Université Catholique de Louvain

Faculté des Sciences Appliquées

Unité de Gestion Industrielle

An Integrated Affine Jump Diffusion Framework to Manage Power Portfolios in a Deregulated Market

Thèse présentée en vue de l'obtention du grade de
Docteur en Sciences Appliquées par

Michel Culot

Promoteur: Professeur Y. Smeers

Louvain-la-Neuve, Janvier 2003.

Université Catholique de Louvain

Faculté des Sciences Appliquées

Unité de Gestion Industrielle

An Integrated Affine Jump Diffusion Framework to Manage Power Portfolios in a Deregulated Market

Thèse présentée en vue de l'obtention du grade de
Docteur en Sciences Appliquées par

Michel Culot

Promoteur: Professeur Y. Smeers

Louvain-la-Neuve, Janvier 2003.

Le Jury:

Professeur Michel Gevers (Président)

Université Catholique de Louvain

Faculté des Sciences Appliquées - Département d'Ingénierie Mathématique

Professeur Yves Smeers (Promoteur)

Université Catholique de Louvain

Faculté des Sciences Appliquées - Département d'Ingénierie Mathématique

Professeur Georges Bastin

Université Catholique de Louvain

Faculté des Sciences Appliquées - Département d'Ingénierie Mathématique

Professeur Léon Bolle

Université Catholique de Louvain

Faculté des Sciences Appliquées - Département de Mécanique

Professeur Michel Denuit

Université Catholique de Louvain

Institut de Statistique

Professor Freddy Delbaen

ETH - Zurich

Department of Mathematics

Professor Darrell Duffie

Stanford University

The Graduate School of Business

To my wife, Catherine, and my children, Aline and Chloé, for laughing when I bumped into walls,
and to my parents for not laughing when I bumped into walls.

"There is nothing more practical than a good theory."
Ludwig Boltzmann

Acknowledgements

First of all, I would like to express my gratitude to Yves Smeers, my supervisor. I had an enormous benefit from the numerous discussions we had, and I thank him for his patience. He provided me with unfailing guidance. His contribution is essential to the accomplishment of this work.

I am also very thankful to the members of my jury for taking the time to read the manuscript despite a very busy schedule, and for the important suggestions presented. I owe special gratitude to the Professor Léon Bolle who sadly passed away a few weeks before the defense. He was a member of my guidance committee and followed and supported with great interest this project form the very beginning until the end despite the severe distress of the illness.

The research work for this thesis has been done in two stages. First at the Center for Operations Research and Econometrics (CORE) during the first year (1997-98), then with the Risk Asset and Liability Management (RALM) department of Electrabel. I would like to acknowledge the financial support of the SSTC during my first year at CORE, and I am very grateful to Jacqueline Boucher, head of RALM, for supporting this research.

These two working environments gave me the chance to encounter many different wonderful people I would like to thank for their many advices, but above all for their friendship. In particular, I would like to mention Olivier Daxhelet, Emmanuel

Canon, Thierry Leboulengé, Stéphane Métens, Dean Travers, Cécile Cordier, Christian Hafner, Hugues Marchand and Marko Loparic. Benjamin Carton de Wiart, Sébastien de Menten de Horne, and Valérie Goffin deserve a special tribute for their invaluable contribution.

I also had the chance to benefit from the great experience of Professor Engle whom I would like to thank very much for the incredible insight he offered me.

Finally, I would like to express my gratitude to my high school teacher, Mr. Hosselet, from the Ecole Abbatiale de Maredsous, who initiated me to the beauty of mathematics.

Abstract

Electricity markets around the world have gone through, or are currently in, a restructuring phase. The opening of electricity markets to competition dramatically changes the electric utilities environment by introducing the notion of risk. As a consequence, new activities have now revealed strong competitive stakes. It is indeed crucial for utilities to accurately value complex exotic contracts and physical power plants, and to manage the risk of their positions through hedging. We propose here an integrated Affine Jump Diffusion framework that coherently deals with all of these issues.

Contents

Chapter 1

Introduction

Electricity markets around the world have gone through, or are currently in, a restructuring phase. The opening of electricity markets to competition dramatically changes the electric utilities environment by introducing the notion of risk. As a consequence, new activities have appeared or have now revealed strong competitive stakes.

A necessary foundation for all of these activities is a representation of the uncertainty in the level of future power and fuel prices. In Chapter 2, a general multi-variate Affine Jump Diffusion (AJD) setting is defined to model the uncertainty on the electricity prices and the different fuel prices. This model characterizes the so called "physical" statistical description of uncertainty. Discussions of the applicability of AJD to energy prices can be found in Johnson and Barz (1999) [25], Deng(1999) [13], Duffie (1999) [16], Pindyck (1998) [36] ...

When pricing contingent claims, however, one must also take into account a risk-premium that "coherently" prices the risks that are associated to them. "Coherently" is to be understood here as "in a way that does not allow for arbitrage opportunities". Though arbitrage opportunities might exist during some short periods of time, it is a well accepted fact that they should not be allowed for in models. Allowing for arbitrage opportunities would indeed have undesired implications in many applications such as portfolio selection because it would imply the possibility of an infinite gain with no risk ! Instead, it is preferred to have "no arbitrage opportunities" models and to interpret as an indication of a possible arbitrage opportunity any discrepancy between the prices given by the model and the observed ones. The last section of Chapter 2 is thus dedicated to "coherently" introduce a price of risk to value any contingent claim.

Next, it is shown in chapter 3 how powerful the Affine Jump Diffusion framework is to price complex exotic options. This power essentially comes from the fact that the conditional characteristic function of an AJD vector is given by the solution of a system of complex-valued ODE's as shown in Duffie, Pan and Singleton (1999) [15]. The importance of stochastic processes whose characteristic function is known in closed form is pointed out in Bakshi and Madan (1999) [1] who show that *"by translating or differentiating the characteristic function, one can synthesize the values of exponential and polynomial of the uncertainty"*.

Pricing formulas are developed for a wide variety of exotic options considered as relevant in the energy (see Kaminski et al. (1999) [27]) in Chapter 4. In this chapter, the machinery is also applied to the "real options" valuation of a multi-fuel power plant. Furthermore, the valuation of interruptible contracts with an early notification option is considered. Kamat and Oren (2000) have indeed shown that these contracts provide a perfect hedge for customers that can curtail loads in response to high spot prices and can mitigate curtailment losses when the curtailment decision is made with sufficient lead time. This valuation problem involves the pricing of compound options. In this chapter, it is shown that compound options can be valued by using the framework that has been developed, even in the multi-variate case.

The "real option" valuation of multi-fuel power plants has been chosen as an example of implementation because it nicely illustrates the need to use evaluation techniques that fully capture the value of flexibility offered by power plants, and how this can be done in a deregulated environment. This issue is currently particularly crucial as many mergers and acquisitions have been taken place in the European energy industry. On the other hand, the interruptible contracts with early notification option example shows how, in a deregulated market, demand management problems can be solved by introducing new structures of contracts that provide adequate incentives to consumers, and how consumers can benefit from them.

Some exotic options, such as Take-or-Pay options, however, do not have a closed-form solution. Their valuation requires a dynamic programming algorithm based on a grid. In Chapter 5, we develop such a grid algorithm that uses all the power of AJD's numerical tractability by using Fast Fourier transforms.

Finally, risk-management is dealt with in Chapter 6. Once a contract has been valued and eventually a position taken, it is indeed very important to be able to manage its risk by constructing a hedge against it. Hedging against one risk factor consists of taking an opposite position in proportion to the impact of this risk factor. Hedging is often conducted through "Greeks" which measure this impact. It is shown how Greeks can be efficiently computed in the AJD framework.

A statistical estimation procedure of the parameters of the Affine Jump Diffusion (AJD) describing the power and fuel dynamics in the "physical" probability measure together with the "coherent" risk-premiums is developed in Chapter 7. This estimation procedure builds upon the Kalman Filter algorithm.

Chapter 8 concludes the thesis and indicates some areas of further research.

Chapter 2

Uncertainty and Risk-Premiums

2.1 Introduction

When pricing a derivatives contract, one usually uses the arbitrage pricing approach. That is, if it is possible to construct a dynamically adapted portfolio that will perfectly replicate the payoff of the derivative contract, then the argument of absence of arbitrage forces the price of this derivative to be equal to the price of the replicating portfolio. If it is always possible to build a dynamically adapted portfolio that will perfectly replicate any payoff, then the market is said to be complete, there will only be a single no-arbitrage price for any contingent claim, and there exists a probability measure called the risk-neutral probability measure equivalent to the physical probability measure under which the no-arbitrage price of any contingent claim is equal to the expectation of its payoff discounted at the risk-free rate.

Electricity, however, is a quite challenging commodity for analysts who want to price derivatives on it. First of all, storage possibilities are very limited and quite expensive. Only large hydro systems do have this possibility. Even then, because reservoirs and turbine capacities are not infinite, this possibility is only partial. Lack of storage capacities prevents us from using no arbitrage arguments for derivatives pricing because one cannot create a replicating trading strategy involving the spot price. Furthermore, the dynamics of electricity spot price are very complex and quite far from what one usually assumes for financial assets. Mean-reversion, stochastic volatility and jump behaviors are the joined consequences of the existence of a long term equilibrium price towards which the price is assumed to revert, together with the lack of storage capacities, a rather inelastic demand curve, a marginal cost curve rising sharply, and frequent failures in the production or transmission capacities.

Next to the physical spot market, most electricity markets have a well developed financial market with futures contracts traded. Though the spot price is not an asset that can be used in a replicating portfolio (because of its lack of storability), futures contracts, on the other hand, are regular financial contracts that can be traded and used in a replicating strategy.

This information contained in the futures prices should not be disregarded when pricing derivatives on the spot price. Similarly, one should not disregard the spot price dynamics information when constructing a futures model or pricing derivatives on futures contracts. Clearly, there is a need for a coherent unifying framework which allows to price derivatives on both the spot and the futures contracts.

Modelling spot and futures prices dynamics, estimating the parameters of their stochastic processes, and pricing contingent claims on both of them are closely related issues. Though each of these activities have been historically approached separately, it is now recognized that strong links are tying them together. Completeness of the market is in fact the key issue relating these activities.

In other words, one wonders if the dynamics of spot and futures contracts of different maturities have any reason to be linked with one another if there are only a few players that have the possibility to buy energy today, to store it (at a quite high cost and with some physical limits on the volumes), and to sell it back later on ?

Furthermore, one wonders if, for any claim contingent on the spot or futures prices, there is only one or an infinity of prices that will prevent any arbitrage opportunity.

The issue of the completeness of a market is always closely related to the choice of a stochastic process to model its securities. If we consider a market composed of a single security and a deterministic risk-free bond, this market will be complete if the security's dynamics are modelled with a single factor diffusive stochastic process, but it will be incomplete if its dynamics are modelled with a multi-factor diffusive stochastic process. It will also be incomplete if we introduce a jump risk.

As a result, the answer to the questions above will depend on wether one only considers as starting point a stochastic process for the spot price dynamics and considers futures contracts as contingent claims on the spot price with no assumption on their stochastic process, i.e. the spot price is the only primary traded security. If this is the case, then only spot price historical data are used for the parameter estimation of the stochastic process, and the next question one should ask itself is how far do energy commodities lie from a tradable asset? Four different approaches are considered in the literature (see Seppi (2001) [42]):

- Energy commodities are non-dividend paying assets:

  ∗ The market is then complete if the model is univariate and diffusive.

  ∗ Under the risk-neutral probability measure, the energy spot price must have a drift equal to the risk-free rate of return, so that we cannot reproduce any mean reversion.

- Energy commodities are dividend paying assets:

  ∗ The market is then complete if the model is univariate and diffusive.

  ∗ Under the risk-neutral probability measure, the energy spot price must have a drift equal to the difference between the risk-free rate of return and the convenience yield.

$*$ Mean-reversion in the spot price dynamics under the risk-neutral probability measure can be achieved if the convenience yield is modelled such that this difference is negative (positive) when the spot price exceeds (lies under) a given threshold.

- Energy commodities are not assets, but, rather are in some cases such as electricity, consumption goods with no storability:

  $*$ The market is then incomplete.

  $*$ One must assume a price of risk to price contingent claims.

  $*$ The drift in the risk-neutral probability measure is the drift in the physical probability measure minus the market price of risk. It is thus possible to reproduce some mean-reverting behavior under the risk-neutral probability measure.

  $*$ Contingent claims prices are, however, internally consistent, in the sense that they will not provide any arbitrage opportunities and that they will be valued based on the same price of risk.

If, however, one considers, as a starting point, different stochastic processes that are assumed to describe both the spot and futures contracts with different maturities (possibly with some constraints to ensure the absence of arbitrage opportunity),i.e. both spot and futures contracts are considered as primary traded securities, then these stochastic processes should be estimated on both the spot and futures prices historical time series. In this case, since futures contracts are regular financial contracts that can be traded, as long as the model remains diffusive and that values of the risk-factors can be uniquely determined by the observation of the different futures prices, the market will be complete.

With these categories in mind, we can now review the different stochastic models one can find in the literature to describe the dynamics of energy commodity prices, how they are estimated and the assumptions that are made to price contingent claims with them.

The first attempt to model energy prices was a simple application of the Black and Scholes (1973) [3] model. The spot price is assumed to be a Geometric Brownian Motion (GBM) and the model is estimated on only spot price data. Under this model, energy commodities are considered as non-dividend paying assets. The market is thus complete.

The next step was introduced by Bernnan (1991) [2], who sees energy commodities as dividend-paying assets. The spot price dynamics are assumed to follow a GBM with a convenience yield function of the spot price level. Again, the model is estimated only on spot price data. This framework also implies a complete market.

Then Pilipovic (1998) [35] sees energy commodities as consumption goods. The spot price dynamics are assumed to follow a two factor diffusive mean-reverting process. The model is estimated only on spot price data. In this framework, the market is incomplete, so one needs to assume some risk premiums to price contingent claims. These prices of risk are calibrated to match the forward curve level.

The first to consider both spot and forward contracts as primary traded securities are Gibson and Schwartz (1990) [18]. As Bernnan (1991) [2], they consider energy commodities as dividend-paying assets with a convenience yield modelled as a stochastic risk factor. The introduction of an additional source of risk makes the spot market (considered separately) incomplete. However, as mentioned earlier, if one considers as primary traded securities both spot and forward markets, then the market will be complete and the dynamics of the forward price will be determined by the choice of a risk factor. This implies that the model must be estimated on both spot and forward data, simultaneously. Gibson and Schwartz 1990) [18] perform this estimation via a Kalman Filter recursive algorithm.

Later on, Schwartz and Smith (2000) [41] proposed to consider energy commodities as consumption goods. The dynamics of the spot price is assumed to follow a two factor diffusive mean-reverting process. Again, they consider as primary traded commodities both spot and forward contracts. As a result, the market will be complete and the dynamics of the forward prices are completely determined by the choice of some risk-premiums. In their paper, they further show that this model is a rotation of the stochastic convenience yield model of Gibson and Schwartz (1990) [18] so that it can be considered as equivalent in terms of spot and futures dynamics, but more intuitive in terms of representation.

Manoliu and Tompaidis (1999) [29] also consider energy commodities as consumption goods. They take as primary traded securities both spot and forward contracts and propose a spot price model that is the sum of several mean-reverting risk factors. Assuming a price of risk specification, the dynamics of the forward contracts are then completely determined. The model is estimated on both spot and forward historical data via a Kalman Filter recursive algorithm. If the number of risk-factors considered is less than or equal to the number of forward contracts traded, then the market is complete.

Clewlow and Strickland (1999) [5] focus on the forward contracts. The model considers a non-parametric variance term structure of the forward curve and is estimated via a principal components decomposition of the variance-covariance matrix of the forward term structure. If there aren't more risk-factors than forward contracts, then the market is complete. However, the spot price dynamics implied by such a model is, in general, not Markovian and thus very difficult to use.

Deng (1999) [13] starts from the spot price dynamics that is assumed to follow a multi-factor Affine Jump Diffusion Process and is the only primary traded commodity. The model is estimated on the spot price historical data. In this framework, however, the market is not complete and assumptions of prices of risk must be made to price derivatives. Deng considers a price of risk for only the diffusive sources of risk. Jump risks are assumed not to be priced. These prices of risk are calibrated to match some contingent claims prices.

Finally, this survey would not be complete without mentioning a quite different approach developed by Routledge, Seppi and Spatt (1998) [38] based on an equilibrium model and storage theory. We will keep this approach aside as it is quite far from what is developed in this work. There is, however, an incremental contribution of our AJD framework over this equilibrium model because it is numerically much more tractable in terms of parameter estimation, and in terms of derivatives pricing.

In this chapter, we consider energy commodities as consumption goods. We model the dynamics of the spot price by an Affine Jump Diffusion Process but consider as starting point stochastic processes for both the spot and forward contracts prices. It has been argued above that in this case, if one specifies a stochastic process for the spot price and a risk-premium associated to each factor of risk then the stochastic process for the forward prices were completely determined and the market is complete (at least if the values of the risk-factors can be uniquely determined by the observation of the different futures prices and if there is no jump). We now develop this point a bit further.

It is a well known fact that if the futures market admits no arbitrage opportunity, then at least one so called "equivalent martingale" probability measure must exist. Under this new probability measure, futures prices are martingales, that is their price at time t is their expected price at any time t', with $t' > t$ (see Harrison and Kreps (1979) [20], Cox, Ingersoll and Ross (1981) [6], Musiela and Rutowski (1998) [31], and Delbaen and Schachermayer (2001) [12]). Furthermore, it is quite obvious that the no-arbitrage condition also implies that any contingent claim must converge towards the value of its payoff at maturity. In particular, futures prices must converge to the spot price at maturity.

From these two properties, it is quite trivial to show (from the property of repeated expectations) that futures prices must, at any time, be the expected value of the spot price at their maturity under an equivalent martingale probability measure. This close link between the spot and the futures prices' dynamics, implied only from no-arbitrage, considerably restricts the candidate stochastic models for describing the futures prices' dynamics.

A first problem immediately arises from this restriction. What if, once we have chosen a given stochastic process to describe the dynamics of the spot price under the physical probability measure, there isn't any equivalent probability measure under which all the futures prices observed are the expectation of the spot price at their maturity ? We are then unable to write a stochastic process describing the dynamics of all the futures prices under the physical probability measure that will not allow for some arbitrage opportunities.

As explained above, we do not want to build models that allow for some arbitrage opportunities. Instead, we will rather build an arbitrage-free model that might not be able to explain all of the futures prices movements but that will be chosen such as to "best fit" the observed spot and futures prices dynamics.

First, we define a given class of equivalent probability measures that prices all sources of uncertainty. We then choose, within this class, the particular martingale probability measure, Q^*, that "best fits" the observed historical prices of available financial contracts through a statistical estimation procedure.

To statistically estimate this particular equivalent martingale probability measure, we start by characterizing the Radon-Nikodym derivative of all the candidate equivalent martingale probability measures. This defines the particular class of equivalent probability measures considered. As mentioned above, this characterization should allow one to price every source of uncertainty, including stochastic volatility and jumps. In a first step, we leave the characterization of the probability distribution of the jumps of the Radon-Nikodym derivative free, allowing for total freedom in pricing of the risks implied by the stochastic arrivals of jumps and by the distribution of their size. We give the relationship between the mean arrival rates and the probability distribution of the jump processes under the actual probability measure P, and their counterparts under the equivalent martingale measure Q, according to the probability distribution chosen for the jumps of the Radon-Nikodym derivative. Then, in a next step, we consider particular cases such as the Gaussian case developed by Pan (2001) [34] that only prices the risk associated with the distribution of the sizes of the jumps. Another particular case studied links the pricing of both risks (arrival rate and size) associated to jumps through a single vector. This latter case will be further adopted for the characterization of the Radon-Nikodym derivative process.

The simultaneous statistical estimation procedure of the parameters of the AJD describing the spot price process in the physical probability measure and the parameters of the characterized Radon-Nikodym derivative, based on both spot and forward historical prices, are developed in Chapter 7.

If the resulting model describing the dynamics of the futures prices under the physical probability measure implies a complete futures market, then the equivalent martingale probability measure considered above, Q^*, is in fact unique and any other additional contingent claim must be priced as the expectation under Q^* of its payoff discounted at the risk-free rate. This is indeed the only price that will satisfy the no arbitrage opportunity condition.

If the implied futures market is incomplete, the number of equivalent probability measures under which futures prices are martingales is infinite. Pricing all the other contingent claims as the expectation under Q^* of their payoff discounted at the risk-free rate will result in prices that do not allow for arbitrage opportunities. These prices, however are not the only ones that will satisfy the no arbitrage opportunity condition. All the prices obtained as the expectation, under any one of the infinity of equivalent martingale probability measures, of their payoffs discounted at the risk-free rate will satisfy the no arbitrage opportunity condition.

As shown further, choosing Q^* within this infinity of equivalent martingale probability measures, will however keep the Affine Jump Diffusion property of the spot and futures prices under the equivalent martingale probability measure. This will then enable us to make use of the great computation power of the Affine Jump Diffusion framework to compute the expectation of the payoffs of many different kinds of contingent claims on the spot and on futures contracts. Furthermore, this assumption can be interpreted as keeping the same utility function when pricing futures contracts or any other financial contract.

Our approach is thus along the lines of Schwartz and Smith (2000) [41] and Manoliu and Tompaidis (1999) [29], though we expand their models to the entire family of multi-factor Affine Jump Diffusion processes, allowing us to go beyond the Gaussian framework developed by these authors. In comparison to Deng (1999) [13], our approach differs from the fact that we consider simultaneously spot and forward historical information in the estimation of our parameters and that we introduce a pricing of jump risks.

This chapter will be divided as follows. A first section will set the AJD framework describing the "physical" probability measure. Section 2 will present the characterization of the Radon-Nikodym derivative. With this characterization, it is further shown that the stochastic process describing the dynamics of the risk factors remains an AJD under the implied equivalent probability measure.

2.2 Description of the "physical" probability measure by an AJD

Affine Jump Diffusions are well equipped to model most of the features one finds in Energy price dynamics:

- Mean reversion denotes the fact that Energy prices tend to revert towards a "long-term" equilibrium price given by a marginal cost of production;
- Due to a lack of storability and the sensitivity of the offer and demand functions on weather conditions, Energy prices dynamics show some seasonality;
- Jumps and spikes are the result of common breakdowns of transmission lines or power plants;
- Negative prices are sometimes possible because of switching on and off costs and ramp-up rates or, in large hydro-systems during high-water periods (Note, however, that negative prices are only possible in the "level" model developed hereunder);
- Finally, the very steep slope of the marginal cost curve at high prices results in some heteroscedasticity in the Energy price dynamics;

All of these features can indeed be modelled by AJD's.

Two types of models are considered here. The first model works in terms of returns and describes the logarithm of the spot price as a multi variate AJD. The second model works in terms of spot price levels and describes the level of the spot price as a multi variate AJD.

2.2.1 The "return" model

Consider a multi factor model of the type

$$S_t = e^{\gamma_{0,1}+\gamma_1.X_t} \tag{2.1}$$

where

- S_t is the electricity spot price;
- X_t is the vector of risk-factors. These risk-factors may be identifiable with observable variables such as a fuel price, a meteorological variable, or a measure of the economic activity, but this is not mandatory. They may as well be considered as instrumental variables used to best model the dynamics of the spot price. This is the approach that will be adopted in what follows. These instrumental variables, however, can be given, a posteriori, some substantive interpretation such as a long term equilibrium cost of production level or trend. Finally, some elements of this vector of risk factors can be interpreted as the instantaneous variance of other risk factors as will be further noted in the section relative to stochastic volatility models.

We fix a probability space $(\Omega, \mathcal{F}, P)$, where P will further be referred to as the "physical" probability measure, and an information filtration $(\mathcal{F}_t)$, and suppose that X is a Markov process in some state space $D \subset \mathbb{R}^n$, solving the stochastic differential equation

$$dX_t = \mu(X_t, t)\ dt + \sigma(X_t, t)\ dW_t + dZ_t \tag{2.2}$$

where :

- W is a Brownian motion in $\mathbb{R}^n$;
- $\mu : \{D, [0, \infty)\} \to \mathbb{R}^n$;

- $\sigma : \{D, [0, \infty)\} \to \mathbb{R}^{n \times n}$;
- Z is a pure jump process whose jumps have a fixed probability distribution ν on $\mathbb{R}^n$ and arrive with intensity $\{\lambda(X_t, t) : t \geq 0\}$, for some $\lambda : \{D, [0, \infty)\} \to [0, \infty)$.

For $c \in \mathbb{C}^n$, the set of n-tuples of complex numbers, we let $\theta(c) = \int_{\mathbb{R}^n} \exp(c.z) d\nu(z)$ whenever the integral is well defined.

We impose an affine structure on μ, $\sigma\sigma^T$, and λ in that all of these are assumed to be affine in X, and we further fix an affine discount-rate function $R : \{D, [0, \infty)\} \to \mathbb{R}$:

- $\mu(X_t, t) = K_0(t) + K_1(t).X_t$, for $K(t) = (K_0(t), K_1(t)) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}$;
- $(\sigma(X_t, t)\sigma(X_t, t)^T)_{i,j} = (H_0(t))_{i,j} + (H_1(t))_{i,j}.X_t$, for $H(t) = (H_0(t), H_1(t)) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n \times n}$;
- $\lambda(X_t, t) = l_0(t) + l_1(t).X_t$, for $l(t) = (l_0(t), l_1(t)) \in \mathbb{R} \times \mathbb{R}^n$;
- $R(X_t, t) = \rho_0(t) + \rho_1(t).X_t$, for $\rho(t) = (\rho_0(t), \rho_1(t)) \in \mathbb{R} \times \mathbb{R}^n$.

2.2.2 The "level" model

In the "level" model, the spot price is given by a multi-factor model of the type

$$S_t = \gamma_{0,2} + \gamma_2.X_t \quad , \tag{2.3}$$

where, again,

- S_t is the electricity spot price;
- X_t is the vector of risk-factors.

The vector X of risk-factors is assumed to follow the same AJD dynamics as in the previous "return" model.

2.3 Specific Examples

In this section, we propose different models for energy commodities with an increasing level of sophistication.

2.3.1 The One Factor Model

In the one factor model, spot prices are first made stationary by extracting a weekly pattern, a linear trend and a seasonal cycle. The residual is then assumed to follow a simple mean reverting process.

$$\log(Spot_t) = Weekly\_Pattern(t) + \theta(t) + X_t \tag{2.4}$$

$$\theta(t) = a_0 + b\ t + c_1 \cos(2\pi t/365) + c_2 \sin(2\pi t/365) \tag{2.5}$$

$$dX_t = -\kappa X_t dt + \sigma dW_t \tag{2.6}$$

We chose a "return" model, that is, the risk factor models the de-trended logarithm of the spot price. Advantages of this kind of models are that

- the spot price level will be more volatile during periods of high prices than during periods of low prices. This is indeed supported by the shape of the marginal cost curve which is much steeper at high prices than at low prices.
- the spot price always remains positive no matter how big its variance is.

On the other hand, the last advantage can in fact be seen as a disadvantage as we know that in some very rare cases energy prices can become negative for some short periods of time. However, "level" models that would allow for negative prices typically over estimate the probability of negative price occurrences in long term simulations, where the variance of the distribution becomes high.

2.3.2 The Two Factor Model

In the two factor model, the level of the trend towards which the spot price is reverting is itself a stochastic state variable:

$$\log(Spot_t) = Weekly\_Pattern(t) + \theta(t) + X_1(t) \tag{2.7}$$

$$\theta(t) = a_0 + b\ t + c_1 \cos(2\pi t/365) + c_2 \sin(2\pi t/365) \tag{2.8}$$

$$dX_1(t) = \kappa_1(X_2(t) - X_1(t))dt + \sigma_1 dW_1(t) \tag{2.9}$$

$$dX_2(t) = -\kappa_2 X_2(t)dt + \sigma_2 dW_2(t) \tag{2.10}$$

$$E[dW_1(t)dW_2(t)] = \rho dt \tag{2.11}$$

Adding a stochastic mean of reversion is crucial as it allows to

- model a high level of volatility of the spot price with a quite strong speed of reversion without setting to zero the volatility of longer term futures prices;
- introduce a correlation between the spot or short term futures and the long term futures that is lower than 1.

2.3.3 The Three Factor Model

Finally, in the three factor model, not only the level but also the slope of the trend is made stochastic:

$$Spot_t = Weekly\_Pattern(t) + \theta(t) + X_1(t) \tag{2.12}$$

$$\theta(t) = a_0 + c_1 \cos(2\pi t/365) + c_2 \sin(2\pi t/365) \tag{2.13}$$

$$dX_1(t) = \kappa_1(X_2(t) + b_t - X_1(t))\ dt + \sigma_1\ dW_1(t) \tag{2.14}$$

$$dX_2(t) = -\kappa_2 X_2(t)\ dt + \sigma_2\ dW_2(t) \tag{2.15}$$

$$db_t = X_3(t)\ dt \tag{2.16}$$

$$dX_3(t)_t = \kappa_3(\overline{\mu} - X_3(t))\ dt + \sigma_3\ dW_3(t) \tag{2.17}$$

$$E[dW_1(t)\ dW_2(t)] = \rho_{12}\ dt \tag{2.18}$$

$$E[dW_1(t)\ dW_3(t)] = \rho_{13}\ dt \tag{2.19}$$

$$E[dW_2(t)\ dW_3(t)] = \rho_{23}\ dt \tag{2.20}$$

Introducing a mean-reverting process to model the slope of the trend is indeed a quite intuitive way to add a volatility component to the volatility term structure of spot and futures prices that does not vanish to zero as the maturity of the futures goes to infinity.

2.3.4 Modelling stochastic volatility

In the models developed above, the AJD risk factors had deterministic volatilities and correlations. By introducing a non-zero H_1 tensor, it is however possible to model stochastic volatilities and correlations.

Dai and Singleton (1999) [11] showed that, for a model to remain Affine, it must be possible to express its volatility, $\sigma(t)$, under the form

$$\sigma(t) = \Sigma(t)\sqrt{S(t)} \tag{2.21}$$

where $\Sigma(t)$ is a deterministic matrix and $S(t)$ is a diagonal matrix with

$$[S(t)]_{i,i} = a_i + b_i \cdot X_t \tag{2.22}$$

We can thus introduce some stochastic volatilities and correlations to the two factor AJD process developed above and rewritten here:

$$dX(t) = \begin{pmatrix} -\kappa_1 & \kappa_1 \\ 0 & -\kappa_2 \end{pmatrix} X(t)\ dt + \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\ \sigma_2 \end{pmatrix} dW_t \tag{2.23}$$

by introducing the vector $Y(t) = (X^T(t), h_1(t), h_2(t))^T$ and defining the process

$$
\begin{aligned}
dY(t) &= \left(\begin{pmatrix} 0 \\ 0 \\ \kappa_3 \bar{h}_3 \\ \kappa_4 \bar{h}_4 \end{pmatrix} + \begin{pmatrix} -\kappa_1 & \kappa_1 & 0 & 0 \\ 0 & -\kappa_2 & 0 & 0 \\ 0 & 0 & -\kappa_3 & 0 \\ 0 & 0 & 0 & -\kappa_4 \end{pmatrix} X(t) \right) dt \\
&\quad + \begin{pmatrix} \alpha_1 & \alpha_2 & 0 & 0 \\ \beta_1 & \beta_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{h_1(t)} & 0 & 0 & 0 \\ 0 & \sqrt{h_2(t)} & 0 & 0 \\ 0 & 0 & \xi_1 & 0 \\ 0 & 0 & 0 & \xi_2 \end{pmatrix} dW_t
\end{aligned}
$$

This is indeed an AJD process. The instantaneous variance of the process is given by

$$
\sigma(t)(\sigma(t))^T = \begin{pmatrix} \alpha_1^2 h_1(t) + \alpha_2^2 h_2(t) & \alpha_1\ \beta_1\ h_1(t) + \alpha_2\ \beta_2 h_2(t) & 0 & 0 \\ \alpha_1\ \beta_1\ h_1(t) + \alpha_2\ \beta_2 h_2(t) & \beta_1^2 h_1(t) + \beta_2^2 h_2(t) & 0 & 0 \\ 0 & 0 & \xi_1^2 & 0 \\ 0 & 0 & 0 & \xi_2^2 \end{pmatrix}
$$

We see that the variances of the initial risk factors, $X(t)$, are now linear combinations of mean-reverting processes and are thus stochastic. The problem is that the AJD specifications do not allow that their correlation remain constant. Their correlation, $\rho(t)$, is indeed now given by the expression

$$
\rho(t) = \frac{\alpha_1\ \beta_1\ h_1(t) + \alpha_2\ \beta_2 h_2(t)}{\sqrt{\alpha_1^2 h_1(t) + \alpha_2^2 h_2(t)}\sqrt{\beta_1^2 h_1(t) + \beta_2^2 h_2(t)}} \tag{2.24}
$$

If we set $\alpha_2 = 0$ and $\beta_1 = 0$, such that $h_1(t)$ and $h_2(t)$ can be identified as the instantaneous variance of $X_1(t)$ and $X_2(t)$, respectively; then their correlation drops to zero. There is no other way to introduce a correlation between risk-factors $(X(t))$ with stochastic volatility than letting their instantaneous variance be linear combinations of the same state-variables $(h(t))$, implying a correlation that also becomes stochastic.

Note, however, that if we set $\alpha_2 = 0$ and $\beta_2 = 0$, such that the instantaneous variance of $X_1(t)$ and $X_2(t)$ are both proportional to $h_1(t)$, then their correlation becomes one, which is constant.

2.3.5 Modelling Jumps and Spikes

In the three models studied above, we focused on the diffusion part of the spot price dynamics. As was already mentioned above, energy (and commodity) prices, however, show some jumps and spikes in their dynamics.

Jumps are defined here as sudden, non-continuous, price changes in the spot price. Spikes are the combinations of two jumps of same magnitude, but opposite directions. That is, when a positive jump occurs, often because of a plant or transmission line break-down, a negative jump of the same amplitude occurs as soon as the break-down is repaired and the system comes back to a normal regime.

Intuitively, it is easy to see that Affine Jump Diffusions are well equipped to model jumps by adding a jump process to the risk-factors' stochastic differential equation. The spiking behavior, however, is not as direct to model. The problem is that if we use the above type of models, the speed of reversion of the risk-factors is typically too little to pull back the level of the spot price to a normal level, once a jump has occurred, in a time period that is short enough. On the other hand, if the speed of the reversion of the risk-factors is enhanced to adequately represent the spiking behavior, then this level of reversion rate will be too large to adequately model the diffusion dynamics of the risk-factors. The usual way to model spikes is thus to add an additional risk factor with only a jump component (no diffusion) and a very strong mean-reverting rate. That is, we add to the spot price equation (common to all three models studied above) an additional jump factor:

$$Spot_t = Weekly\_Pattern(t) + \theta(t) + X_1(t) + X_t^J \tag{2.25}$$

$$dX_t^J \quad = \quad -\kappa_J X_t^J + dJ_t \tag{2.26}$$

In this model, the jumps induced by X_t^J are brought back to zero through an exponentially decaying continuous function, with a decay factor set by the (usually very high) speed of reversion factor κ_J. The problem with this type of models is that the exponential decay function does not really reflect what is observed on the markets. Instead, the markets show jumps in the spot price levels that stay for a moment (the time required to fix the break-down) and then jump back to a more "normal" level.

We will thus propose a spiking model that can reproduce this kind of dynamics. As previously, we add to the spot price equation spiking variables. This time, however, the spiking levels are divided into n discrete levels, and a spiking variable is associated to each discrete level :

$$Spot_t \quad = \quad Weekly\_Pattern(t) + \theta(t) + W_t + \sum_{i=1}^{n} X_i^J(t) \tag{2.27}$$

The dynamics of each spiking variable are then modelled as the combination of two jump processes:

$$dX_i^J(t) = dJ_i^{up}(t) + dJ_i^{down}(t) \tag{2.28}$$

$$\tag{2.29}$$

where

- $J_i^{up}(t)$ and $J_i^{down}(t)$ are pure jump processes with a constant jump size, $Spike_i$ and $-Spike_i$, respectively;
- $J_i^{up}(t)$ and $J_i^{down}(t)$ have arrival rate equal to

$$\lambda_i^{up}(t) \quad = \quad l_i^{up}(t)\left(1 - X^J(t).(Spike)^{-1}\right) \tag{2.30}$$

$$\lambda_i^{down}(t) \quad = \quad l_i^{down}(t)X^J(t).(Spike)^{-1} \tag{2.31}$$

where X^J is the vector of jump risk factors, and $Spike$ is the vector of discrete levels of the spikes.

It is easy to see that this model remains an AJD model, and the empirical spiking behavior is well reproduced. More precisely, the outcome of such a model is that there will only be at most one jump at a time. The arrival rates of the jumps (fixed by the coefficients $l_i^{up}(t)$), and the time during which the spikes remain (fixed by the coefficients $l_i^{down}(t)$) can be different according to the spiking level.

2.4 The risk-neutral probability measure

2.4.1 Introduction

As mentioned in the introduction of this chapter, the most common practice, up to now, was to consider as starting point the spot price stochastic process only. When pricing derivatives, the price(s) of risk was (were) then either deduced from a no arbitrage argument if the market was complete or calibrated to match a forward curve level if the market was not complete.

Another approach that is sometimes found is to calibrate the parameters of the stochastic process in the risk-neutral probability measure directly by matching a forward curve and some option prices.

The theory behind the link between the actual probability distributions of actual asset prices and their equivalent probability measures is now well established in a diffusive framework. A good reference on the topic is Duffie (1996) [14]. In her paper, Pan (2001) addresses the issue for a Jump-diffusion framework but remains in a Gaussian setting.

In this section, we expand her model to a more general framework and show the restrictions that guarantee that our process remains an Affine Jump Diffusion under the equivalent probability measure implied by some risk-premiums.

The simultaneous estimation of the AJD parameters under the physical probability measure and the risk-premiums will be discussed in Chapter 7.

2.4.2 The model

In what follows, prices of futures at time t with maturity T will be noted $f(t, T)$.

Definition 2.1 Equivalent probability measures. *Two probability measures Q and P are equivalent probability measures on $(\Omega, \mathcal{F})$ if, for any event A, $P(A) = 0$ if and only if $Q(A) = 0$.*

Definition 2.2 Martingales. *A process M_t is a martingale under the probability measure Q if for any t, $M_t = E^Q[M_T|\mathcal{F}_t]$, for any $T > t$.*

Without going into the details of the arbitrage theory, we will just assume here the absence of arbitrage opportunities, and the existence of an equivalent probability measure Q, under which futures prices are martingales. The interested reader is referred to Musiela and Rutowski(1998)[31], for example, for a detailed discussion of the conditions under which the absence of arbitrage opportunity implies the existence of such a martingale measure. However, the existence of such a martingale measure suffices to imply the absence of arbitrage opportunities.

Furthermore, quite trivially, the absence of arbitrage opportunity implies that futures prices must, at their maturity, converge towards the spot price, that is, $f(T,T) = S_T$.

As a result, we have that futures prices must be the expectation of the spot price at their maturity under an equivalent probability measure Q. This indeed derives directly from the property of repeated expectations.

Proposition 2.1 (From Bayes' Rule): *if P and Q are equivalent probability measures, then there exists a strictly positive martingale ξ such that, for any $\mathcal{F}_t$-measurable random variable, W_t:*

$$E^Q\left[W_T|\mathcal{F}_t\right] = \frac{E^P\left[W_T\xi_T|\mathcal{F}_t\right]}{\xi_t} \tag{2.32}$$

This strictly positive martingale ξ is often referred to as the density process of Q with respect to P.$\square$

Proof: See, for example, Duffie (1996) [14].

Making some arbitrary assumptions on the dynamics of the state price density process, and without going into any discussion of an underlying equilibrium model, we will now propose a given process for ξ with some restrictions on its parameters. In **Proposition 2.2** we show that this process is indeed a strictly positive martingale and that it is thus a suitable candidate to define an equivalent probability measure from **Proposition 2.1**. Next, in **Proposition 2.3**, we will show that an AJD in a physical probability measure P remains an AJD in the equivalent probability measure Q implied by ξ.

Having an AJD process in the equivalent probability measure Q is crucial to be able to use the pricing engines developed in Chapter 3. Furthermore, conserving the AJD property in both probability measures Q and P will be very important in Chapter 7, when developing a procedure for the simultaneous estimation of the parameters of the AJD process under the physical probability measure P and parameters of the process ξ defining the equivalent probability measure Q.

