}{\theta+1}\int_{\underline{x}}^{\bar{x}} f(x,\sigma_b(x))dU_1(x).
$$

For $\underline{x}, \underline{y} \to 0$ and $\bar{x}, \bar{y} \to \infty$ we obtain the integrals of f over $\mathcal{D}_1$ and $\mathcal{D}_2$.

Now we consider the second case where we have $\theta \leq U_2(0+)/U_1(0+)$ and $\theta \geq U_1(0+)/U_2(0+)$. The following procedure is parallel with the one in the first case. Define $y_a^* = \varsigma_a(0)$ and $y_b^* = \varsigma_b(0)$, i.e.

$$
\theta U_1(0+) = U_2(y_a^*) \quad \text{and} \quad U_1(0+) = \theta U_2(y_b^*).
$$

We have then $U_2(y_a^*) = \theta U_1(0+) = \theta^2 U_2(y_b^*) \geq U_2(y_b^*)$, i.e. $y_a^* \leq y_b^*$. Let us now suppose that $y_a^* < y_b^*$. Then there is $\underline{x} > 0$ such that $\theta U_1(\underline{x}) > U_2(y_b^*)$. Specifying $\bar{x} > \underline{x}$ and $\bar{y} > y_b^* > y_a^* > \underline{y}$ we may then use the grid given by the points

$$
x_i = \underline{x} + \frac{i}{n}(\bar{x} - \underline{x}), \quad i = 0, \dots, n,
$$

and

$$
y_j = \begin{cases} \underline{y} + \frac{j}{p}(y_a^* - \underline{y}), & , \quad j = 0, \dots, p, \\ y_a^* + \frac{j-p}{q}(y_b^* - y_a^*), & , \quad j = p+1, \dots, p+q, \\ y_b^* + \frac{j-p-q}{m}(\bar{y} - y_b^*, & , \quad j = p+q+1, \dots, p+q+m. \end{cases}
$$

Overlapping again $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{D}_a$ and $\mathcal{D}_b$ by some rectangles $\tilde{\mathcal{X}}$, $\tilde{\mathcal{Y}}$, $\tilde{\mathcal{D}}_2^i$ and $\tilde{\mathcal{D}}_2^i$ which are defined exactly the same way as in the consideration of the first case. Take again some points $(\xi_i, \xi) \in \tilde{\mathcal{X}}_i$ and $(\eta, \eta_j) \in \tilde{\mathcal{Y}}_j$. The integral of f over $\mathcal{X}$ is then the limit of the sum

$$\sum_{i=0}^{n} f(\xi_i, \xi) \left[\min(\theta U_1(x_{i+1}), U_2(\underline{y})) + \min(U_1(x_{i+1}), \theta U_2(\underline{y}))\right.$$
$$- \min(\theta U_1(x_i), U_2(\underline{y})) - \min(U_1(x_i), \theta U_2(\underline{y}))$$
$$\left.- \theta U_1(x_{i+1}) - U_1(x_{i+1}) + \theta U_1(x_i) + U_1(x_i)\right],$$

which turns out to be zero for every n. For the integration over $\mathcal{Y}$ we have the following relations:

$$\begin{array}{llll} \theta U_1(\underline{x}) < U_2(y) & \text{and} & U_1(\underline{x}) < \theta U_2(y) & \text{for} \quad \underline{y} < y < y_a^*, \\ \theta U_1(\underline{x}) > U_2(y) & \text{and} & U_1(\underline{x}) < \theta U_2(y) & \text{for} \quad y_a^* \le y \le y_b^*, \\ \theta U_1(\underline{x}) > U_2(y) & \text{and} & U_1(\underline{x}) > \theta U_2(y) & \text{for} \quad y_b^* < y < \bar{y}. \end{array}$$

This entails that

$$\frac{1}{\theta+1} \sum_{j=0}^{p+q+m} f(\eta, \eta_j) \left[\min(\theta U_1(\underline{x}), U_2(y_{j+1})) + \min(U_1(\underline{x}), \theta U_2(y_{j+1}))\right.$$
$$- \min(\theta U_1(\underline{x}), U_2(y_j)) - \min(U_1(\underline{x}), \theta U_2(y_j))$$
$$\left.-U_2(y_{j+1}) - \theta U_2(y_{j+1}) + U_2(y_j) + \theta U_2(y_j)\right]$$
$$= \quad -\sum_{j=0}^{p} f(\eta, \eta_j)[U_2(y_{j+1} - U_2(y_j)] - \frac{\theta}{\theta+1} \sum_{j=p}^{p+q} f(\eta, \eta_j)[U_2(y_{j+1} - U_2(y_j)],$$

which converges to

$$-\int_{\underline{y}}^{y_a^*} f(0, y) U_2(dy) - \frac{\theta}{\theta+1} \int_{y_a^*}^{y_b^*} f(0, y) U_2(dy).$$

