POLITECNICO
MILANO 1863

Politecnico di Milano

Department of Mathematics

Doctoral Programme in Mathematical Models and Methods in Engineering

Additive normal tempered stable process: a new way to model the implied volatility surface

Author:
Michele Azzone

Supervisor:
Prof. Roberto Baviera

Year 2021, Circle XXXIV

Abstract

In this thesis, we introduce a new class of pure jump additive processes for modeling equity implied volatility surfaces: the additive normal tempered stable process (ATS). We derive its short-time-to-maturity asymptotics. We show that ATS accurately calibrate the volatility level and skew on different days (nine-year dataset). This framework allows pricing European options with classical closed-form methods (e.g., Lewis formula) , on the one hand, and exotic derivatives with fast Monte Carlo schemes , on the other hand.

The ATS is a simple additive process for equity index derivatives. A process is said to be an additive process if it presents independent (but not-stationary) increments. In particular, we present in detail the application of Normal Tempered Stable processes (e.g., NIG and VG) with time-dependent parameters. It accurately fits the equity index volatility surfaces in the whole time range of quoted instruments, including options with small time-horizon (days) and long time-horizon (years).

We introduce the model via its characteristic function; this allows using Fourier pricing techniques. We show that even if the model loses the classical stationarity property of Lévy processes, it presents interesting scaling properties for the calibrated parameters. The two power-law scaling parameters are β, related to the variance of jumps, and δ, related to the smile asymmetry. In option market data, we observe that $\beta = 1$ and $\delta = -1/2$; we build a statistical test that confirms this power-law scaling result.

We examine the short-time-to-maturity behavior of the ATS. As emphasized by empirical studies, a negative skew inversely proportional to the square root of the time-to-maturity characterizes the equity implied volatility. We prove that the implied volatility of these additive processes is consistent, in the short time, with the equity market observed characteristics if and only if $\beta = 1$ and $\delta = -1/2$.

We design a new fast Monte Carlo scheme also for ATS. The scheme leverages on the independence of the increments for additive processes and is based on the ATS characteristic function. We prove some bounds on the method biases and we test its performances against a classic Gaussian approximation method (based on the ATS Lévy measure).

Finally, we test the quality of the calibration on historical S&P 500 and EURO STOXX 50 options prices on a large dataset composed by nine years of closing prices. To calibrate the process on equity market data, we introduce a new technique to recover the implicit discount factor (and forward prices) in the derivative market using only European put and call prices: this discount is grounded in actual transactions in active markets. The (unique) forward contract -built using the put-call parity relation- contains information about the market discount factor: by no-arbitrage conditions, we identify the implicit interest rate such that the forward contract value does not depend on the strike.
Key Words: Additive process, volatility surface, calibration, skew, small-time, forward price, put-call parity, implied interest rate, cost-of-funding, synthetic forward, Monte Carlo

Contents

Abstract **i**

0 Introduction **1**

1 Additive normal tempered stable processes for equity derivatives and power law scaling **9**
1.1 Introduction . . . 9
1.2 The model . . . 10
1.3 Model calibration and power law scaling . . . 14
1.3.1 Dataset . . . 14
1.3.2 Calibration . . . 16
1.3.3 Scaling properties . . . 19
1.4 Model selection and robustness tests . . . 22
1.4.1 Model selection via statistical tests . . . 22
1.4.2 Robustness tests . . . 23
1.5 Conclusions . . . 29

2 Short-time implied volatility of additive normal tempered stable processes **31**
2.1 Introduction . . . 31
2.2 The ATS implied volatility . . . 34
2.3 Short-time ATM implied volatility . . . 39
2.4 Short-time skew . . . 47
2.5 Main Result . . . 53
2.6 Conclusions . . . 53

3 A fast Monte Carlo scheme for additive processes and option pricing **54**
3.1 Introduction . . . 54
3.2 Overview of the MC method for additive process . . . 55
3.2.1 Lewis CDF via FFT . . . 57
3.2.2 CDF error sources . . . 58
3.3 The simulation method . . . 60
3.3.1 Simulation error sources: truncation and interpolation . . . 60
3.3.2 A simulation benchmark: the Gaussian approximation . . . 62
3.4 Numerical results . . . 64
3.4.1 European options: accuracy . . . 65
3.4.2 European options: computational time . . . 69
3.4.3 Discretely monitoring options . . . 70

3.5 Conclusion 71

4 The equity derivatives market: Synthetic forwards and cost-of-funding 73
4.1 Introduction 73
4.2 The methodology and the dataset 75
4.3 S&P 500 and EURO STOXX 50 implicit interest rates 77
4.4 ATS calibration on multiple volatility surfaces 81
4.5 Conclusions 85

5 Conclusions 87

Acknowledgement 89

Bibliography 90

Appendices 100
.1 Proofs of Chapter 1 101
.2 ATS Parameter estimation 105
.3 Additive subordination 108
.4 Basic properties of ATS 114
.5 Short-time limits 117
.6 Proofs of Chapter 3 122
.7 An application to commodities 125

Chapter 0

Introduction

A derivative product is a security whose value depends on another security. Derivative products are now actively traded in most exchanges and their trading volume is often above the volume of the underlying securities. A major branch of quantitative finance is dedicated to options pricing and hedging (see e.g., Hull 2003). A European call (put) option gives the right to buy (sell) for a certain price the underlying security at a certain maturity date. Hence, the call option has discounted pay-off $B_t \left(F_0(t)\, e^{X_t} - F_0(t)\, e^{x}\right)^+$ (and $B_t \left(F_0(t)\, e^{x} - F_0(t)\, e^{X_t}\right)^+$ is the discounted payoff for the corresponding put) where $F_0(t)$ is the underlying forward price[1] with maturity t at time 0, $F_0(t)e^{X_t}$ is the underlying forward price at time t, t is the option time-to-maturity, K is the option strike price, $x := \ln \frac{K}{F_0(t)}$ the asset log-moneyness and B_t the deterministic discount factor between 0 and t. Notice that the price, at time $t = 0$, of an European option depends directly from the law of the forward exponent X_t.

One of the first models that link option prices with the underlying forward dynamic is the world-renowned Black (1976) model. In this framework, the forward exponent X_t is modeled as follows

$$X_t = -\frac{I^2 t}{2} + I\, W_t \ ,$$

where W_t is a Brownian motion with drift and I the volatility parameter. A simple close formula is available for European call option prices:

$$B_t\, F_0(t)\, \mathbb{E}\left[\left(e^{X_t} - e^{x}\right)^+\right] = B_t \left(F_0(t) N(d_1) - e^{x} N(d_2)\right) \quad ,$$

where N is the Gaussian cumulative distribution function and

$$d_{1,2} = \frac{\log(-x)}{I\sqrt{t}} \pm \frac{I\sqrt{t}}{2} \quad .$$

Market prices are available for European options, thus it is possible to invert the formula above to get the market-implied I, the so-called implied volatility. In the simple Black (1976) model implied volatility, obtained from options on the same underlying, should be constant over time-to-maturity t and strike price K. However, this is not the case. In figure 0.1, as an example, we plot the S&P 500 implied volatility surface (i.e. the implied volatility at different log-moneyness and maturities) at the 30^{th} of May 2013.

[1] For the sake of simplicity we define the options directly w.r.t. the forward price and not w.r.t. the spot price.

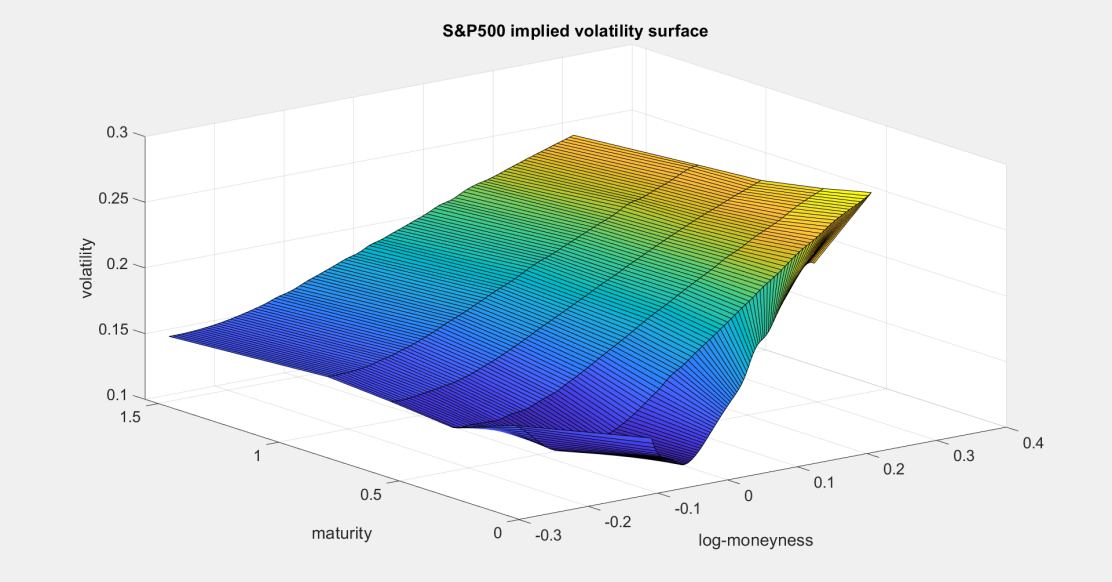

Figure 0.1: Implied volatility surface of S&P 500 at the 30^{th} of May 2013 for time-to-maturities that goes from three weeks to two years.

Typically, volatility surfaces of underlying in the same asset class are similar. We use the S&P 500, which is one of the most liquid equity indexes worldwide, to present the three main characteristics of equity implied volatility.

First, the level of volatility, i.e. the volatility around the at-the-money (ATM), is not constant in time. Second, the *skew*, i.e. the derivative ATM of the implied volatility w.r.t. the strike, is negative and increasing w.r.t. the time to maturity. Third, around the ATM the implied volatility is convex. Numerous model specifications for the forward exponents have been proposed to explain European options market prices; hereinafter we focus on pure jump processes.

Pure jump Lévy processes are a powerful modeling solution that provides parsimonious models consistent with option prices and underlying prices. Mantegna and Stanley (1995) explore the possibility that a pure jump Lévy process explains the power-law scaling phenomena observed in the S&P 500 returns.

Lévy process standard definition follows.

Definition 0.0.1. *(Sato (1999a), Def.1.6, p.3)*
A cadlag stochastic process $\{X_t\}_{t\geq 0}$ *is a Lévy process if and only if*

1. *For any choice of* $n \geq 1$ *and* $0 \leq t_0 < t_1 ... < t_n$*, the random variables* $X_{t_1} - X_{t_0}$*,* $X_{t_2} - X_{t_1}$*,...,*$X_{t_n} - X_{t_{n-1}}$ *are independent.*
2. $X_0 = 0$*.*
3. *The law of the random variable* $X_{t+s} - X_s$ *does not depend on* s*.*
4. $\{X_t\}_{t\geq 0}$ *is stochastically continuous:*

$$\forall \epsilon > 0, \lim_{h \to 0} P\left[|X_{t+h} - X_t| > \epsilon\right] = 0 \ .$$

One of the first applications to option pricing dates back to the seminal work of Madan and Seneta (1990). Lévy processes parsimoniously describe some key features of the market volatility

surface. Moreover, this model class admits a simple closed formula for European derivatives (Carr and Madan 1999, Lewis 2001). The class of Lévy normal tempered stable processes (LTS) appears to be relatively flexible and parsimonious. LTS are pure jump processes, obtained via the well-established Lévy subordination technique (see, e.g., Cont and Tankov 2003, Schoutens 2003). Most of the applications involve two processes in the LTS family: Normal Inverse Gaussian (NIG) (Barndorff-Nielsen 1997) and Variance Gamma (VG) (Madan *et al.* 1998).

The Lévy normal tempered stable processes f_t are usually introduced time-changing a Brownian motion with drift W_t with a Lévy tempered stable subordinator Z_t (e.g. Inverse Gaussian or Gamma process) independent from W_t.

$$f_t = -\left(\eta + \frac{1}{2}\right)\sigma^2\, Z_t + \sigma\, W_{Z_t} + \varphi t \qquad \forall t \in [0,T]\ ,$$

where where η, σ are two real parameters ($\eta \in \mathbb{R}, \sigma \in \mathbb{R}^+$), while the φ is obtained by imposing the martingale condition on the forward price. LTS are characterized by three parameters: σ, which controls the level of the volatility; k, which is related to the convexity of the surface; and η, which is linked to the *skew*.

Pure jump processes generally describe underlying dynamics more parsimoniously than standard jump-diffusion processes. Both the infinitesimal jumps and the diffusion term contribute to the process quadratic variation (see, e.g., Asmussen and Rosiński 2001). The effects of the two components are qualitatively similar: when calibrating the model to the plain vanilla option market, it is rather difficult to disentangle the two components and several sets of parameters achieve similar results (this could cause over-fitting).

Unfortunately, pure jump Lévy models fail to calibrate accurately the entire volatility surface, i.e. they do not reproduce the implied volatilities that are observed in the market data at different time horizons with sufficient precision (see, e.g., Cont and Tankov 2003, Ch.14). Lévy normal tempered stable processes are pure jump models with independent and stationary increments. Why should we consider stationary increments when modeling implied volatility? Model's stationarity is a feature that significantly simplifies the model's tractability, however, it is rather difficult to justify from a financial point of view.

One can select additive processes to overcome this problem. They are an extension of Lévy processes characterized by independent but not stationary increments.

The standard definition of additive process follows.

Definition 0.0.2. *(Cont and Tankov (2003), Def.14.1, p.455)*
A cadlag stochastic process on $\mathbb{R}$ $\{X_t\}_{t\geq 0}$ *is an additive process if and only if:*

1. *For any choice of* $n \geq 1$ *and* $0 \leq t_0 < t_1 ... < t_n$*, the random variables* $X_{t_1} - X_{t_0}$*,* $X_{t_2} - X_{t_1}$*,...,*$X_{t_n} - X_{t_{n-1}}$ *are independent.*
2. X_0*=0.*
3. $\{X_t\}_{t\geq 0}$ *is stochastically continuous:*

$$\forall \epsilon > 0, \lim_{h\to 0} P\left[|X_{t+h} - X_t| > \epsilon\right] = 0\ .$$

Notice that Lévy processes are additive processes by definition: a Lévy process is an additive process that also has stationary increments, i.e. a Lévy process satisfies all conditions of **Definition 0.0.2** but it also satisfies condition (3) of **Definition 0.0.1**.

The following holds for every additive process: for every fixed time t, it is always possible to define a Lévy process that at time t has the same law as the additive process. Thanks to this

feature additive processes preserve several properties (both analytical and numerical) of Lévy processes.

The theory behind additive processes is well-established (Sato 1999a) but relatively few applications in quantitative finance are available. A first application of additive processes to option pricing is proposed by Carr *et al.* (2007), who apply Sato processes (Sato 1991) to derivative modeling (see also Eberlein and Madan 2009, for an application to structured products). Benth and Sgarra (2012) consider additive processes, which they call time-inhomogeneous Lévy processes, in the electricity market. In their paper, the electricity spot price is characterized by Ornstein-Uhlenbeck processes, driven by additive processes. More recently, Li *et al.* (2016) have introduced a larger class of additive processes. Their work focuses on additive subordination, which (they show) is a useful technique for constructing time inhomogeneous Markov processes with an analytically tractable characteristic function. This technique is a natural generalization of Lévy subordination.

As shown in figure 0.1 the equity implied volatility surface (e.g., S&P 500 surface) is not flat. Specifically, the short-time, or short-time-to-maturity, negative skew is proportionally inverse to the square root of the time-to-maturity. The empirical study of the term structure of the equity skew dates back to the seminal paper of Carr and Wu (2003): they find that the S&P 500 short-time skew is, on average, asymptotic to $-0.25/\sqrt{t}$. Fouque *et al.* (2004) arrive at a similar conclusion considering only options with short-time-to-maturity (i.e. up to three months). Finding a model with this characteristic has been the holy grail of short-time literature in recent years. We show that the ATS, a pure jump additive process, reproduces the market power scaling skew.

Most of the models used in options pricing admit no close formula for the implied volatility. For this reason, the literature on short-time implied volatility and skew of jump-diffusion processes is quite vast. Both the ATM behavior (see, e.g., Alòs *et al.* 2007, Roper 2009, Muhle-Karbe and Nutz 2011, Andersen and Lipton 2013, Figueroa-López *et al.* 2016) and the out-of-the-money (OTM) behavior of implied volatility (see, e.g., Tankov 2011, Figueroa-López and Forde 2012, Mijatović and Tankov 2016, Figueroa-López *et al.* 2018) are discussed in detail. For a Lévy process, the ATM implied volatility is determined uniquely by the diffusion term; it goes to zero as the time-to-maturity goes to zero if there is no diffusion term, i.e., for a pure-jump process. For this reason, pure jumps Lévy processes fail to reproduce the market short-time smile: the short-time implied volatility is strictly positive in all financial markets.

Muhle-Karbe and Nutz (2011) have discussed how, for a relatively broad class of additive models, the ATM behavior at small-time is the same as the corresponding Lévy. We analyze the ATM implied volatility and skew for a class of pure jump additive processes (the ATS) that, differently from the Lévy case, is consistent with the equity market smile.

We introduce a new (pure-jump) class of additive processes through their characteristic function. We call them additive normal tempered stable (ATS) processes. ATS processes (in general) cannot be built via a time-change of a Lévy process as in the additive subordination of Li *et al.* (2016) and are not Sato processes. However, there is a subclass of ATS obtained via additive subordination and a subclass of Sato ATS process. This class of additive process can "exactly" calibrate the term structure of observed implied volatility surfaces while maintaining the parsimony of LTS.

The ATS calibration is excellent, on average, two orders of magnitude better than the corresponding LTS in terms of mean squared error. We discuss in detail how the calibrated time-dependent parameters present an interesting and statistically relevant self-similar behavior compatible with a power-law scaling subcase of ATS. Specifically, among all allowed power laws, the power scaling parameters β -related to the variance of jumps- is close to one, while the power

scaling δ -related to the skew- is statistically consistent with minus one half. Moreover, we have verified that these results are robust over time. We study the behavior of the ATM implied volatility and skew for the ATS process, deriving an extension of the Hull and White (1987, Eq.(7), p.4) formula. Thanks to this pricing formula, on the one hand, we build some relevant bounds for ATM volatility; on the other hand, we obtain an expression for the skew via the implicit function theorem. We prove that only the scaling parameters estimated from market data ($\beta = 1$ and $\delta = -1/2$) admit a finite short-time implied volatility and a short-time skew proportionally inverse to the square root of the time-to-maturity.

The calibration results and the short-time property of the ATS are remarkable. It is easy to price European payoff via Fourier-based methods, but we need a Monte Carlo scheme for more complex derivatives. Monte Carlo simulation of pure jumps Lévy processes is not straightforward because, as in the case of the ATS, there is no closed formula for the cumulative distribution function (CDF) of the process increments.

Eberlein and Madan (2009) implemented a simulation method for a specific class of additive processes, Sato processes. The method they adapt to this class of additive processes, builds upon a well-known jump simulation technique developed for Lévy processes, that can be found in many excellent textbooks (see e.g., Cont and Tankov 2003, Asmussen and Glynn 2007). To the best of our knowledge, this is the unique Monte Carlo scheme developed for additive processes.

The jump simulation algorithms can be divided into two steps. First, one should truncate small jumps below a certain threshold, and second, simulate the finite number of independent jumps; finally, it is possible to apply the Asmussen and Rosiński (2001) Gaussian approximation to substitute small jumps with a Brownian component: we use this method as a benchmark technique to compare numerical results.
We introduce a fast Monte Carlo technique for additive processes, in general, and for the ATS, in particular. This scheme is based on the numerical inversion of the cumulative distribution function. To simulate an additive process, as in the Lévy case, is not straightforward because, in general, the CDF of process and of the process increments is not known explicitly.

Different methods have been proposed for sampling from a generic characteristic function (see e.g., the seminal paper of Bohman 1970).

In the financial literature, these techniques have been considered in the Lévy case, where it is possible to leverage on the stationary of increments (see e.g., Glasserman and Liu 2010, Chen *et al.* 2012, Ballotta and Kyriakou 2014). These techniques are stable and efficient: they employ different specifications of characteristic function numerical inversion to obtain an estimation of the CDF. We use the fast Fourier transform method for the numerical inversion as proposed by Lee (2004) and then applied to MC option pricing in the studies of Chen *et al.* (2012) and Ballotta and Kyriakou (2014). Relative to this literature, our main contribution lies in a detailed discussion of the three sources of error in the derivative price expectations and showing how to improve the two most significant ones.

We desire to calibrate our model via a "calibration cascade": first on the most liquid derivatives contracts and then on the less liquid ones. For this reason, we first focus and interest rates and forward prices (by far more liquid than options) and then on options.

The Overnight Index Swap (OIS) curve has emerged as a possible candidate for the risk-free curve for discounting after the great financial crisis. The OIS is a swap derived from the unsecured interbank overnight rate (OR), for example, the Euro Short Term Rate (STR) rate for Euros and the Effective Federal Fund Rate (EFFR) for US dollars. This OR is the best estimation of a risk-free rate, and it is the interest rate most commonly paid on collateral. Moreover, there are numerous points in favor of the OIS curve: it is a curve based on liquid swaps and bootstrap of

the discounting curve is as simple as the well-established 3M-Libor or 6M-Libor bootstrap (see, e.g., Ron 2000).

We focus on market makers who operate in a given exchange-traded derivative market. Probably they will use the OIS curve for discounting derivatives, and will add a spread that summarizes other risks or costs not included in the "risk-free" rate. We call this spread "cost-of-funding" because it is the additional implicit cost in operating in this specific market.

We can pose our research question from the perspective of the market maker: which cost-of-funding (if any) will I pay when operating in a liquid exchange-traded derivatives market? Answering this question can provide both operational and management insights. First, for their daily activity, the dealers should monitor an indicator on this spread: they need a discounting curve in line with other market participants; second, this spread has important consequences on the management of a financial firm. To understand the "market" cost-of-funding for each business unit is a piece of relevant information for management within a financial firm. We propose an elementary indicator that can monitor this funding cost in real-time and point out possible stress in funding liquidity. We use the market implicit interest rate (composed of the "risk-free" OIS rate plus the cost-of-funding) to discount option prices and to build forward prices in our dataset.

To test the robustness of the ATS performances on a comprehensive dataset -nine years of closing prices- we need forward prices and discount factors synchronized with option prices. In general, interest rates used in derivative pricing are not "risk-free" because contingent claim evaluation should depend on the risks of the investment and, in particular, on the funding risk and the risk of default of one of the two counterparties in the derivative contract. When dealing with exchange-traded derivatives, the situation should be more straightforward: the presence of a clearinghouse with margin calls allows neglecting the market participants' default risk. For this reason, we can use this market-implied discount factor (built via the funding indicator) to calibrate the ATS.

This new family of additive processes aims to solve two key open problems in the equity market literature.
First, it is a pure jump additive process that calibrates the term structure and skew of the volatility surface parsimoniously. On a comprehensive dataset, we will show that it significantly outperforms pure jumps Lévy processes (LTS) and pure jump additive processes (self-similar). The ATS parsimony is due to a power-law scaling characteristic that arises in calibrated parameters. Surprisingly, it holds both for short and long maturities (substantially different participants characterize option markets on short and long maturities). As we have already mentioned, market data are consistent with $\beta = 1$ and $\delta = -1/2$. By considering a power-law scaling ATS with fixed β and δ, after the volatility term structure is taken into account, the surface is calibrated with just two free parameters.
Second, differently from the other pure jump models introduced in the literature, the short-time implied volatility and skew of the ATS are consistent with empirical studies. We will prove that ATS has a finite and constant ATM implied volatility and a skew that is inversely proportional to the square root of the time-to-maturity if and only if $\beta = 1$ and $\delta = -1/2$. We can explain the goodness of the ATS calibration under this light. Our educated guess is that market data are consistent with $\beta = 1$ and $\delta = -1/2$ because these are the unique parameters such that the ATS implied volatility is coherent at short-time. Moreover, we emphasize that ATS is numerically tractable. The characteristic function is explicit and European options are priced with Fourier methods. This allows a fast calibration of the volatility surface. We will also introduce a fast simulation method for the ATS that is based on the independence of increments. Finally, we

spend a few words on the cost-of-funding indicator. The indicator was devised for calibrating a large dataset of equity volatility surfaces with synchronized interest rates and forward prices. However, we think that this methodology can provide interesting insights into different option markets. It is an elementary indicator that can monitor in real-time the funding cost and point out possible stress in funding liquidity. We will show that the S&P 500 market has, on average, a cost-of-funding of 33bps. while EURO STOXX 50, on average, has no cost-of-funding.

The thesis is divided into four chapters and a conclusion chapter; an introduction opens each chapter. The results in chapter 1 and part of the results in chapter 4 have already been published respectively, in Azzone and Baviera (2021a) and Azzone and Baviera (2021d). The structure of the chapters follows.

Chapter 1 introduces the ATS and analyzes the power-law scaling arising from market data. Section 1.2 introduces the model: we prove that there exists a new family of additive processes as the natural extension of the corresponding Lévy processes. Section 1.3, describes the dataset used in the calibration, the calibration results for ATS, LTS, and Sato processes. Furthermore, it introduces an interesting scaling property of the calibrated parameters. Section 1.4, shows that LTS and Sato processes fail to reproduce some stylized facts observed in market data, which are adequately described by ATS processes and presents a robustness analysis. Section 1.5 concludes. Moreover, appendix .1 contains the proofs for this chapter, appendix .2 describes the ATS parameter estimations, and appendix .3 proves that it is not possible to build the ATS via additive subordination.

Chapter 2 analyzes the ATS short-time asymptotics. Section 2.2 presents the ATS power scaling process and the extension of the Hull and White formula. Section 2.3 defines the implied volatility problem and analyzes the short-time ATM implied volatility $\hat{\sigma}_t$. Section 2.4 computes the short-time limit of the skew term $\hat{\xi}_t$. Section 2.5 presents the major result: the ATS process is consistent with the equity market if and only if $\beta = 1$ and $\delta = -1/2$. Section 2.6 concludes. In the appendices, we report some technical lemmas that we use in this chapter: on basic properties in appendix .4 and on short-time limits in appendix .5.

Chapter 3 introduces a fast Monte Carlo scheme for additive processes. Section 3.2 overviews the method and recalls both Lewis (2001) formula for CDF and the error source in the numerical approximation: we discuss the optimal selection of the integration path. Section 3.3 describes the proposed simulation method and presents the other main error source in MC option pricing: the interpolation method in numerical inversion. We also discuss how to generalize the GA method for additives in an efficient way. Section 3.4 presents numerical results for a large class of pure-jump additive processes in the case of both European options (where analytical pricing methods are available), and some discretely monitoring path-dependent options. Section 3.5 concludes. The proofs of this chapter are in appendix .6.

Chapter 4 is dedicated to the market implicit interest rate and the results of the ATS calibration on a nine-year dataset. Section 4.2 shows the methodology to find the implicit interest rates using only option prices and describes the dataset. Section 4.3 infers the S&P 500 and the EURO STOXX 50 implicit discount factor and the corresponding cost-of-funding. Section 4.4 reports the calibration results: on the nine-year dataset and on both implied volatility surfaces ATS has significantly better performances than alternatives. Section 4.5 concludes. An analysis of the ATS performances on a commodity volatility surface is available in appendix .7.

Finally, chapter 5 concludes by summarizing the key finding and contributions of the thesis.

Chapter 1

Additive normal tempered stable processes for equity derivatives and power law scaling

In this chapter, we introduce the additive normal tempered stable process (ATS): a simple additive process for equity index derivatives. The model generalizes Lévy Normal Tempered Stable processes (e.g., NIG and VG) with time-dependent parameters. It accurately fits the equity index volatility surfaces in the whole time range of quoted instruments, including options with small time-horizon (days) and long time-horizon (years).

We introduce the model via its characteristic function. This allows using classical Fourier pricing techniques. We discuss the calibration issues in detail and we show that, in terms of mean squared error, calibration is on average two orders of magnitude better than both Lévy and Sato processes alternatives. We show that even if the model loses the classical stationarity property of Lévy processes, it presents interesting scaling properties for the calibrated parameters.

The results in this chapter have already been published in Azzone and Baviera (2021a).

1.1 Introduction

As already discussed in the thesis introduction, Lévy processes are a powerful modeling solution that provides parsimonious models consistent with option prices and with underlying asset prices. LTS are pure jump[1] processes, obtained via the well-established Lévy subordination technique (see, e.g., Cont and Tankov 2003, Schoutens 2003). Specifically, the LTS are characterized by three parameters: σ, which controls the average level of the volatility surface; k, which is related to the convexity of the implied volatility surface; and η, which is linked to the volatility *skew*.

Unfortunately, the recent literature has shown that these models do not reproduce the implied volatilities that are observed in the market data at different time horizons with sufficient precision (see, e.g., Cont and Tankov 2003, Ch.14). Lévy normal tempered stable processes are pure jump models with independent and stationary increments. The key question is as follows: is it reasonable to consider stationary increments when modeling implied volatility? Jump stationarity is a feature that significantly simplifies the model's characteristics but it is rather difficult to justify *a priori* from a financial point of view. For example, a *market maker* in the option market does not consider the consequences of a jump to be equivalent on options with different maturities. He cares about the amount of trading in the underlying required to replicate the

[1]The relevance of pure jump dynamics in the equity and commodity asset classes has been discussed in the recent literature, see, e.g., Ornthanalai (2014), Li and Linetsky (2014), Ballotta and Rayée (2018).

option after a jump arrival. The impact of such a jump on the hedging policy is inhomogeneous with option maturity.[2] Hence, *a priori*, it is not probable that a stationary model can adequately describe implied volatilities.

Additive processes have been proposed to overcome this problem. Additive processes are an extension of Lévy processes that consider independent but not stationary increments. Given an additive process, for every fixed time t, it is always possible to define a Lévy process that at time t has the same law as the additive process. This feature allows us to maintain several properties (both analytical and numerical) of the Lévy processes. In the introduction of the thesis we have already mentioned the few additive process applications in quantitative finance.

In this chapter, introduce a new class of (pure-jump) additive processes through their characteristic function which are named additive normal tempered stable (ATS) processes. ATS processes (in general) cannot be obtained via a time-change as in the additive subordination of Li *et al.* (2016) and are not Sato processes. There is a subclass of ATS obtained via additive subordination and a subclass of Sato ATS process. In appendix .3 we prove that it is not possible to introduce the ATS by additive subordination. The main advantage of this new class of models is the possibility to "exactly" calibrate the term structure of observed implied volatility surfaces, while maintaining the parsimony of LTS.

We provide a calibration example of the ATS on the S&P 500 and EURO STOXX 50 implied volatility surfaces of the 30^{th} of May 2013. The ATS calibration is on average two orders of magnitude better than the corresponding LTS in terms of mean squared error. We show that the calibrated time-dependent parameters present an interesting and statistically relevant self-similar behavior compatible with a power-law scaling subcase of ATS. Moreover, we have verified that these results are robust over time.

The main contributions of this chapter are threefold. First, we introduce a new broad family of additive processes, which we call additive normal tempered stable (ATS) processes.
Second, we calibrate the ATS processes on S&P 500 and EURO STOXX 50 volatility surfaces. We show that ATS have better calibration features (in terms of both the Mean Squared Error and the Mean Absolute Percentage Error) than LTS and Sato processes.
Finally, we consider a re-scaled ATS process via a time-change based on the implied volatility term structure. We show that the calibrated parameters exhibit a self-similar behavior w.r.t. the new time. The statistical relevance of these scaling properties is verified.

The rest of the chapter is organized as follows. In section 1.2, we introduce the model: we prove that there exists a new family of additive processes as the natural extension of the corresponding Lévy processes. In section 1.3, we describe the dataset used in the calibration, the calibration results for ATS, LTS, and Sato processes and an interesting scaling property of the calibrated parameters. In section 1.4, we show that LTS and Sato processes fail to reproduce some stylized facts observed in market data, which are adequately described by ATS processes and we present a robustness analysis. Finally, section 1.5 concludes.

1.2 The model

In this section, we introduce the ATS process, as a natural extension of the LTS process, that, on the one hand, maintains the increments' independence as in the corresponding Lévy process, and, on the other hand, allows for time-inhomogeneous parameters. First, we prove a sufficient

[2]Gamma is the Greek measure that quantifies the amount of this hedging and, generally, it decreases with time-to-maturity.

condition for the existence of ATS processes (**Theorem 1.2.1**) and the martingale property for the corresponding forward process (**Proposition 1.2.2**). Then, we introduce the power-law scaling ATS as a subcase of a generic ATS (**Theorem 1.2.3**): we show in the next section that this model accurately describes the implied volatility surface. Finally, we prove a key model feature: the model allows to reproduce a generic volatility term structure (**Proposition 1.2.4**).

Lévy normal tempered stable processes (LTS) are commonly used in the financial industry for derivative pricing. According to this modeling approach, the forward with expiry T is an exponential Lévy; i.e.

$$F_t(T) := F_0(T)\ \exp(f_t)\ , \tag{1.2.1}$$

with f_t a LTS

$$f_t = -\left(\frac{1}{2}+\eta\right)\ \sigma^2\ Z_t + \sigma\ W_{Z_t} + \varphi\, t \qquad \forall t \in [0,T]\ ,$$

where η, σ are two real parameters ($\eta \in \mathbb{R}, \sigma \in \mathbb{R}^+$), while φ is obtained by imposing the martingale condition on $F_t(T)$.[3] W_t is a Brownian motion and Z_t is a Lévy tempered stable subordinator independent from the Brownian motion with variance per unit of time k. Examples of LTS subordinators are the Inverse Gaussian process for NIG or the Gamma process for VG.

It is possible to write the characteristic function of f_t as

$$\mathbb{E}\left[e^{iuf_t}\right] = \mathcal{L}_t\left(iu\left(\frac{1}{2}+\eta\right)\sigma^2 + \frac{u^2\sigma^2}{2};\ k,\ \alpha\right) e^{iu\varphi\, t}\ , \tag{1.2.2}$$

where $\alpha \in [0,1)$ is the LTS *index of stability* and $\mathcal{L}_t$ is the Laplace transform of Z_t

$$\ln \mathcal{L}_t\left(u;\ k,\ \alpha\right) := \begin{cases} \dfrac{t}{k}\dfrac{1-\alpha}{\alpha}\left\{1-\left(1+\dfrac{u\ k}{1-\alpha}\right)^{\alpha}\right\} & \text{if } \ 0<\alpha<1 \\ -\dfrac{t}{k}\ln\left(1+u\ k\right) & \text{if } \ \alpha = 0 \end{cases} . \tag{1.2.3}$$

This theory is well known and can be found in many excellent textbooks (see, e.g., Cont and Tankov 2003, Schoutens 2003).

As already discussed in the Introduction, LTS processes do not properly describe short and long maturities at the same time, while they allow an excellent calibration for a fixed maturity. For this reason, we would like to select a process that allows independent but non-stationary increments: i.e. an additive process. The simplest way to obtain this modeling feature is to consider an additive process with a characteristic function of the same form of (1.2.2) but with time-dependent parameters

$$\mathbb{E}\left[e^{iuf_t}\right] = \mathcal{L}_t\left(iu\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 + \frac{u^2\sigma_t^2}{2};\ k_t,\ \alpha\right) e^{iu\varphi_t t}\ , \tag{1.2.4}$$

where σ_t, k_t are continuous on $[0,\infty)$ and η_t, φ_t are continuous on $(0,\infty)$ with $\sigma_t > 0$, $k_t \geq 0$ and $\varphi_t\, t$ goes to zero as t goes to zero. $\alpha \in [0,1)$ as in the LTS case.

In **Theorem 1.2.1**, we prove that this process exists if some conditions on σ_t, η_t and k_t are satisfied.

[3] A parametrization scheme of the drift in terms of η can be suitable for some applications: η controls the volatility *skew*. In particular, it can be proven that for $\eta = 0$ the smile is symmetric, i.e. the implied volatility *skew* is zero (see, e.g., Baviera 2007, Prop. p.21).

By **Definition 0.0.2**, an additive process is a càdlàg stochastic process on $\mathbb{R}$ $\{X_t\}_{t\geq 0}$, with $X_0 = 0$ a.s. and characterized by independent increments and stochastic continuity. It can be proved that the distribution of an additive process at time t is infinitely divisible. (A_t, ν_t, γ_t) is the generating triplet that characterizes the additive process $\{X_t\}_{t\geq 0}$. A_t, ν_t and γ_t are called respectively the diffusion term, the Lévy measure and the drift term (see Sato 1999a, pp.38-39).[4]

Sato (1999a, Th.9.8, p.52) proves a powerful link between a system of infinitely divisible probability distributions and the existence of an additive process. In particular, Sato requires two main classes of conditions on the generating triplet: i) some conditions of monotonicity, necessary to avoid meaningless negative diffusion term or negative Lévy measure for the process increments and ii) some continuity conditions, in order to obtain the stochastic continuity.

We use Sato (1999a, Th.9.8, p.52) to prove the main theoretical results in this chapter: there exists a family of additive processes with characteristic function (1.2.4).

Theorem 1.2.1. Sufficient conditions for existence of ATS
There exists an additive process $\{f_t\}_{t\geq 0}$ with the characteristic function (1.2.4) if the following two conditions hold.

1. *$g_1(t)$, $g_2(t)$, and $g_3(t)$ are non decreasing, where*

$$
\begin{aligned}
g_1(t) &:= (1/2 + \eta_t) - \sqrt{(1/2 + \eta_t)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)} \\
g_2(t) &:= -(1/2 + \eta_t) - \sqrt{(1/2 + \eta_t)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)} \\
g_3(t) &:= \frac{t^{1/\alpha}\sigma_t^2}{k_t^{(1-\alpha)/\alpha}} \sqrt{(1/2 + \eta_t)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)} \ ;
\end{aligned}
$$

2. *Both $t\,\sigma_t^2\,\eta_t$ and $t\,\sigma_t^{2\alpha}\,\eta_t^{\alpha}\,/k_t^{1-\alpha}$ go to zero as t goes to zero.*

Proof. See appendix .1 □

Let us emphasize that the conditions of **Theorem 1.2.1** are quite general. Condition 1 ensures the monotonicity of the ATS jump measure while condition 2 that jump measure and drift term go to zero at short time. In the market, we observe only a limited number of maturities, and thus, there is a large set of functions of time that reproduce market data and satisfy the conditions of the theorem. Furthermore, we prove in **Theorem 1.2.3**, that the conditions of **Theorem 1.2.1** are satisfied by a simple sub-case of ATS with power-law scaling η_t and k_t and constant sigma. Finally, we also prove in **Proposition 1.2.4** that ATS models allow a generic volatility term structure σ_t.

In a similar way to the LTS case, it is possible to consider a forward price $F_t(T)$ (1.2.1) as the exponential of the ATS process $\{f_t\}_{t\geq 0}$ with the characteristic function (1.2.4). The deterministic function of time φ_t can be chosen s.t. the process $F_t(T)$ satisfies the martingale property, as shown in the next proposition.

Proposition 1.2.2. Martingale property
The forward $\{F_t(T)\}_{t\geq 0}$, modeled via an exponential additive process characterized by an ATS process $\{f_t\}_{t\geq 0}$ is a martingale if and only if

$$
\varphi_t\, t = -\ln \mathcal{L}_t\left(\sigma_t^2\eta_t;\ k_t,\ \alpha\right) \quad , \tag{1.2.5}
$$

where $\mathcal{L}_t$ is the Laplace transform in (1.2.3).

[4] In this thesis, the notation follows closely the one in Sato (1999a).

Proof. See appendix .1 □

We introduce a sub-case of ATS, determined by self-similar functions of time. In subsection 1.3.3, we show that this family of processes accurately describes market implied volatility surfaces. Power-law scaling functions of time allow us to rewrite **Theorem 1.2.1** conditions as simple inequalities on the scaling parameters.

Theorem 1.2.3. Power-law scaling ATS
There exists an ATS with

$$k_t = \bar{k}\, t^{\beta}, \qquad \eta_t = \bar{\eta}\, t^{\delta}, \qquad \sigma_t = \bar{\sigma}\ ,$$

where $\alpha \in [0, 1)$, $\bar{\sigma}, \bar{k}, \bar{\eta} \in \mathbb{R}^+$, *and* $\beta, \delta \in \mathbb{R}$ *that satisfy the following conditions:*

1. $0 \leq \beta \leq \dfrac{1}{1 - \alpha/2}$;
2. $-\min\left(\beta, \dfrac{1 - \beta\,(1-\alpha)}{\alpha}\right) < \delta \leq 0$;

where the second condition reduces to $-\beta < \delta \leq 0$ *for* $\alpha = 0$.

Proof. See appendix .1 □

The admissible region for ATS power law scaling parameters is plotted in Figure 1.1.

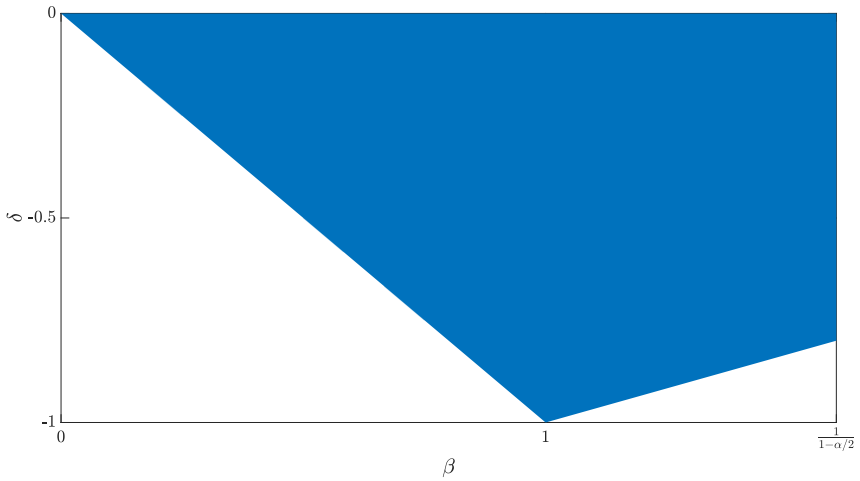

Figure 1.1: Admissible region for ATS power law scaling parameters

It is interesting to observe that the LTS case falls in the subcase described by this theorem. This corresponds to the case with both k_t and η_t time independent; that is, β and δ equal to zero.

The following result allows us to obtain a new additive process from a known one with a deterministic time change.

Proposition 1.2.4. Deterministic time change of additive process
Given an additive process $\{X_t\}_{t\geq 0}$ *and a real continuous increasing function of time* r_t *s.t.* $r_0 = 0$, *then* $\{X_{r_t}\}_{t\geq 0}$ *is an additive process.*

Proof. See appendix .1 □

Thanks to **Proposition 1.2.4**, it is possible to extend the ATS power-law scaling sub-case to a case with time-dependent σ_t. Indeed, if $\sigma_t^2 t$ is non decreasing, we can use it to time-change a power-law scaling ATS without losing the property of independent increments: being able to reproduce the volatility term structure is an important feature from a practitioner perspective.

In the next section, we show that the ATS model introduced in this section accurately describes volatility surfaces observed in the equity derivative market.

1.3 Model calibration and power law scaling

In this section, we show that the ATS processes achieve excellent calibration results on the S&P 500 and EURO STOXX 50 volatility surfaces; moreover, we show that power-law scaling parameters are observed in market data.

First, after having described the dataset, we illustrate the model calibration procedure and compare the performance of ATS processes with some benchmarks (LTS processes and Sato processes in Carr *et al.* (2007)). Then, we outline some statistical evidence that the market-implied volatility surface is compatible with a power-law scaling of ATS parameters.

1.3.1 Dataset

We analyze all quoted S&P 500 and EURO STOXX 50 option prices observed at 11:00 am NT on the 30^{th} of May 2013. The dataset is composed of real market quotes (no smoothing or interpolation[5]). Let us recall that the options on these two indices are the most liquid options in the equity market at the world level. For both indices, options expire on the third Friday of March, June, September, and December in the front year and June and December in the next year. In the EURO STOXX 50 case also December contracts for the following three years are available. The dataset includes the risk-free interest rate curves bootstrapped from (USD and EUR) OIS curves. Financial data are provided by Bloomberg. The dataset contains all bid/ask prices for both call and put. The strikes are in a regular grid for each available maturity. We exclude options that do not satisfy two simple liquidity thresholds. We discard options whose price is less than 10% the minimum difference in the grid of market strikes (the so-called *penny options*) and options with bid-ask over bid bigger than 60%. The last condition filters out strikes for which either a bid or an ask price is missing.

We use the synthetic forward, as forward price, because this allows a perfect synchronization with option prices and, for several maturities, it identifies the most liquid forward in the market. The synthetic forward price is obtained for every maturity, from very liquid options as the (algebraic) mean of the lowest forward ask and the highest forward bid.

We implement a simple iterated algorithm that identifies the synthetic forward price at a given maturity T: let us briefly describe it. We start selecting the call and put options with strike price nearest to the spot price for the shortest maturity or to the previous maturity forward price for the next maturities. We compute forward bid, ask, and mid prices for that strike price. We consider the options with the nearest superior strike. If the forward mid-price computed previously falls within the new bid-ask interval, then the updated forward bid is the highest value among the two forward bids, while the updated ask is the lowest value among the two ask prices. The updated forward mid-price corresponds to the mean of the updated bid and

[5]The practice of some market-data providers is to smooth or interpolate market data with proprietary tools. By using real market data we eliminate the risk of fitting our model on the provider's model.

ask. Then, we consider the nearest inferior strike and iterate the same procedure comparing the updated forward mid-price with the new bid/ask prices relative to this new strike. This procedure is iterated with the next superior strike and then with the next inferior strike, and so on for all the options present at that maturity T.