Proposition 2.2 *The process ξ_t solving the stochastic differential equation*

$$dln(\xi_t) = (k_0 + k_1.X_t)dt + \eta_1.dW_t + dZ_t^* \tag{2.33}$$

where

(a) *Z_t^* is a jump process whose jumps arrivals are completely determined by those of Z_t in equation (2.2). When there is a jump of size z on Z_t, then there is a jump of size z_ξ on Z_t^*.*

(b) *$\sigma\eta_1 = h_0 + h_1 X_t$, with $h_0 \in \mathbb{R}^n$, and $h_1 \in \mathbb{R}^{n\times n}$;*

(c) *$\eta_1^T\eta_1 = g_0 + g_1.X_t$, with $g_0 \in \mathbb{R}$, and $g_1 \in \mathbb{R}^n$;*

where in these two last constraints, and in what follows, the eventual dependence of $\eta_1(X_t)$ and $\sigma(X_t)$ on X_t is not written explicitely for some notational facilities. These two last constraints imply that if σ is a function of X_t, then η_1 is also a function of X_t such that (b) and (c) are satisfied. Furthermore, if σ is not function of X_t (such that $\sigma\sigma^T = H_0$) , then satisfying (b) and (c) imply that η_1 cannot be function of X_t neither. These constraints are needed in **Proposition 2.3** *to ensure that X_t remains an AJD under the equivalent probability measure Q implied by the process ξ_t.*

(d) $k_0 = -\frac{1}{2}g_0 - l_0(\theta^*([0,\ldots,0,1]) - 1)$;

(e) $k_1 = -\frac{1}{2}g_1 - l_1(\theta^*([0,\ldots,0,1]) - 1)$;

(f) $\theta^*(c) = \int_{\mathbb{R}^{n+1}} e^{c.[z',z_\xi]'} d\nu^{**}(\{z, z_\xi\})$, *with* $c \in \mathbb{C}^{n+1}$

(g) $\nu^{**}(\{z, z_\xi\})$ *is the joint probability distribution function of the jumps of* Z_t *and* Z_t^*

is a strictly positive martingale under P and is thus an admissible candidate to define an equivalent probability measure Q through equation (2.32).

Proof : Consider the augmented vector $Y_t = \begin{pmatrix} X_t \\ ln(\xi_t) \end{pmatrix}$. We have that :

$$\begin{aligned} dY_t = d\begin{pmatrix} X_t \\ ln(\xi_t) \end{pmatrix} &= \left(\begin{pmatrix} K_0 \\ k_0 \end{pmatrix} + \begin{pmatrix} K_1 & \mathbf{0} \\ k_1^T & 0 \end{pmatrix} Y_t\right) dt + \begin{pmatrix} \sigma \\ \eta_1^T \end{pmatrix} dW_t \\ &\quad + \begin{pmatrix} dZ_t \\ dZ_t^* \end{pmatrix} \end{aligned} \tag{2.34}$$

From Duffie, Pan and Singleton (1999) [16], we know that for any $a \in \mathbb{C}^{n+1}$,

$$E^P[e^{a.Y_T}|\mathcal{F}_t] = e^{\alpha(t)+\beta(t).Y_t} \tag{2.35}$$

where :

$$\begin{cases} \dot{\beta}_t = -\begin{pmatrix} K_1 & \mathbf{0} \\ k_1^T & 0 \end{pmatrix}^T \beta_t - \frac{1}{2}\beta_t^T H_1^* \beta_t - [l_1^T, 0]^T(\theta^*(\beta_t) - 1) \\ \dot{\alpha}_t = -\begin{pmatrix} K_0 \\ k_0 \end{pmatrix}^T \beta_t - \frac{1}{2}\beta_t^T H_0^* \beta_t - l_0(\theta^*(\beta_t) - 1) \end{cases} \tag{2.36}$$

where, $c^T H_1 c$ denotes the vector in C^{n+1} with k-th element $\sum_{i,j} c_i (H_1)_{i,j,k} c_j$; and with boundary conditions

$$\alpha(T) = 0 \tag{2.37}$$

$$\beta(T) = a \tag{2.38}$$

and where

$$H_0^* = \begin{pmatrix} H_0 & h_0 \\ h_0^T & g_0 \end{pmatrix} \tag{2.39}$$

$$\begin{array}{rcll} (H_1^*)_{i,j,k} & = & (H_1)_{i,j,k} & for \quad i,j,k = 1 \ldots n \\ (H_1^*)_{i,j,k} & = & 0 & for \quad i,j = 1 \ldots n+1; k = n+1 \\ (H_1^*)_{i,j,k} & = & (h_1)_{i,k} & for \quad i,k = 1 \ldots n; j = n+1 \\ (H_1^*)_{i,j,k} & = & (h_1)_{j,k} & for \quad j,k = 1 \ldots n; i = n+1 \\ (H_1^*)_{i,j,k} & = & (g_1)_k & for \quad i,j = n+1; k = 1 \ldots n \end{array} \tag{2.40}$$

So that we have :

$$E^P[\xi_T|\mathcal{F}_t] = E^P[e^{[0,\ldots,0,1]Y_T}|\mathcal{F}_t] = e^{\alpha(t)+\beta(t).Y_t} \tag{2.41}$$

From the definition of a martingale, ξ_t will be a martingale if we have

$$E^P[\xi_T|\mathcal{F}_t] = E^P[e^{[0,\ldots,0,1]Y_T}|\mathcal{F}_t] = e^{\alpha(t)+\beta(t).Y_t} = \xi_t \tag{2.42}$$

This will be the case if $\alpha(t) = 0$ and $\beta(t) = [0,\ldots,0,1]^T$ which, in turn, from the boundary conditions (2.37) and (2.38) on α and β, will be the case if $\dot{\alpha}_t = 0$ and $\dot{\beta}_t = 0$.

From the system of equations (2.36), $\dot{\alpha}_t = 0$ and $\dot{\beta}_t = 0$ if

$$k_0 = -\frac{1}{2}g_0 - l_0(\theta^*([0,\ldots,0,1]) - 1) \tag{2.43}$$

$$k_1 = -\frac{1}{2}g_1 - l_1(\theta^*([0,\ldots,0,1]) - 1) \tag{2.44}$$

which is indeed the case. The strictly positive character of ξ is guaranteed by the fact that the stochastic process is written in terms of the logarithm of ξ.$\square$

Furthermore, the process ξ_t defined by (2.33) prices all sources of uncertainty. The diffusion risk is priced by the vector η_1, and the jump risk is priced by the probability distribution of z_ξ.

As mentioned earlier, the form of the vector η_1 is however constrained by the form of the volatility matrix σ through the equations

- $\sigma\eta_1 = h_0 + h_1 X_t$, with $h_0 \in \mathbb{R}^n$, and $h_1 \in \mathbb{R}^{n\times n}$;
- $\eta_1^T\eta_1 = g_0 + g_1.X_t$, with $g_0 \in \mathbb{R}$, and $g_1 \in \mathbb{R}^n$;

This will be illustrated in the following example.

Consider an asset S_t with stochastic volatility

$$\begin{aligned} d\begin{pmatrix} ln(S_t) \\ \sigma_t^2 \end{pmatrix} &= \left(\begin{pmatrix} \mu \\ \kappa\bar{\theta} \end{pmatrix} + \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & -\kappa \end{pmatrix}\begin{pmatrix} ln(S_t) \\ \sigma_t^2 \end{pmatrix}\right)dt \\ &\quad + \begin{pmatrix} \sigma_t & 0 \\ 0 & \omega\sigma_t \end{pmatrix} dW_t \end{aligned} \tag{2.45}$$

The vector of prices of risk η_1 is here constrained to the form

$$\eta_1 = \begin{pmatrix} \eta_{1,1}\sigma_t \\ \eta_{1,2}\sigma_t \end{pmatrix} \tag{2.46}$$

so that the two constraints are indeed satisfied :

$$\begin{pmatrix} \sigma_t & 0 \\ 0 & \omega\sigma_t \end{pmatrix}\begin{pmatrix} \eta_{1,1}\sigma_t \\ \eta_{1,2}\sigma_t \end{pmatrix} = \begin{pmatrix} 0 & \eta_{1,1} \\ 0 & \omega\eta_{1,2} \end{pmatrix}\begin{pmatrix} ln(S_t) \\ \sigma_t^2 \end{pmatrix} \tag{2.47}$$

and

$$\begin{pmatrix} \eta_{1,1}\sigma_t & \eta_{1,2}\sigma_t \end{pmatrix} \begin{pmatrix} \eta_{1,1}\sigma_t \\ \eta_{1,2}\sigma_t \end{pmatrix} = \begin{pmatrix} 0 & \eta_{1,1}^2 + \eta_{1,2}^2 \end{pmatrix} \begin{pmatrix} ln(S_t) \\ \sigma_t^2 \end{pmatrix} \tag{2.48}$$

That is, the price of risk is forced to be a linear function of the volatility.

Proposition 2.3 *Let X_t be an AJD vector and ξ an AJD scalar whose dynamics in the probability measure P are respectively given by the stochastic differential equations (2.2) and (2.33). Let Q be the equivalent probability measure implied by ξ through the equation (2.32). Then X_t remains an AJD in Q.*

Proof : A random variable or vector is completely determined by its characteristic function. Thus, all we need to prove is that there exists an AJD process which implies the desired characteristic function of X_t in Q.

In the probability measure Q, the characteristic function of X_T conditional on the information available at time t is given by :

$$\begin{aligned} \chi(u,t,T) &= E^Q\left[e^{iu.X_T}|\mathcal{F}_t\right] && (2.49)\\ &= \frac{E^P\left[e^{iu.X_T+ln(\xi_T)}|\mathcal{F}_t\right]}{\xi_t} && (2.50)\\ &= \frac{e^{\alpha(t)+\beta(t).Y_t}}{\xi_t} && (2.51) \end{aligned}$$

where α and β are the solutions of the complex valued system of ODE's (2.36) subject to the boundary conditions

$$\begin{aligned} \alpha(T) &= 0 && (2.52)\\ \beta(T) &= [iu^T, 1]^T && (2.53) \end{aligned}$$

We observe that the $(n+1)^{th}$ element of $\dot{\beta}_t$ is zero for any t, so that $\beta_{n+1}(t) = 1$.

Equation (2.51) then reduces to

$$\chi(u,t,T) = e^{\alpha(t)+[\beta_1(t),\ldots,\beta_n(t)]X_t} \tag{2.54}$$

and the complex valued system of ODE's (2.36) reduces to

$$\begin{cases} \begin{pmatrix} \dot\beta_1(t) \\ \cdots \\ \dot\beta_n(t) \\ 0 \end{pmatrix} & = -\begin{pmatrix} K_1^T & k_1 \\ \mathbf{0}^T & 0 \end{pmatrix} \begin{pmatrix} \beta_1(t) \\ \cdots \\ \beta_n(t) \\ 1 \end{pmatrix} \\ & -\frac{1}{2} \begin{pmatrix} \beta_1(t) \\ \cdots \\ \beta_n(t) \\ 1 \end{pmatrix}^T H_1^* \begin{pmatrix} \beta_1(t) \\ \cdots \\ \beta_n(t) \\ 1 \end{pmatrix} \\ & -\begin{pmatrix} l_1 \\ 0 \end{pmatrix} (\theta^*([\beta_1(t), \ldots, \beta_n(t), 1]) - 1) \\ \dot\alpha_t & = -\left(K_0^T, k_0\right) \begin{pmatrix} \beta_1(t) \\ \cdots \\ \beta_n(t) \\ 1 \end{pmatrix} \\ & -\frac{1}{2} \begin{pmatrix} \beta_1(t) \\ \cdots \\ \beta_n(t) \\ 1 \end{pmatrix}^T \begin{pmatrix} H_0 & h_0 \\ h_0^T & g_0 \end{pmatrix} \begin{pmatrix} \beta_1(t) \\ \cdots \\ \beta_n(t) \\ 1 \end{pmatrix} \\ & -l_0(\theta^*([\beta_1(t), \ldots, \beta_n(t), 1]) - 1) \end{cases} \tag{2.55}$$

which can be rewritten as

$$
\left\{
\begin{array}{rcl}
\begin{pmatrix} \dot\beta_1(t) \\ \dots \\ \dot\beta_n(t) \end{pmatrix} & = & -K_1^T \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix} + l_1\left(\theta^*([0,\dots,0,1]) - 1\right) \\
 & & -\frac{1}{2} \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix}^T H_1 \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix} - h_1^T \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix} \\
 & & -l_1(\theta^*([\beta_1(t),\dots,\beta_n(t),1]) - 1) \\
\dot\alpha_t & = & -K_0^T \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix} + l_0\left(\theta^*([0,\dots,0,1]) - 1\right) \\
 & & -\frac{1}{2} \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix}^T H_0 \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix} - h_0^T \begin{pmatrix} \beta_1(t) \\ \dots \\ \beta_n(t) \end{pmatrix} \\
 & & -l_0(\theta^*([\beta_1(t),\dots,\beta_n(t),1]) - 1)
\end{array}
\right. \tag{2.56}
$$

with the boundary conditions

$$\alpha(T) = 0 \tag{2.57}$$

$$\begin{pmatrix} \beta_1(T) \\ \dots \\ \beta_n(T) \end{pmatrix} = iu \tag{2.58}$$

Consider, on the other hand, the following AJD stochastic process candidate to describe the dynamics of X_t in Q,

$$dX_t = \left(\tilde K_0(t) + \tilde K_1(t) X_t\right)\ dt + \tilde\sigma(X_t, t)\ d\tilde W_t + d\tilde Z_t \tag{2.59}$$

where :

- $\tilde W$ is a Brownian motion in $\mathbb{R}^n$;

- $\tilde{\sigma}(X_t, t)^T \tilde{\sigma}(X_t, t) = \tilde{H}_0 + \tilde{H}_1 X_t$
- $\tilde{Z}$ is a pure jump process whose jumps have a fixed probability distribution $\tilde{\nu}$ on $\mathbb{R}^n$ and arrive with intensity $\tilde{\lambda}(X_t, t) = \tilde{l}_0 + \tilde{l}_1.X_t$.

and define the function $\tilde{\theta}(c)$ with $c \in \mathbb{C}^n$

$$\tilde{\theta}(c) = \int_{\mathbb{R}^n} e^{c.z} d\tilde{\nu}(z) \tag{2.60}$$

This stochastic process (2.59) will indeed be a successful candidate if we can choose the values of its parameters such that the conditional characteristic function that it implies, $\tilde{\chi}(u, t, T)$, matches the conditional characteristic function $\chi(u, t, T)$. Again, we can write,

$$\tilde{\chi}(u, t, T) = E^Q \left[e^{iu.X_T} | \mathcal{F}_t \right] \tag{2.61}$$

$$= e^{\tilde{\alpha}(t) + \tilde{\beta}(t).X_t} \tag{2.62}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are the solutions of a system of complex valued ODE's

$$\begin{cases} \dot{\tilde{\beta}}_t = -\tilde{K}_1^T \tilde{\beta}_t - \frac{1}{2} \tilde{\beta}_t^T \tilde{H}_1 \tilde{\beta}_t - \tilde{l}_1 (\tilde{\theta}(\tilde{\beta}_t) - 1) \\ \dot{\tilde{\alpha}}_t = -\tilde{K}_0^T \tilde{\beta}_t - \frac{1}{2} \tilde{\beta}_t^T \tilde{H}_0 \tilde{\beta}_t - \tilde{l}_0 (\tilde{\theta}(\tilde{\beta}_t) - 1) \end{cases} \tag{2.63}$$

subjected to the boundary conditions

$$\tilde{\alpha}(T) = 0 \tag{2.64}$$

$$\tilde{\beta}(T) = iu \tag{2.65}$$

We observe that the two conditional characteristic functions do indeed match, that is $\tilde{\chi}(u, t, T) = \chi(u, t, T)$, if and only if

- $\tilde{K}_0 = K_0 + h_0$
- $\tilde{K}_1 = K_1 + h_1$

- $\tilde{H}_0 = H_0$
- $\tilde{H}_1 = H_1$
- $\tilde{l}_0(\tilde{\theta}(c) - 1) = l_0(\theta^*([c^T, 1]^T) - \theta^*([0, \ldots, 0, 1]^T))$
- $\tilde{l}_1(\tilde{\theta}(c) - 1) = l_1(\theta^*([c^T, 1]^T) - \theta^*([0, \ldots, 0, 1]^T))$

These two last constraints imply some relationships between the arrival rates and the probability density distribution of the jumps in both probability measures.

Indeed, defining the constant ω,

$$\omega = \int_{\mathbb{R}} e^{z_\xi} d\nu_\xi(z_\xi) \tag{2.66}$$

where $\nu_\xi(z_\xi)$ is the probability distribution function of the jumps of Z_t^*, the two last conditions then imply that

$$\frac{\tilde{l}_0}{l_0} = \frac{(\tilde{l}_1)_j}{(l_1)_j} = \omega \qquad\qquad \forall j = 1, \ldots, n \tag{2.67}$$

and that

$$\int_{\mathbb{R}^n} e^{c.z} d\tilde{\nu}(z) = \frac{1}{\omega} \int_{\mathbb{R}^{n+1}} e^{c.z + z_\xi} d\nu^{**}(\{z', z_\xi\}') \tag{2.68}$$

Finally, this last equation is equivalent to writing that

$$\tilde{f}(z) = \frac{1}{\omega} \int_{\mathbb{R}} e^{z_\xi} f^{m\xi}(z_\xi | z) f(z) dz_\xi \tag{2.69}$$

where $\tilde{f}(z)$ is the probability density function of the jumps $\tilde{Z}$ in Q, $f(z)$ is the probability density function of the jumps Z in P, and $f^{m\xi}(z_\xi|z)$ is the marginal probability density function of Z^* in P conditional on a jump of size z of Z.

This completes the proof.□

We observe that the coefficient of the affine function defining the instantaneous variance of the AJD processes under the probability measures P and Q are the same, i.e., $H_0 = \tilde{H}_0$ and $H_1 = \tilde{H}_1$.

This, however, does not necessarily imply that the variance of an underlying S at a given time T conditional on the information available at time t , with $t < T$, will be the same under the two probability measures. If the instantaneous variance of the underlying, σ_t^2 is itself an element of the vector of state variables X_t, that is, if we have a stochastic volatility model like in the example presented earlier, then the risk premium on the volatility , $\eta_{1,2}\sigma_t$, will make the mean of reversion of the instantaneous variance either higher or lower (according to the sign of $\eta_{1,2}$) in the probability measure Q compared to its mean in the probability measure P. The conditional variances of the underlying can thus be different (either higher or lower) in both probability measures.

From equation (2.67), we observe that the mean arrival rate of the jumps are the same in both probability measures P and Q if and only if $\omega = 1$.

Furthermore, from equation (2.69) if the jumps of Z_t^* are independent of the jumps of Z_t, then the probability distributions of the jumps of Z_t (in P) and of $\tilde{Z}_t$ (in Q) will be the same.

Several particular cases can be considered for the distribution of the jumps of Z_t^*. Pan (2001) [34] has, for example, recently stated that when the jumps of Z_t and Z_t^* are normally distributed, one can keep the arrival rates of the jumps in both probability measures equal by restricting the mean, μ_ξ, and the standard deviation σ_ξ of the jumps of Z_t^* such that

$$\mu_\xi = -\frac{\sigma_\xi^2}{2} \tag{2.70}$$

It is indeed easy to show that in this case, $\omega = 1$. Furthermore, she considers the jumps of Z_t to be normally distributed with mean μ_J and covariance matrix Cov_J, and she considers the covariance matrix of the joint distribution of the jumps of Z_t and Z_t^* to be equal to Σ^*. She then states, and one can see from equation (2.68), that the jumps of $\tilde{Z}_t$ are normally distributed with mean $\tilde{\mu}_J = \mu_J + \rho\sigma_J\sigma_\xi$, where $\rho\sigma_J\sigma_\xi$ is the vector of the n first elements of the last column of Σ^*.

Another particular case is to let the size of the jumps of Z_t^* be completely determined by the size of the jumps of Z_t and by a vector $\eta_2 \in \mathbb{R}^n$ such that when a jump of size z, $z \in \mathbb{R}^n$, occurs on Z_t, then a jump of size $z.\eta_2$ occurs on Z_t^*.

In this case, the factor ω is given by

$$\omega = \int_{\mathbb{R}^n} e^{\eta_2 . z} d\nu(z) \tag{2.71}$$

and the arrival rates of the jumps in both probability measures are thus modified.

The probability density functions are also modified such that

$$\tilde{f}(z) = \frac{1}{\omega} e^{\eta_2 . z} f(z) \tag{2.72}$$

as can be derived from equation (2.68). This is a convenient way to price both the risk associated to the arrival of the jumps and the risk associated with the distribution of the jumps when they occur through a single vector η_2. Furthermore, this framework can be applied for any jump distribution of Z_t.

We will thus further adopt this framework to price the jump risk. As a result, for any vectors of prices of risk η_1 and η_2 (with the constraints on the form of η_1 presented earlier), if the futures prices are set equal to the conditional expectation of the spot price under the equivalent probability measure Q defined by (2.32), then our model will be arbitrage-free. Furthermore, if the vector of risk-factors X_t is an AJD in the physical probability measure P, then it remains an AJD in the equivalent probability measure Q.

In Chapter 7, we show how to estimate η_1 and η_2 so that the futures prices computed as explained above best match the historical futures prices times series. We will further refer to the equivalent probability measure Q implied by these estimated η_1 and η_2 as the "risk-neutral" probability measure.

Chapter 3

Option Pricing: A Novel General Framework

3.1 Introduction

A conjunction of several factors have created the conditions for explosive growth in energy-related derivatives. First financial institutions, facing some difficulties in creating and expanding new lines of business, have shown a growing interest in trading energy commodities and financial and physical options related to energy.

Next, and more importantly, producers and end-users of energy are searching for risk-management tools to mitigate the risks they are now facing since the deregulation of the energy markets.

Finally, the size of the energy markets suggest that they can support the infrastructure necessary for derivative transactions of considerable volume, the size of the markets also creates enough room for a large number of market makers.

In their paper, Kaminski et al. (1999) [27] have defined as "exotic", options with assumptions that differ from the traditional options originally studied by Black and Scholes (1973) [3], Merton (1973) [30] and Cox, Ross and Rubinstein (1979) [8]. These early models apply to call and put options with payoffs that are defined as the difference between the price of the underlying instrument at time of maturity and the strike price or zero if it makes no sense to exercise. Furthermore, the dynamics of the price of the underlying are assumed to follow a Geometric Brownian Motion (GBM).

Exotic options can differ in terms of the assumptions regarding the dynamics of the price of the underlying (as discussed in Chapter 2) and in terms of the payoffs considered.

Kaminski et al. (1999) [27] have identified several classes of exotic options of interest in the energy industry:

- **Path-dependent options**, the most popular of which are the **Asian options** with payoffs that are function of averages of the underlying price during a certain period. This is due to the fact that energy contracts are often indexed or refer to averages of prices during a given period so that market players cannot take advantage of the relative lack of liquidity of some energy markets to exert some market power and distort the price for a short period of time at maturity of the option.

- **Multi-commodity options** refer to options with payoffs that are function of the prices of more than one underlying. These options have been extensively used because of the exposure of the different actors in the energy industry to several sources of uncertainty on top of energy prices such as other commodity prices (i.e., aluminium, copper,...), some fuel prices (i.e., natural gas, fuel oil, coal,...), interest rates,... These options also include **Spread options** where the option is written on the difference between the prices of two commodities. More generally, spread options can be written on the difference between the prices of the same commodity at two different locations (location spreads), the prices of the same commodity at two different points in time (calendar spreads), the prices of inputs to and outputs from, a production process (production spreads or spark-spreads), prices of different grades of the same commodity (quality spreads). When the strike price of a spread option is set to zero, we get a so-called **Option to exchange one asset for another**. Basket options refer to options on the weighted average of the prices of a basket of different commodities. Finally, **Options on the minimum or the maximum of two commodities** are also a special case of multi-commodity options.

- **Compound options** are options written on options. A particular case of compound options is to be found in interruptible contracts with early notification further studied in details.

- **Hybrid options** are options that combine aspects of the exotics mentioned above. One of the most common practices in the energy markets is to settle against an average of prices over some time period. Asian options are also often combined with spread options to give rise to Asian spread options.

- **Take-or-pay options** are options where the buyer agrees to purchase between a minimum v and a maximum V units of a commodity (natural gas or electricity) every day or every hour for the length of the contract. The purchase (strike) price for the commodity is fixed at $K/unit. The total volume for the contract period has a minimum V_{min} and a maximum V_{max}. On the NordPool such take-or-pay options are called **Brukstid** contracts.

Pricing methodologies for these exotic options include Monte-Carlo methods, Grid methods (i.e., binomial or trinomial trees), and finite difference methods according to the type of exotic option to be priced.

All of these methodologies, however suffer some drawbacks

- Monte-Carlo methods are very easy to implement but are very computationally intensive. Furthermore, the accurate computation of the greeks of option prices is quite tedious. Finally, they are not well suited for options with endogenous exercise policies such as take or pay options.

- Grid methods are difficult to implement with stochastic processes departing from the pure Brownian motion. They also suffer computation burdens when the size of the vector of underlying state variables exceed two or three.

- Finite difference methods face the same problem of computation burden when the size of the underlying state variables vector increases.

As a result, in order to overcome these problems, researchers have traditionally devoted their time in finding closed form solutions to the pricing of some of these exotic options under the assumption that the underlying's dynamics was driven by a given stochastic process. However, each time the stochastic process was changed, the quest for a closed form solution to the pricing problem had to be restarted again.

A great step towards generality was achieved when attention has been given to characteristic functions of the underlying. Generalizing the work of Stein and Stein (1991)[47], and Heston (1993)[22], and adapting some results of Shephard (1991) [43], Bakshi and Madan (2000) [1] are the first to have used the power of characteristic functions for the pricing of contingent claims. They indeed reduced the option pricing problem to the estimation of Arrow Debreu securities prices under modified equivalent probability measures. These securities are easily priced for exponential or polynomial functions of the underlying variables because their characteristic functions are obtainable through simple translation or differentiation of the original underlying variables' characteristic function, respectively.

A similar approach has also been developed by Duffie, Pan and Singleton (1999) [15] in the special case of Affine Jump Diffusion (AJD) processes. They indeed showed that the characteristic function of a vector of AJDs could be computed by solving a system of complex valued ordinary differential equations and obtained the same pricing formulae as Bakshi and Madan (1999) [1] through an approach based on an Inverse Fourier Stieltjes Transform with respect to the (log) strike price.

Kammat and Oren (2000) [26] extended the work of Duffie, Pan and Singleton (1999) [15] to show that the joined characteristic function of the underlying variables at different times in the future could also be obtained by solving a system of complex valued ordinary differential equations. They then extended the work of Bakshi and Madan (2000) [1] to value options with payoffs that are functions of affine functions of the underlying at two different dates.

Finally, Carr and Madan (1999) [4] adapted the approach developed by Duffie, Pan and Singleton (1999) [15] to take advantage of the power of fast Fourier transforms algorithms to compute at low cost options' prices with different strike prices.

The novel methodology to price Contingent Claims developed in this chapter and its applications exposed in the Chapters 4 and 5 are from Culot, Menten de Horne and Smeers (2001) [10]. This methodology is designed to price any arbitrary payoff satisfying only weak constraints. In the particular cases where the payoff function can be expressed as the product of (a) a polynomial of the underlying variables, (b) an exponential affine function of the underlying variables and (c) any function of the underlying variables with an analytical Fourier transform, than closed form solutions are provided in Chapter 4. This covers all of the exotic options mentioned above but the take or pay kind of options. The results of Bakshi and Madan (1999) [1], Duffie, Pan and Singleton (1999) [15], and Kammat and Oren (2000) can, for example be shown to fall into this category. Similarly, the fast Fourier transforms introduced by Carr and Madan (1999) [4] to price options with different strike prices at low cost can also be retrieved from our framework.

When the payoff cannot be decomposed into such a product of functions, we propose, in Chapter 5, a numerical implementation of our methodology based on fast discrete cosine transforms. This provides an alternative to the approach developed by Van Steenkiste and Foresi (1999) [48]. Our approach is also of particular interest in dynamic programming algorithms, which are needed to price take or pay kind of options, and where the payoff is only known at some discrete points on a grid.

3.2 The theoretical framework

3.2.1 The concept

A review of notations, concepts and results about Fourier theory and distribution theory introduces the Parseval identity at the heart of the new methodology. Further details on distribution theory can be found in Willem (1995) [49] or Safarov (1996) [39].

3.2.1.1 The Schwartz space $\mathcal{S}(\mathbb{R}^n)$

Definition 3.1 *A multi-index $\gamma \in \mathbb{N}^n$ can be used as a short-hand in two cases:*

1. *if $x \in \mathbb{R}^n$, then x^γ is a short-hand for $x_1^{\gamma_1} \dots x_n^{\gamma_n}$*
2. *if $f : \mathbb{R}^n \to \mathbb{C}$, then $\partial_x^\gamma f(x)$ is a short-hand for $\frac{\partial^{\gamma_1} \dots \partial^{\gamma_n}}{\partial x_1^{\gamma_1} \dots \partial x_n^{\gamma_n}} f(x)$*

Definition 3.2 *A function $u : \mathbb{R}^n \to \mathbb{C}$ belongs to the Schwartz space $\mathcal{S}(\mathbb{R}^n)$, $u \in \mathcal{S}(\mathbb{R}^n)$, if $u \in \mathcal{C}^\infty$ and*

$$\|u\|_{\alpha,\beta} \triangleq \sup_{x \in \mathbb{R}^n} |x^\beta \partial_x^\alpha u(x)| < \infty \tag{3.1}$$

for all multi-indices α, β.

An example of a function belonging to the Schwartz space is $e^{-|x|^2}$. On the other hand, heavy-tailed probability density functions with tails going to zero as power functions do not belong to this space.

Definition 3.3 *Here is a list of continuous linear applications from $\mathcal{S}(\mathbb{R}^n)$ in itself:*

1. *The translations :* $\tau_a u(x) \triangleq u(x-a)$*, where* $a \in \mathbb{R}^n$*;*
2. *The dilatations/rotations :* $\partial_A u(x) \triangleq u(A^{-1}x)$*, where* $A \in \mathbb{R}^{n\times n}$ *is invertible; in particular, we note* $\partial_{-I} u \triangleq \check{u}$*;*
3. *The derivations:* $\partial_x^\alpha u(x)$*, where* $\alpha \in \mathbb{N}^n$*;*
4. *The Fourier Transform :* $\hat{u}(\omega) \triangleq \mathfrak{F}(u)(\omega) \triangleq \int_{\mathbb{R}^n} u(x)e^{-i2\pi x\omega}dx$*.*

Properties 3.1 *The following properties of Fourier transforms of* $u \in \mathcal{S}(\mathbb{R}^n)$*are standard results*

1. $\hat{\hat{u}} = \check{u}$ *;*
2. *The translations :* $\mathfrak{F}(\tau_a u) = e^{-2i\pi y.a}\hat{u}$*, and* $\tau_a \hat{u} = \mathfrak{F}\left(e^{2i\pi y.a}u\right)$*;*
3. *The dilatations/rotations :* $\mathfrak{F}(\partial_A u) = |\det(A)|\partial_{A^{-T}}\hat{u}$*, and* $\partial_A \hat{u} = |\det(A)|\mathfrak{F}\left(\partial_{A^{-T}}u\right)$*;*
4. *The derivations:* $\mathfrak{F}(\partial^\alpha u) = (2i\pi y)^\alpha \hat{u}$*, and* $\partial^\alpha \hat{u} = \mathfrak{F}\left((-2i\pi y)^\alpha u\right)$ *.*

3.2.1.2 The tempered distributions space $\mathcal{S}'(\mathbb{R}^n)$

The dual space of the Schwartz space is the tempered distributions space.

Definition 3.4 *The space of tempered distributions is defined as*

$$\mathcal{S}'(\mathbb{R}^n) \triangleq \{f : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C} : f \textit{ is linear and continuous}\} \tag{3.2}$$

The application of the tempered distribution f on the function $u \in \mathcal{S}(\mathbb{R}^n)$ is noted $\langle f, u \rangle$.

The next proposition defines an important subspace of $\mathcal{S}'(\mathbb{R}^n)$.

Proposition 3.1 *Let $f \in \mathcal{P}(\mathbb{R}^n)$ with*

$$\mathcal{P}(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{C} \mid (\exists m \in \mathbb{N})(\exists p \in [1, \infty[) : (1 + |x|^2)^{-m} f \in L^p(\mathbb{R}^n)\} \quad (3.3)$$

then the linear form

$$\mathcal{S}(\mathbb{R}^n) \to \mathbb{C} : u \mapsto \int_{\mathbb{R}^n} f(x)u(x)dx \quad (3.4)$$

defines a tempered distribution. We will associate to each function satisfying (3.3) its tempered distribution and we will note

$$\langle f, u \rangle \triangleq \int_{\mathbb{R}^n} f(x)u(x)dx \quad \forall f \in \mathcal{P}(\mathbb{R}^n) \quad (3.5)$$

Proof: See Willem (1995) [49], p. 145.

However, there exist tempered distributions that cannot be expressed in an integral form as in (3.5). A well known example being the dirac distribution

$$\forall u \in \mathcal{S}(\mathbb{R}^n) \quad \langle \delta, u \rangle \triangleq u(0) \quad (3.6)$$

The active variable will be subscripted when explicit forms of functions are used. For example, we will note $\langle f, e^{a.x}u \rangle_x$.

Definition 3.5 *The space of tempered functions is defined as*

$$\begin{aligned} \mathcal{O}_M(\mathbb{R}^n) \triangleq \{g \in \mathcal{C}^\infty(\mathbb{R}^n) : (\forall \gamma \in \mathbb{N}^n)(\exists c \geq 0)(\exists m \in \mathbb{N}^n) \\ (\forall x \in \mathbb{R}^n) : |\partial^\gamma g(x)| \leq c(1 + |x|^2)^m\} \end{aligned} \quad (3.7)$$

Definition 3.6 *The following list of continuous linear applications on $\mathcal{S}'(\mathbb{R}^n)$ are defined by transposition of continuous linear applications on $\mathcal{S}(\mathbb{R}^n)$:*

1. *The translations :* $\langle \tau_a f, u \rangle \triangleq \langle f, \tau_{-a} u \rangle$*, where* $a \in \mathbb{R}^n$*;*

2. *The dilatations/rotations:* $\langle \partial_A f, u \rangle \triangleq |\det(A)| \langle f, \partial_{A^{-1}} u \rangle$*, where* $A \in \mathbb{R}^{n \times n}$ *is invertible; in particular, we note* $\partial_{-I} f \triangleq \check{f}$*;*

3. *The derivations:* $\langle \partial^\alpha f, u \rangle \triangleq (-1)^{|\alpha|} \langle f, \partial^\alpha u \rangle$*, where* $\alpha \in \mathbb{N}^n$*;*

4. *The Fourier Transform (a.k.a. Parseval identity):* $\langle \hat{f}, u \rangle \triangleq \langle f, \hat{u} \rangle$.

5. *Multiplication by* $v \in \mathcal{O}_M(\mathbb{R}^n)$*:* $\langle v f, u \rangle \triangleq \langle f, v u \rangle \quad \forall u \in \mathcal{S}(\mathbb{R}^n), \forall v \in \mathcal{O}_M(\mathbb{R}^n)$

Example 3.1 *Some examples of Fourier transforms for a tempered distribution* $f \in \mathcal{S}'(\mathbb{R}^n)$*:*

1. $\hat{\hat{f}} = \check{f}$

2. $\check{\hat{f}} = \hat{\check{f}}$

3. *The translations :* $\mathfrak{F}(\tau_a f) = e^{-2i\pi y.a} \hat{f}$*, and* $\tau_a \hat{f} = \mathfrak{F}\left(e^{2i\pi y.a} f\right)$

4. *The dilatations/rotations :* $\mathfrak{F}(\partial_A f) = |\det(A)| \partial_{A^{-T}} \hat{f}$*, and* $\partial_A \hat{f} = |\det(A)| \mathfrak{F}\left(\partial_{A^{-T}} f\right)$

5. *The derivations:* $\mathfrak{F}(\partial^\alpha f) = (2i\pi y)^\alpha \hat{f}$*, and* $\partial^\alpha \hat{f} = \mathfrak{F}\left((-2i\pi y)^\alpha f\right)$*,*

6. *The Dirac distribution:* $\hat{1} = \delta, \quad \hat{\delta} = 1$.

3.2.1.3 The Parseval identity

The Parseval identity is the key result that opens a new perspective on the classical problem of European option pricing.

Indeed, this problem is essentially equivalent to the computation of the expectation of a nonlinear payoff f of a random variable with probability density function (p.d.f.) p or:

$$E[f(X)] = \int_{\mathbb{R}^n} f(x) p(x) dx \tag{3.8}$$

However, the p.d.f. has not often an analytical form and even a numerical form is tricky to obtain. On the contrary, the characteristic function $\chi(\omega) \triangleq E[e^{-i2\pi\omega.X}]$ of the random variable is available in a much wider number of cases.

The problem is that the payoff f is in the space domain whereas the characteristic function χ is defined in the frequency domain. A simple reasoning about the relation (3.8) will however bring an elegant solution to reconcile the two.

Theorem 3.1 *If $f \in \mathcal{P}(\mathbb{R}^n)$ and $p \in \mathcal{S}(\mathbb{R}^n)$, then*

$$E[f(X)] = \langle \hat{f}, \check{\chi} \rangle \tag{3.9}$$

where $\chi = \hat{p}$.

Proof: The integral form of the expectation leads to an expression with a distribution :

$$E[f(X)] = \langle f, p \rangle \tag{3.10}$$

and using properties of the Fourier transforms, it follows that:

$$\begin{aligned} E[f(X)] &= \langle f, p \rangle && (3.11)\\ &= \langle f, \check{\hat{\hat{p}}} \rangle \quad \text{by the Fourier transform Property 3.1-1} && (3.12)\\ &= \langle \widehat{f}, \check{\hat{p}} \rangle \quad \text{by the Parseval identity (Definition 3.6-4)} && (3.13) \end{aligned}$$

which is the desired result. $\square$

Example 3.2 *Moments from the characteristic function.*
The moment of order m of a random variable is defined as $E[X^m]$. The Theorem 3.1 and some algebra give the well-known result:

$$\begin{aligned} E[X^m] &= \langle x^m, p \rangle \\ &= \langle 1, x^m p \rangle \end{aligned}$$

$$
\begin{aligned}
&= \langle \hat{1}, \partial_{-I}\mathfrak{F}(x^m p)\rangle \\
&= \langle \delta, \partial_{-I}\left((-2i\pi)^{-m}\partial^m \hat{p}\right)\rangle \\
&= \langle \delta, (2i\pi)^{-m}\partial^m \check{p}\rangle \\
&= \frac{1}{(2i\pi)^m}\partial^m \chi(0)
\end{aligned} \tag{3.14}
$$

Example 3.3 *The cumulative probability distribution (c.p.d.) of an uni-dimensional random variable.*
The c.p.d. of X *is defined as* $P(a) = \int_{-\infty}^{a} p(x)dx$*, or equivalently as* $E[\mathbf{1}_{(x\leq a)}]$*. Using the Theorem 3.1 and the knowledge of the Fourier transform of the unitstep function* $\mathfrak{F}\left(\mathbf{1}_{(x\geq 0)}\right) = \frac{1}{2}\left(\frac{-i}{\pi\omega} + \delta\right)$*, it follows that:*

$$
\begin{aligned}
E[\mathbf{1}_{(x\leq a)}] &= \langle \mathbf{1}_{(x\leq a)}, p\rangle \\
&= \langle \tau_a \partial_{-1}\mathbf{1}_{(x\geq 0)}, p\rangle \\
&= \langle \mathbf{1}_{(x\geq 0)}, \partial_{-1}\tau_{-a} p\rangle \\
&= \left\langle \mathbf{1}_{(x\geq 0)}, \partial_{-1}\left(\widehat{\widehat{\partial_{-1}\tau_{-a}p}}\right)\right\rangle \\
&= \left\langle \mathbf{1}_{(x\geq 0)}, \widehat{\widehat{\tau_{-a}p}}\right\rangle \\
&= \langle \widehat{\mathbf{1}_{(x\geq 0)}}, \widehat{\tau_{-a}p}\rangle \\
&= \left\langle \frac{1}{2}\left(\frac{-i}{\pi\omega} + \delta\right), e^{2i\pi\omega a}\hat{p}\right\rangle_{\omega} \\
&= \frac{1}{2}\left(\chi(0) + \langle \frac{-i}{\pi\omega}, e^{2i\pi\omega a}\chi\rangle_{\omega}\right) \\
&= \frac{\chi(0)}{2} - \int_0^{\infty} \frac{Im[e^{2i\pi\omega a}\chi(\omega)]}{\pi\omega} d\omega
\end{aligned} \tag{3.15}
$$

3.2.2 The option pricing framework

The Parseval identity and its applications to compute expectations will now be applied more specifically to the pricing of options with a fixed exercise date.

We fix a probability space $(\Omega, \mathcal{F}, Q)$ and an information filtration $(\mathcal{F}_t) = \{\mathcal{F}_t : t \geq 0\}$, and suppose that X, a vector of underlying variables is a process in some sate space $D \subset \mathbb{R}^n$ defined on the filtration $(\mathcal{F}_t)$.

Let $f \in \mathcal{P}(\mathbb{R}^n)$, we will consider here payoff functions that can be expressed as

$$V(X_T) = f(X_T) \sum_i (c_i X_T^{\alpha_i}) e^{a.X_T} \tag{3.16}$$

with $c_i \in \mathbb{R}$, $a \in \mathbb{R}^n$, and where the α_i denote multi-indexes.

Let $r(X_t, t) : \{\mathbb{R}^n, \mathbb{R}_+\} \to \mathbb{R}$ denote the instantaneous risk-free rate of return and $B(t)$ be a risk-free bond with $dB_t = B_t r(X_t, t)dt$.