Again, we can let θ run to zero and obtain the integral of f over $\mathcal{Y}$.
Integration over $\mathcal{D}_a$ and $\mathcal{D}_b$ is done in the same manner as in the first case if one takes into account that both integrations are in terms of x yielding for $(\xi_i^1, \eta_i^1) \in \tilde{\mathcal{D}}_a^i$ and $(\xi_i^2, \eta_i^2) \in \tilde{\mathcal{D}}_b^i$

$$\frac{-\theta}{\theta+1} \sum_{i=1}^{n} f(\xi_i^1, \eta_i^1)[U_1(x_{i+1} - U_1(x_i)]$$

for $\mathcal{D}_a$, and for $\mathcal{D}_b$ we have

$$\frac{-1}{\theta+1} \sum_{i=1}^{n} f(\xi_i^2, \eta_i^2)[U_1(x_{i+1} - U_1(x_i)].$$

Hence, these expressions converge to

$$\frac{-\theta}{\theta+1} \int_{\underline{x}}^{\bar{x}} f(x, \varsigma_a(x)) U_1(dx) \quad \text{resp.} \quad \frac{-1}{\theta+1} \int_{\underline{x}}^{\bar{x}} f(x, \varsigma_b(x)) U_1(dx).$$

as $n \to \infty$. The limit cases of $\theta \to 1$ and $\theta \to \frac{U_2(0+)}{U_1(0+)}$ can simply be calculated by plugging in the values of θ in the solution of the second case because of the weak convergence of K_θ as $\theta \to 1$ resp. $\theta \to \frac{U_2(0+)}{U_1(0+)}$. Weak convergence follows from Proposition 1.32. □

C.2 The integrals over K_{θ^+,θ^-}

To represent the integrals we need to distinguish some cases depending upon the relation between p_1, λ_1, p_2, and λ_2. This is a bit cumbersome because it is not recommendable to maintain the w.l.o.g. assumption of the main text.

The sets $\mathcal{D}_1^+$ and $\mathcal{D}_2^+$ are

$$\begin{aligned}\mathcal{D}_1^+ &= \{(x,y) \in \mathbb{R}_+^2 | \theta^+ p_1\lambda_1 e^{-\lambda_{1+}x} = p_2\lambda_2 e^{-\lambda_{2+}y}\} \quad \text{and}\\ \mathcal{D}_2^+ &= \{(x,y) \in \mathbb{R}_+^2 | p_1\lambda_1 e^{-\lambda_{1+}x} = \theta^+ p_2\lambda_2 e^{-\lambda_{2+}y}\}\end{aligned}$$

from which we have that

$$\begin{aligned}\sigma_a(y) &= \frac{\lambda_{2+}}{\lambda_{1+}}y + \frac{1}{\lambda_{1+}}\log\left(\theta^+\frac{p_1\lambda_1}{p_2\lambda_2}\right)\\ \varsigma_a(x) &= \frac{\lambda_{1+}}{\lambda_{2+}}x + \frac{1}{\lambda_{2+}}\log\left(\frac{1}{\theta^+}\frac{p_2\lambda_2}{p_1\lambda_1}\right)\\ \sigma_b(y) &= \frac{\lambda_{2+}}{\lambda_{1+}}y + \frac{1}{\lambda_{1+}}\log\left(\frac{1}{\theta^+}\frac{p_1\lambda_1}{p_2\lambda_2}\right)\\ \varsigma_b(x) &= \frac{\lambda_{1+}}{\lambda_{2+}}x + \frac{1}{\lambda_{2+}}\log\left(\theta^+\frac{p_2\lambda_2}{p_1\lambda_1}\right)\end{aligned}$$

The expression

$$\int_0^\infty \int_0^\infty e^{izx+iz'y} K_{\theta^+,\theta^-}(dx,dy)$$

is equal to
Case I: $\theta^+ > \frac{p_1\lambda_1}{p_2\lambda_2}$ and $\theta^+ > \frac{p_2\lambda_2}{p_1\lambda_1}$.

$$\begin{aligned}&= \frac{1}{\theta^+ + 1}\int_0^\infty e^{izx+iz'\varsigma_b(x)} K_1^+(dx) + \frac{1}{\theta^+ + 1}\int_0^\infty e^{iz\sigma_a(y)+iz'y} K_2^+(dy)\\ &\quad + \frac{\theta^+}{\theta^+ + 1}\int_0^{\sigma_a(0)} e^{izx} K_1^+(dx) + \frac{\theta^+}{\theta^+ + 1}\int_0^{\varsigma_b(0)} e^{iz'y} K_2^+(dy)\\ &= \frac{1}{\theta^+ + 1}\Big\{\Phi_b^\varsigma(\theta^+; z, z') + \Phi_a^\sigma(\theta^+; z, z')\\ &\quad + \frac{\theta^+ p_1\lambda_1\lambda_{1+}}{\lambda_{1+} - iz}[1 - e^{-(\lambda_{1+}-iz)\sigma_a(0)}] + \frac{\theta^+ p_2\lambda_2\lambda_{2+}}{\lambda_{2+} - iz'}[1 - e^{-(\lambda_{2+}-iz')\varsigma_b(0)}]\Big\}\end{aligned}$$

where

$$\begin{aligned}
\Phi_a^\sigma(\theta^+; z, z') &= \frac{p_2\lambda_2\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{2+} - iz'\lambda_{1+}} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\theta^+ \frac{p_1\lambda_1}{p_2\lambda_2}\right)\right], \\
\Phi_b^\varsigma(\theta^+; z, z') &= \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\theta^+ \frac{p_2\lambda_2}{p_1\lambda_1}\right)\right].
\end{aligned}$$