In Figure 1.2, we show, for a given underlying and a given maturity, the values considered in the forward price construction and the value selected by the procedure.

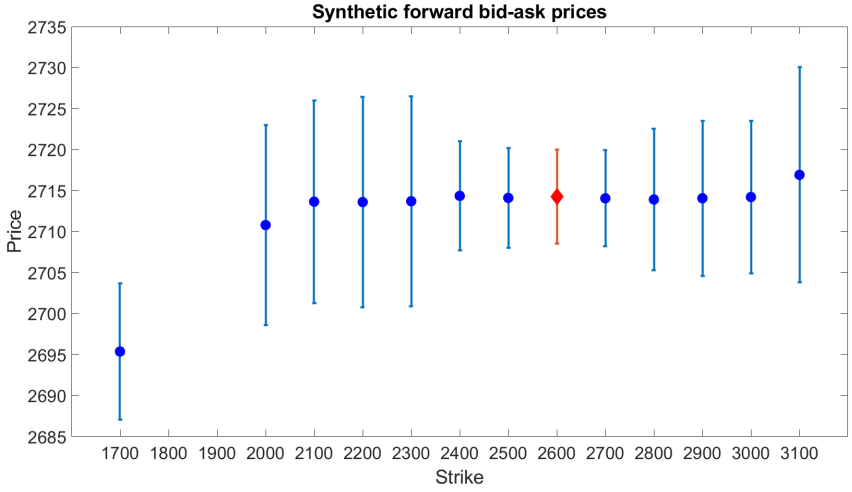

Figure 1.2: EURO STOXX 50 synthetic forward prices on the 30^{th} of May 2013 at 11 am NT for the JUN14 maturity: bid, ask, and mid forward prices. Only prices not discarded by the two liquidity criteria are shown in the figure. According to the algorithm described in the text also the price related to the strike 1700 is discarded from the forward price computation. We show in red the corresponding forward bid-ask prices and with a diamond, the selected forward price $F_0(T)$ relative to this expiry.

In Figure 1.3, we plot the bid, ask, and mid synthetic forward prices for the different maturities available for the S&P 500 and the EURO STOXX 50.

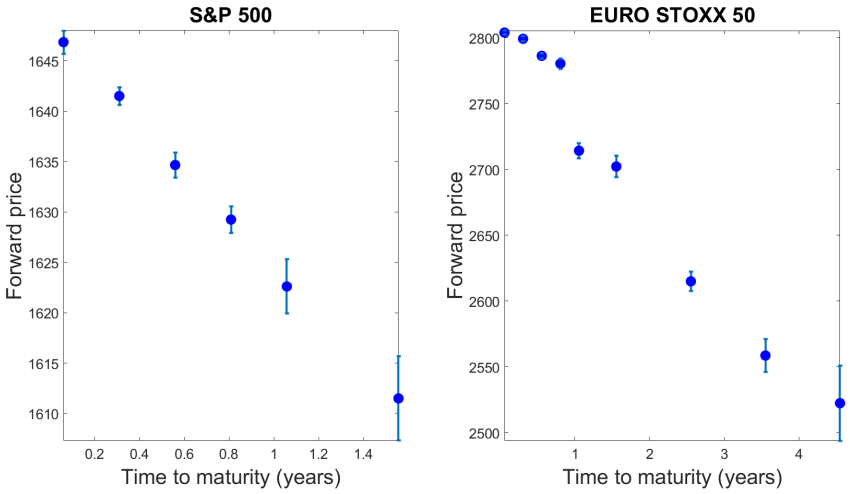

Figure 1.3: Term structure of the synthetic forward prices (dots) on the 30^{th} of May 2013: we report also the observed bid and ask prices (respectively down and up bars) for every maturity. On the left hand side, we plot the S&P 500 index case and, on the right hand side, the EURO STOXX 50 index case.

1.3.2 Calibration

We calibrate the ATS following the procedure discussed by Cont and Tankov (2003, Ch.14, pp.464-465). We cut the volatility surface into slices, each one containing options with the same maturity, and calibrate each slice separately. Hereinafter, we focus on $\alpha = 1/2$ (NIG) and $\alpha = 0$ (VG), which are the two (ATS and Sato) generalizations of the two most frequently used LTS processes. For every fixed maturity T, it is possible to define a new Lévy normal tempered stable process such that, at time T, its marginal distribution is equal to the marginal distribution of an ATS. A different Lévy NIG and VG is calibrated for every different maturity and the three time-dependent parameters k_T, η_T, σ_T are obtained. The calibration is performed imposing the conditions of monotonicity of **Theorem 1.2.1**.

Beneath the ATS processes, we consider the calibration of the standard Lévy processes and of the (four parameters) Sato processes proposed by Carr *et al.* (2007).[6] We remind that the latter are additive and self-similar processes (see, e.g., Sato 1991). Call option prices, with strike K and maturity T, are computed using the Lewis (2001) formula

$$C_T(x) = B_T \; F_0(T) \left\{ 1 - e^{x/2} \int_{-\infty}^{\infty} \frac{dz}{2\pi} e^{iz\,x} \phi^c \left(-z - \frac{i}{2} \right) \frac{1}{z^2 + \frac{1}{4}} \right\} \; , \tag{1.3.1}$$

where $\phi^c(u)$ is the characteristic function of f_T, $x := \ln K/F_0(T)$ is the *moneyness*, and B_T is the discount factor between value date and T.

The calibration is performed by minimizing the Euclidean distance between model and market prices. The simplex method is used to calibrate every maturity of the ATS process. For Lévy processes and Sato processes, because standard routines for global minimum algorithms are not satisfactory[7], we consider a differential evolution algorithm together with a multi-start simplex method.

[6] We underline that, in both cases (LTS and Sato), model parameters are obtained through a global calibration of the whole volatility surface.

[7] For these processes, we observe multiple local minimums.

The calibration performance is reported in Table 1.1 in terms of Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE).[8] It is possible to observe that Sato processes slightly improve Lévy performance, as reported in the literature (see, e.g Carr *et al.* 2007), while the ATS processes improvement is, on average, above two orders of magnitude. Although we present the results for VG and NIG, similar results can be obtained for all ATS processes with $\alpha \in [0, 1)$. The worst results are observed in the VG case.

		MSE			MAPE		
Index	Model	Lévy	Sato	ATS	Lévy	Sato	ATS
S&P 500	NIG	4.56	1.92	**0.02**	3.13%	1.47%	**0.23**%
S&P 500	VG	8.49	2.20	**0.24**	4.31%	1.62%	**0.79**%
Euro Stoxx 50	NIG	22.15	9.87	**0.10**	1.75%	0.75%	**0.09**%
Euro Stoxx 50	VG	55.81	9.22	**0.35**	2.85%	0.73%	**0.21**%

Table 1.1: Calibration performance for the S&P 500 and EURO STOXX 50 in terms of MSE and MAPE. In the NIG ($\alpha = 1/2$) and VG ($\alpha = 0$) cases, we consider the standard Lévy process, the Sato process, and the corresponding ATS process. Sato processes perform better than Lévy processes but ATS improvement is far more significant: two orders of magnitude for MSE and one order of magnitude for MAPE.

Figure 1.4 shows the differences of MSE w.r.t. the different times-to-maturity for S&P 500 volatility surface calibrated with a NIG process. Sato and Lévy LTS have a MSE of the same order of magnitude, while the improvement of ATS is of two orders of magnitude and particularly significant at short-time. The short time improvement in implied volatility calibration is particularly evident, as shown in Figure 1.5.

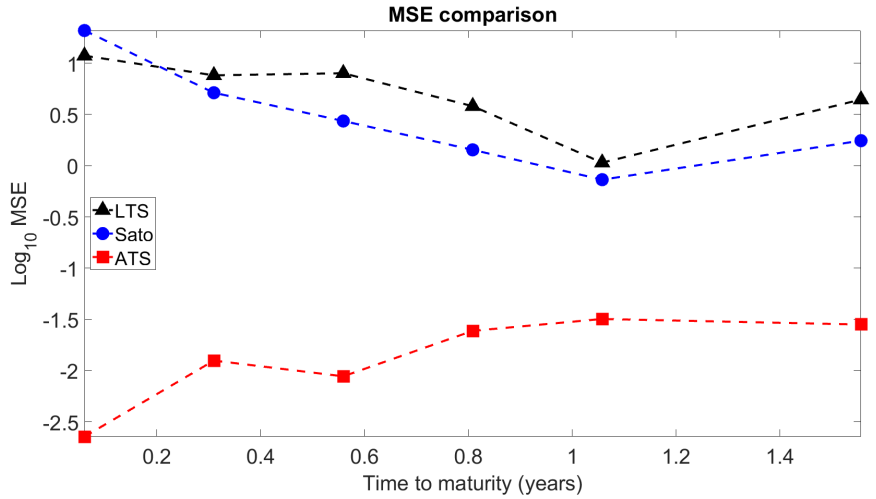

Figure 1.4: MSE w.r.t. the different times-to-maturity (in years) for S&P 500 volatility surface calibrated with a NIG process. Sato (circles) and Lévy (triangles) have a MSE of the same order of magnitude, while the improvement of ATS (squares) is of two orders of magnitude and particularly significant at short-time.

In Figure 1.5, we plot the market implied volatility and the volatility replicated via ATS, LTS, and Sato processes at the 22 days (on the left) and 9 months and 21 days (on the right)

[8] Calibrated model parameters are available upon request.

maturities. We observe that the ATS implied volatility is the closest to the market implied volatility in any case and it significantly improves both LTS and Sato processes, particularly for small maturities. Similar results hold for all other ATS.

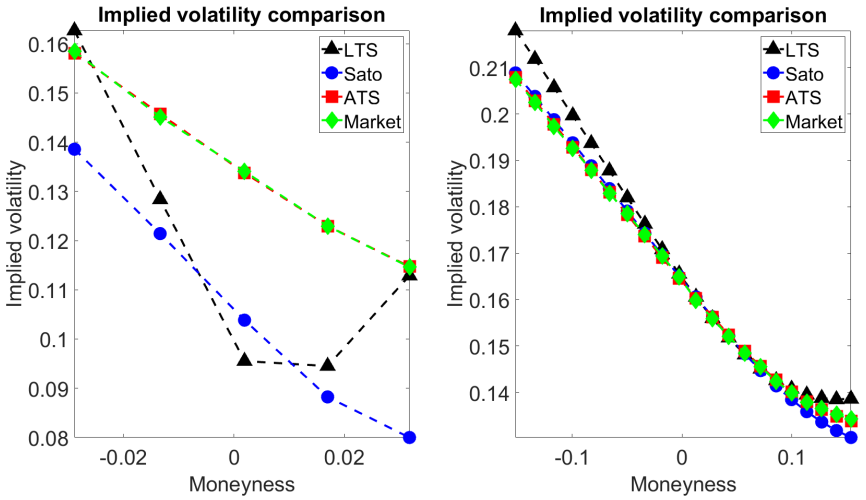

Figure 1.5: Implied volatility smile for S&P 500 for a given time-to-maturity: 22 days (on the left) and 9 months and 21 days (on the right). The NIG ATS process, Sato process, and LTS process implied volatilities are plotted together with the market-implied volatility. ATS reproduces the smile significantly better than the alternatives, the improvement is particularly evident for small maturities.

In Figure 1.6, we plot the market and the ATS implied volatility *skew* for EURO STOXX 50 w.r.t. the times-to-maturity. We observe that the calibrated ATS replicates accurately the market implied volatility *skew*. In chapter, 2, we show that ATS can replicate market skew exactly in the short time.

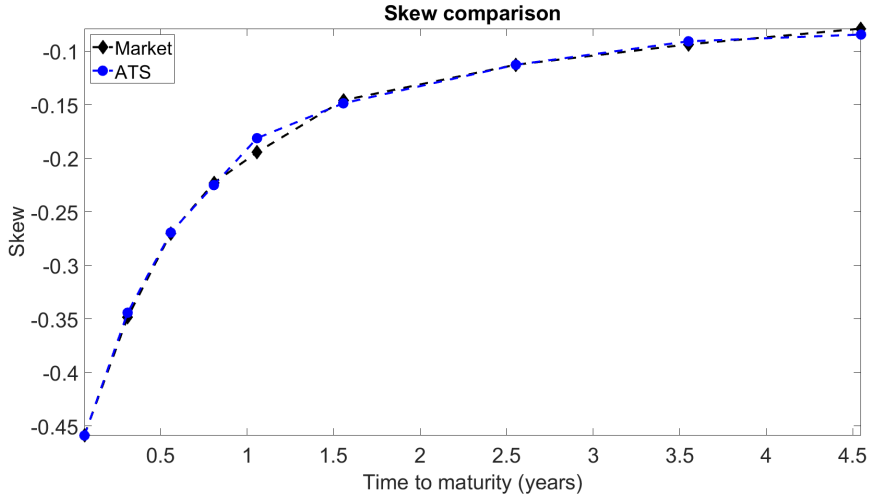

Figure 1.6: The market and the NIG ATS implied volatility *skew* for EURO STOXX 50 w.r.t. the times-to-maturity. ATS replicates the market implied volatility *skew* behavior.

The calibration results of ATS are startling. In particular, we have a model that reproduces

"exactly" the volatility term structure observed in the market.

It is useful to stop and comment. The ATS model allows us to calibrate the surface slice-by-slice; we have only to impose the monotonicity conditions of **Theorem 1.2.1**. With the slice-by-slice approach, we use 3 parameters for every expiry (e.g., 18 parameters in the S&P case). We have observed that ATS outperforms Lévy and Sato processes, the benchmark pure-jump processes in the literature. It could seem unfair to compare the calibration results of ATS with a Lévy (3 parameters) and a Sato (4 parameters).

In the next subsection, we show that this family of additive processes combines parsimony with the desired property of a perfect fit of the volatility term structure: we show that, once the term structure has been taken into account, only 2 free parameters allow a detailed calibration of the whole volatility surface.

1.3.3 Scaling properties

In this subsection, we show that power-law scaling parameters are observed in market data. This stylized fact is extremely relevant: we observe statistical evidence that the market-implied volatility surface is compatible with a power-law scaling ATS of **Theorem 1.2.3**.

We introduce a new ATS process, w.r.t. the time $\theta := T\sigma_T^2$.
We define $\hat{k}_\theta := k_T\sigma_T^2$ and $\hat{\eta}_\theta := \eta_T$. We call $\hat{f}_\theta$ the ATS with parameters $\hat{k}_\theta$, $\hat{\eta}_\theta$ and $\sigma_\theta := 1$. Notice that for every maturity $\hat{f}_\theta$ has the same characteristic function of the calibrated ATS f_T. We analyze the re-scaled parameters, in both S&P 500 and EURO STOXX 50 cases. We observe a self-similar behavior of $\hat{k}_\theta$ and $\hat{\eta}_\theta$; that is,

$$\begin{cases} \hat{k}_\theta &= \bar{k}\theta^\beta \\ \hat{\eta}_\theta &= \bar{\eta}\theta^\delta \end{cases}, \tag{1.3.2}$$

where $\bar{k}$, $\bar{\eta}$ are positive constants and β, δ are real constant parameters. To investigate this behavior and to infer the value of the scaling parameters we consider equations (1.3.2) in the log-log scale.

In Figures 1.7 and 1.8 we plot the weighted regression lines and the observed time-dependent parameters $\ln\hat{k}_\theta$ and $\ln\hat{\eta}_\theta$ with their confidence intervals for S&P 500 and EURO STOXX 50. The confidence intervals are two times the standard deviations of $\ln\hat{k}_\theta$, of $\ln\hat{\eta}_\theta$ and of $\ln\theta$. In appendix .2, we discuss the estimation of the standard deviations via a confidence interval propagation technique and the selection of the weights.

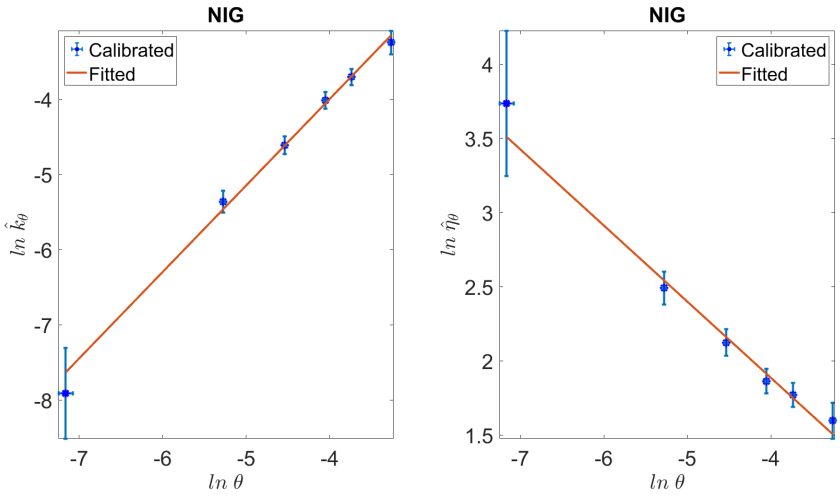

Figure 1.7: Weighted regression line and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ w.r.t. $\ln \theta$ for the NIG calibrated model for S&P 500. We plot a confidence interval equal to two times the corresponding standard deviation. Notice that confidence intervals on $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ are one order of magnitude wider than confidence intervals on $\ln \theta$. The scalings of $\hat{k}_\theta$ and $\hat{\eta}_\theta$ in (1.3.2) are statistically consistent with $\beta = 1$ and $\delta = -1/2$. The values of θ correspond to times-to-maturity that go from 22 days to two years.

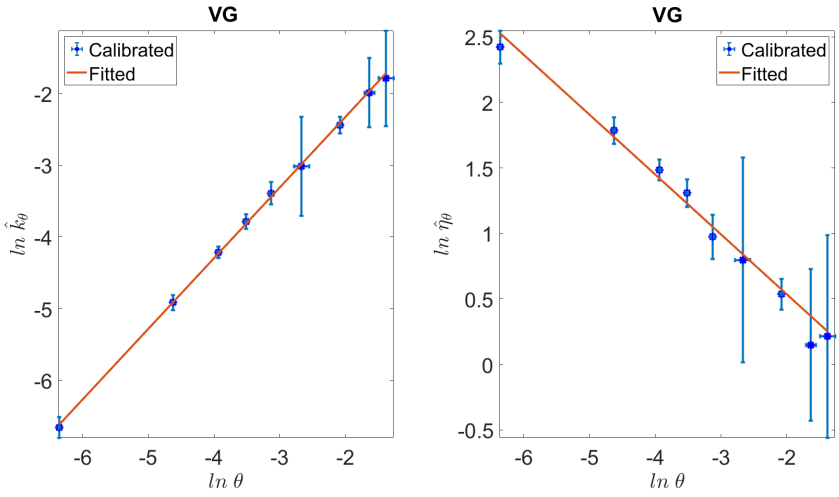

Figure 1.8: Weighted regression line and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ w.r.t. $\ln \theta$ for the VG model calibrated on EURO STOXX 50. We plot a confidence interval equal to two times the corresponding standard deviation. Notice that confidence intervals on $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ are one order of magnitude wider than confidence intervals on $\ln \theta$. The scalings of $\hat{k}_\theta$ and $\hat{\eta}_\theta$ in (1.3.2) are statistically consistent with $\beta = 1$ and $\delta = -1/2$. The values of θ correspond to times-to-maturity that go from 22 days to five years.

We have observed what seems to be a stylized fact of this model class: both $\hat{\eta}_\theta$ and $\hat{k}_\theta$ scale with power-law. The same scaling laws are observed both for short time-horizon (days) and long

time-horizon (few years) options. The fitted regression lines provide us with an estimation of β and δ. Moreover, let us emphasize that the scaling parameters appear qualitatively compatible to $\beta = 1$ and $\delta = -\frac{1}{2}$ in all observed cases.

We can build a test on the regression coefficient to determine whether there is statistical evidence that our hypotheses are consistent with market data. The estimated scaling parameters together with the p-value of these statistical tests are reported in Table 1.2.

Surface	Model	Parameter	Parameter's Value	p-value
S&P 500	NIG	β	1.10	0.228
S&P 500	NIG	δ	-0.47	0.705
S&P 500	VG	β	1.01	0.758
S&P 500	VG	δ	-0.43	0.057
EURO STOXX 50	NIG	β	1.02	0.816
EURO STOXX 50	NIG	δ	-0.44	0.472
EURO STOXX 50	VG	β	0.99	0.690
EURO STOXX 50	VG	δ	-0.45	0.195

Table 1.2: Scaling parameters calibrated from S&P 500 and EURO STOXX 50 volatility surfaces for NIG ($\alpha = 1/2$) and VG ($\alpha = 0$). Parameter estimates are provided together with the p-values of the statistical tests that verify whether it is possible to accept the null hypotheses $\beta = 1$ and $\delta = -\frac{1}{2}$.

In all cases, we accept the null hypotheses ($\beta = 1$ and $\delta = -\frac{1}{2}$) with a 5% confidence level. Notice that all p-values, except the S&P 500 VG δ, are above 19%. It is remarkable to observe that $\beta = 1$ and $\delta = -\frac{1}{2}$ are the scaling parameters that, as we prove in chapter 2 are consistent with market implied volatility and skew. That is a constant and finite short-time ATM volatility and a skew that is proportional to the inverse of the square root of the time-to-maturity.

In Table 1.3 we report an estimation of the parameter $\bar{k}$ and $\bar{\eta}$.

Surface	Model	Parameter	Parameter's Value	p-value
S&P 500	NIG	$\bar{k}$	1.50	0.022
S&P 500	NIG	$\bar{\eta}$	0.98	0.015
S&P 500	VG	$\bar{k}$	1.01	0.001
S&P 500	VG	$\bar{\eta}$	0.91	0.000
EURO STOXX 50	NIG	$\bar{k}$	0.68	0.023
EURO STOXX 50	NIG	$\bar{\eta}$	1.21	0.021
EURO STOXX 50	VG	$\bar{k}$	0.98	0.000
EURO STOXX 50	VG	$\bar{\eta}$	0.72	0.000

Table 1.3: $\bar{k}$ and $\bar{\eta}$ calibrated from S&P 500 and EURO STOXX 50 volatility surfaces. Parameter estimates are provided together with the p-values of the statistical tests that verify whether it is possible to accept the null hypothesis $\bar{k} = 0$ and $\bar{\eta} = 0$.

We have statistical evidence that in all cases $\bar{k}$ and $\bar{\eta}$ are positive (we reject the null hypotheses of $\bar{k} = 0$ and $\bar{\eta} = 0$ with a 5% confidence level). From these results and from Figure 1.6 it is possible to infer a connection between a positive $\bar{\eta}$ and the observed negative *skew*.

It is interesting to observe that these estimated parameters satisfy the inequalities of **Theorem 1.2.3** for the existence of a power-law scaling ATS $\hat{f}_\theta$. Moreover, the re-scaled process is additive w.r.t. the "real" time T. This fact is a consequence of the properties of the volatility term

structure σ_T (it is always observed on real data that $\sigma_T^2 T$ is non-decreasing) and of **Proposition 1.2.4**. This proposition states that if $\left\{\hat{f}_\theta\right\}_{\theta\geq 0}$ is an additive process then $\left\{\hat{f}_{T\sigma_T^2}\right\}_{T\geq 0}$ is an additive process w.r.t. T.[9]

1.4 Model selection and robustness tests

In this section, we show two additional results for the proposed process class. First, we compare the ATS with two other additive processes already present in the financial literature and propose some statistical tests able to select the most adequate modeling description of the implied volatility surface. Then, we show that the results, described in detail in the previous section, appear robust over time.

1.4.1 Model selection via statistical tests

In this subsection, we compare ATS with two classes of additive processes already present in the financial literature, the Sato processes (see, e.g., Carr *et al.* 2007) and the additive processes constructed via additive subordination (see, e.g., Li *et al.* 2016). The comparison is among processes that have the marginal distribution of normal tempered stable type: i.e. with the Sato processes NIGSSD and VGSSD and with the sub-class of ATS constructed through additive subordination. An ATS process can be obtained as a Brownian motion subordinated with an additive subordinator, as in Li *et al.* (2016), if and only if η_T is constant.[10] We discuss two features: one related to the η_T parameter and another to the skewness and to the excess kurtosis of the calibrated exponential forward.

A first test is built to verify the adequacy of Sato processes. Given the *index of stability* for the model (e.g., chosen α in the Normal Tempered Stable model), it is possible to compute skewness and kurtosis (see, e.g., Cont and Tankov 2003, p.129). For example the ATS NIG skewness is

$$\frac{\mathbb{E}\left[\left(f_T - \mathbb{E}\left[f_T\right]\right)^3\right]}{\left(Var(f_T)\right)^{\frac{3}{2}}} = -\frac{3\sigma_T^4\left(\eta_T+\frac{1}{2}\right)k_T T + 2{\sigma_T}^6\left(\eta_T+\frac{1}{2}\right)^3 k_T^2 T}{\left({\sigma_T}^2 T + k_T T {\sigma_T}^4\left(\eta_T+\frac{1}{2}\right)^2\right)^{\frac{3}{2}}} \ .$$

A Sato process has skewness and kurtosis constant over time, as it can be deduced by definition (see, e.g., Carr *et al.* 2007).

We analyze the term structure of these higher-order moments observed in our dataset adopting the same procedure of Konikov and Madan (2002). For both indices, we observe a linear behavior of skewness and kurtosis w.r.t. the squared root of the maturity as shown in Figure 1.9 in the NIG case. In the Figure, we have plotted also the confidence interval chosen equal to two times the standard deviation, respectively, of the skewness and the kurtosis (cf. appendix .2 for the methodology adopted to obtain these standard deviations).

The statistical test is simple. We perform a linear regression statistical analysis of the higher moments behavior w.r.t. the square root of the time-to-maturity $\sqrt{T}$: we reject the null hypothesis of no slope in all of the cases that we analyze (both indices and both tempered stable models; that is, NIG and VG) with p-values of the order of 10^{-16}. Similar results hold in all ATS cases.

[9] We have also considered a global calibration of the implied volatility surfaces considering the power-law scaling parameters in (1.3.2) with $\beta = 1$ and $\delta = -0.5$. The results are of the same order of magnitude of Table 1.1.

[10] Proof available in appendix .3.

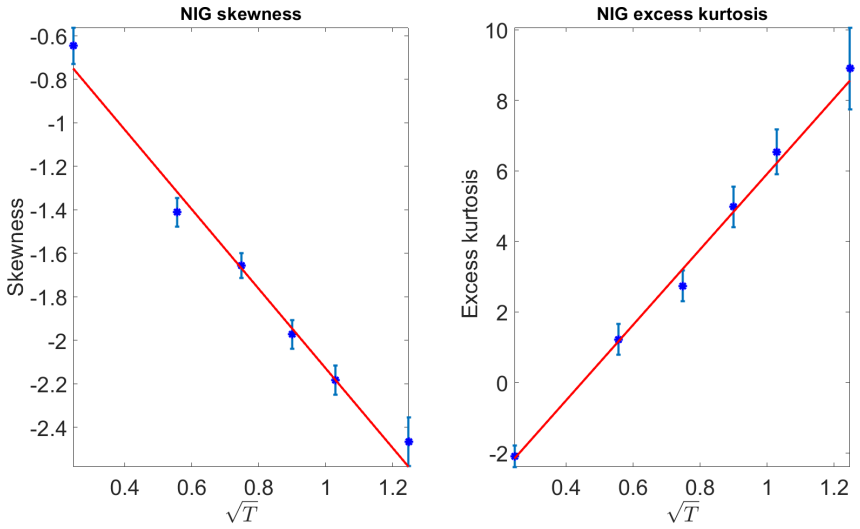

Figure 1.9: Observed time-dependent skewness (kurtosis) w.r.t. $\sqrt{T}$ for the NIG calibrated model on S&P 500 volatility surfaces. We plot a confidence interval equal to two times the standard deviation. The behavior is not consistent with a Sato process.

The other statistical test aims to verify the adequacy of additive processes obtained through additive subordination (Li *et al.* 2016) in volatility surface calibration. As already mentioned the ATS process, when $\hat{\eta}_\theta$ is equal to a constant $\bar{\eta}$, falls within this class.

In Figure 1.7 and 1.8 we have already shown the time scaling $\hat{\eta}_\theta$. We can statistically test the null hypothesis of constant $\hat{\eta}_\theta$. For both volatility surfaces and for both tested tempered stable models (NIG and VG) we reject the null hypothesis of a constant $\hat{\eta}_\theta$ with p-values below 10^{-7}. As already observed, ATS processes are characterized by a power-law scaling in $\hat{\eta}_\theta$.

1.4.2 Robustness tests

In this subsection, we perform a robustness analysis of the results in section 1.3. We repeat the analysis on four other days, both on the S&P 500 and EURO STOXX 50 volatility surfaces. We show that the excellent calibration features of the ATS and the power-law scaling properties, observed on the 30^{th} of May 2013, arise also in these other dates.

In these robustness tests, we use bid and ask close prices for the S&P 500 and EURO STOXX 50 options on the 29^{th} of November 2012 (6 months before the date of the analysis of section 1.3, the 30^{th} of May 2013), the 27^{th} of February 2013 (3 months before), the 30^{th} of August 2013 (3 months after), and the 29^{th} November 2013 (6 months after).[11] The dataset includes the bootstrapped risk-free rate curve. The data is provided by Reuters Datastream (option data) and Reuters Eikons (rate data). Let us observe that close prices are, in general, less accurate than open market prices (the ones used for the analysis in section 1.3).

In Table 1.4, we report calibration performances for the S&P 500 and EURO STOXX 50 in terms of MSE and MAPE on the four dates considered. In the NIG and VG cases, we consider

[11] These are the penultimate business days of November 2012, February 2013, August 2013, and November 2013.

the standard Lévy process, the Sato process, and the corresponding ATS process. In all considered cases, Sato processes perform better than Lévy processes but ATS improvement is far more significant: on average, two orders of magnitude for MSE and one order of magnitude for MAPE. These results appear coherent with the ones reported in Table 1.1.

		MSE			MAPE			MSE			MAPE		
Index	Model	Lévy	Sato	ATS	Lévy	Sato	ATS	Lévy	Sato	ATS	Lévy	Sato	ATS
		29^{th} of November 2012 (-6 months)						27^{th} of February 2013 (-3 months)					
S&P 500	NIG	4.78	1.15	**0.38**	2.93%	1.36%	**0.60**%	10.77	4.31	**0.52**	3.65%	3.30%	**0.66**%
S&P 500	VG	11.04	1.00	**0.38**	4.46%	1.36%	**0.71**%	18.28	3.74	**0.48**	4.81%	2.16%	**0.69**%
Euro Stoxx 50	NIG	20.64	19.73	**0.26**	2.39%	2.29%	**0.18**%	54.54	19.99	**0.15**	3.79%	2.47%	**0.15**%
Euro Stoxx 50	VG	34.51	20.65	**0.41**	3.05%	2.38%	**0.31**%	90.81	19.25	**0.38**	4.91%	2.47%	**0.24**%
		30^{th} of August 2013 (+3 months)						29^{th} November 2013 (+6 months)					
S&P 500	NIG	8.27	1.08	**0.18**	3.29%	1.21%	**0.12**%	10.23	1.42	**0.01**	3.50%	1.32%	**0.09**%
S&P 500	VG	18.37	0.98	**0.37**	4.95%	1.16%	**0.20**%	16.80	1.36	**0.09**	4.52%	1.30%	**0.35**%
Euro Stoxx 50	NIG	40.98	5.03	**1.53**	2.68%	0.93%	**0.44**%	24.22	12.03	**0.50**	2.38%	1.73%	**0.27**%
Euro Stoxx 50	VG	59.23	4.81	**0.64**	3.26%	0.94%	**0.32**%	57.25	12.66	**0.91**	3.75%	1.77%	**0.45**%

Table 1.4: Calibration performance for the S&P 500 and EURO STOXX 50 in terms of MSE and MAPE on the 29^{th} of November 2012 (6 months before the date of the analysis), the 27^{th} of February 2013 (3 months before), the 30^{th} of August 2013 (3 months after), and the 29^{th} November 2013 (6 months after). In the NIG and VG cases, we consider the standard Lévy process, the Sato process, and the corresponding ATS process, as in Table 1.1. In all considered cases, Sato processes perform better than Lévy processes but ATS improvement is far more significant: two orders of magnitude for MSE and one order of magnitude for MAPE.

In Figure 1.10, we plot the MSE w.r.t. the different times-to-maturity (in years) for S&P 500 volatility surface calibrated with a NIG process on the four considered dates. We observe that Sato (circles) and Lévy (triangles) have a MSE of the same order of magnitude, while the improvement of ATS (squares) is, on average, of two orders of magnitude and particularly significant at short-time.

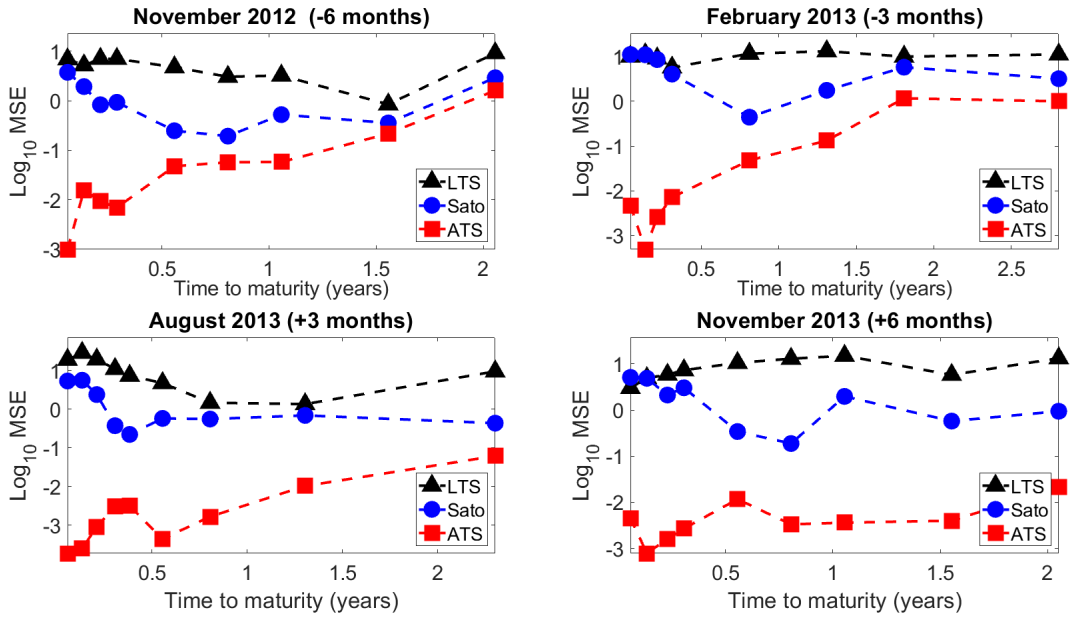

Figure 1.10: MSE w.r.t. the different times-to-maturity (in years) for S&P 500 volatility surface calibrated with a NIG process on the 29^{th} of November 2012 (6 months before), the 27^{th} of February 2013 (3 months before), the 30^{th} of August 2013 (3 months after), and the 29^{th} November 2013 (6 months after). Sato (circles) and Lévy (triangles) have a MSE of the same order of magnitude, while the improvement of ATS (squares) is, on average, of two orders of magnitude and particularly significant at short-time.

In Figures 1.11 and 1.12, we plot the implied volatility smile for S&P 500 on the 29^{th} of November 2012 (time-to-maturity of 22 days on the left and of 9 months and 22 days on the right) and on the 29^{th} of November 2013 (time-to-maturity of 21 days on the left and of 9 months and 21 days on the right). The NIG ATS process, Sato process, and LTS process implied volatility are plotted together with the market-implied volatility. As in the case of the 30^{th} of May 2013 (cf. Figure 1.5) the ATS reproduces the smiles significantly better than the alternatives, the improvement is particularly evident for small maturities.

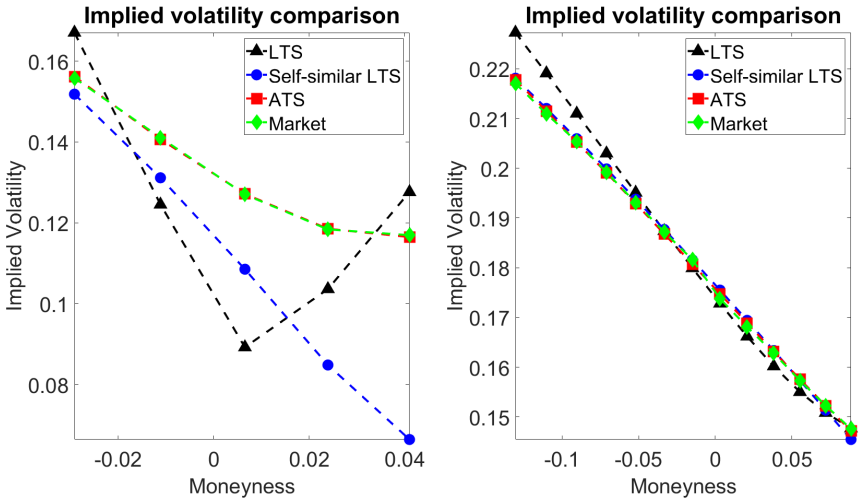

Figure 1.11: Implied volatility smile for S&P 500 at the 29^{th} of November 2012 (6 months before) for a given time-to-maturity: 22 days (on the left) and 9 months and 22 days (on the right). The NIG ATS process, Sato process, and LTS process implied volatility are plotted together with the market-implied volatility. ATS reproduces the smile significantly better than the alternatives, the improvement is particularly evident for small maturities.

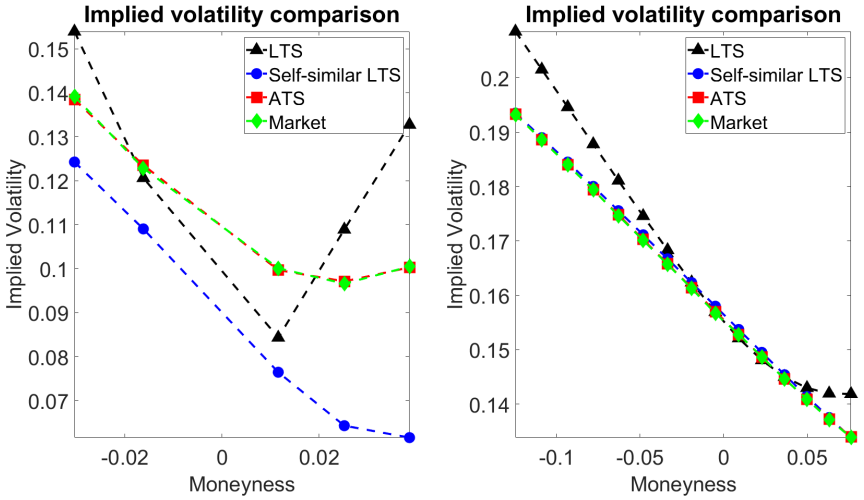

Figure 1.12: Implied volatility smile for S&P 500 at the 29^{th} of November 2013 (6 months after) for a given time-to-maturity: 21 days (on the left) and 9 months and 21 days (on the right). The NIG ATS process, Sato process, and LTS process implied volatility are plotted together with the market-implied volatility. ATS reproduces the smile significantly better than the alternatives, the improvement is particularly evident for small maturities.

In Figure 1.13, we plot the market and the ATS NIG implied volatility *skew* for EURO STOXX 50 w.r.t. the times-to-maturity on the 27^{th} of February and on the 30^{th} of August 2013. In both cases, the calibrated ATS replicates accurately the market implied volatility *skew*, as already observed in section 1.3 for the 30^{th} of May 2013. Similar results hold for the other two dates and in the S&P 500 case.

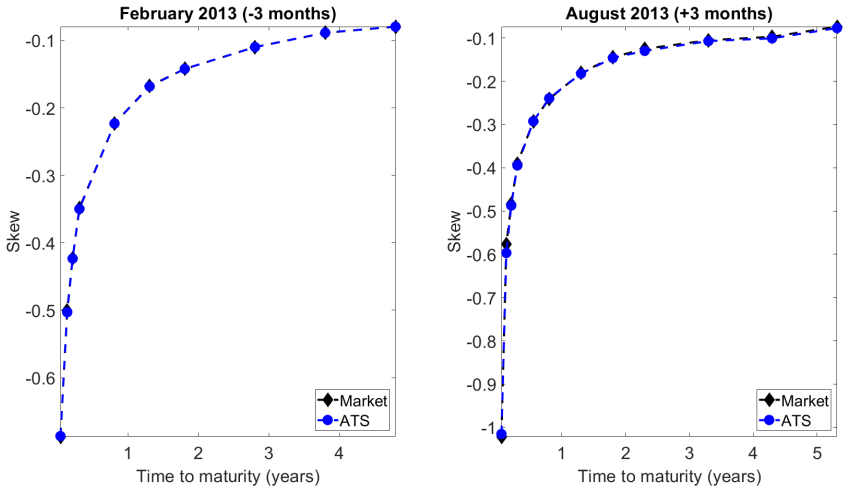

Figure 1.13: The market and the ATS NIG implied volatility *skew* for EURO STOXX 50 w.r.t. the times-to-maturity on the 27^{th} of February and on the 30^{th} of August 2013. Again, ATS replicates the market implied implied volatility *skew*.

In Figures 1.14 and 1.15, we plot the weighted regression lines and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ with their confidence intervals for S&P 500 and EURO STOXX 50 on the 29^{th} of November 2012 (on the top) and on the 29^{th} of November 2013 (on the bottom).[12] The confidence intervals are two times the standard deviations of $\ln \hat{k}_\theta$, of $\ln \hat{\eta}_\theta$ and of $\ln \theta$. In both days, the observed scalings of $\hat{k}_\theta$ and $\hat{\eta}_\theta$ are equivalent to the ones observed in Figures 1.7 and 1.8.

[12]Results for the 27^{th} of February 2013 and the 30^{th} of August 2013 are available upon request.

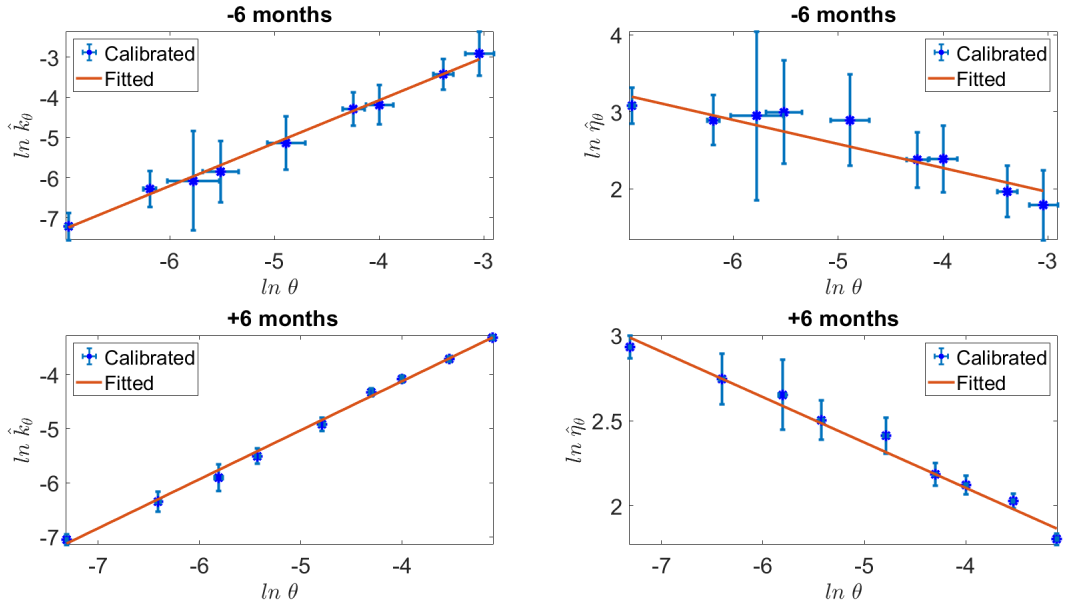

Figure 1.14: Weighted regression line and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ w.r.t. $\ln \theta$ for the NIG calibrated model for S&P 500 on the 29^{th} of November 2012 (on the top) and on the 29^{th} of November 2013 (on the bottom). We plot a confidence interval equal to two times the corresponding standard deviation. Notice that, also in this case, confidence intervals on $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ are one order of magnitude wider than confidence intervals on $\ln \theta$. The observed scalings of $\hat{k}_\theta$ and $\hat{\eta}_\theta$ are equivalents to the ones observed in Figures 1.7 and 1.8. The values of θ correspond to times-to-maturity that go from tree weeks to two years and a half (all available maturities on Reuters Datastream dataset).