Let $g(\{X_T, B_T\}; t, T) : \{\mathbb{R}^{n+1}; \mathbb{R}_+, \mathbb{R}_+\} \to \mathbb{R}$ denote the joined probability density function, under the probability measure Q, of the random $(n+1)$-dimensional vector $\{X, B\}$ at time T conditional on the information available at time t, with $t \leq T$. The set of probability density functions will be restricted here to probability density functions satisfying the constraint that the product $g(\{X_T, B_T\}; t, T)e^{a.X_T}$ is in $\mathcal{S}(\mathbb{R}^n)$. The "discounted" probability density function will be noted

$$g^*_{t,T}(X_T) = \int_{\mathbb{R}} \frac{B_t}{y} g(\{X_T, y\}; t, T) dy \tag{3.17}$$

where y is an integration variable denoting possible outcomes of the stochastic variable B_T.

Theorem 3.2 *Assuming that Q is the risk-neutral probability measure, the price at time t, of a contingent claim with the payoff at time T as defined above is formally given by*

$$c(t) \equiv E^Q\left[e^{-\int_t^T r(X_s,s)ds} f(X_T) \sum_i (c_i X_T^{\alpha_i}) e^{a.X_T} | \mathcal{F}_t\right] \tag{3.18}$$

$$
\begin{aligned}
&= \sum_i c_i E^Q \left[e^{-\int_t^T r(X_s,s)ds} f(X_T) X_T^{\alpha_i} e^{a.X_T} | \mathcal{F}_t \right] &(3.19)\\
&= \sum_i c_i \langle \hat{f}, \frac{1}{(i2\pi)^{\alpha_i}} \partial^{\alpha_i} \mathfrak{F}\left(e^{-a.X_T} \check{g}^*_{t,T}(X_T) \right) \rangle &(3.20)
\end{aligned}
$$

Proof: Each expectation can be written as an integral

$$
\begin{aligned}
&E^Q \left[e^{-\int_t^T r(X_s,s)ds} f(X_T) X_T^{\alpha_i} e^{a.X_T} | \mathcal{F}_t \right] \\
&= \int_{\mathbb{R}^n} f(x) x^{\alpha_i} e^{a.X_T} \int_{\mathbb{R}} \frac{B_t}{y} g(\{x,y\}; t,T) dy dx \\
&= \int_{\mathbb{R}^n} f(x) x^{\alpha_i} e^{a.x} g^*_{t,T}(x) dx \\
&= \langle f, X_T^{\alpha_i} e^{a.X_T} g^*_{t,T} \rangle_{X_T} &(3.21)
\end{aligned}
$$

where x and y are integration variables denoting possible outcomes of the stochastic variables X_T and B_T, respectively.

From the Parseval identity (Definition 3.6-4), this is equal to

$$\langle f, X_T^{\alpha_i} e^{a.X_T} g^*_{t,T} \rangle_{X_T} = \langle \widehat{f}, \partial_{-I} \mathfrak{F}\left(X_T^{\alpha_i} e^{a.X_T} g^*_{t,T} \right) \rangle_{X_T} \tag{3.22}$$

which, from the **Properties 3.1** of translation and differentiation of the Fourier transforms and from the restriction set on the probability density function $g(\{X_t, B_t\}; t, T)$, it is trivial to show that

$$\langle \widehat{f}, \partial_{-I} \mathfrak{F}\left(X_T^{\alpha_i} e^{a.X_T} g^*_{t,T} \right) \rangle_{X_T} = \langle \widehat{f}, \frac{1}{(i2\pi)^{\alpha_i}} \partial^{\alpha_i} \mathfrak{F}\left(e^{-a.X_T} \check{g}^*_{t,T}(X_T) \right) \rangle \tag{3.23}$$

This concludes the proof $\square$.

Note that the expectation can only be computed if $e^{a.X_T} g^*_{t,T}(X_T) \in \mathcal{S}(\mathbb{R}^n)$. Furthermore, in the case $g^*_{t,T}$ is analytic, the result of the theorem can further be simplified into

$$c(t) \equiv \langle \widehat{f}, \partial_{-I} \mathfrak{F}\left(X_T^{\alpha_i} e^{a.X_T} g^*_{t,T} \right) \rangle_{X_T} = \langle \widehat{f}, \frac{1}{(i2\pi)^{\alpha_i}} \partial^{\alpha_i} \left(\check{\check{g}}^*_{t,T}(X_T) \right) (\omega + \frac{ia}{2\pi}) \rangle \tag{3.24}$$

3.2.3 Using Fast Fourier Transforms to compute options with different strike prices

As mentioned in the introduction, Carr and Madan (1999)[4] have developed an approach that will reduce the number of computations when evaluating European call or put option prices at different strike price levels by using Fast Fourier Transforms algorithms. We show in this section that these results can also be retrieved in our framework.

The setting introduced by Carr and Madan (1999)[4] assumes the following: let $C_T(k)$ be the desired value of a T maturity call option with log strike k. Let the risk-neutral density of the terminal log price s_T be $g_T(s)$. The characteristic function of this density is defined by:

$$\chi_T(u) \equiv \int_{-\infty}^{\infty} e^{-i2\pi us} g_T(s)ds \tag{3.25}$$

The value of the call is then given by:

$$C_T(k) \equiv e^{-rT} \int_{-\infty}^{\infty} \left(e^s - e^k\right) \mathbf{1}_{s\geq k}\, g_T(s)ds \tag{3.26}$$

where r denotes the risk-free rate (assumed constant).

Because Carr and Madan (1999)[4] restrict themselves to Fourier transforms of square integrable functions, they have to introduce a dampening factor. They thus compute the value of a "dampened" call price, $c_T(k) \in \mathcal{L}^2$ defined by:

$$c_T(k) \equiv e^{\alpha k} C_T(k) \qquad , \alpha > 0 \tag{3.27}$$

We show here that the same formula can be obtained in our framework for the dampened call price.

More interestingly, our framework only requires the payoff to be a distribution and not necessarily a square integrable function. Hence the development can be done without the dampening factor artefact.

3.2.3.1 Development with a dampening factor

Combining equations (3.26) and (3.26), the value of the dampened call price can indeed be expressed as

$$\begin{aligned} c_T(k) &= e^{-rT} \int_{-\infty}^{\infty} e^{\alpha k} \left(e^s - e^k\right) \mathbf{1}_{s \geq k}\, g_T(s) ds && (3.28) \\ &= e^{-rT} \left\langle e^{\alpha k} \left(e^s - e^k\right) \mathbf{1}_{s \geq k}, g_T(s) \right\rangle_s && (3.29) \\ &= e^{-rT} \left\langle e^{\alpha k - (\alpha+1)s} \left(e^s - e^k\right) \mathbf{1}_{s \geq k}, e^{(\alpha+1)s} g_T(s) \right\rangle_s && (3.30) \\ &= e^{-rT} \left\langle \left(e^{-\alpha(s-k)} - e^{-(\alpha+1)(s-k)}\right) \mathbf{1}_{s \geq k}, e^{(\alpha+1)s} g_T(s) \right\rangle_w && (3.31) \end{aligned}$$

which, from the Parseval Identity, becomes

$$c_T(k) = e^{-rT} \left\langle \partial_{-1} \mathfrak{F} \left(\left(e^{-\alpha(s-k)} - e^{-(\alpha+1)(s-k)}\right) \mathbf{1}_{s \geq k} \right), \mathfrak{F} \left(e^{(\alpha+1)s} g_T(s) \right) \right\rangle \quad (3.32)$$

Let us now develop the two sides of the application:

$$\mathfrak{F} \left(e^{(\alpha+1)s} g_T(s) \right)(w) = \chi_T(w - i(\alpha+1)) \quad (3.33)$$

and

$$\begin{aligned} &\partial_{-1} \mathfrak{F} \left(\left(e^{-\alpha(s-k)} - e^{-(\alpha+1)(s-k)}\right) \mathbf{1}_{s \geq k} \right)(w) \\ &\qquad = \int_k^{\infty} \left(e^{-\alpha(s-k)} - e^{-(\alpha+1)(s-k)}\right) e^{i2\pi w s} ds \\ &\qquad = e^{i2\pi w k} \left(-\frac{1}{i2\pi w - \alpha} + \frac{1}{i2\pi w - (\alpha+1)} \right) \\ &\qquad = \frac{e^{i2\pi w k}}{\alpha + \alpha^2 - i2\pi w(2\alpha+1) - (2\pi w)^2} && (3.34) \end{aligned}$$

Finally, we get

$$\begin{aligned} c_T(k) &= e^{-rT} \left\langle \frac{e^{i2\pi w k}}{\alpha + \alpha^2 - i2\pi w(2\alpha+1) - (2\pi w)^2}, \chi_T\left(w - i(\alpha+1)\right) \right\rangle_w \\ &= e^{-rT} \left\langle e^{i2\pi w k}, \frac{\chi_T\left(w - i(\alpha+1)\right)}{\alpha + \alpha^2 - i2\pi w(2\alpha+1) - (2\pi w)^2} \right\rangle_w \\ &= e^{-rT} \int_{-\infty}^{\infty} e^{i2\pi w k} \frac{\chi_T\left(w - i(\alpha+1)\right)}{\alpha + \alpha^2 - i2\pi w(2\alpha+1) - (2\pi w)^2} dw \\ &= e^{-rT} 2 \int_0^{\infty} Re\left[e^{i2\pi w k} \frac{\chi_T\left(w - i(\alpha+1)\right)}{\alpha + \alpha^2 - i2\pi w(2\alpha+1) - (2\pi w)^2} \right] dw && (3.35) \end{aligned}$$

which is just a direct Fourier transform and lends itself quite easily to an application of the FFT to compute its value for different k's.

Quite trivially, the values of the European calls, $C_T(k)$, are then retrieved by a simple application of equation (3.27). Furthermore, it is easy to show that the prices of European put options with different strike prices can also be computed along the same lines.

3.2.3.2 Development without a dampening factor

This time, we directly develop the expression of the call price:

$$\begin{aligned}
C_T(k) &\equiv e^{-rT}\int_{-\infty}^{\infty}\left(e^s - e^k\right)\mathbf{1}_{s\geq k}\, g_T(s)ds \\
&= \left\langle \left(e^{s} - e^k\right)\mathbf{1}_{s\geq k}, g_T(s)\right\rangle_s \\
&= \left\langle \left(1 - e^{k-s}\right)\mathbf{1}_{s\geq k}, e^{s} g_T(s)\right\rangle_s \\
&= \left\langle \tau_k\left(\left(1 - e^{-s}\right)\mathbf{1}_{s\geq 0}\right), e^{s} g_T(s)\right\rangle_s \\
&= \left\langle \mathfrak{F}\left(\tau_k\left(\left(1 - e^{-s}\right)\mathbf{1}_{s\geq 0}\right)\right), \partial_{-I}\mathfrak{F}\left(e^{s} g_T(s)\right)\right\rangle_w \\
&= \left\langle e^{-2i\pi kw}\left(\frac{\delta}{2} - \frac{i}{2\pi w} - \frac{i}{i-2\pi w}\right), \tau_{\frac{i}{2\pi}}\check{\hat{g}}_T\right\rangle_w \\
&= \frac{\hat{g}_T\left(\frac{i}{2\pi}\right)}{2} + \int_{-\infty}^{\infty} e^{-2i\pi kw}\left(-\frac{i}{2\pi w} - \frac{i}{i-2\pi w}\right)\hat{g}_T(-w + \frac{i}{2\pi})dw \\
&= \frac{\hat{g}_T\left(\frac{i}{2\pi}\right)}{2} + \int_{-\infty}^{\infty} e^{-2i\pi kw}\frac{\hat{g}_T(-w+\frac{i}{2\pi})}{i2\pi w - (2\pi w)^2}\,dw \\
&= \frac{\hat{g}_T\left(\frac{i}{2\pi}\right)}{2} + 2\int_0^{\infty} Re\left[e^{-2i\pi kw}\frac{\hat{g}_T(-w+\frac{i}{2\pi})}{i2\pi w - (2\pi w)^2}\right]dw
\end{aligned} \tag{3.36}$$

where the second term is a regular Fourier transform with respect to the log strike price k, which can be approximated by its discrete equivalent and computed via an FFT algorithm for different levels of the log strike price k.

3.2.3.3 The FFT algorithm

Both equations (3.35) and (3.36) have an expression of the type:

$$f_q(k) = \int_0^\infty Re\left[e^{(-1)^q 2i\pi kw} F_q(w)\right] dw \tag{3.37}$$

where q is equal to 0 in (3.35) and 1 in (3.36).

This expression will be approximated by

$$\begin{aligned} f_q(k) &\approx \int_0^a Re\left[e^{(-1)^q 2i\pi kw} F_q(w)\right] dw \\ &\approx \frac{a}{N}\sum_{j=0}^{N-1} Re\left[e^{(-1)^q 2i\pi k(j+\frac{1}{2})\frac{a}{N}} F_q((j+1/2)a/N)\right] \\ &= \frac{a}{N} Re\left[e^{(-1)^q 2i\pi \frac{k}{2}\frac{a}{N}} \sum_{j=0}^{N-1} e^{(-1)^q 2i\pi kj\frac{a}{N}} F_q((j+1/2)a/N)\right] \end{aligned} \tag{3.38}$$

where a centered rectangular rule has been used to approximate the integral by a sum.

The FFT is an efficient algorithm for computing the sum

$$W(s) = \sum_{j=0}^{N-1} e^{i2\pi bjs/N} X(j) \qquad \text{for } s = 0, \ldots, N-1, \tag{3.39}$$

where W and X are vectors of size N and b is a parameter of the FFT, with $b = 1$ or $b = -1$. The algorithm reduces the number of multiplications from an order of N^2 to that of $N \log_2(N)$. We will note: $W = FFT(X, b)$.

Let $[k_{min}, k_{max} - (k_{max} - k_{min})/N]$ be the range of log strike prices for which we want to compute the option price, with the constraint that $(k_{max} - k_{min})a = N$. Noting $k_s = k_{min} + s(k_{max} - k_{min})/N$ for $s = 0, \ldots, N-1$, the sum in (3.38) can be written

$$\begin{aligned} S(s) \;=\; \sum_{j=0}^{N-1} \Bigg(&\exp\left[(-1)^q 2i\pi s \frac{(k_{max} - k_{min})}{N} j \frac{a}{N}\right] \\ &\exp\left[(-1)^q 2i\pi k_{min} j \frac{a}{N}\right] F_q\left((j+1/2)a/N\right)\Bigg) \end{aligned} \tag{3.40}$$

can be computed, for each s, through the FFT algorithm

$$S = FFT(F_q^*, (-1)^q) \tag{3.41}$$

where F_q^* is the N dimensional vector with

$$(F_q^*)_j = e^{(-1)^q 2i\pi k_{min} j \frac{a}{N}} F_q((j + \frac{1}{2})\frac{a}{N}), \qquad \text{for } j = 0, \ldots, N-1 \tag{3.42}$$

The choice of a must be such that the characteristic function of the underlying variable, $\hat{g}_T(w)$ is negligible for $|w| > a$. As a rule of thumb, one can notice that, in a Gaussian setting, if $g_T(x)$ can be considered as negligible for $|x - \mu_T| > n\,\sigma_T$, then the characteristic function $\hat{g}_T(w)$ will also be negligible for $|w| > n/\sigma_T$.

Chapter 4

Option Pricing: Closed Form Solutions

4.1 Introduction

This chapter is organized around some key transforms that are special cases of the framework developed in **Chapter 3** applied to the Affine Jump Diffusion framework. In what follows, the "risk-factors" dynamics are thus assumed to be described by the AJD stochastic differential equation (2.2) under the "risk-neutral" probability measure Q.

Each time a key transform is introduced, we show how this key transform can be used as a building block to price different exotic options.

4.2 Futures and forwards

Futures and forwards are the most commonly traded financial products in energy markets. They are agreements on the price at which the commodity will be exchanged at some date in the future (maturity). Futures or forward contracts are just agreements on some future settlement rules and do not require any cash payment at the time the positions are taken.

On commodity markets, it is usual to refer to forward contracts when they are private contracts between two parties that are not standardized and settled at the end of the contract. Futures contracts, on the other hand denote standardized contracts that are traded on an exchange and settled daily.

The daily settlement of futures contracts is done through a clearing house that will adjust a trader's margin account to reflect the price changes occurred during the trading day. The holder of a long (short) position in a futures contract will thus receive (pay) at the end of every day the difference between the Futures closing price and its closing price at the end of the previous trading day. The cash settlement will thus happen continuously in the case of a futures contract and instantaneously at its maturity in the case of a forward contract. As will be shown later, this difference in cash settlements will be responsible for the price difference between the two contracts, but this difference will vanish in the case interest rates remain deterministic.

4.2.1 Elementary Transform

The elementary transform that is used to price futures and forward contracts is the key transform over which every other elementary transform that will be developed further builds upon. This is because it is in fact a generalization of the characteristic function of the underlying state variables that we denoted in the previous chapter as $\frac{1}{(i2\pi)^{\alpha_i}}\partial^{\alpha_i}\left(\check{\hat{g}}^*_{t,T}(X_T)\right)(\omega+\frac{ia}{2\pi})$ with only first order derivatives, that is $||\alpha_i||=1$. It has originally been developed by Duffie, Pan and Singleton (1999) [15].

Proposition 4.1 (from Duffie, Pan and Singleton (1999) [15]): *The transform $\phi(d_0, d_1, a, b, X_t, t, T)$, $\{\mathbb{R}, \mathbb{R}^n, \mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, [0,\infty), [0,\infty)\} \to \{\mathbb{R}\}$, defined as*

$$\begin{aligned}\phi(d_0, d_1, a, b, X_t, t, T) \quad &= \quad E^Q\Big[\exp\Big(-\int_t^T R(X_s, s)ds\Big) \\ &\qquad \Big(d_0 + d_1.X(T)\Big)e^{(a+ib)X(T)}\Big|\mathcal{F}_t\Big] \qquad (4.1)\end{aligned}$$

is given by

$$\phi(d_0, d_1, a, b, X_t, t, T) \equiv (A(t) + B(t).X_t)\, e^{\alpha(t)+\beta(t)X_t} \qquad (4.2)$$

where $A(t), B(t), \alpha(t)$ and $\beta(t)$ are solutions to the system of complex valued ODE's

$$\left\{\begin{array}{l}\dot{\beta}_t = \rho_1 - K_1^T\beta_t - \frac{1}{2}\beta_t^T H_1\beta_t - l_1(\theta(\beta_t) - 1) \\ \dot{\alpha}_t = \rho_0 - K_0^T\beta_t - \frac{1}{2}\beta_t^T H_0\beta_t - l_0(\theta(\beta_t) - 1) \\ -\dot{B}_t = K_1^T B_t + \beta_t^T H_1 B_t + l_1\nabla\theta(\beta_t)B_t \\ -\dot{A}_t = K_0^T B_t + \beta_t^T H_0 B_t + l_0\nabla\theta(\beta_t)B_t\end{array}\right. \qquad (4.3)$$

with the boundary conditions $B(T) = d_1, A(T) = d_0, \alpha(T) = 0$ and $\beta(T) = a + ib$, and where $\nabla\theta(c)$ is gradient of $\theta(c)$ with respect to $c \in \mathbb{C}^n$. $\square$

Note that for some notational conveniences, the dependence of $A(t), B(t), \alpha(t)$ and $\beta(t)$ on T and the boundary conditions has not been written explicitly.

4.2.2 Products Priced

4.2.2.1 Forwards

As mentioned earlier, this elementary transform will enable us to price forward contracts. We indeed have that

Proposition 4.2 *If the spot price is modelled by either the return model (2.1) or the level model (2.3), with X_t being an AJD under the risk-neutral probability model Q, then the forward price, $F(t,T)$ is a special case of the transform (7.2), where the expectation is taken under Q.*

Proof : By definition, a forward contract is an agreement between two parties to buy or sell an asset at a certain time in the future for a certain price. This price is referred to as the forward price and noted $F(t,T)$ where t is the time where the agreement takes place and T is the time of the transaction that has been agreed upon.

For such an agreement to be possible, the value of holding such a contract should be zero, so that

$$E^Q[e^{\int_t^T -R(X_s,s)ds}(Spot_t - F(t,T))|\mathcal{F}_t] = 0 \tag{4.4}$$

where $R(X_t,t)$ is the risk-free rate; and thus

$$F(t,T) = \frac{E^Q[e^{\int_t^T -R(X_s,s)ds}Spot_T|\mathcal{F}_t]}{E^Q[e^{\int_t^T -R(X_s,s)ds}|\mathcal{F}_t]} \tag{4.5}$$

For the level model, from equation (4.5), we have that

$$\begin{aligned}
F(t,T) &= \frac{E^Q[e^{\int_t^T -R(X_s,s)ds}\,(\gamma_{0,2} + \gamma_2.X_T)\,|\mathcal{F}_t]}{E^Q[e^{\int_t^T -R(X_s,s)ds}|\mathcal{F}_t]} && (4.6)\\
&= \frac{\phi(\gamma_{0,2},\gamma_2,0,0,t,T)}{\phi(1,0,0,0,t,T)} \\
&= \frac{(A_F(t) + B_F(t).X_t)\,e^{\alpha(t)+\beta(t)X_t}}{e^{\alpha(t)+\beta(t)X_t}} && (4.7)\\
&= A_F(t) + B_F(t).X_t
\end{aligned}$$

For the return model, from equation (4.5), we have that

$$F(t,T) = \frac{E^Q[e^{\int_t^T -R(X_s,s)ds}e^{\gamma_{0,1}+\gamma_1 X_T}|\mathcal{F}_t]}{E^Q[e^{\int_t^T -R(X_s,s)ds}|\mathcal{F}_t]} \tag{4.8}$$

$$
\begin{aligned}
&= \frac{\phi(e^{\gamma_{0,1}}, 0, \gamma_1, 0, t, T)}{\phi(1, 0, 0, 0, t, T)} \\
&= e^{\gamma_{0,1}} e^{\alpha_F(t) + \beta_F(t).X_t}
\end{aligned}
$$

□

Proposition 4.3 *Suppose the spot price is modelled by the return model (2.1) with X_t being an AJD under the risk-neutral probability model Q, then $ln(F(t,T))$ follows an AJD in Q; furthermore, if X_t also follows an AJD in the "physical" probability measure P, then $ln(F(t,T))$ follows an AJD in P as well. Suppose the spot price is modelled by the level model (2.3), with X_t being an AJD under the risk-neutral probability model Q, then $F(t,T)$ follows an AJD in Q; furthermore, if X_t also follows an AJD in the "physical" probability measure P, then $F(t,T)$ follows an AJD in P as well.*

Proof : For the return model, we have from (4.8) that

$$ln(F(t,T)) = \gamma_{0,1} + \alpha_F(t) + \beta_F(t).X_t \tag{4.9}$$

so that by differentiation, we get

$$dln(F(t,T)) = \dot{\alpha}_F(t)dt + \dot{\beta}_F(t).X_t dt + \beta_F(t)dX_t \tag{4.10}$$

Replacing dX_t by either equation (2.2) or (2.59) according to whether we want the stochastic process of $ln(F(t,T))$ under the "physical" or the risk-neutral probability measure, respectively, we observe that in any case, $ln(F(t,T))$ is indeed an AJD.

For the level model, we have from (4.6) that

$$F(t,T) = A_F(t) + B_F(t).X_t \tag{4.11}$$

so that by differentiation, we get

$$dF(t,T) = \dot{A}_F(t)dt + \dot{B}_F(t).X_t dt + B_F(t)dX_t \tag{4.12}$$

Replacing dX_t by either equation (2.2) or (2.59) according to whether we want the stochastic process of $F(t,T)$ under the "physical" or the risk-neutral probability measure, respectively, we observe that in any case, $F(t,T)$ is indeed an AJD. □

4.2.2.2 Futures

Proposition 4.4 *If the spot price is modelled by either the return model (2.1) or the level model (2.3), with X_t being an AJD under the risk-neutral probability model Q, then the futures price, $f(t,T)$ is a special case of the transform (7.2), where the risk-free rate is set to zero, and the expectation is taken under Q.*

Proof : It is a quite standard result in finance that if the market is arbitrage-free, then futures contracts are martingales under the risk-neutral probability measure Q (see, for example, Musiela and Rutowski (1997) [31]). That is,

$$E^Q[f(t+\tau,T)|\mathcal{F}_t] = f(t,T), \qquad \forall \tau, 0 \leq \tau \leq T-t \tag{4.13}$$

Furthermore, it is quite trivial that if the market is arbitrage-free then

$$f(T,T) = Spot_T \tag{4.14}$$

It follows from equations (4.13) and (4.14), and from the property of iterated expectations that

$$f(t,T) = E^Q[S_T|\mathcal{F}_t] \tag{4.15}$$

For the level model, from equation (4.5), we have that

$$\begin{aligned} f(t,T) &= E^Q[\gamma_{0,2} + \gamma_2.X_T|\mathcal{F}_t] \qquad (4.16) \\ &= \phi(\gamma_{0,2}, \gamma_2, 0, 0, t, T) \\ &= A_f(t) + B_f(t).X_t \end{aligned}$$

For the return model, from equation (4.5), we have that

$$\begin{aligned} f(t,T) &= E^Q[e^{\gamma_{0,1}+\gamma_1 X_T}|\mathcal{F}_t] \\ &= \phi(e^{\gamma_{0,1}},0,\gamma_1,0,t,T) \\ &= e^{\gamma_{0,1}}e^{\alpha_f(t)+\beta_f(t).X_t} \end{aligned} \tag{4.17}$$

where we have set $\rho_0 = 0$ and $\rho_1 = 0$ in the computation of (7.2). $\square$

Proposition 4.5 *Suppose the spot price is modelled by the return model (2.1) with X_t being an AJD under the risk-neutral probability model Q, then $ln(f(t,T))$ follows an AJD in Q; furthermore, if X_t also follows an AJD in the "physical" probability measure P, then $ln(f(t,T))$ follows an AJD in P as well. Suppose the spot price is modelled by the level model (2.3), with X_t being an AJD under the risk-neutral probability model Q, then $f(t,T)$ follows an AJD in Q; furthermore, if X_t also follows an AJD in the "physical" probability measure P, then $f(t,T)$ follows an AJD in P as well.*

Proof : For the return model, we have from (4.17) that

$$ln(f(t,T)) = \gamma_{0,1} + \alpha_f(t) + \beta_f(t).X_t \tag{4.18}$$

so that by differentiation, we get

$$dln(f(t,T)) = \dot{\alpha}_f(t)dt + \dot{\beta}_f(t).X_t dt + \beta_f(t)dX_t \tag{4.19}$$

Replacing dX_t by either equation (2.2) or (2.59) according to whether we want the stochastic process of $ln(f(t,T))$ under the "physical" or the risk-neutral probability measure, respectively, we observe that in any case, $ln(f(t,T))$ is indeed an AJD.

For the level model, we have from (4.6) that

$$f(t,T) = A_f(t) + B_f(t).X_t \tag{4.20}$$

so that by differentiation, we get

$$df(t,T) \quad = \quad \dot{A}_f(t)dt + \dot{B}_f(t).X_t dt + B_f(t)dX_t \tag{4.21}$$

Replacing dX_t by either equation (2.2) or (2.59) according to whether we want the stochastic process of $f(t,T)$ under the "physical" or the risk-neutral probability measure, respectively, we observe that in any case, $f(t,T)$ is indeed an AJD. □

It is quite trivial to observe that in the case the risk-free interest rate is deterministic ($\rho_1 = 0$), then futures and forward prices are the same. In any case, both products' prices are given by an affine or exponential affine function of the state variables so that if one is able to price an option on one product, it is always possible to price options on the other product as well. In the remainder of the text, we will focus on options on futures and will omit the underscores f denoting the pricing of futures. The computation of options on forwards will be assumed to derive trivially from these formulas.

4.3 European Options

Considering a spot price given by the return model, our objective is to show here how to price an European option with a payoff at maturity, T given by

$$Payoff(T) = \max\{e^{\gamma_{0,1}+\gamma_1.X_T} - e^{\gamma_{0,2}+\gamma_2.X_T}; 0\} \tag{4.22}$$

where X is an AJD in the risk-neutral probability measure Q.

For the level model, we will show how to price an European option with a payoff at maturity, T, given by

$$Payoff(T) = \max\{\gamma_{0,1} + \gamma_1.X_T - (\gamma_{0,2} + \gamma_2.X_T); 0\} \tag{4.23}$$

where X is an AJD in the risk-neutral probability measure Q.

4.3.1 A preliminary general result

We start by introducing the expectation $c(t)$ in a general (not necessarily AJD) statistical framework

$$\begin{aligned} c(t) &\equiv E^Q\left[e^{-\int_t^T R(X_s,s)ds}\mathbf{1}_{b.X_T\leq y}X_T^{\alpha}e^{a.X_T}|\mathcal{F}_t\right] \\ &= \left\langle \mathbf{1}_{b.X_T\leq y}, X_T^{\alpha}e^{a.X_T}g^*_{t,T}\right\rangle_{X_T} \end{aligned} \tag{4.24}$$

where $g^*_{t,T}(X_T)$ is the "discounted" joined probability density function, under the probability measure Q, of the vector of state variables X_T conditional on the information available at time t.

Proposition 4.6 *The value of the expectation described above is given by*

$$c(t) = \left\langle \frac{1}{2}e^{-i2\pi\omega_1 y}\left(\frac{-i}{\pi\omega_1} + \delta_{e_1}\right), \left(\frac{1}{(i2\pi)^{\alpha_i}}\partial^{\alpha_i}\mathfrak{F}\left(e^{-a.X_T}\check{g}^*_{t,T}(X_T)\right)\right)(b\omega_1)\right\rangle_{\omega_1} \tag{4.25}$$

Proof: Introducing a change of variable, $Y = BX$ with B a matrix of rank n and with its first line equal to b^T, this can be rewritten

$$c(t) \;=\; \left\langle \mathbf{1}_{Y_1(T)\leq y}, \partial_B \left(Y_T^\alpha e^{a.Y_T} g^*_{t,T}\right) |\det(B^{-1})| \right\rangle_{Y_T} \tag{4.26}$$

From (3.18), this is equivalent to

$$\begin{aligned} c(t) &= \left\langle \mathfrak{F}(\mathbf{1}_{Y_1(T)\leq y}), \partial_{-I} \mathfrak{F}\left(\partial_B \left(Y_T^\alpha e^{a.Y_T} g^*_{t,T}\right) |\det(B^{-1})|\right)\right\rangle \\ &= \left\langle \frac{1}{2} e^{-i2\pi\omega_1 y}\left(\frac{-i}{\pi\omega_1} + \delta_{e_1}\right) \prod_{j=2}^n \delta_{e_j}, \partial_{B^{-T}} \partial_{-I} \mathfrak{F}\left(Y_T^\alpha e^{a.Y_T} g^*_{t,T}(Y_T)\right)\right\rangle_\omega \\ &= \left\langle \frac{1}{2} e^{-i2\pi\omega_1 y}\left(\frac{-i}{\pi\omega_1} + \delta_{e_1}\right) \prod_{j=2}^n \delta_{e_j}, \partial_{B^{-T}} \left(\frac{1}{(i2\pi)^{\alpha_i}} \partial^{\alpha_i} \mathfrak{F}\left(e^{-a.X_T} \check{g}^*_{t,T}(X_T)\right)\right)\right\rangle_\omega \\ &= \left\langle \frac{1}{2} e^{-i2\pi\omega_1 y}\left(\frac{-i}{\pi\omega_1} + \delta_{e_1}\right), \left(\frac{1}{(i2\pi)^{\alpha_i}} \partial^{\alpha_i} \mathfrak{F}\left(e^{-a.X_T} \check{g}^*_{t,T}(X_T)\right)\right)(b\omega_1)\right\rangle_{\omega_1} \end{aligned} \tag{4.27}$$

□

This expectation further simplifies itself to all the pricing formulas developed in Bakshi and Madan (1999) [1] as will be shown in the case of Affine Jump Diffusion state variables.

4.3.2 Application to AJDs: the elementary transform

The following elementary transform allows to price European options with payoffs given by either (4.6) or by (4.23) according to whether we are in a return or a level model

Proposition 4.7 *The transorm $\tilde{G}(y, d_0, d_1, a, b, X_t, t, T)$, $\{\mathbb{R}, \mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, [0, \infty), [0, \infty)\} \to \{\mathbb{R}\}$, defined as*

$$\begin{aligned} \tilde{G}(y; d_0, d_1, a, b, X_t, t, T) &= E^Q\Big[\exp\left(-\int_t^T R(X_s, s) ds\right) \\ &\qquad (d_0 + d_1.X(T)) e^{a.X(T)} \mathbf{1}_{b.X(T)\leq y} |\mathcal{F}_t\Big] \end{aligned} \tag{4.28}$$

is given by

$$\tilde{G}(y; d_0, d_1, a, b, X_t, t, T) = \frac{\phi(d_0, d_1, a, 0, X_t, t, T)}{2} - \frac{1}{\pi}\int_0^\infty \frac{Im\left[\phi(d_0, d_1, a, vb, X_t, t, T)e^{-ivy}\right]}{v}dv \quad (4.29)$$

Proof : The transform can be written

$$\tilde{G}(y; d_0, d_1, a, b, X_t, t, T) = \left\langle \mathbf{1}_{b.X(T)\le y}, e^{-\int_t^T R(X_s,s)ds}(d_0 + d_1.X(T))e^{a.X(T)}g^*_{t,T}(X_T)\right\rangle_{X_T} \quad (4.30)$$

where $g^*_{t,T}(X_T)$ is the "discounted" joined probability density function of the vector of state variables X_T conditional on the information available at time t.

Introducing a change of variable, $Y = BX$ with B a matrix of rank n and with its first line equal to b^T, this can be rewritten

$$\tilde{G}(y; d_0, d_1, a, b, X_t, t, T) = \left\langle \mathbf{1}_{Y_1(T)\le y}, \partial_B\left((d_0 + d_1.Y_T)e^{a.Y_T}g^*_{t,T}(Y_T)|\det(B^{-1})|\right)\right\rangle_{Y_T} \quad (4.31)$$

From equation (4.27), we get

$$\begin{aligned}
&\tilde{G}(y; d_0, d_1, a, b, X_t, t, T) &(4.32)\\
&= \left\langle \mathfrak{F}(\mathbf{1}_{Y_1(T)\le y}), \partial_{-I}\mathfrak{F}\left(\partial_B\left((d_0 + d_1.Y_T)e^{a.Y_T}g^*_{t,T}(Y_T)|\det(B^{-1})|\right)\right)\right\rangle\\
&= \left\langle \mathfrak{F}(\mathbf{1}_{Y_1(T)\le y}), \partial_{-I}\partial_{B^{-T}}\mathfrak{F}\left((d_0 + d_1.Y_T)e^{a.Y_T}g^*_{t,T}(Y_T)\right)\right\rangle\\
&= \left\langle \frac{1}{2}e^{-i2\pi\omega_1 y}\left(\frac{-i}{\pi\omega_1} + \delta_{e_1}\right)\prod_{j=2}^n \delta_{e_j}, \partial_{B^{-T}}\left(d_0 + d_1.\frac{1}{i2\pi}\partial^{\{1,\ldots,1\}}\right)\check{\hat{g}}^*_{t,T}\left(\omega + \frac{a}{i2\pi}\right)\right\rangle_\omega\\
&= \left\langle \frac{1}{2}e^{-i2\pi\omega_1 y}\left(\frac{-i}{\pi\omega_1} + \delta_{e_1}\right), \left(\left(d_0 + d_1.\frac{1}{i2\pi}\partial^{\{1,\ldots,1\}}\right)\check{\hat{g}}^*_{t,T}\right)(b\omega_1 + \frac{a}{i2\pi})\right\rangle_{\omega_1}
\end{aligned}$$

which, from the definition of the transform $\phi(d_0, d_1, a, b, X, t, T)$ (7.1), can be written

$$\tilde{G}(y; d_0, d_1, a, b, X_t, t, T) \quad (4.33)$$

$$
\begin{aligned}
&= \frac{\phi(d_0, d_1, a, 0, X_t, t, T)}{2} \\
&\quad - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{i\phi(d_0, d_1, a, 2\pi\omega b, X_t, t, T) e^{i2\pi\omega y}}{\omega} d\omega \\
&= \frac{\phi(d_0, d_1, a, 0, X_t, t, T)}{2} \\
&\quad - \frac{1}{\pi} \int_{0}^{\infty} \frac{Im\left[\phi(d_0, d_1, a, vb, X_t, t, T) e^{-ivy}\right]}{v} dv
\end{aligned}
$$

□

This is exactly the result Duffie, Pan and Singleton(1999) [15] arrived at through a Fourier-Stieltjes transform of $\tilde{G}$ with respect to the strike price. Furthermore, it is also the multi-variate generalization of the univariate result of Bakshi and Madan (1999) [1].

4.3.3 Products Priced

We first show how to price European options with payoffs given by either (4.6) or (4.23), then we will show how different products can be seen as special cases of such options.

Proposition 4.8 (from Duffie, Pan and Singleton (1999) [15]): *Let a European option with a payoff at maturity T given by equation (4.6) (return model), with X_t an AJD in the risk- neutral probability measure Q. The price at time t of such an European option is given by a sum of linear functions of special cases of the transform $\tilde{G}$, (7.3).*

Proof : By definition of the risk-neutral probability measure, the value at time t of such a contingent claim, $c(t)$, is given by

$$
c(t) \quad = \quad E^Q\left[e^{-\int_t^T R(X_s)ds} \max\{e^{\gamma_{0_1}+\gamma_1.X_T} - e^{\gamma_{0,2}+\gamma_2.X_T}; 0\} | \mathcal{F}_t\right] \tag{4.34}
$$

$$
\begin{aligned}
&= E^Q\left[e^{-\int_t^T R(X_s)ds} e^{\gamma_{0_1}+\gamma_1.X_T} \mathbf{1}_{(\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-\gamma_{0,2}}|\mathcal{F}_t\right] \\
&\quad -E^Q\left[e^{-\int_t^T R(X_s)ds} e^{\gamma_{0_2}+\gamma_2.X_T} \mathbf{1}_{(\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-\gamma_{0,2}}|\mathcal{F}_t\right] \\
&= \tilde{G}(\gamma_{0,1}-\gamma_{0,2}, e^{\gamma_{0,1}}, 0, \gamma_1, \gamma_2-\gamma_1, X_t, t, T) \\
&\quad -\tilde{G}(\gamma_{0,1}-\gamma_{0,2}, e^{\gamma_{0,2}}, 0, \gamma_2, \gamma_2-\gamma_1, X_t, t, T)
\end{aligned}
$$

□

Proposition 4.9 (from Duffie, Pan and Singleton (1999) [15]): *Let a European option with a payoff at maturity T given by equation (4.23) (level model), with X_t an AJD in the risk- neutral probability measure Q. The price at time t of such an European option is given by a sum of linear functions of special cases of the transform $\tilde{G}$, (7.3).*

Proof : By definition of the risk-neutral probability measure, the value at time t of such a contingent claim, $c(t)$, is given by

$$
\begin{aligned}
c(t) &= E^Q\left[e^{-\int_t^T R(X_s)ds} \max\{\gamma_{0,1}+\gamma_1.X_T-(\gamma_{0,2}+\gamma_2.X_T);0\}|\mathcal{F}_t\right] \quad (4.35) \\
&= E^Q\left[e^{-\int_t^T R(X_s)ds} \left((\gamma_{0,1}-\gamma_{0,2})+(\gamma_1-\gamma_2).X_T\right) \mathbf{1}_{(\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-\gamma_{0,2}}|\mathcal{F}_t\right] \\
&= \tilde{G}(\gamma_{0,1}-\gamma_{0,2}, \gamma_{0,1}-\gamma_{0,2}, \gamma_1-\gamma_2, 0, \gamma_2-\gamma_1, X_t, t, T) \quad \square
\end{aligned}
$$

4.3.3.1 European calls and puts on the spot

By letting $\gamma_2 = 0$, the payoffs (4.6) for the return model and (4.23) for the level model become the payoffs of a classical European call with strike price $e^{\gamma_{0,2}}$ or $\gamma_{0,2}$, respectively.

Furthermore, by letting $\gamma_1 = 0$, the payoffs (4.6) for the return model and (4.23) for the level model become the payoffs of a classical European put with strike price $e^{\gamma_{0,1}}$ or $\gamma_{0,1}$, respectively.

4.3.3.2 European calls and puts on futures (and forwards)

From **Proposition 4.4**, the price of a futures contract is given by an (exponential) affine function of the underlying factors X_t when the spot price is modelled by a (return) level model.

As a result, letting $\gamma_2 = 0$, and replacing $\gamma_{0,1}$ in equations (4.6) and (4.23) by $\gamma_{0,1} + \alpha(T_{option}, T_{futures})$ from equation (4.18) in the return model, and $A(T_{option}, T_{futures})$ from equation (4.20) in the level model, and replacing also γ_1 in equations (4.6) and (4.23) by $\beta(T_{option}, T_{futures})$ from equation (4.18) in the return model, and $B(T_{option}, T_{futures})$ from equation (4.20) in the level model, we get the payoff of a standard call option with maturity T_{option} on futures with maturity $T_{futures}$, and with a strike price $e^{\gamma_{0,2}}$ or $\gamma_{0,2}$, respectively. Notice that we made explicit here the dependence of α and A on both time of evaluation and time of maturity of the futures contract.