Case II: $\theta^+ > \frac{p_1\lambda_1}{p_2\lambda_2}$ and $\theta^+ < \frac{p_2\lambda_2}{p_1\lambda_1}$.

$$\begin{aligned}
&= \frac{\theta^+}{\theta^+ + 1} \int_0^\infty e^{izx + iz'\varsigma_a(x)} K_1^+(dx) + \frac{1}{\theta + 1} \int_0^\infty e^{izx + iz'\varsigma_b(x)} K_1^+(dx) \\
&\quad + \frac{\theta^+}{\theta^+ + 1} \int_{\varsigma_a(0)}^{\varsigma_b(0)} e^{iz'y} K_2^+(dy) + \int_0^{\varsigma_a(0)} e^{iz'y} dK_2^+(dy) \\
&= \frac{1}{\theta^+ + 1} \Bigg\{ \theta^+ \Phi_a^\varsigma(\theta^+; z, z') + \Phi_b^\varsigma(\theta^+; z, z') \\
&\quad \frac{\theta^+ p_2\lambda_2\lambda_{2+}}{\lambda_{2+} - iz'} [e^{-(\lambda_{2+} - iz')\varsigma_a(0)} - e^{-(\lambda_{2+} - iz')\varsigma_b(0)}] \Bigg\} + \frac{p_2\lambda_2\lambda_{2+}}{\lambda_{2+} - iz'} [1 - e^{-(\lambda_{2+} - iz')\varsigma_a(0)}]
\end{aligned}$$

where

$$\begin{aligned}
\Phi_a^\varsigma(\theta^+; z, z') &= \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\frac{1}{\theta^+} \frac{p_2\lambda_2}{p_1\lambda_1}\right)\right], \\
\Phi_b^\varsigma(\theta^+; z, z') &= \frac{p_1\lambda_1\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{1+} - iz'\lambda_{2+}} \exp\left[\frac{iz'}{\lambda_{2+}} \log\left(\theta^+ \frac{p_2\lambda_2}{p_1\lambda_1}\right)\right].
\end{aligned}$$

Case III: $\theta^+ < \frac{p_1\lambda_1}{p_2\lambda_2}$ and $\theta^+ > \frac{p_2\lambda_2}{p_1\lambda_1}$.

$$\begin{aligned}
&= \frac{\theta^+}{\theta^+ + 1} \int_0^\infty e^{iz\sigma_b(y) + iz'y} K_2^+(dy) + \frac{1}{\theta^+ + 1} \int_0^\infty e^{iz\sigma_a(y) + iz'y} K_2^+(dy) \\
&\quad \frac{\theta^+}{\theta^+ + 1} \int_{\sigma_b(0)}^{\sigma_a(0)} e^{izx} K_1^+(dx) + \int_0^{\sigma_b(0)} e^{izx} K_1^+(dx) \\
&= \frac{1}{\theta^+ + 1} \Bigg\{ \theta^+ \Phi_b^\sigma(\theta^+; z, z') + \Phi_a^\sigma(\theta^+; z, z') \\
&\quad \frac{\theta^+ p_1\lambda_1\lambda_{1+}}{\lambda_{1+} - iz} [e^{-(\lambda_{1+} - iz)\sigma_b(0)} - e^{-(\lambda_{1+} - iz)\sigma_a(0)}] \Bigg\} + \frac{p_1\lambda_1\lambda_{1+}}{\lambda_{1+} - iz} [1 - e^{-(\lambda_{1+} - iz)\sigma_b(0)}]
\end{aligned}$$

where

$$\begin{aligned}
\Phi_a^\sigma(\theta^+; z, z') &= \frac{p_2\lambda_2\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{2+} - iz'\lambda_{1+}} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\theta^+ \frac{p_1\lambda_1}{p_2\lambda_2}\right)\right], \\
\Phi_b^\sigma(\theta^+; z, z') &= \frac{p_2\lambda_2\lambda_{1+}\lambda_{2+}}{\lambda_{1+}\lambda_{2+} - iz\lambda_{2+} - iz'\lambda_{1+}} \exp\left[\frac{iz}{\lambda_{1+}} \log\left(\frac{1}{\theta^+} \frac{p_1\lambda_1}{p_2\lambda_2}\right)\right]
\end{aligned}$$