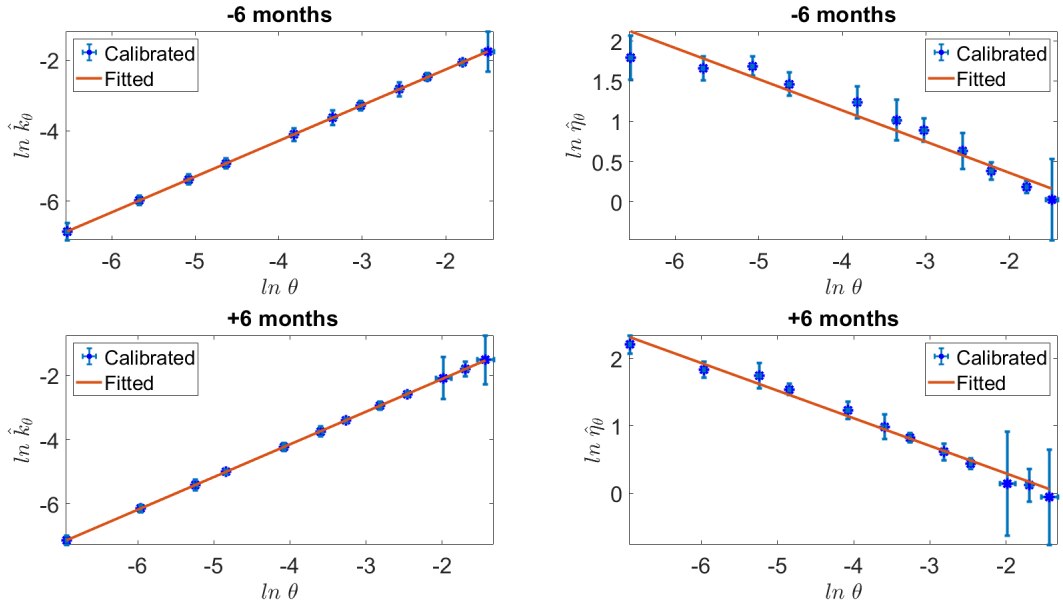

Figure 1.15: Weighted regression line and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ w.r.t. $\ln \theta$ for the VG calibrated model for EURO STOXX 50 on the 29^{th} of November 2012 (on the top) and on the 30^{th} of November 2013 (on the bottom). We plot a confidence interval equal to two times the corresponding standard deviation. Notice that, also in this case, confidence intervals on $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ are one order of magnitude wider than confidence intervals on $\ln \theta$. The observed scalings of $\hat{k}_\theta$ and $\hat{\eta}_\theta$ are equivalents to the ones observed in Figures 1.7 and 1.8. The values of θ correspond to times-to-maturity that go from tree weeks to five years (all available maturities on Reuters Datastream dataset).

The results presented in Table 1.4 and in Figures 9-14 are equivalent to the one of section 1.3. This analysis confirms the robustness, over a one-year time interval, of the excellent calibration performances and the power scaling behavior of the ATS.

1.5 Conclusions

In this chapter, we introduce a new broad family of stochastic processes that we call additive normal tempered stable processes (ATS). An interesting subcase of ATS presents a power-law scaling of the time-dependent parameters.

We have considered all quoted options on S&P 500 and EURO STOXX 50 at 11:00 am NT on the 30^{th} of May 2013. The dataset considers options with a time-to-maturity starting from three weeks and up to several years. We calibrate the ATS processes on the options of both indices, showing that ATS processes present better calibration features than LTS and Sato processes. The observed improvement of ATS is even of two orders of magnitude in terms of MSE, as reported in Table 1.1. ATS replicates accurately market implied volatility term structure and *skew* as observed in Figures 1.5 and 1.6.
The quality of ATS calibration results looks quite incredible. In subsection 1.3.3, we have shown that once the volatility term structure has been taken into account, the whole implied volatility surface is calibrated accurately with only two free parameters.

We also construct a re-scaled ATS process via a time-change based on the implied volatility term structure. We show that the re-scaled process calibrated parameters exhibit a power-law behavior. The statistical relevance of the scaling properties is discussed in detail.

We have compared some model characteristics with the two alternative additive processes present in the financial literature. These two classes fail to reproduce some stylized facts observed in market data, which are adequately described by ATS processes.
Finally, we have verified the robustness, over a one-year time interval, of the excellent calibration performances and the power scaling behavior of the ATS. In section 4.4, we discuss the ATS calibration performances on a nine-year dataset.

For future research, two main promising directions appear evident. First, it can be interesting to extend ATS processes to the commodity asset class, in general, and to the oil (Shiraya and Takahashi 2011, Kyriakou *et al.* 2016) and freight markets (Prokopczuk 2011, Nomikos *et al.* 2013), in particular; this model extension should allow for mean-reversion and seasonality patterns in prices, which are typically found in empirical studies (see, e.g., Benth *et al.* 2014, and references therein). We include a preliminary analysis in appendix .7. Second, it would be worthy to develope a fast and reliable Monte Carlo scheme for ATS processes and to study pricing techniques for exotic derivatives (e.g., generalizing the techniques for path-dependent exotics products, as Asian options, in Fusai and Meucci 2008, Černỳ and Kyriakou 2011, Fusai and Kyriakou 2016, through the characteristic function of the ATS increments). We develop two Monte Carlo schemes in chapter 3.

Chapter 2

Short-time implied volatility of additive normal tempered stable processes

Empirical studies have emphasized that the equity implied volatility is characterized by a negative skew inversely proportional to the square root of the time-to-maturity.

We examine the short-time-to-maturity behavior of the implied volatility smile for pure jump exponential additive processes introduced in the previous chapter. An excellent calibration of the equity volatility surfaces is achieved by the power-law scaling ATS. The two power-law scaling parameters are β, related to the variance of jumps, and δ, related to the smile asymmetry. It has been observed, in option market data (cf. chapter 1), that $\beta = 1$ and $\delta = -1/2$.
In this chapter, we prove that the implied volatility of these additive processes is consistent, in the short-time, with the equity market empirical characteristics if and only if $\beta = 1$ and $\delta = -1/2$.

These results have already been presented in Azzone and Baviera (2021c).

2.1 Introduction

Which characteristics of the implied volatility surface should be reproduced by an option pricing model? A *stylized fact* that characterizes the equity market is a downward slope in terms of strike, i.e. a negative skew, we recall that the skew is the at-the-money (ATM) derivative of the implied volatility w.r.t. the moneyness.[1] Specifically, the short-time[2] negative skew is proportionally inverse to the square root of the time-to-maturity. The first empirical study of the equity skew dates back to Carr and Wu (2003): they find that the S&P 500 short-time skew is, on average, asymptotic to $-0.25/\sqrt{t}$. Fouque *et al.* (2004) arrive at a similar conclusion considering only options with short-time-to-maturity (i.e. up to three months). In this chapter, we show that the power law scaling ATS, which also calibrates accurately the whole equity volatility surface, reproduces the power scaling market skew.

A vast literature on short-time implied volatility and skew is available for jump-diffusion processes. Both the ATM (see, e.g., Alòs *et al.* 2007, Roper 2009, Muhle-Karbe and Nutz 2011, Andersen and Lipton 2013, Figueroa-López *et al.* 2016) and the OTM implied volatility (see, e.g., Tankov 2011, Figueroa-López and Forde 2012, Mijatović and Tankov 2016, Figueroa-López

[1] The moneyness is the logarithm of the strike price over the forward price. For a description of the equity volatility surface and a definition of skew, see, e.g., Gatheral (2011).
[2] We refer to short-time-to-maturity.

et al. 2018) are analyzed. For a jump-diffusion Lévy process, the ATM implied volatility is determined uniquely by the diffusion term; it goes to zero as the time-to-maturity goes to zero if there is no diffusion term, i.e. for a pure-jump process. For this reason, pure jumps Lévy processes are not suitable to reproduce the market short-time smile, because the short-time implied volatility is strictly positive in all financial markets.
Muhle-Karbe and Nutz (2011) have shown that, for a relatively broad class of additive models, the ATM behavior at small-time is the same as the corresponding Lévy. In this chapter, we analyze the ATM implied volatility and skew for a class of pure jump additive processes that is consistent with the equity market smile, differently from the Lévy case: this is the main theoretical contribution of this study.

In chapter 1, we have introduced the ATS, a pure jump additive extension of the well-known Lévy normal tempered stable process (for a comprehensive description of this set of Lévy processes, see, e.g., Cont and Tankov 2003, Ch.4). In the following, we will focus on the short time behavior of this process.

Pure jump processes present a main advantage w.r.t. jump-diffusion models: they generally describe underlying dynamics more parsimoniously. In a jump-diffusion, both small jumps and the diffusion term describe little changes in the process (see, e.g., Asmussen and Rosiński 2001). Because both components of the jump-diffusion process are qualitatively similar, when calibrating the model to the plain vanilla option market, it is rather difficult to disentangle the two components and several sets of parameters achieve similar results.

The power-law scaling ATS which is characterized by two key time-dependent parameters –the variance of jumps per unit of time, k_t, and the asymmetry parameter, η_t– that present a power scaling w.r.t. the time-to-maturity t. The excellent calibrating performances of this class of processes has been shown in chapter 1 and will be discussed also in chapter 2 and appendix .7. On the one hand, this class of pure jump additive processes allows calibrating the S&P 500 and EURO STOXX 50 implied volatility surfaces with great accuracy, reproducing "exactly" the term structure of the equity market implied volatility surfaces. On the other hand, the observed reproduction of the skew term structure appears remarkable.
Moreover, an interesting self-similar characteristic w.r.t. the time-to-maturity, in the form of (1.3.2), arises. Specifically, among all allowed power laws , the power scaling index of k_t, β, is close to one, while the power scaling index of η_t, δ, is statistically consistent with minus one half (see, e.g., Azzone and Baviera 2021a).

Consider an option price with strike K and time-to-maturity t. We define $I_t(x)$ the model implied volatility, where $x := \log \frac{K}{F_0(t)}$ is the moneyness and $F_0(t)$ is the underlying forward price with time-to-maturity t. In particular, we consider the *moneyness degree* y, s.t. $x =: y\sqrt{t}$, introduced by Medvedev and Scaillet (2006). It has been observed that the *moneyness degree* y can be interpreted as the distance of the option moneyness from the forward price in terms of the Black Brownian motion standard deviation (see, e.g., Carr and Wu 2003, Medvedev and Scaillet 2006). The implied volatility w.r.t y is

$$\mathcal{I}_t(y) := I_t(y\sqrt{t}) \ ,$$

and its first order Taylor expansion w.r.t. y in $y = 0$ is

$$\mathcal{I}_t(y) = \mathcal{I}_t(0) + y \left.\frac{d\mathcal{I}_t(y)}{dy}\right|_{y=0} + o(y) =: \hat{\sigma}_t + y\,\hat{\xi}_t + o(y) \ .$$

We call $\hat{\xi}_t$ the skew term. We define $\hat{\sigma}_0$ and $\hat{\xi}_0$ as the the limits for t that goes to zero of $\hat{\sigma}_t$ and $\hat{\xi}_t$. Their financial interpretation is straightforward: $\hat{\sigma}_0$ corresponds to the short-time ATM

implied volatility, while $\hat{\xi}_0$ is related to the short-time skew, because it is possible to write the skew as

$$\left.\frac{dI_t(x)}{dx}\right|_{x=0} = \frac{\hat{\xi}_t}{\sqrt{t}} \quad .$$

In Figure 2.1, we present an example of the short-time implied volatility and the skew for the S&P 500 at a given date, the 22^{nd} of June 2020 (the business day after a quadruple witching Friday[3]). On the left, we plot the one month (blue circles), two months (red squares), three months (orange stars), and four months (purple triangles) market implied volatility w.r.t. the *moneyness degree* y: we observe a positive and finite short-time $\hat{\sigma}_t$. On the right, we plot the market skew w.r.t. the time t: it appears to be well described by a fit $O\left(\sqrt{\frac{1}{t}}\right)$.

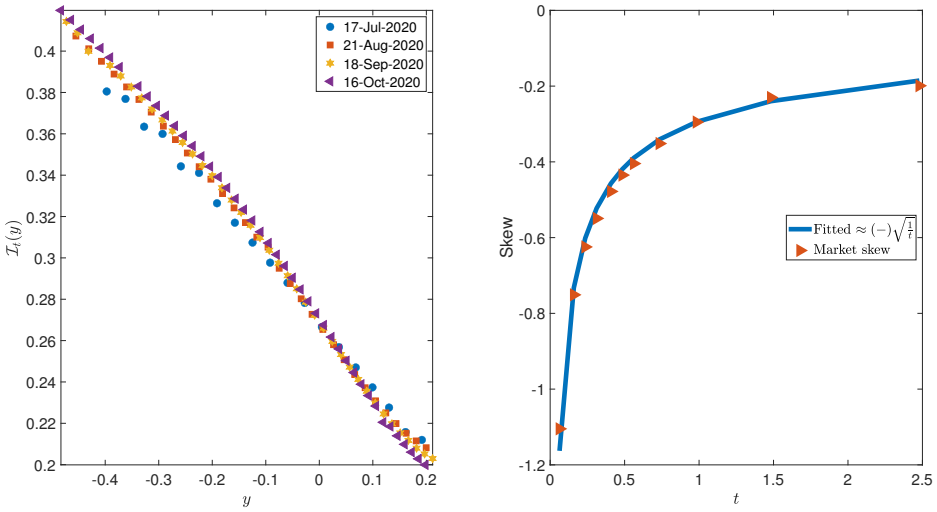

Figure 2.1: Example of the S&P 500 short-time implied volatility and skew on the 22^{nd} of June 2020. On the left, we plot the one month (blue circles), two months (red squares), three months (orange stars), and four months (purple triangles) market implied volatility w.r.t. the *moneyness degree* y. We observe a positive short-time $\hat{\sigma}_t$. On the right, we plot the market skew w.r.t. the time t and the fitted $\approx (-)\sqrt{\frac{1}{t}}$.

As already observed in some empirical studies (see, e.g., Carr and Wu 2003, Fouque *et al.* 2004), equity market data are compatible with a positive and finite $\hat{\sigma}_0$ and a negative and finite $\hat{\xi}_0$, that leads to a skew proportionally inverse to the square root of the time-to-maturity. We aim to present a pure-jump model with these features.

We study the behavior of $\hat{\sigma}_t$ and $\hat{\xi}_t$ for the ATS process, deriving, in (2.2.4, 2.2.5), an extension of the Hull and White (1987, Eq.(7), p.4) formula (see, e.g., Alòs *et al.* 2007, for another application of this formula to the short-time case). This formula leads to two results: on the one hand, we build some relevant bounds for $\hat{\sigma}_t$; on the other hand, we obtain an expression for $\hat{\xi}_t$ in (2.4.1) via the Dini's implicit function theorem (Dini 1907, pp.197-241).

[3]A quadruple witching Friday is the third Friday of the months of March, June, September and December: in this quarterly date, stock options, stock futures, equity index futures, and equity index options all expire on the same day.

There are three main contributions in this chapter. First, we deduce for a family of pure-jump additive processes, the ATS, the behavior of the short-time ATM implied volatility $\hat{\sigma}_t$ and skew term $\hat{\xi}_t$ over the region of admissible parameters (see **Theorem 1.2.3**). Second, we prove that only the scaling parameters observed in market data ($\beta = 1$ and $\delta = -1/2$) are compatible with a finite short-time implied volatility and a short-time skew proportionally inverse to the square root of the time-to-maturity. Third, we prove that there exists a pure-jump additive process (an exponential ATS) that presents the two features observed in market data: not only a finite and positive short-time implied volatility but also a power scaling skew.

The rest of the chapter is organized as follows. Section 2.2 presents the ATS power scaling process and the extension of the Hull and White formula. Section 2.3 defines the implied volatility problem and analyzes the short-time ATM implied volatility $\hat{\sigma}_t$. Section 2.4 computes the short-time limit of the skew term $\hat{\xi}_t$. Section 2.5 presents the major result: the ATS process is consistent with the equity market if and only if $\beta = 1$ and $\delta = -1/2$. Finally, section 2.6 concludes. In the appendices, we report some technical lemmas that we use in this chapter: on basic properties in appendix .4 on short-time limits in appendix .5.

2.2 The ATS implied volatility

In this section, we recall the characteristic function of the power-law scaling additive normal tempered stable process (ATS), introduced in chapter 1, and we introduce the notation employed in the chapter. We also introduce a sequence of random variables (2.2.3) with the same distribution of the ATS for any fixed time t; we use these random variables to study the short-time implied volatility.
We discuss the volatility smile at small-maturity produced by this forward model and determine the power laws of the ATS parameters that are consistent with the market data i.e. which choices of β and δ reproduce the market short-time features mentioned above.

We define a sequence of positive random variables S_t via its Laplace transform. The random variable S_t appears in the definition of the random variable f_t, that is used to model a forward contract of the underlying of interest.

Definition 2.2.1. Definition of $\{S_t\}_{t\geq 0}$
Let $\{S_t\}_{t\geq 0}$ be a sequence of positive random variables s.t. the logarithm of the Laplace transform of tS_t is

$$\ln \mathcal{L}_t\left(u;\ k_t,\ \alpha\right) := \ln \mathbb{E}\left[e^{-utS_t}\right] = \begin{cases} \dfrac{t}{k_t}\dfrac{1-\alpha}{\alpha}\left\{1-\left(1+\dfrac{u\ k_t}{(1-\alpha)}\right)^{\alpha}\right\} & \text{if } 0<\alpha<1 \\ -\dfrac{t}{k_t}\ln\left(1+u\ k_t\right) & \text{if } \alpha = 0 \end{cases}, \ t\geq 0$$

where $k_t := \bar{k}t^{\beta}$ and $\bar{k},\ \beta \in \mathbb{R}^+$.

Notice that, by the Laplace transform, we can compute any moment of S_t. The first two are

1. $\mathbb{E}\left[S_t\right] = 1$;
2. $Var\left[S_t\right] = k_t/t$.

Definition 2.2.2. Definition of $\{f_t\}_{t\geq 0}$
Let $\{f_t\}_{t\geq 0}$ be a sequence of random variables with characteristic function s.t.

$$\mathbb{E}\left[e^{iuf_t}\right] := \mathcal{L}_t\left(iut\left(\frac{1}{2}+\eta_t\right)\bar{\sigma}^2 + t\frac{u^2\bar{\sigma}^2}{2};\ k_t,\ \alpha\right)e^{i\,u\,\varphi_t\,t}\ ,\ t\geq 0 \tag{2.2.1}$$

where

$$\eta_t := \bar{\eta} t^{\delta} \text{ and } \varphi_t\, t := -\ln \mathcal{L}_t \left(t\bar{\sigma}^2 \eta_t;\ k_t,\ \alpha \right) \quad . \tag{2.2.2}$$

$\bar{\sigma},\ \bar{\eta} \in \mathbb{R}^+$ *and* $\delta \in \mathbb{R}$.

Notice that the characteristic function is the same as the power-law scaling additive normal tempered stable process (ATS) in chapter 1. We also define $F_0(t)$, the forward contract at time 0 with maturity t, and model the same forward contract at maturity as

$$F_t(t) := F_0(t)\ e^{f_t} \quad .$$

We report the result on the existence of an additive process with characteristic function (2.2.1), cf. **Theorem** 1.2.3.

Theorem 2.2.3. Power-law scaling ATS
There exists an additive process with the same characteristic function of (2.2.1), where $\alpha \in [0, 1)$ *and* $\beta, \delta \in \mathbb{R}$ *with either* $\beta = \delta = 0$ *or*

1. $0 \leq \beta \leq \dfrac{1}{1 - \alpha/2}$;

2. $-\min\left(\beta, \dfrac{1 - \beta\,(1-\alpha)}{\alpha}\right) < \delta \leq 0$;

where the second condition reduces to $-\beta < \delta \leq 0$ *for* $\alpha = 0$ □

The region of admissible values for the scaling parameters β and δ is shown in Figure 2.2.
In particular, we mention that, $\forall\, \alpha \in [0, 1)$, the scaling parameters observed in the market, $\{\delta = -1/2,\ \beta = 1\}$, are always inside the ATS admissible region. In Figure 2.2, we plot the admissible region for the scaling parameters β and δ. In this chapter, we prove that the ATS implied volatility at short-time is qualitatively different for different sets of scaling parameters. We separate the admissible region into five Cases:
Case 1 (grey area):

$$\left\{\beta < 1,\ \ -\min\left(\frac{1}{2}, \beta\right) < \delta \leq 0\right\} \cup \{\delta = \beta = 0\} \quad .$$

Case 2 (orange area):

$$\left\{-\min\left(\beta, \frac{1 - \beta(1-\alpha)}{\alpha}\right) < \delta < -\frac{1}{2}\ \max\left(\beta, 1\right)\right\} \quad .$$

Case 3 (light green area):

$$\left\{\beta \geq 1,\ \ -\frac{\beta}{2} \leq \delta \leq 0\right\} \setminus \left\{\beta = 1,\ \ \delta = -\frac{1}{2}\right\} \quad .$$

Case 4 (continuous dark green line):

$$\left\{\beta < 1,\ \ \delta = -\frac{1}{2}\right\} \quad .$$

Case 5 (red dot):

$$\left\{\beta = 1,\ \ \delta = -\frac{1}{2}\right\} \quad .$$

Notice that Case 3 includes all its boundaries, identified by the green circles, with the exception of the point $\{\beta = 1,\ \delta = -1/2\}$ (red); Case 1 includes just its upper boundary (but it does not include its lower boundary), identified by the grey squares.
The main objective of this chapter is to prove that the 5 Cases correspond to different behaviors of the implied volatility in the short-time. A summary of the ATS short-time behavior, and in particular of the ATM value $\hat{\sigma}_0$ and of the skew term $\hat{\xi}_0$, w.r.t. the different Cases is available in Table 2.1.

Case 1	$\hat{\sigma}_0 = 0$		
Case 2	$\hat{\sigma}_0 = \infty$		
Case 3	$\hat{\sigma}_0 > 0$	and	$\hat{\xi}_0 = 0$
Case 4	$\hat{\sigma}_0 > 0$	and	$\hat{\xi}_0 = -\sqrt{\frac{\pi}{2}}$
Case 5	$\hat{\sigma}_0 > 0$	and	$\hat{\xi}_0 < 0$

Table 2.1: Summary of ATS short-time implied volatility behavior in the five Cases. We observe that only the last 3 Cases admit a positive (and finite) implied volatility.

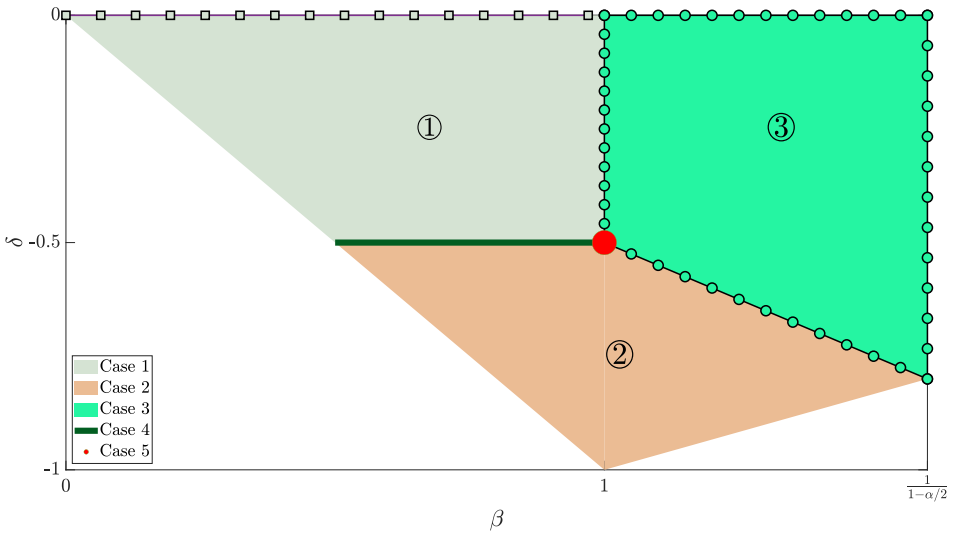

Figure 2.2: ATS admissible region for the scaling parameters. We separate the region in five Cases. i) Case 1 (grey area) with $\hat{\sigma}_0 = 0$. ii) Case 2 (orange area) with $\hat{\sigma}_0 = \infty$. iii) Case 3 (light green area) with finite $\hat{\sigma}_0$ and $\hat{\xi}_0 = 0$. iv) Case 4 (continuous dark green line) with finite $\hat{\sigma}_0$ and $\hat{\xi}_0 = -\sqrt{\frac{\pi}{2}}$. v) Case 5 (red dot) with finite $\hat{\sigma}_0$ and negative and finite $\hat{\xi}_0$.
Notice that Case 3 includes all its boundaries, identified by the green circles, with the exception of the point $\{\beta = 1,\ \delta = -1/2\}$ (red), that corresponds to Case 5. Moreover, Case 1 includes just its upper bound, identified by the grey squares. We emphasize that for all α in $[0, 1)$ the point $\{\beta = 1,\ \delta = -1/2\}$ is inside the admissible region.

It is also useful to provide the same result dividing the region for the admissible values of **Theorem 1.2.3**, in terms of β and δ. A summary of the ATS short-time behavior, w.r.t. the scaling parameters β and δ in the additive process admissible region is reported in Table 2.2.

	$0<\beta<1$	$\beta=1$	$\beta>1$
$\delta > -\frac{1}{2}$	$\hat{\sigma}_0 = 0$	$\hat{\xi}_0 = 0$	$\hat{\xi}_0 = 0$
$\delta = -\frac{1}{2}$	$\hat{\xi}_0 = -\sqrt{\frac{\pi}{2}}$	$\boldsymbol{\hat{\xi}_0 < 0}$	$\hat{\xi}_0 = 0$
$\delta < -\frac{1}{2}$	$\hat{\sigma}_0 = \infty$	$\hat{\sigma}_0 = \infty$	$\hat{\sigma}_0 = \infty$ or $\hat{\xi}_0 = 0$

Table 2.2: Summary of ATS short-time implied volatility behavior, w.r.t. the scaling parameters β and δ in the additive process admissible region. The ATM implied volatility σ_0 is positive (and finite) when we report the value for $\hat{\xi}_0$.

It can be proven that, for every time t, the random variable

$$f_t = -\left(\eta_t + \frac{1}{2}\right)\,\bar{\sigma}^2\,S_t\,t + \,\bar{\sigma}\sqrt{S_t t}\,g + \varphi_t t \tag{2.2.3}$$

has the characteristic function in (2.2.1), where g is a standard normal random variable independent from S_t. The proof is the same as in the Lévy case, but with time dependent parameters, and it is obtained by direct computation of $\mathbb{E}[e^{i\,u\,f_t}]$, conditioning w.r.t. S_t. The f_t in (2.2.3) is then equivalent in law to the ATS process at maturity t; thus, we can use this expression of f_t to compute the price of European options.

Consider a European call option discounted payoff $B_t\left(F_0(t)\,e^{f_t} - F_0(t)\,e^{x}\right)^+$ (and $B_t\left(F_0(t)\,e^{x} - F_0(t)\,e^{f_t}\right)^+$ is the discounted payoff for the corresponding put) where t is option time-to-maturity, K option strike price, $x := \ln\frac{K}{F_0(t)}$ the asset moneyness and B_t the deterministic discount factor between 0 and t. We can write the expected European call and put option price at time zero conditioning on S_t

$$\begin{aligned} C_t(x) &= B_t\,F_0(t)\,\mathbb{E}\left[\left(e^{f_t} - e^{x}\right)^+\right] \\ &= B_t\,F_0(t)\,\mathbb{E}\left[e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} N\left(\frac{-x}{\bar{\sigma}\sqrt{S_t t}} + l_t^{S_t} + \frac{\bar{\sigma}\sqrt{S_t t}}{2}\right) - e^{x} N\left(\frac{-x}{\bar{\sigma}\sqrt{S_t t}} + l_t^{S_t} - \frac{\bar{\sigma}\sqrt{S_t t}}{2}\right)\right] \end{aligned} \tag{2.2.4}$$

$$\begin{aligned} P_t(x) &= B_t\,F_0(t)\,\mathbb{E}\left[\left(e^{x} - e^{f_t}\right)^+\right] \\ &= B_t\,F_0(t)\,\mathbb{E}\left[e^{x} N\left(\frac{x}{\bar{\sigma}\sqrt{S_t t}} - l_t^{S_t} + \frac{\bar{\sigma}\sqrt{S_t t}}{2}\right) - e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} N\left(\frac{x}{\bar{\sigma}\sqrt{S_t t}} - l_t^{S_t} - \frac{\bar{\sigma}\sqrt{S_t t}}{2}\right)\right], \end{aligned} \tag{2.2.5}$$

where

$$l_t^z := -\bar{\sigma}\eta_t\sqrt{z\,t} + \frac{\varphi_t\sqrt{t}}{\bar{\sigma}\sqrt{z}} \tag{2.2.6}$$

and N is the standard normal cumulative distribution.

Equations (2.2.4) and (2.2.5) are crucial in the deduction of chapter's key results: let us stop and comment. First, let us notice that we can consider option prices with $F_0(t) = 1$ and $B_t = 1$ without any loss of generality: we are interested in the implied volatility and these two quantities cancel out from both sides of the implied volatility equation. Second, let us emphasize that the quantity inside the expected values are, in both equations (2.2.4) and (2.2.5), positive. Finally, it can be useful to mention that a similar result has been obtained by Hull and White (1987) for options on an asset with stochastic volatility.

We can re-write equations (2.2.4) and (2.2.5) w.r.t to the *moneyness degree* y (cf. Introduction)

$$C_t(y\sqrt{t}) = \mathbb{E}\left[e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} N\left(-\frac{y}{\bar{\sigma}\sqrt{S_t}} + l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - e^{y\sqrt{t}} N\left(-\frac{y}{\bar{\sigma}\sqrt{S_t}} + l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right]$$

$$P_t(y\sqrt{t}) = \mathbb{E}\left[e^{y\sqrt{t}}N\left(\frac{y}{\bar{\sigma}\sqrt{S_t}} - l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t}N\left(\frac{y}{\bar{\sigma}\sqrt{S_t}} - l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right] \ ,$$

and we can define $c_t(S_t, y)$ and $p_t(S_t, y)$ such that

$$\begin{aligned}\mathbb{E}[c_t\left(S_t, y\right)] &:= C_t(y\sqrt{t}) \\ \mathbb{E}[p_t\left(S_t, y\right)] &:= P_t(y\sqrt{t}) \ .\end{aligned}$$

Black (1976) option prices w.r.t. y are

$$\begin{aligned}c_t^B(\mathcal{I}_t(y), y) &= N\left(-\frac{y}{\mathcal{I}_t(y)} + \frac{\mathcal{I}_t(y)\sqrt{t}}{2}\right) - e^{y\sqrt{t}}N\left(-\frac{y}{\mathcal{I}_t(y)} - \frac{\mathcal{I}_t(y)\sqrt{t}}{2}\right) \\ p_t^B(\mathcal{I}_t(y), y) &= e^{y\sqrt{t}}N\left(\frac{y}{\mathcal{I}_t(y)} + \frac{\mathcal{I}_t(y)\sqrt{t}}{2}\right) - N\left(\frac{y}{\mathcal{I}_t(y)} - \frac{\mathcal{I}_t(y)\sqrt{t}}{2}\right) \ ,\end{aligned}$$

where $\mathcal{I}_t(y)$ is the implied volatility w.r.t. the *moneyness degree*.
The implied volatility equation for the call options is

$$\mathbb{E}[c_t\left(S_t, y\right)] = c_t^B(\mathcal{I}_t(y), y) \tag{2.2.7}$$

and the one for the put option is

$$\mathbb{E}[p_t\left(S_t, y\right)] = p_t^B(\mathcal{I}_t(y), y) \ . \tag{2.2.8}$$

In the following lemma, we prove that $\hat{\sigma}_t\sqrt{t}$ goes to zero at short-time following an approach similar to Alòs *et al.* (2007, Lemma 6.1, p.580), who considered a generalization of the Bates model.

Lemma 2.2.4.
For the ATS, at short-time,

$$\hat{\sigma}_t\sqrt{t} = o(1) \ .$$

Proof. For an ATM put (i.e. when $y = 0$), the left-hand side of equation (2.2.8) is equal to $\mathbb{E}\left[\left(1 - e^{f_t}\right)^+\right]$. In the region of admissible scaling parameters, f_t goes to zero in distribution because its characteristic function in (2.2.1) goes to one. Hence, by the dominated convergence theorem, $\mathbb{E}\left[\left(1 - e^{f_t}\right)^+\right]$ goes to zero at short time. For $y = 0$, the right-hand side of equation (2.2.8) becomes

$$p_t^B(\hat{\sigma}_t, 0) = N\left(\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) \ ,$$

that goes to zero if and only if $\hat{\sigma}_t\sqrt{t}$ goes to zero. □

Thanks to this lemma, ATM and at short-time, we can rewrite the right hand side of (2.2.7) and (2.2.8) as

$$c_t^B(\hat{\sigma}_t, 0) = p_t^B(\hat{\sigma}_t, 0) = \hat{\sigma}_t\sqrt{\frac{t}{2\pi}} + o\left(\hat{\sigma}_t\sqrt{t}\right) \ , \tag{2.2.9}$$

where the asymptotic expansion holds because $N'(0) = \sqrt{\frac{1}{2\pi}}$, with N' the standard normal probability density function.

2.3 Short-time ATM implied volatility

In this section, we study the behavior of $\hat{\sigma}_t$ at short-time for the ATS. The idea of the proofs is simple. Equation (2.2.9) is the short-time asymptotic expansion of the ATM Black call and put prices. We can study the short-time behavior of the ATS model price in (2.2.7) and (2.2.8).

1. If the model price (left hand side in (2.2.7) and (2.2.8)) goes to zero faster than $\sqrt{t}$, then $\hat{\sigma}_0 = 0$ (Case 1).
2. If the model price goes to zero slower than $\sqrt{t}$, then $\hat{\sigma}_0 = \infty$ (Case 2).
3. If the model price goes to zero as $\sqrt{t}$, then $\hat{\sigma}_0$ is finite (Cases 3, 4, 5).

The idea of the proofs is the following. In Case 1 we bound the model price from above and we prove that it is $o\left(\sqrt{t}\right)$. In Case 2 we bound the model price from below and we show that it converges to zero slower than $\sqrt{t}$. Finally in the remaining Cases we build upper and lower bounds for the model price and prove that both bounds are $O\left(\sqrt{t}\right)$. Furthermore, the proofs are divided into some sub-cases that correspond to particular ranges of the parameters β and δ : we indicate with bold characters the range at the beginning of each sub-case.

Proposition 2.3.1.
For Case 1: $\begin{cases} \beta < 1 \ \ \& \ \ -\min\left(\frac{1}{2}, \beta\right) < \delta \leq 0 \quad or \\ \beta = \delta = 0 \end{cases}$,
the implied volatility is s.t.

$$\hat{\sigma}_0 = 0 \ .$$

Proof.

$$\boldsymbol{\beta < 1 \ \& \ -\min\left(\frac{1}{2}, \beta\right) < \delta \leq 0 \text{ or } \beta = \delta = 0}$$

We bound $c_t\left(S_t, 0\right)$ from above as follows.

$$\begin{aligned} c_t\left(S_t, 0\right) =& N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - \left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} - 1\right) N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) \\ \leq& \sqrt{\frac{t}{2\pi}}\ \bar{\sigma}\ \sqrt{S_t} + e^{\varphi_t t} - 1 \ . \end{aligned} \tag{2.3.1}$$

In the equality we have just added and subtracted the quantity $N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)$. The inequality holds because, by definition of standard normal cumulative distribution function,

$$N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) = \frac{1}{\sqrt{2\pi}}\int_{l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}}^{l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}} dz\ e^{-z^2/2} \leq \sqrt{\frac{t}{2\pi}}\ \bar{\sigma}\sqrt{S_t} \ , \tag{2.3.2}$$

and because we bound from above the product $\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} - 1\right) N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)$ with the (positive) maxima of both factors.
We bound the expected value of $c_t\left(S_t, 0\right)$ as

$$\mathbb{E}[c_t\left(S_t, 0\right)] \leq \mathbb{E}\left[\sqrt{\frac{t}{2\pi}}\bar{\sigma}\sqrt{S_t}\right] + e^{\varphi_t t} - 1 = \sqrt{\frac{t}{2\pi}}\ \bar{\sigma}\ \mathbb{E}[\sqrt{S_t}] + o\left(\sqrt{t}\right) = o\left(\sqrt{t}\right) \ .$$

The first equality holds because $e^{\varphi_t t}-1=O\left(\varphi_t t\right)=o\left(\sqrt{t}\right)$ and the last equality because $\mathbb{E}[\sqrt{S_t}]$ converges to zero at short-time (see **Lemma .4.5**).

Summarizing, the upper bound to the ATS ATM price in (2.2.7) is $o\left(\sqrt{t}\right)$. From (2.2.9) we have that the Black price is $O\left(\hat{\sigma}_t\sqrt{t}\right)$. Thus,

$$\hat{\sigma}_0=0 \qquad \square$$

Proposition 2.3.2.
For Case 2: $-\min\left(\beta,\frac{1-\beta(1-\alpha)}{\alpha}\right)<\delta<-\frac{1}{2}\max\left(\beta,1\right)$,

$$\hat{\sigma}_0=\infty \ .$$

Proof.
We divide the proof in two sub-cases.

$$\boldsymbol{\beta\leq 1 \ \& \ -\beta<\delta<-\frac{1}{2}}$$

Consider the left hand side of equation (2.2.7). We compute the derivative of $c_t(z,y)$ w.r.t. z in $y=0$.

$$\begin{aligned}\frac{\partial c_t(z,0)}{\partial z}=&-t\bar{\sigma}^2\eta_t e^{\varphi_t t-t\bar{\sigma}^2\eta_t z}N\left(l_t^z+\bar{\sigma}\frac{\sqrt{zt}}{2}\right)\\&-\left(\frac{\varphi_t\sqrt{t}}{2\bar{\sigma}z^{3/2}}+\frac{\sqrt{t}\bar{\sigma}\eta_t}{2\sqrt{z}}\right)\left(e^{\varphi_t t-t\bar{\sigma}^2\eta_t z}N'\left(l_t^z+\bar{\sigma}\frac{\sqrt{zt}}{2}\right)-N'\left(l_t^z-\bar{\sigma}\frac{\sqrt{zt}}{2}\right)\right)\\&+\frac{\sqrt{t}\bar{\sigma}}{4\sqrt{z}}\left(e^{\varphi_t t-t\bar{\sigma}^2\eta_t z}N'\left(l_t^z+\bar{\sigma}\frac{\sqrt{zt}}{2}\right)+N'\left(l_t^z-\bar{\sigma}\frac{\sqrt{zt}}{2}\right)\right) \ . \qquad (2.3.3)\end{aligned}$$

At short-time, for a given $z\in\left(0,\frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right)$, $l_t^z=\frac{\sqrt{t}\bar{\sigma}\eta_t}{\sqrt{z}}\left(-z+\frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right)>0$. We observe that $e^{\varphi_t t-t\bar{\sigma}^2\eta_t z}=1+o(1)$ and $\lim_{t\to 0}l_t^z=\infty$ due to **Lemma .4.6** point 1; then, $N\left(l_t^z+\bar{\sigma}\frac{\sqrt{zt}}{2}\right)=1+o(1)$. Thus,

$$\frac{\partial c_t(z,0)}{\partial z}=-t\bar{\sigma}^2\eta_t+o(t\eta_t) \ ,$$

because the first term goes to zero as $t\eta_t$, while the second and the third terms go to zero as $N'(\sqrt{t}\,\eta_t)$ (i.e. as a negative exponential). Thus, for sufficiently small t, $c_t(z,0)$ is decreasing w.r.t. z in $(0,\frac{\varphi_t}{\bar{\sigma}^2\eta_t})$. We emphasize that the right extreme of the interval is increasing to one for sufficiently small t, see **Lemma .4.6** points 2 and 3.
Fix $\tau>0$ and $S^*\in(0,\frac{\varphi_\tau}{\bar{\sigma}^2\eta_\tau})$; for any $t<\tau$

$$\begin{aligned}&\mathbb{E}\left[c_t(S_t,0)\right]\\&\geq c_t(S^*,0)\mathbb{P}\left(S_t\leq S^*\right)\\&\geq\left\{N\left(l_t^{S^*}+\bar{\sigma}\frac{\sqrt{S^*t}}{2}\right)-N\left(l_t^{S^*}-\bar{\sigma}\frac{\sqrt{S^*t}}{2}\right)+\left(\varphi_t t-t\bar{\sigma}^2\eta_t S^*\right)N\left(l_t^{S^*}+\bar{\sigma}\frac{\sqrt{S^*t}}{2}\right)\right\}\mathbb{P}\left(S_t\leq S^*\right)\\&\geq\left(\varphi_t t-t\bar{\sigma}^2\eta_t S^*\right)N\left(l_t^{S^*}+\bar{\sigma}\frac{\sqrt{S^*t}}{2}\right)\mathbb{P}\left(S_t\leq S^*\right)=O(t\eta_t) \ .\end{aligned}$$

The first inequality holds because $c_t(z,0)$ is positive for any $z\geq 0$ and because we bound from below the expected value with its minimum in the interval $(0,S^*)$ multiplied by the probability

of the interval, $\mathbb{P}\left(S_t \leq S^*\right)$. The second inequality is due to the fact that $e^x \geq x+1$. Finally, the last inequality holds because, by definition of the standard normal cumulative distribution function,

$$N\left(l_t^z + \bar{\sigma}\frac{\sqrt{zt}}{2}\right) - N\left(l_t^z - \bar{\sigma}\frac{\sqrt{zt}}{2}\right) \geq 0 \;\;,\;\; z \in \mathbb{R}^+ \;\;. \tag{2.3.4}$$

Recall that $N\left(l_t^{S^*} + \bar{\sigma}\frac{\sqrt{S^*t}}{2}\right) = 1 + o(1)$; notice that $\mathbb{P}\left(S_t \leq S^*\right)$ is constant for $\beta = 1$ and goes to one, by **Lemma .4.4** point 1, for $\beta < 1$. This proves the last equality.
Notice that $t\,\eta_t$ goes to zero slower than $\sqrt{t}$ ($\delta < -0.5$), then the ATM call price goes to zero slower than $\sqrt{t}$.

$$\boldsymbol{\beta > 1 \;\&\; -\frac{1-\beta(1-\alpha)}{\alpha} < \delta < -\frac{\beta}{2}}$$

There exists q such that $(\beta-1)/2 < q < -\delta - 1/2$. We bound the ATM put price (2.2.5) from below for a sufficiently small t

$$\begin{aligned}
&\mathbb{E}[p_t\left(S_t, 0\right)] \\
&\geq \mathbb{E}\left[\mathbb{1}_{S_t \geq 1+t^q}\left(N\left(-l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - N\left(-l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) + N\left(-l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\left(1 - e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t}\right)\right)\right] \\
&\geq \mathbb{E}\left[\mathbb{1}_{S_t \geq 1+t^q} N\left(-l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\left(1 - e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t}\right)\right] \\
&\geq \mathbb{P}(S_t \geq 1 + t^q)\frac{1}{3}\left(1 - e^{\varphi_t t - t\bar{\sigma}^2\eta_t(1+t^q)}\right) =: M_t\left(t^{1+q}\bar{\sigma}^2\eta_t + t\bar{\sigma}^4\eta_t^2 k_t/2\right) \\
&\geq M_t\, t^{1+q}\bar{\sigma}^2\eta_t \;\;.
\end{aligned} \tag{2.3.5}$$

The first inequality holds because $p_t\left(S_t, 0\right)$ is non negative and because we have added and subtracted the term $N\left(-l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)$. The second because the difference between the standard normal cumulative distribution functions is non negative, analogously to (2.3.4). The third because, for $S_t \in [1, \infty)$, $1 - e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t}$ is positive and non decreasing in S_t; moreover, for a sufficiently small t, $N\left(-l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) > 1/3$ because

$$\lim_{t\to 0} N\left(-l_t^z - \bar{\sigma}\frac{\sqrt{zt}}{2}\right) \geq 1/2 \;\;,\;\; z \in [1, \infty) \;\;.$$

The quantity M_t is defined in (2.3.5). At short-time $M_t = 1/6 + o(1)$ because i) by **Lemma .5.3**, $\mathbb{P}(S_t \geq 1 + t^q)$ goes to $1/2$ as t goes to zero, and ii) by **Lemma .4.6** point 1,

$$1 - e^{\varphi_t t - t\bar{\sigma}^2\eta_t(1+t^q)} = \left(t^{1+q}\bar{\sigma}^2\eta_t + t\bar{\sigma}^4\eta_t^2 k_t/2\right)\left(1 + o(1)\right) \;\;.$$

Notice that $t^{1+q}\eta_t$ goes to zero slower than $\sqrt{t}$, then ATM put price goes to zero slower than $\sqrt{t}$.

$$\boldsymbol{\text{Case 2: } -\min\left(\beta, \frac{1-\beta(1-\alpha)}{\alpha}\right) < \delta < -\max\left(\frac{\beta}{2}, \frac{1}{2}\right)}$$

Summing up, for both sub-cases, $\beta \leq 1$ & $-\beta < \delta < -1/2$ and $\beta > 1$ & $-\frac{1-\beta(1-\alpha)}{\alpha} < \delta < -\beta/2$, the lower bounds on the ATM option prices in (2.2.7) and (2.2.8) go to zero slower than $\sqrt{t}$. Moreover, from (2.2.9) we have that the Black price is $O\left(\hat{\sigma}_t\sqrt{t}\right)$. Then,

$$\hat{\sigma}_0 = \infty \qquad \square$$

Proposition 2.3.3.
For Case 3: $\beta \geq 1$ & $\delta \geq -\beta/2$, with the exception of the point $\{\beta = 1, \delta = -1/2\}$,

$$\hat{\sigma}_0 \text{ is finite } .$$

Proof.
We split the proof in three sub-cases. For each sub-case we build an upper and a lower bound, on the model price, and we demonstrate that both bounds are $O\left(\sqrt{t}\right)$ and then, that $\hat{\sigma}_0$ is finite.

$$\boldsymbol{\beta > 1 \ \& -\frac{\beta}{2} <= \delta < -\frac{1}{2}}$$

Upper bound.