Similarly, letting $\gamma_1 = 0$, and replacing $\gamma_{0,2}$ in equations (4.6) and (4.23) by $\gamma_{0,1} + \alpha(T_{option}, T_{futures})$ from equation (4.18) in the return model, and $A(T_{option}, T_{futures})$ from equation (4.20) in the level model, and also replacing γ_2 in equations (4.6) and (4.23) by $\beta(T_{option}, T_{futures})$ from equation (4.18) in the return model, and $B(T_{option}, T_{futures})$ from equation (4.20) in the level model, we get the payoffs of a standard put option with maturity T_{option} on futures with maturity $T_{futures}$, and with strike price $e^{\gamma_{0,1}}$ or $\gamma_{0,1}$, respectively.

4.3.3.3 Spread Options

Equations (4.6) and (4.23), as such, are the payoffs of Spread options with a strike price set to zero in the return and in the level model framework, respectively. According to the interpretation given to the risk-factors X_t, the spread can either be a spark spread, or a locational spread.

One can trivially account for a non-zero strike price in the level model (4.23) by incorporating to either $\gamma_{0,1}$ or $\gamma_{0,2}$ the strike price. Unfortunately, this does not hold in the payoff equation for the return model (4.6).

4.4 Multi-Criteria European Options

Multi-Criteria European Options are options with a payoff at maturity that can be decomposed in a sum of affine or exponential affine functions of the state variables, each element of the sum being conditional on several criteria expressed as a product of step functions. The payoff at maturity T can thus be expressed as

$$Payoff(T) = \sum_{i=1}^{n} \Big(d_{i,0} + d_{i,1}.X(T)\Big) e^{a_i.X(T)} \prod_{j=1}^{m} \mathbf{1}_{b_{i,j}.X(T) \leq y_{i,j}} \tag{4.36}$$

Examples of such contracts include options on the minimum or the maximum of n assets, or the replication of Multi-Fuel Power Plants where the ramp-up rates have been neglected.

4.4.1 Preliminary general result

We start by introducing the expectation $c(t)$ in a general (not necessarily AJD) statistical framework:

$$c(t) = E^Q \left[e^{-\int_t^T R(X_s,s)ds} X_T^{\alpha} e^{a.X_T} \prod_{k=1}^{m} \mathbf{1}_{b_k.X_T \leq y_k} \middle| \mathcal{F}_t \right] \tag{4.37}$$

Proposition 4.10 *Let m denote the number of criteria (unit-step functions) in the decomposition of the payoff function, and n, the size of the vector of underlying variables X, then*

$$c(t) = \left\langle \prod_{k=1}^{m} \frac{1}{2} e^{-i2\pi\omega_k y_k} \left(\frac{-i}{\pi\omega_k} + \delta_{e_k} \right), \left(\frac{1}{(i2\pi)^{\alpha}} \partial^{\alpha} \mathfrak{F} \left(e^{-a.X_T} \breve{g}^*_{t,T}(X_T) \right) \right) (A^T\omega) \right\rangle_{\omega} \tag{4.38}$$

where A is a $m \times n$ matrix with its k^{th} line, $A_k = b_k$, and $\omega \in \mathbb{R}^m$.

Proof: Let us first consider the case where $m \leq n$ and introduce the change of variable $Y = B_\epsilon X$ where $Y \in \mathbb{R}^n$, and $B_\epsilon = B + \epsilon$, with $B, \epsilon \in \mathbb{R}^{n\times n}$, and with $B = \begin{pmatrix} A \\ \mathbf{0} \end{pmatrix}$, where the k^{th} row of $A[m \times n]$, $A_k = b_k$. The matrix ϵ is further constrained such that B_ϵ is invertible.

$$\begin{aligned}
c(t) &= \lim_{||\epsilon||\to 0} \left\langle \prod_{k=1}^{m} \mathbf{1}_{(B_\epsilon)_k X_T \leq y_k}, X_T^\alpha e^{a.X_T} g_{t,T}^* \right\rangle_{X_T} && (4.39)\\
&= \lim_{||\epsilon||\to 0} \left\langle \prod_{k=1}^{m} \mathbf{1}_{Y_k(T) \leq y_k}, \partial_{B_\epsilon} \left(Y_T^\alpha e^{a.Y_T} g_{t,T}^* \right) |\det(B_\epsilon^{-1})| \right\rangle_{Y_T} \\
&= \left\langle \prod_{k=1}^{m} \mathfrak{F}\left(\mathbf{1}_{Y_k(T)\leq y_k}\right) \prod_{j=m+1}^{n} \delta_{e_j}, \left(\frac{1}{(i2\pi)^\alpha} \partial^\alpha \mathfrak{F}\left(e^{-a.Y_T} \check{g}_{t,T}^*(Y_T)\right)\right)(B^T\omega) \right\rangle_\omega \\
&= \left\langle \prod_{k=1}^{m} \frac{1}{2} e^{-i2\pi\omega_k y_k} \left(\frac{-i}{\pi\omega_k} + \delta_{e_k} \right), \left(\frac{1}{(i2\pi)^\alpha} \partial^\alpha \mathfrak{F}\left(e^{-a.Y_T} \check{g}_{t,T}^*(Y_T)\right)\right)(A^T\omega) \right\rangle_\omega
\end{aligned}$$

If $m > n$, the vector of state variables X is extended to $X' = \begin{pmatrix} X \\ \widetilde{X} \end{pmatrix}$, where $\widetilde{X}$ is of size $m-n$. The change of variable $Y = B_\epsilon X'$ is also introduced, where $Y \in \mathbb{R}^m$, and $B_\epsilon = B + \epsilon$, with $B, \epsilon \in \mathbb{R}^{m\times m}$, and with $B = (A \quad \mathbf{0})$, where the k^{th} row of $A[m \times n]$, $A_k = b_k$. The matrix ϵ is further constrained such that B_ϵ is invertible.

Let $\widetilde{g}_\Sigma : \mathbb{R}^{m-n} \to \mathbb{C}$, denote the function $\widetilde{g}_\Sigma(x) = |2\pi\Sigma|^{\frac{-1}{2}} e^{-(x^T\Sigma^{-1}x)}$. It is easy to show that $\widetilde{g}_\Sigma \in \mathcal{P}(\mathbb{R}^{m-n})$, and that

$$\widehat{\widetilde{g}}_\Sigma(\mathbf{0}) = 1 \tag{4.40}$$

We then have that

$$\begin{aligned}
c(t) &= \lim_{||\epsilon||\to 0} \left\langle \prod_{k=1}^{m} \mathbf{1}_{(B_\epsilon^T)_k X_T' \leq y_k}, X_T^\alpha e^{a.X_T} g_{t,T}^*(X_T) \widetilde{g}_\Sigma(\widetilde{X}_T) \right\rangle_{X_T'} \\
&\equiv \lim_{||\epsilon||\to 0} \left\langle \prod_{k=1}^{m} \mathbf{1}_{(B_\epsilon^T)_k X_T' \leq y_k}, h(X_T') \right\rangle_{X_T'}
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{||\epsilon||\to 0} \left\langle \prod_{k=1}^{m} \mathbf{1}_{Y_k(T)\le y_k}, \partial_{B_\epsilon} h(Y_T) |\det(B_\epsilon^{-T})| \right\rangle_{Y_T} \\
&= \lim_{||\epsilon||\to 0} \left\langle \prod_{k=1}^{m} \mathfrak{F}\left(\mathbf{1}_{Y_k(T)\le y_k}\right), \partial_{B_\epsilon^{-T}} \check{\widehat{h}} \right\rangle_{Y_T} \qquad (4.41)
\end{aligned}
$$

Noticing that

$$
\widehat{h}\left(\{\omega', \omega''\}\right) = \mathfrak{F}\left(Y_T^{\alpha} e^{a.Y_T} g_{t,T}^{*}(Y_T)\right)(\omega') \quad \widehat{\widehat{g}}_{\Sigma}(\omega'') \qquad (4.42)
$$

with $\omega' \in \mathbb{R}^n$, and $\omega'' \in \mathbb{R}^{m-n}$, we then have that

$$
\begin{aligned}
\widehat{h}(B^T\omega) &= \mathfrak{F}\left(X_T^{\alpha} e^{a.X_T} g_{t,T}^{*}(X_T)\right)\left(A^T\omega\right) \quad \widehat{\widehat{g}}_{\Sigma}(\mathbf{0}^T\omega) \\
&= \mathfrak{F}\left(X_T^{\alpha} e^{a.X_T} g_{t,T}^{*}(X_T)\right)\left(A^T\omega\right) \qquad (4.43)
\end{aligned}
$$

such that equation (4.41) can be written

$$
\begin{aligned}
c(t) &= \left\langle \prod_{k=1}^{m} \mathfrak{F}\left(\mathbf{1}_{Y_k(T)\le y_k}\right), \left(\frac{1}{(i2\pi)^{\alpha}} \partial^{\alpha} \mathfrak{F}\left(e^{-a.Y_T} \check{g}_{t,T}^{*}(Y_T)\right)\right)\left(A^T\omega\right) \right\rangle \\
&= \left\langle \prod_{k=1}^{m} \frac{1}{2} e^{-i2\pi\omega_k y_k} \left(\frac{-i}{\pi\omega_k} + \delta_{e_k}\right), \left(\frac{1}{(i2\pi)^{\alpha}} \partial^{\alpha} \tau_{\frac{-a}{i2\pi}} \check{\hat{g}}_{t,T}^{*}\right)\left(A^T\omega\right) \right\rangle_{\omega} \qquad (4.44)
\end{aligned}
$$

This concludes the proof. $\square$

4.4.2 Application to AJDs: the elementary transform

A straight application of the previous result to the AJD setting gives:

Proposition 4.11 *The transform $G(y_1, y_2, d_0, d_1, a, b_1, b_2, X_t, t, T)$, $\{\mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{R}^n, \mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, [0,\infty), [0,\infty)\} \to \{\mathbb{R}\}$, defined as*

$$
\begin{aligned}
&G(y_1, y_1, d_0, d_1, a, b_1, b_2, X_t, t, T) = \\
&E^{Q}\Big[\exp\left(-\int_t^T R(X_s, s)ds\right) \\
&\qquad \Big(d_0 + d_1.X(T)\Big) e^{a.X(T)} \prod_{j=1}^{2} \mathbf{1}_{b_j.X(T)\le y_j} \Big| \mathcal{F}_t\Big] \qquad (4.45)
\end{aligned}
$$

is given by

$$
\begin{aligned}
G(y_1, y_2, d_0, d_1, a, b_1, b_2, X_t, t, T) &= \frac{\phi(d_0, d_1, a, 0, X_t, t, T)}{4} \qquad (4.46)\\
&- \frac{1}{2\pi} \sum_{j=1}^{2} \int_0^\infty \frac{Im\left[e^{-iwy_j}\phi(d_0, d_1, a, wb_j, X_t, t, T)\right]}{w} dw \\
&+ \frac{1}{2\pi^2} \int_0^\infty \int_0^\infty \left\{ \frac{Re\left[e^{-i(w_1y_1 - w_2y_2)}\phi(d_0, d_1, a, w_1b_1 - w_2b_2, X_t, t, T)\right]}{w_1w_2} \right. \\
&\left. - \frac{Re\left[e^{-i(w_1y_1 + w_2y_2)}\phi(d_0, d_1, a, w_1b_1 + w_2b_2, X_t, t, T)\right]}{w_1w_2} \right\} dw_1 dw_2
\end{aligned}
$$

□

This is a two-dimensional extension of the results in Duffie, Pan and Singleton (1999) [15] that had already been derived in Kamat and Oren (2000) [26]. It is quite obvious that it can be generalized to any higher dimension.

4.4.3 Products Priced

4.4.3.1 Option on the min or the max of n assets

We consider here the case of a call on the maximum of three assets. The generalization to other situations (min of n assets or a put...) is straightforward.

Let the payoff of such a call be given by:

$$\left(\max\left(S_1(T), S_2(T), S_3(T)\right) - K\right)^+ \qquad (4.47)$$

We start by assuming a return model for the state variables such that, defining the vector :

$$X_t \equiv \begin{pmatrix} \log(S_1(t)) \\ \log(S_2(t)) \\ \log(S_3(t)) \end{pmatrix} \equiv \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \qquad (4.48)$$

we assume that the dynamics of this vector under the risk neutral probability measure is well described by an AJD.

The price of the call, C is given by:

$$C = E^{\chi}\left[\exp\left(-\int_t^T R(X_s,s)ds\right)\left(\max\Big(S_1(T),S_2(T),S_3(T)\Big)-K\right)^+\Big|\mathcal{F}_t\right] \tag{4.49}$$

This can further be expressed as:

$$\begin{aligned}
C &= E^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right)e^{X_1(T)} \\
&\qquad \mathbf{1}_{X_2(T)-X_1(T)\le 0}\,\mathbf{1}_{X_3(T)-X_1(T)\le 0}\,\mathbf{1}_{-X_1(T)\le -\log(K)}\Big|\,\mathcal{F}_t\Big] \\
&+ E^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right)e^{X_2(T)} \\
&\qquad \mathbf{1}_{X_1(T)-X_2(T)\le 0}\,\mathbf{1}_{X_3(T)-X_2(T)\le 0}\,\mathbf{1}_{-X_2(T)\le -\log(K)}\Big|\,\mathcal{F}_t\Big] \\
&+ E^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right)e^{X_3(T)} \\
&\qquad \mathbf{1}_{X_1(T)-X_3(T)\le 0}\,\mathbf{1}_{X_2(T)-X_3(T)\le 0}\,\mathbf{1}_{-X_3(T)\le -\log(K)}\Big|\,\mathcal{F}_t\Big] \\
&- KE^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right) \\
&\qquad \mathbf{1}_{X_2(T)-X_1(T)\le 0}\,\mathbf{1}_{X_3(T)-X_1(T)\le 0}\,\mathbf{1}_{-X_1(T)\le -\log(K)}\Big|\,\mathcal{F}_t\Big] \\
&- KE^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right) \\
&\qquad \mathbf{1}_{X_1(T)-X_2(T)\le 0}\,\mathbf{1}_{X_3(T)-X_2(T)\le 0}\,\mathbf{1}_{-X_2(T)\le -\log(K)}\Big|\,\mathcal{F}_t\Big] \\
&- KE^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right) \\
&\qquad \mathbf{1}_{X_1(T)-X_3(T)\le 0}\,\mathbf{1}_{X_2(T)-X_3(T)\le 0}\,\mathbf{1}_{-X_3(T)\le -\log(K)}\Big|\,\mathcal{F}_t\Big]
\end{aligned} \tag{4.50}$$

Obviously, all terms of (4.50) are special cases of (4.45).

Similarly, considering a level model for the state variables, we define the vector

$$X_t \equiv \begin{pmatrix} S_1(t) \\ S_2(t) \\ S_3(t) \end{pmatrix} \equiv \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \tag{4.51}$$

and assume that the dynamics of this vector under the risk neutral probability measure is well described by an AJD.

The price of the call, C can then further be expressed as:

$$\begin{aligned} C \;=\; & E^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right)(X_1(T)-K) \\ & \quad \mathbf{1}_{X_2(T)-X_1(T)\le 0}\,\mathbf{1}_{X_3(T)-X_1(T)\le 0}\,\mathbf{1}_{-X_1(T)\le -K}\Big|\,\mathcal{F}_t\Big] \\ + \; & E^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right)(X_2(T)-K) \\ & \quad \mathbf{1}_{X_1(T)-X_2(T)\le 0}\,\mathbf{1}_{X_3(T)-X_2(T)\le 0}\,\mathbf{1}_{-X_2(T)\le -K}\Big|\,\mathcal{F}_t\Big] \\ + \; & E^{\chi}\Big[\exp\left(-\int_t^T R(X_s,s)ds\right)(X_3(T)-K) \\ & \quad \mathbf{1}_{X_1(T)-X_3(T)\le 0}\,\mathbf{1}_{X_2(T)-X_3(T)\le 0}\,\mathbf{1}_{-X_3(T)\le -K}\Big|\,\mathcal{F}_t\Big] \end{aligned} \tag{4.52}$$

Obviously, all terms of (4.52) are again special cases of the transform G (4.45).

4.4.3.2 Real Options Valuation of a Multi-Fuel Power Plant

A multi-fuel power plant has the flexibility to choose among different fuels to produce electricity. At each time period during its economic life-time, the power plant will choose the most profitable fuel according to the market conditions that will occur at that time.

We consider here a power plant that can run either on natural gas or fuel oil. At each time period, t, if we neglect the costs of switching fuels and the ramp-up or down rates, we can express its payoff as:

$$Payoff(t) = Capacity.\Big(S_e(T) - \min\Big(H_g S_g(T), H_o S_o(T)\Big)\Big)^+ \tag{4.53}$$

where :

- $Capacity$ stands for the turbine capacity of the power plant [Mwh];
- $S_e(t)$ stands for the electricity spot price at time t [$/Mwh];
- $S_g(t)$ stands for the nat. gas spot price at time t [$/Mjoules];
- $S_c(t)$ stands for the oil spot price at time t [$/Mjoules];
- H_g stands for the power plant's nat. gas Heat Rate [Mwh/Mjoules];
- H_c stands for the power plant's oil Heat Rate [Mwh/Mjoules].

Such a power plant can thus be seen as a collection of "options to exchange one asset for another" with a generic payoff at time T given by:

$$Payoff(T) = max\{S_1(T) - min\{H_2 S_2(T); H_3 S_3(T)\}; 0\} \tag{4.54}$$

where S_1, S_2 and S_3 stand for the prices of three different assets.

Again, we will consider a return model and a level model for the assets. The return model writes

$$S_i(t) = e^{\gamma_{0,i} + \gamma_i . X_t} \tag{4.55}$$

The level model writes

$$S_i(t) = \gamma_{0,i} + \gamma_i . X_t \tag{4.56}$$

In both cases, the vector X_t is assumed to follow an AJD under the risk-neutral probability measure Q.

Proposition 4.12 *Let, $S_i(t)$ be a set of assets described either by a return model of the type (4.55), or a level model of the type (4.56), with X_T following an AJD process under the risk-neutral probability measure Q. Contingent claims with payoffs of the type (4.54) have a value at time t, $c(t)$, that can be expressed as the sum of affine functions of the transform G, (4.45).*

Proof : For the return model, by definition of the risk-neutral probability measure, we have that

$$\begin{aligned}
c(t) &= E^Q\Big[e^{-\int_t^T R(X_s)ds}\max\Big\{e^{\gamma_{0_1}+\gamma_1.X_T} - \\
&\qquad min\{H_2 e^{\gamma_{0,2}+\gamma_2.X_T}; H_3 e^{\gamma_{0,3}+\gamma_3.X_T}\}; 0\Big\}|\mathcal{F}_t\Big] \qquad (4.57)\\
&= E^Q\Big[e^{-\int_t^T R(X_s)ds} e^{\gamma_{0_1}+\gamma_1.X_T}\mathbf{1}_{(\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-(ln(H_2)+\gamma_{0,2})}\\
&\qquad \mathbf{1}_{(\gamma_3-\gamma_1).X_T\le\gamma_{0,1}-(ln(H_3)+\gamma_{0,3})}|\mathcal{F}_t\Big]\\
&\quad -E^Q\Big[e^{-\int_t^T R(X_s)ds} H_2 e^{\gamma_{0_2}+\gamma_2.X_T}\mathbf{1}_{(\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-(ln(H_2)+\gamma_{0,2})}\\
&\qquad \mathbf{1}_{(\gamma_2-\gamma_3).X_T\le ln(H_3)+\gamma_{0,3}-(ln(H_2)+\gamma_{0,2})}|\mathcal{F}_t\Big]\\
&\quad -E^Q\Big[e^{-\int_t^T R(X_s)ds} H_3 e^{\gamma_{0_3}+\gamma_3.X_T}\mathbf{1}_{(\gamma_3-\gamma_1).X_T\le\gamma_{0,1}-(ln(H_3)+\gamma_{0,3})}\\
&\qquad \mathbf{1}_{(\gamma_3-\gamma_2).X_T\le ln(H_2)+\gamma_{0,2}-(ln(H_3)+\gamma_{0,3})}|\mathcal{F}_t\Big]\\
&= G\Big(\gamma_{0,1}-(ln(H_2)+\gamma_{0,2}), ln(H_3)+\gamma_{0,3}-(ln(H_2)+\gamma_{0,2}), e^{\gamma_{0_1}}, 0,\\
&\qquad \gamma_1, \gamma_2-\gamma_1, \gamma_3-\gamma_1, X_t, t, T\Big)\\
&\quad -G\Big(\gamma_{0,1}-(ln(H_2)+\gamma_{0,2}), ln(H_3)+\gamma_{0,3}-(ln(H_2)+\gamma_{0,2}), H_2 e^{\gamma_{0,2}}, 0,\\
&\qquad \gamma_2, \gamma_2-\gamma_1, \gamma_2-\gamma_3, X_t, t, T\Big)\\
&\quad -G\Big(\gamma_{0,1}-(ln(H_3)+\gamma_{0,3}), ln(H_2)+\gamma_{0,2}-(ln(H_3)+\gamma_{0,3}), H_3 e^{\gamma_{0,3}}, 0,\\
&\qquad \gamma_3, \gamma_3-\gamma_1, \gamma_3-\gamma_2, X_t, t, T\Big)
\end{aligned}$$

For the level model, by definition of the risk-neutral probability measure, we have that

$$
\begin{aligned}
c(t) &= E^Q\Big[e^{-\int_t^T R(X_s)ds}\max\Big\{\gamma_{0_1}+\gamma_1.X_T- \\
&\qquad min\{H_2(\gamma_{0,2}+\gamma_2.X_T);H_3(\gamma_{0,3}+\gamma_3.X_T)\};0\Big\}|\mathcal{F}_t\Big] \qquad (4.58)\\
&= E^Q\Big[e^{-\int_t^T R(X_s)ds}\gamma_{0_1}+\gamma_1.X_T\mathbf{1}_{(H_2\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-H_2\gamma_{0,2}} \\
&\qquad \mathbf{1}_{(H_3\gamma_3-\gamma_1).X_T\le\gamma_{0,1}-H_3\gamma_{0,3}}|\mathcal{F}_t\Big] \\
&\quad -E^Q\Big[e^{-\int_t^T R(X_s)ds}H_2(\gamma_{0_2}+\gamma_2.X_T)\mathbf{1}_{(H_2\gamma_2-\gamma_1).X_T\le\gamma_{0,1}-H_2\gamma_{0,2}} \\
&\qquad \mathbf{1}_{(H_2\gamma_2-H3\gamma_3).X_T\le H_3\gamma_{0,3}-H_2\gamma_{0,2}}|\mathcal{F}_t\Big] \\
&\quad -E^Q\Big[e^{-\int_t^T R(X_s)ds}H_3(\gamma_{0_3}+\gamma_3.X_T)\mathbf{1}_{(H_3\gamma_3-\gamma_1).X_T\le\gamma_{0,1}-H_3\gamma_{0,3}} \\
&\qquad \mathbf{1}_{(H_3\gamma_3-H_2\gamma_2).X_T\le H_2\gamma_{0,2}-H_3\gamma_{0,3}}|\mathcal{F}_t\Big] \\
&= G\Big(\gamma_{0,1}-H_2\gamma_{0,2},\gamma_{0,1}-H_3)\gamma_{0,3},\gamma_{0_1},\gamma_1, \\
&\qquad 0,H_2\gamma_2-\gamma_1,H_3\gamma_3-\gamma_1,X_t,t,T\Big) \\
&\quad -G\Big(\gamma_{0,1}-H_2\gamma_{0,2},\gamma_{0,1}-H_3\gamma_{0,3},H_2\gamma_{0_2},H_2\gamma_2, \\
&\qquad 0,H_2\gamma_2-\gamma_1,H_3\gamma_3-H_2\gamma_2,X_t,t,T\Big) \\
&\quad -G\Big(\gamma_{0,1}-H_3\gamma_{0,3},H_2\gamma_{0,2}-H_3\gamma_{0,3},H_3\gamma_{0,3},H_3\gamma_3, \\
&\qquad 0,H_3\gamma_3-\gamma_1,H_3\gamma_3-H_2\gamma_2,X_t,t,T\Big)
\end{aligned}
$$

□.

Obviously, it is quite trivial to extend equation (4.54) to the case where there are more than three different assets.

4.4.3.2.1 Example of implementation :

With the deregulation of energy markets, the whole European energy industry is being re-structured. Mergers, and acquisitions have occurred and will most probably continue. In such a context, it is essential to accurately value power plants and to apply valuation techniques that capture the whole potential offered by flexible power plants in a deregulated environment. Common flexibilities offered by power plants include the possibility to start and shut down, or switch from one fuel to another according to the circumstances encountered.

Traditional valuation techniques such as the Net Pesent Value (NPV), because they only consider a single "most likely" central scenario for electricity and fuel prices, typically fail to capture the full value of this flexibility. On the other hand, considering power plants as options to buy the less expensive fuel and sell electricity, does capture the full extent of the flexibility value of such multi-fuel power plants. In fact, this will be an upper-bound for the value of the power plant because the ramp-up and -down rates and switching fuel costs are being ignored.

In this section, we implement the transforms developed above to value single- and multi-fuel power plants using this real options approach and show the added value of such an approach compared to a static NPV approach that does not capture any flexibility value.

We chose as a reference for the electricity price, the prices quoted on the German EEX market. Natural gas prices and crude prices will be given by the NBP and Brent prices both quoted on the IPE market in London. It is assumed here that the price of the fuel oil that is used in power plants is directly related to the crude's price by a factor of 0.8 (Fuel Oil = 0.8 Crude).

In Chapter 7, we provide AJD models for these commodities and estimates of their parameters. Those AJD models have two factors for NBP and Brent, and three factors for the EEX, one extra factor being used to model the spiking behavior of the electricity prices.

The value of the power plant, however, is not only dependent upon the marginal distribution of each commodity, but rather upon their joint distribution. As a result, we need to introduce a correlation among the risk-factors of each commodity.

This has been done by keeping the parameter estimators that had been estimated on each market, individually, and by computing the correlations between the estimated time series of the different risk-factors that could be retrieved from the Kalman Filters. Estimating the correlations between the risk-factors in a second stage dramatically decreases the computation burdens but does not provide efficient estimators compared to a single step estimation of all the parameters based on the historical time series of spot and futures prices of all the different markets. However, it will provide consistent estimators.

It appeared that only the long-term factors of the different commodities were statistically significantly correlated, showing that the commodities face independent shocks around long term means of reversion that are correlated.

We will thus add to the models detailed in Chapter 7 a correlation matrix between the diffusion shocks of the long-term factors.

$$Corr(EEX, Brent, NBP) = \begin{pmatrix} 1 & 0.292 & 0.327 \\ 0.292 & 1 & 0.786 \\ 0.327 & 0.786 & 1 \end{pmatrix} \tag{4.59}$$

Based on the information available on May 15^{th} 2002, Figure 4.4.1 plots the expected value of the electricity price during the period 2003-2005 (highest seasonal curve with a weekly pattern), together with the expected cost of production with a classical natural gas power plant (lowest seasonal curve) or a fuel oil power plant (highest curve with no seasonality), assuming for both power plants an efficiency of 40%. The graph has been zoomed on the first two weeks in Figure 4.4.2 to show the weekly structure of the data.

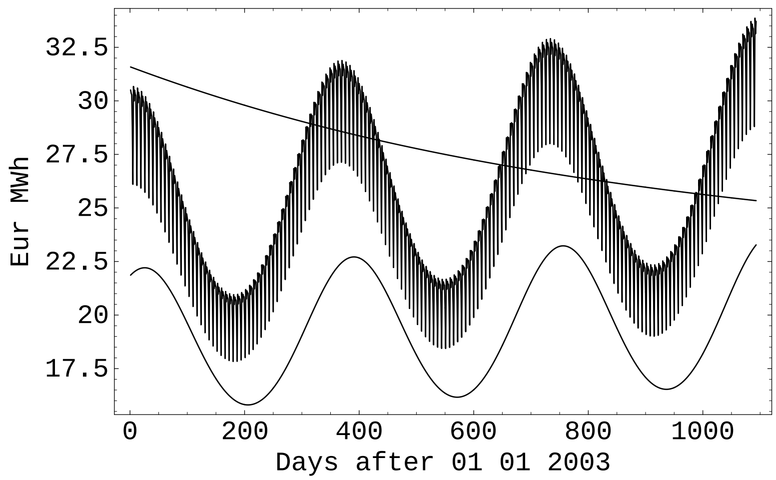

Figure 4.4.1: Expected electricity price (highest seasonal curve with a weekly pattern) and generation costs with a natural gas plant (lowest seasonal curve) and fuel oil plant (highest curve with no seasonality)

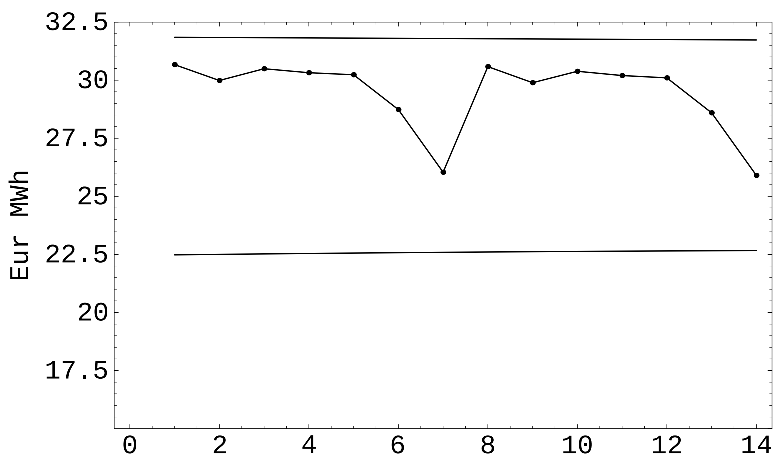

Figure 4.4.2: Zoom on the first two weeks of Figure 4.4.1

Based on these curves, a static Net Present Value approach would compute, at each time step, the expected payoff from operating a power plant just by subtracting to the expectation of the electricity price the expected cost of production.

This approach would provide the expected payoffs plotted in Figure 4.4.3 for the natural gas power plant, Figure 4.4.4 for the fuel oil power plant, and in Figure 4.4.5 for the multi-fuel power plant (it is the same as for the natural gas power plant since expected production costs are always lower with natural gas than with fuel oil).

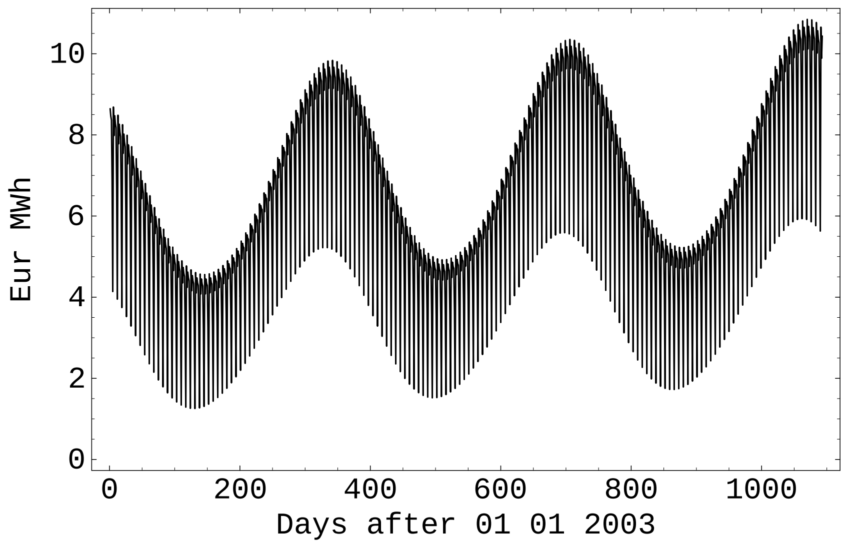

Figure 4.4.3: Expected payoff arising from operating a natural gas power plant using the static NPV approach

The Real Options approach, on the other hand, recognizes that one is not obliged to take the decision now for each time step in the future about whether to operate or shut down the power plant. Instead, it is possible to dynamically manage the power plant and decide to operate or shut down a power plant according to the economic conditions that will be revealed. Equivalently, for a multi-fuel power plant, the choice of the fuel can also be done dynamically. Applying this approach, the expected payoffs can be estimated by computing the transforms $\tilde{G}$ of **Proposition 4.7** with the appropriate parameters, for the single fuel power plants, and, for the multi-fuel power plant, by computing the transforms G of **Proposition 4.11**, as was shown in **Proposition 4.12**. The expected payoffs have now increased to levels

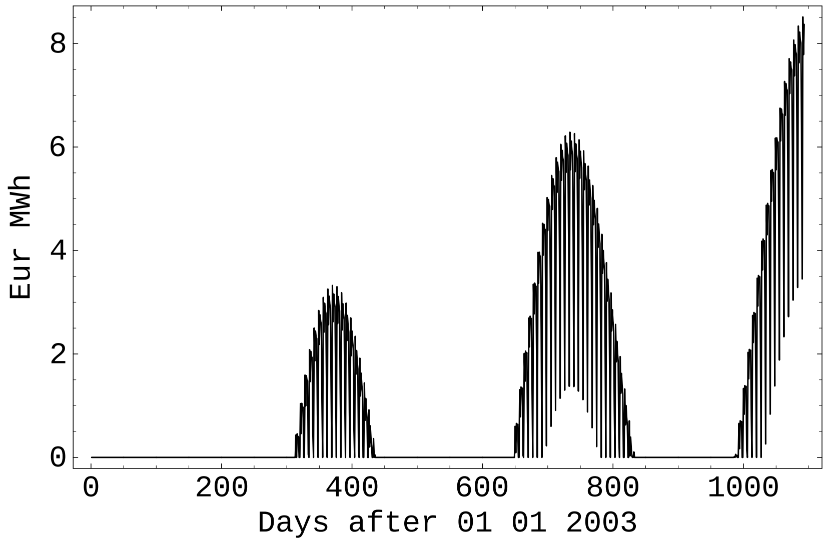

Figure 4.4.4: Expected payoff arising from operating a fuel oil power plant using the static NPV approach

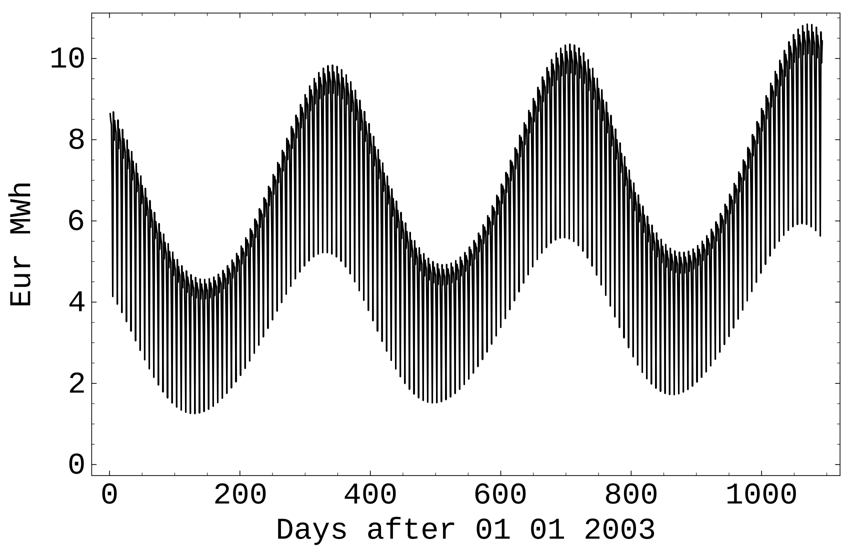

Figure 4.4.5: Expected payoff arising from operating a multi-fuel power plant using the static NPV approach

as plotted in Figure 4.4.6 for the natural gas power plant, Figure 4.4.7 for the fuel oil power plant, and in Figure 4.4.8 for the multi-fuel power plant. Note that the step pattern of the graphs are only due to the fact that, in order to gain some computation time, the expectations have only been estimated once for every day of the week of every block of 4 weeks and have been assumed to be constant during that block.

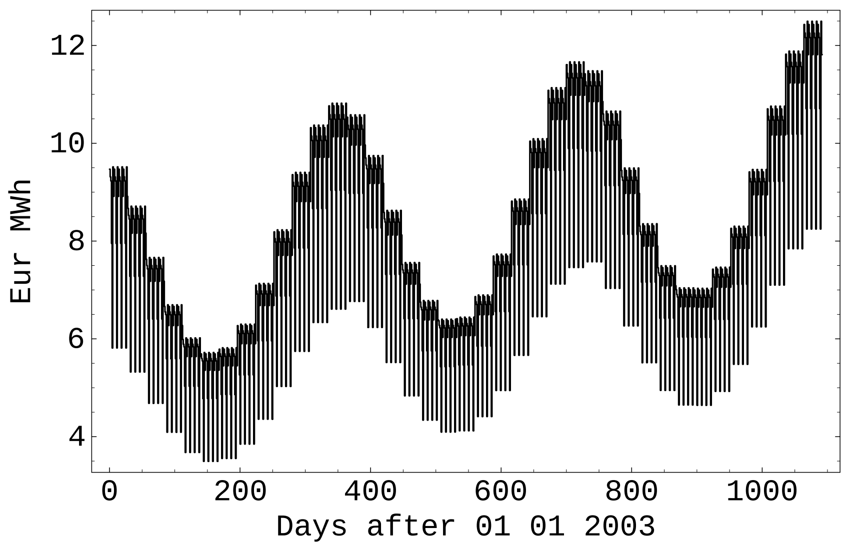

Figure 4.4.6: Expected payoff arising from operating a natural gas power plant using the real options approach

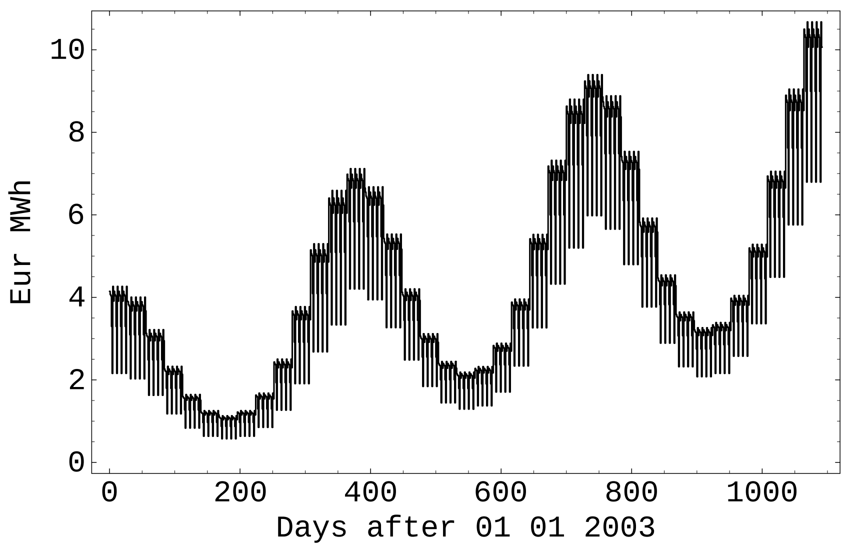

Figure 4.4.7: Expected payoff arising from operating a fuel oil power plant using the real options approach

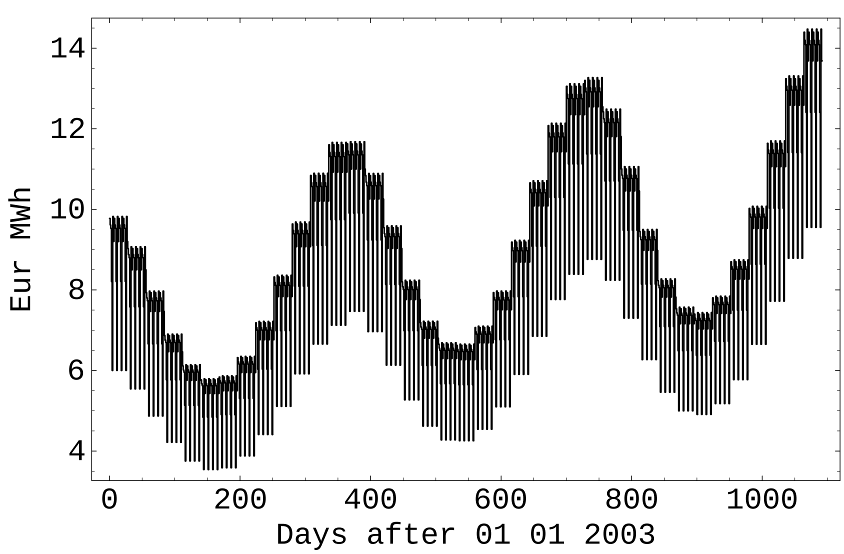

Figure 4.4.8: Expected payoff arising from operating a multi-fuel power plant using the real options approach

The gain in expected payoffs of a multi-fuel power plant compared to a single fuel natural gas power plant are plotted in Figure 4.4.9, and compared to a single fuel fuel oil plant in Figure 4.4.10.