Likewise, the sets $\mathcal{D}_1^-$ and $\mathcal{D}_2^-$ are

$$
\begin{aligned}
-\mathcal{D}_1^- &= \{(x,y) \in \mathbb{R}_+^2 | \theta^-(1-p_1)\lambda_1 e^{-\lambda_{1-}x} = (1-p_2)\lambda_2 e^{-\lambda_{2-}y}\} \quad \text{and} \\
-\mathcal{D}_2^- &= \{(x,y) \in \mathbb{R}_+^2 | (1-p_1)\lambda_1 e^{-\lambda_{1-}x} = \theta^-(1-p_2)\lambda_2 e^{-\lambda_{2-}y}\}
\end{aligned}
$$

from which we have that

$$
\begin{aligned}
\sigma_a(y) &= \frac{\lambda_{2-}}{\lambda_{1-}} y + \frac{1}{\lambda_{1-}} \log\left(\theta^- \frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}\right) \\
\varsigma_a(x) &= \frac{\lambda_{1-}}{\lambda_{2-}} x + \frac{1}{\lambda_{2-}} \log\left(\frac{1}{\theta^-} \frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}\right) \\
\sigma_b(y) &= \frac{\lambda_{2-}}{\lambda_{1-}} y + \frac{1}{\lambda_{1-}} \log\left(\frac{1}{\theta^-} \frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}\right) \\
\varsigma_b(x) &= \frac{\lambda_{1-}}{\lambda_{2-}} x + \frac{1}{\lambda_{2-}} \log\left(\theta^- \frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}\right)
\end{aligned}
$$

We have for the expression

$$
\int_{-\infty}^0 \int_{-\infty}^0 e^{izx+iz'y} K_{\theta^+,\theta^-}(dx,dy) = \int_0^\infty \int_0^\infty e^{i(-z)x+i(-z')y} K'_{\theta^+,\theta^-}(dx,dy)
$$

Case I: $\theta^- > \frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}$ and $\theta^- > \frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}$.

$$
\begin{aligned}
= \; & \frac{1}{\theta^- + 1} \Bigg\{ \bar{\Phi}_b^\varsigma(\theta^-; z, z') + \bar{\Phi}_a^\sigma(\theta^-; z, z') \\
& + \frac{\theta^-(1-p_1)\lambda_1\lambda_{1-}}{\lambda_{1-} + iz}[1 - e^{-(\lambda_{1-}+iz)\sigma_a(0)}] + \frac{\theta^+(1-p_2)\lambda_2\lambda_{2-}}{\lambda_{2-} + iz'}[1 - e^{-(\lambda_{2-}+iz')\varsigma_b(0)}] \Bigg\}
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{\Phi}_a^\sigma(\theta^-; z, z') &= \frac{(1-p_2)\lambda_2\lambda_{1-}\lambda_{2-}}{\lambda_{1-}\lambda_{2-} + iz\lambda_{2-} + iz'\lambda_{1-}} \exp\left[-\frac{iz}{\lambda_{1-}} \log\left(\theta^- \frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}\right)\right], \\
\bar{\Phi}_b^\varsigma(\theta^-; z, z') &= \frac{(1-p_1)\lambda_1\lambda_{1-}\lambda_{2-}}{\lambda_{1-}\lambda_{2-} + iz\lambda_{1-} + iz'\lambda_{2-}} \exp\left[-\frac{iz'}{\lambda_{2-}} \log\left(\theta^- \frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}\right)\right].
\end{aligned}
$$

Case II: $\theta^- > \frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}$ and $\theta^- < \frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}$.

$$
\begin{aligned}
= \; & \frac{1}{\theta^- + 1} \Bigg\{ \theta^- \bar{\Phi}_a^\varsigma(\theta^-; z, z') + \bar{\Phi}_b^\varsigma(\theta^-; z, z') \\
& \frac{\theta^+(1-p_2)\lambda_2\lambda_{2-}}{\lambda_{2-} + iz'}[e^{-(\lambda_{2-}+iz')\varsigma_a(0)} - e^{-(\lambda_{2-}+iz')\varsigma_b(0)}] \Bigg\} + \\
& \frac{(1-p_2)\lambda_2\lambda_{2-}}{\lambda_{2-} + iz'}[1 - e^{-(\lambda_{2-}+iz')\varsigma_a(0)}]
\end{aligned}
$$

where

$$\bar{\Phi}_a^{\varsigma}(\theta^-; z, z') = \frac{(1-p_1)\lambda_1\lambda_{1-}\lambda_{2-}}{\lambda_{1-}\lambda_{2-} + iz\lambda_{1-} + iz'\lambda_{2-}} \exp\left[-\frac{iz'}{\lambda_{2-}} \log\left(\frac{1}{\theta^-}\frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}\right)\right],$$

$$\bar{\Phi}_b^{\varsigma}(\theta^-; z, z') = \frac{(1-p_1)\lambda_1\lambda_{1-}\lambda_{2-}}{\lambda_{1-}\lambda_{2-} + iz\lambda_{1-} + iz'\lambda_{2-}} \exp\left[-\frac{iz'}{\lambda_{2-}} \log\left(\theta^-\frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}\right)\right].$$

Case III: $\theta^- < \frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}$ and $\theta^- > \frac{(1-p_2)\lambda_2}{(1-p_1)\lambda_1}$.