Let us split the expected value of the ATS call in two parts

$$\begin{aligned}
&\mathbb{E}[c_t\left(S_t, 0\right)] \\
&= \mathbb{E}\left[N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right] + \mathbb{E}\left[N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} - 1\right)\right] \\
&=: A_1(t) + A_2(t) \ .
\end{aligned}$$

We prove that both parts are bounded from above by quantities $O\left(\sqrt{t}\right)$. The first expected value is s.t.

$$A_1(t) \leq \sqrt{\frac{t}{2\pi}}\,\bar{\sigma}\,\mathbb{E}[\sqrt{S_t}] = O\left(\sqrt{t}\right) \ , \tag{2.3.6}$$

where the inequality holds true because of (2.3.2) and $\sqrt{t}\,\mathbb{E}[\sqrt{S_t}] = O\left(\sqrt{t}\right)$ because, by **Lemma .4.5** point 1, $\mathbb{E}[\sqrt{S_t}]$ goes to one as t goes to zero.
Let us study the term $A_2(t)$.

$$\begin{aligned}
A_2(t) &< \mathbb{E}\left[\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t S_t} - 1\right)\mathbb{1}_{S_t < \varphi_t/(\bar{\sigma}^2\eta_t)}\right] \\
&= \sqrt{\frac{t}{2\pi k_t}}\int_0^{\varphi_t/(\bar{\sigma}^2\eta_t)} dz e^{-\frac{t(z-1)^2}{2k_t}}\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} - 1\right) &(2.3.7)\\
&\quad + \int_0^{\varphi_t/(\bar{\sigma}^2\eta_t)} dz\left(\mathcal{P}_{S_t}(z) - \sqrt{\frac{t}{2\pi k_t}}\,e^{-\frac{t(z-1)^2}{2k_t}}\right)\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} - 1\right) &(2.3.8)\\
&\leq O\left(t^{\delta+(\beta+1)/2}\right) \ ,
\end{aligned}$$

where $\mathcal{P}_{S_t}$ is the law of S_t. The first inequality is true because the quantity inside the expected value is positive on $\left(0, \frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right)$ and negative elsewhere. The equality is obtained by adding and subtrancting the same expected value for a Gaussian random variable. We prove the second inequality in two steps, showing that both (2.3.7) and (2.3.8) are bounded by $O\left(t^{\delta+(\beta+1)/2}\right)$.
First, we consider (2.3.7)

$$\begin{aligned}
&\sqrt{\frac{t}{2\pi k_t}}\int_0^{\varphi_t/(\bar{\sigma}^2\eta_t)} dz\, e^{-\frac{t(z-1)^2}{2k_t}}\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} - 1\right) \\
=&\frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}_t} dw\, e^{-\frac{w^2}{2}}\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t(1+w\sqrt{k_t/t})} - 1\right) &(2.3.9)\\
\leq&\frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}_t} dw\, e^{-\frac{w^2}{2}} e^{-\sqrt{t}\bar{\sigma}^2\eta_t\sqrt{k_t}w} - \frac{1}{\sqrt{2\pi}}\int_{\mathcal{A}_t} dw\, e^{-\frac{w^2}{2}} &(2.3.10)
\end{aligned}$$

$$=e^{t\bar{\sigma}^4\eta_t^2 k_t/2} N\left(\sqrt{t}\bar{\sigma}^2\eta_t\sqrt{k_t} + \left(\frac{\varphi_t}{\bar{\sigma}^2\eta_t} - 1\right)\sqrt{\frac{t}{k_t}}\right) - N\left(\left(\frac{\varphi_t}{\bar{\sigma}^2\eta_t} - 1\right)\sqrt{\frac{t}{k_t}}\right) + o\left(t^{\delta+(\beta+1)/2}\right) \tag{2.3.11}$$

$$=\sqrt{\frac{t}{2\pi}}\bar{\sigma}^2\eta_t\sqrt{k_t} + \frac{t\bar{\sigma}^4\eta_t^2 k_t}{4} + o\left(t^{\delta+(\beta+1)/2}\right) = O\left(t^{\delta+(\beta+1)/2}\right) \ , \tag{2.3.12}$$

where $\mathcal{A}_t \equiv \left\{w \in \mathbb{R} : -\sqrt{\frac{t}{k_t}} < w < \left(\varphi_t/(\bar{\sigma}^2\eta_t) - 1\right)\sqrt{\frac{t}{k_t}}\right\}$. Equality (2.3.9) is due to a change of the integration variable $w := (z-1)/\sqrt{k_t/t}$, equality (2.3.10) to the fact that, by **Lemma .4.6**, $e^{\varphi_t t - t\bar{\sigma}^2\eta_t} < 1$. Equality (2.3.11) to a change of variable $m := w + \sqrt{t}\bar{\sigma}^2\eta_t\sqrt{k_t}$ and to the fact that both $N\left(-\sqrt{\frac{t}{k_t}}\right)$ and $N\left(\sqrt{t}\bar{\sigma}^2\eta_t\sqrt{k_t} - \sqrt{\frac{t}{k_t}}\right)$ go to zero faster that any power of t. Finally, (2.3.12) holds true because of the Taylor expansion of N in zero.

Second, we consider (2.3.8)

$$\begin{aligned}
&\left|\int_0^{\varphi_t/(\bar{\sigma}^2\eta_t)} dz \left(\mathcal{P}_{S_t}(z) - \sqrt{\frac{t}{2\pi k_t}}\, e^{-\frac{t(z-1)^2}{2k_t}}\right)\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} - 1\right)\right| \\
\leq &\left|-\left(\mathbb{P}(S_t < 0) - N\left(-\sqrt{\frac{t}{k_t}}\right)\right)\left(e^{\varphi_t t} - 1\right)\right| \\
&+ \left|\int_0^{\varphi_t/(\bar{\sigma}^2\eta_t)} dz \left(\mathbb{P}(S_t < z) - N\left((z-1)\sqrt{\frac{t}{k_t}}\right)\right)\bar{\sigma}^2\,\eta_t\, t\, e^{\varphi_t t - t\bar{\sigma}^2\eta_t z}\right| \\
\leq &2\frac{2-\alpha}{1-\alpha}\sqrt{\frac{k_t}{t}}\left(e^{\varphi_t t} - 1\right) = O\left(t^{\delta+(\beta+1)/2}\right) \ .
\end{aligned}$$

The first inequality is due to integration by part and to the triangular inequality. The second inequality is a consequence of Jensen inequality and of **Lemma .5.3**.

Lower bound.

As discussed in the proof of **Proposition 2.3.2**, for a sufficiently small t, $c_t(S_t, 0)$ is decreasing for $S_t \in \left(0, \frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right)$ hence,

$$\begin{aligned}
\mathbb{E}\left[c_t(S_t,0)\right] &\geq \mathbb{E}\left[c_t(S_t,0)\mathbb{1}_{S_t \leq \varphi_t/(\bar{\sigma}^2\eta_t)}\right] \\
&\geq c_t\left(\frac{\varphi_t}{\bar{\sigma}^2\eta_t}, 0\right)\mathbb{P}\left(S_t \leq \frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right) = \sqrt{\frac{\varphi_t t}{8\pi\eta_t}} + o\left(\sqrt{t}\right) = O\left(\sqrt{t}\right) \ .
\end{aligned}$$

The first inequality is because $c_t(S_t, 0)$ is non negative and the second is because we bound the expected value from below with the minimum of $c_t(S_t, 0)$ multiplied by the probability of the interval $\left(0, \frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right)$. The equality holds because, by **Lemma .5.3**,

$$\lim_{t\to 0}\mathbb{P}\left(S_t \leq \frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right) = \frac{1}{2} \ ,$$

and

$$c_t\left(\frac{\varphi_t}{\bar{\sigma}^2\eta_t}, 0\right) = N\left(\sqrt{\frac{\varphi_t\, t}{4\bar{\sigma}\eta_t}}\right) - N\left(-\sqrt{\frac{\varphi_t\, t}{4\bar{\sigma}\eta_t}}\right) = \sqrt{\frac{\varphi_t\, t}{2\pi\bar{\sigma}\eta_t}} + o\left(\sqrt{t}\right) \ ,$$

with $\sqrt{\frac{\varphi_t t}{8\pi\bar{\sigma}\eta_t}} = O\left(\sqrt{t}\right)$.

$$\boldsymbol{\beta > 1 \ \& \ \delta = -\frac{1}{2}}$$

Upper bound.

The upper bound on the ATS call price is the same to the one of the previous sub-case $-\beta/2 \leq \delta < -1/2,\ \beta > 1$.

Lower bound.

We bound the put price from below. It exist $H > 1$ such that for a sufficiently small t

$$\begin{aligned}\mathbb{E}\left[p_t\left(S_t,0\right)\right] &\geq \mathbb{E}\left[p_t\left(S_t,0\right)\mathbb{1}_{S_t\in[1,H]}\right] \geq \mathbb{E}\left[\left(N\left(-l_t^{S_t}+\bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)-N\left(-l_t^{S_t}-\bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right)\mathbb{1}_{S_t\in[1,H]}\right] \\ &\geq \mathbb{E}\left[N'\left(-l_t^{S_t}+\frac{\bar{\sigma}\sqrt{S_t t}}{2}\right)\bar{\sigma}\sqrt{S_t t}\,\mathbb{1}_{S_t\in[1,H]}\right] \\ &\geq N'\left(\bar{\sigma}\bar{\eta}-\frac{\varphi_t\sqrt{t}}{\bar{\sigma}}+\frac{\bar{\sigma}\sqrt{t}}{2}\right)\bar{\sigma}\sqrt{t}\,\mathbb{P}(S_t\in[1,H]) = \sqrt{\frac{t}{8\pi}}\bar{\sigma}+o\left(\sqrt{t}\right)\ .\end{aligned}$$

The first inequality holds because $p_t(S_t,0)$ is non negative. The second because $e^{\varphi_t t-t\bar{\sigma}^2\eta_t S_t} < 1$ in $[1,H]$. The third inequality is due to the fact that we bound from above the difference

$$N\left(-l_t^{S_t}+\bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)-N\left(-l_t^{S_t}-\bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)$$

with the standard normal law evaluated in the maximum between the two (positive) arguments multiplied by the difference of the two arguments. Notice that $\bar{\sigma}\sqrt{S_t t}$ is a positive quantity almost surely. The last inequality holds because, by **Lemma .5.4**, it exists $H > 1$ s.t. the quantity inside the expected value is increasing in $[1,H]$ for a sufficiently small t. The equality is because, by **Lemma .5.3**, $\mathbb{P}(S_t\in[1,H])$ goes to $1/2$ as t goes to zero and

$$\lim_{t\to 0} N'\left(\bar{\sigma}\bar{\eta}-\frac{\varphi_t\sqrt{t}}{\bar{\sigma}}+\frac{\bar{\sigma}\sqrt{t}}{2}\right) = \frac{1}{\sqrt{2\pi}}\ .$$

$$\boldsymbol{\beta \geq 1 \ \& \ -\frac{1}{2} < \delta \leq 0}$$

Upper bound.

We can bound $c_t(S_t,0)$ from above as in (2.3.1).
We bound the ATS option price as

$$\mathbb{E}[c_t(S_t,0)] \leq \mathbb{E}\left[\frac{1}{\sqrt{2\pi}}\bar{\sigma}\sqrt{S_t t}\right]+e^{\varphi_t t}-1 \leq O\left(\sqrt{t}\right)\ .$$

The last inequality holds because, by Jensen inequality with concave function $\sqrt{*}$, $\mathbb{E}[\sqrt{S_t}] \leq \sqrt{\mathbb{E}[S_t]} = 1$ and because, by **Lemma .4.6** point 1, $e^{\varphi_t t}-1 = o\left(\sqrt{t}\right)$.

Lower bound.

To bound $c_t(z,0)$ from below we have to study its derivative in (2.3.3). Notice that, at short-time, $l_t^z = O\left(\sqrt{t}\eta_t\right) = o(1)$, due to **Lemma .4.6** point 1, and to the fact that $\delta > -1/2$. Moreover, again due to **Lemma .4.6** point 1, $e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} = 1 + O(t\eta_t)$. Then, we have
i) The negative first term at short-time is $o\left(\sqrt{t}\right)$

$$-t\bar{\sigma}^2\eta_t e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} N\left(l_t^z + \bar{\sigma}\frac{\sqrt{zt}}{2}\right) = O(t\eta_t) = o\left(\sqrt{t}\right) \quad .$$

ii) The second term at short-time is $o\left(\sqrt{t}\right)$

$$\begin{aligned}
&\left(\frac{\varphi_t\sqrt{t}}{2\bar{\sigma}z^{3/2}} + \frac{\sqrt{t}\bar{\sigma}\eta_t}{2\sqrt{z}}\right)\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} N'\left(l_t^z + \bar{\sigma}\frac{\sqrt{zt}}{2}\right) - N'\left(l_t^z - \bar{\sigma}\frac{\sqrt{zt}}{2}\right)\right) \\
=&O\left(\sqrt{t}\eta_t\right)\frac{e^{-(l_t^z)^2/2 - \bar{\sigma}^2 zt/8}}{\sqrt{2\pi}}\left((1+O(t\eta_t))\left(1 - \frac{l_t^z\bar{\sigma}\sqrt{zt}}{2} + o(t\eta_t)\right) - \left(1 + \frac{l_t^z\bar{\sigma}\sqrt{zt}}{2} + o(t\eta_t)\right)\right) \\
=&O(\eta_t^2 t^{3/2}) = o\left(\sqrt{t}\right) \quad ,
\end{aligned}$$

because

$$N'\left(l_t^z \pm \bar{\sigma}\frac{\sqrt{zt}}{2}\right) = e^{-(l_t^z)^2/2 - \bar{\sigma}^2 zt/8}\left(1 \pm \frac{l_t^z\bar{\sigma}\sqrt{zt}}{2} + o(t\eta_t)\right) \quad .$$

iii) The positive third term at short-time is $O\left(\sqrt{t}\right)$

$$\frac{\sqrt{t}\bar{\sigma}}{4\sqrt{z}}\left(e^{\varphi_t t - t\bar{\sigma}^2\eta_t z} N'\left(l_t^z + \bar{\sigma}\frac{\sqrt{zt}}{2}\right) + N'\left(l_t^z - \bar{\sigma}\frac{\sqrt{zt}}{2}\right)\right) = \sqrt{\frac{t}{8\pi\, z}}\,\bar{\sigma} + o\left(\sqrt{t}\right) \quad .$$

Summarizing, the leading term in (2.3.3), at short-time, is the third one, which is positive. Hence, for a fixed $z > 0$ and for sufficiently small t, $c_t(z,0)$ is increasing; thus, we can bound the expected value from below

$$\begin{aligned}
\mathbb{E}\left[c_t(S_t,0)\right] \geq \mathbb{E}\left[c_t(S_t,0)\mathbb{1}_{S_t\in[1/2,3/2]}\right] &> c_t\left(\frac{1}{2},0\right)\mathbb{P}\left(S_t \in \left[\frac{1}{2},\frac{3}{2}\right]\right) \\
&> \left\{N\left(l_t^{1/2} + \bar{\sigma}\sqrt{\frac{t}{8}}\right) - N\left(l_t^{1/2} - \bar{\sigma}\sqrt{\frac{t}{8}}\right)\right\}\mathbb{P}\left(S_t \in \left[\frac{1}{2},\frac{3}{2}\right]\right) \\
&> N'\left(l_t^{1/2} + \bar{\sigma}\sqrt{\frac{t}{8}}\right)\bar{\sigma}\sqrt{\frac{t}{2}}\mathbb{P}\left(S_t \in \left[\frac{1}{2},\frac{3}{2}\right]\right) \\
&= \left(\bar{\sigma}\sqrt{\frac{t}{4\pi}} + o\left(\sqrt{t}\right)\right)\mathbb{P}\left(S_t \in \left[\frac{1}{2},\frac{3}{2}\right]\right) = O\left(\sqrt{t}\right) \quad .
\end{aligned}$$

The first inequality holds because $c_t(S_t,0)$ is non negative. The second because, for a sufficiently small t, $c_t(S_t,0)$ is increasing. The third is true because, for sufficiently small t, $e^{\varphi_t t - t\bar{\sigma}^2\eta_t/2} > 1$, by **Lemma .4.6** point 3. The forth is due to the fact that the difference of the standard normal cumulative distribution functions can be bounded from below by the (positive) maximum of the two arguments multiplied by the (positive) difference of the two arguments. The equality is due to the fact that $\mathbb{P}\left(S_t \in \left[\frac{1}{2},\frac{3}{2}\right]\right)$ is constant if $\beta = 1$ and goes to 1 at short-time if $\beta > 1$ because, by **Lemma .4.4** point 2, S_t goes to one in distribution at short-time.

Case 3: $\beta \geq 1$ & $-\frac{\beta}{2} \leq \delta \leq 0 \setminus \beta = 1$, $\left\{\delta = -\frac{1}{2}\right\}$

Summing up, in all sub-cases the upper bound and the lower bounds of the ATS option prices in (2.2.7) and (2.2.8) are $O(\sqrt{t})$. Moreover, from (2.2.9) we have that the Black price is $O\left(\hat{\sigma}_t\sqrt{t}\right)$. Thus,

$$\hat{\sigma}_0 \text{ is finite} \qquad \square$$

Proposition 2.3.4.
For Cases 4 and 5: $\beta \leq 1$ & $\delta = -\frac{1}{2}$,

$$\hat{\sigma}_0 \textit{ is finite } .$$

Proof.

$$\boldsymbol{\beta \leq 1 \ \& \ \delta = -\frac{1}{2}}$$

Upper bound.

We can bound $c_t(S_t, 0)$ from above as in (2.3.1).
We bound the ATS option price as

$$\mathbb{E}[c_t(S_t,0)] \leq \mathbb{E}\left[\frac{1}{\sqrt{2\pi}}\bar{\sigma}\sqrt{S_t t}\right] + e^{\varphi_t t} - 1 = O\left(\sqrt{t}\right) \ .$$

The equality holds because, by Jensen inequality with concave function $\sqrt{*}$, $\mathbb{E}[\sqrt{S_t}] \leq \sqrt{\mathbb{E}[S_t]} = 1$ and because, by **Lemma .4.6** point 1, $e^{\varphi_t t} - 1 = O\left(\sqrt{t}\right)$.

Lower bound.

We bound $c_t\left(S_t, 0\right)$ from below as:

$$\begin{aligned} c_t\left(S_t,0\right) &\geq \mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}\left(N\left(l_t^{S_t} + \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) - N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right) + \left(\varphi_t t - t\bar{\sigma}^2\eta_t S_t\right)/2\right) \\ &\geq \mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}\left(\varphi_t t - t\bar{\sigma}^2\eta_t S_t\right)/2 \ . \end{aligned}$$

The first inequality is because $c_t\left(S_t, 0\right)$ is non negative, because $e^x \geq x + 1$, and because the normal cumulative distribution function evaluated in a positive quantity is above 1/2. The second holds because the difference between the two normal cumulative function is non negative.

$$\begin{aligned} \mathbb{E}[c_t\left(S_t,0\right)] &\geq \mathbb{E}\left[\mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}(\varphi_t t - t\bar{\sigma}^2\eta_t S_t)/2\right] \\ &= \sqrt{t}\,\bar{\sigma}^2\,\bar{\eta}/2\ \mathbb{E}\left[\mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}(-S_t + \varphi_t/(\bar{\sigma}^2\eta_t))\right] = O\left(\sqrt{t}\right). \end{aligned}$$

The last equality is due to the fact that

$$\mathbb{E}\left[\mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}(-S_t + \varphi_t/(\bar{\sigma}^2\eta_t))\right] = \varphi_t/(\bar{\sigma}^2\eta_t)\mathbb{P}(S_t < \varphi_t/(\bar{\sigma}^2\eta_t)) - \mathbb{E}[S_t\mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}] \quad (2.3.13)$$

can be bounded from below with a positive constant for sufficiently small t. This fact can be deduced for $\beta \leq 1$. We prove it separately for the two cases $\beta < 1$ and $\beta = 1$.
For $\beta < 1$, let us observe that, at short-time,

$$0 \leq \mathbb{E}[S_t\mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}] \leq \mathbb{E}[S_t\mathbb{1}_{S_t<1}] = o(1) \ ,$$

because, by point 2 of **Lemma .4.6**, $\varphi_t/(\bar{\sigma}^2\eta_t) < 1$ and, by definition of convergence in distribution, at short-time $\mathbb{E}[S_t \mathbb{1}_{S_t<1}] = o(1)$, because, by **Lemma .4.4** point 1, S_t converges in distribution to 0. Moreover, at short-time, $\varphi_t/(\bar{\sigma}^2\eta_t)\mathbb{P}(S_t < \varphi_t/(\bar{\sigma}^2\eta_t)) = 1 + o(1)$, by point 1 of **Lemma .4.6** and by point 1 of **Lemma .4.4**.
For $\beta = 1$, we remind that the law of S_t does not depend from t and we observe that the limit of (2.3.13) for t that goes to zero is positive

$$\lim_{t\to 0} \left\{\varphi_t/(\bar{\sigma}^2\eta_t)\mathbb{P}(S_t < \varphi_t/(\bar{\sigma}^2\eta_t)) - \mathbb{E}[S_t \mathbb{1}_{S_t<\varphi_t/(\bar{\sigma}^2\eta_t)}]\right\} = \mathbb{P}(S_t < 1) - \mathbb{E}[S_t \mathbb{1}_{S_t<1}] > 0 \ ,$$

where the last inequality is due to the fact that S_t has unitary mean and finite variance $\bar{k}$.

Summarizing, as in **Proposition 2.3.3** the upper and lower bounds of the ATM prices in (2.2.7) are $O\left(\sqrt{t}\right)$. From (2.2.9), we have that the Black price is $O\left(\hat{\sigma}_t\sqrt{t}\right)$. Thus,

$$\hat{\sigma}_0 \ \text{ is finite} \qquad \square$$

In the propositions above, we have proven that $\hat{\sigma}_0$ is finite only in Cases 3, 4 and 5. Only for these Cases we study the short-time skew in the next section.

2.4 Short-time skew

In this section, we focus on the skew term $\hat{\xi}_t$ for the ATS when $\hat{\sigma}_0$ is finite. We obtain an expression of $\hat{\xi}_t$ in **Lemma 2.4.1** and study its short-time limit.
In the introduction, we have mentioned that the implied volatility skew observed in the equity market is negative and it goes to zero as one over the square root of t. This behavior is equivalent to a negative and finite $\hat{\xi}_0$. In this section, we prove that $\hat{\xi}_0$ is zero in Case 3 (**Proposition 2.4.2**) and is negative and finite in Cases 4 and 5 (**Proposition 2.4.3**). Moreover, Case 5 identifies the unique parameters' set where $\hat{\xi}_0$ can be a generic value that it is possible to calibrate from market data.

Lemma 2.4.1. *The skew term $\hat{\xi}_t$ is*

$$\hat{\xi}_t = \frac{N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - \mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right]}{N'\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right)} \ . \tag{2.4.1}$$

Proof. Applying the implicit function theorem to the implied volatility equation for the call option (2.2.7) we obtain the derivative of the implied volatility w.r.t y

$$\frac{\partial \mathcal{I}_t(y)}{\partial y} = \frac{\frac{\partial \mathbb{E}[c_t(S_t,y)]}{\partial y} - \frac{\partial c_t^B(\mathcal{I}_t(y),y)}{\partial y}}{\frac{\partial c_t^B(\mathcal{I}_t(y),y)}{\partial \mathcal{I}_t(y)}} \ .$$

We prove the thesis by computing the three partial derivatives separately.

$$\frac{\partial \mathbb{E}\left[c_t\left(S_t,y\right)\right]}{\partial y} = -\sqrt{t}e^{\sqrt{t}\,y}\mathbb{E}\left[N\left(-\frac{y}{\bar{\sigma}\sqrt{S_t}} + l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right]$$
$$\frac{\partial c_t^B\left(\mathcal{I}_t(y),y\right)}{\partial y} = -\sqrt{t}e^{\sqrt{t}\,y}N\left(-\frac{y}{\mathcal{I}_t(y)} - \frac{\mathcal{I}_t(y)\sqrt{t}}{2}\right)$$

$$\frac{\partial c_t^B\left(\mathcal{I}_t(y),y\right)}{\partial \mathcal{I}_t(y)} = \sqrt{t}e^{\sqrt{t}\,y}N'\left(-\frac{y}{\mathcal{I}_t(y)} - \frac{\mathcal{I}_t(y)\sqrt{t}}{2}\right) \quad .$$

Notice that it is possible to exchange the expected value w.r.t. S_t and the derivative w.r.t. y using the Leibniz rule because the law of S_t does not depend from y. By substituting $y = 0$ and reminding that $\mathcal{I}_t(0) = \hat{\sigma}_t$, we get (2.4.1) □

Notice that, because of **Lemma 2.2.4**, the denominator of $\hat{\xi}_t$ in (2.4.1), $N'\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right)$, goes to $\frac{1}{\sqrt{2\pi}}$ at short-time. To study the short-time behavior of $\hat{\xi}_t$ it is sufficient to consider only the numerator of equation (2.4.1)

$$N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - \mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right] \quad .$$

Proposition 2.4.2.
For Case 3: $\beta \geq 1$ & $-\beta/2 \leq \delta \leq 0$, with the exception of the point $\{\beta = 1,\ \delta = -1/2\}$, the skew term is

$$\hat{\xi}_0 = 0 \quad .$$

Proof.
We divide the proof in two sub-cases.

$$\boldsymbol{\beta = 1\ \&\ -\frac{1}{2} < \delta \leq 0}$$

We study the numerator of $\hat{\xi}_t$ in (2.4.1).

$$\lim_{t\to 0}\left\{N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - \mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right] = 0\right\} \quad .$$

We compute the limit thanks to the dominated convergence theorem because the law of S_t does not depend on t and $l_t^z = o(1)$ in this sub-case.

$$\boldsymbol{\beta > 1\ \&\ -\frac{\beta}{2} \leq \delta \leq 0}$$

We want to prove that

$$\mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right] = \frac{1}{2} + o\left(1\right) \quad . \tag{2.4.2}$$

The equality holds because

$$\begin{aligned}
&\mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right] \\
=&\sqrt{\frac{t}{2\pi k_t}}\int_0^\infty dz\, e^{-t\frac{(z-1)^2}{2k_t}} N\left(l_t^z - \bar{\sigma}\frac{\sqrt{zt}}{2}\right) &(2.4.3)\\
&+\int_0^\infty dz\left(\mathcal{P}_{S_t}(z) - \sqrt{\frac{t}{2\pi k_t}}\, e^{-t\frac{(z-1)^2}{2k_t}}\right) N\left(l_t^z - \bar{\sigma}\frac{\sqrt{zt}}{2}\right) \quad , &(2.4.4)
\end{aligned}$$

where $\mathcal{P}_{S_t}$ is the distribution of S_t. We study the quantities in (2.4.3) and (2.4.4) separately. First, we consider (2.4.3)

$$
\begin{aligned}
&\lim_{t\to 0}\sqrt{\frac{t}{2\pi k_t}}\int_0^\infty dz\; e^{-t\frac{(z-1)^2}{2k_t}} N\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\\
=&\lim_{t\to 0}\frac{1}{\sqrt{2\pi}}\int_{-\sqrt{\frac{k_t}{t}}}^\infty dw\; e^{-\frac{w^2}{2}} N\left(\frac{\varphi_t\sqrt{t}}{\bar\sigma\sqrt{1+w\sqrt{k_t/t}}}-\bar\sigma\eta_t\sqrt{t\left(1+w\sqrt{k_t/t}\right)}-\bar\sigma\frac{\sqrt{t\left(1+w\sqrt{k_t/t}\right)}}{2}\right)\\
=&\lim_{t\to 0}\frac{1}{\sqrt{2\pi}}\int_{-\sqrt{\frac{k_t}{t}}}^\infty dw\; e^{-\frac{w^2}{2}} N\left(\bar\sigma\eta_t\sqrt{t}\left(1-w\sqrt{k_t/t}/2\right)-\bar\sigma\eta_t\sqrt{t}\left(1+w\sqrt{k_t/t}/2\right)+O\left(\sqrt{t}\right)\right)\\
=&\frac{1}{\sqrt{2\pi}}\lim_{t\to 0}\int_{\mathbb{R}} dw\; e^{-\frac{w^2}{2}}\left\{N\left(-\bar\sigma\bar\eta\sqrt{\bar k}t^{\delta+\beta/2}w\right)-\frac{1}{2}+\frac{1}{2}\right\}=\frac{1}{2}\ .
\end{aligned}
$$

The first equality is obtained via a change of the integration variable ($w := \sqrt{t}(z-1)/\sqrt{k_t}$). The second equality is due to the asymptotic of $\varphi_t t$ in **Lemma .4.6** point 1. The third equality holds because of the dominated convergence theorem. The last is trivial because $\left[N\left(-\bar\eta\sqrt{\bar k}t^{\delta+\beta/2}w\right)-1/2\right]$ is odd w.r.t. w.
Second, we consider (2.4.4)

$$
\begin{aligned}
&\int_0^\infty dz\left(\mathcal{P}_{S_t}(z)-\sqrt{\frac{t}{2\pi k_t}}\,e^{-t\frac{(z-1)^2}{2k_t}}\right)N\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\\
=&\left(\mathbb{P}(S_t<0))-N\left(-\sqrt{\frac{t}{k_t}}\right)\right)\\
&+\int_0^\infty dz\left(\mathbb{P}(S_t<z)-N\left((z-1)\sqrt{\frac{t}{k_t}}\right)\right)N'\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\left(\frac{\varphi_t\sqrt{t}}{2\bar\sigma\, z^{3/2}}+\frac{\bar\sigma\sqrt{t}\eta_t+\bar\sigma\sqrt{t}/2}{2\sqrt{z}}\right)=o(1)\ .
\end{aligned}
$$

The first equality is due to integration by part. The second to the fact that i) $\mathbb{P}(S_t<0)=0$, ii) $N\left(-\sqrt{\frac{t}{k_t}}\right)$ go to zero as t goes to zero, and iii)

$$
\begin{aligned}
&\left|\int_0^\infty dz\left(N\left((z-1)\sqrt{\frac{t}{k_t}}\right)-\mathbb{P}(S_t<z)\right)N'\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\left(\frac{\varphi_t\sqrt{t}}{2\bar\sigma\, z^{3/2}}+\frac{\bar\sigma\sqrt{t}\eta_t+\bar\sigma\sqrt{t}/2}{2\sqrt{z}}\right)\right|\\
\le&\frac{2-\alpha}{1-\alpha}\sqrt{\frac{k_t}{t}}\int_0^\infty dz N'\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\left(\frac{\varphi_t\sqrt{t}}{2\bar\sigma\, z^{3/2}}+\frac{\bar\sigma\sqrt{t}\eta_t+\bar\sigma\sqrt{t}/2}{2\sqrt{z}}\right)\\
=&\frac{2-\alpha}{1-\alpha}\sqrt{\frac{k_t}{t}}=O\left(\sqrt{\frac{k_t}{t}}\right)\ ,
\end{aligned}
$$

where the inequality is due to **Lemma .5.3** and the first equality is due the fact that

$$
\int_0^\infty dz N'\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\left(\frac{\varphi_t\sqrt{t}}{2\bar\sigma\, z^{3/2}}+\frac{\bar\sigma\sqrt{t}\eta_t+\bar\sigma\sqrt{t}/2}{2\sqrt{z}}\right)=-\left.N\left(l_t^z-\bar\sigma\frac{\sqrt{zt}}{2}\right)\right|_0^\infty=1\ .
$$

This proves (2.4.2).
It is now possible to compute the short-time limit of the skew term

$$
\lim_{t\to 0}\left\{\left(N\left(-\frac{\hat\sigma_t\sqrt{t}}{2}\right)-\mathbb{E}\left[N\left(l_t^{S_t}-\bar\sigma\frac{\sqrt{S_t t}}{2}\right)\right)\right]\right\}=0 \qquad \square
$$

Proposition 2.4.3.
For Case 4: $\beta < 1$ and $\delta = -1/2$, the skew term is

$$\hat{\xi}_0 = -\sqrt{\frac{\pi}{2}} \ .$$

For Case 5: $\beta = 1$ and $\delta = -1/2$ the skew term is

$$\hat{\xi}_0 = -\sqrt{\frac{\pi}{2}}\, \mathbb{E}\left[erf\left(\bar{\sigma}\bar{\eta}\, r(S_t)\right)\right] \ , \tag{2.4.5}$$

where $r(S_t) := \sqrt{2}(1/\sqrt{S_t} - \sqrt{S_t})$.

Proof.
We prove separately the two Cases.

$$\boldsymbol{\delta = -\frac{1}{2}\ \&\ \beta < 1}$$

Thanks to **Lemma .5.2**, the limit of the numerator of $\hat{\xi}_t$ in (2.4.1) can be computed simply,

$$\lim_{t\to 0}\left(N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - \mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right]\right) = -\frac{1}{2} \ .$$

Thus,

$$\hat{\xi}_0 = \lim_{t\to 0}\frac{N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - \mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right]}{N'\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right)} = -\sqrt{\frac{\pi}{2}} \ .$$

$$\boldsymbol{\delta = -\frac{1}{2}\ \&\ \beta = 1}$$

We compute the limit in $t = 0$ of the numerator of $\hat{\xi}_t$ in (2.4.1)

$$\lim_{t\to 0}\left(N\left(-\frac{\hat{\sigma}_t\sqrt{t}}{2}\right) - \mathbb{E}\left[N\left(l_t^{S_t} - \bar{\sigma}\frac{\sqrt{S_t t}}{2}\right)\right]\right) = \mathbb{E}\left[1/2 - N\left(\bar{\sigma}\bar{\eta}\left(1/\sqrt{S_t} - \sqrt{S_t}\right)\right)\right] \ .$$

We obtain the equality thanks to the dominated convergence theorem, because the law of S_t is constant in time. We recall that $erf(z) = 2N(z/\sqrt{2}) - 1$, substituting in (2.4.1), we obtain (2.4.5) □

Equation (2.4.5) is one of the major results of the thesis. Let us stop and comment.
First, let us notice that $\hat{\xi}_0$ in (2.4.5), is a generic function of the couple of positive parameters $\bar{\sigma}\bar{\eta}$ and $\bar{k}$; in particular the erf function is odd in its argument and $r : \mathbb{R}^+ \to \mathbb{R}$. Moreover $\hat{\xi}_0$ depends on the parameter $\alpha \in [0, 1)$ that selects the truncated additive process of interest.
Second,

$$-\sqrt{\frac{\pi}{2}} \le \hat{\xi}_0 \le 0$$

i.e. the minimum value for the skew term is $-\sqrt{\pi/2}$, its value in Case 4.
To show the upper bound, we can rewrite

$$\mathbb{E}\left[erf\left(\bar{\sigma}\bar{\eta}\, d(S_t)\right)\right] = \int_0^\infty dz\, \mathcal{P}_{S_t}(z)\, erf\left(\bar{\sigma}\bar{\eta}\, r(z)\right) = \int_0^1 dz\, \left(\mathcal{P}_{S_t}(z) - \frac{\mathcal{P}_{S_t}(z)}{z^2}\right) erf\left(\bar{\sigma}\bar{\eta}\, r(z)\right) \ ,$$

where the second equality is due to the change of variable $w = 1/z$, and second to $r(1/w) = -r(w)$ and to the fact that $erf(z)$ is odd. We also observe that $erf\left(\bar{\sigma}\bar{\eta}\, r(z)\right) > 0$ in $(0,1)$.
For the two cases where the distribution of S_t is known analytically $\alpha = 0$ (VG) and $\alpha = 1/2$ (NIG), we can prove that the skew term $\hat{\xi}_0$ in (2.4.5) is negative for non zero $\bar{\sigma}\bar{\eta}$ and $\bar{k}$ (for the expression of the Gamma and Inverse Gaussian laws see, e.g., Cont and Tankov 2003, Ch.4, p.128). In both cases we can prove that $\left(\mathcal{P}_{S_t}(z) - \frac{\mathcal{P}_{S_t}(1/z)}{z^2}\right) > 0$ in $(0,1)$; recall that $\mathcal{P}_{S_t}(z)$ does not depend from time because $\beta = 1$.
In the $\alpha = 0$ case, S_t has the law of a Gamma random variable

$$\mathcal{P}_{S_t}(z) - \frac{\mathcal{P}_{S_t}(1/z)}{z^2} = \frac{1}{\bar{k}^{1/\bar{k}}\Gamma(1/\bar{k})} z^{1/\bar{k}} e^{-z/\bar{k}} \left(1 - \frac{e^{-1/\bar{k}\,(1/z - z)}}{z^{2/\bar{k}}}\right) > 0 \ ,$$

where the inequality is true in $(0,1)$ because $1 - \frac{e^{-1/\bar{k}\,(1/z-z)}}{z^{2/\bar{k}}} > 0$ or equivalently $1/z - z + 2\ln z > 0$. The last inequality is trivial $\forall z \in (0,1)$, because it is equal to zero for $z = 1$ and its derivative is negative.
In the $\alpha = 1/2$ case, S_t has the law of an Inverse Gaussian random variable

$$\mathcal{P}_{S_t}(z) - \frac{\mathcal{P}_{S_t}(1/z)}{z^2} = \frac{1}{\sqrt{2\pi\bar{k}}} e^{-r(z)^2/(2\bar{k})} \left(\frac{1}{z^{3/2}} - \frac{1}{\sqrt{z}}\right) > 0 \ ,$$

where the inequality is true because $\frac{1}{z^{3/2}} - \frac{1}{\sqrt{z}} > 0, \forall z \in (0,1)$.

In all other cases, we compute numerically the skew term $\hat{\xi}_0$ for different admissible values of $\bar{k}, \bar{\sigma}\bar{\eta} \in \mathbb{R}^+$ and $\alpha \in [0,1)$, by means of inversion of the characteristic function of S_t, showing that it is either negative or equal to zero. In Figure 2.3, we plot the numerical estimation of the skew term for $\bar{\sigma}\bar{\eta}$ and $\bar{k}$ below 3 (an interval in line with the situation generally observed in market data) and for a grid of four values of α ($\alpha = 0, 1/4, 1/2, 3/4$); in all cases the skew term $\hat{\xi}_0$ looks rather similar: equal to zero on the boundaries ($\bar{k} = 0$ and $\bar{\sigma}\bar{\eta} = 0$), a negative quantity in all other cases and a decreasing function w.r.t. both $\bar{k}$ and $\bar{\sigma}\bar{\eta}$. In Figure 2.4, we plot also the skew term for the same four values of α, varying $\bar{k}$ with $\bar{\sigma}\bar{\eta} = 1$ (on the left) and varying $\bar{\sigma}\bar{\eta}$ for $\bar{k} = 1$ (on the right): all plots look rather similar with a decreasing $\hat{\xi}_0$.

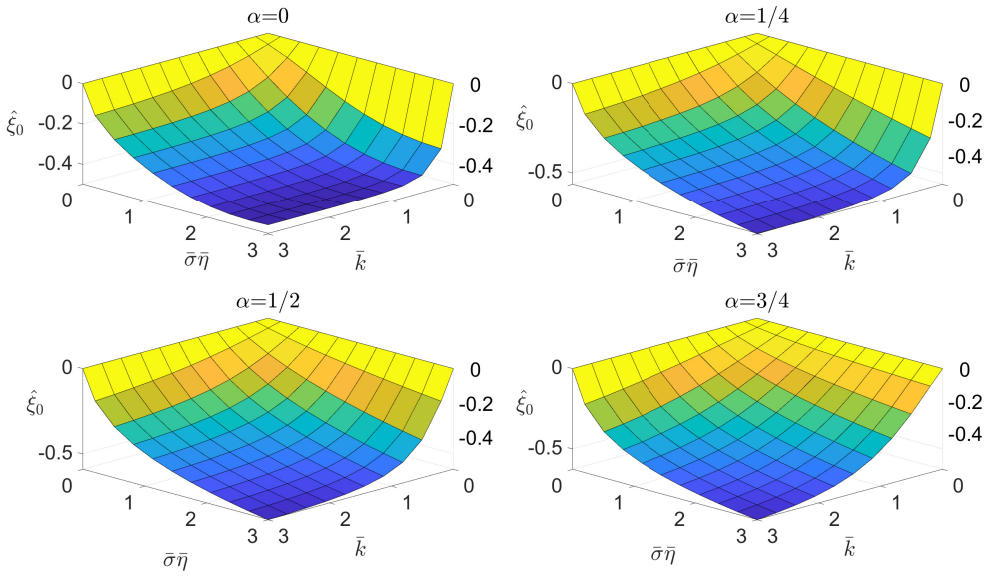

Figure 2.3: ATS skew term $\hat{\xi}_0$ for $\{\beta = 1, \delta = -1/2\}$. We report $\hat{\xi}_0$ for four values of α: $\alpha = 0$ in the upper left corner, $\alpha = 1/4$ in the upper right corner, $\alpha = 1/2$ in the lower left corner and $\alpha = 3/4$ in the lower right corner. We plot the skew for $\bar{k}, \bar{\sigma}\bar{\eta} \in [0, 3]$. In all cases the skew is negative and decreasing w.r.t. $\bar{k}$ and $\bar{\sigma}\bar{\eta}$.

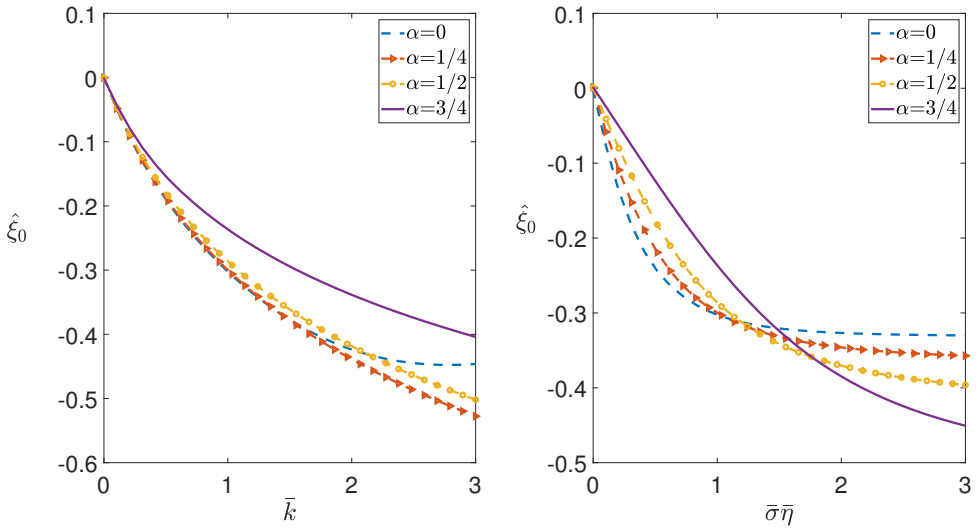

Figure 2.4: ATS skew term $\hat{\xi}_0$ for $\beta = 1$ and $\delta = -1/2$ for $\alpha = 0$ (dashed blue line), $\alpha = 1/4$ (red triangles), $\alpha = 1/2$ (orange circles) and, $\alpha = 3/4$ (continuous violet line). We plot the skew for $\bar{k} \in [0, 3]$ with $\bar{\sigma}\bar{\eta} = 1$ (on the left) and for $\bar{\sigma}\bar{\eta} \in [0, 3]$ for $\bar{k} = 1$ (on the right). In all cases the skew is decreasing w.r.t. $\bar{k}$ and $\bar{\sigma}\bar{\eta}$.

Finally, let us emphasize that the limits of $\hat{\xi}_0$ are zero for $\bar{\sigma}\bar{\eta}$ and $\bar{k}$ that go to zero.
On the one hand, recall that the law of S_t, $\mathcal{P}_{S_t}$, does not depend of $\bar{\sigma}\bar{\eta}$. By the dominated

convergence theorem with bound $\mathcal{P}_{S_t}$, we have that

$$\lim_{\bar{\sigma}\bar{\eta}\to 0} \mathbb{E}\left[erf\left(\bar{\sigma}\bar{\eta}\, r(S_t)\right)\right] = 0 \ .$$

On the other hand, by Kijima (1997, Th.B.9, p.308), we have that S_t converges in distribution to 1 as $\bar{k}$ goes to zero because

$$\lim_{\bar{k}\to 0} \mathcal{L}_t(u;\ k_t,\ \alpha) = \lim_{t\to 0} e^{\frac{1}{\bar{k}}\frac{1-\alpha}{\alpha}\left\{1-\left(1+\frac{u\,\bar{k}}{(1-\alpha)}\right)^{\alpha}\right\}} = e^{-u} \ .$$

We are computing the expected value of a bounded function of S_t that does not depend of $\bar{k}$. Thus, by definition of convergence in distribution,

$$\lim_{\bar{k}\to 0} \mathbb{E}\left[erf\left(\bar{\sigma}\bar{\eta}\sqrt{2}\,\left(1/\sqrt{S_t} - \sqrt{S_t}\right)\right)\right] = 0 \ .$$

2.5 Main Result

In the following theorem, we present the main results of this chapter and one of the main results of the thesis. We prove that if and only if $\beta = 1$ and $\delta = -\frac{1}{2}$ the ATS has a positive and constant short-time implied volatility $\hat{\sigma}_0$ and a negative and constant short-time skew $\hat{\xi}_0$. We point out that a finite skew w.r.t. y correspond to a skew that goes as $\frac{1}{\sqrt{t}}$ at short-time w.r.t. the moneyness x. The proof is based on the propositions of sections 2.3 and 2.4.

Theorem 2.5.1. *The ATS short-time implied volatility behaves as described in Table 2.1.*

Proof. We prove that, for Case 1, $\hat{\sigma}_0 = 0$ in **Proposition 2.3.1**. We prove that, for Case 2, $\hat{\sigma}_0 = \infty$ in **Proposition 2.3.2**. We prove that, for Cases 3, 4 and 5, $\hat{\sigma}_0$ is finite in **Proposition 2.3.3** and **Proposition 2.3.4**.
Moreover, in **Proposition 2.4.2** we show that, for Case 3, $\hat{\xi}_0 = 0$ and in **Proposition 2.4.3** we demonstrate that, for Case 4, $\hat{\xi}_0 = -\sqrt{\frac{\pi}{2}}$ and that, for Case 5, $\hat{\xi}_0$ is negative and finite. □

2.6 Conclusions

An excellent calibration of the equity implied volatility surface has been achieved by the ATS, a class of power-law scaling additive processes, cf. chapter 1 and 4. This class of processes builds upon the power-law scaling parameters β, related to the variance of jumps, and δ related to the smile asymmetry.