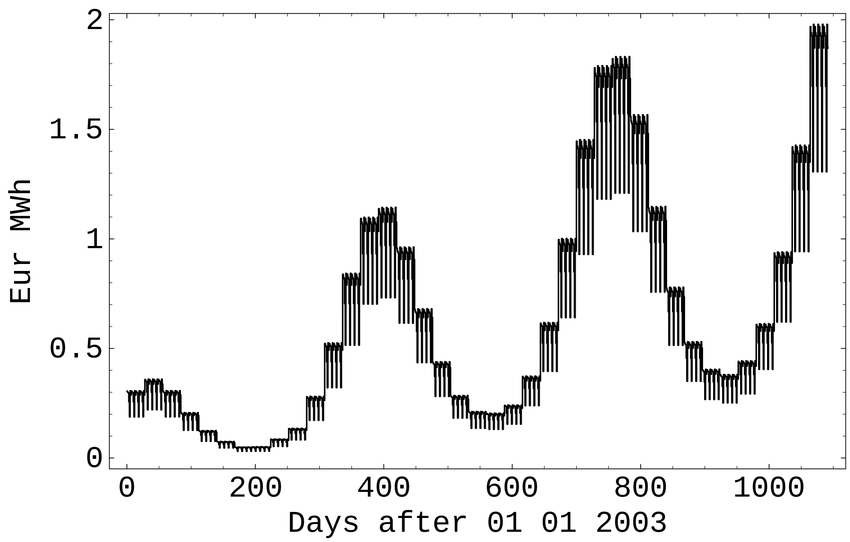

Figure 4.4.9: The gain in expected payoffs of a multi-fuel power plant compared to a single fuel natural gas power plant using the real options approach

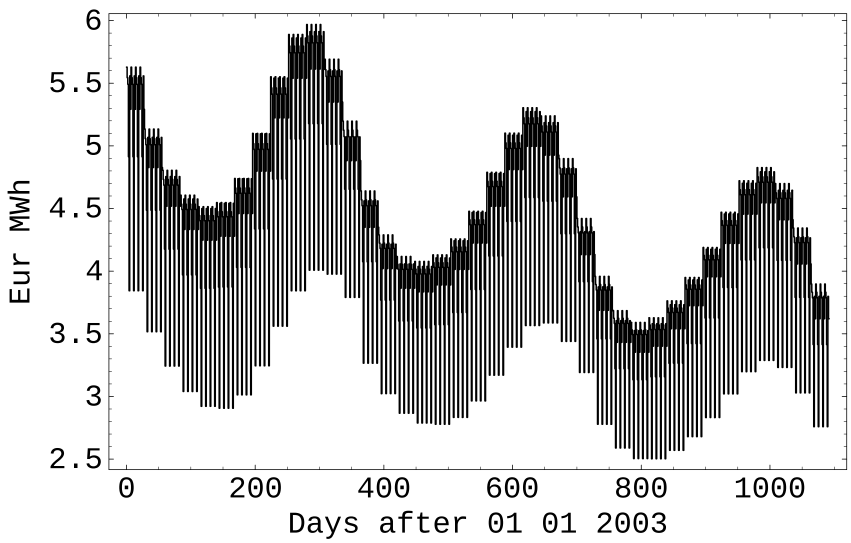

Figure 4.4.10: The gain in expected payoffs of a multi-fuel power plant compared to a single fuel fuel oil power plant using the real options approach

4.5 Asian Contracts

Averages of prices during a fixed period are very common in energy pricing clauses. This can be dealt with by noticing the following.

Proposition 4.13 *Let $I(t) = \int_0^t \gamma.X_s ds$. If X_t is an AJD under a probability measure Q, then $\begin{pmatrix} X \\ I \end{pmatrix}$ is also an AJD under the probability measure Q.*

Proof : By differentiation, we get

$$d(I_t) = \gamma.X_t dt \tag{4.60}$$

so that if X_t is an AJD under a probability measure Q, then $\begin{pmatrix} X \\ I \end{pmatrix}$ is also an AJD under the probability measure Q.$\square$

Obviously, the same holds under the probability measure P.

4.5.1 Elementary Transform

The following multi-temporal transform will make it possible to deal with Asian contracts.

Proposition 4.14 (extended from Kamat and Oren (2000) [26]) : *Let X_t be an AJD vector of size n under the probability measure P, as defined by (2.2), ρ_0 be the constant continuously compounded risk-free rate. The transform $\Phi(d_0, d_1, d_2, a_1, a_2, t, T_1, T_2)$, $\{\mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, \mathbb{C}^n, \mathbb{C}^n, [0\infty), [0\infty), [0\infty)\} \rightarrow \{\mathbb{C}\}$, with $t \leq T_1 \leq T_2$ defined as*

$$\begin{aligned} \Phi(d_0, d_1, d_2, a_1, a_2, t, T_1, T_2) &= \\ & E^P\left[e^{\int_t^{T_2} \rho_0 ds}\Big(d_0 + d_1.X_{T_1} + d_2.X_{T_2}\Big)e^{a_1.X_{T_1} + a_2.X_{T_2}}|\mathcal{F}_t\right] \end{aligned} \tag{4.61}$$

is given by

$$\begin{aligned}\Phi(d_0, d_1, d_2, a_1, a_2, t, T_1, T_2) &= (A_1(t) + B_1(t).X_t)e^{\alpha_1(t)+\beta_1(t).X_t} \\ &\quad +(A_2(t) + B_2(t).X_t)e^{\alpha_2(t)+\beta_2(t).X_t}\end{aligned} \tag{4.62}$$

where $A_1, B_1, \alpha_1, \beta_1$ *and* $A_2, B_2, \alpha_2, \beta_2$ *are solutions of systems of complex valued ODE's.*

Proof : From the definition of the transform Φ (4.61), $e^{-\rho_0 t}\Phi_t$ will be a martingale under the probability measure P. Applying Ito's lemma and replacing Φ by its conjectured form (4.62), we have that $A_1, B_1, \alpha_1, \beta_1$ and $A_2, B_2, \alpha_2, \beta_2$ satisfy the system of complex valued ODE's (7.4).

For the boundary conditions, consider Φ_{T_1}

$$\begin{aligned}\Phi(d_0, d_1, d_2, a_1, a_2, T_1, T_1, T_2) &= e^{-\int_t^{T_1} \rho_0 ds} \\ &\quad \Big((d_0 + d_1.X_{T_1})e^{a_1.X_{T_1}} E^P\left[e^{a_2.X_{T_2}}|\mathcal{F}_{T_1}\right] \\ &\quad +e^{a_1.X_{T_1}} E^P\left[d_2.X_{T_2}e^{a_2.X_{T_2}}|\mathcal{F}_{T_1}\right]\Big)\end{aligned} \tag{4.63}$$

From equations (7.1) and (7.2), this can be expressed as

$$\begin{aligned}\Phi(d_0, d_1, d_2, a_1, a_2, T_1, T_1, T_2) &= e^{-\int_t^{T_1} \rho_0 ds} \\ &\quad \Big((d_0 + d_1.X_{T_1})e^{a_1.X_{T_1}} e^{\alpha^*(T_1)+\beta^*(T_1).X_{T_1}} \\ &\quad +e^{a_1.X_{T_1}}(A^{**}(T_1) + B^{**}(T_1).X_{T_1})e^{\alpha^{**}(T_1)+\beta^{**}(T_1).X_{T_1}}\Big)\end{aligned} \tag{4.64}$$

This implies the boundary conditions

$$A_1(T_1) = e^{-\int_t^{T_1} \rho_0 ds} d_0 \tag{4.65}$$

$$B_1(T_1) = e^{-\int_t^{T_1} \rho_0 ds} d_1 \tag{4.66}$$

$$\alpha_1(T_1) = \alpha^*(T_1) \tag{4.67}$$

$$\beta_1(T_1) = \beta^*(T_1) + a_1 \tag{4.68}$$

and

$$
\begin{aligned}
A_2(T_1) &= e^{-\int_t^{T_1}\rho_0 ds}A^{**}(T_1) && (4.69)\\
B_2(T_1) &= e^{-\int_t^{T_1}\rho_0 ds}B^{**}(T_1) && (4.70)\\
\alpha_2(T_1) &= \alpha^{**}(T_1) && (4.71)\\
\beta_2(T_1) &= \beta^{**}(T_1) + a_1 && (4.72)
\end{aligned}
$$

□

As mentioned, this proposition is an extension of the work of Kamat and Oren (2000) [26] who gave the solution for the particular case where $d_0 = 1$, $d_1 = 0$ and $d_2 = 0$. In this case, they show that the solution can be expressed in a simpler form

$$\Phi(d_0 = 1, d_1 = 0, d_2 = 0, a_1, a_2, t, T_1, T_2) = e^{\alpha(t)+\beta(t).X_t} \tag{4.73}$$

4.5.2 Products Priced

4.5.2.1 Asian futures and forwards

In this section, we will restrict our framework to a constant instantaneous risk-free rates of return, ρ_0. As a result futures and forward contracts have the same price so that we will only refer to futures to denote both kinds of contract.

Proposition 4.15 *Consider a spot price given by the level model (2.3), with X_t an AJD under the risk-neutral probability measure Q, then the price at time t of a futures contract with a delivery period going from time T_1 to time T_2, $f(t, T_1, T_2)$, is given by a special case of the transform Φ.*

Proof : Consider the augmented vector $Y_t = \begin{pmatrix} X_t \\ I_t \end{pmatrix}$ where $I(t) = \int_0^t e^{-\rho_0 s}\gamma_2.X_s ds$. From **Proposition 4.13**, Y_t is also an AJD under Q.

From the no-arbitrage argument, the futures price will then be given by the solution of the equation

$$E^Q\Big[\int_{T_1}^{T_2} e^{-\rho_0(s-t)}\Big(Spot(s) - f(t,T_1,T_2)\Big)ds\Big|\mathcal{F}_t\Big] = 0 \tag{4.74}$$

where ρ_0 denotes the risk-free rate of return.

This can be written as

$$\begin{aligned}
f(t,T_1,T_2) &= \frac{E^Q\Big[\int_{T_1}^{T_2} e^{-\rho_0(s-t)}Spot(s)ds\Big|\mathcal{F}_t\Big]}{E^Q\Big[\int_{T_1}^{T_2} e^{-\rho_0(s-t)}ds\Big|\mathcal{F}_t\Big]} \qquad (4.75)\\
&= \gamma_{0,2} + \frac{\rho_0}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} E^Q\left[\Big(I(T_2) - I(T_1)\Big)\Big|\mathcal{F}_t\right]\\
&= \gamma_{0,2} + \frac{\rho_0}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} E^Q\left[[0,\ldots,0,1]Y(T_2) - [0,\ldots,0,1]Y(T_1)|\mathcal{F}_t\right]\\
&= \gamma_{0,2} + \frac{\rho_0}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} \Phi(0,-[0,\ldots,0,1]^T,[0,\ldots,0,1]^T,0,0,t,T_1,T_2)\\
&= \gamma_{0,2} + \frac{\rho_0}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} \Big(A_1(t) + A_2(t) + \Big(B_1(t) + B_2(t)\Big)X_t\Big)\\
&\equiv A(t) + B(t).X(t)
\end{aligned}$$

where ρ_0 in (4.61) is set to 0.□

Proposition 4.16 *Consider a spot price given by the return model (2.1), with X_t an AJD under the risk-neutral probability measure Q, then the price at time t of a futures contract with a delivery period going from time T_1 to time T_2, $f(t,T_1,T_2)$, can be approximated by a special case of the transform Φ. The approximation consists in replacing an arithmetic average by its geometric average.*

Proof : Consider the augmented vector $Y_t = \begin{pmatrix} X_t \\ I_t \end{pmatrix}$ where $I(t) = \int_0^t \gamma_1.X_s ds$. From **Proposition 4.13**, Y_t is also an AJD under Q.

As previously, the no-arbitrage argument guarantees that

$$f(t,T_1,T_2) = \frac{E^Q\Big[\int_{T_1}^{T_2} e^{-\rho_0(s-t)}Spot(s)ds\Big|\mathcal{F}_t\Big]}{E^Q\Big[\int_{T_1}^{T_2} e^{-\rho_0(s-t)}ds\Big|\mathcal{F}_t\Big]} \tag{4.76}$$

$$\begin{aligned}
&= \frac{E^Q\left[\int_{T_1}^{T_2} e^{-\rho_0 s} e^{\gamma_{0,1}+\gamma_1 X(s)} ds|\mathcal{F}_t\right]}{E^Q\left[\int_{T_1}^{T_2} e^{-\rho_0 s} ds\Big|\mathcal{F}_t\right]} \\
&\approx \frac{\rho_0 e^{\gamma_{0,1}}}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} E^Q\left[e^{\int_{T_1}^{T_2}(-\rho_0 s+\gamma_1 X(s))ds}|\mathcal{F}_t\right] \\
&= \frac{\rho_0 e^{\gamma_{0,1}-\rho_0(T_2^2-T_1^2)/2}}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} E^Q\left[e^{I(T_2)-I(T_1)}|\mathcal{F}_t\right] \\
&= \frac{\rho_0 e^{\gamma_{0,1}-\rho_0(T_2^2-T_1^2)/2}}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} E^Q\left[([0,\ldots,0,1]Y(T_2) - [0,\ldots,0,1]Y(T_1))|\mathcal{F}_t\right] \\
&= \frac{\rho_0 e^{\gamma_{0,1}-\rho_0(T_2^2-T_1^2)/2}}{e^{-\rho_0 T_1} - e^{-\rho_0 T_2}} \Phi(1,0,0,-[0,\ldots,0,1]^T,[0,\ldots,0,1]^T,t,T_1,T_2) \\
&= e^{\alpha_1(t)+\beta_1(t).X_t}
\end{aligned}$$

where ρ_0 in (4.61) is set to 0.□

4.5.2.2 European Options on Asian futures and forwards

Consider an call option with strike price K, and maturity T on an Asian futures contract on the average of the spot price from T_1 to T_2 with $T < T_1 < T_2$.

The payoff of such an option is thus given by

$$Payoff(T) = max\left\{f(T,T_1,T_2) - K; 0\right\} \tag{4.77}$$

where $f(T,T_1,T_2)$ denotes the price at time T of an asian futures contract on the average of the spot price from T_1 to T_2.

Let us first consider the case of a spot price modelled by the level model (2.3), with X_t an AJD under the risk-neutral probability measure Q. The average of the spot price from T_1 to T_2 will thus refer to the arithmetic average.

From **Proposition 4.15**, the price at time t of such an Asian futures contract is given by

$$f(t,T_1,T_2) = A(t) + B(t).X(t) \tag{4.78}$$

The value of the call option, $c(t)$, is thus given by:

$$
\begin{aligned}
c(t) &= E^Q\Big[e^{-\int_t^T R(s)ds}\Big(B(T).X(T) + A(T) - K\Big)\mathbf{1}_{-B(T).X(T) \le A(T) - K}\Big|\mathcal{F}_t\Big] \\
&= \tilde{G}(A(T) - K, A(T) - K, B(T), 0, -B(T), X_t, t, T) \qquad (4.79)
\end{aligned}
$$

If we now consider the case of a spot price modelled by the return model (2.1), with X_t an AJD under the risk-neutral probability measure Q. The average of the spot price from T_1 to T_2 will then refer to the geometric average.

From **Proposition 4.16**, the price at time t of such an Asian futures contract is given by

$$
f(t, T_1, T_2) = e^{\alpha(t) + \beta(t).X(t)} \qquad (4.80)
$$

The value of the call option, $c(t)$, is thus given by:

$$
\begin{aligned}
c(t) &= E^Q\Big[e^{-\int_t^T R(s)ds} e^{\alpha(T) + \beta(T).X(T)}\mathbf{1}_{-\beta(T).X(T) \le \alpha(T) - ln(K)}\Big|\mathcal{F}_t\Big] \qquad (4.81) \\
&\quad -KE^Q\Big[e^{-\int_t^T R(s)ds}\mathbf{1}_{-\beta(T).X(T) \le \alpha(T) - ln(K)}\Big|\mathcal{F}_t\Big] \\
&= \tilde{G}(\alpha(T) - ln(K), e^{\alpha(T)}, 0, e^{\beta(T)}, -\beta(T), X_t, t, T) \\
&\quad -K\tilde{G}(\alpha(T) - ln(K), 1, 0, 0, -\beta(T), X_t, t, T)
\end{aligned}
$$

4.6 Options with Multi-Temporal Payoffs

Options with Multi-Temporal Payoffs include options with payoffs that are functions of the state variables' values at different times in the future.

4.6.1 Elementary Transform

Proposition 4.17 *The transform G, (4.45), can be extended into the transform* $\Gamma(y_1, \ldots, y_m, d_0, d_{1,1}, d_{1,2}, a_1, a_2, b_{1,1}, b_{2,1}, \ldots, b_{1,m}, b_{2,m}, X_t, t, T_1, T_2)$, $\{\mathbb{R}, \ldots, \mathbb{R}, \mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, \ldots, \mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^n, [0,\infty), [0,\infty), [0,\infty)\} \to \{\mathbb{R}\}$, *defined as*

$$\begin{aligned}\Gamma(y_1, \ldots, y_m, d_0, d_{1,1}, d_{1,2}, a_1, a_2, b_{1,1}, b_{2,1}, \ldots, b_{1,m}, b_{2,m}, X_t, t, T_1, T_2) &= \\ E^P\Big[\exp\left(-\int_t^T \rho_0 ds\right)(d_0 + d_{1,1}.X(T_1) + d_{1,2}.X(T_2)) \\ e^{a_1.X(T_1)+a_2.X(T_2)} \prod_{j=1}^m \mathbf{1}_{b_{1,j}.X(T_1)+b_{2,j}.X(T_2) \le y_j} |\mathcal{F}_t\Big]\end{aligned} \tag{4.82}$$

which can also be solved in closed form, up to a m-dimensional integration.

Proof : Consider the case where $m = 2$, extrapolation to any higher number of constraints is indeed again straightforward. Following the same steps as for the proof of **Proposition 4.11**, we get

$$\Gamma(y_1, y_2, d_0, d_{1,1}, d_{1,2}, a_1, a_2, b_{1,1}, b_{2,1}, b_{1,2}, b_{2,2}, X_t, t, T_1, T_2) = \tag{4.83}$$

$$\frac{\Phi(d_0, d_{1,1}, d_{1,2}, a_1, a_2, X_t, t, T_1, T_2)}{4}$$
$$-\frac{1}{2\pi}\sum_{j=1}^{2}\int_0^\infty \frac{Im\left[e^{-iwy_j}\Phi(d_0, d_{1,1}, d_{1,2}, a_1 + iwb_{1,j}, a_2 + iwb_{2,j}, X_t, t, T_1, T_2)\right]}{w}dw$$
$$+\frac{1}{2\pi^2}\int_0^\infty\int_0^\infty \left\{ \frac{Re\left[\begin{array}{r} e^{-i(w_1y_1 - w_2y_2)}\Phi(d_0, d_{1,1}, d_{1,2}, a_1 + iw_1b_{1,1} - iw_2b_{1,2}, \\ a_2 + iw_1b_{2,1} - iw_2b_{2,2}, X_t, t, T_1, T_2)\end{array}\right]}{w_1w_2} \right.$$

$$-\frac{Re\left[\begin{array}{r} e^{-i(w_1y_1+w_2y_2)}\Phi(d_0,d_{1,1},d_{1,2},a_1+iw_1b_{1,1}+iw_2b_{1,2}, \\ a_2+iw_1b_{2,1}+iw_2b_{2,2},X_t,t,T_1,T_2) \end{array}\right]}{w_1w_2}\left.\vphantom{\int}\right\}dw_1dw_2$$

$\square$.

Finally, it is quite trivial to extend the transforms Γ and Φ to the case where there are more than two maturities T_1 and T_2.

4.6.2 Products Priced

4.6.2.1 Asian Options

Proposition 4.18 *Consider the spot price modeled by the level model (2.3), with X_t an AJD under the risk-neutral probability measure Q, then the price at time t of an (arithmetic) Asian call option, $c(t)$, with a payoff at maturity T_2 given by*

$$Payoff(T_2) = max\left\{\frac{1}{T_2-T_1}\int_{T_1}^{T_2} Spot_s ds - K; 0\right\} \tag{4.84}$$

can be expressed as a sum of special cases of the transform Γ if the risk-free rate is assumed to be constant and equal to ρ_0.

Proof Consider the augmented vector $Y_t = \begin{pmatrix} X_t \\ I_t \end{pmatrix}$ where $I(t) = \int_0^t \gamma_2.X_s ds$. From **Proposition 4.13**, Y_t is also an AJD under Q.

The price of the arithmetic Asian call can now be expressed as

$$\begin{aligned} c(t) &= E^Q\left[e^{-\rho_0(T_2-t)}max\left\{\frac{1}{T_2-T_1}\int_{T_1}^{T_2}(\gamma_{0,2}+\gamma_2.X_s)ds - K; 0\right\}|\mathcal{F}_t\right] \qquad (4.85) \\ &= E^Q\Big[\frac{e^{-\rho_0(T_2-t)}}{T_2-T_1}\Big(\gamma_{0,2}(T_2-T_1)+Y_{n+1}(T_2)-Y_{n+1}(T_1)-K(T_2-T_1)\Big) \\ &\qquad \mathbf{1}_{Y_{n+1}(T_1)-Y_{n+1}(T_2)\leq\gamma_{0,2}(T_2-T_1)-K}|\mathcal{F}_t\Big] \end{aligned}$$

which is a special case of Γ. $\square$

Proposition 4.19 *Consider the spot price modeled by the return model (2.1), with* X_t *an AJD under the risk-neutral probability measure* Q*, then the price at time* t *of a (geometric) Asian call option,* $c(t)$*, with a payoff at maturity* T_2 *given by*

$$Payoff(T_2) = max\left\{e^{\frac{1}{T_2-T_1}\int_{T_1}^{T_2} ln(Spot_s)ds} - K; 0\right\} \tag{4.86}$$

can be expressed as a sum of special cases of the transform Γ *if the risk-free rate is assumed to be constant and equal to* ρ_0*.*

Proof Consider the augmented vector $Y_t = \begin{pmatrix} X_t \\ I_t \end{pmatrix}$ where $I(t) = \int_0^t \gamma_1.X_s ds$. From **Proposition 4.13**, Y_t is also an AJD under Q.

The price of the arithmetic Asian call can now be expressed as

$$\begin{aligned} c(t) &= E^Q\left[e^{-\rho_0(T_2-t)} max\left\{e^{\frac{1}{T_2-T_1}\int_{T_1}^{T_2}(\gamma_{0,1}+\gamma_1.X_s)ds} - K; 0\right\}|\mathcal{F}_t\right] \qquad (4.87) \\ &= E^Q\Big[e^{-\rho_0(T_2-t)} e^{\frac{1}{T_2-T_1}(\gamma_{0,2}(T_2-T_1)+Y_{n+1}(T_2)-Y_{n+1}(T_1))} \\ &\qquad\qquad \mathbf{1}_{Y_{n+1}(T_1)-Y_{n+1}(T_2)\leq\gamma_{0,2}(T_2-T_1)-ln(K)}|\mathcal{F}_t\Big] \\ &\quad -E^Q\left[e^{-\rho_0(T_2-t)} K \mathbf{1}_{Y_{n+1}(T_1)-Y_{n+1}(T_2)\leq\gamma_{0,2}(T_2-T_1)-ln(K)}|\mathcal{F}_t\right] \end{aligned}$$

where all the elements in (4.87) are special cases of the transform Γ. □

Similarly, it is trivial to show that Asian put options are also decomposable into a sum of special cases of the transform Γ.

4.6.2.2 Discretely Monitored Barrier Options

Barrier options are options where the payoff depends on whether the underlying asset's price reaches a certain level during a certain period of time.

A number of different types of barrier options regularly trade in the over-the-counter market. They are attractive to some market participants because they are less expensive than the corresponding regular options. These barrier options can be classified as either *knock-out options* or *knock-in options*. A knock-out option is an option that ceases to exist when the underlying asset price reaches a certain barrier. A knock-in option is an option that comes into existence only when the underlying asset price reaches a barrier.

Discretely monitored Barrier options are barrier options where the underlying price is only monitored (to check whether the barrier has been reached) at some pre-specified discrete times in the future.

We will consider here the case of a discretely monitored *down-and-out* call option. The extension to any other kind of discretely monitored barrier option is trivial. A discretely monitored *down-and-out* call option is an option with payoff at maturity T equal to $Spot(T) - K$, where K is the strike price, unless the spot price has gone under a barrier level H_i at any of the monitoring times $\tau_i < T$, in which case the option has a value of zero.

The price of this option, $c(t)$, can be expressed as

$$c(t) = E^Q\Big[e^{-\rho_0 T}\Big(Spot(T) - K\Big)\mathbf{1}_{Spot(T)>K}\prod_i \mathbf{1}_{Spot(\tau_i)>H_i}\Big|\mathcal{F}_t\Big] \tag{4.88}$$

Considering a spot price given by the level model (2.3), this can be rewritten as

$$\begin{aligned} c(t) \quad &= \quad E^Q\Big[e^{-\rho_0 T}\Big(\gamma_2.X(T) + \gamma_{0,2} - K\Big)\mathbf{1}_{-\gamma_2.X(T)<\gamma_{0,2}-K} \\ &\qquad\qquad \prod_i \mathbf{1}_{-\gamma_2.X(\tau_i)<\gamma_{0,2}-H_i}\Big|\mathcal{F}_t\Big] \end{aligned} \tag{4.89}$$

which is indeed a special case of the transform Γ.

Considering now a spot price given by the return model(2.1), equation (4.88) can then be rewritten as

$$\begin{aligned} c(t) \quad &= \quad E^Q\Big[e^{-\rho_0 T}e^{\gamma_{0,1}+\gamma_1.X(T)}\mathbf{1}_{-\gamma_1.X(T)<\gamma_{0,1}-ln(K)} \\ &\qquad\qquad \prod_i \mathbf{1}_{-\gamma_1.X(\tau_i)<\gamma_{0,1}-ln(H_i)}\Big|\mathcal{F}_t\Big] \end{aligned}$$

$$-e^{-\rho_0 T} K E^Q \left[\mathbf{1}_{-\gamma_1 . X(T) < \gamma_{0,1} - ln(K)} \prod_i \mathbf{1}_{-\gamma_1 . X(\tau_i) < \gamma_{0,1} - ln(H_i)} \middle| \mathcal{F}_t \right] \quad (4.90)$$

4.7 Interruptible Contracts with Early Notification Option

Interruptible contracts with early notification options are in fact special cases of Options with Multi-Temporal Payoffs. They do not necessitate the development of a new transform to price them, however, because of their relative complex structure, we develop them separately in this section.

Oren (1996) [33] has shown that a forward contract bundled with an appropriate double call option provides a "perfect hedge" for customers that can curtail loads in response to high spot prices and can mitigate their curtailment losses when the curtailment decision is made with sufficient lead time. Kamat and Oren (2000) [26] give pricing formulas for this double call contract under the assumption of a mean reverting price process with jumps. In this paper we extend the possible price processes to the entire Affine Jump Diffusion family (including multi factors and stochastic volatility).

A double call option contract is defined by two exercise times T_1 and T_2 together with two exercise prices k_1 and k_2. Its payoff at time T_2 is the maximum of the spot at time T_2 minus the time two exercise price k_2 and zero, unless it has already been exercised at time T_1, in which case its payoff is the futures price at time T_1 with maturity at time T_2 minus the time one exercise price k_1.

Combining a long position in a futures contract with maturity at time T_2, $f(t, T_2)$, with a short position in a double call contract is thus equivalent to being supplied by an interruptible contract with an additional early notification option. If we choose the exercise prices k_1 and k_2 equal to the forward price, $f(t, T_2)$, minus the premium of the double call, and plus the shortage costs with and without notification, x_1 and x_2, respectively ($k_1 = f(t, T_2) - premium + x_1$ and $k_2 = f(t, T_2) - premium + x_2$), this combination allows customers to save the premium of the double call contract on their electricity supply costs by letting them benefit of their ability to curtail loads in response to high spot prices and to mitigate their curtailment losses when the curtailment decision is made with sufficient lead time.

By taking a long position in a forward contract, $f(t, T_2)$, a consumer is indeed sure to be delivered at time T_2 at the forward price. Let us now consider the case of a consumer that would agree to be curtailed if prices became too high and who would face a shortage cost of x_1 or x_2, according to whether the curtailment is done with early notification or not, respectively.

This consumer could reduce the delivery price of the forward during normal conditions (no curtailment) by selling a double call option and thus receiving the value of the double call premium.

Furthermore, if the double call has strike prices equal to the shortage costs plus the forward price, and minus the premium of the double call ($k_1 = f(t, T_2) - premium + x_1$ and $k_2 = f(t, T_2) - premium + x_2$), the consumer would then be reimbursed exactly its shortage costs, x_1 and x_2. Indeed, if the curtailment is done with early notification (at time T_1), the consumer must pay the difference between the forward price at that time and the first strike price: $Payoff(T_1) = f(T_1, T_2) - k_1$. If he then enters into a short position in the future $f(T_1, T_2)$, his payoff at time T_2 will then be $Payoff(T_2) = Spot(T_2) - f(t, T_2) - (Spot(T_2) - f(T_1, T_2))$. Summing up his payoffs at times t, T_1 and T_2, we get: $premium + Payoff(T_1) + Payoff(T_2) = premium + k1 - f(t, T_2) = x_1$, which is indeed his shortage cost with early notification.

If, on the other hand, the double call is exercised at time T_2 (with no early notification), than the consumer's payoff at time T_2 is given by: $Payoff(T_2) = Spot(T_2) - f(t, T_2) - (Spot(T_2) - k_2) = k_2 - f(t, T_2)$. Adding to this the double call premium received at time t, we gat a total payoff equal to $premium + k_2 - f(t, T_2) = x_2$, which is his shortage cost without early notification.

In the next sections we will show how double calls can be priced using the methodology we developed for Options with Multi-Temporal Payoffs.

4.7.1 Double call on futures

Consider the spot price follows an AJD model under the risk-neutral probability measure as defined in Chapter 2.

Let us define the function $\bar{k}(X_{T_1})$, solution of the equation:

$$\bar{k}(X_{T_1}) - k_1 = C_{T_2}\Big(k_2|X_{T_1}\Big) \tag{4.91}$$

where $C_{T_2}\Big(k_2|X_{T_1}\Big)$ is the value at time T_1 of a call on the spot price with maturity T_2 and a strike price k_2.

Clearly, this function divides the state-space in two regions : if the futures price at time T_1 with maturity at time T_2 lies over this value, it is optimal to exercise the option at time T_1, otherwise, it is optimal to keep the second option alive.

The value of the double call at time T_1 is thus:

$$f(T_1, T_2) - k_1 \qquad \text{if} \qquad f(T_1, T_2) > \bar{k}(X_{T_1}) \tag{4.92}$$

$$C_{T_2}(k_2|X_{T_1}) \qquad \text{if} \qquad f(T_1, T_2) \leq \bar{k}(X_{T_1}) \tag{4.93}$$

Again, by definition of the risk neutral probability measure, the price of the double call, $C_{T_1,T_2}(k_1, k_2|X_t)$, is then given by :

$$\begin{aligned} C_{T_1,T_2}(k_1, k_2|X_t) &= e^{-r(T_1-t)}E^Q\Big[(f(T_1, T_2) - k_1)\mathbf{1}_{f(T_1,T_2)>\bar{k}(X_{T_1})} \\ &+ e^{-r(T_2-T_1)}E^Q\Big[(f(T_2, T_2) - k_2)\mathbf{1}_{f(T_2,T_2)>k_2}|\mathcal{F}_{T_1}\Big]\mathbf{1}_{f(T_1,T_2)\leq\bar{k}(X_{T_1})}\Big|\mathcal{F}_t\Big] \end{aligned} \tag{4.94}$$

If the spot price model is a "return" model, we consider the following linearization of the threshold function

$$ln(\bar{k}(X_{T_1})) \simeq \delta_{0,1} + \delta_1 X_{T_1} \tag{4.95}$$

where the parameters $\delta_{0,1}$ and δ_1 can be computed as shown in **Chapter 5**.

We can now rewrite equation (4.94) as

$$\begin{aligned} C_{T_1,T_2}(k_1,k_2|Y(t)) &= e^{-r(T_1-t)}E^Q\Big[\left(e^{\alpha(T_1)+\beta(T_1)X_{T_1}}-k_1\right) \\ &\qquad\qquad \mathbf{1}_{\alpha(T_1)+\beta(T_1)X_{T_1}>\delta_{0,1}+\delta_1 X_{T_1}}\Big|\mathcal{F}_t\Big] \\ &\quad +e^{-r(T_2-t)}E^Q\Big[\left(e^{\gamma_{0,1}+\gamma_1.X_{T_2}}-k_2\right)\mathbf{1}_{\gamma_{0,1}+\gamma_1.X_{T_2}>k_2} \\ &\qquad\qquad \mathbf{1}_{\alpha(T_1)+\beta(T_1)X_{T_1}\leq\delta_{0,1}+\delta_1 X_{T_1}}\Big|\mathcal{F}_t\Big] \\ &\equiv e^{-r(T_1-t)}R1+e^{-r(T_2-t)}R2 \end{aligned} \tag{4.96}$$

where $R1$ and $R2$ are special cases of the building blocks $\tilde{G}$ and Γ, respectively.

If the spot price model is a "level" model, we consider the following liberalization of the threshold function

$$\bar{k}(X_{T_1}) \simeq \delta_{0,2}+\delta_2 X_{T_1} \tag{4.97}$$

where, again, the parameters $\delta_{0,2}$ and δ_2 can be computed as shown in **Chapter 5**.

We can now rewrite equation (4.94) as

$$\begin{aligned} C_{T_1,T_2}(k_1,k_2|Y(t)) &= e^{-r(T_1-t)}E^Q\Big[(A(T_1)+B(T_1)X_{T_1}-k_1) \\ &\qquad\qquad \mathbf{1}_{A(T_1)+B(T_1)X_{T_1}>\delta_{0,2}+\delta_2 X_{T_1}}\Big|\mathcal{F}_t\Big] \\ &\quad + e^{-r(T_2-t)}E^Q\Big[(\gamma_{0,2}+\gamma_2.X_{T_2}-k_2)\,\mathbf{1}_{\gamma_{0,2}+\gamma_2.X_{T_2}>k_2} \\ &\qquad\qquad \mathbf{1}_{A(T_1)+B(T_1)X_{T_1}\leq\delta_{0,2}+\delta_2 X_{T_1}}\Big|\mathcal{F}_t\Big] \\ &\equiv e^{-r(T_1-t)}R3+e^{-r(T_2-t)}R4 \end{aligned} \tag{4.98}$$

where $R3$ and $R4$ are special cases of the building blocks $\tilde{G}$ and Γ, respectively.

4.7.2 Note on the linearization of the threshold function

The error arising from approximating the threshold function $\bar{k}(X_{T_1})$ by an affine function results in a sub-optimal exercise policy at time T_1. The estimated value of the double call is thus a lower bound for the true value.

Furthermore, this error can be assumed to be quite small since the region in the state-space where the exercise choice is not optimal remains quite close to the "frontier" where the two alternatives (exercise or not) result in the same value. (See Singleton and Umantsev (2002) [46] for a discussion on the adequacy of a linear threshold function approximation to price options on coupon bonds and swaptions in affine term structure models).

If, however, this approximation is not satisfactory, it is possible to replace it by a piece-wise affine function where the different regions, Ω_j, would be determined by inequalities of the type:

$$\Omega_j : \{X(T_1) \quad | \quad a_{i,j} \leq X_i(T_1) \leq b_{i,j}, \quad i = 1 \dots n\}$$

Equation (4.94) for the return model will then be replaced by

$$\begin{aligned} C_{T_1,T_2}(k_1, k_2|X(t)) &= e^{-r(T_1-t)} E^Q \Big[\left(e^{\alpha(T_1)+\beta(T_1)X(T_1)} - k_1\right) \\ &\qquad \prod_j \mathbf{1}_{\alpha(T_1)+\beta(T_1)X(T_1)>\delta_{0,j}+\delta_j X(T_1)} \\ &\qquad\qquad \prod_i \mathbf{1}_{-X_i(T1)\leq -a_{i,j}} \mathbf{1}_{X_i(T1)\leq b_{i,j}} \Big| \mathcal{F}_t \Big] \\ &+ e^{-r(T_2-t)} E^Q \Big[\left(e^{\gamma_{0,1}+\gamma_1 . X_{T_2}} - k_2\right) \mathbf{1}_{\gamma_{0,1}+\gamma_1 . X_{T_2}>k_2} \\ &\qquad \prod_j \mathbf{1}_{\alpha(T_1)+\beta(T_1)X(T_1)>\delta_{0,j}+\delta_j X(T_1)} \\ &\qquad\qquad \prod_i \mathbf{1}_{-X_i(T1)\leq -a_{i,j}} \mathbf{1}_{X_i(T1)\leq b_{i,j}} \Big| \mathcal{F}_t \Big] \end{aligned} \tag{4.99}$$

which is a "multi-criteria option" and can thus be solved by following the same steps as previously.

Similarly, equation (4.98) for the level model will be replaced by

$$\begin{aligned} C_{T_1,T_2}(k_1,k_2|X(t)) &= e^{-r(T_1-t)}E^Q\Big[(A(T_1)+B(T_1)X_{T_1}-k_1) \\ & \prod_j \mathbf{1}_{A(T_1)+B(T_1)X(T_1)>\delta_{0,j}+\delta_j X(T_1)} \\ & \prod_i \mathbf{1}_{-X_i(T1)\le -a_{i,j}}\mathbf{1}_{X_i(T1)\le b_{i,j}}\Big|\mathcal{F}_t\Big] \\ &+ e^{-r(T_2-t)}E^Q\Big[(\gamma_{0,2}+\gamma_2.X_{T_2}-k_2)\,\mathbf{1}_{\gamma_{0,2}+\gamma_2.X_{T_2}>k_2} \\ & \prod_j \mathbf{1}_{\alpha(T_1)+\beta(T_1)X(T_1)>\delta_{0,j}+\delta_j X(T_1)} \\ & \prod_i \mathbf{1}_{-X_i(T1)\le -a_{i,j}}\mathbf{1}_{X_i(T1)\le b_{i,j}}\Big|\mathcal{F}_t\Big] \end{aligned} \tag{4.100}$$

which is also a "multi-criteria option".

4.7.3 Example of implementation

In this section we provide numerical illustrations of values of double call options. We will value these options on the German EEX market which is one of the principal reference for energy prices in North Western Europe. The model and the parameter estimations are detailed in Chapter 7. The computations are done based on the information available on May 15^{th} 2002. At that date, the price of a baseload forward with delivery on Friday November 15^{th} 2002 (3 months later) is 33.72 Eur/MWh. This is the sure price at which, on May 15 2002, a consumer can reserve one Megawatt hour of baseload energy for 3 months later (November 15^{th} 2002).

Assuming this consumer is ready to be curtailed in the event of very high spot prices on November 15^{th} 2002, he can then sell a call option that matures on November 15^{th} 2002, and with a strike price of, for example, 40 Eur/MWh. He would then receive the premium of this option, which is here equal to 1.95 Eur/MWh. Having sold this call option, the consumer has now reduced its fixed price of delivery by

1.95 Eur/MWh in the case where he has not been curtailed (if the spot price on November 15 2002 is lower than 40 Eur/MWh). If, on the other hand, he is curtailed (if the spot price on November 15^{th} 2002 is greater than 40 Eur/MWh), he will then receive from his forward contract the spot price minus the forward price of 33.72 Eur/MWh, and will have to pay, because of the call option he sold, the spot price minus the strike price of 40 Eur/MWh. In total, he will thus have gotten the call option premium (1.95 Eur/MWh) plus the difference between the strike price and the forward's price (40 Eur/MWh - 33.72 Eur/MWh), which is equal to 1.95 Eur/MWh + 7.28 Eur/MWh = 9.23 Eur/MWh. In any case, this will be of the advantage of the consumer assuming he has shortage costs equal to or lower than 9.23 Eur/MWh.

Let us now assume that, if he is notified 2 weeks (14 days) ahead, the consumer can considerably lower its shutdown costs to, for example, 3.57 Eur/MWh. He can now sell, instead of a simple call option, a double call option with the two maturities being November 1st 2002 and November 15^{th} 2002, and with the two strike prices, 35 Eur/MWh and 40 Eur/MWh. The premium of this double call is 2.29 Eur/MWh, so that the consumer has increased his discount when there is no curtailment from 1.95 Eur/MWh to 2.29 Eur/MWh, and in the case he is curtailed, he will get 3.57 Eur/MWh (= 35 Eur/MWh - 33.72 Eur/MWh + 2.29 Eur/MWh), or 9.57 Eur/MWh (= 40 Eur/MWh - 33.72 Eur/MWh + 2.29 Eur/MWh), according to whether it is done with or without early notification, respectively. We observe that those remunerations indeed cover the consumer's shutdown costs with and without notification, respectively.

Applying the formulae of equation (4.98), Figure 4.7.1 shows how the value of the premium of double calls with the same strike prices (35 Eur/MWh and 40 Eur/MWh) is influenced by the early notification period. We see that if the consumer can lower his shutdown costs to the same level with a shorter notification period, then he will earn a greater premium.