$$= \frac{1}{\theta^- + 1}\left\{\theta^-\bar{\Phi}_b^{\sigma}(\theta^-; z, z') + \bar{\Phi}_a^{\sigma}(\theta^-; z, z')\right.$$
$$\left.\frac{\theta^-(1-p_1)\lambda_1\lambda_{1+}}{\lambda_{1+} + iz}[e^{-(\lambda_{1-}+iz)\sigma_b(0)} - e^{-(\lambda_{1+}+iz)\sigma_a(0)}]\right\} +$$
$$\frac{(1-p_1)\lambda_1\lambda_{1-}}{\lambda_{1-} + iz}[1 - e^{-(\lambda_{1-}+iz)\sigma_b(0)}]$$

where

$$\bar{\Phi}_b^{\sigma}(\theta^-; z, z') = \frac{(1-p_2)\lambda_2\lambda_{1-}\lambda_{2-}}{\lambda_{1-}\lambda_{2-} + iz\lambda_{2-} + iz'\lambda_{1-}} \exp\left[-\frac{iz}{\lambda_{1-}} \log\left(\frac{1}{\theta^-}\frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}\right)\right],$$

$$\bar{\Phi}_a^{\sigma}(\theta^-; z, z') = \frac{(1-p_2)\lambda_2\lambda_{1-}\lambda_{2-}}{\lambda_{1-}\lambda_{2-} + iz\lambda_{2-} + iz'\lambda_{1-}} \exp\left[-\frac{iz}{\lambda_{1-}} \log\left(\theta^-\frac{(1-p_1)\lambda_1}{(1-p_2)\lambda_2}\right)\right].$$

C.3 Proof of Lemma 4.13

Step 1: First we remark that for fixed x_2 $f_1^{1,0}$ satisfies the ordinary differential equation

$$f_1^{1,0} = \frac{\partial}{\partial x_1}\frac{\omega_1 e^{x_1}}{\sum_{i=1}^2 \omega_i e^{x_i}} = f_1 - (f_1)^2 = f_1(1-f_1).$$

Likewise we have $f_2^{0,1} = f_2(1-f_2)$. Equation (4.4.8) is obtained by

$$\begin{aligned} F^{s,0} = f_1^{s-1} &= \frac{\partial^{s-2}}{\partial x_1^{s-2}}(f_1(1-f_1)) \\ &= \sum_{j=0}^{s-2}\binom{s-2}{j} f_1^{j,0}\frac{\partial^{s-2-j}}{\partial x_1^{s-2-j}}(1-f_1) \\ &= f_1^{s-2}(1-f_1) - \sum_{j=0}^{s-3}\binom{s-2}{j} f_1^{j,0} f_1^{s-2-j,0}, \end{aligned}$$

and (4.4.9) is obtained in the same way.
Step 2: From the relation $\partial f_2/\partial x_1 = -f_2 f_1$ we obtain

$$\frac{\partial^s}{\partial x_1^s}\frac{\partial^t}{\partial x_2^t}F(x_1,x_2) = \frac{\partial^s}{\partial x_1^s}\frac{\partial^{t-1}}{\partial x_2^{t-1}}f_2 = \frac{\partial^{t-1}}{\partial x_2^{t-1}}\frac{\partial^{s-1}}{\partial x_1^{s-1}}\frac{\partial f_2}{\partial x_1} = -\frac{\partial^{t-1}}{\partial x_2^{t-1}}\frac{\partial^{s-1}}{\partial x_1^{s-1}}f_2 f_1.$$

Hence equation (4.4.10) in the following way:

$$\begin{aligned}
F^{s,t} &= f_2^{s,t-1} = \frac{\partial^s}{\partial x_1^s}\frac{\partial^{t-1}}{\partial x_2^{t-1}} f_2 \\
&= -\frac{\partial^{t-1}}{\partial x_2^{t-1}} \sum_{j=0}^{s-1} \binom{s-1}{j} f_2^{j,0} f_1^{s-1-j,0} \\
&= -\sum_{j=0}^{s-1} \binom{s-1}{j} \sum_{k=0}^{t-1} \binom{t-1}{k} f_2^{j,k} f_1^{s-1-j,t-1-k} \\
&= -\sum_{j=0}^{s-1}\sum_{k=0}^{t-1} \binom{s-1}{j}\binom{t-1}{k} f_2^{j,k} f_1^{s-1-j,t-1-k}.
\end{aligned}$$

Observing that $f_1^{s-1-j,t-1-k} = f_2^{s-j,t-2-k}$ leads to

$$\begin{aligned}
&= -\sum_{j=0}^{s-1}\sum_{k=0}^{t-2} \binom{s-1}{j}\binom{t-1}{k} f_2^{j,k} f_1^{s-1-j,t-1-k} - \sum_{j=0}^{s-1} \binom{s-1}{j} f_2^{j,t-1} f_1^{s-1-j,0} \\
&= -\sum_{j=0}^{s-1}\sum_{k=0}^{t-2} \binom{s-1}{j}\binom{t-1}{k} f_2^{j,k} f_2^{s-j,t-2-k} - \sum_{j=0}^{s-1} \binom{s-1}{j} f_2^{j,t-1} f_1^{s-1-j,0}.
\end{aligned}$$