First, for this family of pure-jump additive processes we have obtained the behavior of the short-time ATM implied volatility $\hat{\sigma}_t$ and the skew term $\hat{\xi}_t$ over the region of admissible parameters (cf. **Theorem 1.2.3**). We get this result by constructing some relevant bounds for $\hat{\sigma}_t$ and obtaining the expression of $\hat{\xi}_t$, cf. equation (2.4.1), via the implicit function theorem.
Second, we have proven that only the scaling parameters observed in empirical analysis ($\beta = 1$ and $\delta = -1/2$) are compatible with the implied volatility observed in the equity market (cf. **Theorem 2.5.1**). Hence, we have demonstrated that it exists a pure-jump additive process (an exponential ATS) that, differently from the Lévy case, presents the two features observed in market data: not only a finite and positive short-time implied volatility but also a short-time skew proportionally inverse to the square root of the time-to-maturity.

Chapter 3

A fast Monte Carlo scheme for additive processes and option pricing

In this chapter, we present a fast Monte Carlo scheme for additive processes.

We analyze in detail error sources and propose a technique that reduces the two major sources of error. We compare our result with a benchmark method: the jump simulation with Gaussian approximation.
We show an application to additive normal tempered stable processes, to the class of additive processes introduced in chapters 1 and 2. Numerical results are relevant. The algorithm is an accurate tool for pricing path-dependent discretely-monitoring options with errors of one basis point or below. The scheme is also fast: the computational time is of the same order of magnitude of standard algorithms for Brownian motions.

These results have already been presented in Azzone and Baviera (2021b).

3.1 Introduction

In this chapter, we introduce a fast Monte Carlo simulation technique for additive processes. In option pricing, Monte Carlo methods are attractive because they do not require significant modifications when the payoff structure of the derivative changes. Additive processes are becoming the new frontier in equity derivatives for their ability, on the one hand, to reproduce accurately market data in model calibration, and on the other hand, to keep the process rather elementary (see e.g., chapter 1 and Li *et al.* 2016, Carr and Torricelli 2021). A process $X = \{X(t)\}_{t\geq 0}$ is said to be an additive process, if it presents independent (but not-stationary) increments and satisfies $X(0) = 0$ a.s.; stationarity is the main difference with Lévy processes.
For most additive processes, the law of increments is not known explicitly, but analytic expressions exist for the characteristic functions thanks to the celebrated Lévy–Khintchine formula (Sato 1999a). Given any such characteristic function for increments, this chapter aims to describe an efficient and accurate algorithm for Monte Carlo simulations of the increments and to compute the prices of a class of discretely monitoring path-dependent options.

Up to our knowledge, the unique Monte Carlo (MC) scheme developed for a specific class of additive processes, Sato processes, has been introduced by Eberlein and Madan (2009). They generalize to this class of additive processes, a well-known jump simulation technique developed for Lévy processes, that can be found in many excellent textbooks (see e.g., Cont and Tankov 2003, Asmussen and Glynn 2007). It entails truncating small jumps below a certain threshold and then simulating the finite number of independent jumps; finally, the Asmussen and Rosiński

(2001) Gaussian approximation (hereinafter GA) can be applied to substitute small jumps with a diffusive term: this has become a benchmark technique to compare numerical results.

In this chapter, we propose a MC technique for additive processes based on a numerical inversion of the cumulative distribution function (CDF). Monte Carlo simulation of additive processes is not straightforward because, in general, the CDF of process increments is not known explicitly. Since the seminal paper of Bohman (1970) general methods have been developed for sampling from Fourier transforms and even some specific methods for some distributions (e.g. stable distributions) that do not require numerical inversion (Samorodnitsky and Taqqu 1994, Sec.1.7, p.41).
In the financial literature, these techniques have been considered when the transition probability density of the underlying asset dynamics is not known explicitly; they have been developed specifically in the Lévy case, where it is possible to leverage on the stationary increments (see e.g., Glasserman and Liu 2010, Chen *et al.* 2012, Ballotta and Kyriakou 2014). These techniques are reliable and efficient: they build upon the characteristic function numerical inversion to obtain an estimation of the CDF. Specifically, we use the fast Fourier transform (FFT) method for the numerical inversion as proposed by Lee (2004) and then applied to MC option pricing in the studies of Chen *et al.* (2012) and Ballotta and Kyriakou (2014). Relative to this literature, our contribution lies in analyzing the three sources of error that arise in estimating derivative price expectations and showing how to improve the two largest ones.

The main contributions of this chapter are three. First, we propose a Monte Carlo simulation technique for additive processes based on FFT. Second, we improve the two main sources of numerical error in existing techniques to accelerate convergence, using both an analytic property of Fourier inversion in the complex plane and a spline method for CDF numerical inversion. Finally, we point out that the proposed technique is accurate and fast: *i*) we compare with traditional GA simulations showing that it is at least one and a half orders of magnitude faster whatever time horizon we consider and *ii*) we observe that, when pricing some discretely monitoring path-dependent options, the computing time has the same order of magnitude as standard algorithms for Brownian motions.

The rest of the chapter is organized as follows. In section 3.2, we overview the method and recall both Lewis (2001) formula for CDF and the error source in the numerical approximation: we discuss the optimal selection of the integration path. In section 3.3, we describe the proposed simulation method and present the other main error source in MC option pricing: the interpolation method in numerical inversion. We also discuss how to generalize the GA method for additives in an efficient way. Section 3.4 presents numerical results for a large class of pure-jump additive processes in the case of both European options (where analytic pricing methods are available), and some discretely monitoring path-dependent options. Section 3.5 concludes. We report the proofs in appendix .6.

3.2 Overview of the MC method for additive process

Pure jump asset pricing models based on additive processes have enjoyed remarkable popularity in recent years. At least for two main reasons. First, they allow a highly tractable closed-form approach with simple analytic expression for European options following Lewis (2001). This formula is computable as fast as the standard Black-Scholes one. Second, additive processes provide an adequate calibration to the implied volatility surface of equity derivatives, as well as they reproduce *stylized facts* as the time scaling of skew in volatility smile, see chapters 1 and 2. In this section, we describe a third reason in favor of these models: they allow a simple, accurate, and fast numerical scheme for path-dependent option valuation. We extend to additive processes

the preceding literature on Lévy processes' simulation techniques and we discover that, thanks to this Monte Carlo scheme, it is not a challenge to price exotic derivatives as Asian contracts or barrier options with discretely monitored barriers.

The simulation of a discrete sample path of an additive process reduces to simulating from the distribution of the process increment between time s and time $t > s$. Lévy process simulation is based on time-homogeneity of the jump process: the characteristic function of an increment is the same as the characteristic function of the process itself at time $t = 1$, re-scaled by the time interval $(t - s)$ of interest.

In this chapter, we extend the preceding analysis on Lévy processes by i) presenting an explicit method for additive processes from their characteristic function and ii) analyzing the explicit bound for the total estimation bias. In the Lévy case, thanks to process time homogeneity, the properties of the process characteristic function are immediately extended to its increments. For example, the characteristic function (also of increments) is analytic in a horizontal strip and the purely imaginary points on the boundary of the strip of regularity are singular points (cf. Lukacs 1972, Th.3.1, p.12). This identification of process characteristic function and increments' characteristic function is not anymore valid for additive processes. However, this chapter shows that the analyticity strip depends on time and that it is possible to build, in a simple way, a numerical scheme for additive processes requiring an additional condition.

Our method is based on three key observations. First, computing a CDF $P(x)$ corresponds to pricing a digital option: this can be done efficiently in the Fourier space. This step can be crucial, as already highlighted by Ballotta and Kyriakou (2014), the Fourier formula presents some numerical instabilities due to the presence of a pole in the origin. They propose a regularization that leads to an additional numerical error. We propose a different approach that is based on the Lee (2004) formula which presents two significant advantages. On the one hand, this technique is exact (thus, no numerical error is associated with it), and, on the other hand, it allows selecting the integration path that reduces the numerical error in the discretization of the CDF.
Second, the Lewis (2001) formula for the CDF can be viewed as an inverse Fourier transform method that can be approximated with a fast Fourier transform (FFT) technique: Lewis-FFT computes multiple values of the CDF simultaneously in a very efficient way.
Finally, knowing CDF approximation $\hat{P}$, we can sample from this distribution by inverting the CDF, i.e. by setting $X = \hat{P}^{-1}(U)$, with U an uniform r.v. in $[0, 1]$. Thus, simulating a r.v. via a numerical CDF (i.e. coupling the Fourier transform with a Monte Carlo simulation), requires a numerical inversion that is realized via an interpolation method. Following Glasserman and Liu (2010), due to its simplicity a linear interpolation of the CDF is chosen in the existing financial literature (see e.g., Chen *et al.* 2012, Feng and Lin 2013). We propose the spline as interpolation rule because the computational cost is very similar while the bias associated with the two interpolation rules is significantly different: the upper bound of the bias can be estimated for a given grid spacing γ, and, as we discuss in section 3.3, it should be at least γ^2 smaller for the spline interpolation. In extensive numerical experiments we observe that, on the one hand, the error decreases even faster as a power of γ than predicted by the upper bound, thanks to the additional properties of the interpolated functions, and on the other, it becomes negligible for the grids that are selected in practice.
Due to these three main ingredients (Lewis formula, FFT, and Spline interpolation) that play a crucial role in the proposed Monte Carlo simulation technique, we call the method Lewis-FFT-S.

The Lewis-FFT-S method extends the Eberlein and Madan (2009) technique to any additive process being significantly faster: we show that the proposed Monte Carlo is much faster than any jump-simulation method even considering the Asmussen and Rosiński (2001) Gaussian approximation. Analyzing in detail the numerical errors related to the methodology, we design an

algorithm that increases both accuracy and computational efficiency. To the best of our knowledge, the proposed scheme is the first application in financial engineering of the MC simulation based on Lewis formula and FFT, when the underlying is governed by an additive process.

In the next subsection, we also recall explicit and computable expressions for the error estimates.

3.2.1 Lewis CDF via FFT

The proposed MC method simulates from the characteristic function of the additive increments. Due to the Lévy-Khintchine formula, the characteristic function

$$\phi_t(u) := \mathbb{E}\, e^{i\, u\, f_t}$$

of an additive process f_t admits a closed-form expression. Furthermore, as already mentioned in the Introduction, according to Lukacs (1972, Th.3.1, p.12), process characteristic function is an analytic function in a horizontal strip delimited on the imaginary axis by two values. Similarly to Lee (2004), we define these values $p_t^- > 0$ and $-(p_t^+ + 1) < -1$, s.t. the characteristic function is analytic when $\Im(u) \in (-(p_t^+ + 1), p_t^-)$.
We observe that for Lévy processes, the increment $f_t - f_s$ has the same distribution as f_Δ, where $\Delta = t - s$: the same property does not hold for additive processes, due to the time inhomogeneity. For an additive process, the characteristic function of an increment $f_t - f_s$ between times s and $t > s$ is

$$\phi_{s,t}(u) = \mathbb{E}\, e^{i\, u\, (f_t - f_s)} = \frac{\mathbb{E}\, e^{i\, u\, f_t}}{\mathbb{E}\, e^{i\, u\, f_s}}\,,$$

due to the independent increment properties of additive processes, then

$$\ln \phi_{s,t} = \ln \phi_t - \ln \phi_s\,.$$

Assumption 1. p_t^+ and p_t^- are non increasing in t and $p_t^+ \geq p_t^- > 0$ ♣

Thanks to Lukacs theorem and under Assumption 1, we are able to easily identify the strip of regularity in the case of interest: the characteristic function of an increment $f_t - f_s$ is analytic when $\Im(u) \in (-(p_t^+ + 1), p_t^-)$. Lewis (2001) obtains the CDF, shifting the integration path within the characteristic function horizontal analyticity strip. The shift is $-i\, a$ with a a real constant s.t. $a \in (-p_t^-, 1 + p_t^+)$. Lewis deduces this formula using the properties of contour integrals in the complex plane.
If Assumption 1 holds, the CDF $P(x)$ of an additive process increment is (cf. Lee 2004, Th.5.1)

$$P(x) = R_a - \frac{e^{-ax}}{\pi} \int_0^\infty du\, Re\left[\frac{e^{-iux}\phi_{s,t}(u - ia)}{i\, u + a}\right] \tag{3.2.1}$$

where

$$R_a = \begin{cases} 1 & 0 < a < p_t^+ \\ \frac{1}{2} & a = 0 \\ 0 & p_t^- < a < 0 \end{cases}\,.$$

The case with no shift ($a = 0$) is the Hilbert transform: it has been considered in several studies in the financial literature on MC pricing (see, e.g., Ballotta and Kyriakou 2014, Chen *et al.* 2012). In the Hilbert transform case, the singularity in zero in the integration path should be taken into account as a Cauchy principal value; as already emphasized by Ballotta and Kyriakou

(2014), the method could be not robust enough for applications in the financial industry: they have suggested a regularization technique that introduces an additional error source, while the Lewis method we consider here is exact.

In the following, we focus on $a > 0$ and we approximate the Fourier transform with a discrete Fourier transform $\hat{P}(x)$

$$\hat{P}(x) := 1 - \frac{e^{-ax}}{\pi} \sum_{l=0}^{N-1} Re\left[\frac{e^{-i(l+1/2)hx}\phi_{s,t}((l+1/2)h - i\,a)}{i\,(l+1/2)h + a}\right] \quad ,$$

where h is the step size in the Fourier domain and N is the number of points in the grid. In subsection 3.2.2, we will discuss how the choice of $a > 0$ improves the CDF error.

We wish to obtain the CDF function for a large number of values in a regular grid with step size γ. A computationally efficient algorithm is the fast Fourier transform (see Lee 2004, for a detailed analysis of the method in derivative pricing): it involves Toeplitz matrix-vector multiplication (see e.g., Press *et al.* 1992, Ch.12) and relies on an additional requirement for N, whose simplest choice is $N = 2^M$ with $M \in \mathbb{N}$; hereinafter, we consider an N within this set.
The main advantage of the method is that the computational complexity of the FFT method is $O(N \log_2 N)$ when computing one time-increment. Moreover, with an FFT, it holds the relationship

$$\gamma\, h = \frac{2\pi}{N} \quad ;$$

i.e., for a given number N of grid points, the step size in the Fourier domain h fixes the step size γ.[1]

3.2.2 CDF error sources

The numerical Fourier inversion is subject *i*) to a discretization error, because the integrand is evaluated only at the grid points, and *ii*) to a range error, because we approximate with a finite sum.

Assumption 2. $\forall\, t > s \geq 0$ there exists $B > 0$, $b > 0$ and $\omega > 0$ such that, for sufficiently large u, the following bound for the absolute value of the characteristic function holds

$$|\phi_{s,t}(u - i\,a)| < Be^{-b\,u^{\omega}} \quad , \qquad \forall a \in (0, p_t^+) \qquad \clubsuit$$

Leveraging on Assumption 2, we can estimate the explicit bound for the bias in terms of the step size h and the number of grid points N, as shown in the next proposition. This result improves the known bounds for numerical errors when computing the CDF (3.2.1), via a discrete Fourier transform, and indicates an integration path that minimizes this error bound.

Proposition 3.2.1. *If Assumptions 1 and 2 hold, then*

1. *the numerical error $|P(x) - \hat{P}(x)|$ for the CDF is bounded by*

$$\mathcal{E}_{h,M}^{CDF}(x) = \frac{e^{-x\,p_t^+/2}}{\omega\; b^{1/\omega}} \frac{1}{Nh} \Gamma\left[\frac{1}{\omega}, b\,(N\,h)^{\omega}\right] + \frac{e^{-\pi p_t^+/h} + e^{-\pi p_t^+/h - p_t^+\,x}\phi_{s,t}(-i\,p_t^+)}{1 - e^{-2\pi p_t^+/h}} \quad , \quad (3.2.2)$$

[1] To avoid this constraint, one can consider the fractional fast Fourier transform (Chourdakis 2005) instead of the standard FFT. We have verified that the additional computational costs of the former method do not look justified in the CDF simulation described in this chapter.

where $\Gamma(z,u)$ *is the upper incomplete gamma function and*

$$\frac{1}{N\,h}\Gamma\left[1/\omega, b\,(N\,h)^{\omega}\right] = O\left((N\,h)^{-\omega} e^{-b\,(N\,h)^{\omega}}\right) \quad ;$$

2. the (optimal) bound holds selecting the shift a *in (3.2.1) equal to* $p_t^+/2$.

Proof. See Appendix .6 □

The first term of $\mathcal{E}^{CDF}_{h,M}(x)$ accounts for the range error in the numerical inversion, while the second one accounts for the discretization error.[2]
In the financial literature, error estimations have been proposed when approximating a CDF via a discrete Fourier Transform (see e.g., Lee 2004, Chen *et al.* 2012, Ballotta and Kyriakou 2014). The bound in **Proposition 3.2.1** extends these results to the Lewis-FFT case, showing how to select the optimal integration path in Lewis formula (3.2.1) to minimize the exponential decay of the error. Selecting the optimal path, CDF error is even better than the one proposed by Chen *et al.* (2012)[3] deduced via the sinc expansion technique.

We wish to get a small approximation error increasing N and decreasing h. However, let us observe that, if one takes the limit $h \to 0$ and $N \to \infty$ keeping Nh fixed, then the range error bound does not decrease. Thus, our interest is to select $h = h(N)$ so that the discretization and range errors are of the same order. Expression (3.2.2) allows us to determine the size h and the number N such that the two sources of CDF error are comparable: we can impose that $\exp(-\pi p_t^+/h) = \exp(-b\,(N\,h)^{\omega})$, i.e. we select

$$h(N) = \left(\frac{\pi\, p_t^+}{b}\frac{1}{N^{\omega}}\right)^{1/\omega + 1} \quad .$$

We define

$$\mathcal{E}^{CDF}_{M}(x) := \mathcal{E}^{CDF}_{h(2^M),M}(x) \tag{3.2.3}$$

the error in this case. $\mathcal{E}^{CDF}_{M}(x)$ in (3.2.3), is the relevant estimation of the CDF error that we use in practice: with this selection of h, the total CDF error is $O(N^{-\omega/(1+\omega)})\exp(-bN^{\omega/(1+\omega)})$ and decays almost exponentially as we increase N; moreover, the step size $\gamma = 2\pi/(hN) = O(N^{-1/(1+\omega)})$.

From this result, it is possible to understand the reason why we choose $a > 0$. It is possible to prove, following the same steps of **Proposition 3.2.1**, that in the $a < 0$ case the leading term in $\mathcal{E}^{CDF}_{M}(x)$ is $\exp(-\pi p_t^-/h)$. Thanks to Assumption 1, the discretization error for the same discretization step h, is better in the $a > 0$ case.

[2]It is possible also to obtain an error bound even when Assumption 2 does not hold. Equation (3.2.2) can be extended to the case where the characteristic function has an asymptotical polynomial decay $|\phi_{s,t}(u - i\,a)| \leq B\,|u|^{-b}$, with $b > 0$: in this case the range error decays only polynomialy due to the polynomial decay of the characteristic function (see e.g., Ballotta and Kyriakou 2014, eq.(14), p.1099). However, in practice, when pricing exotic derivatives, the exponential decay of the characteristic function is a good reason for model selection.

[3]In theorem 2.1. of Chen *et al.* (2012) the leading term in the discretization error goes as $\max(e^{-\pi\, p_t^-/h}, e^{-\pi\,(p_t^+ + 1)/h})$. Thanks to Assumption 1, the discretization error in (3.2.2), that goes as $e^{-\pi\, p_t^+/h}$ is better than the one proposed by Chen *et al.* (2012).

3.3 The simulation method

Knowing the CDF approximation $\hat{P}$, we can sample from this distribution by inverting $\hat{P}$, i.e. by setting $X = \hat{P}^{-1}(U)$, with U an uniform r.v. in $[0, 1]$.
From the Fourier inversion, we have an estimate of $\hat{P}$ on a grid of N points with step γ.
As pointed out by Glasserman and Liu (2010, sec.3, pp.1614-1615) an adequate inversion requires to impose that $\hat{P}$ is increasing and inside the interval [0,1].
It is convenient to work with a subset of this grid. We truncate the CDF between $x_0 < 0$ and $x_{\mathcal{K}} > 0$ and consider the equally spaced grid (with step γ) $x_0 < x_1 < ... < x_{\mathcal{K}}$ with $\mathcal{K} < N$.
Simulating a r.v. via a numerical CDF (i.e. coupling the Fourier transform with a MC simulation), requires a numerical inversion that is realized via an interpolation method.
As already discussed in section 3.2, differently from the existing financial literature (see e.g., Glasserman and Liu 2010, Chen *et al.* 2012, Feng and Lin 2013) the proposed method is based on spline interpolation. In the next subsection, we discuss the key idea behind this choice of the interpolation method.

3.3.1 Simulation error sources: truncation and interpolation

Besides the numerical inversion error, there are other two error sources: truncation and interpolation of the CDF.
Let us consider the expected value $\mathbb{E}V(f_t - f_s)$, with $V(x)$ a derivative contract with a pay-off differentiable everywhere except in n_V points. It can be proven, similarly to Chen *et al.* (2012, Th.4.3, p.14:11), that the pricing error[4] using the Lewis-FFT method with linear interpolation is

$$\mathcal{E} := \int_{-\infty}^{\infty} dx\, V(x)\, [p(x) - \hat{p}(x)] \tag{3.3.1}$$

$$< \left(|V(x_0)| + |V(x_{\mathcal{K}})| + (2\mathcal{K} + n_V) \sup_{x \in (x_0, x_{\mathcal{K}})} |V(x)| + 2 \sup_{x \in (x_0, x_{\mathcal{K}})} |V'(x)| \right) \mathcal{E}_M^{CDF}(x_0) \tag{3.3.2}$$

$$+ \frac{\phi_{s,t}^-}{2\pi} \left(\frac{|V(x_{\mathcal{K}})| e^{x_{\mathcal{K}} p_t^-}}{|p_t^-|} + \int_{x_{\mathcal{K}}}^{\infty} dx\, V(x) e^{x\, p_t^-} \right) + \frac{\phi_{s,t}^+}{2\pi} \left(\frac{V(x_0) e^{x_0 (p_t^+ + 1)}}{p_t^+ + 1} + \int_{-\infty}^{x_0} dx\, V(x) e^{x\, (p_t^+ + 1)} \right) \tag{3.3.3}$$

$$+ \frac{\gamma^2}{2\pi} (x_{\mathcal{K}} - x_0) \sup_{x \in (x_0, x_{\mathcal{K}})} |V'(x)| \int_{\mathbb{R}} |du\, u\, \phi_{s,t}(u)| \ , \tag{3.3.4}$$

where $p(x)$ is the probability density function of $f_t - f_s$, $\hat{p}$ its estimation and

$$\phi_{s,t}^- := \lim_{a \to p_-} \int_{\mathbb{R}} du\, |\phi_{s,t}(u - ia)| \quad \& \quad \phi_{s,t}^+ := \lim_{a \to p_+} \int_{\mathbb{R}} du\, |\phi_{s,t}(u - ia)| \ .$$

The components of the bias error (3.3.1) when pricing a derivative are three: an error related to the numerical approximation of the CDF (3.3.2), a truncation error (3.3.3) and an interpolation error (3.3.4). Let us consider each error source separately.

First, the error related to the numerical approximation of the CDF in (3.3.2) is proportional to $\mathcal{E}_M^{CDF}(x_0)$: we have discussed in the previous section how to select h in order to minimize it.

[4]The upper bound on the bias $\mathcal{E}$ can be trivially extended to a payoff with a finite number n of monitoring times. The most relevant case, for $n = 1$, will be discussed in detail in subsection 3.4.1.

Second, we can always choose x_0 and $x_{\mathcal{K}}$ s.t. the truncation error is negligible.
We select a symmetric interval ($x_{\mathcal{K}} = -x_0$), a standard choice in the literature, and x_0 the nearest point to $-D\sqrt{t-s}$ on the grid in which the CDF $\hat{P}$ is estimated, with $D > 0$.
Why do we select an x_0 that depends on the time-interval $t - s$? It can be easily explained with a graph. In Figure 3.1, as an example, we plot the one-day and one-year normalized probability density functions of the additive process used in the numerical experiments of section 3.4. As expected, the one-day density is significantly more concentrated around zero than the one-year density when considering a constant x (on the right). Conversely, the ranges of the two densities look similar when considering the rescaled moneyness $x/\sqrt{t-s}$. Thus, the range of probability densities scales approximately with $\sqrt{t-s}$.
Moreover, we choose $D = 5$. In extensive numerical experiments, we observe that for values below x_0 (or above $x_{\mathcal{K}}$) the probability density function is lower than 10^{-9} whatever time horizon is considered and its contribution to the price is negligible for all practical purposes.

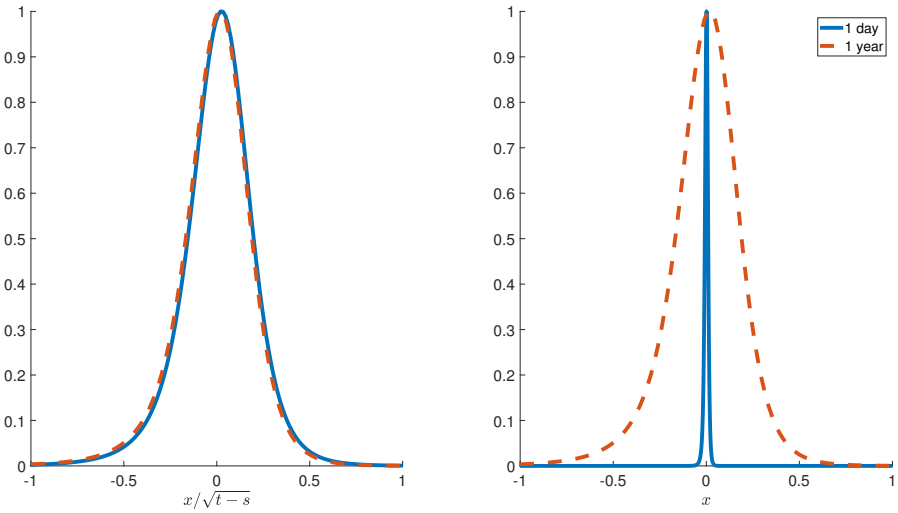

Figure 3.1: One-day and one-year normalized probability density functions of the additive process that we use in the numerical experiments of section 3.4 with $s = 0$. On the right, we see that, as expected, the one-day density is significantly more concentrated around zero than the one-year density. Conversely, on the left, we see that the ranges of the two densities are the same w.r.t. to the rescaled moneyness $x/\sqrt{t-s}$. Note that both probability density functions have been divided by their respective maximum for visualization purposes.

Finally, the bias associated with the linear interpolation, when computing the option value, is quadratic in the grid spacing γ; this turns out to be the most significant source of error, in most cases, as shown in the next section. It is well known that linear interpolation error goes as γ^2 (see e.g., Quarteroni *et al.* 2007, eq.(8.26), p.339). For this reason, in this chapter, we propose a spline interpolation method. In this latter case, it is known that the bias goes, at least, as γ^4 as shown in Hall and Meyer (1976).

As already emphasized by Glasserman and Liu (2010, sec.3, p.1615), to sample X from $\hat{P}(x)$ with a linear interpolation, after having generated U, a r.v. uniformly distributed in (0,1), one should

1. select the index j for which $\hat{P}(x_{j-1}) \leq U < \hat{P}(x_j)$;

2. for each j determine the linear interpolation coefficients $c^L_{0,j}$ and $c^L_{1,j}$

$$c^L_{0,j} := \frac{x_j\,\hat{P}(x_j) - x_{j-1}\,\hat{P}(x_{j-1})}{\hat{P}(x_j) - \hat{P}(x_{j-1})} \quad \text{and} \quad c^L_{1,j} := \frac{\gamma}{\hat{P}(x_j) - \hat{P}(x_{j-1})}$$

3. compute

$$X = c^L_{0,j} + c^L_{1,j}\,U \quad .$$

We discuss the computational cost of each step when sampling N_{sim} observations. The first step relies on a nearest neighborhood algorithm with an average computational cost proportional to $N_{sim} \times \log_2 N_{sim}$ (see, e.g., Cormen *et al.* 2001, p.11)[5]. The second step cost is proportional to $6 \times \mathcal{K}$. Finally the last step is proportional to N_{sim}.
Whereas step 1 is shared by both interpolation methods steps 2 and 3 differ between spline and linear interpolations. In step 2, the additional computational cost of considering spline interpolation instead boils down to the cost of solving a $K+1$-dimensional linear system with a tridiagonal matrix to determine the spline coefficients $\{c^S_{q,j}\}^3_{q=0}$, cf. Quarteroni *et al.* (2007, Ch.8), i.e. the cost is $8 \times K - 7$ (Quarteroni *et al.* 2007, Ch.7, p.391). As for step 3, the cost of computing the spline interpolation of U is still proportional to N_{sim}. It is clear that for a sufficiently large number of simulations N_{sim} and for $N_{sim} >> K$, for both methods, the most relevant contribution in the computational cost is the one due to step 1, the nearest neighborhood algorithm.

We perform numerical experiments to compare linear and spline interpolation. We observe that, if the number of simulations is significantly above the grid dimension K, the spline interpolation is as expensive as the linear interpolation. Moreover, in table 3.1, we compare the computational cost of linear interpolation and spline interpolation. We consider a grid of size $K = 10^4$ and $N_{sim} = 10^5$ simulations. In this case, spline simulation's cost is just 10% more than linear simulation's. The case considered in table 3.1 is a particularly unfavorable situation when comparing spline interpolation with linear interpolation: a large grid size $K = 10^4$ and a small number of simulations $N_{sim} = 10^5$. In this case steps 1, 2 and 3 computational times are comparable while, in practice, most of the computational costs are absorbed by the nearest neighborhood algorithm. For reasonable values of M (e.g. for $M \leq 15$), the dimension of the grid K is always well below 10^4. Thus, for all values of K and N_{sim} used in practice the incremental cost between Lewis-FFT (with linear interpolation) and Lewis-FFT-S (with spline interpolation) is negligible.

3.3.2 A simulation benchmark: the Gaussian approximation

In this subsection, we show how to generalize the GA method for additive processes in an efficient way when a monotonicity property holds for the Lévy measure and then the Ziggurat method (Marsaglia *et al.* 2000) can be applied.
A generic additive process may have an infinite number of jumps, most of them being small, over an arbitrary finite time horizon, making the simulation of such a process often nontrivial. Defining ν_t the additive process jump measure Sato (as in 1999a, Def.8.2, p.38), the jump

[5]The computational cost estimation is for the *merge sort* algorithm. Since *merge sort* is a recursive algorithm it could be necessary, for memory efficiency, to recur to *an insertion sort* algorithm which computational cost is roughly proportional to N^2_{sim} (see, e.g., Cormen *et al.* 2001, p.11).

Algorithm	Nearest neighborhood	Linear interpolation	Spline interpolation
time [ms]	1.08	1.13	1.27

Table 3.1: Computational cost in milliseconds [ms] for nearest neighborhood, linear interpolation, and spline interpolation. We consider a grid size $K = 10^4$ and $N_{sim} = 10^5$ simulations. Even considering a low number of simulations and a grid size K one order of magnitude above what is used in practice (in the Lewis-FFT-S case K is of order 10^3) spline simulation's cost is just 10% more than linear simulation's.

measure of the additive process increment $f_t - f_s$ is $\nu_t - \nu_s$.

Eberlein and Madan (2009), in their study on simulation of additive processes, consider only a class of additive processes (Sato processes): their approach consists in discarding the small jumps that in absolute value are below a given threshold ϵ. It is well known, in the Lévy case, that such an approach is accurate only if there are not too many small jumps (see e.g., Cont and Tankov 2003). Alternatively, the small jump component of an additive process may be approximated by a Brownian motion (Asmussen and Rosiński 2001).

Once the jump measure of the increment (between time s and time $t > s$) is truncated, we have i) to draw a Poisson number of positive and negative jumps and ii) to simulate separately positive jumps from the probability density $m^+_{s,t}$ and negative jumps from the probability density $m^-_{s,t}$, where

$$m^+_{s,t}(x) := \mathbb{I}_{x>\epsilon} \frac{\nu_t(x) - \nu_s(x)}{\int_\epsilon^\infty dz(\nu_t(z) - \nu_s(z))} \quad \& \quad m^-_{s,t}(x) := \mathbb{I}_{x<-\epsilon} \frac{\nu_t(x) - \nu_s(x)}{\int_\epsilon^\infty dz(\nu_t(z) - \nu_s(z))} \ .$$

To sample positive and negative jumps is extremely costly because often it is not possible to compute explicitly the integrals of $m^+_{s,t}$ and $m^-_{s,t}$.

Assumption 3. $m^+_{s,t}(x)$ is non increasing in x and $m^-_{s,t}(x)$ is non decreasing in x $\forall s, t$ s.t. $0 < s < t$ ♣

A faster methodology -for sampling from a known distribution without inverting numerically its integrals- is based on the Ziggurat method of Marsaglia *et al.* (2000). This method is applicable to probability density functions that are bounded and monotonic. We can apply the algorithm separately to negative and positive jumps. Having truncated small jumps the density functions are bounded, we need to ask the conditions of monotonicity listed in Assumption 3. The Ziggurat method covers a probability density with N_{ret} rectangles with equal area and a base strip. The base strip contains the tail of the probability density, it is built s.t. it has the same area of the rectangles. The method is composed of two building blocks: first, the rectangles with equal area are identified; second, the random variable is simulated either from a rectangle or from the base strip. Only in the latter case, an inversion of the integral is needed. N_{ret} is a key parameter because it controls the trade-off, in terms of computational time, between this inversion and building the rectangles.

With respect to Eberlein and Madan (2009), to reduce the bias of the method, we also consider the Gaussian approximation of Asmussen and Rosiński (2001).

β	δ	$\bar{k}$	$\bar{\eta}$	$\bar{\sigma}$
1	-1/2	1	1	0.2

Table 3.2: ATS parameters used in all numerical simulations. These selected parameters are consistent with the ones observed in market data.

3.4 Numerical results

Financial applications provide an important motivation for this study. We show that the proposed Monte Carlo technique for additive processes can price path-dependent options fast and accurately. The computational time is comparable to the case with simple Brownian motion dynamics.
We are interested in simulating a discrete sample path of the process over a finite time horizon: we are only concerned about the values of an additive process on such a discrete-time grid. This arises from situations where only discrete values of the process are concerned as in Chen *et al.* (2012), Ballotta and Kyriakou (2014) (e.g., they consider discrete barrier, lookback, and Asian options).
The case of an additive normal tempered stable (ATS) is discussed in detail. ATS processes present several advantages: they calibrate accurately equity implied volatility surfaces and, in particular, they capture volatility skews, cf. chapter 1.
The Lewis-FFT-S method and the GA benchmark can be used for the ATS because, in the next proposition, we prove that Assumptions 1, 2, and 3 hold for this additive process.

Proposition 3.4.1. *For an ATS process with $\alpha \in (0,1)$, Assumptions 1, 2 and 3 hold.*

Proof. See appendix .6 □

In particular, for the numerical example, we focus on the power-law scaling ATS, cf. **Theorem 1.2.3**, that is characterized by the parameters

$$k_t = \bar{k}\, t^{\beta}, \qquad \eta_t = \bar{\eta}\, t^{\delta}, \qquad \sigma_t = \bar{\sigma}\ ,$$

where $\bar{\sigma}, \bar{k}, \bar{\eta} \in \mathbb{R}^+$, and $\beta, \delta \in \mathbb{R}$. For all numerical experiments, we use the parameters reported in table 3.2: these parameters are consistent with the ones observed in market data.

To evaluate the Lewis-FFT-S performances we consider plain vanilla and exotic derivatives at different moneyness x and at different times-to-maturity. Deeply out-of-the-money and in-the-money options are less informative on the method performances, as the option value is close to the intrinsic value. In the rest of the section, to ensure that we verify the performance of the method on options in a relevant range of moneyness x, we consider x in the range $\sqrt{t}(-0.2, 0.2)$, where t is the option time-to-maturity.

In subsection 3.4.1, we show how the Lewis-FFT-S (with spline interpolation) method significantly outperforms the method with linear interpolation for European options.
In subsection 3.4.2, we provide evidence that Lewis-FFT-S is extremely fast and it is less computationally expensive, by at least 1.5 orders of magnitude than the GA method.
In subsection 3.4.3, we price discretely monitored Asian options, lookback options, and Down-and-In options with a time-to-maturity of five years. We also show that the Lewis-FFT-S is particularly efficient. The computational time needed to price path-dependent options with this method is just three times the computational time needed when using standard MC techniques for a geometric Brownian motion.

3.4.1 European options: accuracy

In the following, the Lewis-FFT-S performances are assessed for the ATS process. First, we show that, when using linear interpolation the leading term in (3.3.1) goes as γ^2. Then, we improve the bound by considering spline interpolation (Lewis-FFT-S) and we discuss the excellent performances of the method for the ATS case. Thanks to the FFT approach, Lewis-FFT-S is particularly fast: computational time has the same orders of magnitude of standard algorithm that simulate Brownian motions). Thanks to the spline interpolation, Lewis-FFT-S is also particularly accurate, for 10^7 simulations, and for any $M > 9$ the maximum error is 0.03 bp or below.

We do not desire a method that performs well either only OTM or only ITM. We want a MC that prices accurately options with generic moneyness: for this reason, we consider the 30 European call options with moneyness in a regular grid with range $\sqrt{t}(-0.2, 0.2)$. The numerical error arising from the MC can be estimated easily in the European call case having a closed formula for option's prices in equation1.3.1. The method's error is assessed in terms of the maximum error in absolute value (MAX) and the average over the 30 MC standard deviation (SD).
Monte Carlo error is often decomposed into bias and variance. In this chapter, we aim to reduce the bias error, but it is relevant to take into account also the variance. For a large number of simulations, confidence intervals estimated via a MC are directly linked to this quantity (see, e.g., Glasserman 2004, Ch.1, eq.(1.10), p.10). In our case, since we are considering the average error over 30 call options, we consider the average standard deviation SD as a rough estimate of the variance error in the estimated prices. When the maximum error is below SD we can infer that the error on bias has been dealt with correctly. In all considered cases, SD is of the order of 0.1 basis points.

In Figure 3.2, we plot the three terms that appear in the bias bound of equation (3.3.1) for an ATS with $\alpha = 2/3$ over a one-month time interval. The bound is for Lewis-FFT simulation with linear interpolation varying the number of grid points in the FFT via M s.t. $N = 2^M$. We plot the bounds on the error i) due to the truncation of the CDF (blue circles) in expression (3.3.3), ii) due to the linear interpolation of the CDF (red squares) in expression (3.3.4), and iii) due to the range and discretization error of the FFT inversion (green triangles) in expression (3.3.2). As we have already anticipated in subsection 3.3.1, two are the most relevant error sources: the error originating from the CDF approximation and the one due to the interpolation. The error originating from the truncation is always negligible: at least ten orders of magnitude lower than interpolation error for every M. For the CDF approximation error, as explained in section 3.2, we have suggested an optimal selection of the shift a in the Lewis-FFT approach. The term that we need to tackle is the interpolation one: for $M > 8$ the unique relevant bound is the one on the interpolation error (e.g. for M=10 it is 10 orders of magnitude above all other errors). Similar results holds $\forall \alpha \in (0, 1)$.

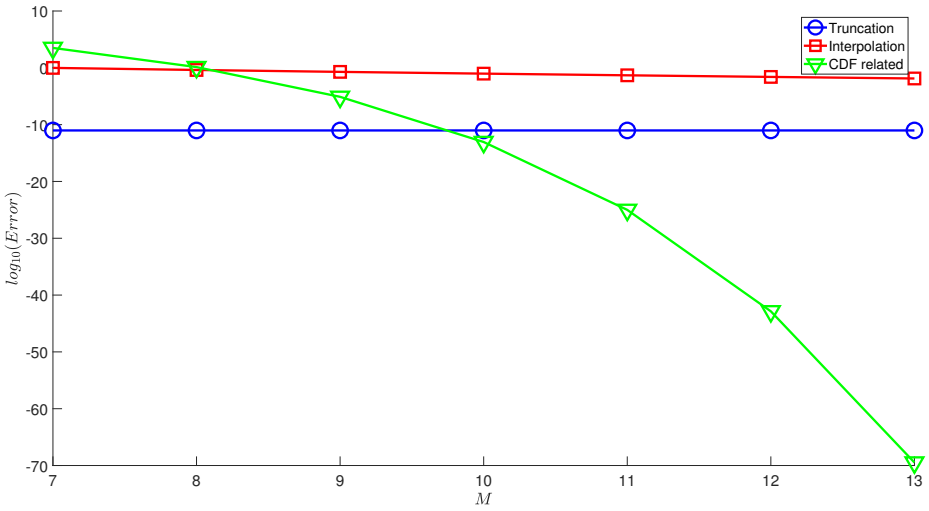

Figure 3.2: One-month European call options error bounds for an ATS ($\alpha = 2/3$) simulated with Lewis-FFT and linear interpolation. We plot the bounds on the three error sources: *i*) the truncation error (3.3.3) (blue circles), *ii*) the error (3.3.4) due to the linear interpolation of the CDF (red squares) and *iii*) the error (3.3.2) related to numerical CDF (green triangles). Let us emphasize that the truncation error is always negligible w.r.t. the linear interpolation error (at least 10 orders of magnitude smaller for every M). Notice that, for $M > 8$, the unique significant term is the bound on the linear interpolation error (e.g. for $M = 10$ it is at least 10 orders of magnitude above all other errors).

As discussed in subsection 3.3.1, to reduce the CDF interpolation error, we consider the spline interpolation for the numerical inversion instead of the linear interpolation. With spline interpolation $\mathcal{E}$ should scale as γ^4 instead of γ^2. In Figure 3.3 and 3.4, we plot the Lewis-FFT maximum error for two different times-to-maturity: the error is for 30 European call options for different values of M using spline (blue circles) and linear (red squares) interpolation. We also plot SD, the average MC standard deviation with a dashed green line. Notice that, for $M > 7$ the spline interpolation error is significantly below the linear interpolation error. Spline interpolation's error improves significantly faster than the linear interpolation's error: for M in the interval (7,10) the maximum error scales as γ^6 for the spline interpolation, this is probably due to the monotonicity and boundness of the CDF, and as γ^2 for the linear interpolation.

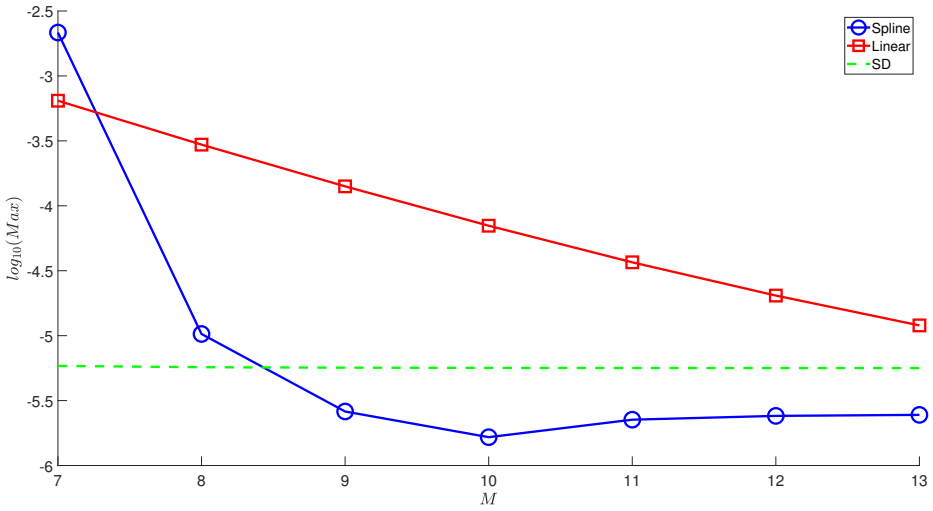

Figure 3.3: Maximum error for different values of M using Lewis-FFT-S (blue circles) and Lewis-FFT with linear interpolation (red squares). The maximum is computed over 30 call options (one-week maturity), with moneyness in the range $\sqrt{t}$(-0.2,0.2). We consider 10^7 simulations and $\alpha = 2/3$. Notice that, for $M > 9$ the spline interpolation error is significantly below the linear interpolation error. Spline interpolation's error improves significantly faster than the linear interpolation's error: for M in the interval (7,10) the maximum error scales, on average, as γ^6 for the spline interpolation and as γ^2 for the linear interpolation. Moreover, the maximum error becomes significantly lower than the average MC standard deviation in a dashed green line.

	M	7	8	9	10	11	12	13
$\alpha = 1/3$	MAX [bp]	5.89	0.27	0.02	0.03	0.03	0.03	0.03
	RMSE [bp]	2.92	0.24	0.01	0.02	0.02	0.02	0.02
	MAPE [%]	1.88	0.18	0.01	0.01	0.01	0.01	0.01
	SD [bp]	0.12	0.12	0.12	0.12	0.12	0.12	0.12
$\alpha = 2/3$	MAX [bp]	42.91	0.20	0.05	0.01	0.02	0.03	0.03
	RMSE [bp]	27.38	0.17	0.03	0.01	0.01	0.02	0.02
	MAPE [%]	11.45	0.12	0.02	0.01	0.01	0.01	0.01
	SD [bp]	0.13	0.11	0.11	0.11	0.11	0.11	0.11

Table 3.3: Lewis-FFT-S algorithm (with spline) performances w.r.t. different metrics using 10^7 trials for $\alpha = 1/3$ and $\alpha = 2/3$: maximum error [bp], RMSE [bp], MAPE [%], SD [bp]. The process is simulated for M that goes from 7 to 13. The metrics are computed for 30 call options (one-month maturity), with moneyness in the range $\sqrt{t}$(-0.2,0.2). We observe that for all $M \geq 9$ the maximum error is 0.03 basis points or below.