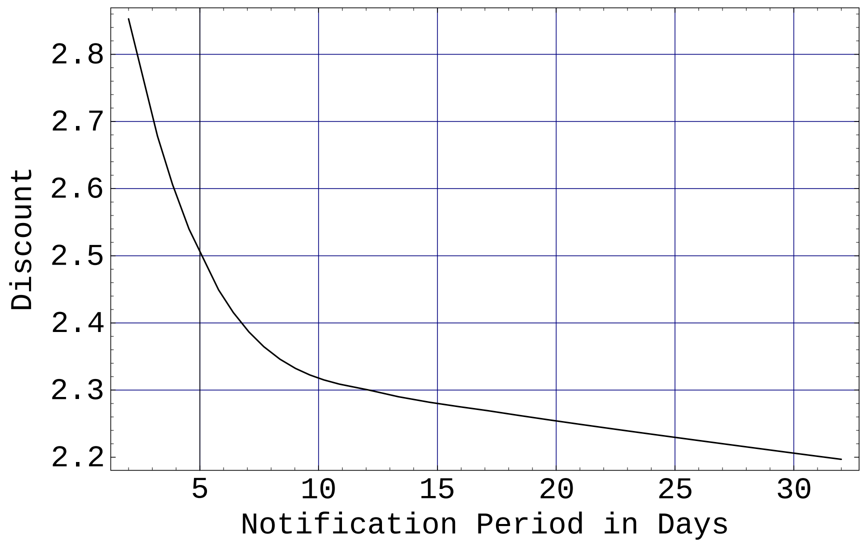

Figure 4.7.1: Double call premium (in Eur/MWh) as a function of the early notification period.

On the other hand, the consumer can also increase the value of its discount if he can further decrease its shutdown costs for the same early notification period. This is illustrated by the Figure 4.7.2 that plots the value of the double call with the same initial 2 weeks notification period and the same second strike price equal to 40 Eur/MWh, but with different first strike prices.

Figure 4.7.2: Double call premium (in Eur/MWh) as a function of the first strike price.

Let us now study the quality of the linear approximation made at time T_1 to determine whether the double call should be exercised or kept alive. We know that it is optimal to exercise the option if the forward price at time T_1 is greater than a threshold $\overline{k}(X_{T_1})$, which is given by equation (4.91). Since the risk-factor describing the spiking behavior of energy prices can only have two values, zero or the spike level, we computed the price of the double call conditional on the fact that there was no spike at time T_1, then conditional ont the fact that there was a spike at time T_1, and weighted the results by their probability of occurrence. Figure 4.7.3 plots the logarithm of this optimal threshold in the case where there is no spike at time T_1. The plot is drawn over the area defined by the 90% confidence intervals for the two risk factors $X_1(T_1)$ and $X_2(T_1)$.

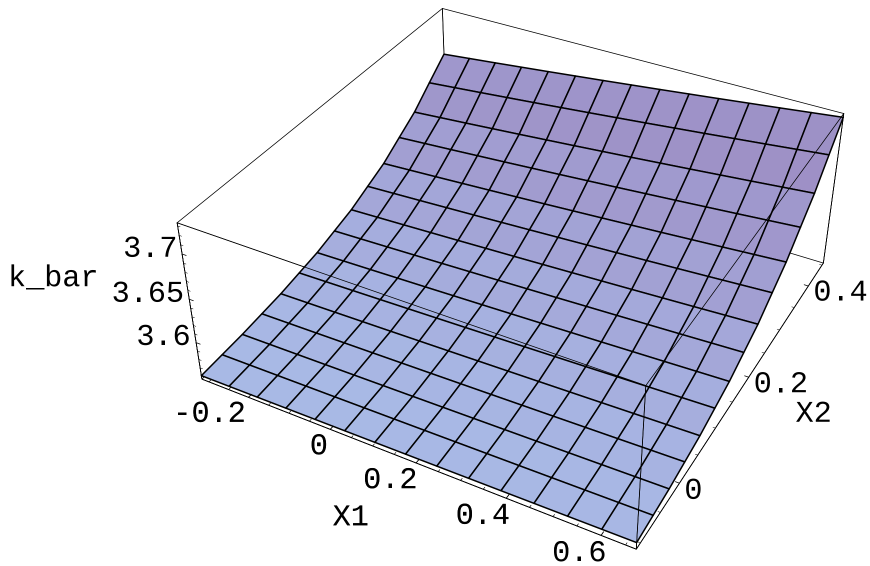

Figure 4.7.3: Logarithm of the optimal threshold level if there is no spike at time T_1.

This threshold will be approximated by the tangent plane drawn in Figure 4.7.4. The tangent is computed at a point where equation (4.91) is an equality and, for example, where $X_1(T_1) = X_2(T_1)$. The area where equation (4.91) is an equality is indeed the area where it is the most crucial to approximate correctly the threshold function.

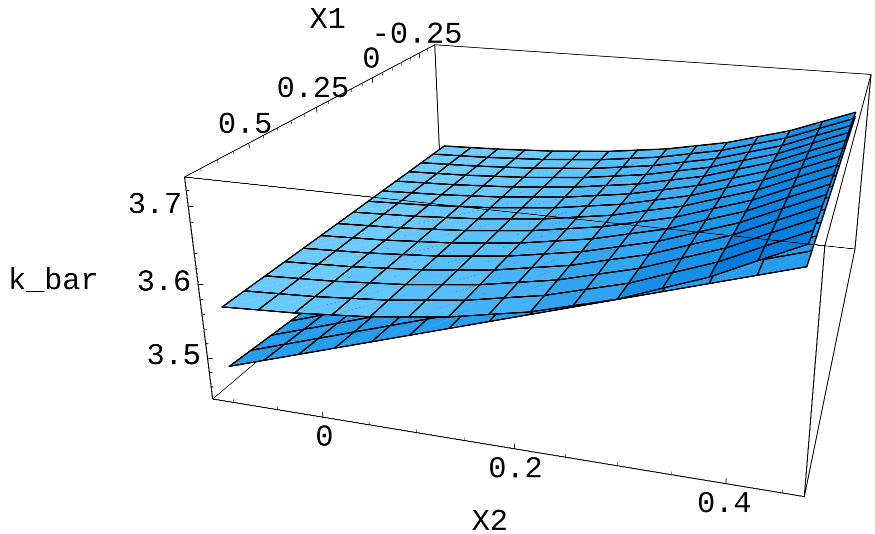

Figure 4.7.4: Logarithm of the Optimal threshold level and its linear approximation if there is no spike at time T_1.

Since the exercise region is defined by the area where the forward price is greater than the threshold, Figure 4.7.5 plots the logarithm of the forward price as a function of the risk factors $X_1(T_1)$ and $X_2(T_1)$, together with the logarithm of the optimal threshold and its linear approximation.

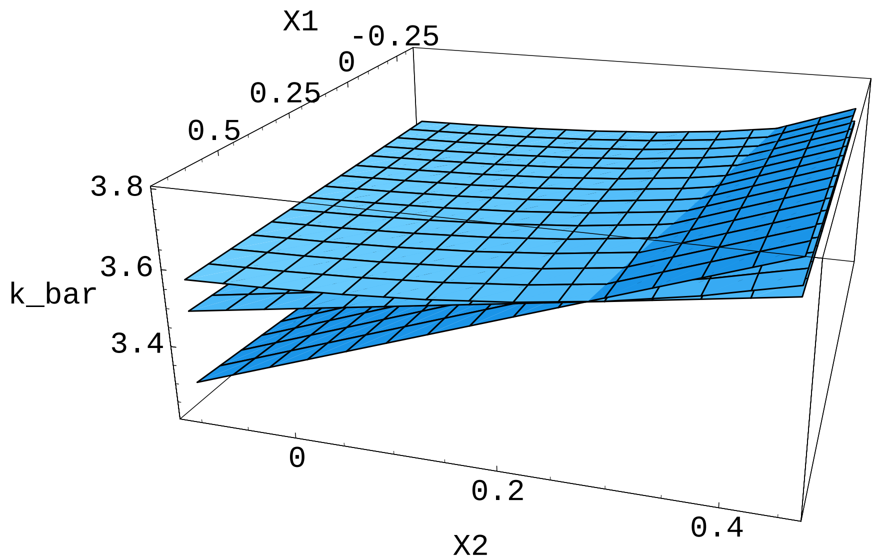

Figure 4.7.5: Logarithm of the forward price together with the logarithm of the Optimal threshold level and its linear approximation, if there is no spike at time T_1.

We see that the exercise region (where the forward is greater than the threshold) defined by the optimal or approximated threshold are almost exactly the same. The area where the approximated threshold implies a non-optimal exercise decision at time T_1 is in fact reduced to a very thin line around the intersection between the plane of the logarithm of the forward price and the surface of the logarithm of the optimal threshold, as shown (with the plotting accuracy that could be reached) in Figure 4.7.6.

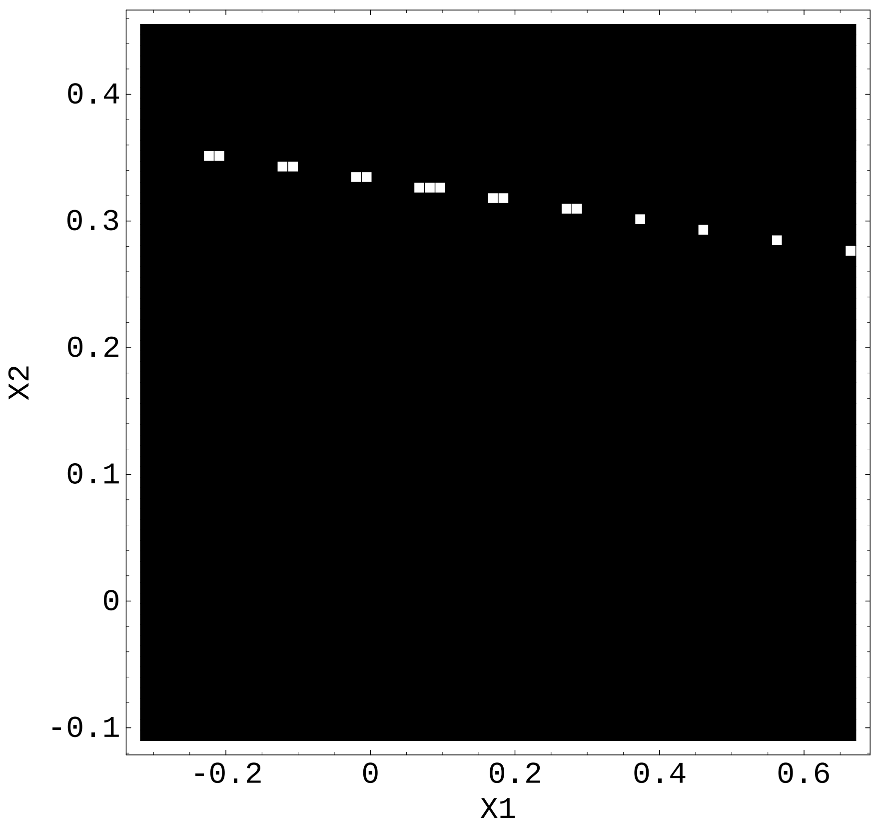

Figure 4.7.6: Region where the approximated threshold implies a non-optimal exercise strategy.

Chapter 5

Option Pricing: Grid Algorithms

5.1 Introduction

Analytic Fourier transforms of payoff functions are not always available. Furthermore, grid algorithms used in a Dynamic Programming (DP) context need to estimate expectations of payoffs that are only known at discrete points of the underlying variables state space. In order to deal with these issues, section 3.2 of this chapter develops a numerical method based on the theoretical results of **Chapter 3** and the discrete Fourier transforms theory. The DP algorithm is reminded in section 3.3.

The numerical approach can be briefly summarized as follows:

1. specify an interval of interest in the domain of the payoff f. This interval is noted $[a, b], a, b \in \Re^n$ and represents the interval $\prod_i [a_i, b_i] \in \Re^n$
2. map this interval to a standard interval and define g equivalently over this interval. This standard interval is $[0, \frac{1}{2}]^n \in \Re^n$
3. symmetrize g and make it periodic (as the fast Fourier transform (FFT) makes the hypothesis that the function is periodic, this avoids discontinuities at the

bound of the interval leading to a faster decay of the Fourier coefficients) around the origin to give g_{sp}

4. apply a fast Fourier transform (FFT) to g_{sp}

5. compute $\langle f, u\rangle = \langle \hat{f}, \check{\hat{u}}\rangle$ in this new framework

Van Steenkiste and Foresi (1999) [48], on the other hand, first compute Arrow-Debreu prices by taking the FFT of the characteristic function of the underlying, and then compute the expected payoff. The main difference with our approach is thus that they compute the FFT on the characteristic function rather than on the payoff function as in our approach. There are several advantages of computing the FFT on the payoff function, rather than on the characteristic function:

- Assuming that the most costly operations in both algorithms are the evaluations of the characteristic function, computing the FFT on the characteristic function will require N evaluations, where N stands for the number of nodes in the grid. If the FFT is computed on the payoff function, however, the number of evaluations of the characteristic function can be dramatically reduced if the grid of the FFT coefficients has a sparse structure, which is often the case.

- Dynamic Programming algorithms will require an FFT computation for each node at each time-step if the FFT is done on the characteristic function, whereas it will only require an FFT computation for each time-step if the FFT is done on the payoff function.

5.2 Discrete Fourier transforms (FFT & FCT)

5.2.1 The interval of interest

The specification of the interval depends heavily on the shape of the integrand in

$$\langle f, u\rangle = \int_{\Re^n} f(x)u(x)\mathrm{dx}; \tag{5.1}$$

Informally, the interval $[a, b]$ should essentially be such that

$$\int_{[a,b]} f(x)u(x)\mathrm{dx} \ = \int_{\Re^n} f(x)u(x)\mathrm{dx} \ + \epsilon \tag{5.2}$$

with ϵ the degree of accuracy needed. Practically, the interval is chosen wide enough empirically.

5.2.2 Mapping the interval

Mapping the interval $[a, b]$ over $[0, \frac{1}{2}]^n$ is not mandatory but enhance notations in the following steps of the reasoning.

Let $g : R^n \to R^n$ be defined as

$$f = \tau_a \partial_{2(b-a)} g \tag{5.3}$$

which has $[0, \frac{1}{2}]^n$ as interval of interest.

Using the result of **Theorem 3.1**, the initial relation becomes (without any approximation):

$$\begin{aligned}
\langle f, u\rangle &= \langle \tau_a \partial_{2(b-a)} g, u\rangle && (5.4)\\
&= \langle g, \partial_{\frac{1}{2(b-a)}} \tau_{-a} u\rangle \prod_i |2(b_i - a_i)| && (5.5)\\
&= \langle \hat{g}, \partial_{-I} \mathfrak{F}\left(\partial_{\frac{1}{2(b-a)}} \tau_{-a} u\right)\rangle \prod_i |2(b_i - a_i)| && (5.6)\\
&= \langle \hat{g}, e^{-2i\pi y\cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}}\rangle_y \prod_i \left|\frac{1}{2(b_i - a_i)}\right| \prod_j |2(b_i - a_i)| && (5.7)\\
&= \langle \hat{g}, e^{-2i\pi y\cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}}\rangle_y && (5.8)
\end{aligned}$$

the challenge being to evaluate $\hat{g}$ from g.

5.2.3 Symmetrization

The numerical method will employ a discrete Fourier transform (DFT) to approximate the Fourier transform of g.

Definition 5.1 *The discrete Fourier transform*

A periodic function h has a Fourier transform (a.k.a. Fourier series in this case) which is a train of δ distributions

$$\hat{h} = \sum_{k \in Z} H_k \tau_k \delta \tag{5.9}$$

H_k are called the Fourier coefficients of h.

Moreover, if h is bandwidth limited ($H_k = 0$ when $|k| > M$) then a DFT on a sample of $2M+1$ equi-distributed points of h in the interval $[-\frac{1}{2}, \frac{1}{2}]$ will give H_k for all $|k| \leqslant M$.

If h is not bandwidth limited, this same DFT will give H_k for all H_k $|k| \leqslant M$ with some aliasing.

Hence, we replace the initial problem with a problem with a payoff made periodic and, to avoid introducing a discontinuity at the limit of the interval, we symmetrize the payoff around the origin before making it periodic.

The resulting payoff $g_{\text{sp}} : \Re^n \to \Re^n$ is defined

$$g_{sp}(x) = \begin{cases} g(y) & \text{if } y \equiv x - 2k\pi \in [0, \frac{1}{2}] \\ g(-y) & \text{if } y \equiv x - 2k\pi \in [-\frac{1}{2}, 0] \end{cases} \tag{5.10}$$

and we approximate the initial integral with

$$\langle g, \partial_{\frac{1}{2(b-a)}} \tau_{-a} u \rangle \backsimeq \langle g_{\text{sp}}, \partial_{\frac{1}{2(b-a)}} \tau_{-a} u \rangle \tag{5.11}$$

The interval of interest will influence the quality of this approximation.

5.2.4 Fast Fourier transform

This section is expressed as if it was a one-dimensional problem. Extension to the n-dimensional problem is trivial.

The objective is to take the DFT of g_{sp} sampled at $2M+1$ points.

This is done equivalently via:

- a fast Fourier transform on the $2M+1$ samples of g_{sp} in $[-\frac{1}{2}, \frac{1}{2}]$,
- a fast cosine transform on $M+1$ samples of g in $[0, \frac{1}{2}]$.

and gives the $2M+1$ lowest symmetric coefficients of the Fourier transform of g_{sp}.

We approximate here

$$\hat{g}_{\mathrm{sp}} = \sum_{k \in Z} G_{\mathrm{sp}_k} \tau_k \delta \simeq \sum_{|k| \leqslant M} \tilde{G}_{\mathrm{sp}_k} \tau_k \delta \tag{5.12}$$

where $\tilde{G}_{\mathrm{sp}_k} \simeq G_{\mathrm{sp}_k}$ due to the aliasing. The influence of M is twofold: the number of Fourier coefficients in the sum (truncation effect) and their quality (aliasing effect).

Remark 5.1 *The symmetrisation is not mandatory but accelerates the decay of the Fourier coefficients giving lower truncation error and a lower aliasing effect and so better performance for a fixed M.*

Remark 5.2 *Even if $2M+1$ is not a power of 2, it is still possible to apply FFT algorithms if $2M+1$ can be factorised with small prime factors.*

Otherwise, a small modification of the methodology enables to work with samples of 2^N points.

Remark 5.3 *The DFT is defined as*

$$\mathrm{DFT}\ (g)_s = \frac{\sum_{r=1}^{n} g_r e^{2\pi i \frac{(r-1)(s-1)}{n}}}{\sqrt{n}} \tag{5.13}$$

and the DCT as

$$\mathrm{DCT}\ (g)_s = \sqrt{\frac{2}{n}} \left(\frac{g_1}{2} + \frac{(-1)^{-1+s} g_{n+1}}{2} + \sum_{j=2}^{n} g_j \cos\left(\pi \frac{(j-1)(s-1)}{n} \right) \right) \tag{5.14}$$

5.2.5 The final step

Finally, we obtain

$$\begin{aligned}
\langle f, u \rangle &= \langle \tau_a \partial_{\frac{b-a}{\pi}} g, u \rangle \\
&= \langle \hat{g}, e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \rangle_y \\
&\backsimeq \langle \hat{g}_{sp}, e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \rangle_y \\
&= \langle \sum\nolimits_{k \in Z^n} G_{\mathrm{sp}_k} \tau_k \delta, e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \rangle_y \\
&= \sum\nolimits_{k \in Z^n} G_{\mathrm{sp}_k} \langle \tau_k \delta, e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \rangle_y \\
&= \sum\nolimits_{k \in Z^n} G_{\mathrm{sp}_k} \left(e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \right)_{y=k} \\
&\backsimeq \sum\nolimits_{|k_i| \leqslant M} G_{\mathrm{sp}_k} \left(e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \right)_{y=k} \\
&\backsimeq \sum\nolimits_{|k_i| \leqslant M} \tilde{G}_{\mathrm{sp}_k} \left(e^{-2i\pi y \cdot a \frac{1}{2(b-a)}} \partial_{2(b-a)} \check{\hat{u}} \right)_{y=k} ,
\end{aligned} \tag{5.15}$$

where the $\tilde{G}_{\mathrm{sp}_k}$ are given (except for a factor) by the FFT or the FCT of g_{sp} .

Remark 5.4 *The sum is done over a square of indices k_i and implies $(2M-1)^n$ terms. This grows exponentially in the number of dimensions (problem a.k.a. the curse of dimensionality).*

However, the structure of the coefficients G_{sp_k} is often sparse and $\left(e^{-2i\pi y \cdot a} \partial_{2(b-a)} \check{\hat{u}}\right)_{y=k}$ can also be band-limited (negligible for some k's).

FFT's that compute only some lower frequency coefficient can be used to lower the complexity of the method in high dimensions by truncating the sum.

5.3 The Dynamic Programming (DP) algorithm

Pricing contracts such as **Take or pay** or **American** options requires to compute an optimal exercise strategy. This exercise strategy will provide the rules to optimally manage the contract based on the history of observable state variables. Longstaff and Schwartz (2001)[28] have proposed a Monte-Carlo algorithm to estimate this strategy. We will briefly discuss here an alternative approach often referred to as Dynamic Programming (DP).

Let S_t denote the state of the contract at time t, the exercise strategy will then provide the optimal S_{t+1} based on S_t and the history of the vector of underlying state variables X_j, $j \leq t$. Note, however, that in a Markov setting (which is the case for AJDs), all the information contained in the history of the vector of underlying state variables X_j, $j \leq t$, is summarized in X_t.

Dynamic Programming is a backward recursive algorithm that will provide this optimal exercise strategy by choosing, within all the $k = 1, \ldots, K$ possible states of the contract at time $t+1$, S^k_{t+1}, conditional on the fact that the contract is in state S_t, the optimal S^k_{t+1} that maximizes

$$\max_{k=1,\ldots,K} \left\{ P\left(S_t \to S^k_{t+1}\right) + E\left[RV(S^k_{t+1}, X_{t+1})|X_t\right] \right\} \equiv RV(S_t, X_t) \tag{5.16}$$

where

- $P\left(S_t \to S^k_{t+1}\right)$ denotes the payoff at time t resulting from the exercise strategy consisting in making the contract go from the state S_t to the S^k_{t+1};
- $RV(S_t, X_t)$ denotes the residual value of the contract in state S_t at time t with the vector of state variables equal to X_t.

When the contract ceases to exist (say at time $T+1$) its residual value is trivially zero. As a result the DP algorithm can be recursively applied starting at time T

and going backwards in time up to time t_0. The value of the contract at this time is then given by its residual value $RV(S_{t_0}, X_{t_0})$.

At every time t, the state-space of the underlying variables X_t is discretized in an equi-distant grid, such that the expectation $E\left[RV(S^k_{t+1}, X_{t+1})|X_t\right]$ can be computed via the equation (5.15).

Chapter 6

Risk Management

The risk-management of a portfolio essentially implies two activities. The computation of different risk measures and hedging.

Hedging against one risk factor consists of taking an opposite position in proportion to the impact of this risk factor. Hedging is often conducted through "Greeks" which measure this impact.

This section shows how the AJD framework can be used to compute Greeks of different contracts.

6.1 Hedging with "Greeks"

As said previously, "Greeks" measure the impact of a change in the risk-factors' value on the contract's value. This impact is measured by computing the first order partial derivatives of the contract's value with respect to the risk-factors' value. We will consider here the first and second order partial derivatives, often referred to as the "delta" and the "gamma" of the contract. Extension to any higher order derivative is straightforward but is rarely used.

We will also consider the effect of time on the contract's value by computing the first order partial derivative of the contract's value with respect to time. This measure is often referred to as the "theta" of the contract.

Since all the contracts considered in the Exotic Options Library are obtained via an affine combination of the "building blocks" developed in section 3.1, all that is needed to be done is to compute the partial derivatives of these building blocks. The partial derivatives of the contracts will then indeed become trivial.

6.1.1 Delta

We will use the subscript notation f_X to refer to the vector with the i^{th} element given by the partial derivative of the function f with respect to the i^{th} element of the vector X.

From the equations in section 3.1, we have that

$$\phi_X(d_0, d_1, a, b, X_t, t, T) = e^{\alpha(t)+\beta(t).X_t}\Big(\beta(t)A(t) + \beta(t)\left(B(t).X_t\right) + B(t)\Big) \tag{6.1}$$

$$\begin{aligned}\tilde{G}_X(y; d_0, d_1, a, b, X_t, t, T) \quad = \quad & \frac{\phi_X(d_0, d_1, a, 0, X_t, t, T)}{2} \\ & -\frac{1}{\pi}\int_0^\infty \frac{Im\left[\phi_X(d_0, d_1, a, v, X_t, t, T)e^{-ivy}\right]}{v}dv\end{aligned} \tag{6.2}$$

$$\begin{aligned}G_X(y_1, y_2; d_0, d_1, a, b_1, b_2, X_t, t, T) \quad = \quad & \frac{\phi_X(d_0, d_1, a, 0, X_t, t, T)}{4} \\ & -\frac{1}{2\pi}\sum_{j=1}^{2}\int_0^\infty \frac{Im\left[e^{-iwy_j}\phi_X(d_0, d_1, a, wb_j, X_t, t, T)\right]}{w}dw\end{aligned} \tag{6.3}$$

$$+\frac{1}{2\pi^2}\int_0^\infty\int_0^\infty\left\{\frac{Re\left[e^{-i(w_1y_1-w_2y_2)}\phi_X(d_0,d_1,a,w_1b_1-w_2b_2,X_t,t,T)\right]}{w_1w_2}\right.$$
$$\left.-\frac{Re\left[e^{-i(w_1y_1+w_2y_2)}\phi_X(d_0,d_1,a,w_1b_1+w_2b_2,X_t,t,T)\right]}{w_1w_2}\right\}dw_1dw_2$$

where extensions to the case where we have more than two constraints are trivial.

$$\begin{aligned}\Phi_X(d_0,d_1,d_2,a_1,a_2,T_1,T_1,T_2) \quad = \quad & \\ & e^{\alpha_1(t)+\beta_1(t).X_t}\Big(\beta_1(t)A_1(t)+\beta_1(t)\left(B_1(t).X_t\right)+B_1(t)\Big) \\ & +e^{\alpha_2(t)+\beta_2(t).X_t}\Big(\beta_2(t)A_2(t)+\beta_2(t)\left(B_2(t).X_t\right)+B_2(t)\Big)\end{aligned} \tag{6.4}$$

$$\Gamma_X(y_1,,y_2,d_0,d_{1,1},d1,2,a_1,a_2,b_{1,1},b_{2,1},b_{1,2},b_{2,2},X_t,t,T_1,T_2) \quad = \tag{6.5}$$

$$\frac{\Phi_X(d_0,d_{1,1},d_{1,2},a_1,a_2,X_t,t,T_1,T_2)}{4}$$
$$-\frac{1}{2\pi}\sum_{j=1}^{2}\int_0^\infty\frac{Im\left[e^{-iwy_j}\Phi_X(d_0,d_{1,1},d_{1,2},a_1+wb_{1,j},a_2+wb_{2,j},X_t,t,T_1,T_2)\right]}{w}dw$$
$$+\frac{1}{2\pi^2}\int_0^\infty\int_0^\infty\left\{\frac{Re\left[\begin{array}{r} e^{-i(w_1y_1-w_2y_2)}\Phi_X(d_0,d_{1,1},d_{1,2},a_1+w_1b_{1,1}-w_2b_{1,2}, \\ a_2+w_1b_{2,1}-w_2b_{2,2},X_t,t,T_1,T_2)\end{array}\right]}{w_1w_2}\right.$$
$$\left.-\frac{Re\left[\begin{array}{r} e^{-i(w_1y_1+w_2y_2)}\Phi_X(d_0,d_{1,1},d_{1,2},a_1+w_1b_{1,1}+w_2b_{1,2}, \\ a_2+w_1b_{2,1}+w_2b_{2,2},X_t,t,T_1,T_2)\end{array}\right]}{w_1w_2}\right\}dw_1dw_2$$

where again, extensions to the case where we have more than two constraints are trivial.

6.1.2 Theta

The first-order partial derivative of the different building blocks with respect to time is given by

$$\begin{aligned} \phi_t(d_0, d_1, a, b, X_t, t, T) &= -e^{\alpha(t)+\beta(t).X_t} \\ & \Big(K_0^T B(t) + \beta^T(t) H_0 B(t) + l_0 \nabla\theta(\beta(t)) B(t) \\ & + X_t.(K_1^T B(t) + \beta^T(t) H_1 B(t) + l_1 \nabla\theta(\beta(t)) B(t))\Big) \\ & + \Big(A(t) + B(t).X_t\Big) \\ & \Big(\rho_0 - K_0^T \beta(t) - \frac{1}{2}\beta^T(t) H_0 \beta(t) - l_0(\theta(\beta(t)) - 1) \\ & + X_t.(\rho_1 - K_1^T \beta(t) - \frac{1}{2}\beta^T(t) H_1 \beta(t) - l_1(\theta(\beta(t)) - 1))\Big) \end{aligned} \tag{6.6}$$

$$\begin{aligned} \tilde{G}_t(y; d_0, d_1, a, b, X_t, t, T) &= \frac{\phi_t(d_0, d_1, a, 0, X_t, t, T)}{2} \\ & - \frac{1}{\pi} \int_0^\infty \frac{Im\left[\phi_t(d_0, d_1, a, v, X_t, t, T) e^{-ivy}\right]}{v} dv \end{aligned} \tag{6.7}$$

$$\begin{aligned} G_t(y_1, y_2; d_0, d_1, a, b_1, b_2, X_t, t, T) &= \frac{\phi_t(d_0, d_1, a, 0, X_t, t, T)}{4} \\ - \frac{1}{2\pi} \sum_{j=1}^{2} \int_0^\infty & \frac{Im\left[e^{-iwy_j} \phi_t(d_0, d_1, a, wb_j, X_t, t, T)\right]}{w} dw \\ + \frac{1}{2\pi^2} \int_0^\infty \int_0^\infty & \Bigg\{ \frac{Re\left[e^{-i(w_1 y_1 - w_2 y_2)} \phi_t(d_0, d_1, a, w_1 b_1 - w_2 b_2, X_t, t, T)\right]}{w_1 w_2} \\ & - \frac{Re\left[e^{-i(w_1 y_1 + w_2 y_2)} \phi_t(d_0, d_1, a, w_1 b_1 + w_2 b_2, X_t, t, T)\right]}{w_1 w_2} \Bigg\} dw_1 dw_2 \end{aligned} \tag{6.8}$$

where extensions to the case where we have more than two constraints are trivial.

$$
\begin{aligned}
\Phi_t(d_0, d_1, d_2, a_1, a_2, T_1, T_1, T_2) \quad = \quad & -e^{\alpha_1(t)+\beta_1(t).X_t} \qquad (6.9) \\
& \Big(K_0^T B_1(t) + \beta_1^T(t) H_0 B_1(t) + l_0 \nabla\theta(\beta_1(t)) B_1(t) \\
& +X_t.(K_1^T B_1(t) + \beta_1^T(t) H_1 B_1(t) + l_1 \nabla\theta(\beta_1(t)) B_1(t))\Big) \\
& +\Big(A_1(t) + B_1(t).X_t\Big) \\
& \Big(\rho_0 - K_0^T \beta_1(t) - \frac{1}{2}\beta_1^T(t) H_0 \beta_1(t) - l_0(\theta(\beta_1(t)) - 1) \\
& +X_t.(\rho_1 - K_1^T \beta_1(t) - \frac{1}{2}\beta_1^T(t) H_1 \beta_1(t) - l_1(\theta(\beta_1(t)) - 1))\Big) \\
& -e^{\alpha_2(t)+\beta_2(t).X_t} \\
& \Big(K_0^T B_2(t) + \beta_2^T(t) H_0 B_2(t) + l_0 \nabla\theta(\beta_2(t)) B_2(t) \\
& +X_t.(K_1^T B_2(t) + \beta_2^T(t) H_1 B_2(t) + l_1 \nabla\theta(\beta_2(t)) B_2(t))\Big) \\
& +\Big(A_2(t) + B_2(t).X_t\Big) \\
& \Big(\rho_0 - K_0^T \beta_2(t) - \frac{1}{2}\beta_2^T(t) H_0 \beta_2(t) - l_0(\theta(\beta_2(t)) - 1) \\
& +X_t.(\rho_1 - K_1^T \beta_2(t) - \frac{1}{2}\beta_2^T(t) H_1 \beta_2(t) - l_1(\theta(\beta_2(t)) - 1))\Big)
\end{aligned}
$$

$$
\Gamma_t(y_1, , y_2, d_0, d_{1,1}, d_{1,2}, a_1, a_2, b_{1,1}, b_{2,1}, b_{1,2}, b_{2,2}, X_t, t, T_1, T_2) \quad = \qquad (6.10)
$$

$$
\begin{aligned}
& \frac{\Phi_t(d_0, d_{1,1}, d_{1,2}, a_1, a_2, X_t, t, T_1, T_2)}{4} \\
& -\frac{1}{2\pi}\sum_{j=1}^{2}\int_0^\infty \frac{Im\,[e^{-iwy_j}\Phi_t(d_0, d_{1,1}, d_{1,2}, a_1 + wb_{1,j}, a_2 + wb_{2,j}, X_t, t, T_1, T_2)]}{w} dw \\
& +\frac{1}{2\pi^2}\int_0^\infty\int_0^\infty \left\{ \frac{Re\left[\begin{array}{r} e^{-i(w_1y_1 - w_2y_2)}\Phi_t(d_0, d_{1,1}, d_{1,2}, a_1 + w_1b_{1,1} - w_2b_{1,2}, \\ a_2 + w_1b_{2,1} - w_2b_{2,2}, X_t, t, T_1, T_2) \end{array}\right]}{w_1w_2} \right.
\end{aligned}
$$

$$- \frac{Re\left[\begin{array}{r} e^{-i(w_1 y_1 + w_2 y_2)}\Phi_t(d_0, d_{1,1}, d_{1,2}, a_1 + w_1 b_{1,1} + w_2 b_{1,2}, \\ a_2 + w_1 b_{2,1} + w_2 b_{2,2}, X_t, t, T_1, T_2) \end{array}\right]}{w_1 w_2} \Bigg\} dw_1 dw_2$$

where again, extensions to the case where we have more than two constraints are trivial.

6.1.3 Gamma

We will use the subscript notation f_{XX} to refer to the matrix with the $(i,j)^{th}$ element given by $\frac{\partial f}{\partial X_i \partial X_j}$.

Differentiating the delta equations, we get

$$\begin{aligned} \phi_{XX}(d_0, d_1, a, b, X_t, t, T) &= \phi(d_0, d_1, a, b, X_t, t, T)\beta(t)\beta^T(t) \\ &\quad + e^{\alpha(t)+\beta(t).X_t}\Big(\beta(t)B^T(t) + B(t)\beta^T(t)\Big) \end{aligned} \tag{6.11}$$

$$\begin{aligned} \tilde{G}_{XX}(y; d_0, d_1, a, b, X_t, t, T) &= \frac{\phi_{XX}(d_0, d_1, a, 0, X_t, t, T)}{2} \\ &\quad - \frac{1}{\pi}\int_0^\infty \frac{Im\left[\phi_{XX}(d_0, d_1, a, v, X_t, t, T)e^{-ivy}\right]}{v} dv \end{aligned} \tag{6.12}$$

$$\begin{aligned} G_{XX}(y_1, y_2; d_0, d_1, a, b_1, b_2, X_t, t, T) &= \frac{\phi_{XX}(d_0, d_1, a, 0, X_t, t, T)}{4} \\ &\quad - \frac{1}{2\pi}\sum_{j=1}^{2}\int_0^\infty \frac{Im\left[e^{-iwy_j}\phi_{XX}(d_0, d_1, a, wb_j, X_t, t, T)\right]}{w} dw \\ &\quad + \frac{1}{2\pi^2}\int_0^\infty\int_0^\infty \Bigg\{ \frac{Re\left[e^{-i(w_1 y_1 - w_2 y_2)}\phi_{XX}(d_0, d_1, a, w_1 b_1 - w_2 b_2, X_t, t, T)\right]}{w_1 w_2} \\ &\quad - \frac{Re\left[e^{-i(w_1 y_1 + w_2 y_2)}\phi_{XX}(d_0, d_1, a, w_1 b_1 + w_2 b_2, X_t, t, T)\right]}{w_1 w_2} \Bigg\} dw_1 dw_2 \end{aligned} \tag{6.13}$$

where extensions to the case where we have more than two constraints are trivial.

$$\begin{aligned}
\Phi_{XX}(d_0,d_1,d_2,a_1,a_2,T_1,T_1,T_2) \;=\;& (A_1(t)+B_1(t).X_t)e^{\alpha_1(t)+\beta_1(t).X_t}\beta_1(t)\beta_1^T(t) \\
&+e^{\alpha_1(t)+\beta_1(t).X_t}\Big(\beta_1(t)B^T(t)+B_1(t)\beta_1(t)\Big) \\
&(A_2(t)+B_2(t).X_t)e^{\alpha_2(t)+\beta_2(t).X_t}\beta_2(t)\beta_2^T(t) \\
&+e^{\alpha_2(t)+\beta_2(t).X_t}\Big(\beta_2(t)B_2^T(t)+B_2(t)\beta_2(t)\Big)
\end{aligned} \tag{6.14}$$

$$\Gamma_{XX}(y_1,,y_2,d_0,d_{1,1},d_{1,2},a_1,a_2,b_{1,1},b_{2,1},b_{1,2},b_{2,2},X_t,t,T_1,T_2) \;=\; \tag{6.15}$$

$$\begin{aligned}
&\frac{\Phi_{XX}(d_0,d_{1,1},d_{1,2},a_1,a_2,X_t,t,T_1,T_2)}{4} \\
&-\frac{1}{2\pi}\sum_{j=1}^{2}\int_0^\infty \frac{Im\,[e^{-iwy_j}\Phi_{XX}(d_0,d_{1,1},d_{1,2},a_1+wb_{1,j},a_2+wb_{2,j},X_t,t,T_1,T_2)]}{w}dw \\
&+\frac{1}{2\pi^2}\int_0^\infty\int_0^\infty \left\{ \frac{Re\left[\begin{array}{r} e^{-i(w_1y_1-w_2y_2)}\Phi_{XX}(d_0,d_{1,1},d_{1,2},a_1+w_1b_{1,1}-w_2b_{1,2}, \\ a_2+w_1b_{2,1}-w_2b_{2,2},X_t,t,T_1,T_2)\end{array}\right]}{w_1w_2} \right. \\
&\qquad \left. -\frac{Re\left[\begin{array}{r} e^{-i(w_1y_1+w_2y_2)}\Phi_{XX}(d_0,d_{1,1},d_{1,2},a_1+w_1b_{1,1}+w_2b_{1,2}, \\ a_2+w_1b_{2,1}+w_2b_{2,2},X_t,t,T_1,T_2)\end{array}\right]}{w_1w_2} \right\} dw_1dw_2
\end{aligned}$$

where again, extensions to the case where we have more than two constraints are trivial.

Chapter 7

Estimation Issues

7.1 Introduction

As mentioned earlier, if one considers the spot price as the only primary traded commodity, then the estimation procedure is usually done in two steps. First one estimates the parameters of the spot price dynamics under the physical probability measure, then (if the market is incomplete) one calibrates the market prices of risk to match a forward curve level.

Another approach that is sometimes used is to only consider the risk-neutral probability measure and to calibrate the parameters of the model in order to match a series of contingent claims. The problem is that we then do not have any information on the prices dynamics under the physical probability measure.

In this chapter we will consider both spot and forward contracts as primary traded commodities. We thus propose an estimation procedure that jointly estimates, based on the historical series of spot and forward contracts, the parameters of the AJD under the "physical" probability measure P, together with the prices of risk η_1 and η_2 determining the risk-neutral probability measure Q. We note this set of parameters to be estimated ψ_0.

The Kalman Filter recursive algorithm is used to compute estimators of the unobserved state variables at every time step, and to compute the likelihood of observing a particular sequence of spot and forward prices. This is done by allowing for some measurement noise in the observation of the spot and forward prices, which can be interpreted as representing errors in the reporting of prices (perhaps due to asynchronous price quotes) or, alternatively, as errors in the model's fit to observed prices. This approach has been used by Pyndick (1997) [36], Schwartz (1997) [40] and Manoliu and Tompaidis (1999) [29]. In this chapter, however, we fully use the AJD framework machinery to write the state-space representation of the model, instead of applying the Euler approximation as in the cited literature. This also enables us to study the whole family of Affine Jump Diffusion processes, which is a broader scope than what was studied by these authors.

Instead of estimating the parameters of steady state structural state-space models by Maximum Likelihood via the error decomposition provided by the Kalman Filter, one could use their reduced form ARIMA (Auto Regressive Integrated Moving Average) models for which approximations to the Maximum Likelihood estimators, such as the Conditional Sum of Squares estimators, exist. In order to apply the method, a general procedure for computing the reduced form AR and MA parameters from the structural parameters is needed. Nerlove et al. (1979, pp. 125-31) [32] give such a procedure but show that it can be quite complex and time consuming.

If the exact Maximum Likelihood is required, however, then there is no practical advantage to estimation via the reduced form, since it is as time consuming to compute the exact ML estimators for an ARIMA model as it is for the corresponding structural model, as was reported by Harvey (1994, p. 187) [21]. In fact, translating ARMA (Auto Regressive Moving Average) or ARMAX (Auto Regressive Moving Average with eXogenous variables) models into state-space models for a Maximum

Likelihood estimation of their parameters is proposed by Shumway and Stoffer (2000, pp. 312-338) [44].