Step 3: The proofs of (4.4.11) and (4.4.12) follow the same line of reasoning, namely

$$\begin{aligned}
F^{s,0,m} &= m\frac{\partial^{s-1}}{\partial x_1^{s-1}} \left\{ [F(x_1,x_2)]^{m-1} \frac{\partial}{\partial x_1} F(x_1,x_2) \right\} \\
&= m\sum_{j=0}^{s-1} \binom{s-1}{j} F^{j,0,m-1} F^{s-j,0}
\end{aligned}$$

and

$$\begin{aligned}
F^{s,t,m} &= m\frac{\partial^s}{\partial x_1^s}\frac{\partial^{t-1}}{\partial x_2^{t-1}} \left\{ [F(x_1,x_2)]^{m-1} \frac{\partial}{\partial x_2} F(x_1,x_2) \right\} \\
&= m\frac{\partial^s}{\partial x_1^s} \left\{ \sum_{j=0}^{t-1} \binom{t-1}{j} \frac{\partial^j}{\partial x_2^j} F(x_1,x_2)^{m-1} \frac{\partial^{t-j}}{\partial x_2^{t-j}} F(x_1,x_2) \right\} \\
&= m\sum_{j=0}^{t-1}\sum_{k=0}^{s} \binom{t-1}{j}\binom{s}{k} \frac{\partial^k}{\partial x_1^k}\frac{\partial^j}{\partial x_2^j} F(x_1,x_2)^{m-1} \frac{\partial^{s-k}}{\partial x_1^{s-k}}\frac{\partial^{t-j}}{\partial x_2^{t-j}} F(x_1,x_2) \\
&= m\sum_{j=0}^{t-1}\sum_{k=0}^{s} \binom{t-1}{j}\binom{s}{k} F^{k,j,m-1} F^{s-k,t-k}.
\end{aligned}$$

C.4 Monte-Carlo pricing with variance reduction in the diffusion model

We implement a variance reduction technique for basket option pricing in the pure diffusion model described in Pellizzari (1998). We briefly describe the idea behind this approach:
Given a European option with payoff $C_T = (\omega_1 S_T^X + \omega_2 S_T^Y - K)^+$ at time $t = T$ and strike price K its approximate price $\hat{C}_T$ is given by

$$\hat{C}_T = e^{-rT} \frac{1}{N} \sum_{j=1}^{N} C_T(\omega_j)$$

where the ω_i, $i = 1, \ldots, N$, are states of nature drawn from Ω. The price could as well be estimated by

$$\bar{C}_T = e^{-rT} \frac{1}{N} \sum_{j=1}^{N} (C_T(\omega_j) - \Gamma(\omega_j) + E[\Gamma])$$

where Γ is a random variable defined on the same probability space.
Both estimators are unbiased, but if $\text{cov}[\Gamma, C_T] \geq \frac{1}{2}\text{var}[\Gamma]$ the estimator $\bar{C}_T$ has smaller variance because

$$\text{var}[\bar{C}_T] = \text{var}[\hat{C}_T] + \frac{1}{N}(\text{var}[\Gamma] - 2\text{cov}[\Gamma, C_T]).$$

This means that a random variable Γ is wanted which is easy to calculate and has a high correlation with the option payoff C_T. Pellizzari (1998) proposes

$$\Gamma_1 = (\omega_1 S_T^X + \omega_2 E[S_T^Y] - K)^+ \quad \text{and} \quad \Gamma_2 = (\omega_1 E[S_T^X] + \omega_2 S_T^Y - K)^+$$

where $E[S_T^X] = S_0^X \exp[T(b_1 + \frac{1}{2}c_1)]$ and $E[S_T^Y] = S_0^Y \exp[T(b_2 + \frac{1}{2}c_2)]$. The expectation of Γ_1 and Γ_2 is readily evaluated by the Black-Scholes formula with S^X resp. S^Y as the underlying security and by plugging in $\omega_2 E[S_T^Y] - K$ resp. $\omega_1 E[S_T^X] - K$ as the corresponding strike prices.
It is worth mentioning that the parameters b_1 and b_2 are adjusted according to equation (4.3.23) to obtain

$$b_1 = r - \frac{1}{2}\sigma_1^2 \quad \text{and} \quad b_2 = r - \frac{1}{2}\sigma_2^2$$

such that the two marginal diffusion processes are martingales.