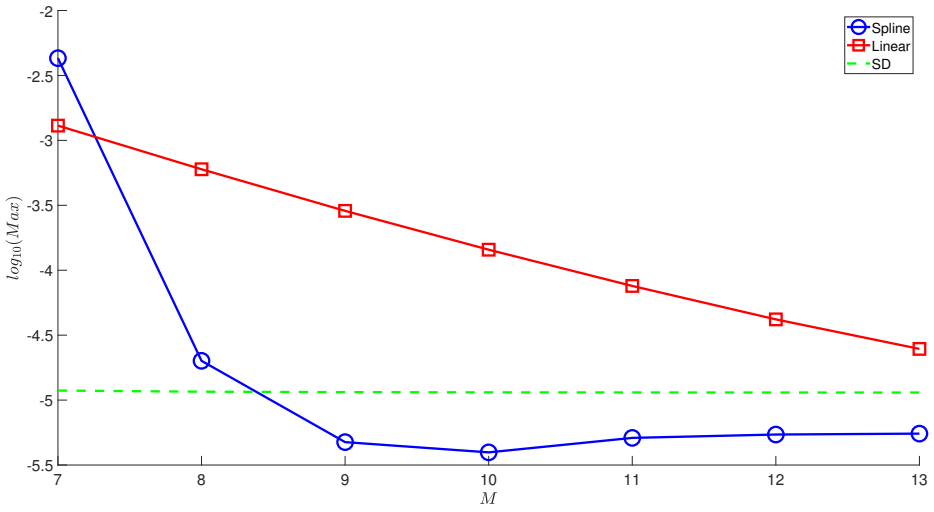

Figure 3.4: As Figure 3.3 but for one-month maturity. Notice that, for $M > 7$ the spline interpolation error is significantly below the linear interpolation error. Spline interpolation's error improves significantly faster than the linear interpolation's error: for M in the interval (7,10) the maximum error scales, on average, as γ^6 for the spline interpolation and as γ^2 for the linear interpolation.

We also desire to estimate the method's error with different metrics: besides MAX we consider the root mean squared error (RMSE) and the mean absolute percentage error (MAPE). In table 3.3, we report the performances of the Lewis-FFT-S algorithm for 10^7 simulations. We consider two values of α for the ATS: $\alpha = 1/3$ and $\alpha = 2/3$. The metrics are computed for 30 call options (one-month maturity) and moneyness in the range $\sqrt{t}$(-0.2,0.2). We observe that for $M \geq 9$ the error is 0.03 basis points or below whatever metric we consider.

The main result of this subsection is that, in the Lewis-FFT-S framework, a Monte Carlo with 10^7 simulations and $M = 13$ provides an accurate pricing tool whatever time-horizon and $\alpha \in (0, 1)$

	M	7	8	9	10	11	12	13
$\alpha = 1/3$	Time [s]	0.23	0.27	0.28	0.28	0.28	0.28	0.29
$\alpha = 2/3$	Time [s]	0.25	0.27	0.28	0.28	0.28	0.28	0.28

Table 3.4: Lewis-FFT-S computational time in MATLAB for simulating the ATS (with $\alpha = 1/3$ and $\alpha = 2/3$) over a one-month time-interval.

we consider.

3.4.2 European options: computational time

In this subsection, we emphasize that the proposed MC method is fast. We compare the Lewis-FFT-S computational cost both with the simplest possible dynamic for the underlying (geometric Brownian motion) and with the methodology that is often considered a benchmark for simulating jump processes (the simulation of jumps via Gaussian approximation).

In table 3.4, we report the performances of the Lewis-FFT-S algorithm for 10^7 simulations. We consider the ATS with $\alpha = 1/3$ and $\alpha = 2/3$. For every choice of M, we have the maximum error [bp], the RMSE [bp],the MAPE [%], the SD [bp], and the computational time [s]. The metrics are computed for 30 call options (one-month maturity), with moneyness in the range $\sqrt{t}$(-0.2,0.2). We observe that for $M \geq 10$ the maximum error is 0.03 basis points or below.

We point out, that Lewis-FFT-S is considerably efficient. In our machine[6], sampling 10^7 observation of a geometric Brownian motion takes approximately 0.08 seconds which is just one-third of the Lewis-FFT-S's computational cost (reported in table 3.4).

In Figure 3.5, we plot the computational time w.r.t. the time-to-maturity in log-log scale for 10^7 simulations with Gaussian Approximation (blue squares) and Lewis-FFT-S (red circles). Time-to-maturity goes from one day to two years. To compare the two methods fairly we need to select M for the Lewis-FFT-S and ϵ for the Gaussian approximation s.t. the two methods provide similar errors. As above, for both methods, we price the 30 call options, with moneyness in the range $\sqrt{t}$(-0.2,0.2). For each time-to-maturity, we select M and ϵ s.t. the maximum error (MAX) is between 1 basis point and 0.1 basis points, and s.t. the Lewis-FFT-S error is always below the Gaussian approximation error. Lewis-FFT-S computational time appears constant as the time-to-maturity increases. While GA computational time improves as the time-to-maturity reduces. However, Gaussian approximation is always more computationally expensive than Lewis-FFT-S by at least 1.75 orders of magnitude. This difference appears remarkable considering that we have verified that Lewis-FFT-S error is always below Gaussian approximation error.

[6]We use MATLAB 2021a on an AMD Ryzen 7 5800H, with 3.2 GHz.

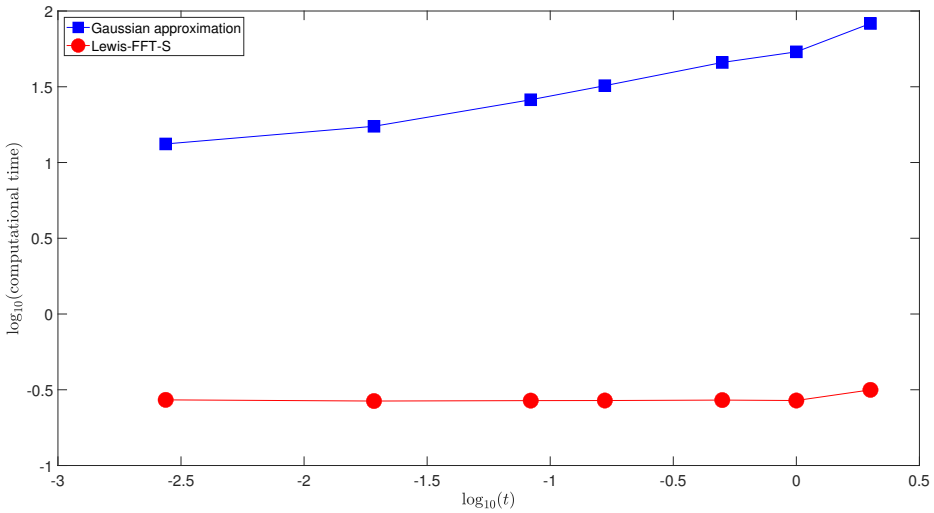

Figure 3.5: Computational time w.r.t. the time-to-maturity in log-log scale for 10^7 simulations with GA (blue squares) and Lewis-FFT-S (red circles) techniques. We price 30 European call options, with moneyness in the range $\sqrt{t}(-0.2, 0.2)$ with GA and Lewis-FFT-S. For each time-to-maturity, between one day to two years, we select M and the threshold ϵ s.t. the maximum error is between 1 basis point and 0.1 basis points and we require that the Lewis-FFT-S error is always below the GA error. The GA computational time improves as the time-to-maturity reduces because a lower number of jumps is involved, while the Lewis-FFT-S simulation depends weakly on the time horizon. We observe that GA is always more computationally expensive than Lewis-FFT-S by at least 1.75 orders of magnitude.

3.4.3 Discretely monitoring options

In this subsection, we price discretely monitored (quarterly) Asian options, lookback options, and Down-and-In options with a time-to-maturity of five years.

For simplicity, we consider the case with no interest rates and dividends: these two deterministic quantities can be easily added to simulated prices without any computational effort.
The Asian call options, the lookback put options and the Down-and-In put options we consider are respectively

$$\left(\sum_{i=0}^{n} e^{f_{t_i}} - e^{-x}\right)^+ ,$$
$$\left(e^{-x} - \min_i e^{f_{t_i}}\right)^+ \text{ and}$$
$$\left(e^{-x} - e^{f_{t_i}}\right)^+ \mathbb{I}_{\min_i e^{f_{t_i}} \geq L} ,$$

where $0 = t_0 < t_1 < ... < t_i < ... < t_n$ with $n = 20$ the monitoring times, f_{t_i} is the ATS price at time t_i, and x is the option moneyness and L is the Down-and-In barrier. We simulate the paths of f_t by simulating the increments $f_{t_i} - f_{t_{i-1}}$.

In table 3.5, we report prices and MC standard deviation of Asian call options and lookback put options. We simulate 10^7 paths of the ATS with $\alpha = 2/3$ and price the discretely monitored

Moneyness	Asian [%]	SD [%]	Lookback [%]	SD [%]	Down-and-In [%]	SD [%]
-0.5	39.79	0.01	3.31	0.00	2.31	0.00
-0.25	24.36	0.01	8.72	0.00	3.98	0.00
0	10.04	0.01	23.07	0.01	6.15	0.01
0.25	2.57	0.00	50.53	0.01	8.95	0.01
0.5	0.55	0.00	86.98	0.01	12.55	0.01

Table 3.5: Prices and MC standard deviation of Asian calls, lookback puts, and Down-and-In put options for different moneyness. We simulate 10^7 paths of the ATS with $\alpha = 2/3$ and price the discretely monitored (quarterly) path-dependent options with time-to-maturity of five years. SD errors are always lower than 1 basis point.

(quarterly) Asian options and lookback options with time-to-maturity five years. We consider options with different moneyness in the range (-0.5,0.5), where $0.5 \approx 0.2\sqrt{t}$. The Down-and-In barrier strike is $L = 0.6$. We use $M = 13$ for the numerical CDF inversion. As pointed out in the previous subsection, the Lewis-FFT-S is extremely efficient: the computational cost of simulating the ATS path is just three times the computational cost of simulating a standard geometric Brownian motion.

3.5 Conclusion

In this chapter, we propose a new Monte Carlo simulation technique for additive processes that leverages on the numerical efficiency of the FFT applied to the Lewis formula for a CDF and on the spline interpolation properties when inverting the CDF. We call Lewis-FFT-S this Monte Carlo simulation. We propose an application to the Additive normal tempered stable (ATS) process introduced in chapter 1 and we show that the Lewis-FFT-S significantly outperforms the GA method. This simulation technique is accurate and fast.

We discuss in detail the accuracy of the method. In Figure 3.2 we discuss the three-components of the bias error in (3.3.1). To accelerate convergence we improve the two main sources of numerical error in (3.3.1): the CDF error in equation (3.3.2), and the interpolation error in equation (3.3.4). First, we consider the Lewis formula, see equation (3.2.1), for the Fourier transform inversion. This eliminates the source of error originating from the pole in the origin (see, e.g., Ballotta and Kyriakou 2014, eq.(4)). Second, we substitute linear interpolation with spline interpolation. In this way, the leading term in the interpolation error improves from γ^2 to at least γ^4. This improvement is particularly evident in Figures 3.3-3.4, where, for $M > 7$, the Lewis-FFT-S (with spline interpolation) maximum error is significantly below the Lewis-FFT version of the method with linear interpolation and it appears to decrease as γ^6 in numerical experiments.

The Lewis-FFT-S is accurate but also fast. As discussed in subsection 3.3.1, for a sufficiently large number of simulations, the increment in computational time due to spline interpolation is negligible. Moreover, as shown in Figure 3.5, the proposed method is at least one and a half orders of magnitude faster than the traditional GA simulations whatever time horizon we consider. We also observe that, when pricing some discretely monitoring path-dependent options, the computational time is of the same order of magnitude as standard algorithms for Brownian motions.

A brief description of the Lewis-FFT algorithm follows.

---

procedure LEWIS-FFT($M, \mathcal{N}_{sim}, FlagSpline$)

COMPUTE $h(M)$, N, γ
COMPUTE $\hat{P}$ with FFT ▷ z_0, z_{N-1} fixed by FFT

FIX x_K nearest point to $5\sqrt{t-s}$ and $x_0 = -x_K$
$\vec{x} = x_0 : \gamma : x_K$ ▷ Grid dimension: $K+1$

SAMPLE a vector U of $\mathcal{N}_{sim}$ uniform r.v. in [0,1]
J = NearestNeighborhood(U, $\hat{P}(\vec{x})$) ▷ Find next element in the grid

if $FlagSpline = True$ **then**
COMPUTE spline interpolation coefficients $\{c^S_{q,J}\}_{q=0}^3$ ▷ Solve tridiagonal linear sistem
X = spline($\hat{P}(\vec{x})$, $\vec{x}$, U,J) ▷ Interpolate on U
else
COMPUTE linear interpolation coefficients $\{c^L_{q,J}\}_{q=0}^1$
X = $c^L_{0,J} + U\, c^L_{1,J}$ ▷ Interpolate on U

Chapter 4

The equity derivatives market: Synthetic forwards and cost-of-funding

This chapter presents the equity market dataset and introduces a new technique to recover the implicit discount factor in the derivative market using only European put and call prices: this discount is grounded in actual transactions in active markets. Moreover, this chapter identifies the implied cost-of-funding, over OIS, of major market players.

Does a liquid equity market allow arbitrage? The key idea is that the (unique) forward contract -built using the put-call parity relation- contains information about the market discount factor: by no-arbitrage conditions we identify the implicit interest rate such that the forward contract value does not depend on the strike.

The procedure is applied to options on S&P 500 and EURO STOXX 50 indices. There is statistical evidence that, in the EURO STOXX 50 market, the implicit interest rate curve coincides with the EUR OIS one, while, in the S&P 500 market, a cost-of-funding of, on average, 34 basis points is added on top of the USD OIS curve.

We use the market implicit interest rates and forward prices to calibrate the ATS on a large dataset. We discuss the excellent calibration performances and the power scaling behavior of the ATS, over a dataset of nine-year length. We show that ATS calibrates the EURO STOXX 50 and S&P 500 implied volatility surfaces significantly better than the Lévy case (more than two order of magnitude).

Part of this chapter has already been published in Azzone and Baviera (2021d).

4.1 Introduction

The term structure of interest rates is a crucial input in the derivative market. It is used for determining the discount rate for expected payoffs in a given currency.
The main research question we consider in this is: when dealing with liquid exchange-traded derivatives, which is the interest rate term structure used by market makers?
In general, interest rates used in derivative pricing are not "risk-free" because contingent claim evaluation should depend on the risks of the investment and in particular, on the funding risk and on the risk of default of one of the two counterparties in the derivative contract.[1]

[1]For a dealer, the expected loss due to a possible default by the counterparty is related to the credit value

When dealing with exchange-traded derivatives the situation should be simpler: the presence of a clearinghouse with margin calls allows neglecting the market participants' default risk. Before the Great Financial Crisis of 2007, the answer was to consider the Libor curve as the discounting term structure (see e.g., Hull and White 2013, p.14). After the crisis, the difference between Libor rates with different tenors enlarged to several tenths of basis points (see, e.g., Henrard 2014, and references therein) making this answer less obvious; more recently the situation has become even more complicated, in particular after July 2017, when the Chief Executive of U.K. Financial Conduct Authority (FCA) in a famous speech (Bailey 2017) increased market expectation that Libor benchmarks will be discontinued within a few years (see, e.g., Henrard 2019, for a clear and exciting discussion on the Libor fallout from a quantitative perspective).

The Overnight Index Swap (OIS) curve has emerged as a possible candidate for the risk-free curve for derivative discounting in the aftermath of the crisis. The OIS is a swap derived from the unsecured interbank overnight rate (OR), which is, for example, the EONIA rate for Euros[2] and the Effective Federal Fund Rate (EFFR) for US dollars[3]. This OR can be considered a good proxy of a risk-free rate and it is the interest rate most commonly paid on margins. Moreover, the OIS curve presents several advantages: it is a curve based on liquid swaps. The bootstrap of the discounting curve is as simple as the well-established pre-crisis methodology (see, e.g., Ron 2000).

The approach of selecting the interest rate term structure from a practitioner perspective appears relatively clear. We are particularly interested in market makers that operate in a given exchange-traded derivative market. Often they consider the OIS curve for discounting, allowing for a spread that accounts for other risks or costs not included in the "risk-free" rate.
We call this spread "cost-of-funding" because it can be seen as the implicit additional cost in operating in this derivative market. We reformulate our research question for these market makers: which is the cost-of-funding (if any) of operating in a liquid exchange-traded derivative market?
The answer to this question has both operational and management implications. On the one side, for their daily activity, the market makers should build and monitor an indicator on this spread, to use a discounting curve in line with other market participants; on the other side, this spread has relevant consequences on the management of a financial firm. If, within a financial institution, the cost-of-funding of a given business unit of market-making is larger than the market, it is rather difficult that this unit can be competitive in the derivative market where it operates. Determine at which cost-of-funding each business unit should operate is a relevant management decision within a financial firm. This chapter introduces an elementary indicator that can monitor in real-time the funding cost and point out a possible stress in funding liquidity.

We consider all options on the EURO STOXX 50 and the S&P 500, respectively the most liquid equity index in the Euro area and in the U.S.A. (see e.g., Dash and Liu 2009, Bai *et al.* 2019, Vo and Daly 2008, Ñíguez 2016).
This study builds over the put-call parity of European options. The idea of using put-call parity to obtain the implied interest rates dates back to Brenner and Galai (1986), who consider

adjustment and the expected gain due to a possible default by the dealer itself is referred to as the debt value adjustment; the funding risk is associated with the funding valuation adjustment. This pricing approach can be found in excellent textbooks (see, e.g., Gregory 2012, Brigo *et al.* 2013).

[2]Substituted by the Euro short-term rate (€STR) starting from the $2^n d$ of October 2019, i.e. after the period of analysis considered in this chapter.

[3]USD OIS market is mainly based on this rate. The OIS swap trading volumes based on the Secured Overnight Financing Rate (SOFR) is negligible w.r.t. the total OIS volume at the time of writing.

underlyings that do not pay dividends. Frankfurter and Leung (1991) and Naranjo (2009) extended this methodology to options on an underlying that pays dividends. To apply their techniques it is necessary to know both the forward prices and the forward dividends. They infer future dividends from realized ones and discuss the differences between the discount factor observed in the market and the discount factor obtained from the LIBOR and Treasury curves.

This study presents an alternative approach that allows us to obtain the implicit interest rates using only option prices and a no-arbitrage condition on an option portfolio known as *synthetic* forward.[4] The implicit interest rates of the S&P 500 and EURO STOXX 50 option markets are computed with a simple technique. Together with OIS discounting term structure, this technique allows a market maker to build an elementary measure of the cost-of-funding that can be obtained instantly from option prices. Moreover, from the market implicit interest rate it is possible to build the underlying forward prices. We will use this technique to calibrate the ATS to a big dataset of option prices (nine-year length) in section 4.4.
In chapter 1, we introduced a new broad family of stochastic processes that we call additive normal tempered stable processes (ATS). An interesting subcase of ATS presents a power-law scaling of the time-dependent parameters.
The model is calibrated on market data on a choosen maturity. In chapter 1, we considered all quoted options on S&P 500 and EURO STOXX 50 at 11:00 am NT on the 30^{th} of May 2013. The dataset contains options with a time-to-maturity starting from three weeks and up to several years. We calibrate the ATS processes on the options of both indices, showing that ATS processes present better calibration features than LTS and Sato processes (Sato 1991). The observed improvement of ATS is even of two orders of magnitude in terms of MSE, as reported in Table 1.1. ATS replicates accurately market implied volatility term structure and *skew* as observed in Figures 1.5 and 1.6.
The quality of ATS calibration results are stunning. In subsection 1.3.3, we have shown that once the volatility term structure has been taken into account, the whole implied volatility surface is calibrated accurately with only two free parameters.
We employ the market implied interest rates and forward prices to calibrate the ATS over a nine-year time interval.
We consider a dataset on options on the S&P 500 and EURO STOXX 50 indexes of 9 years length, of closing market prices. The calibration results are stunning. First, for both indexes the ATS outperforms the Lévy process by at least two orders of magnitude in terms of mean squared error (MSE) and one order of magnitude mean absolute percentage error (MAPE). ATS also outperforms significantly Sato processes in all cases.

The rest of the chapter is organized as follows. Section 4.2 shows the methodology to find the implicit interest rates using only option prices and describes the dataset. Section 4.3 infers the S&P 500 and the EURO STOXX 50 implicit discount factor and the corresponding cost-of-funding. Section 4.4 presents the results of the ATS calibrations on a nine-year dataset. Section 4.5 concludes.

4.2 The methodology and the dataset

This section shows how to obtain the discount factor from market data using only call and put prices. We present the dataset and discuss the data preprocessing techniques.

[4]Synthetic forwards are perfectly synchronized with option prices. There is empirical evidence that, in some markets, they are more reliable than quoted futures (see e.g., Muravyev *et al.* 2013, Hao *et al.* 2020).

The absence of arbitrage condition allows us to write, at value date t_0 and at a fixed maturity T, the put-call parity for European options (see, e.g., Hull 2003, Ch.8 p.174) w.r.t. the forward price $F_{t_0}(T)$ and the strike price K

$$C_T(K) - P_T(K) = \overline{B}_T(F_{t_0}(T) - K) \ , \tag{4.2.1}$$

where $C_T(K)$ and $P_T(K)$ are respectively the European call and put option prices and $\overline{B}_T$ is the market discount factor between value date and T.
Instead of considering a standard forward contract, a trader in this market can mimic this position using call and put options with the same strike price and the same maturity to create a forward position: this position is called *synthetic* forward. The synthetic forwards are frequently traded in the equity derivative markets: they identify –for several maturities– the most liquid forwards in the market.
A synthetic forward $\mathcal{G}_T(K)$ with maturity T is a portfolio that comprises of a long call and a short put at a given strike price K. Forward prices in t_0 with the same maturity T are all equivalent whatever strike K is considered and, due to the no-arbitrage condition, they should have the same price.[5] The market implied discount factor $\overline{B}_T$ is the (unique) factor such that the forward price

$$F_{t_0}(T) = \frac{\mathcal{G}_T(K)}{\overline{B}_T} + K \tag{4.2.2}$$

does not depend on the strike K: this is the main idea of the chapter. This is a linear problem in $\overline{B}$ and F. We discuss its solution in section 4.2.

We consider all quoted S&P 500 and EURO STOXX 50 option prices[6] observed at 11:00 am NT each business day from the 1^{st} of November 2018 to the 19^{th} of July 2019 excluding days from the 20^{th} of December 2018 to the 6^{th} of January 2019 and from the 13^{th} of April 2019 to the 2^{nd} of May 2019. For both indices, the most liquid options expire on the third Friday of the first six months after the value date and then on March, June, September, and December in the front year and June and December in the next year. In the EURO STOXX 50 case also June and December contracts for the following years are available.[7] In Table 4.1 we provide the descriptive statistics of some relevant quantities in the options' dataset. We report the number of strike for each maturity, the straddle position, $C_T(K) + P_T(K)$, and the synthetic forward plus the strike, $\mathcal{G}_T(K) + K$.

[5] We could build an arbitrage position on synthetic forwards with the same maturity and different strikes via the so-called *box* strategy: i.e. a position composed by a long synthetic forward at a given strike and a short synthetic forward at a different one. For this strategy -that is equivalent to a long or short cash position- we can neglect margin (MVA) and capital (KVA) adjustment.

[6] We consider the CBOE European options on the S&P 500 index (option prices are reported by the U.S.A. Options Price Reporting Authority) and the Eurex European options on the EURO STOXX 50 index. Eikon Reuters option chains are respectively *0#SPX*.U* and *0#STXE*.EX*.

[7] For each value date t_0 we observe 10 to 13 liquid synthetic forward maturities in the S&P 500 market and 18 to 19 contracts' maturities in the EURO STOXX 50.

market	quantity	mean	median	std	$q_{0.05}$	$q_{0.95}$
S&P 500	#Strikes	131	95	67	77	250
S&P 500	C+P	585.70	488.00	396.56	130.02	1389.42
S&P 500	$\mathcal{G}_T$+K	2794.00	2796.75	118.61	2583.45	2991.30
EURO STOXX 50	#Strikes	44	46	19	16	73
EURO STOXX 50	C+P	584.59	533.55	331.27	172.90	1226.35
EURO STOXX 50	$\mathcal{G}_T$+K	3201.73	3202.50	170.15	2918.05	3485.50

Table 4.1: Descriptive statistics. Mean, median, standard deviation (std), and quantiles (q) 5%, 95% of some relevant quantities in the options' dataset we analyze. We report the number of strikes for each maturity and value date, the straddle position, $C_T(K) + P_T(K)$, and the synthetic forward plus the strike, $\mathcal{G}_T(K) + K$.

The dataset also includes the OIS rates at 11:00 am NT (USD and EUR) with a time-to-maturity equal to 1-12, 15, 18, and 21 months and 2, 3, 4 and 5 years. The OIS interest rate curve is bootstrapped following the standard methodology (see, e.g., Henrard 2014, Baviera and Cassaro 2015). Eikon Reuters provides all financial data.
The dataset provides call/put bid and ask prices for each available maturity. Data pre-processing criteria are simple: we filter out the options that do not satisfy two basic liquidity criteria and we discard maturities with just one or two strikes. As first liquidity criterion, we filter the so-called "penny options", i.e. options at a very low price. All options, whose value is less than 0.1 (S&P 500 or EURO STOXX 50) index points, fall within this class. Then, options with a wide bid-ask spread are discarded. We filter out options with a ratio ask-bid/ask larger than 60%. This second liquidity criterion excludes strikes for which either bid or ask prices for call and put options are not available.

4.3 S&P 500 and EURO STOXX 50 implicit interest rates

In this section, we infer the market discount factor from option prices and analyze it for the two option markets. We verify whether the market discount factor corresponds to the EUR or USD OIS curve and find statistical evidence that a cost-of-funding of 34 basis points is added to the OIS curve in the S&P 500 case.

In the market, we observe bid and ask prices for every different strike and a fixed maturity. The bid synthetic forward is obtained by selling the call and buying the corresponding put, vice-versa for the ask price. Mid prices are the arithmetic average of bid and ask prices.

$$\begin{cases} \mathcal{G}_T^{bid}(K) & := C^{bid}(K) - P^{ask}(K) \\ \mathcal{G}_T^{ask}(K) & := C^{ask}(K) - P^{bid}(K) \\ \mathcal{G}_T(K) & := \dfrac{\mathcal{G}_T^{bid}(K) + \mathcal{G}_T^{ask}(K)}{2} . \end{cases} \tag{4.3.1}$$

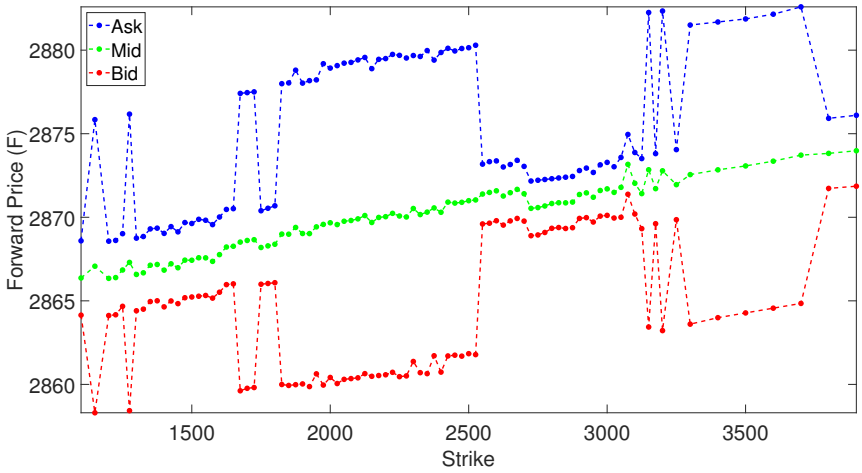

Figure 4.1: S&P 500 forward prices observed at 11:00 am NT of the 1^{st} of April 2019 with maturity on the 21^{st} of March 2020. Forward prices are obtained via synthetic forwards in (4.2.2) for different strikes: we assume $\overline{B}_T = B_T$, the USD OIS discount factor. We plot in red the bid prices, in blue the ask prices, and in green the mid prices. Notice that the mid prices are linear w.r.t. the strike. This fact denotes a market implied discount factor $\overline{B}(t_0, T)$ lower than the USD OIS one, as explained in the text.

In Figure 4.1, we plot an example of one year S&P 500 forward prices $F_{t_0}(T)$ in (4.2.2) using the discount factor B_T of the USD OIS curve obtained via the bootstrap. We can notice a linear behavior w.r.t. the strikes.
By non-arbitrage principle, the forward $F_{t_0}(T)$ should be constant in K. Thanks to equation (4.2.2), we observe from Figure 4.1 that also the ratio $\mathcal{G}_T(K)/B_T$ is a decreasing linear function of K, but with an angular coefficient greater than -1, because it cannot compensate the linear term K in (4.2.2). Hence, in absolute value, the actual angular coefficient of $\mathcal{G}_T(K)/\overline{B}_T$ should be larger that the one of $\mathcal{G}_T(K)/B_T$: we infer that the actual discount $\overline{B}_T$ is lower than the OIS one B_T.

The discount factor used in the market $\overline{B}_T$ can be obtained as the angular coefficient in the linear regression

$$\mathcal{G}_{i,T} = -\overline{B}_T\, K_i + \overline{B}_T\, F_{t_0}(T) + \epsilon_i \qquad i = 1, .., N \tag{4.3.2}$$

for the different strikes $\{K_i\}_{i=1,..,N}$ available at value date t_0 and maturity T, where ϵ_i are some error variables. Its least squares estimation is

$$\overline{B}_T = -\frac{\sum_{i=1}^N (K_i - \hat{K})(\mathcal{G}_{i,T} - \hat{\mathcal{G}}_T)}{\sum_{i=1}^N (K_i - \hat{K})^2} \tag{4.3.3}$$

where

$$\hat{\mathcal{G}}_T := \frac{1}{N}\sum_{i=1}^N \mathcal{G}_{i,T} \quad , \quad \hat{K} := \frac{1}{N}\sum_{i=1}^N K_i\,. \tag{4.3.4}$$

We observe that the regressions are very precise with an R^2 above 0.9995 for all value dates t_0 and all maturities T in the dataset analyzed.

This result is equivalent to state that a spread is added to the USD OIS curve. The funding spread (or cost-of-funding) can be defined in several ways; the simplest one is

$$\bar{s} := \frac{1}{T - t_0} \ln \frac{B_T}{\overline{B}_T} \tag{4.3.5}$$

where B_T has been obtained from the bootstrap of the OIS curve and time intervals are measured according to an $Act/365$ convention.[8] The elementary indicator (4.3.5) allows the market makers to monitor in real-time the cost of funding in the derivative market where they operate; it allows also to detect possible situations of stress in funding liquidity.

We measure this spread for all value dates t_0 and all maturities T in the whole options' dataset. In Figure 4.2 we plot the spread over the USD OIS curve w.r.t. the synthetic forward time-to-maturity (ttm) $T - t_0$. It seems that a spread of 34 basis points is applied to the USD OIS curve for maturities higher than one month.

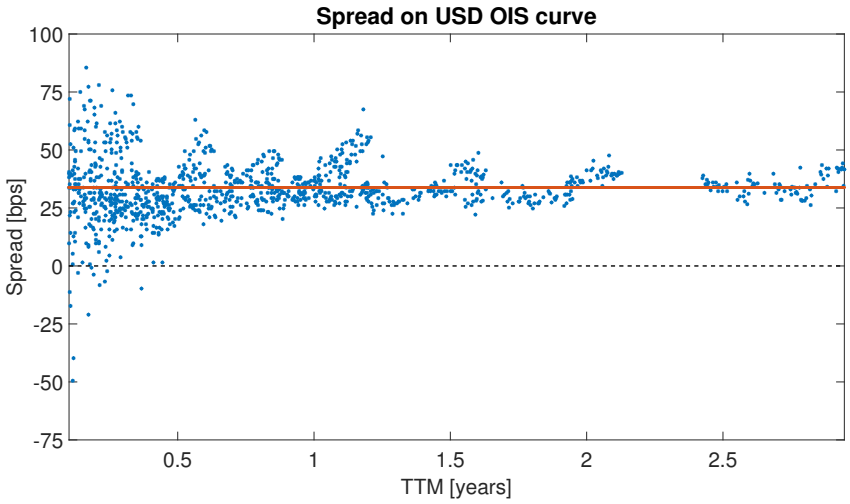

Figure 4.2: Spread over USD OIS in the S&P 500 case. The spread (4.3.5) is plotted against the time-to-maturity (ttm), for ttm longer than one month. The average spread of 34 basis points over the OIS curve seems constant over the different maturities (continuous red line). We observe a higher variance for short term maturities.

We fit the spread $\bar{s}$ as a function of the ttm and we test the statistical significance of the results. We can accept the null hypothesis of no slope with a p-value of 11% and we reject the null hypothesis of zero intercept with a p-value below 10^{-16}. The intercept estimated assuming no slope is of 34 basis points.

We follow the same procedure for the EURO STOXX 50 forward prices, the spread over the EUR OIS curve is reported in Figure 4.3.

[8] This is equivalent –up to a fraction of basis point– to consider a cost-of-funding $\bar{s}$ over the overnight rate OR.

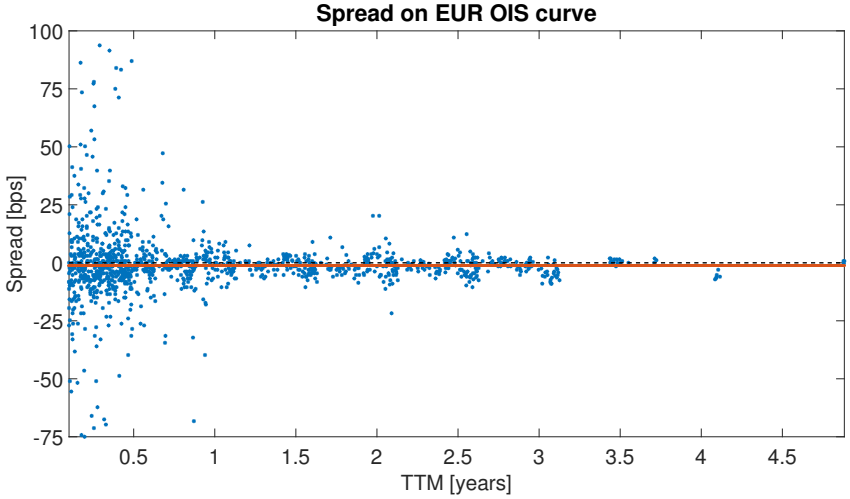

Figure 4.3: Spread over the EUR OIS curve in the EURO STOXX 50 case. The spread (4.3.5) is plotted against the ttm, for ttm larger than one month. The average spread seems to be zero over the different maturities (continuous red line).

We accept the null hypothesis of no intercept with a p-value of 23% and we accept the null hypothesis of no slope with a p-value of 81%. In Table 4.2 we report a summary of the estimated slope and intercept parameters together with the statistical test p-values for both option markets. We can conclude that dealers in the S&P 500 market are subjected to a cost-of-funding, constant w.r.t. the ttm, on average of 34 basis points, the same does not apply for dealers in the EURO STOXX 50 market.

market	**parameter**	**estimate**	**p-value**
S&P 500	Intercept	33	$< 10^{-16}$
S&P 500	Slope	1	0.11
EURO STOXX 50	Intercept	-1	0.23
EURO STOXX 50	Slope	0	0.81

Table 4.2: Spread over OIS. Estimated intercept and slope of the spread over the OIS curve in basis points (USD OIS curve for S&P 500 and EUR OIS curve for EURO STOXX 50). We accept the null hypotheses of no slope for both markets. We refuse the null hypothesis of zero intercept only for the S&P 500 market. There is statistical evidence that dealers in the S&P 500 are subjected to a cost-of-funding, constant w.r.t. the ttm: the intercept estimated assuming no slope is of 34 basis points. No spread is observed for the EURO STOXX 50.

We observe in both Figure 4.2 and 4.3 a higher variance for short maturities. This is due to the fact that only the product of the spread and the time-to-maturity is relevant for the forward: for shorter maturities, the no-arbitrage condition is granted by a larger range of values for the spread.

Four robustness tests are performed. (i) We fit a weighted linear regression (see, e.g., Strutz 2010, Ch.3, p.51) to tackle heteroskedasticity problems. The weights are selected as one over the square of the linear regression residuals. (ii) We change the penny-option and the bid-ask spread thresholds respectively in the range (0, 1) and (30%, 90%) to verify the robustness w.r.t. the excluded strikes. (iii) We extend the analysis window up to the 1^{st} of October 2019 (the last date before the EONIA is discontinued[2]) and limit the analysis to ttm larger than either 6

or 12 months. (iv) We do not discard the maturities with less than three valid strikes. In all robustness tests, results do not change up to a single basis point.

Let us underline that this methodology allows us to determine also the forward price obtained via synthetic forwards. We will use it, in section 4.4, to calibrate the volatility surfaces (on a time range of nine years) with the ATS process. This forward price is obtained from the put-call parity relation (4.2.2) using the $\overline{B}_T$ that includes the cost-of-funding: at a given maturity T, the forward ask price is the lowest forward ask in (4.2.2) and the forward bid price is the highest forward bid. In Figure 4.4 we show an example of the forward bid and ask prices obtained in this way.

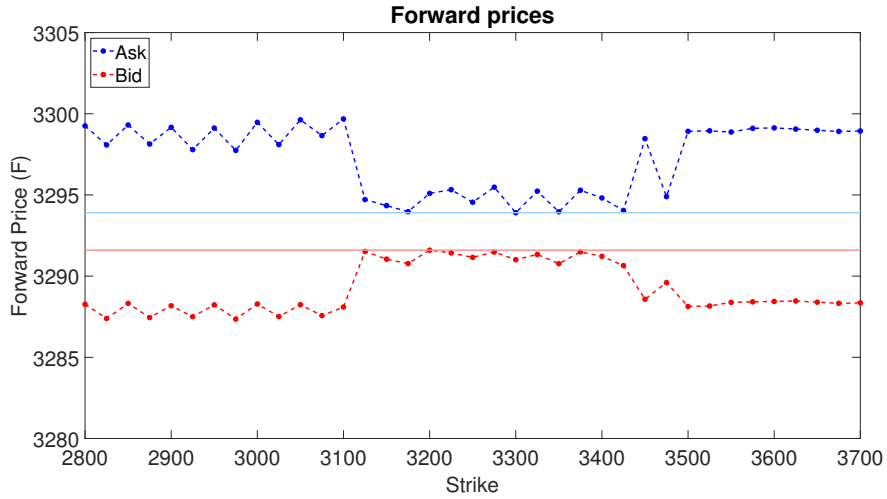

Figure 4.4: Example of the construction of the forward price via synthetic forwards. We show bid (in red) and ask (in blue) forward prices (4.2.2) of the EURO STOXX 50 at the 1^{st} of April 2019 for the 21^{st} of June 2019 maturity. Only prices not discarded by the two liquidity criteria described in the text are considered. We also plot the forward ask price (continuous light blue) and bid price (continuous light red) obtained as the lowest and highest values respectively. Notice that the length of the bid-ask interval changes with the strikes signaling different liquidity for different strikes.

4.4 ATS calibration on multiple volatility surfaces

In this section, we consider a dataset of closing market prices that goes from January 2012 to December 2020. The dataset comprises of options on the S&P 500 and EURO STOXX 50 indexes. The data is provided by Eikons Datastream. Let us observe that close prices are, in general, less accurate than open market prices (the ones used for the analysis in section 1.3).

For both indices, the most liquid options expire on the third Friday of the first six months after the value date and then on March, June, September, and December in the front year and June and December in the next year. In the EURO STOXX 50 case also June and December contracts for the following years are available. For each value date, we have 9 to 13 maturities in the S&P 500 market and 18 to 19 maturities in the EURO STOXX 50.

We filter out the options that do not satisfy the two basic liquidity criteria explained in section 4.2. We filter the so-called "penny options" and options with a ratio ask-bid/ask

larger than 60%. This second liquidity criterion excludes strikes for which either bid or ask prices for call and put options are not available. Moreover, we do not consider maturities with less than four valid strikes and we discard options with B&S delta outside of the range (0.25,0.75).

The ATS is introduced through its characteristic function in equation (1.2.4)

$$\mathbb{E}\left[e^{iuf_t}\right] = \mathcal{L}_t\left(iu\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 + \frac{u^2\sigma_t^2}{2};\ k_t,\ \alpha\right)e^{iu\varphi_t t} \quad , \tag{4.4.1}$$

where σ_t, k_t are continuous on $[0,\infty)$ and η_t, φ_t are continuous on $(0,\infty)$ with $\sigma_t > 0$, $k_t \geq 0$ and $\varphi_t\, t$ goes to zero as t goes to zero. $\alpha \in [0,1)$ as in the LTS case and

$$\ln \mathcal{L}_t\left(u;\ k,\ \alpha\right) := \begin{cases} \dfrac{t}{k}\dfrac{1-\alpha}{\alpha}\left\{1-\left(1+\dfrac{u\,k}{1-\alpha}\right)^{\alpha}\right\} & \text{if } \ 0<\alpha<1 \\ -\dfrac{t}{k}\ln\left(1+u\,k\right) & \text{if } \ \alpha = 0 \end{cases} .$$

Moreover, we impose the martingality of the ATS by setting

$$t\varphi_t := -\log \mathcal{L}_t\left(\eta_t\sigma_t^2;\ k_t,\ \alpha\right) \quad .$$

We calibrate the ATS following the procedure discussed in subsection 1.3.2. We cut the volatility surface into slices, each one containing options with the same maturity, and calibrate each slice separately. Also, in this case, we focus on $\alpha = 1/2$ (NIG) and $\alpha = 0$ (VG), which are the two ATS generalizations of the two most frequently used LTS processes. The three time-dependent parameters k_T, η_T, σ_T are calibrated separately for each maturity but the calibration is performed imposing the conditions of monotonicity of **Theorem 1.2.1**.
As in chapter 1, we consider also the calibration of the standard Lévy processes and the (four parameters) Sato processes proposed by Carr *et al.* (2007). We remind that the latter are additive and self-similar processes (see, e.g., Sato 1991). Call option prices, with strike K and maturity T, are computed using the Lewis (2001) formula

$$C_T(x) = B_T\ F_0(T)\left\{1 - e^{x/2}\int_{-\infty}^{\infty}\frac{dz}{2\pi}e^{iz\,x}\phi^c\left(-z-\frac{i}{2}\right)\frac{1}{z^2+\frac{1}{4}}\right\} \quad , \tag{4.4.2}$$

where $\phi^c(u)$ is the characteristic function of f_T, $x := \ln K/F_0(T)$ is the *moneyness*, and B_T is the discount factor between value date and T.
The calibration is performed by minimizing the distance between model and market prices. The simplex method is used to calibrate every maturity of the ATS process. For Lévy processes and Sato processes, because standard routines for global minimum algorithms are not satisfactory, we consider a multi-start trust-region-reflective method.

In tables 4.3 and 4.4 we report the average calibration performance for the S&P 500 and EURO STOXX 50 in terms, respectively, of MSE and MAPE for the Lévy process, the Sato process, and the ATS process. The period of analysis goes from January 2012 to December 2020. The results are for closing option prices. In the NIG ($\alpha = 1/2$) and VG ($\alpha = 0$) cases, we consider the standard Lévy process, the Sato process, and the corresponding ATS process. Sato processes perform better than Lévy processes but ATS improvement is far more significant: more than two orders of magnitude of MSE and more than one order of magnitude of MAPE.

Index	Model	Lévy	Sato	ATS
S&P 500	NIG	108.37	7.39	**0.56**
S&P 500	VG	120.20	7.24	**0.64**
Euro Stoxx 50	NIG	200.25	22.77	**4.11**
Euro Stoxx 50	VG	214.86	23.51	**4.91**

Table 4.3: Average calibration performance for the S&P 500 and EURO STOXX 50 in terms of MSE for the Lévy process, the Sato process, and the ATS process. The period of analysis goes from January 2012 to December 2020. In the NIG ($\alpha = 1/2$) and VG ($\alpha = 0$) cases, we consider the standard Lévy process, the Sato process, and the corresponding ATS process. Sato processes perform better than Lévy processes but ATS improvement is far more significant: more than two orders of magnitude of MSE.

Index	Model	Lévy	Sato	ATS
S&P 500	NIG	9.59%	5.05%	**0.43**%
S&P 500	VG	12.07%	4.91%	**0.89**%
Euro Stoxx 50	NIG	11.04%	3.98%	**0.31**%
Euro Stoxx 50	VG	9.87%	4.00%	**0.59**%

Table 4.4: Average calibration performance for the S&P 500 and EURO STOXX 50 in terms of MAPE for the Lévy process, the Sato process, and the ATS process. The period of analysis goes from January 2012 to December 2020. In the NIG ($\alpha = 1/2$) and VG ($\alpha = 0$) cases, we consider the standard Lévy process, the Sato process, and the corresponding ATS process. Sato processes perform better than Lévy processes but ATS improvement is far more significant: the MAPE is almost two orders of magnitude below the Lévy case.

In Figure 4.5, we report MSE for the ATS (red circles), the self-similar process (green ex) and the Lévy process (black squares) for S&P 500 closing option prices. The MSE are reported in log10 scale. The period of analysis goes from January 2012 to December 2020. We observe that both in the NIG (on the left) and VG (on the right) case the MSE of the ATS calibrated slice by slice are, on average, more than two orders of magnitude better than the Lévy and more than one order of magnitude better than the self-similar case.