Furthermore, though the Kalman Filter is, as such, designed for a Gaussian framework, we propose in this chapter different ways to extend the use of the Kalman Filters to a non-Gaussian framework. These extensions are done using the state-space representation of the structural models and are not transposable to their reduced form. As a result, we will keep the state-space representation throughout this chapter.

In the case of stochastic volatility models, for example, the Gaussian assumption made in the Kalman Filter recursive algorithm will lead to poor estimators when there is no correlation between the state variables and their volatility. We propose here an efficient algorithm to relax the assumption of normality for models with stochastic volatility that leads to better estimates of the volatility level at each time step.

When the assumption of Gaussian error vectors is relaxed, the computation of the Likelihood function is numerically quite intensive. We then apply the numerical techniques developed in Singleton (2001) [45], based on the expected conditional characteristic function, to speed up the computations without losing too much on the efficiency of the parameter estimators.

In his paper, Singleton (2001) [45] approaches the problem of the estimation of the unobserved state variables by choosing at each time step the vector of unobserved state variables that either perfectly price a subset of observed contingent claim prices or the vector of unobserved state variables that minimizes the mean square pricing error of these observed contingent claims. Using the Kalman Filter is thus an improvement compared to the existing literature in the sense that it applies the numerical techniques based on the expected conditional characteristic function with unobserved state variables estimated via a Kalman Filter.

Though the model is continuous, we decided to use the discrete Kalman Filter instead of its continuous version because the data (spot and futures prices) are observed discretely.

The first section of this chapter focusses on AJD's resulting in Gaussian distributions. The next section indicates some possible research paths to depart from the Gaussian framework and to make the Maximum Likelihood optimization faster. Among these paths, only the Bernoulli approximation has been implemented on the European Energy indeX (EEX) for electricity, on the Brent for crude oil, and on the NBP for natural gas. The other paths developed in this section have not been implemented.

7.2 Estimation procedure : The Kalman Filter

7.2.1 The State Space Representation

In an Affine Jump Diffusion framework, the price of a futures contract, $f(t, T_1, T_2)$ -where t is the pricing time, T_1 is the starting date of the delivery period, and T_2 is the ending date of the delivery period-, or its logarithm is an affine function of the vector of the state variables X_t. In the level model, we indeed have

$$\begin{aligned} f(t, T_1, T_2) &= E^Q\left[\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} Spot(s)ds\Big|\mathcal{F}_t\right] && (7.1)\\ &= E^Q\left[\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} \sum_i \gamma_i X_i(s)ds\Big|\mathcal{F}_t\right] && (7.2)\\ &= \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} A(t,s)ds + \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} B(t,s)ds.X_t && (7.3) \end{aligned}$$

where $A(t,T)$ and $B(t,T)$ are respectively a scalar and a vector, solution of a system of ordinary differential equations obtained by simplifying the transform ϕ (7.2)

$$\left\{ \begin{array}{l} -\dot{B}_t = \widetilde{K}_1^T B_t + \widetilde{l}_1 \nabla\widetilde{\theta}(\overline{0}) B_t \\ -\dot{A}_t = \widetilde{K}_0^T B_t + \widetilde{l}_0 \nabla\widetilde{\theta}(\overline{0}) B_t \end{array} \right. \tag{7.4}$$

with the boundary conditions $B(T) = \gamma$ and $A(T) = 0$, and where $\nabla\widetilde{\theta}(c)$ is gradient of $\widetilde{\theta}(c)$ with respect to $c \in \mathbb{C}^n$.

In the return model, we have

$$\begin{aligned} \log(f(t,T_1,T_2)) &= \log\left(E^Q\left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} Spot(s)ds \Big| \mathcal{F}_t \right] \right) && (7.5) \\ &= \log\left(E^Q\left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} exp\left(\sum_i \gamma_i X_i(s) \right) ds \Big| \mathcal{F}_t \right] \right) && (7.6) \\ &= \log\left(\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} e^{\alpha(t,s) + \beta(t,s).X_t} ds \right) && (7.7) \end{aligned}$$

where $\alpha(t,T)$ and $\beta(t,T)$ are respectively a scalar and a vector of same size as X, solution of a system of ordinary differential equations also obtained by simplifying the transform ϕ (7.2)

$$\left\{ \begin{array}{l} \dot{\beta}_t = -\widetilde{K}_1^T \beta_t - \frac{1}{2}\beta_t^T \widetilde{H}_1 \beta_t - \widetilde{l}_1(\widetilde{\theta}(\beta_t) - 1) \\ \dot{\alpha}_t = -\widetilde{K}_0^T \beta_t - \frac{1}{2}\beta_t^T \widetilde{H}_0 \beta_t - \widetilde{l}_0(\widetilde{\theta}(\beta_t) - 1) \end{array} \right. \tag{7.8}$$

with the boundary conditions $\alpha(T,T) = 0$ and $\beta(T,T) = \gamma$.

If we further make the assumption that we can replace the arithmetic average with its geometric average, we get

$$\log(f(t,T_1,T_2)) \approx \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \alpha(t,s)ds + \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \beta(t,s)ds.X_t \tag{7.9}$$

Note here that we replace the arithmetic average of expectations by its geometric average, which is much more accurate than replacing the expectation of an arithmetic average by the expectation of its geometric average. This is because expectations with different horizons typically do not vary so much, whereas realized prices can vary quite dramatically.

As a result, at each time step we can write the system of so called "measuring equations" :

$$Y_t = d_t + Z_t X_t + \epsilon_t \tag{7.10}$$

where

- Y_t is a vector with the spot price or its logarithm and other futures contracts prices or their logarithm, according to whether the level or return model is used;
- d_t is a vector;
- Z_t is a matrix;
- ϵ_t is a vector of measurement errors normally distributed with mean zero and covariance matrix V_t.

The measurement errors ϵ_t are added to the equations (7.3) and (7.9) to cope with observations that are not exactly contemporaneous or with model shortcomings that make it impossible to reproduce all of the observed prices.

Next to these measurement equations, the Affine Jump Diffusion process can be discretized to give the so called transition equations:

$$X_{t+1} = c_t + T_t X_t + \zeta_t \tag{7.11}$$

where ζ_t is a vector of errors with mean zero and variance covariance matrix $R(X_t, t)$. We will further approximate this vector to be normally distributed.

Noting $A_j(t)$ and $B_j(t)$ the j^{th} line of c_t and T_t, respectively, $A_j(t)$ and $B_j(t)$ can be computed as the solution to the same system of ordinary differential equations (7.4), but with the parameters of the AJD process X_t under the physical probability measure P:

$$\begin{cases} -\dot{B}_{j,t} = K_1^T B_{j,t} + l_1 \nabla\theta(\overline{0}) B_{j,t} \\ -\dot{A}_{j,t} = K_0^T B_{j,t} + l_0 \nabla\theta(\overline{0}) B_{j,t} \end{cases} \tag{7.12}$$

with the boundary conditions $A_j(t+1) = 0$ and $B_j(t+1) = \mathbf{1}_j$, where $\mathbf{1}_j$ is a vector of zeros with only the j^{th} element equal to 1.

The covariance matrix $R(X_t, t)$ is equal to

$$R(X_t, t) = R^0(t) + R^1(t) X_t \tag{7.13}$$

where R_t^0 is a matrix and R_t^1 is a three-dimensional tensor.

By definition, the covariance matrix is indeed given by

$$\begin{aligned} R(X_t,t) &= E[X_{t+1}X_{t+1}^T] - E[X_{t+1}]E[X_{t+1}^T] && (7.14) \\ &= \nabla\nabla^T E[e^{u.X_{t+1}}]|_{u=0} - \nabla E[e^{u.X_{t+1}}]|_{u=0} \nabla^T E[e^{u.X_{t+1}}]|_{u=0} && (7.15) \\ &= \nabla\nabla^T \left(e^{\alpha(u,t,t+1)+\beta(u,t,t+1).X_{t+1}}\right)|_{u=0} \\ &\quad -\nabla \left(e^{\alpha(u,t,t+1)+\beta(u,t,t+1).X_{t+1}}\right)|_{u=0} \\ &\qquad \nabla^T \left(e^{\alpha(u,t,t+1)+\beta(u,t,t+1).X_{t+1}}\right)|_{u=0} && (7.16) \\ &= \nabla\nabla^T \alpha(u,t,t+1)|_{u=0} + (\nabla\nabla^T \beta(u,t,t+1)|_{u=0})^T X_t && (7.17) \\ &\equiv R^0(t) + R^1(t) X_t && (7.18) \end{aligned}$$

Considering that $\alpha(u, t, T)$ and $\beta(u, t, T)$ are solutions of the system of ODE's developed in the transform ϕ (7.2), we finally get

$$\begin{aligned} \dot{R}_{i,j}^1(\tau) &= -K_1^T R_{i,j}^1(\tau) - B_j^T(\tau) H_1 B_i(\tau) - l_1 \nabla\theta(\overline{0}) R_{i,j}^1 \\ &\quad -l_1 B_i^T(\tau) \nabla\nabla^T \theta(\overline{0}) B_j(\tau) && (7.19) \\ \dot{R}_{i,j}^0(\tau) &= -K_0^T R_{i,j}^1(\tau) - B_j^T(\tau) H_0 B_i(\tau) - l_0 \nabla\theta(\overline{0}) R_{i,j}^1 \\ &\quad -l_0 B_i^T(\tau) \nabla\nabla^T \theta(\overline{0}) B_j(\tau) && (7.20) \end{aligned}$$

with boundary conditions $R^1_{i,j}(t+1) = \overline{0}$ and $R^0_{i,j}(t+1) = 0$.

Note that if $H_1 = 0$ and $l_1 = 0$, then $R^1(t) = 0$ and $R(t)$ is no more a function of X_t. Strictly speaking, the Kalman Filter Recursion Algorithm does not allow for a state space representation with a covariance matrix $R(X_t, t)$ function of X_t and thus could not allow AJD's with non zero R^1 and l_1. In what follows, we will however keep the formalism where $R(X_t, t)$ is an affine function of X_t in order to introduce an approximation to the Kalman Filter when R^1 and l_1 are non-zero.

The measurement equations (7.10) and the transition equations (7.11) form the state space representation of the model.

7.2.2 The Kalman Filter Recursion Algorithm

Once a model has been put in a state space form, the Kalman Filter can be used. The Kalman Filter is a recursive procedure for computing the optimal estimator of the state vector X at time t, based on the information available at time t. This information consists of the observations up to and including Y_t.

When the error vectors ϵ_t and ζ_t are assumed to be normally distributed, the estimator of the state vector is optimal in the sense that it minimizes the mean square error. Furthermore, in this Gaussian setting, the estimator yielded by the Kalman filter can be interpreted as the conditional expectation of X_t based on the information available at time t.

We briefly expose the filter's algorithm here. Further information on Kalman Filters can be found in Harvey (1994) [21].

Let $\widehat{X}_{t-1}$ denote the optimal estimator of X_{t-1} based on the observations up to and including Y_{t-1}. Let P_{t-1} denote the covariance matrix of the estimation error, i.e.

$$P_{t-1} = E\Big[(\widehat{X}_{t-1} - X_{t-1})(\widehat{X}_{t-1} - X_{t-1})'\Big] \tag{7.21}$$

Given $\widehat{X}_{t-1}$ and P_{t-1}, the optimal estimator of X_t is given by

$$\widehat{X}_{t|t-1} = T_{t-1}\widehat{X}_{t-1} + c_{t-1} \tag{7.22}$$

while the covariance matrix of the estimation error is

$$P_{t|t-1} = T_{t-1}P_{t-1}T'_{t-1} + R(\widehat{X}_{t-1}, t-1) \tag{7.23}$$

These two equations are known as the *prediction equations*.

As noted earlier, this is only an approximation to the estimation of the covariance matrix of the estimation error when $R(X_t, t)$ is a function of X_t, that is when R_1 and l_1 are non-zero. In what follows, we will bear that approximation for achieving a higher level of generality in the class of AJD processes that can be treated.

Once the new observation, Y_t, becomes available, the estimator of X_t, $\widehat{X}_{t|t-1}$, can be updated. The *updating equations* are

$$\widehat{X}_t = \widehat{X}_{t|t-1} + P_{t|t-1}Z'_tF_t^{-1}\Big(y_t - Z_t\widehat{X}_{t|t-1} - d_t\Big) \tag{7.24}$$

and

$$P_t = P_{t|t-1} - P_{t|t-1}Z'_tF_t^{-1}Z_tP_{t|t-1} \tag{7.25}$$

where

$$F_t = Z_tP_{t|t-1}Z'_t + V_t \tag{7.26}$$

The starting values of the Kalman Filter may be specified in terms of $\widehat{X}_0$ and P_0 or $\widehat{X}_{1|0}$ and $P_{1|0}$. Given these initial conditions, the Kalman filter delivers the optimal estimator of the state vector as each new observation becomes available.

In the next sections, it will be useful to be able to differentiate the estimator $\widehat{X}_t^{\psi}$ with respect to ψ, such that the optimization algorithm in the Maximum Likelihood procedure will be able to use first and second order methods based on numerically computed gradients instead of gradients computed by finite differences.

Proposition 7.1 *The gradient of $\widehat{X}_t^{\psi}$ (where the dependence on ψ has been made explicit) with respect to ψ can be computed recursively. The computations only involve partial derivatives of the matrices and vectors $d_s, Z_s, c_s,$ and T_s with respect to ψ, where $s = 1, \ldots, t$.*

Proof : From equation (7.24), we have that

$$\frac{\partial \widehat{X}_t}{\partial \psi} = \frac{\partial \widehat{X}_{t|t-1}}{\partial \psi} + \frac{\partial P_{t|t-1} Z_t' F_t^{-1} \Big(y_t - Z_t \widehat{X}_{t|t-1} - d_t \Big)}{\partial \psi} \tag{7.27}$$

From equation (7.22), the first term of the right hand side of this equation can further be developed into

$$\frac{\partial \widehat{X}_{t|t-1}}{\partial \psi} = \frac{\partial T_{t-1}}{\partial \psi} \widehat{X}_{t-1} + T_{t-1} \frac{\partial \widehat{X}_{t-1}}{\partial \psi} + \frac{\partial c_{t-1}}{\partial \psi} \tag{7.28}$$

where $\partial \widehat{X}_{t-1} / \partial \psi$ is obtained by taking the result of the computation of (7.27) at time $t-1$, with an initial value of $\partial \widehat{X}_0 / \partial \psi = 0$. The other derivatives are again easily computed by eventually solving additional systems of ordinary equations.

The second term of the right hand side of equation (7.27) can be rewritten as

$$\begin{aligned} \frac{\partial P_{t|t-1} Z_t' F_t^{-1} \Big(Y_t - Z_t \widehat{X}_{t|t-1} - d_t \Big)}{\partial \psi} \quad &= \\ & \frac{\partial P_{t|t-1}}{\partial \psi} Z_t' F_t^{-1} \Big(Y_t - Z_t \widehat{X}_{t|t-1} - d_t \Big) \\ & \quad + P_{t|t-1} \frac{\partial Z_t'}{\partial \psi} F_t^{-1} \Big(Y_t - Z_t \widehat{X}_{t|t-1} - d_t \Big) \\ & \quad - P_{t|t-1} Z_t' F_t^{-1} \frac{\partial F_t}{\partial \psi} F_t^{-1} \Big(Y_t - Z_t \widehat{X}_{t|t-1} - d_t \Big) \\ & \quad + P_{t|t-1} Z_t' F_t^{-1} \Big(- \frac{\partial Z_t}{\partial \psi} \widehat{X}_{t|t-1} - Z_t \frac{\partial \widehat{X}_{t|t-1}}{\partial \psi} - \frac{\partial d_t}{\partial \psi} \Big) \end{aligned} \tag{7.29}$$

where, from equation (7.26), we have that

$$\frac{\partial F_t}{\partial \psi} = \frac{\partial V_t}{\partial \psi} + \frac{\partial Z_t}{\partial \psi} P_{t|t-1} Z_t' + Z_t \frac{\partial P_{t|t-1}}{\partial \psi} Z_t' + Z_t P_{t|t-1} \frac{\partial Z_t'}{\partial \psi} \tag{7.30}$$

and where, from equation (7.23), we have that

$$\frac{\partial P_{t|t-1}}{\partial \psi} = \frac{\partial R_{t-1}}{\partial \psi} + \frac{\partial T_{t-1}}{\partial \psi} P_{t-1} T_{t-1}' + T_{t-1} \frac{\partial P_{t-1}}{\partial \psi} T_{t-1}' + T_{t-1} P_{t-1} \frac{\partial T_{t-1}'}{\partial \psi} \tag{7.31}$$

with $\frac{\partial P_0}{\partial \psi} = 0$, and

$$\begin{aligned} \frac{\partial P_t}{\partial \psi} = \frac{\partial P_{t|t-1}}{\partial \psi} &- \frac{\partial P_{t|t-1}}{\partial \psi} Z_t' F_t^{-1} Z_t P_{t|t-1} \\ &- P_{t|t-1} \frac{\partial Z_t'}{\partial \psi} F_t^{-1} Z_t P_{t|t-1} \\ &+ P_{t|t-1} Z_t' F_t^{-1} \frac{\partial F_t}{\partial \psi} F_t^{-1} Z_t P_{t|t-1} \\ &- P_{t|t-1} Z_t' F_t^{-1} \frac{\partial Z_t}{\partial \psi} P_{t|t-1} \\ &- P_{t|t-1} Z_t' F_t^{-1} Z_t \frac{\partial P_{t|t-1}}{\partial \psi} \end{aligned} \tag{7.32}$$

and all other partial derivatives being again easily computed.□

7.2.3 The Maximum Likelihood estimation

The joint density function of all the observations, $Y_1, \ldots, Y_T$, is given by the function

$$L(\mathbf{Y}; \psi) = \prod_{t=1}^{T} p(Y_t | Y_{t-1}) \tag{7.33}$$

where $p(Y_t|Y_{t-1})$ denotes the distribution of Y_t conditional on the information set at time $t-1$, that is all of the observations Y_i , $i = 1 \ldots t-1$.

If the measurement equation is written as

$$Y_t = Z_t \widehat{X}_{t|t-1} + Z_t (X_t - \widehat{X}_{t|t-1}) + d_t + \epsilon_t \tag{7.34}$$

it can be seen immediately that the conditional distribution of Y_t is normal with mean

$$E[Y_t|Y_{t-1}] = \widetilde{Y}_{t|t-1} = Z_t\widehat{X}_{t|t-1} + d_t \tag{7.35}$$

and a covariance matrix, F_t. For a Gaussian model, therefore, the likelihood function can be written immediately as

$$\log L(\mathbf{Y};\psi) = -\frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|F_t| - \frac{1}{2}\sum_{t=1}^{T}\nu_t^T F_t^{-1}\nu_t \tag{7.36}$$

where

$$\nu_t = Y_t - \widetilde{Y}_{t|t-1} \tag{7.37}$$

and N is the size of each vector of observations Y_t.

Finally, we choose the vector of parameters ψ that will maximize equation (7.36), i.e., the logarithm of the joint density function of all observations.

The gradient of the likelihood function can be written

$$(\log L(\mathbf{Y};\psi))' = -\frac{1}{2}\sum_{t=1}^{T}\frac{(|F_t|)'}{|F_t|} - \frac{1}{2}\sum_{t=1}^{T}(\nu_t^T F_t^{-1}\nu_t)' \tag{7.38}$$

This expression can be computed by noticing that

$$|F_t|' = |F_t|Tr(F_t^{-1}F_t') \tag{7.39}$$

with F_t' given by (7.30) and that

$$(\nu_t^T F_t^{-1}\nu_t)' = (\nu_t^T)'F_t^{-1}\nu_t + \nu_t^T(F_t^{-1})'\nu_t + \nu_t^T F_t^{-1}(\nu_t)' \tag{7.40}$$

with

$$(F_t^{-1})' = -F_t^{-1}F_t'F_t^{-1} \tag{7.41}$$

and with

$$\begin{aligned}(\nu_t)' &= -(\widetilde{Y}_{t|t-1})' && (7.42)\\ &= -\left(Z_t'\widehat{X}_{t|t-1} + Z_t\widehat{X}_{t|t-1}' + d_t'\right) && (7.43)\end{aligned}$$

and where $\widehat{X}_{t|t-1}'$ is given by (7.28).

7.3 Accounting for non-Gaussian errors

The assumption of normality is made twice in the estimation process. First it is made in the Kalman Filter Recursive Algorithm for computing the estimators $\widehat{X}_t$, then it is used to compute the loglikelihood function to be maximized. For many different stochastic processes of the AJD family, however, the assumption of a vector ζ normally distributed in equation (7.11) is quite dramatically invalidated. The next sections show how to relax this approximation in both cases.

7.3.1 The Kalman Filter Recursion Algorithm

Even if the vector ζ is not normally distributed, it is however shown in Harvey (1994) that the Kalman Filter recursive algorithm can still be applied. The estimators $\widehat{X}_t$ yielded will no more be conditional expectations of X_t based on the information available at time t, but, in the case of a deterministic conditional covariance, they will still remain optimal in the sense that, if attention is restricted to estimators which are linear combinations of the observations, then it is the one that minimizes the Mean Square Error. Thus, in this case, $\widehat{X}_t$ is the minimum mean square linear estimator of X_t based on observations up to and including time t.

For some models, however, the conditional covariance matrix is not deterministic, and the estimators yielded will not provide good results.

This will particularly be the case with stochastic volatility models where the observations have a very small vega, i.e., are not sensible to the volatility levels (for example, futures prices are not sensible to the level of volatility), and where there is no or very little correlation between the state variables and the volatility. In this case, the Kalman Filter will indeed not provide any correction to the expected level of the variance at time t once the observations of time t are made available.

We present here an algorithm to overcome this problem and give better estimates of the variance levels at each time step t, $\widehat{V}_t$. Our objective is thus to compute

$$\widehat{V}_t = E[V_t|\mathcal{F}_t] = \int_{\mathbb{R}} v_t f(V_t = v_t|\mathcal{F}_t) dv_t \tag{7.44}$$

where $f(V_t = v_t|\mathcal{F}_t)$ is the probability density function of V_t conditional on all the observations up to and including time t. This can be expressed as

$$E[V_t|\mathcal{F}_t] = \int_{\mathbb{R}} v_t \Big(\int_{\mathbb{R}^M} f(V_t = v_t|X_t = x_t, \mathcal{F}_t) f(X_t = x_t|\mathcal{F}_t) dx_t \Big) dv_t \tag{7.45}$$

where it is assumed that the vector X is of size M and includes all the other state variables but V.

Again, we will approximate this expression with

$$E[V_t|\mathcal{F}_t] \approx \int_{\mathbb{R}} v_t f(V_t = v_t|X_t = \widehat{X}_t, \mathcal{F}_t) dv_t \tag{7.46}$$

which can further be expressed as

$$\begin{aligned}
E[V_t|\mathcal{F}_t] &\approx \int_{\mathbb{R}} v_t f(V_t = v_t|X_t = \widehat{X}_t, \mathcal{F}_{t-1}) dv_t && (7.47)\\
&= \int_{\mathbb{R}} v_t \frac{f(V_t = v_t, X_t = \widehat{X}_t|\mathcal{F}_{t-1})}{f(X_t = \widehat{X}_t|\mathcal{F}_{t-1})} dv_t && (7.48)\\
&= \int_{\mathbb{R}} v_t \frac{\int_{\mathbb{R}^{M+1}} \begin{pmatrix} f(V_t = v_t, X_t = \widehat{X}_t|X_{t-1} = x_{t-1}, V_{t-1} = v_{t-1}) \\ f(X_{t-1} = x_{t-1}, V_{t-1} = v_{t-1}|\mathcal{F}_{t-1}) \end{pmatrix} d\{x_{t-1}, v_{t-1}\}}{\int_{\mathbb{R}^{M+1}} \begin{pmatrix} f(X_t = \widehat{X}_t|X_{t-1} = x_{t-1}, V_{t-1} = v_{t-1}) \\ f(X_{t-1} = x_{t-1}, V_{t-1} = v_{t-1}|\mathcal{F}_{t-1}) \end{pmatrix} d\{x_{t-1}, v_{t-1}\}} dv_t && (7.49)\\
&\approx \int_{\mathbb{R}} v_t \frac{f(V_t = v_t, X_t = \widehat{X}_t|X_{t-1} = \widehat{X}_{t-1}, V_{t-1} = \widehat{V}_{t-1})}{f(X_t = \widehat{X}_t|X_{t-1}, V_{t-1} = \widehat{V}_{t-1})} dv_t && (7.50)
\end{aligned}$$

where, again, an approximation has been done to arrive at the last equation.

The denominator of the last equation can be taken out of the integral. Its computation will only require an integration on $\mathbb{R}^M$ of the characteristic function of X.

Assuming $M = 1$, or that we only consider the realization of one element of X_t (one whose variance is a function of V) for updating the variance level at time t, we now show an efficient way to compute the integral we are left with.

Let us start with

$$\begin{aligned}
&\frac{\partial E[V_t \mathbf{1}_{X_t<y}|X_{t-1} = x_{t-1}, V_{t-1} = v_{t-1}]}{\partial y} \\
&= \int_{\mathbb{R}^{1+1}} v_t \frac{\partial \mathbf{1}_{X_t<y}}{\partial y} f(V_t = v_t, X_t = x_t|X_{t-1} = x_t - 1, V_{t-1} = v_{t-1}) d\{x_t, v_t\} \quad (7.51)\\
&= \int_{\mathbb{R}^{1+1}} v_t \delta_{x_t=y} f(V_t = v_t, X_t = x_t|X_{t-1} = x_t - 1, V_{t-1} = v_{t-1}) d\{x_t, v_t\} \quad (7.52)\\
&= \int_{\mathbb{R}} v_t f(V_t = v_t, X_t = y|X_{t-1} = x_t - 1, V_{t-1} = v_{t-1}) d\{x_t, v_t\} \quad (7.53)
\end{aligned}$$

which is indeed equivalent to the numerator of (7.50) if $y = \widehat{X}_t$.

From the building block of the **Proposition 4.7**, it is easy to show that equation (7.51) is also equal to

$$\begin{aligned}
&\frac{\partial E[V_t \mathbf{1}_{X_t<y}|X_{t-1} = x_t - 1, V_{t-1} = v_{t-1}]}{\partial y} \\
&\quad = \frac{\phi(\{0,1\}, \{0,0\}, \{x_{t-1}, v_{t-1}\}, t-1, t)}{2} \\
&\quad\quad + \frac{1}{\pi} \int_0^\infty Re[\phi(\{0,1\}, \{iu, 0\}, \{x_{t-1}, v_{t-1}\}, t-1, t) e^{-iuy}] du \quad (7.54)
\end{aligned}$$

In the case where $M = 1$, we have thus shown that equation (7.47) could be computed by using only integrals on $\mathbb{R}$.

If $M > 1$, the same approach can be used by solving

$$\frac{\partial^M E[V_t \prod_{j=1}^M \mathbf{1}_{X_{j,t}<y_j}|X_{t-1} = x_t - 1, V_{t-1} = v_{t-1}]}{\prod_{j=1}^M \partial y_j} \quad (7.55)$$

and differentiating the building block of **Proposition 4.11**. This will however require integrals on $\mathbb{R}^M$.

7.3.2 Computing the Log-Likelihood function

The Bernoulli approximation for Jump processes

Obviously, when ζ is not normally distributed, the computation of the log-likelihood function $\log L(\mathbf{Y};\psi)$ does not hold anymore. Considering jumps that are normally distributed, the computation can, however, be improved by replacing the transition equations (7.11) by

$$X_t = c^*_{t-1} + T^*_{t-1} X_{t-1} + \zeta^*(t-1) + J_t \tag{7.56}$$

where

- the Bernoulli approximation has been done, i.e. at each time step there can only be one or no jump;
- the probability of having a jump is given by $(1 - e^{-(\lambda_0 + \lambda_1 X_{t-1})})$;
- when a jump occurs, J_t is normally distributed with mean μ_J and covariance matrix Σ_J, otherwise it is equal to zero;
- the vector c^*_{t-1}, the matrix T^*_{t-1}, and the covariance matrix of the vector $\zeta^*_1(t)$, i.e. R^*_t, are computed as if there was no jump component to the stochastic process of X_t, as indicated earlier.

At each time step, we then compute an extra vector and an extra covariance matrix corresponding to the case where there has been no jump:

$$\widehat{X}^0_{t|t-1} = T^*_{t-1}\widehat{X}_{t-1} + c^*_t \tag{7.57}$$

$$P^0_{t|t-1} = T^*_{t-1} P_{t-1} T^{*'}_{t-1} + R^*_t \tag{7.58}$$

and an extra vector and an extra covariance matrix corresponding to the case where there has been a jump:

$$\widehat{X}^1_{t|t-1} = T^*_{t-1}\widehat{X}_{t-1} + c^*_t + \mu_j \tag{7.59}$$

$$P^1_{t|t-1} = T^*_{t-1} P_{t-1} T^{*'}_{t-1} + R^*_t + \Sigma_J \tag{7.60}$$

The log-likelihood equation can now be written as

$$\log L(\mathbf{Y};\psi) = \sum_{t=1}^{T} \log\left(e^{-(\lambda_0+\lambda_1 X_{t-1})} p_0(Y_t|Y_{t-1}) + (1-e^{-(\lambda_0+\lambda_1 X_{t-1})}) p_1(Y_t|Y_{t-1})\right) \tag{7.61}$$

where

$$p_0(Y_t|Y_{t-1}) = (2\pi)^{-N/2}(|F^0_t|)^{-1/2} e^{-\frac{1}{2}\nu^{0'}_t (F^0_t)^{-1}\nu^0_t} \tag{7.62}$$

$$p_1(Y_t|Y_{t-1}) = (2\pi)^{-N/2}(|F^1_t|)^{-1/2} e^{-\frac{1}{2}\nu^{1'}_t (F^1_t)^{-1}\nu^1_t} \tag{7.63}$$

and where

$$\nu^0_t = Y_t - Z_t \widehat{X}^0_{t|t-1} - d_t \tag{7.64}$$

$$\nu^1_t = Y_t - Z_t \widehat{X}^1_{t|t-1} - d_t \tag{7.65}$$

$$F^0_t = Z_t P^0_{t|t-1} Z'_t + H_t \tag{7.66}$$

$$F^1_t = Z_t P^1_{t|t-1} Z'_t + H_t \tag{7.67}$$

Again, we choose the vector of parameters ψ that will maximize equation (7.61), i.e., the logarithm of the joint density function of all observations.

A general approach based on the conditional characteristic function

The characteristic function of Y_t conditional on the information available at time $t-1$, $\phi_{Y_t}(c)$ is by definition

$$\phi_{Y_t}(c) = E^P[e^{i(c.Y_{t+1})}|\mathcal{F}_t] \tag{7.68}$$

Looking at the measurement equation (7.10)

$$Y_t = d_t + Z_t X_t + \epsilon_t$$

where the vector of measurement errors ϵ_t is independent of X_t, one can write

$$\phi_{Y_t}(c) = \int_{\mathbb{R}^M} \phi_{d_t+Z_t x_t}(c) f(x_t|\mathcal{F}_t) dx_t \;\; \phi_{\epsilon_t}(c) \tag{7.69}$$

where M denotes the size of X and $f(x_t|\mathcal{F}_t)$ denotes the probability density function of X_t conditional on all the observations up to and including t; and where, by definition

$$\begin{aligned}
\phi_{d_t+Z_t X_t}(c) &= E^P[e^{i(c.d_{t+1}+c.Z_{t+1}X_{t+1}))}|X_t = x_t] && (7.70)\\
&= e^{i(c.d_{t+1})} E^P[e^{i(c.Z_{t+1}X_{t+1}))}|X_t = x_t] && (7.71)\\
&= e^{i(c.d_{t+1})} E^P[e^{i(c.Z_{t+1}X_{t+1}))}|X_t = x_t] && (7.72)\\
&= e^{i(c.d_{t+1})} \phi_{X_t}(Z'_{t+1}c) && (7.73)
\end{aligned}$$

In order to simplify the computations, equation (7.69) could be approximated by

$$\phi_{Y_t}(c) \approx \phi_{d_t+Z_t \widehat{X}_t}(c) \phi_{\epsilon_t}(c) \tag{7.74}$$

with

$$\begin{aligned}
\phi_{d_t+Z_t \widehat{X}_t}(c) &= E^P[e^{i(c.d_{t+1}+c.Z_{t+1}X_{t+1}))}|X_t = \widehat{X}_t] && (7.75)\\
&= e^{i(c.d_{t+1})} E^P[e^{i(c.Z_{t+1}X_{t+1}))}|X_t = \widehat{X}_t] && (7.76)\\
&= e^{i(c.d_{t+1})} E^P[e^{i(c.Z_{t+1}X_{t+1}))}|X_t = \widehat{X}_t] && (7.77)\\
&= e^{i(c.d_{t+1})} \phi_{\widehat{X}_t}(Z'_{t+1}c) && (7.78)
\end{aligned}$$

As a result, we have that the probability density function of Y_{t+1} conditional on Y_t, $f(Y_{t+1}|Y_t)$ is given by

$$f(Y_{t+1}|Y_t) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \phi_{\widehat{X}_t}(Z'_{t+1}c) \phi_{\epsilon_t}(c) e^{i(c.d_{t+1}-Y_{t+1})} dc \tag{7.79}$$

where N denotes the size of Y_t.

For large values of N, this approach will thus unfortunately become computationally very intensive.

The vector ψ is then estimated by maximizing the objective function

$$LogLikelihood = \sum_{t=1}^{T-1} \log f(Y_{t+1}|Y_t) \tag{7.80}$$

7.3.3 Revisiting the Kalman Filter Algorithm with the Bernoulli approximation for jumps

One can use the same Bernoulli approximation for the jumps that has been made to improve the Kalman Filter Algorithm. Consider the two probability densities $p_0(Y_t|Y_{t-1})$ and $p_1(Y_t|Y_{t-1})$ computed previously. If $p_0(Y_t|Y_{t-1}) < p_1(Y_t|Y_{t-1})$ then one can assume that there has not been any jump between the times $t-1$ and t. In this case, one can replace equations (7.22) and (7.23) by the equations (7.57) and (7.58).

Similarly, if $p_0(Y_t|Y_{t-1}) > p_1(Y_t|Y_{t-1})$ then one can assume that there has been a jump between the times $t-1$ and t. In this case, one can replace equations (7.22) and (7.23) by the equations (7.59) and (7.60).

The rest of the algorithm stays unchanged.

7.3.4 An Expected Conditional Characteristic Function estimator

Along the lines of Singleton (2001) [45], we begin by constructing an estimator that is asymptotically equivalent to the Maximum Likelihood estimator, using the Expected Conditional Characteristic Function (ECCF). Let ψ_0 denote the vector of size p of data generating parameters to be estimated. Let Z_T^{∞} denote the class of "continuous grid" estimators defined as follows : each $z \in Z_T^\infty$ indexes an estimator $\psi_{\infty T}^z$ associated with the instrument function $z_t(u) : \mathbb{R}^N \to \mathbb{C}^p$, where $\mathbb{C}$ denotes the complex numbers, with $z_t(u) \in I_t$, $z_t(u) = \overline{z}_t(-u)$, $t = 1, \ldots, T$, where I_t is the σ-algebra generated by Y_t; and the estimator $\psi_{\infty T}^z$ of ψ_0 satisfies

$$\frac{1}{T}\sum_t \int_{\mathbb{R}^N} z_t(u)[e^{iu'Y_{t+1}} - \phi_{Y_t}(u, \psi_{\infty T}^z)]du = \mathbf{0} \tag{7.81}$$

Proposition 7.2 (from Singleton (2001)[45]): *Under regularity conditions, $\psi_{\infty T^z} \in Z_T^\infty$ is consistent. Furthermore, the optimal index in Z_T^∞, in the sense of giving the smallest asymptotic covariance matrix among continuous-grid ECCF estimators, is*

$$z_{\infty T}^* = \frac{1}{(2\pi)^N}\int_{\mathbb{R}^N} \frac{\partial \log f(Y_{t+1}|Y_t, \psi_0)'}{\partial \psi} e^{-iu'Y_{t+1}} dY_{t+1} \tag{7.82}$$

Finally, the continuous-grid estimator based on the index $z_{\infty t}^*$ is asymptotically equivalent to the ML estimator based on the true conditional density function of Y_t.

Proof: See Singleton (2001)[45] □.

From a practical perspective, the ECCF estimator $z_{\infty t}^*$ has no computational advantages over the ML estimator because the index z_∞^* cannot be computed without a priori knowledge of the conditional density function. Accordingly, we proceed to develop a computationally tractable estimator that is consistent and "nearly" as efficient as the ML estimator. For notational simplicity, we set $N = 1$ (Y_t is one dimensional) in the remainder of this section. The basic idea is to approximate the integral

$$\int_{\mathbb{R}} z_t(u)[e^{iuY_{t+1}} - \phi_{Y_t}(u, \psi)]du \tag{7.83}$$

with the sum over a finite grid in $\mathbb{R}$.

More precisely, we fix the interval $[-K\tau, K\tau] \subset \mathbb{R}$ and divide it into $(2K+1)$ equally spaced intervals of width τ. Let Z_T^K denote the class of FG-ECCF estimators ψ_{KT}^z that solve

$$\frac{1}{T}\sum_t \tau \sum_{k=-K}^{K} z_t(k\tau)[e^{ik\tau Y_{t+1}} - \phi_{Y_t}(k\tau, \psi_{KT}^z)]du = 0 \tag{7.84}$$

for $z_t(u) \in \mathbb{C}^p$ with the same characteristics as in the case of continuous-grid estimation. Solving (7.84) is equivalent to solving

$$\frac{1}{T}\sum_t \tau \sum_{k=-K}^{K} Re\left[z_t(k\tau)[e^{ik\tau Y_{t+1}} - \phi_{Y_t}(k\tau, \psi_{KT}^z)]\right] du = 0 \tag{7.85}$$

This expression, in turn, can be simplified further by letting

$$\begin{aligned} \widetilde{\epsilon}_{K,t+1}(\psi)' \;\equiv\; & \Big(cos(\tau Y_{t+1}) - Re\phi_{Y_t}(\tau,\psi), \ldots, cos(K\tau Y_{t+1}) - Re\phi_{Y_t}(K\tau,\psi), \\ & sin(\tau Y_{t+1}) - Im\phi_{Y_t}(\tau,\psi), \ldots, sin(K\tau Y_{t+1}) - Im\phi_{Y_t}(K\tau,\psi)\Big), \end{aligned} \tag{7.86}$$

and $\widetilde{z}_{Kt}$ denote the $p \times 2K$ real matrix with the $Re[z_t(k\tau)]$ and $Im[z_t(k\tau)]$ being the first K and second K columns, respectively. Then (7.85) becomes

$$\frac{1}{T}\sum_t \widetilde{z}_{Kt}\widetilde{\epsilon}_{K,t+1}(\psi_{KT}^z) = 0 \tag{7.87}$$

Proposition 7.3 (from Singleton (2001)[45]): *The optimal index $\widetilde{z}_{Kt}^* \in Z_T^K$, in the sense of giving the smallest covariance matrix of the finite-grid ECCF estimator, is*

$$\widetilde{z}_{Kt}^* = \Phi_t^{K'} \times \{\Sigma_t^K\}^{-1} \tag{7.88}$$

where the $2K \times p$ matrix Φ_t^K is $\partial\widetilde{\epsilon}_{K,t+1}(\psi_0)/\partial\psi$ and $\Sigma_t^K \equiv E[\widetilde{\epsilon}_{K,t+1}\widetilde{\epsilon}'_{K,t+1}|Y_t]$

Proof: See Singleton (2001)[45] □.

The elements of the matrix Σ_t^K are known in closed-form as functions of the real and imaginary parts of $\phi_{Y_t}(k\tau)$. The elements of the matrix of derivatives of $\epsilon_{K,t+1}$ with respect to ψ involve only the derivatives of $Re[\phi_{Y_t}(k\tau, \psi, \widehat{X}_t^{\psi})]$ and $Im[\phi_{Y_t}(k\tau, \psi, \widehat{X}_t^{\psi})]$ with respect to ψ, where we made explicit the dependence of $\widehat{X}_t^{\psi}$ on ψ.

Indeed, making the same approximation as in equation (7.74), the characteristic function $\phi_{Y_t}(k\tau, \psi, \widehat{X}_t^{\psi})$ can be written

$$\begin{aligned} \phi_{Y_t}(k\tau, \psi, \widehat{X}_t^{\psi}) &\approx \phi_{d_t + Z_t \widehat{X}_t}(k\tau, \psi, \widehat{X}_t^{\psi}) \phi_{\epsilon_t}(k\tau, \psi) && (7.89) \\ &= e^{i(k\tau d_{t+1})} \phi_{\widehat{X}_t}(Z'_{t+1} k\tau, \psi, \widehat{X}_t^{\psi}) \phi_{\epsilon_t}(k\tau, \psi) && (7.90) \\ &= e^{i(k\tau d_{t+1})} e^{\alpha(t) + \beta(t).\widehat{X}_t^{\psi}} \phi_{\epsilon_t}(k\tau, \psi) && (7.91) \end{aligned}$$

All of the elements in this last equation but $\widehat{X}_t^{\psi}$ can easily be derived with respect to ψ by eventually solving additional systems of ordinary equations. The partial derivative of $\widehat{X}_t^{\psi}$ with respect to ψ is computed in **Proposition 7.1**.