Notation

$-A$	$\{x \in \mathbb{R}^n \| -x \in A\}$, $A \subset \mathbb{R}^n$
A'	transposed of a matrix A
A^i	i-th row of a matrix A, interpreted as a column vector
A^{ij}	element in the i-th row and j-th column of a matrix A
ATM	at-the-money (for an option)
$[a,b]$ (n-box)	$[a^1,b^1] \times [a^2,b^2] \times \ldots \times [a^n,b^n]$, $a=(a^1,\ldots,a^n)$, $b=(b^1,\ldots,b^n)$
$B_\delta(x)$	$\{y \in \mathbb{R}^d : \|y-x\| < \delta\}$
$B_{\delta_1,\delta_2}(x_1,x_2)$	$\{(y_1,y_2) \in \mathbb{R}^2 : \|y_1-x_1\| < \delta_1, \|y_2-x_2\| < \delta_2\}$
$\mathbb{C}$	complex numbers
$\mathrm{cov}(X,Y)$	covariance of two random variables X and Y
$\mathrm{Dom}C$	domain of a function C
e^X	$(e^{X^1},\ldots,e^{X^n})'$ for $X \in \mathbb{R}^n$
$\mathbf{e}^X$	$\mathrm{diag}(e^{X^1},\ldots,e^{X^n})$ for $X \in \mathbb{R}^n$
$F^{s,t,m}(x,y)$	$\frac{\partial^s}{\partial x^s}\frac{\partial^t}{\partial y^t}[F(x,y)]^m,\ m \in \mathbb{N}_0$
$\Im(z)$	imaginary part von $z \in \mathbb{C}$
ITM	in-the-money (for an option)
$\mathbb{N}$	$\{1,2,3,\ldots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$
OTM	out-of-the-money (for an option)
$\psi^{s,t}(z,z')$	$\frac{\partial^s}{\partial z^s}\frac{\partial^t}{\partial z'^t}\psi(z,z')$
$\bar{\mathbb{R}}$	extended real line $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$
$\mathrm{Ran}C$	range of a function C
$\Re(z)$	real part of $z \in \mathbb{C}$
$\mathrm{var}(X)$	variance of a random variable X
$\|x\|$	any norm of $x=(x^1,\ldots,x^n) \in \mathbb{R}^n$ for $n \geq 2$
$\|x\|$	$\sqrt{x^2}$ for $x \in \mathbb{R}$
$\|x\|_\infty$	maximum norm $\max\{\|x^1\|,\ldots,\|x^n\|\}$ of $x=(x^1,\ldots,x^n) \in \mathbb{R}^n$
X	stochastic process $(X_t)_{t\geq 0}$ on $(\Omega,\mathcal{F},\mathbf{P})$
$X=(b,c,K)_P$	Lévy process with characteristic triplet b, c and K under the probability measure $\mathbf{P}$
$\bar{z}$	the complex conjugate $x-iy$ of $z=x+iy \in \mathbb{C}$

Bibliography

M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1972.

L. Arnold. *Stochastische Differentialgleichungen*. R. Oldenbourg Verlag München Wien, 1973.

O.E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modeling. *Scandinavian Journal of Statistics*, 24:1–13, 1997.

O.E. Barndorff-Nielsen. Normal inverse gaussian distributions and stochastic volatility modeling. Research Report no. 18, Centre for Mathematical Physics and Stochastics, University of Aarhus, 2000.

O.E. Barndorff-Nielsen and A. Lindner. Some aspects of Lévy copulas. Working Paper, available from `http://www-m4.mathematik.tu-muenchen.de/m4/pers/lindner/`, 2004.

S. Barnett. *Matrices: Methods and Applications*. Oxford University Press, 1990.

D.S. Bates. The crash of '87: Was it expected? the evidence from options markets. *The Journal of Finance*, 46(3):1009–1044, 1991.

H. Bauer. *Maß-und Integrationstheorie*. Walter de Gruyter, 1992.

T. Björk. Interest rate theory. In W. Runggaldier, editor, *Financial Mathematics*, volume 1656 of *Lecture Notes in Mathematics*. Springer, 1992.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

P. Blæsild and M.K. Sørensen. 'hyp' - a computer program for analyzing data by means of the hyperbolic distribution. Research report no.248, Department of Theoretical Statistics, University of Aarhus, 1992.

M. Blix. Pricing basket options by use of the Edgeworth expansion. Working Paper, Department of Mathematical Statistics, Lund University, Sweden, 1998.

W. Breymann, A. Dias, and P. Embrechts. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, 3(1):1–16, 2003.

I.N. Bronstein, K.A. Semendjajew, G. Musiol, and H. Mühlig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 1993.

P. Carr, H. Geman, D.P. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. Working Paper, 2000a.

P. Carr, K. Lewis, and D.P. Madan. On the nature of options. Working Paper, 2000b.

A. Černý. *Mathematical Techniques in Finance*. Princeton University Press, 2004a.

A. Černý. The risk of optimal, continuously rebalanced hedging strategies and its efficient evaluation via Fourier transform. Working Paper. Tanaka Business School, Imperial College London, 2004b.

T. Chan. Pricing contingent claims on stocks driven by Lévy processes. *The Annals of Applied Probability*, 9(2):504–528, 1999.

T. Choulli, L. Krawczyk, and C. Stricker. $\mathcal{E}$-martingales and their applications in mathematical finance. *The Annals of Probability*, 26(2):853–876, 1998.

D.B. Colwell and R.J. Elliott. Discontinuous asset prices and non-attainable claims. *Mathematical Finance*, 3(3):295–308, 1993.

R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

R. Cont and P. Tankov. Calibration of jump-diffusion option pricing models: a robust non-parametric approach. *Journal of Computational Finance*, 7(3):1–49, 2004a.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, 2004b.

C.J. Corrado and T. Su. Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *The European Journal of Finance*, 3: 73–85, 1997.

G. Deelstra, J. Liinev, and M. Vanmaele. Pricing of arithmetic basket options by conditioning. *Insurance: Mathematics and Economics*, 34:55–77, 2004.

G. Doetsch. *Handbuch der Laplace-Transformation*, volume 1. Birkhäuser, Basel, 1950.

E. Eberlein and J. Jacod. On the range of options prices. *Finance and Stochastics*, 1:131–140, 1997.

E. Eberlein and A. Papapantoleon. Symmetries and pricing of exotic options in Lévy models. Working Paper, University of Freiburg, 2004.