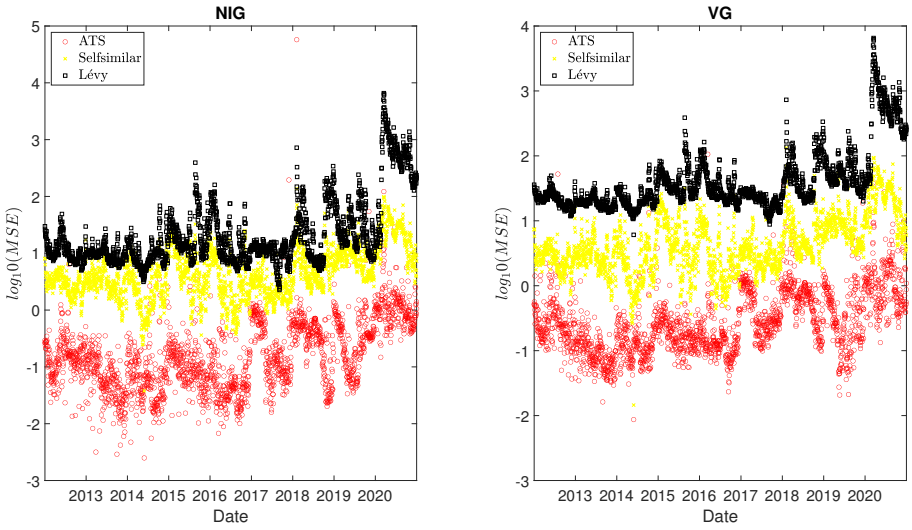

Figure 4.5: MSE for the ATS (red circles), the self-similar process (green ex) and the Lévy process (black squares) for S&P 500 closing option prices. The MSE are reported in log10 scale. The period of analysis goes from from January 2012 to December 2020. We observe that both in the NIG (on the left) and VG (on the right) case the MSE of the ATS calibrated slice by slice are, on average, more than two orders of magnitude better than the Lévy and more than one order of magnitude better that the self-similar case.

In Figure 4.6, we report the MSE for the ATS (red circles), the self-similar process (green ex), and the Lévy process (black squares) for EURO STOXX 50 closing option prices. The MSE are reported in log10 scale. The period of analysis goes from January 2012 to December 2020. We observe that both in the NIG (on the left) and VG (on the right) case the MSE of the ATS calibrated slice by slice are, on average, more than two orders of magnitude better than the Lévy and more than one order of magnitude better that the self-similar case.

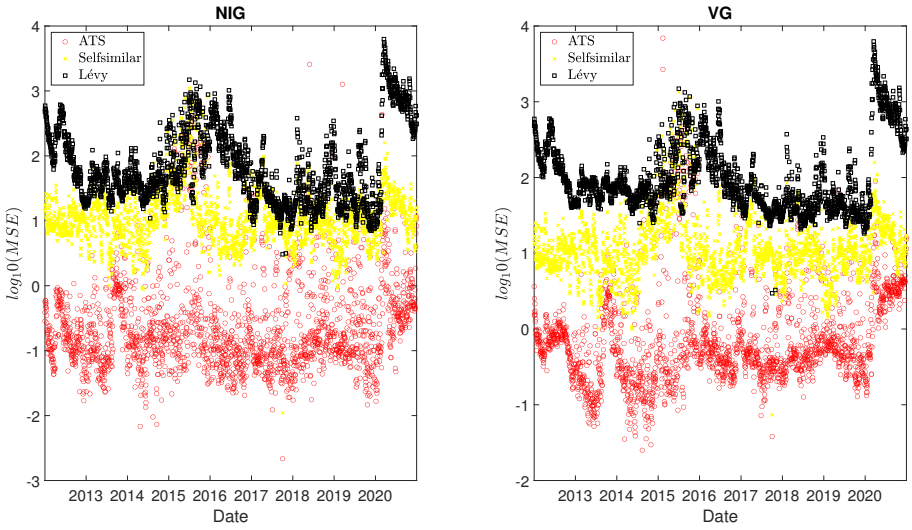

Figure 4.6: MSE for the ATS (red circles), the self-similar process (green ex) and the Lévy process (black squares) for EURO STOXX 50 closing option prices. The MSE are reported in log10 scale. The period of analysis goes from from January 2012 to December 2020. We observe that both in the NIG (on the left) and VG (on the right) case the MSE of the ATS calibrated slice by slice are, on average, more than two orders of magnitude better than the Lévy and more than one order of magnitude better that the self-similar case.

4.5 Conclusions

Which discount factor should be used in exchange-traded derivatives? This study exploits the implications of the put-call parity to develop a methodology that allows us to recover the discount factor implied by option prices on S&P 500 and EURO STOXX 50 indices. A dealer in the option market can use this technique to real-time monitor the funding spreads of market players. The implicit discount factor is the one such that a forward contract, built using the put-call parity relation, does not depend on the strike. We compute the S&P 500 and EURO STOXX 50 option markets' implicit discount factors and evaluate the cost-of-funding over the curve obtained bootstrapping OIS derivatives. We have statistical evidence of a cost-of-funding of, on average, 34 basis points on top of the USD OIS curve in the S&P 500 case and no cost-of-funding for the EURO STOXX 50 case. This cost-of-funding is constant for all liquid maturities up to several years for both markets.

Hence, the natural question is: why do we observe a spread over USD and no spread over EUR OIS? The reason should be sought in the differences between the two underlying money markets. Let us remind that the FED Target range indicates only some target rates, while in Europe the corridor system denotes the real rates at which ECB serves as lender of last resort to the financial system. In the USD market the two rates, the collateralized one (SOFR) and the uncollateralized one (EFFR), differ for a spread that can be significant in several days. A dysfunctional repo market, indicated by a sharp spike in the SOFR, has been observed several days in the analyzed period;[9] besides, a disruption in the repo market has been signaled by

[9]For example, during the period of analysis starting from the 1^{st} of November 2018 and ending on the 1^{st} of

the well-known repo blow-up of the 17^{th} of September 2019, when the spread between the two fixings (SOFR and EFFR) topped to almost 3 percent and prompted the Federal Reserve to inject tens of billions of dollars of reserves into money markets (see e.g., Barrett and Hamilton 2020, Tilford *et al.* 2019). While, as we have already underlined in the introduction, the OIS market is mainly based on the EFFR, the volumes in the money market are mostly concentrated on the repo rate, with SOFR volumes more than ten times larger than EFFR ones (see, e.g., Schrimpf and Sushko 2019). These funding disruptions in the USD money market, not observed in the EUR market, suggest that market players could require a spread over OIS as a premium for the additional liquidity risk observed in this market. As for future research, one main promising direction is evident: it could be interesting to understand whether this funding spread is connected to the implied/historical volatility on the two indices.

The implied discount factor and the synthetic forward prices are used to calibrate the ATS over a nine-year dataset. The excellent calibration performance presented in chapter 1 can be observed also for every date of this nine-year dataset (c.f. tables 4.3-4.4 and figures 4.5-4.6)

October 2018, apart from the end of months (and, in particular, the End-of-Year), spikes outside the FED Fund range are observed in 15 days. Let us notice that, in all these days, SOFR is always larger than the upper side of the range.

Chapter 5

Conclusions

In this thesis, we introduce a new broad family of stochastic processes that we call additive normal tempered stable processes (ATS). This is a family of pure jump additive processes that has excellent calibrations features and reproduces the short-time equity skew. The ATS maintains several properties (both analytic and numerical) of the LTS processes. Among these, we have the simple and fast Lewis formula for European pay-off in (1.3.1); this formula allows an efficient calibration. We discuss the four key advantages of ATS when modeling the implied volatility in equity markets.

The first argument in favor of the ATS is a high calibration quality.
We consider a dataset of closing market prices that goes from January 2012 to December 2020. The dataset comprises of options on the S&P 500 and EURO STOXX 50 indexes with a time-to-maturity starting from one day and up to several years. The observed improvement of ATS calibration, w.r.t. to LTS (Lévy) and Sato alternatives, is stunning: at least two orders of magnitude in terms of MSE and one order of magnitude of MAPE, as reported in tables 4.3-4.4 and Figures 4.5-4.6. Moreover, ATS perfectly reproduces market implied volatility term structure and *skew* as observed in Figures 1.5-1.6.

The second argument in favor of the ATS is that it is a parsimonious pure jump process.
A pure jump process, by construction, describes underlying dynamics more parsimoniously (see, e.g., Asmussen and Rosiński 2001) than a standard jump diffusion process. For this reason, in general, a pure jump process calibration performances are not extraordinary (see, e.g., Cont and Tankov 2003, Ch.14). The ATS is parsimonious also because a power-law scaling characteristic arises in calibrated parameters as we can observe in Figures 1.7-1.8 and 1.14-1.15 . Market data are consistent with $\beta = 1$ (linked to the variance of jumps) and $\delta = -1/2$ (linked to the skew). Surprisingly, this characteristic holds both for short and long maturities: different players characterize option markets on short and long maturities. By considering a power-law scaling ATS with fixed β and δ, after the volatility term structure is taken into account, the surface is calibrated with just two free parameters.

The third argument in favor of the ATS is that a pure jump process, differently from the Lévy case, reproduces the equity short time implied volatility thanks to the scaling properties observed in market data.
We discover that this power-law scaling characteristic is linked to the ATS short time asymptotic behavior.
We investigate the short-time ATM implied volatility and skew over the region of admissible power-law scaling parameters (cf. **Theorem 1.2.3**). As discussed in the introduction of chapter

2, in the short time, the equity market is characterized by a positive and finite ATM implied volatility and a negative skew that goes as $t^{-1/2}$. We prove that only the scaling parameters estimated from market data ($\beta = 1$ and $\delta = -1/2$) are compatible with the implied volatility observed in the equity market (cf. **Theorem 2.5.1**). We can link the ATS parsimony to the short-time behavior under this light. The parameters $\beta = 1$ and $\delta = -1/2$ are observed in market data because they are the unique configuration that is compatible with the equity short time characteristics.

Finally, the fourth element in favor of the ATS is that a fast and accurate Monte Carlo method exists for path-dependent exotics.
Since the probability density function of ATS increments is not explicit, we have to design a simulation algorithm to price path-dependent exotics. We propose a new technique for additive processes that builds upon the efficiency of the FFT, the Lewis formula for a CDF, and the spline interpolation properties when inverting the CDF. We call Lewis-FFT-S this Monte Carlo simulation.
In Figure 3.2 we analyze the three-components of the bias error in (3.3.1). We improve the method's convergence by using spline interpolation for the numerical inversion. In this way, the leading term in the bias improves from γ^2 to at least γ^4, where γ is the step of the interpolation grid. We emphasize this improvement in Figures 3.3-3.4, where, for $M > 7$, the Lewis-FFT-S (with spline interpolation) clearly outperforms the Lewis-FFT version of the method with linear interpolation. The bias appears to decrease as γ^6 in numerical experiments. The Lewis-FFT-S is accurate but also fast. As discussed in subsection 3.3.1, the additional computational time due to spline interpolation is negligible. Moreover, the method computational time is of the same order of magnitude as standard algorithms for Brownian motions.

This thesis is a modeling breakthrough in equity derivatives that could change significantly the financial engineering and the risk management of this sector in the near future.

Acknowledgements

I thank my supervisor R. Baviera for his guidance, his patience and his availability during these "new normal" years. I thank E. Alòs, F. Benth, P. Carr and J. Gatheral for enlightening discussions on the topic of this thesis. I thank also C. Sgarra, D. Marazzina, E. Barucci, G. Giuffra, G. Guatteri, G. Ding, M.P. Gregoratti, J. Guyon, M. Fukasawa, P. Spreij, and all participants to WSMF 2019 in Lunteren, to VCMF 2019 in Vienna, to AFM 2020 in Paris and to new Challenges in Quantitative Finance 2021 in Barcelona. I am grateful to M. Beretta (Gardena Capital), A. Cassaro (Goldman Sachs), L. Fusco (Unicredit), F. Gregori (BNPP), P. Liberti (Cerved), E. Mercuri (Illimity), G. Polidoro (Mediobanca), M. Potente (Société Général), T. Santagostino Baldi (Morgan Stanley), M. Spadaccino (Illimity), L. Viola (Arca SGR) and A. Vite (Unicredit).

Bibliography

Abate, J. and Whitt, W., 1992. The fourier-series method for inverting transforms of probability distributions, *Queueing systems*, 10 (1), 5–87.

Abramowitz, M. and Stegun, I.A., 1948. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55, US Government printing office.

Alòs, E., León, J.A., and Vives, J., 2007. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility, *Finance and Stochastics*, 11 (4), 571–589.

Andersen, L. and Lipton, A., 2013. Asymptotics for exponential lévy processes and their volatility smile: survey and new results, *International Journal of Theoretical and Applied Finance*, 16 (1), 1350001.

Asmussen, S. and Glynn, P.W., 2007. *Stochastic simulation: algorithms and analysis*, vol. 57, Springer.

Asmussen, S. and Rosiński, J., 2001. Approximations of small jumps of Lévy processes with a view towards simulation, *Journal of Applied Probability*, 38 (2), 482–493.

Azzone, M. and Baviera, R., 2021a. Additive normal tempered stable processes for equity derivatives and power-law scaling, *Quantitative Finance*, 1–18.

Azzone, M. and Baviera, R., 2021b. A fast monte carlo scheme for additive processes and option pricing, *arXiv preprint arXiv:2112.08291*.

Azzone, M. and Baviera, R., 2021c. Short-time implied volatility of additive normal tempered stable processes, *arXiv preprint arXiv:2108.02447*.

Azzone, M. and Baviera, R., 2021d. Synthetic forwards and cost of funding in the equity derivative market, *Finance Research Letters*, 41, 101841.

Bai, Y., Pan, Z., and Liu, L., 2019. Improving futures hedging performance using option information: Evidence from the S&P 500 index, *Finance Research Letters*, 28, 112–117.

Bailey, A., 2017. The future of LIBOR. speech at Bloomberg London., *Available online: https://www.fca.org.uk/news/speeches/the-future-of-libor*.

Ballotta, L. and Kyriakou, I., 2014. Monte Carlo simulation of the CGMY process and option pricing, *Journal of Futures Markets*, 34 (12), 1095–1121.

Ballotta, L. and Rayée, G., 2018. Smiles & smirks: a tale of factors, *Available at SSRN 2980349*.

Barndorff-Nielsen, O.E., 1997. Normal inverse Gaussian distributions and stochastic volatility modelling, *Scandinavian Journal of Statistics*, 24 (1), 1–13.

Barrett, E. and Hamilton, J., 2020. Why the U.S. repo market blew up and how to fix it, *Bloomberg*, (6^{th} of January).

Bartle, R.G. and Bartle, R.G., 1995. *The elements of integration and Lebesgue measure*, vol. 27, Wiley Online Library.

Baviera, R., 2007. Gigi model: a simple stochastic volatility approach for multifactor interest rates, *Available at SSRN 977479*.

Baviera, R. and Cassaro, A., 2015. A note on dual-curve construction: Mr. Crab's Bootstrap, *Applied Mathematical Finance*, 22 (2), 105–132.

Benth, F.E., Kholodnyi, V.A., and Laurence, P., 2014. Quantitative energy finance, *Modelling, pricing, and hedging in energy and commodity markets, Springer*.

Benth, F.E. and Sgarra, C., 2012. The risk premium and the Esscher transform in power markets, *Stochastic Analysis and Applications*, 30 (1), 20–43.

Bethea, R.M., 2018. *Statistical methods for engineers and scientists*, Routledge.

Black, F., 1976. The pricing of commodity contracts, *Journal of financial economics*, 3 (1-2), 167–179.

Bohman, H., 1970. A method to calculate the distribution function when the characteristic function is known, *BIT Numerical Mathematics*, 10 (3), 237–242.

Brace, A., G̨atarek, D., and Musiela, M., 1997. The market model of interest rate dynamics, *Mathematical Finance*, 7 (2), 127–155.

Brenner, M. and Galai, D., 1986. Implied interest rates, *Journal of Business*, 493–507.

Briat, C., 2011. Convergence and equivalence results for the jensen's inequality—application to time-delay and sampled-data systems, *IEEE Transactions on Automatic Control*, 56 (7), 1660–1665.

Brigo, D., Morini, M., and Pallavicini, A., 2013. *Counterparty credit risk, collateral and funding: with pricing cases for all asset classes*, vol. 478, John Wiley & Sons.

Carr, P., Geman, H., Madan, D.B., and Yor, M., 2007. Self-decomposability and option pricing, *Mathematical Finance*, 17 (1), 31–57.

Carr, P. and Madan, D., 1999. Option valuation using the fast Fourier transform, *Journal of Computational Finance*, 2 (4), 61–73.

Carr, P. and Torricelli, L., 2021. Additive logistic processes in option pricing, *Finance and Stochastics*, 25, 689–724.

Carr, P. and Wu, L., 2003. The finite moment log stable process and option pricing, *The Journal of Finance*, 58 (2), 753–777.

Černỳ, A. and Kyriakou, I., 2011. An improved convolution algorithm for discretely sampled Asian options, *Quantitative Finance*, 11 (3), 381–389.

Chen, Z., Feng, L., and Lin, X., 2012. Simulating lévy processes from their characteristic functions and financial applications, *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 22 (3), 1–26.

Chourdakis, K., 2005. Option pricing using the fractional FFT, *Journal of computational finance*, 8 (2), 1–18.

Cont, R. and Tankov, P., 2003. *Financial Modelling with jump processes*, Chapman and Hall/CRC Financial Mathematics Series.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C., 2001. *Introduction to algorithms*, MIT Press.

Cressie, N., Davis, A.S., Folks, J.L., and Folks, J.L., 1981. The moment-generating function and negative integer moments, *The American Statistician*, 35 (3), 148–150.

Dash, S. and Liu, B., 2009. Capturing the index effect via options, *The Journal of Trading*, 4 (2), 72–78.

Dini, U., 1907. *Lezioni di analisi infinitesimale*, vol. 1, Fratelli Nistri.

Durrett, R., 2019. *Probability: theory and examples*, vol. 49, Cambridge university press.

Eberlein, E. and Madan, D.B., 2009. Sato processes and the valuation of structured products, *Quantitative Finance*, 9 (1), 27–42.

Feng, L. and Lin, X., 2013. Inverting analytic characteristic functions and financial applications, *SIAM Journal on Financial Mathematics*, 4 (1), 372–398.

Figueroa-López, J.E. and Forde, M., 2012. The small-maturity smile for exponential Lévy models, *SIAM Journal on Financial Mathematics*, 3 (1), 33–65.

Figueroa-López, J.E., Gong, R., and Houdré, C., 2016. High-order short-time expansions for atm option prices of exponential lévy models, *Mathematical Finance*, 26 (3), 516–557.

Figueroa-López, J.E., Gong, R., and Lorig, M., 2018. Short-time expansions for call options on leveraged ETFs under exponential Lévy models with local volatility, *SIAM Journal on Financial Mathematics*, 9 (1), 347–380.

Forde, M. and Jacquier, A., 2009. Small-time asymptotics for implied volatility under the heston model, *International Journal of Theoretical and Applied Finance*, 12 (06), 861–876.

Fouque, J.P., Papanicolaou, G., Sircar, R., and Solna, K., 2004. Maturity cycles in implied volatility, *Finance and Stochastics*, 8 (4), 451–477.

Frankfurter, G.M. and Leung, W.K., 1991. Further analysis of the put-call parity implied risk-free interest rate, *Journal of Financial Research*, 14 (3), 217–232.

Fukasawa, M., 2017. Short-time at-the-money skew and rough fractional volatility, *Quantitative Finance*, 17 (2), 189–198.

Fusai, G. and Kyriakou, I., 2016. General optimized lower and upper bounds for discrete and continuous arithmetic Asian options, *Mathematics of Operations Research*, 41 (2), 531–559.

Fusai, G. and Meucci, A., 2008. Pricing discretely monitored Asian options under lévy processes, *Journal of Banking & Finance*, 32 (10), 2076–2088.

Gatheral, J., 2011. *The volatility surface: a practitioner's guide*, vol. 357, John Wiley & Sons.

Gatheral, J., Hsu, E.P., Laurence, P., Ouyang, C., and Wang, T.H., 2012. Asymptotics of implied volatility in local volatility models, *Mathematical Finance*, 22 (4), 591–620.

Gatheral, J., Jaisson, T., and Rosenbaum, M., 2018. Volatility is rough, *Quantitative Finance*, 18 (6), 933–949.

George, T.J., Kaul, G., and Nimalendran, M., 1991. Estimation of the bid–ask spread and its components: A new approach, *The Review of Financial Studies*, 4 (4), 623–656.

Glasserman, P., 2004. *Monte Carlo methods in financial engineering*, vol. 53, Springer.

Glasserman, P. and Liu, Z., 2010. Sensitivity estimates from characteristic functions, *Operations Research*, 58 (6), 1611–1623.

Gregory, J., 2012. *Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets*, Wiley Finance.

Hall, C.A. and Meyer, W.W., 1976. Optimal error bounds for cubic spline interpolation, *Journal of Approximation Theory*, 16 (2), 105–122.

Hao, J., He, F., Liu-Chen, B., and Li, Z., 2020. Price discovery and its determinants for the Chinese soybean options and futures markets, *Finance Research Letters*, in press.

Henrard, M., 2014. *Interest Rate Modelling in the Multi-curve Framework: Foundations, Evolution, and Implementation*, Palgrave Macmillan.

Henrard, M.P., 2019. LIBOR fallback and quantitative finance, *Risks*, 7 (3), 1–15.

Hull, J. and White, A., 1987. The pricing of options on assets with stochastic volatilities, *The Journal of Finance*, 42 (2), 281–300.

Hull, J.C., 2003. *Options futures and other derivatives*, Pearson Education India.

Hull, J.C. and White, A., 2013. LIBOR vs. OIS: The derivatives discounting dilemma, *The Journal of Investment Management*, 11 (3), 14–27.

Jacod, J., 2007. Asymptotic properties of power variations of Lévy processes, *ESAIM: Probability and Statistics*, 11, 173–196.

Kao Edward, P., 1997. *An introduction to stochastic processes, 1ˆ'éd*, Belmont, Calif: Duxbury Press, 438p.

Kijima, M., 1997. *Markov processes for stochastic modeling*, vol. 6, CRC Press.

Konikov, M. and Madan, D.B., 2002. Option pricing using variance gamma Markov chains, *Review of Derivatives Research*, 5 (1), 81–115.

Küchler, U. and Tappe, S., 2013. Tempered stable distributions and processes, *Stochastic Processes and their Applications*, 123 (12), 4256–4293.

Kyriakou, I., Pouliasis, P.K., and Papapostolou, N.C., 2016. Jumps and stochastic volatility in crude oil prices and advances in average option pricing, *Quantitative Finance*, 16 (12), 1859–1873.

Lee, R.W., 2004. Option pricing by transform methods: extensions, unification and error control, *Journal of Computational Finance*, 7 (3), 51–86.

Lewis, A.L., 2001. A simple option formula for general jump-diffusion and other exponential Lévy processes, *Mimeo*.

Li, J., Li, L., and Mendoza-Arriaga, R., 2016. Additive subordination and its applications in finance, *Finance and Stochastics*, 20 (3), 589–634.

Li, L. and Linetsky, V., 2014. Time-changed ornstein–uhlenbeck processes and their applications in commodity derivative models, *Mathematical Finance*, 24 (2), 289–330.

Longstaff, F.A., 1995. Option pricing and the martingale restriction, *The Review of Financial Studies*, 8 (4), 1091–1124.

Loomis, L.H. and Sternberg, S., 1990. *Advanced calculus*, Jones and Bartlett Publishers.

Lucia, J.J. and Schwartz, E.S., 2002. Electricity prices and power derivatives: Evidence from the nordic power exchange, *Review of Derivatives Research*, 5 (1), 5–50.

Lukacs, E., 1972. A survey of the theory of characteristic functions, *Advances in Applied Probability*, 4 (1), 1–37.

Madan, D.B., Carr, P.P., and Chang, E.C., 1998. The variance gamma process and option pricing, *Review of Finance*, 2 (1), 79–105.

Madan, D.B. and Seneta, E., 1990. The variance gamma (VG) model for share market returns, *Journal of Business*, 511–524.

Mantegna, R.N. and Stanley, H.E., 1995. Scaling behaviour in the dynamics of an economic index, *Nature*, 376 (6535), 46–49.

Marsaglia, G., Tsang, W.W., *et al.*, 2000. The ziggurat method for generating random variables, *Journal of statistical software*, 5 (8), 1–7.

Medvedev, A. and Scaillet, O., 2006. Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility, *The Review of Financial Studies*, 20 (2), 427–459.

Mijatović, A. and Tankov, P., 2016. A new look at short-term implied volatility in asset price models with jumps, *Mathematical Finance*, 26 (1), 149–183.

Muhle-Karbe, J. and Nutz, M., 2011. Small-time asymptotics of option prices and first absolute moments, *Journal of Applied Probability*, 48 (4), 1003–1020.

Muravyev, D., Pearson, N.D., and Broussard, J.P., 2013. Is there price discovery in equity options?, *Journal of Financial Economics*, 107 (2), 259–283.

Naranjo, L., 2009. Implied interest rates in a market with frictions, *Available at SSRN 1308908*.

Ñíguez, T.M., 2016. Evaluating monthly volatility forecasts using proxies at different frequencies, *Finance Research Letters*, 17, 41–47.

Nomikos, N.K., Kyriakou, I., Papapostolou, N.C., and Pouliasis, P.K., 2013. Freight options: Price modelling and empirical analysis, *Transportation Research Part E: Logistics and Transportation Review*, 51, 82–94.

Ornthanalai, C., 2014. Lévy jump risk: Evidence from options and returns, *Journal of Financial Economics*, 112 (1), 69–90.

Petersen, M.A. and Fialkowski, D., 1994. Posted versus effective spreads: Good prices or bad quotes?, *Journal of Financial Economics*, 35 (3), 269–292.

Petrov, V., 1975. Sums of independent random variables, *Yu. V. Prokhorov. V. StatuleviCius (Eds.)*.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., 1992. *Numerical recipes in C: The Art of Scientific Computing*, Cambridge university press.

Prokopczuk, M., 2011. Pricing and hedging in the freight futures market, *Journal of Futures Markets*, 31 (5), 440–464.

Quarteroni, A., Sacco, R., and Saleri, F., 2007. *Numerical mathematics*, vol. 37, Springer.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of Finance*, 39 (4), 1127–1139.

Ron, U., 2000. A practical guide to swap curve construction, *Bank of Canada Working Paper*, (17), 1–25.

Roper, M., 2009. *Implied volatility: small time to expiry asymptotics in exponential Lévy models*, Ph.D. thesis, Thesis, University of New South Wales.

Rudin, W., 1976. *Principles of mathematical analysis*, 4.2, McGraw-Hill New York.

Rudin, W., 1986. *Real and complex analysis*, Tata McGraw-Hill education.

Ryan, T.P., 2008. *Modern regression methods*, vol. 655, John Wiley & Sons.

Samorodnitsky, G. and Taqqu, M., 1994. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman & Hall, London.

Sato, K.I., 1991. Self-similar processes with independent increments, *Probability Theory and Related Fields*, 89 (3), 285–300.

Sato, K.I., 1999a. *Lévy processes and infinitely divisible distributions*, Cambridge University Press.

Sato, K.I., 1999b. *Lévy processes and infinitely divisible distributions*, Cambridge university press.

Schoutens, W., 2003. *Lévy processes in finance*, Wiley.

Schoutens, W., Simons, E., and Tistaert, J., 2003. A perfect calibration! now what?, *The best of Wilmott*, 281.

Schrimpf, A. and Sushko, V., 2019. Beyond libor: a primer on the new benchmark rates, *BIS Quarterly Review*, (March), 29–52.

Seber, G. and Wild, C., 1989. *Nonlinear regression, 768 pp*, Wiley, New York.

Shiraya, K. and Takahashi, A., 2011. Pricing average options on commodities, *Journal of Futures Markets*, 31 (5), 407–439.

Sørensen, T.M. and Benth, F.E., 2013. Lévy process simulation by stochastic step functions, *SIAM Journal on Scientific Computing*, 35 (5), A2207–A2224.

Stoll, H.R., 1969. The relationship between put and call option prices, *The Journal of Finance*, 24 (5), 801–824.

Strutz, T., 2010. *Data fitting and uncertainty (A practical introduction to weighted least squares and beyond)*, Vieweg Teubner.

Tankov, P., 2011. Pricing and hedging in exponential Lévy models: review of recent results, *Paris-Princeton Lectures on Mathematical Finance 2010*, 319–359.

Taylor, J., 1997. *Introduction to error analysis, the study of uncertainties in physical measurements*, University Science Books, Sausalito, CA.

Tilford, C., Rennison, J., Noonan, L., Smith, C., and Greeley, B., 2019. Repo: How the financial markets' plumbing got blocked, *Financial Times*, (26^{th} of November).

Urbanik, K., 1993. Moments of sums of independent random variables, *in: Stochastic Processes*, Springer, 321–328.

Vo, X.V. and Daly, K., 2008. Volatility amongst firms in the Dow Jones EURO STOXX 50 index, *Applied Financial Economics*, 18 (7), 569–582.

York, D., 1968. Least squares fitting of a straight line with correlated errors, *Earth and Planetary Science Letters*, 5, 320–324.

Notation & Shorthands

Symbol	Description
A_t	diffusion term of the additive process $\{X_t\}_{t\geq 0}$
B_T	discount factor between value date and T
$\overline{B}_T$	market implied discount factor between value date and T
$\mathbb{B}(\mathbb{R})$	Borel sigma algebra on $\mathbb{R}$
$c_t^B\left(\mathcal{I}_t(y), y\right))$	Black call option price
$c_t(S_t, t)$	quantity inside the ATS call expected value
$C_t\left(x\right)$	call option price at value date with maturity t and moneyness x
$C_t\left(K\right)$	call option price at value date with maturity t and strike K
D	FFT simulation regularization parameter
$\{f_t\}_{t\geq 0}$	ATS process that models the forward exponent
$\mathcal{E}$	total error when pricing the derivative with payoff V
$\mathcal{E}_{h,M}^{CDF}$	CDF error bound as a function of the grid size h and of M
$\mathcal{E}_{M}^{CDF}$	CDF error bound when h is s.t. the two sources of error are comparable
$\left\{\hat{f}_\theta\right\}_{\theta\geq 0}$	re-scaled ATS process w.r.t. the time $\theta = \sigma_T^2\, T$
$F_t(T)$	price at time t of a Forward contract with maturity T
g	standard normal random variable
$\mathcal{G}_t(K)$	Synthetic forward with maturity t and strike K
h	grid size in the Fourier domain
$I_t(x)$	Black implied volatility with maturity t and moneyness x
$\mathcal{I}_t(y)$	Black implied volatility with maturity t and *moneyness degree* y
$\mathbb{1}_*$	indicator function of the set $*$
k	variance of jumps of LTS
k_t	variance of jumps of ATS
$\hat{k}_\theta$	re-scaled variance of jumps of ATS
$\bar{k}$	constant part of variance of jumps of ATS
K	option strike price
$\mathcal{K}$	dimension of the CDF interpolation grid
l_t^z	quantity defined in equation (2.2.6)
M	integer number s.t. N is the number of grid points
$\mathcal{L}_t$	Laplace transform of the subordinator Z_t in (1.2.3)
L	Down-and-In barrier strike

Symbol	Description
$m^+_{s,t}$	probability density of positive jumps
$m^-_{s,t}$	probability density of negative jumps
M	integer number s.t. N is the number of grid points
N	number of grid points ($N = 2^M$)
n	number of monitoring time in path dependent derivatives
n_v	number of points in which $V(x)$ is not differentiable
N_sim	Number of MC simulation
$N(*)$	standard normal cumulative distribution function evaluated in $*$
$p^B_t\left(\mathcal{I}_t(y), y\right))$	Black put option price
$p_t(S_t, t)$	quantity inside the ATS put expected value
p^-_t	upper bound of $\phi_{s,t}$ strip of regularity
p^+_t	$-(p^+_t + 1)$ is the lower bound of $\phi_{s,t}$ strip of regularity
$P(x)$	model CDF of the increment between the times s and t
$\hat{P}(x)$	numerical approximation of the CDF of the increment between the times s and t
$P_t\left(x\right)$	put option price at value date with maturity t and moneyness x
$P_t\left(K\right)$	put option price at value date with maturity t and strike K
$\mathcal{P}_{S_t}$	probability density function of S_t
$\bar{s}$	market implied cost-of-funding
S_t	random variable s.t. tS_t has Laplace transform $\mathcal{L}_t$
T	option time-to-maturity
V	derivative payoff
W_t	Brownian motion
x	option moneyness
$(x_0, x_\mathcal{K})$	interval in which the CDF is interpolated
$\{Z_t\}_{t\geq 0}$	Lévy subordinator
α	*Index of stability*: tempered stable parameter of ATS, $\alpha \in [0, 1)]$
β	scaling parameter of k_θ
γ_t	drift term of additive process $\{X_t\}_{t\geq 0}$
γ	length of CDF interpolation grid
$\Gamma(*)$	Gamma function evaluated in $*$
δ	scaling parameter of η_θ
ϵ	jump simulation small jump truncation parameter
φ	deterministic drift term of LTS
φ_t	deterministic drift term of ATS
ϕ^c	characteristic function of the forward exponent
$\phi_{s,t}$	characteristic function of ATS increment between time s and t
η	*skew* parameter of LTS

Symbol	Description
η_t	*skew* parameter of ATS
$\hat{\eta}_\theta$	re-scaled *skew* parameter of ATS
$\bar{\eta}$	constant part of the ATS *skew* parameter
ν_t	Lévy measure of the additive process $\{X_t\}_{t\geq 0}$
$\hat{\xi}_t$	implied volatility skew term
$\hat{\xi}_0$	short-time skew term, i.e. limit for t that goes to zero of $\hat{\xi}_t$
σ	diffusion parameter of LTS
σ_t	diffusion parameter of ATS
$\hat{\sigma}_t$	ATM implied volatility, equal to $\mathcal{I}_t(0)$
$\hat{\sigma}_0$	short-time ATM implied volatility, i.e. limit for t that goes to zero of $\hat{\sigma}_t$
$\hat{\sigma}_\theta$	re-scaled diffusion parameter of ATS, equal to one
$\bar{\sigma}$	constant diffusion parameter of ATS
θ	re-scaled maturity, defined as $\sigma_T^2\, T$

Shorthands

Symbol	Description
a.s.	almost surely
ATS	additive normal tempered stable process
ATM	at-the-money
bp	basis points
CDF	cumulative distribution function
EFFR	effective federal fund rate
EONIA	Euro overnight indexed average rate
FFT	fast Fourier transform
GA	Gaussian approximation technique
ITM	in-the-money
LTS	Lévy normal tempered stable process
MSE	mean squared errors
MAPE	mean absolute percentage error
MAX	maximum error
MC	Monte Carlo
ms	milliseconds
NIG	Normal inverse Gaussian process
OIS	overnight indexed swap
OTM	out-of-the-money
OR	overnight rate
r.v.	random variable
RMSE	root mean squared errors
SD	average MC prices standard deviation
SOFR	secured overnight financing rate
STR	Euro short term rate
TSS	additive tempered stable subordinator
VG	Variance Gamma process
w.r.t.	with respect to

Appendices

.1 Proofs of Chapter 1

We start proving a technical **Lemma** that we use in the proof of **Theorem 1.2.1**.

Lemma .1.1.
If $\lim_{t\to 0} t\,\sigma_t^2\,\eta_t = 0$, *then*

$$\lim_{t\to 0} \int_0^\infty ds \frac{t}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{k_t}\right)^{1-\alpha} \left(\frac{e^{-(1-\alpha)\ s/k_t}}{s^{1+\alpha}}\right) \int_{|x|<1} dx \frac{x}{\sqrt{2\pi s}\sigma_t} e^{-\left(\frac{x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} = 0 \ .$$

Proof.

$$\begin{aligned}
&\left| \int_{|x|<1} dx \frac{x}{\sqrt{2\pi s}\sigma_t} e^{-\left(\frac{x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} \right| \\
\leq &\left| \int_{|x|<1} dx \frac{x}{\sqrt{2\pi s}\sigma_t} e^{-\left(\frac{x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} + \int_1^\infty dx \frac{x}{\sqrt{2\pi s}\sigma_t} \left(e^{-\left(\frac{x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} - e^{-\left(\frac{-x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} \right) \right| \\
= &\sigma_t^2 s \left| \frac{1}{2} + \eta_t \right| \ .
\end{aligned}$$

The inequality is due to the fact that both terms inside the right-hand side absolute value are positive if $(1/2+\eta_t)$ is positive and are negative if $(1/2+\eta_t)$ is negative. Now it is possible to write the following bound

$$\begin{aligned}
&\left| \int_0^\infty ds \frac{t}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{k_t}\right)^{1-\alpha} \left(\frac{e^{-(1-\alpha)\ s/k_t}}{s^{1+\alpha}}\right) \int_{|x|<1} dx \frac{x}{\sqrt{2\pi s}\sigma_t} e^{-\left(\frac{x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} \right| \\
&\leq \sigma_t^2 \left| \frac{1}{2} + \eta_t \right| \int_0^\infty ds \frac{t}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{k_t}\right)^{1-\alpha} \left(\frac{e^{-(1-\alpha)\ s/k_t}}{s^{\alpha}}\right) = t\sigma_t^2 \left| \frac{1}{2} + \eta_t \right| \ ,
\end{aligned}$$

where the last equality is due to the definition of $\Gamma(1-\alpha)$. We prove the thesis by the squeeze theorem □

Proof of Theorem 1.2.1
The idea of this proof is to show that there exists an additive process with the characteristic function in (1.2.4) using the result in Sato (1999a, Th.9.8, p.52).
At any given time $t > 0$ the characteristic function in (1.2.4) is the characteristic function of a LTS (1.2.2), at time t, with parameters $k = k_t$, $\eta = \eta_t$, $\sigma = \sigma_t$ and $\varphi = \varphi_t$. Hence, we have an expression for the generating triplet of (1.2.4) (see, e.g., Cont and Tankov 2003, eq. 4.24, p.130)

$$\begin{cases}
A_t &= 0 \\
\gamma_t &= \int_0^\infty ds \frac{t}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{k_t}\right)^{1-\alpha} \left(\frac{e^{-(1-\alpha)\ s/k_t}}{s^{1+\alpha}}\right) \int_{|x|<1} dx \frac{x}{\sqrt{2\pi s}\sigma_t} e^{-\left(\frac{x+s\sigma_t^2(\eta_t+1/2)}{\sqrt{s}\sigma_t}\right)^2} + t\varphi_t \\
\nu_t(x) &= \frac{tC\,(\alpha, k_t, \sigma_t, \eta_t)}{|x|^{1/2+\alpha}} e^{-(\eta_t+1/2)x} K_{\alpha+1/2}\left(|x| \sqrt{(1/2+\eta_t)^2 + 2(1-\alpha)/(k_t\,\sigma_t^2)} \right)
\end{cases} ,$$

with

$$C\left(\alpha, k_t, \sigma_t, \eta_t\right) := \frac{2}{\Gamma(1-\alpha)\sqrt{2\pi}} \left(\frac{1-\alpha}{k_t}\right)^{1-\alpha} \sigma_t^{2\alpha} \left(\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(k_t\,\sigma_t^2)\right)^{\alpha/2+1/4} ,$$

and

$$K_\nu(z) := \frac{e^{-z}}{\Gamma\left(\nu+\frac{1}{2}\right)} \sqrt{\frac{\pi}{2z}} \int_0^\infty e^{-s} s^{\nu-1/2} \left(\frac{s}{2z}+1\right)^{\nu-1/2} ds$$

is the modified Bessel function of the second kind (see, e.g., Abramowitz and Stegun 1948, Ch.9 p.376). For $t = 0$, as usual in additive processes, we set $\gamma_0 = 0$, $A_0 = 0$ and $\nu_0 = 0$.

First, we verify that $\nu_t(x)$ is a non decreasing function of t. It is possible to identify two expressions in the jump measure $\nu_t(x)$

$$e^{-x(1/2+\eta_t)-|x|\sqrt{(1/2+\eta_t)^2+2(1-\alpha)/(\sigma_t^2 k_t)}} \tag{.1.1}$$

$$\frac{t^{1/\alpha}\sigma_t^2}{k_t^{(1-\alpha)/\alpha}} \left(\frac{s}{|x|} + \sqrt{\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)}\right) \ . \tag{.1.2}$$

We point out that expression (.1.2) is inside the integral and depends on the integration variable $s \geq 0$. If these two expressions, (.1.1) and (.1.2), are non decreasing w.r.t. t for any t, x and $s \geq 0$ then the jump measure is non decreasing. Expression .1.1 is non decreasing because g_1 and g_2 are non decreasing by hypothesis 1. Hypothesis 1 on g_1 and g_2 also implied that the squared root in (.1.2) is non increasing for any t and then, because condition 1 on g_3 holds, the prefactor $\frac{t^{1/\alpha}\sigma_t^2}{k_t^{(1-\alpha)/\alpha}}$ is non decreasing (even multiplied by $s/|x|$). Thus, (.1.2) is non decreasing for any t, x and $s \geq 0$ because it is the sum of a non decreasing function and g_3, a non decreasing function by hypothesis. This proves that $\nu_t(x)$ is non decreasing in t.

Second, we prove that $\lim_{t\to 0} \nu_t(x) = 0$ for $x \neq 0$; this is equivalent to demonstrate that (.1.1) or (.1.2) go to zero as t goes to zero. We show that this happens in all possible cases. We first consider the case where

$$\lim_{t\to 0} k_t > 0 \quad \text{and} \quad \lim_{t\to 0} |(1/2+\eta_t)| < \infty \ . \tag{.1.3}$$

In this case is evident that expression (.1.2) goes to zero for small t. Otherwise, when (.1.3) is not true, we have to distinguish two further cases depending on whether

$$\lim_{t\to 0} (1/2+\eta_t)\, k_t \sigma_t^2 = 0 \tag{.1.4}$$

holds. If (.1.4) is true expression (.1.1) goes to zero, otherwise, because of condition 2 on $t\,\eta_t^\alpha\,\sigma_t^{2\alpha}/k_t^{1-\alpha}$, expression (.1.2) goes to zero. This proves that $\lim_{t\to 0} \nu_t(x) = 0$ for any $x \neq 0$.

We can now check whether the triplet satisfies the conditions in Sato (1999a, Th.9.8, p.52).

1. The triplet has no diffusion term.
2. ν_t is not decreasing in t.

3. The continuity of $\nu_t(B)$ and γ_t , where $B \in \mathbb{B}(\mathbb{R}^+)$ and $B \subset \{x : |x| > \epsilon > 0\}$, is obvious for $t > 0$: it is a natural consequence of the composition of continuous functions. For $t = 0$ we have to prove that the limits of $\nu_t(B)$ and γ_t are 0. We have already proven that $\nu_t(x)$ is non decreasing in t and that $\lim_{t\to 0} \nu_t(x) = 0$, $\forall x \neq 0$. The convergence of $\nu_t(B)$ to 0 is due to the dominated convergence theorem. The convergence of γ_t is because of **Lemma .1.1** and because $t\varphi_t$ goes to zero by definition of ATS. □

Proof of Proposition 1.2.2
A forward contract, valued in t with delivery in T, is $F_t(T) = F_0(T)\ e^{f_t}$, also for an ATS, as in (1.2.1) for the LTS.

Let us prove the sufficient condition. If the forward is martingale

$$\mathbb{E}\left[F_t(T)|\mathcal{F}_0\right] = F_0(T) \quad .$$

This is equivalent to impose that

$$\mathbb{E}\left[e^{f_t}\middle|\mathcal{F}_0\right] = 1\ , \tag{.1.5}$$

that is, the characteristic function of f_t computed in $-i$ is equal to one. From equation (1.2.4)

$$\mathbb{E}[e^{f_t}|\mathcal{F}_0] = \mathcal{L}_t\left(\left(\eta_t + \frac{1}{2}\right)\sigma_t^2 - \frac{\sigma_t^2}{2};\ k_t,\ \alpha\right)e^{\varphi_t t} = \mathcal{L}_t\left(\sigma_t^2\eta_t;\ k_t,\ \alpha\right)e^{\varphi_t t} \quad . \tag{.1.6}$$

Imposing the condition (.1.5), we get φ_t.

Let us prove the necessary condition in two steps. First, given φ_t by equation (1.2.5) we prove that $\mathbb{E}[e^{f_t}|\mathcal{F}_0] = 1, \forall t \geq 0$. This fact is a consequence of equation (.1.6).
Second, we check the martingale condition; that is, $\forall s, t$ s.t $0 \leq s \leq t$

$$\mathbb{E}\left[F_t(T)|\mathcal{F}_s\right] = F_0(T)\,\mathbb{E}\left[e^{f_t - f_s + f_s}\middle|\mathcal{F}_s\right] = e^{f_s}F_0(T)\,\mathbb{E}\left[e^{f_t - f_s}\right] = F_s(T)\mathbb{E}\left[e^{f_t - f_s}\right] \quad .$$

The proposition is proven once we prove that $\mathbb{E}\left[e^{f_t - f_s}\right] = 1$.
This equality holds because f_t is additive; that is, process increments are independent

$$\mathbb{E}\left[e^{f_t}|\mathcal{F}_0\right] = \mathbb{E}\left[e^{f_t - f_s}|\mathcal{F}_0\right]\ \mathbb{E}\left[e^{f_s}|\mathcal{F}_0\right]\ ,$$

then

$$\mathbb{E}\left[e^{f_t - f_s}\right] = \mathbb{E}\left[e^{f_t - f_s}|\mathcal{F}_0\right] = \frac{\mathbb{E}\left[e^{f_t}|\mathcal{F}_0\right]}{\mathbb{E}\left[e^{f_s}|\mathcal{F}_0\right]} = 1$$ □

Proof of Theorem 1.2.3
We check that the power-law sub-case of ATS satisfies the two conditions of **Theorem 1.2.1**. First, we verify that $g_1(t)$, $g_2(t)$, and $g_3(t)$ are non decreasing.