7.4 Example of implementation

For illustration purposes, we have applied the numerical routines developed above to three main references in the energy industry in Europe: electricity quoted at the Frankfurt based European Energy Exchange (EEX), natural gas quoted at the International Petroleum Exchange (IPE) of London for natural gas in the UK natural gas grid at the National Balancing Point (NBP), and the Brent crude also quoted on the International Petroleum Exchange (IPE) of London. There has been, however, no attempt here to try to estimate market risk premiums so that the processes are the same under both the true and risk-neutral probability measure. The estimation results, since they are based on futures prices observations, were indeed quite robust to estimate the level of expected prices in the risk-neutral world, but the repartition of this expected level in the risk-neutral world between a risk premium and an expected level in the true probability measure turned out not to be as robust. One possible reason is that the length of the spot price time series considered was probably not long enough for a stable estimation.

7.4.1 The European Energy Exchange (EEX)

The Frankfurt-based European Energy Exchange (EEX) is one of the largest energy market in Continental North Western Europe. The German electricity market is the biggest in continental Europe by number of players and generation capacity. It is also the fastest to open up, with immediate 100% full customer choice, though UK and parts of Scandinavia were opened earlier. Germany's electricity market is indeed huge (with annual power consumption about 550 TWh and generation capacity of 125 GW, half of which is coal fired, 17% nuclear, and 18% gas-fired) and centrally located in the heart of Europe, with wires and gas pipes connected to the rest of the continent.

The daily average spot price from the 02/03/2001 to the 27/06/2002, expressed in Eur/MWh, is plotted in the next figure.

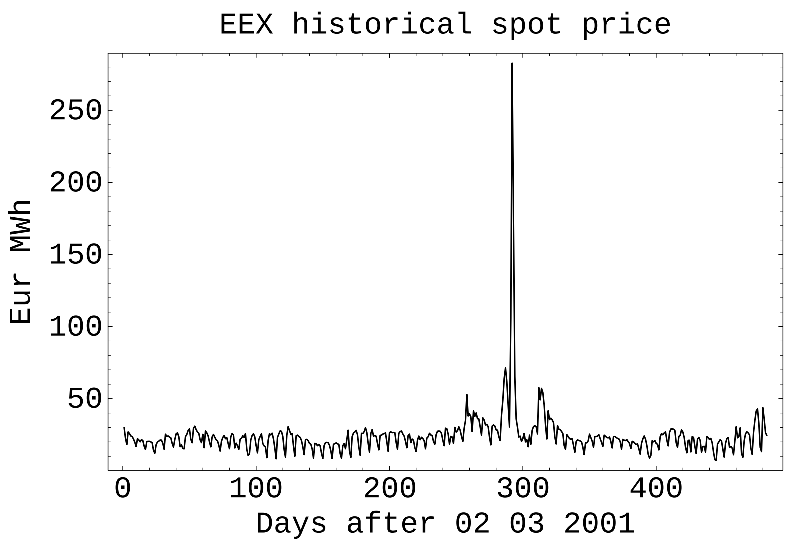

Figure 7.4.1: Historical daily average spot price for the EEX

A weekly pattern is clearly recognizable and will be thus included in the model. It is also quite obvious to identify the occurrence of a spike from 17/12/2001 to 19/12/2001 (a bit less than 300 days after 02/03/2001 on the graph). There has, however, been only one spike, so we will only use one spiking variable in our model with parameters that will be calibrated on its level, duration and frequency of occurrence.

Unfortunately, the seasonality is not really obvious from Figure 7.4.1. In figure 7.4.2, we thus super-imposed on a single graph the futures' term structure that has been observed at each date. We again easily recognize the spike, but the presence of a seasonality indeed becomes much clearer. This is a good illustration on how important the information revealed by the futures contracts is to accurately model the spot price.

Finally, the spot price returns, even after having removed the spikes and patterns,

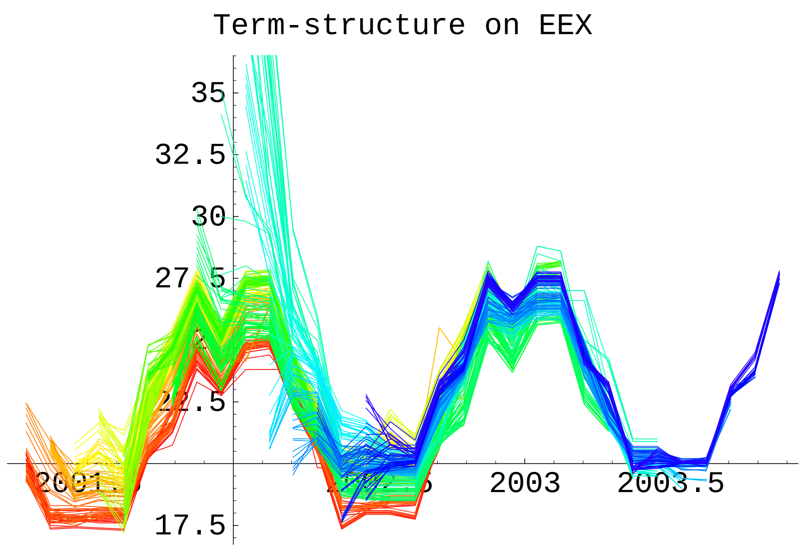

Figure 7.4.2: Evolution of the futures' term structure for the EEX

still have a kurtosis of 4.55. This excess kurtosis will be modelled by adding jump processes to the risk factors.

In summary, the main characteristics of the EEX spot price are the presence of a weekly and seasonal pattern, and the presence of jumps and spikes. Such characteristics are reproduced in the following model for the daily average spot price model, expressed in Eur/MWh:

$$
\begin{aligned}
\log(Spot_t) &= Seas\_pattern(t) + Weekly\_pattern(t) + X_1(t) + X^S(t) \\
Seas\_pattern(t) &= a + b\ t + c_1 \cos(2\pi(t - c_2)/365) \\
Weekly\_pattern(t) &= \begin{cases} d_{mon} & \text{if } t \in \text{Monday} \\ d_{tue} & \text{if } t \in \text{Tuesday} \\ d_{wed} & \text{if } t \in \text{Wednesday} \\ d_{thu} & \text{if } t \in \text{Thursday} \\ d_{fri} & \text{if } t \in \text{Friday} \\ d_{sat} & \text{if } t \in \text{Saturday} \\ d_{sun} & \text{if } t \in \text{Sunday} \end{cases}
\end{aligned}
$$

The dynamics of the Jump-Diffusive risk factors are given by:

$$
\begin{aligned}
dX_1(t) &= \kappa_1(X_2(t) - X_1(t))\ dt + \sigma_1\ dW_1(t) + dJ_1^{up}(t) + dJ_1^{down}(t) \\
dX_2(t) &= -\kappa_2 X_2(t)\ dt + \rho\sigma_2\ dW_1(t) + \sqrt{1-\rho^2}\ \sigma_2\ dW_2(t) + dJ_2^{up}(t) + dJ_2^{down}(t)
\end{aligned}
$$

where $J_i^{up}(t)$ and $J_i^{down}(t)$ are pure jump processes with constant equal arrival rates, λ_i, and with jump sizes normally distributed with mean μ_i^J and $-\mu_i^J$, respectively, and equal standard deviations, σ_i^J. The up and down jumps are thus assumed to be symmetric. This assumption is not so well supported by the data (the spot returns without the spikes still have a skewness of 0.9) and could certainly be relaxed in further developments, but it allows to limit the number of parameters to be estimated by the Maximum Likelihood algorithm.

The dynamics of the spiking risk factor is given by:

$$
dX^S(t) = dJ_s^{up}(t) + dJ_s^{down}(t)
$$

where

- $J_s^{up}(t)$ and $J_s^{down}(t)$ are pure jump processes with a constant jump size, $Spike$ and $-Spike$, respectively;
- $J_s^{up}(t)$ and $J_s^{down}(t)$ have arrival rate equal to

$$
\lambda_s^{up}(t) = l_s^{up}(t)\left(1 - X^S(t)(Spike)^{-1}\right) \tag{7.92}
$$

$$
\lambda_s^{down}(t) = l_s^{down}(t) X^S(t)(Spike)^{-1} \tag{7.93}
$$

  where X^S is the spiking risk factors, and $Spike$ is the level of the spikes.

In order to limit the number of parameters to be estimated via the Maximum Likelihood procedure, the day-level parameters are estimated by Minimum Mean Square Error on the historical spot data and kept fixed in the Maximum Likelihood.

The likelihood function to be optimized, however, is not very well behaved and has many different local optima. It is thus crucial to start the optimization procedure with starting values that are not too far from the optimal solution, so that the optimization algorithm does not get trapped into local optima.

Starting values for the parameters of the deterministic trend, $\theta(t)$, can be computed by minimizing the mean square error on the historical spot and futures data of figures 7.4.1 and 7.4.2.

Starting values for the speed of reversion and volatilities, κ_i and σ_i, can be computed by matching the futures' historical volatility term structure computed on a historical time series of their returns where the jumps and spikes have been filtered away. The filter algorithm computes the standard deviation of the returns and checks whether the proportion of returns with absolute value greater than 2.575 times the standard deviation is greater than 1%. If it is greater, then these returns are considered as jumps and filtered away, the standard deviation of the new data-set is recomputed and the algorithm repeats itself until the appropriate proportion of returns that are greater than 2.575 times the standard deviation (in absolute value) is reached. The starting value for the arrival rates of the jumps on the first factor is computed by matching the frequency of jumps implied by the proportion of returns that have been identified as jumps on the spot price, and the starting value for the arrival rates of the jumps on the second factor is computed by matching the frequency of jumps implied by the proportion of returns that have been identified as jumps on a long-term futures contract. Finally, the starting values for the expected jump size and standard deviation can be computed by looking at the distribution of the returns that have been filtered away.

The futures' volatility term structure is indeed a crucial element that is very important to reproduce because it determines the standard deviation of the spot price (or more precisely its logarithm)'s distribution at a given maturity. This curve is in fact so important to match that it could be given some additional weight in the Maximum Likelihood objective function. We decided to do this by constraining the Maximum Likelihood algorithm.

The decision to constrain the Maximum Likelihood algorithm is further supported by the observation that non-constrained Maximum Likelihood estimators tended to under-estimate the futures' volatility term structure. This can be easily understood by recognizing that the model will tend to model the futures' volatility through the volatility of the risk-factors only up to the level that can be explained by only two risk-factors, the rest of the volatility will be explained by the measurement errors' volatility, very similarly to what happens in a principal components decomposition of a variance covariance matrix, where, if the rank of the covariance matrix is greater than 2, only a proportion of the total variances can be explained by the first two principal components. Measurement errors, however, are then left aside and neglected when using the estimated model. Constraining the Maximum Likelihood to match the futures' volatility term structure will thus correct the estimation for this bias when the number of risk-factors is not sufficient to fully describe the joined dynamics of the futures of all maturities.

Special attention should be given to the choice of the constraints that are imposed on the Maximum Likelihood algorithm. Indeed, it may not be possible to perfectly match the whole historical futures' volatility term structure, because the model does not have enough degrees of freedom. The Maximum Likelihood would thus result in a constrained optimization problem with unfeasible constraints.

The approach that has been followed here is to divide the set of parameters to be estimated into two subsets, $\psi = \{\psi^1, \psi^2\}$, and to maximize the Loglikelihood function under the constraint that the sum of squared errors between the model implied and historical volatilities is minimized with respect to the parameter subset ψ^2.

The arbitrary part of this approach resides in how to split the set of parameters to be estimated. We found that using $\psi^2 = \{\sigma_1, \sigma_2\}$ provided a very good trade-off between the fit of the futures' volatility term structure and the likelihood achieved.

More rigorously, one can gain some intuition on the constraint that has been imposed by looking at the optimization problem via the Lagrange multipliers. The initial (unfeasible) constrained optimization problem, where one seeks to perfectly fit the historical future's volatility term structure, can indeed be formalized as follows:

$$\begin{aligned} \max_{\psi} \quad & f(\psi) \\ s.t. \quad & \left(g_i^{mod}(\psi) - g_i^{hist}\right) = 0 \end{aligned} \tag{7.94}$$

where $f(\psi)$ denotes the log-likelihood function to be maximized; g_i^{hist} denotes the historical daily volatility of the spot price ($i = 0$) or of the futures contract with maturity i, $i = 1...n$, and $g_i^{mod}(\psi)$ the daily volatility implied by the model.

Solving this constrained optimization problem is equivalent to solving

$$\min_{\lambda} \left\{ \max_{\psi} \left\{ f(\psi) + \sum_{i=0}^{n} \lambda_i \left(g_i^{mod}(\psi) - g_i^{hist}\right) \right\} \right\} \tag{7.95}$$

where λ is the vector of Lagrange multipliers.

As mentioned above, it is however not possible to satisfy all the constraints, and equation (7.95) will thus go to $-\infty$. We will now show that the approach that has been proposed is in fact equivalent to introducing some constraints between the different λ_i so that the minimization problem now gets bounded. Writing the optimality conditions of the optimization problem on the parameter subset ψ^2 to best match by minimum mean square error the futures' volatility term structure curve, the proposed methodology is equivalent to

$$\begin{aligned} \max_{\psi} \quad & f(\psi) \\ s.t. \quad & \sum_{i=0}^{n} 2\ \nabla_{\psi^2}\left(g_i^{mod}(\psi)\right)\left(g_i^{mod}(\psi) - g_i^{hist}\right) = \mathbf{0} \end{aligned} \tag{7.96}$$

Again, this constrained optimization problem has another formulation using Lagrangian multipliers, λ^*:

$$\min_{\lambda^*}\left\{\max_{\psi}\left\{f(\psi)+\sum_{j=1}^{m}\lambda_j^*\sum_{i=0}^{n}\ 2\ \nabla_{\psi_j^2}\left(g_i^{mod}(\psi)\right)\left(g_i^{mod}(\psi)-g_i^{hist}\right)\right\}\right\} \tag{7.97}$$

where m denotes the size of the subset of parameters ψ^2.

Comparing (7.95) with (7.97), we observe that we have in fact constrained the λ_i's such that:

$$\lambda_i=\sum_{j=1}^{m}\lambda_j^*\ 2\ \nabla_{\psi_j^2}\left(g_i^{mod}(\psi)\right) \tag{7.98}$$

The optimization problem (7.97) has been solved by using a purely numerical implementation of a Quasi-Newton algorithm (gradients are computed numerically).

Note that instead of calibrating the parameters ψ^2 to fit a historical futures' volatility term structure, one could try to fit option (call or put) prices on futures with different maturities, which would allow to fit forward looking volatilities (implied by the option prices) rather than historical ones.

Taking the spot and futures prices from 02/03/2001 to 27/06/2002, and setting the reference date $t=0$ the 01/01/2001 with time expressed in days, the parameters estimates (and their standard deviation) defining the different patterns are provided in Table 7.1.

The standard deviations of the parameter estimates for ψ^1 have been estimated by computing the Fisher's Information matrix, assuming ψ^2 fixed. We also reported, for information, the standard deviations of the parameters estimates for $\psi^2=\{\sigma_1,\sigma_2\}$, estimated via a Fisher's Information matrix computed as if the Likelihood was unconstrained. Standard deviations of the weekly pattern parameter estimates are not available because these parameter have been kept fixed in the Maximum Likelihood estimation.

Parameter	Estimate	Standard deviation
a	3.083	(0.22)
b	0.0000824	(0.00014)
c_1	0.201	(0.024)
c_2	0.0[1]	(9.39)
d_{mon}	1.0388	N.A.
d_{tue}	1.0159	N.A.
d_{wed}	1.0333	N.A.
d_{thu}	1.0279	N.A.
d_{fri}	1.0253	N.A.
d_{sat}	0.9475	N.A.
d_{sun}	0.8839	N.A.

Table 7.1: Parameter estimates for the EEX trends.

The parameter estimates defining the jump-diffusion risk factors are reported in Table 7.2.

Parameter	Estimate	Standard deviation
κ_1	0.1853	(0.031)
σ_1	0.07823	(0.005)
κ_2	0.00437	(0.0016)
σ_2	0.0171	(0.0094)
ρ	0.0[2]	N.A.
μ_1^J	0.295	(0.027)
μ_2^J	0.0022	(0.48)
σ_1^J	0.13	(0.038)
σ_2^J	0.10	(0.219)
λ_1	0.0828	(0.032)
λ_2	0.0021	(0.28)

Table 7.2: Parameter estimates for the EEX jump-diffusion risk factors

The standard deviations of the estimated parameters are quite low, excepted for the parameters defining the jumps on the second factor.

[1]Actual estimate was -1.00018 that has been fixed to 0.0 given the large standard deviation of the estimate: 9.39.

[2]The correlation has been fixed to zero for numerical stability reasons

The spike parameters have been estimated assuming the spikes could be identified just by looking at a plot of the historical spot prices and by equating their average level, their frequency of arrival and expected duration. They have then been kept fixed in the Maximum Likelihood estimation. Their estimated values are reported in Table7.3.

Parameter	Estimate
$Spike$	1.511
l^{up}	0.0000734
l^{down}	0.3025

Table 7.3: Parameter estimates for the EEX spiking risk factor

The autocorrelogram of the spot price residuals plotted in figure 7.4.3 shows that there is no autocorrelation that is statistically significant (at a 95% confidence level), so that most of the structure in the spot price dynamics has indeed been adequately modelled.

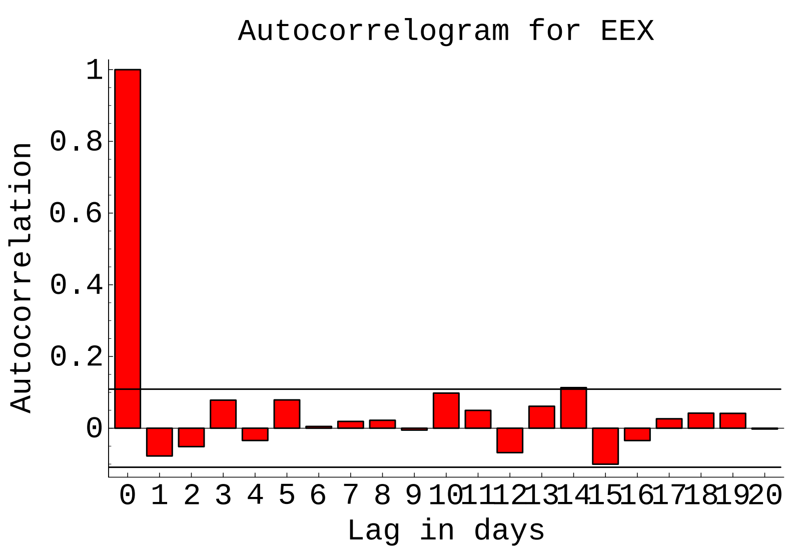

Figure 7.4.3: Autocorrelogram of the EEX spot price residuals

7.4.2 The NBP natural gas market

Based on the historical time series of NBP spot prices in pence per therm plotted on figure 7.4.4, we decided to include only positive and negative jumps on the first risk factor but no spikes like in the electricity markets. We did not allow for a jump process on the second risk-factor because the estimation of these parameters encountered some convergence problems.

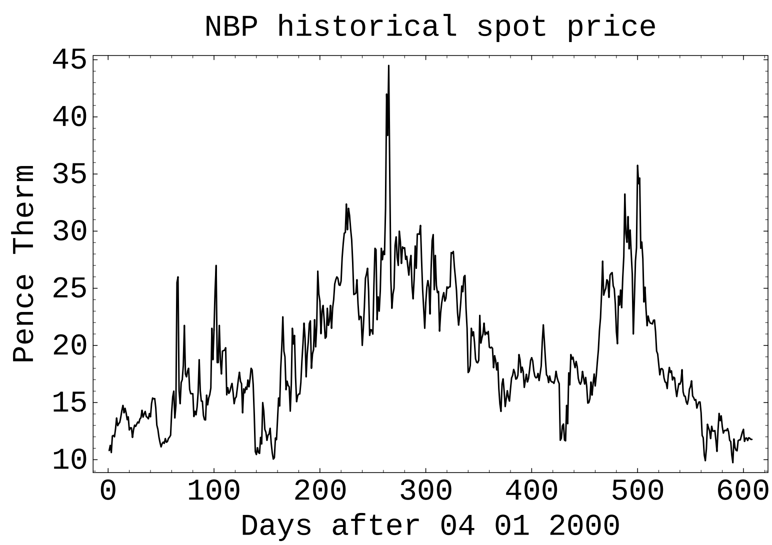

Figure 7.4.4: Historical daily spot price for NBP

Again, the seasonality can be made observable by looking at the super-imposed futures' term structures that have been observed form 04/01/1999 to 14/06/2002 and plotted in figure 7.4.5.

We thus chose a two-factor process with jumps on the first factor that are normally distributed with mean zero, and a significant variance. This specification of jump processes is a parsimonious way of allowing for symmetric positive and negative jumps in a single jump process. Note that the assumption of symmetric positive and negative jumps holds better here as the skewness of the spot price returns is

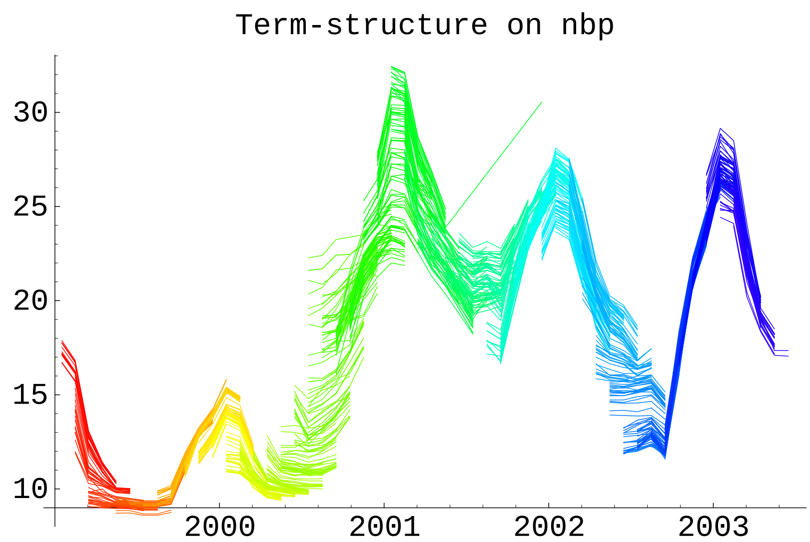

Figure 7.4.5: Evolution of the futures' term structure for the NBP

only -0.0046.

The spot price model, expressed in British pence per therm, can be formalized as follows:

$$\begin{aligned}\log(Spot_t) &= Seas\_pattern(t) + X_1(t)\\ Seas\_pattern(t) &= a + c_1\cos(2\pi(t-c_2)/365)\end{aligned}$$

with

$$\begin{aligned}dX_1(t) &= \kappa_1(X_2(t) - X_1(t))\ dt + \sigma_1\ dW_1(t) + dJ_1(t)\\ dX_2(t) &= -\kappa_2 X_2(t)\ dt + \rho\sigma_2\ dW_1(t) + \sqrt{1-\rho^2}\ \sigma_2\ dW_2(t)\end{aligned}$$

where J_1 is a pure jump processes with jumps that are normally distributed with mean zero and standard deviation σ_1^J, and constant arrival rate equal to λ_1.

The linear drift has been fixed to zero based on a review of the literature on long term forecasts of fuel prices. See for example [24]. We indeed believe that it is more accurate to rely on such long term forecasts rather than on such a short historical time series for the estimation of the trend.

We disregarded the data before 2000 because the volatility increased with the changes in the continental gas market that took place in 2000 (and their impact through the interconnector). Taking thus the spot and futures prices from 04/01/2000 to 14/06/2002, and setting the reference date $t = 0$ the 01/01/2001 with time expressed in days, the parameter estimates and their standard deviation obtained via the same constrained Maximum Likelihood estimation procedure as for the electricity model are provided in Table 7.4.

Parameter	Estimate	Standard deviation
a	3.437	(0.86)
c_1	0.176	(0.009)
c_2	23	(3.4)
κ_1	0.0502	(0.0074)
σ_1	0.0498	(0.001)
κ_2	0.0[3]	($1.7\ 10^{-5}$)
σ_2	0.0111	(0.00016)
ρ	-0.213	(0.014)
λ_1	0.0683	(0.01)
σ_1^J	0.249	(0.01)

Table 7.4: Parameter estimates for the NBP natural gas model

We observe that all standard deviations are quite low.

The autocorrelogram of the spot price residuals is plotted in figure 7.4.6. The autocorrelation level of the two first lags indicate that there probably still is some structure in the spot price dynamics that could be modelled.

[3]The original parameter estimate was in fact $-3.55\ 10^{-6}$, but has been set to zero given the great standard deviation of this estimate, and because models with negative speed of reversion are instable.

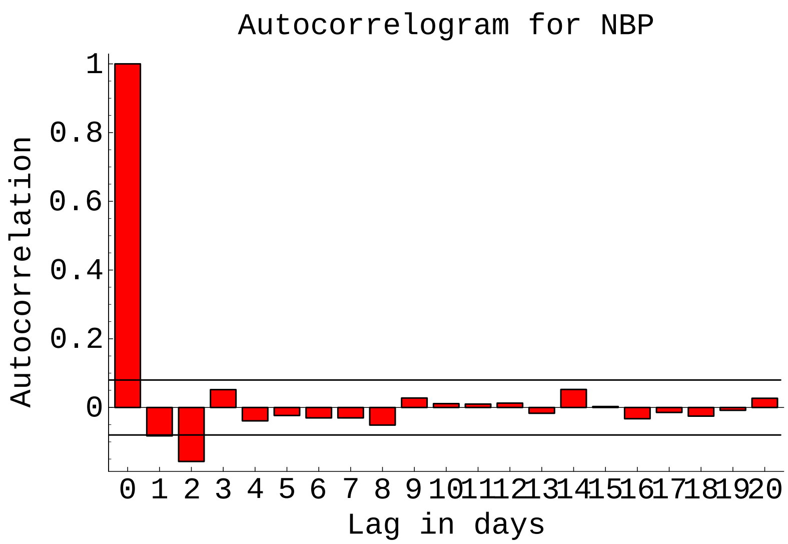

Figure 7.4.6: Historical daily spot price for NBP

7.4.3 The London Brent

The London Brent crude oil historical spot price, in US$ per barrel is plotted on figure 7.4.7. The dynamics are very similar to the NBP spot price dynamics. There is, however, no seasonality on the Brent. The same model (but without any seasonality) has thus be estimated on this market, using the same "constrained" Maximum Likelihood estimation procedure.

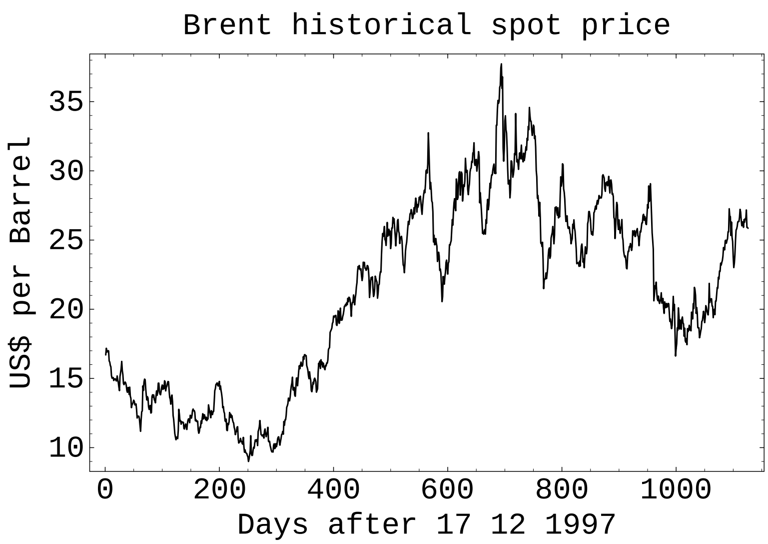

Figure 7.4.7: Historical daily spot price for Brent

Taking the spot and futures prices from 17/12/1997 to 17/05/02, and setting the reference date $t = 0$ the 01/01/2001 with time expressed in days, the parameters estimates and their standard deviations are provided in Table 7.5.

Parameter	Estimate	Standard deviation
a	2.588	(0.04)
κ_1	0.00141	$(1.6\ 10^{-5})$
σ_1	0.000413	(0.00034)
κ_2	0.0171	$(1.8\ 10^{-5})$
σ_2	0.0162	(0.0009)
ρ	0.0259	(0.04)
λ_1	0.02722	(0.0068)
λ_2	0.03721	(0.03)
σ_1^J	0.0652	(0.014)
σ_2^J	0.0462	(0.003)

Table 7.5: Parameter estimates for the Brent model

The estimations for the correlation parameter ρ and the jump parameters associated to the second factor have relatively big standard deviations, indicating a lack of accuracy in their estimated values.

The autocorrelogram of the Brent spot price residuals is plotted in figure 7.4.8.

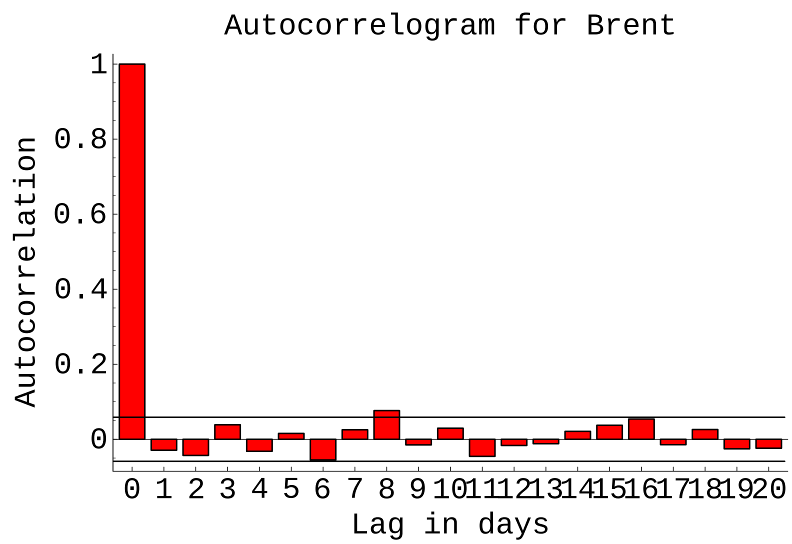

Figure 7.4.8: Historical daily spot price for Brent

Since only the autocorrelation of lag 8 is slightly significant, the structure of the Brent spot price dynamics can be considered as adequately modelled.

Chapter 8

Conclusion

Based on the Affine Jump Diffusion setting, this paper has provided a unified framework that addresses critical issues that power companies face in deregulated markets.

In Chapter 2, we studied the relationships between the modelling activities, the estimation activities, and the pricing activities. We showed that if futures prices were considered as primary commodities and were modelled as such, then these futures could be used in replication strategies to price other contingent claims. Furthermore, no-arbitrage arguments constrained these models to satisfy the property that futures were equal to the expected spot price at their maturity under an equivalent probability measure. We thus proposed a multi-factor Affine Jump Diffusion process to model the spot price in the physical probability measure and a parametrization of the state process defining a class of equivalent martingale probability measures. This parametrization was chosen such that the AJD property of the spot price model was preserved under both probability measures, and such that the sources of risk arising from the diffusion and the jump components of the stochastic model were both priced. We believe that a new insight has been provided in this chapter on how to coherently build an AJD framework to model spot and futures prices and price contingent claims on them.

In Chapter 3, we introduced a novel framework to estimate expectations when only the characteristic function of the underlying variables is known (as opposed to their probability density function) this framework is based on the distribution theory and uses the Parseval Identity as key concept.

The framework developed in Chapter 3 was then applied in Chapter 4 to the option pricing problem and provided closed form solutions to a wide variety of exotic options when their payoff have a known Fourier transform. In particular, the options formulae developed by Bakshi and Madan (1999) [1] and Duffie, Pan and Singleton (1999) [15] can be shown to be special cases of our framework. We also retrieve the Fast Fourier Transforms introduced by Carr and Madan (1999) [4] to price options with different strike prices at low cost. The closed form solutions, however, require some numerical integrations that are sometimes multi-dimensional. Some further attention could thus be devoted to apply efficient numerical integration routines.

Some options, however, do not have Fourier transforms that are known in closed form. In Chapter 5, we thus adapted the theoretical framework of chapter 3 to develop a numerical method that can be applied to options with arbitrary payoffs. The method builds on a cosine decomposition of the payoff function on its interval of interest. The coefficients of the cosines are computed by a Fast Fourier Transform or a Fast Cosine Transform. The Parseval Identity is then applied in this new framework. The method is also of particular interest in grid algorithms such as the Dynamic Programming that is required to price options with an endogenous exercise strategy. The method involves Fast Fourier Transforms or Fast Cosine Transforms with as many dimensions as there are risk-factors. For high dimensional problems the computational burdens become significative. One could think of further developing the methodology by taking advantage of the sparse structure of the Fourier transforms and use FFT or FCT algorithms that only compute the lower frequencies.

The novel theoretical framework put in light by the Parseval identity led to a unification of the closed form solutions developed up to now in the literature. The original numerical approach that has been derived from it allows to fully use the power of Fast Fourier Transforms to efficiently price a wide range of options for which closed form solutions do not exist.

In Chapter 6, we show how to compute partial derivatives, often referred to as Greeks, of the option prices with respect to the underlying state variables or other parameters of the stochastic model. These partial derivatives can be used bo build hedging portfolios by taking positions in other contracts with opposite sensitivity with respect to these variables.

Finally, Chapter 7 uses the Kalman Filter recursive algorithm to estimate the parameters of the AJD process and the risk factors based on historical prices of spot and futures contracts. In this area, further developments in the optimization routines would allow to estimate the parameters of increasingly refined models. Adaptations to the Kalman Filter to allow for non-Gaussian errors and its introduction in the Expected Conditional Characteristic Function estimation method are our main contributions in this chapter.

Bibliography

[1] Bakshi G. and Madan D. (1999), *"Spanning and Derivative-Security Valuation"*, Forthcoming in the Jounal of Financial Economics.

[2] Bernnan, M. (1991), *The Price of Convenience and the Valuation of Commodity Contingent Claims"*, in D. Lund and B. Oksendal (eds.), *Stochastic Models and Option Models*, Elsevier Science Publishers.

[3] Black, F. and Scholes, M. (1973), *The Pricing of Options and Corporate Liabilities"*, Journal of Political Economy, 81 (May-June 73),673-54.

[4] Carr P. and Madan D. (1999), *"Option Valuation and the Fast Fourier Transform"*, Journal of Computational Finance, June 1999.

[5] Clewlow, L., and Strickland, Ch. (1999), *A Multi-Factor Model for Energy Derivatives Risk Management"*, Working Paper, University of Technology, Sydney.

[6] Cox, Ingersoll, and Ross (1981), *"The relation between forward prices and futures prices"*, J. Fin. Econom. 9,321-346.

[7] Cox, Ingersoll, and Ross (1985), *"An intertemporal general equilibrium model of asset prices"*, Econometrica, 53, 363-384.

[8] Cox, Ross, and Rubinstein (1979) *"option pricing: a simplified approach"*, J. Finan. Econom. 7, 229-263.

[9] Culot M. and Smeers Y. (2000), *"Multi-Criteria Options : An Analytical Pricing Tool"*, Talk given at the "Quantitative Risk Management in Finance Conference" held at Carnegie Mellon University.

[10] Culot M., de Menten de Horne S. and Smeers Y. (2000), *"Option Pricing with Grid Algorithms Using Fast Fourier Transforms"*, Working Paper, Electrabel.

[11] Dai, Q. and Singleton, K. (1999), *Specification Analysis of Affine Term Structure Models*, Working Paper.

[12] (2001), *"Applications to Mathematical Finance"*, Working Paper, Eodgenssische Technische Hochschule, Zrich.

[13] Deng S. (1999), *"Stochastic Models of Energy Commodity Prices and their Applications : Mean-Reversion with Jumps and Spikes"*, Industrial and System Engineering, Georgia Institute of Technology, Atlanta.

[14] Duffie D. (2001), *"Dynamic Asset Pricing Theory"*, Princeton University Press, Princeton, 3^{rd} Edition.

[15] Duffie D., Pan J. and Singleton K. (1999), *"Transform Analysis and Asset Pricing for Affine Jump-Diffusions"*, Mineo, Graduate School of Business, Stanford University, Stanford, California.

[16] Duffie D.(1999), *"Volatility in Energy Prices"*,in Managing Energy Price Risk (second edition), RISK books,London.

[17] Duffie D., Filipovic D., and Schachermayer W. (2001), *"Affine processes and applications in finance"*, Working Paper, Graduate School of Business, Stanford University.

[18] Gibson, R., and Schwartz, E. (1990), *Stochastic Convenience Yield and the Pricing of Oil Contingent Claims*", Journal of Finance, 45 (July), 959-976.

[19] Hamilton (1994), *"Time Series"*, Princeton University Press.

[20] Harrison, J.M., and Kreps, D.M. (1979), *Martingales and arbitrage in multiperiod securities markets"*, J. Econom. Theory 20, 381-408.

[21] Harvey A.C. (1994), *"Forecasting, structural time series models and the Kalman filter"*, Cambridge University Press.

[22] Heston, S.L. (1993) *"A closed-form solution for options with stochastic volatility with applications to bond and currency options"*, Rev. Fin. Stud. 6, 327-343.

[23] Hull and White (1987), *"The pricing of options on assets with stochastic volatilities"*, The Journal of Finance, 42 281-300.

[24] IEO2002: Energy Information Administration, Annual Energy Outlook 2002, DOE/EIA-0383(2002) (Washington, DC, December 2001)

[25] Johnson and Barz (1999).

[26] Kamat R. and Oren Sh. S. (2000), *"Exotic Options for Interruptible Electricity Supply Contracts"*, POWER paper 071, Energy Institute, University of California, Berkeley.

[27] Kaminski V., Gibnev S. and Pinnamaneni K. (1999), *"Energy Exotic Options"*, in Managing Energy Price Risk (second edition), RISK books,London.

[28] Longstaff F.A., and Schwartz E.S. (2001) *"Valuing American Options by Simulation: A Simple Least-Squares Approach"*, The Review of Financial Studies, Spring 2001, Vol. 14, No. 1, pp 113-147.

[29] Manoliu M. and Tompaidis, S. (1999), *"Energy Futures Prices: Term Structure Models with Kalman Filter Estimation"*, Working Paper.

[30] Merton, R. (1973), *"The Theory of Rational Option Pricing"*, Bell Journal of Economics and Management Science, 4 (Spring), pp. 141-83.

[31] M. Musiela and M. Rutkowski (1998), *"Martingale Methods in Financial Modelling"*, Springer.

[32] Nerlove, M., D.M. Grether and J.L. Carvalho (1979), *"Analysis of Economic Time Series"*, New York: Acamemic Press.

[33] Oren Sh. S. (1996), *The Role of Financial Instrument in a Competitive Electricity Market"*, Proceedings: 1996 EPRI Conference on Innovative Approaches to Electricity Pricing.

[34] Jun Pan (2001),*The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study*, Working Paper.

[35] Pilipovic, D. (1998),*Energy Risk*, McGraw Hill, New York.

[36] Pindyck R.S. (1999), *"The Long-Run Evolution of Energy Prices"*, The Energy Journal, vol 20, n2, pp. 1-27.

[37] RISK (1999), *"Energy Modelling and the Management of Uncertainty"*, RISK books, London.

[38] Routledge, Seppi, and Spatt (1998), *"The "Spark Spead": An Equilibrium Model of Cross-Commodity Price Relationships in Electricity"*, Working Paper, Graduate School of Industrial Administration.

[39] Safarov, Yu., (1996) *"Distributions, Fourier Transforms and Microlocal Analysis"*, Working Paper.

[40] Schwartz E.S. (1997), *"The Stochaztic Behavior of Commodity Prices: Implications for Valuation and Hedging"*, The Journal of Finance, Vol. LII, NO. 3, July.

[41] Scwartz E.S., and Smith, J. (1997), *Short-Term Variations and Long-Term Dynamics in Commodity Prices"*, Working Paper, UCLA.

[42] Seppi, Duane (2001) *"Understanding Stochastic Processes and their Application to the Energy Industry"*, Energy & Power Risk Management.

[43] Shephard, N. (1991), *From Characteristic Function to a Distribution Function: A Simple Framework for Theory"*, Econometric Theory 7, 519-529.

[44] Shumway, R.H., and D.S. Stoffer (2000) *"Time Series Analysis and Its Applications"*, Springer Texts in Statistics, Springer-Verlag New York, Inc.

[45] Singleton (2001), *Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function*, Working Paper, Forthcoming in the Journal of Econometrics.

[46] Singleton K.J., and Umantsev L. (2002), *Pricing Coupon-Bond Options and Swaptions in Affine Term Structure Models"*, Working Paper, Graduate School of Business, Stanford University.

[47] Stein, E.M., and Stein, J.C. (1991) *Stock price distributions with stochastic volatility: an analytic approach*, Rev. Finan. Stud. 4, 727-752.

[48] Van Steenkiste R.J., and Foresi S. (1999), *"Arrow-Debreu Prices for Affine Models"*, Working Paper, Goldman Sachs Asset Management.

[49] Willem (1995), *Analyse Harmonique Réelle*, Hermann.