E. Eberlein and F. Özkan. Time consistency of lévy models. *Quantitative Finance*, 3:40–50, 2003.

P. Embrechts, F. Lindskog, and A. McNeil. Modelling dependence with copulas and applications to risk management. Working Paper, Department of Mathematics, ETH Zürich, 2001.

F. Esche. *Two Essays on Incomplete Markets*. PhD thesis, Technical University of Berlin , School of Mathematics and Natural Sciences, 2004.

D. Flamouris and D. Giamouridis. Valuing exotic derivatives with jump diffusions: The case of basket options. Working Paper, available from `http://papers.ssrn.com`, 2004.

H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In M.H.A. Davis and R.J. Elliott, editors, *Applied Stochastic Analysis*, pages 389–414. Gordon and Breach, New York, 1991.

H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics in Honor of Gérard Debreu*, pages 205–223. North Holland, Amsterdam, 1986.

D. Gentle. Basket weaving. *Risk*, 6(6):51–52, 1993.

Á. Gereben. Extracting market expectations from option prices: an application to over-the-counter new zealand dollar options. Reserve Bank of New Zealand Discussion Paper, 2002.

P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2004.

T. Goll and J. Kallsen. Optimal portfolios for logarithmic utility. *Stochastic Processes and their Applications*, 89:31–48, 2000.

T. Goll and L. Rüschendorf. Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics*, 5:557–581, 2001.

H. Gut. *An Intermediate Course in Probability*. Springer, 1995.

J. Harrison and S. Pliska. Martingales and stochastic tntegrals in the theory of tontinuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.

C.B. Huynh. Back to baskets. *Risk*, 7(5):59–61, 1994.

J. Jacod. *Calcul Stochastique et Problèmes de Martingales*. Springer, 1979.

J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 2003.

R. Jarrow and A. Rudd. Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10:347–369, 1982.

J.L. Jensen. *Saddlepoint Approximations*. Clarendon Press, Oxford, 1995.

N. Ju. Pricing Asian and basket options via Taylor expansion. *Journal of Computational Finance*, 5(3):1–35, 2002.

J. Kallsen. Optimal portfolios for exponential Lévy processes. *Mathematical Methods of Operations Research*, 51:357–374, 2000.

U. Keller. *Realistic Modelling of Financial Derivatives*. PhD thesis, Universität Freiburg i.Br., 1997.

K. Königsberger. *Analysis 2*. Springer, 1997.

S.G. Kou. A jump-diffusion model for option pricing with three properties: Leptokurtic feature, volatility smile, and analytical tractability. Working Paper, Columbia University, 1999.

S.G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8): 1086–1101, 2002.

P. Laurence and T.-H. Wang. Valuing and hedging basket options without distributional assumptions. Working Paper, available from `http://www.finance-and-physics.org/Papers/Risk1Ver18.pdf`, 2001.

A. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. Working Paper, available from `http://ideas.repec.org/p/vsv/svpubs/explevy.html`, 2001.

D.B. Madan and F. Milne. Option pricing with V.G. martingale components. *Mathematical Finance*, 1(4):39–55, 1991.

B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 63:394–419, 1963.

R.C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

Y. Miyahara. Minimal entropy martingale measures of jump-type price processes in complete assets markets. *Asia-Pacific Financial Markets*, 6:97–113, 1999.

Y. Miyahara. [Geometric Lévy process & MEMM] pricing model and related estimation problems. *Asia-Pacific Financial Markets*, 8:45–60, 2001.

R.B. Nelsen. *An Introduction to Copulas*. Springer, 1999.

P. Pellizzari. Efficient Monte Carlo pricing of basket options. Working Paper, University of Venice, Italy, 1998.

K. Prause. *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*. PhD thesis, Universität Freiburg i.Br., 1999.

P. Protter. *Stochastic Integration and Differential Equations*. Springer, 1995.

S. Raible. *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. PhD thesis, Universität Freiburg i.Br., 2000.

D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, 2001.

T.H. Rydberg. The Normal Inverse Gaussian Lévy process: Simulation and approximation. *Commun. Statist.-Stochastic Models*, 13(4):887–910, 1997.

K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

W. Schoutens. *Lévy Processes in Finance*. John Wiley & Sons, 2003.

M. Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37:339–363, 1991.

M. Schweizer. Approximating random variables by stochastic integrals. *The Annals of Probability*, 22(3):1536–1575, 1994.

M. Schweizer. On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis and Applications*, 13:573–599, 1995.

M. Schweizer. A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanic, and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, pages 538–574. Cambridge University Press, 2001.

A. Stuart and J.K. Ord. *Kendall's Advanced Theory of Statistics*. Charles Griffin & Company Limited, London, 1987.

P. Tankov. Dependence structure of spectrally positive multidimensional Lévy processes. Working Paper, Ecole Polytechnique, Palaiseau, France, 2003.

J. Wannenwetsch. Bewertung von Portfolio-Optionen in einem Sprung-Diffusionsmodell mit Lévy-Copulas. Master's thesis, Universität Bonn, 2004.

G.N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, 1966.

X. Zhang. *Analyse Numérique des Options Américaines dans un Modèle de Diffusion avec Sauts*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 1994.