1. $$g_1(t) = (1/2 + \bar{\eta}t^\delta) - \sqrt{(1/2 + \bar{\eta}t^\delta)^2 + 2\bar{\sigma}^2(1-\alpha)/(\bar{k}t^\beta)}$$

   is non decreasing because its derivative w.r.t. t is always greater or equal than zero for any $t \geq 0$.

$$\frac{d}{dt}\left((1/2 + \bar{\eta}t^\delta) - \sqrt{(1/2 + \bar{\eta}t^\delta)^2 + \frac{2(1-\alpha)t^{-\beta}}{\bar{\sigma}^2\bar{k}}}\right) \geq 0$$

$$\frac{1-\alpha}{2\bar{\sigma}^2\bar{k}}\left(\frac{\beta t^{-\delta-\beta/2}}{\bar{\eta}\delta}\right)^2-\frac{\beta t^{-\delta}}{2\bar{\eta}\delta}-\frac{\beta}{\delta}\geq 1\ .$$

The last inequality is verified for any t if and only if $\beta \geq -\delta$. The inequality holds due to the hypotheses $\delta \leq 0$ and $\beta > -\delta$.

2. $$g_2(t) = -(1/2+\bar{\eta}t^{\delta}) - \sqrt{(1/2+\bar{\eta}t^{\delta})^2 + 2\bar{\sigma}^2(1-\alpha)/(\bar{k}t^{\beta})}$$
is non decreasing for any $t \geq 0$: it is the sum of two non decreasing functions because of the conditions $\beta \geq 0$ and $\delta \leq 0$.

3. $$g_3(t) = \frac{\sqrt{\bar{\sigma}^4 t^{2/\alpha-2\beta(1-\alpha)/\alpha}\,(1/2+\bar{\eta}t^{\delta})^2 + 2t^{-\beta+2/\alpha-2\beta(1-\alpha)/\alpha}\bar{\sigma}^2(1-\alpha)/(\bar{k})}}{\bar{k}^{(1-\alpha)/\alpha}}$$
is non decreasing for any $t \geq 0$: it is the sum of three non decreasing functions of t (positive powers) elevated to a positive power because of the conditions $\beta \leq \frac{1}{1-\alpha/2}$ and $\delta > \frac{\beta(1-\alpha)-1}{\alpha}$.

Second, we verify that $t\,\sigma_t^2\,\eta_t$ and $t\,\sigma_t^{2\,\alpha}\,\eta_t^{\alpha}/k_t^{1-\alpha}$ go to zero. The expressions $t^{1+\delta}\,\bar{\sigma}^2\,\bar{\eta}$ and $t^{1+\delta\alpha-\beta(1-\alpha)}\,\bar{\sigma}^{2\alpha}\,\bar{\eta}^{\alpha}/\bar{k}^{1-\alpha}$ go to zero as t goes to zero because of the conditions $\delta > -\min\left(\beta, \frac{1-\beta(1-\alpha)}{\alpha}\right)$ and $\beta \leq \frac{1}{1-\alpha/2}$ □

Proof of Proposition 1.2.4
We prove the thesis using the definition of additive process (Cont and Tankov 2003, Def.14.1 p.455).

1. By hypothesis $r_0 = 0$ and by definition of additive process $X_0 = 0$ almost surely. Thus, $X_{r_0} = 0$ almost surely.
2. Independence of increments follows from the monotonicity of r_t.
3. Stochastic continuity w.r.t. time follows from stochastic continuity of the additive process and continuity of the function r_t □

.2 ATS Parameter estimation

In physics and engineering, all measurements are subject to some uncertainties or "errors". Error analysis is a vital part of any quantitative study (see, e.g., Taylor 1997). In this appendix, we estimate pricing errors and "propagate" them to model parameters. This is a crucial passage to verify the quality of the proposed model.

First, we estimate pricing errors. In finance, the idea of considering the bid-ask spread in market prices as a sort of measurement error of "true" prices is well known and goes back to the seminal paper of Roll (1984). He considers the price $y = y^* + q(y_{ask} - y_{bid})/2$, where y is the observed price, y^* the unobserved true price, and q a binomial r.v. that takes value in $\{-1, 1\}$ with equal probability, where -1 corresponds to the bid price and $+1$ to the ask price. Modeling the uncertainty with such a distribution, the relation between bid-ask spread and price standard deviation Σ_y is $\Sigma_y = (y_{ask} - y_{bid})/2$. More recently, George *et al.* (1991) propose an extended formulation of the price $y = y^* + \pi q(y_{ask} - y_{bid})/2$, where π is the unobserved proportion of the spread due to the so-called order processing cost; π is estimated from market data as a value 0.8 and in all cases analyzed in George *et al.* (1991) is observed a value greater than 0.5. Conservatively, π can be chosen as 0.5, obtaining the relation $\Sigma_y = (y_{ask} - y_{bid})/4$.
Another possibility, in the plain vanilla option market for equity indices that we consider in this study, is to model the true price y as a Gaussian random variable with a mean equal to the mid-market price $(y_{ask} + y_{bid})/2$ and bid and ask prices chosen as symmetric quantiles. This represents more closely what is observed in this derivative market. On the one hand, it is standard for a market player to pass through an options broker to work the order. Generally, real trades are closer to the mid-market than to bid/ask prices (see, e.g., Petersen and Fialkowski 1994). On the other hand, it is not sure that a market player trades within the bid-ask spread. In some rare cases, a trade can take place at a price higher (lower) than the ask (bid) price: it can happen because the bid-ask enlarges due to sudden movements in the underlying or in presence of a very large trade, such as the hedging of a large exposure. It is rather difficult to estimate how rare these events are. They can happen roughly around the 5% of the cases (i.e. $y_{ask} - y_{bid} \simeq 2 \times 1.96\ \Sigma_y$).
For this reason, in this thesis, we consider the measurement error in prices as Gaussian and related to the bid-ask spread via $\Sigma_y = (y_{ask} - y_{bid})/4$. With this choice, the relation between prices standard deviation and the bid-ask spread is equal to the one obtained by George *et al.* (1991).

Second, we "propagate" to model parameters this measurement error in prices. In applied statistics, the propagation of uncertainties is a standard technique (see, e.g., Taylor 1997, Ryan 2008). We briefly recall some main results present in the literature for the models (.2.1), (.2.3) and (.2.5) considered; then we describe the calibration procedure adopted in the thesis.
Consider the linear model

$$y = Zg + \epsilon \ , \tag{.2.1}$$

where $y \in \mathbb{R}^n$ is the response vector, $Z \in \mathbb{R}^{n\times(r+1)}$ is the explanatory variables matrix, $\epsilon \sim N_n(0, \Sigma)$, $\Sigma \in \mathbb{R}^{n\times n}$ is the diagonal response vector variance-covariance matrix, $g \in \mathbb{R}^{r+1}$ is the unobserved coefficient vector. We indicate with $N_n(\mu, \Sigma)$ an n-dimensional Gaussian distribution with mean μ and variance Σ. We perform a weighted linear regression with weights $W \in \mathbb{R}^{n\times n}$, a diagonal matrix. The least square solution is

$$\hat{g} = (Z'WZ)^{-1} Z'WY \ ,$$

where $Y \in \mathbb{R}^n$ is the observed response vector (see, e.g., Ryan 2008, Ch.3, pp.115-116). Thus, $\hat{g}$ is the Gaussian linear combination of Gaussian random variables:

$$\hat{g} \sim N_{r+1}\left(g, \left(Z'WZ\right)^{-1} Z'W\Sigma WZ'\left(Z'WZ\right)^{-1}\right) \quad . \tag{.2.2}$$

In the weighted non-linear regression case, it is possible to obtain a similar result (see, e.g., Seber and Wild 1989, Ch.2, pp.21-24). Consider the model

$$y_i = f(g, z_i) + \epsilon_i \tag{.2.3}$$

where y_i is the i^{th} component of the response vector $y \in \mathbb{R}^n$, ϵ_i is the i^{th} component of the error vector $\epsilon \sim N_n\left(0, \Sigma\right)$, z_i is the i^{th} row of the explanatory variables' matrix. Similarly, the coefficients of a non-linear regression are:

$$\hat{g} \sim N_{r+1}\left(g, \left(F'WF\right)^{-1} F'W\Sigma W'F\left(F'WF\right)^{-1}\right) \quad , \tag{.2.4}$$

where $F \in \mathbb{R}^{n\times(r+1)}$ is s.t. its (i,j) element is

$$F^{i,j} = \frac{\partial f}{\partial g_j}|_{g,z_i}$$

and g_j is the j^{th} component of g.

In the literature, the case that takes into account Gaussian correlated errors on both the response vector and the explanatory variables is available for the fitting of a straight line (see, e.g., York 1968). Consider the model

$$y_i = a + b(z_i + \epsilon_{z_i}) + \epsilon_{y_i} \quad , \tag{.2.5}$$

with y_i and z_i subjected to Gaussian errors with variance Σ_{z_i} and Σ_{y_i} and covariance Σ_{z_i,y_i}. The estimated slope and intercept $\hat{a}$ and $\hat{b}$ can be obtained through a fast iterative procedure. In the first order approximation

$$\begin{aligned} \hat{a} &\sim N\left(a, \Sigma_a\right) \\ \hat{b} &\sim N\left(b, \Sigma_b\right) \quad , \end{aligned} \tag{.2.6}$$

where the expressions of Σ_a and Σ_b are reported in York (1968, 1^{st} equation in p.324).

In this thesis, the calibration procedure is divided into two steps.
First, for a given maturity T, we deal with the non-linear problem and we calibrate from market data the three time-dependent parameters k_T, σ_T and η_T on options with different strikes. The distribution of the estimated parameters can be obtained using equation (.2.4). We construct Σ through all observed bid and ask prices at the given maturity: the diagonal value is equal to $\left(y_{ask} - y_{bid}\right)^2/16$. The matrix of weights W, as standard in the option market, is chosen as the identity matrix because the bid-ask spread does not differ significantly in the market prices in the calibration dataset. Consequently, the calibration results of different models can be easily compared as shown in section 1.3, where we compare ATS with LTS and Sato models. As result of this step, we obtain a variance-covariance matrix $\Sigma_T \in \mathbb{R}^{3\times3}$ of the estimated parameters $(k_T, \sigma_T^2, \eta_T)$ for every maturity T.
Then, to estimate the scaling parameters of model (1.3.2), we rewrite the parameters definition w.r.t. $\theta := T\sigma_T^2$ in log-log scale as

$$\ln \hat{k}_\theta = \ln \bar{k} + \beta \ln \theta$$

$$\ln \hat{\eta}_\theta = \ln \bar{\eta} + \delta \ln \theta \ .$$

The estimated variance and covariance of $\ln \hat{k}_\theta$, $\ln \hat{\eta}_\theta$ and $\ln \theta$ are obtained by a first-order expansion

$$\begin{cases} Var\left(\ln \hat{k}_\theta\right) &= \dfrac{\Sigma_T^{1,1}}{k_T^2} + \dfrac{\Sigma_T^{2,2}}{\sigma_T^4} + 2\dfrac{\Sigma_T^{2,1}}{k_T \sigma_T^2} \\ Var\left(\ln \hat{\eta}_\theta\right) &= \dfrac{\Sigma_T^{3,3}}{\eta_T^2} \\ Var\left(\ln \theta\right) &= \dfrac{\Sigma_T^{2,2}}{\sigma_T^4} \end{cases} \quad \begin{cases} Cov\left(\ln \hat{k}_\theta, \ln \theta\right) &= \dfrac{\Sigma_T^{2,2}}{\sigma_T^4} + \dfrac{\Sigma_T^{2,1}}{k_T \sigma_T^2} \\ Cov\left(\ln \hat{\eta}_\theta, \ln \theta\right) &= \dfrac{\Sigma_T^{2,3}}{\eta_T \sigma_T^2} \end{cases},$$

where T is the maturity corresponding to the θ of interest. The distributions of the estimated parameters β, δ, $\bar{k}$, and $\bar{\eta}$ are the one identified in equation (.2.6). The weights selected in the minimization procedure (see York 1968, equation (1), p.320) are $1/Var(\ln \hat{k}_\theta)$ in the regression on $\ln \hat{k}_\theta$ and $1/Var(\ln \hat{\eta}_\theta)$ in the regression on $\ln \hat{\eta}_\theta$. The weights of the explanatory variable $\ln \theta$ are $1/Var(\ln \theta)$.

Finally, from the confidence intervals for k_T and η_T we can also get the confidence intervals of the skewness and the excess kurtosis at a given maturity. We are able to obtain skewness and excess kurtosis of ATS thanks to the identity in law with LTS (for the moments of LTS see, e.g., Cont and Tankov 2003, p.129). The linear regression of these two higher moments, w.r.t. the squared root of time, is realized by computing the Gaussian errors (.2.2), in the first-order approximations, of skewness and excess kurtosis.

.3 Additive subordination

In this appendix we give an equivalent definition of additive subordinator (see Li *et al.* 2016). We prove that it is possible to obtain an ATS process via additive subordination if and only if η_t is constant. The additive process is a generalization of the Lévy process. For an accurate description, from a probabilistic point of view, see Sato (1999a). Li *et al.* (2016) introduce the theoretical foundation of additive subordination. They show that it is possible to subordinate a Lévy process with an additive subordinator (a non-decreasing additive process) obtaining a new additive process.
The ATS have excellent performances in applications to equity market modelling (cf. chapters 1 and 4). This class of processes has been introduced through its characteristic function in equation (1.2.4):

$$E\left[e^{iuf_t}\right] = \mathcal{L}_t\left(iu\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 + \frac{u^2\sigma_t^2}{2};\ k_t, \alpha\right) e^{iu\varphi_t t}\ ,$$

where σ_t, k_t are continuous on $[0,\infty)$ and η_t, φ_t are continuous on $(0,\infty)$ with $\sigma_t > 0$, $k_t \geq 0$ and $\varphi_t\, t$ goes to zero as t goes to zero. $\ln \mathcal{L}_t$ is defined as

$$\ln \mathcal{L}_t\left(u;\ k,\ \alpha\right) := \begin{cases} \dfrac{t}{k}\dfrac{1-\alpha}{\alpha}\left\{1-\left(1+\dfrac{u\,k}{1-\alpha}\right)^{\alpha}\right\} & \text{if } \ 0<\alpha<1 \\ -\dfrac{t}{k}\ln\left(1+u\,k\right) & \text{if } \ \alpha = 0 \end{cases},$$

with $\alpha \in [0,1)$. There are strong evidences that, in the equity market, k_t and η_t are non-constant, cf. chapters 1, and 2.
Lévy normal tempered stable process (LTS) is usually built through Lévy subordination (see e.g., Cont and Tankov 2003, Ch.4). Two notorious examples of the Lévy normal tempered stable family are the Normal Inverse Gaussian (NIG) (Barndorff-Nielsen 1997) and the Variance Gamma (VG) (Madan *et al.* 1998), which are obtained via two different Lévy subordinators. The Lévy normal tempered stable processes f_t are usually introduced as

$$f_t = -\left(\eta+\frac{1}{2}\right)\sigma^2\ Z_t + \sigma\ W_{Z_t} + \varphi t \qquad \forall t \in [0,T]\ , \tag{.3.1}$$

where η, σ are two real parameters ($\eta \in \mathbb{R}, \sigma \in \mathbb{R}^+$), while the φ is obtained by imposing the martingale condition on the forward price. W_t is a Brownian motion and Z_t is a Lévy tempered stable subordinator independent from the Brownian motion, such as an Inverse Gaussian process for NIG or a Gamma process for VG.

We would desire to build the LTS extension also via a subordinator (cf. equation (.3.1) for the LTS); i.e.

$$f_t = -\left(\eta_t+\frac{1}{2}\right)\ \sigma_t^2\ Z_t + \ \sigma_t W_{Z_t} + \varphi_t t \tag{.3.2}$$

where $\sigma_t^2 Z_t$ is an additive subordinator independent from the Brownian motion path.

In this appendix (i)we give an equivalent definition of additive subordinator (that links the properties of the subordinators to its characteristic function) (ii)we define the additive tempered stable subordinator (TSS), a natural extension of the Lévy tempered stable subordinator (iii) we prove that it is possible to define an ATS via additive subordination, as in (2.2.3) if and only if η_t is constant.

In this section we give the definition of additive process and additive subordinator, we define the TSS and we prove the main result of the appendix in **Theorem .3.8**.

We consider again the standard definition of additive process.

Definition .3.1. Additive process *(see, e.g., Cont and Tankov 2003, Def.14.1 p.455).*
A càdlàg stochastic process on $\mathbb{R}$ $\{X_t\}_{t\geq 0}$, $X_0 = 0$ *a.s. is an additive process if and only if it satisfies the following conditions:*

1. *Independent increments: for every positive real increasing sequence* $t_0, ..., t_n$ *the random variables* $X_{t_1} - X_{t_0}, ..., X_{t_n} - X_{t_n-1}$ *are independent;*
2. *Stochastic continuity:*
$$\forall \epsilon > 0, \lim_{h\to 0} P\left[|X_{t+h} - X_t| > \epsilon\right] = 0$$ □

(A_t, ν_t, γ_t) are the generating triplets that characterize the additive process $\{X_t\}_{t\geq 0}$. A_t , ν_t and γ_t are known respectively as the diffusion term, the Lévy measure and the drift term (see Sato 1999a, pp.38-39).

We recall the important result of Sato (1999a) that links the existence of an additive process with appropriate conditions on its generating triplets. We use these results to prove the existence of the TSS.

Theorem .3.2. Main additive property.
Let $\{g_t\}_{t\geq 0}$ *be a system of infinitely divisible probability measures on* $\mathbb{R}$ *with generating triplets* (A_t, ν_t, γ_t) *satisfying the following conditions. Then, there exists, uniquely up to identity in law, an additive process* $\{X_t\}_{t\geq 0}$ *on* $\mathbb{R}$ *s.t.* X_t *has law* g_t *for* $t \geq 0$.

1. $A_0 = 0$, $\nu_0 = 0$, $\gamma_0 = 0$;
2. *Given* t, s *s.t.* $0 \leq s \leq t$ *then* $A_s \leq A_t$ *and* $\nu_s(B) \leq \nu_t(B)$, $B \in \mathbb{B}(\mathbb{R})$;
3. *Given* t, s *s.t.* $0 \leq s \leq t$ *then as* $s \to t$, $A_s \to A_t$, $\nu_s(B) \to \nu_t(B)$ *and* $\gamma_s \to \gamma_t$, *where* $B \in \mathbb{B}(\mathbb{R})$ *and* $B \subset \{x : |x| > \epsilon > 0\}$.

Conversely, suppose that $\{X_t\}_{t\geq 0}$ *is an additive process in law on* $\mathbb{R}$. *Let* (A_t, ν_t, γ_t) *be the generating triplet of its infinitely divisible distribution* $\{g_t\}_{t\geq 0}$. *Then, conditions 1, 2 and 3 are satisfied.*

Proof. See Sato (1999a), Th.9.8 p.52. □

We introduce the additive subordinator as is standard in the literature (see e.g., Li *et al.* 2016, p.3). In **Proposition .3.4** we present an equivalent definition of an additive subordinator that links its properties to its generating triplet. This definition is the extension for additive process of the equivalent definition for Lévy subordinator (see, e.g., Cont and Tankov 2003, Prop.3.10 p. 100)

Definition .3.3. Additive subordinator.
An additive subordinator is an a.s. positive and a.s. non decreasing additive process. □

As is done in chapter 1 (see e.g., Cont and Tankov 2003), we consider ν_t a Lévy measure defined on $\mathbb{R}$ that can be identified with the Lévy density $\nu_t(x)$ s.t. $\int_B \nu_t(x)dx = \nu_t(B)$ $\forall B \in \mathbb{B}(\mathbb{R})$ and $B \subset \{x : |x| > \epsilon > 0\}$.

Proposition .3.4. Equivalent characterization of additive subordinator.
An additive process $\{Z_t\}_{t\geq 0}$ *is an additive subordinator if and only it is infinitely divisible distribution is s.t., for every fixed time* t, $A_t = 0$, $b_t := \gamma_t - \int_{0\leq x\leq 1} x\,\nu_t(dx) \geq 0$ *non decreasing and* ν_t *s.t. i)* $\int_{\mathbb{R}} (|x| \wedge 1)\,\nu_t(dx) < \infty$, *ii)* $\nu_t((-\infty, 0]) = 0$. *Moreover* $\{Z_t\}_{t\geq 0}$ *has characteristic function*

$$\ln E\left[e^{iuZ_t}\right] = ib_t u + \int_{x>0} \left(e^{iux} - 1\right)\nu_t(x)dx \ . \tag{.3.3}$$

Proof. This proof reminds the ones in Cont and Tankov (2003) for Lévy subordination (Cor.3.1 and Prop.3.10, pp.84-85). The only if part is proven by Li *et al.* (2016, proof of Prop. 2.1).
We prove the if part. We show that given the conditions on the characteristic triplet equation (.3.3) holds.
Define $L(x) := \mathbb{1}_{|x|\leq 1} x\nu_t(x)$ and $M(x) := \left(e^{iux} - 1\right)\nu_t(x)$. We have that

$$\ln E\left[e^{iuZ_t}\right] = i\gamma_t u + \int_{\mathbb{R}} \left(e^{iux} - 1 - \mathbb{1}_{|x|\leq 1} iux\right)\nu_t(x)\,dx = i\gamma_t u + \int_{\mathbb{R}} \left(iuL(x) + M(x)\right)dx \ .$$

The first equality is due to the definition of the characteristic function of an additive process with no diffusion. $L(x)$ is integrable w.r.t. x thanks to the conditions on ν_t. The sum of $iuL(x)$ and $M(x)$ is integrable, because $E\left[e^{iuZ_t}\right]$ is a well defined characteristic function, thus $M(x)$ is integrable too. We can split the integral and check the thesis defining $b_t := \gamma_t - \int_{0\leq x\leq 1} x\nu_t(dx)$. This proves equation (.3.3).
Given r, t s.t. $0 \leq r < t$ we prove that $Z_t - Z_r$ is an a.s. non negative random variable. By definition of additive process we can easily compute the characteristic function of $Z_t - Z_r$

$$E\left[e^{iuZ_r)}\right] E\left[e^{iu(Z_t - Z_r)}\right] = E\left[e^{iuZ_t}\right]$$

or equivalently

$$E\left[e^{iu(Z_t - Z_r)}\right] = E\left[e^{iuZ_t}\right] / E\left[e^{iuZ_r}\right]$$

and using equation (.3.3) we obtain an explicit formula for its exponent

$$\ln E\left[e^{iu(Z_t - Z_r)}\right] = i\bar{b}u + \int_{x>0} \left(e^{iux} - 1\right)\bar{\nu}(x)dx \ ,$$

where $\bar{b} := b_t - b_r$ and $\bar{\nu}(x) := \nu_t(x) - \nu_r(x)$. We observe that, by definition of b_t, $\bar{b}$ is non negative and that $\bar{\nu}$ is a jump measure with no negative jumps ($\nu_t(x) - \nu_r(x)$ is a non negative function with value on $[0, \infty)$ thanks to **Theorem .3.2**). The random variable $Z_t - Z_r$ is a.s. non negative because it has the same characteristic function of a Lévy subordinator with drift $\bar{b}$ and jump measure $\bar{\nu}$ at time $t = 1$ (cf. Cont and Tankov 2003, p.84); thus, Z_t is non decreasing in time. A non decreasing additive process satisfies the conditions of **Definition .3.3**. □

We prove that it is possible to define a new additive process via additive subordination; a general version of this result is given by Li *et al.* (2016, Th. 3.1).

Theorem .3.5. Additive subordination. *Given* $\{X_t\}_{t\geq 0}$ *a Lévy process and* $\{Z_t\}_{t\geq 0}$ *an additive subordinator, then* $\{X_{Z_t}\}_{t\geq 0}$ *is an additive process*

Proof. We prove the thesis verifying the three conditions of an additive process in **Definition .3.1**. The proof extends the one in Cont and Tankov (2003, Th.4.2, p.120) on Lévy subordination.

1. Condition 1 holds by **Definition .3.1**. For the processes $\{Z_t\}_{t\geq 0}$ and $\{X_t\}_{t\geq 0}$, $Z_0 = X_0 = 0$ almost surely. Thus, $X_{Z_0} = 0$ almost surely.

2. We prove the independence of increments. Let $\mathcal{F}_Z$ be the sigma-algebra generated by the process $\{Z_t\}_{t\geq 0}$; for any increasing time sequence $t_0, t_1, \ldots, t_N$, let us write the characteristic function of the vector of increments:

$$E\left[e^{i\sum_{i=1}^{N} u_j\left(X_{S_{t(j)}} - X_{S_{t(j-1)}}\right)}\right] = E\left[E\left[\prod_{j=1}^{N} e^{iu_j\left(X_{S_{t(j)}} - X_{S_{t(j-1)}}\right)}\middle|\mathcal{F}_Z\right]\right]$$

$$= E\left[\prod_{j=1}^{N} E\left[e^{iu_j\left(X_{S_{t(j)}} - X_{S_{t(j-1)}}\right)}\middle|\mathcal{F}_s\right]\right] = E\left[\prod_{j=1}^{N} e^{\left(S_{t(j)} - S_{t(j-1)}\right)\psi(u_j)}\right] \tag{.3.4}$$

$$= \prod_{j=1}^{N} E\left[e^{\left(S_{t(j)} - S_{t(j-1)}\right)\psi(u_j)}\right] = \prod_{j=1}^{N} E\left[e^{iu_j\left(X_{S_{t(j)}} - X_{S_{t(j-1)}}\right)}\right] \ , \tag{.3.5}$$

   where equality (.3.4) is due to the independence of $\{X_t\}_{t\geq 0}$ increments and to the characteristic function of the Lévy process; equality (.3.5) to the independence of $\{Z_t\}_{t\geq 0}$ increments.

3. Stochastic continuity w.r.t. time follows from stochastic continuity of the two processes.

□

We define the TSS through its generating triplets and we prove its existence. The selected characteristic function is the one of a Lévy tempered stable subordinator (see e.g., Cont and Tankov 2003, p.127) but with a time-dependent parameter k_t and multiplied by σ_t^2.

Definition .3.6. The process $\sigma_t^2 Z_t$.
The TSS is an additive subordinator $\left\{\sigma_t^2 Z_t\right\}_{t\geq 0}$ characterized by the triplets $(0, \mathcal{V}_t, \Gamma_t)$

$$\begin{cases} \mathcal{V}_t(x) & := \dfrac{t\sigma_t^{2\alpha}}{\Gamma(1-\alpha)}\left(\dfrac{1-\alpha}{k_t}\right)^{1-\alpha}\left(\dfrac{e^{-(1-\alpha)\ x/(\sigma_t^2 k_t)}}{x^{1+\alpha}}\right)\mathbb{1}_{x>0} \\ \Gamma_t & := \displaystyle\int_0^1 x\ \mathcal{V}_t(x)\ dx \ , \end{cases} \tag{.3.6}$$

where $\alpha \in [0,1)$, $t \in \mathbb{R}^+$. σ_t and k_t are positive continuous functions of time s.t.

1. $t\ \sigma_t^2$ *is $o(1)$ for small t;*
2. $\dfrac{t}{k_t^{1-\alpha}}\ \sigma_t^{2\alpha}$ *is $o(1)$ for small t and non-decreasing;*
3. $\sigma_t^2\ k_t$ *is non-decreasing;*

Theorem .3.7. $\sigma_t^2\ Z_t$ is an additive subordinator.
The additive tempered stable subordinator exists and has $b_t = 0$.

Proof. We check whether the conditions of **Proposition .3.4** on the generating triplet of an additive subordinator are satisfied by $\{\sigma_t Z_t\}_{t\geq 0}$. Let us observe that there is no diffusion term and

$$\int_0^\infty (|x| \wedge 1)\, \mathcal{V}_t(x)\, dx \leq \int_0^\infty x \mathcal{V}_t(x)\, dx$$
$$= t\sigma_t^{2\alpha} \int_0^\infty \frac{1}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{k_t}\right)^{1-\alpha} \frac{e^{-(1-\alpha)\frac{x}{\sigma_t^2 k_t}}}{x^\alpha} dx = t\sigma_t^2 \quad , \quad (.3.7)$$

where the last equality is due to the definition of $\Gamma(1-\alpha)$. Moreover $\mathcal{V}_t(-\infty, 0) = 0$ and b_t is null by direct substitution of Γ_t in the formula of **Proposition .3.4**.

We show that $\left\{\sigma_t^2 Z_t\right\}_{t\geq 0}$ is an additive process using **Theorem .3.2**; that is, we check whether the triplet introduced in **Definition .3.6** satisfies the theorem conditions.

1. The triplet has no diffusion term.
2. $\mathcal{V}_t$ is not decreasing in t because $\sigma_t^{2\alpha} t / k_t^{1-\alpha}$ and $\sigma_t k_t$ are non-decreasing functions of t (see **Definition .3.6**).
3. For $t > 0$ the continuity of $\mathcal{V}_t(B)$, where $B \in \mathbb{B}(\mathbb{R}^+)$ and $B \subset \{x : |x| > \epsilon > 0\}$, is due to the composition of continuous functions. For $t = 0$ we can extend $\mathcal{V}_t(B)$ and Γ_t to 0 since both converge to 0 as $t \to 0$. The convergence of Γ_t to 0 is due to Γ_t positiveness and to the condition $\Gamma_t \leq t$ (see equation (.3.7)). The convergence of $\mathcal{V}_t(B)$ to 0 is due to the dominated convergence theorem. We observe that, $\forall x \in \mathbb{R}^+$ s.t. $|x| > \epsilon > 0$, $\mathcal{V}_t(x)$ is finite and a decreasing function in t.

An additive process that satisfies the conditions on the triplet of **Proposition .3.4** is a subordinator. □

Theorem .3.8. Main result.
It is possible to construct via additive subordination an additive process with characteristic function (1.2.4) if and only if η_t is constant.

Proof. Consider the process

$$f_t = W\left(\sigma_t^2 Z_t\right) - \left(\frac{1}{2} + \eta_t\right) \sigma_t^2 \quad ,$$

where $W(t)$ is a standard Brownian motion. We can compute f_t characteristic function conditioning w.r.t. $\sigma_t^2 Z_t$

$$E\left[e^{iuf_t}\right] = E\left[e^{-u^2\sigma_t^2 Z_t - iu(\eta_t + 1/2)\sigma_t^2 Z_t}\right] e^{iu\varphi_t t} = \mathcal{L}\left(k_t, t, \alpha, iu\left(\frac{1}{2} + \eta_t\right)\sigma_t^2 + \frac{u^2\sigma_t^2}{2}\right) e^{iu\varphi_t t}$$

Notice that the characteristic function matches the expression of the ATS characteristic function (1.2.4). If η_t is constant the process $\{f_t\}_{t\geq 0}$ is additive by **Theorem .3.5** . It is obtained subordinating the Lévy process $W(t) - (1/2 + \bar{\eta})t$ with the additive subordinator $\sigma_t^2 Z_t$. If η_t is non constant we select t, r s.t. $t > r$ and $\eta_t \neq \eta_r$ and show that $f_t - f_r$ is not independent from f_r.

$$\frac{E\left[e^{iu_1(f_t - f_r) + u_2 f_r}\right]}{e^{iu_1(t\phi_t - r\phi_r) + iu_2 r\phi_r}} = E\left[e^{-\left(u_1^2/2 + iu_1/2\right)\left(\sigma_t^2 Z_t - \sigma_r^2 Z_r\right) - iu_1(\eta_t\sigma_t Z_t - \eta_r\sigma_r Z_r) - (u_2^2/2 + iu_2/2)\sigma_r^2 Z_r - iu_2\eta_r\sigma_r^2 Z_r}\right]$$
$$= E\left[e^{-\left(u_1^2/2 + iu_1/2 + iu_1\eta_t\right)\left(\sigma_t^2 Z_t - \sigma_r^2 Z_r\right)}\right] E\left[e^{-(u_2^2/2 + iu_2/2)\sigma_r^2 Z_r - iu_2\eta_r\sigma_r^2 Z_r - iu_1\sigma_r^2 Z_r(\eta_t - \eta_r)}\right]$$

$$
\begin{aligned}
&= \frac{\mathcal{L}\left(k_t, t, \alpha, iu_1\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 + \frac{u_1^2\sigma_t^2}{2}\right)}{\mathcal{L}\left(k_r, r, \alpha, iu_1\left(\frac{1}{2}+\eta_t\right)\sigma_r^2 + \frac{u_1^2\sigma_r^2}{2}\right)}\mathcal{L}\left(k_r, r, \alpha, iu_2\left(\frac{1}{2}+\eta_r\right)\sigma_r^2 + \frac{u_2^2\sigma_r^2}{2} - iu_1\sigma_r^2(\eta_t-\eta_r)\right) \\
&\neq \frac{\mathcal{L}\left(k_t, t, \alpha, iu_1\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 + \frac{u_1^2\sigma_t^2}{2}\right)}{\mathcal{L}\left(k_r, r, \alpha, iu_1\left(\frac{1}{2}+\eta_r\right)\sigma_r^2 + \frac{u^2\sigma_t^2}{2}\right)}\mathcal{L}\left(k_r, r, \alpha, iu_2\left(\frac{1}{2}+\eta_r\right)\sigma_r^2 + \frac{u_2^2\sigma_r^2}{2}\right) \\
&= \frac{E\left[e^{iu_1(f_t-f_r)}\right]E\left[e^{u_2 f_r}\right]}{e^{iu_1(t\phi_t - r\phi_r)+iu_2 r\phi_r}} \quad ,
\end{aligned}
$$

where the first inequality is obtained conditioning w.r.t. $\sigma_t Z_t$ and $\sigma_r Z_r$, the second one to the fact that $\{\sigma_t^2 Z_t\}_{t\geq 0}$ is an additive process and the last two to the property of the characteristic function of an additive process increments. This proves the thesis.

□

.4 Basic properties of ATS

We report some useful results for the proofs in section 2.3. In the following lemmas we consider S_t of **Definition 2.2.1** with Laplace transform $\mathcal{L}_t(u;\ k_t;\ \alpha)$, at a given time $t > 0$. The proofs that follow are for the $\alpha \in (0,1)$ case. Similar proofs hold in the $\alpha = 0$ case.

Lemma .4.1. *Let $s \in (0,1)$, then*

$$\mathbb{E}\left[S_t^s\right] = \int_0^\infty \frac{\mathcal{L}_t(u;\ k_t;\ \alpha) - 1}{\Gamma(-s)u^{s+1}} du \ , \tag{.4.1}$$

where Γ is the Gamma function.

Proof. By elementary calculus and Fubini's Theorem (see, e.g., Urbanik 1993, Lemma 4, p.325) □

Lemma .4.2. *Let n be a positive integer, then*

$$\mathbb{E}[S_t^{-n}] = \Gamma(n)^{-1} \int_0^\infty u^{n-1} \mathcal{L}_t(u;\ k_t;\ \alpha) du \ .$$

Proof. By elementary calculus and Fubini's Theorem (see, e.g., Cressie *et al.* 1981, Ch.2, p.148) □

Lemma .4.3.

1. *For all $t > 0$, $c \geq 1$ and $u \geq 0$*

$$1 - \mathcal{L}_t(u;\ k_t,\ \alpha) \leq 1 - e^{-cu} \ .$$

2. *If $\beta \geq 1$, $\mathcal{L}_t(u;\ k_t,\ \alpha)$ is non decreasing in t.*

Proof. Let us observe that

$$1 - \mathcal{L}_t(u;\ k_t,\ \alpha) \leq 1 - e^{-cu}$$
$$\frac{t}{k_t}\frac{1-\alpha}{\alpha}\left\{\left(1 + \frac{u\ k_t}{(1-\alpha)t}\right)^\alpha - 1\right\} - cu \leq 0 \ .$$

The last inequality is true for any $c \geq 1$ and $u \geq 0$ because the left hand side is null in $u = 0$ and its first order derivative w.r.t. u is negative:

$$\frac{1}{\left(1 + \frac{k_t u}{t(1-\alpha)}\right)^{1-\alpha}} - c < 0 \ .$$

This proves the first point.

We demonstrate that the logarithm of $\mathcal{L}_t(u;\ k_t,\ \alpha)$ is not decreasing. Consider a positive t, $s \in (0,t)$ and

$$h(u;\ s,\ t) := \frac{t}{k_t}\left\{1 - \left(1 + \frac{u\ k_t}{(1-\alpha)t}\right)^\alpha\right\} - \frac{s}{k_r}\left\{1 - \left(1 + \frac{u\ k_r}{(1-\alpha)s}\right)^\alpha\right\} \ .$$

We observe that $h(0;\ s,\ t) = 0$ and the first order derivative

$$\frac{\partial h(u;\ s,\ t)}{\partial u} = \frac{1}{\left(1 + \frac{k_r u}{s(1-\alpha)}\right)^{1-\alpha}} - \frac{1}{\left(1 + \frac{k_t u}{t(1-\alpha)}\right)^{1-\alpha}}$$

is non negative $\forall u > 0$ because k_t/t is non decreasing in t, if $\beta > 1$, and is constant in t, if $\beta = 1$. Thus, $h(u;\ s,\ t) \geq 0$, $\forall u \geq 0$, and $\mathcal{L}_t(u;\ k_t,\ \alpha)$ is non decreasing w.r.t. t. This proves point 2 □

Lemma .4.4.

1. *If $\beta < 1$ S_t goes to zero in distribution as t goes to zero.*
2. *If $\beta > 1$ S_t goes to one in distribution as t goes to zero.*
3. *If $\beta = 1$ the distribution of S_t does not depend from t.*

Proof. Recall that convergence in the Laplace transform implies convergence in distribution (see, e.g., Kijima 1997, Th.B.9, p.308).
We compute the limit of S_t Laplace transform for $\beta < 1$. By using the fact that k_t/t goes to infinity as t goes to zero we obtain

$$\lim_{t\to 0} \mathcal{L}_t(u;\ k_t,\ \alpha) = \lim_{t\to 0} e^{\frac{t}{k_t}\frac{1-\alpha}{\alpha}\left\{1-\left(1+\frac{u\ k_t}{(1-\alpha)t}\right)^{\alpha}\right\}} = 1 \ .$$

Thus, S_t converges in distribution to the constant zero. This proves point 1.

We compute the limit of S_t Laplace transform for $\beta > 1$. By using the fact that k_t/t goes to zero as t goes to zero we obtain

$$\lim_{t\to 0} \mathcal{L}_t(u;\ k_t,\ \alpha) = \lim_{t\to 0} e^{\frac{t}{k_t}\frac{1-\alpha}{\alpha}\left\{1-\left(1+\frac{u\ k_t}{(1-\alpha)t}\right)^{\alpha}\right\}} = e^{-u} \ .$$

Thus, S_t converges in distribution to the constant one. This proves point 2.

Point 3 follows from the fact that, if $\beta = 1$, $\mathcal{L}_t(u;\ k_t,\ \alpha)$ is constant in t

□

Lemma .4.5.

$$\lim_{t\to 0} \mathbb{E}[\sqrt{S_t}] = \begin{cases} 0 & \text{if } \beta < 1 \\ 1 & \text{if } \beta > 1 \\ D & \text{if } \beta = 1 \end{cases} , \tag{.4.2}$$

where D is a positive constant.

Proof. Recall that S_t is a positive r.v. and $\mathbb{E}[S_t] = 1$. Then, its moment of order $1/2$ is finite. By **Lemma .4.1**

$$\mathbb{E}\left[\sqrt{S_t}\right] = \int_0^\infty \frac{\mathcal{L}_t(u;\ k_t,\ \alpha) - 1}{\Gamma(-1/2)u^{3/2}} du \ ,$$

where $\frac{-1}{\Gamma(-1/2)} \approx 3.45$. By **Lemma .4.3** point 1 with $c = 2$, the positive quantity $(1 - \mathcal{L}_t(u;\ k_t,\ \alpha))/u^{3/2}$ is lower or equal than $(1 - e^{-2u})/u^{3/2}$. Thus,

$$0 \leq \mathbb{E}\left[\sqrt{S_t}\right] \leq \frac{-1}{\Gamma(-1/2)} \int_0^\infty \frac{1 - e^{-2u}}{u^{3/2}} du = \frac{-4}{\Gamma(-1/2)} \int_0^\infty \frac{e^{-2u}}{u^{1/2}} du = \sqrt{2} \ , \tag{.4.3}$$

where the first equality is obtained from integration by parts and the second from the definition of Γ. Inequality (.4.3) has two consequences. First, if $\beta = 1$,

$$\lim_{t\to 0} \mathbb{E}[\sqrt{S_t}] = \mathbb{E}[\sqrt{S_t}] := D \leq \sqrt{2} \ , \tag{.4.4}$$

because, by **Lemma .4.4** point 3, $\mathbb{E}[\sqrt{S_t}]$ is constant w.r.t. to time. Second, we can apply the dominated convergence theorem to (.4.2) for all values of β. Recall that the limits for t that goes to zero of $\mathcal{L}_t(u;\ k_t,\ \alpha)$ for $\beta < 1$ and for $\beta > 1$ are computed in the proof of **Lemma .4.4**.
If $\beta < 1$

$$\lim_{t\to 0} \mathbb{E}[\sqrt{S_t}] = \lim_{t\to 0} \frac{-1}{\Gamma(-1/2)} \int_0^\infty \frac{1 - \mathcal{L}_t(u;\ k_t,\ \alpha)}{u^{3/2}} du = 0 \ . \tag{.4.5}$$

If $\beta > 1$

$$\begin{aligned} \lim_{t\to 0} \mathbb{E}[\sqrt{S_t}] &= \lim_{t\to 0} \frac{-1}{\Gamma(-1/2)} \int_0^\infty \frac{1 - \mathcal{L}_t(u;\ k_t,\ \alpha)}{\Gamma(-1/2) u^{3/2}} du \\ &= \frac{-1}{\Gamma(-1/2)} \int_0^\infty \frac{1 - e^{-u}}{u^{3/2}} du = \frac{-2}{\Gamma(-1/2)} \int_0^\infty \frac{e^{-u}}{u^{1/2}} du = 1 \ , \end{aligned} \tag{.4.6}$$

where the third equality is obtained from integration by parts and the third by the definition of Γ. Equalities (.4.4), (.4.5) and (.4.6) prove the thesis

□

Lemma .4.6.
*Consider φ_t in (2.2.2). For every β and δ in the additive process boundaries of **Theorem 1.2.3***

1. $$\varphi_t t = t\bar{\sigma}^2 \eta_t - t\bar{\sigma}^4 \eta_t^2 k_t/2 + O\left(t\eta_t^3 k_t^2\right) \ , \tag{.4.7}$$
   where the second term $t\bar{\sigma}^4\eta_t^2 k_t/2$ goes to zero faster than $t\bar{\sigma}^2\eta_t$ as t goes to zero.

2. $$\frac{\varphi_t}{\bar{\sigma}^2 \eta_t} \leq 1 \ .$$

3. $$\lim_{t\to 0} \frac{\varphi_t}{\bar{\sigma}^2 \eta_t} = 1 \ , \ \textit{for} \ \ \delta > -\min(1, \beta) \ .$$

Proof. We prove the asymptotic expansion (.4.7). In the additive process boundaries of **Theorem 1.2.3** at least either $\beta = \delta = 0$ or $\delta > -\min(1, \beta)$. In the former case (.4.7) is trivial. In the latter, thanks to (2.2.2), both $t\eta_t = t^{1+\delta}\bar{\eta}$ and $\eta_t k_t = t^{\beta+\delta}\bar{\eta}\bar{k}$ go to zero as t goes to zero. Using the Taylor series expansion

$$\begin{aligned} \varphi_t t &= \frac{t(1-\alpha)}{k_t} \left\{ \frac{\bar{\sigma}^2\, \eta_t\, k_t}{1-\alpha} - \frac{\bar{\sigma}^4\, \eta_t^2\, k_t^2}{2(1-\alpha)} + O\left(\eta_t^3\, k_t^3\right) \right\} \\ &= t\, \bar{\sigma}^2\, \eta_t - t\, \bar{\sigma}^4\, \eta_t^2\, k_t/2 + O\left(t\, \eta_t^3\, \kappa_t^2\right) \ . \end{aligned}$$

This proves point 1.
We prove that $\varphi_t/(\bar{\sigma}^2\eta_t) \leq 1$. We substitute the definition of φ_t in (2.2.2), for $\alpha > 0$, in (.4.7) and we get

$$\varphi_t/(\bar{\sigma}^2\eta_t) = \frac{(1-\alpha)}{\alpha\bar{\sigma}^2\eta_t k_t} \left(\left(1 + \frac{\bar{\sigma}^2\eta_t k_t}{1-\alpha} \right)^\alpha - 1 \right) \leq 1 \ . \tag{.4.8}$$

We define $z := \frac{\bar{\sigma}^2 \eta_t k_t}{1-\alpha}$. Then, (.4.8) is equivalent to

$$(1+z)^\alpha \leq 1 + \alpha z \ ,$$

which is a well known inequality. This proves point 2.

Point 3 is straightforward, given point 1, because, if $\delta > -\min(1, \beta)$, $\eta_t k_t$ goes to zero as t goes to zero □

.5 Short-time limits

Lemma .5.1. *Consider a family of positive random variables X_t s.t. $\lim_{t\to 0} X_t = X$ in distribution and a sequence of functions $g_t(z) \geq 0$ and uniformly bounded s.t. $\lim_{t\to 0} g_t(z) = g(z)$.*
If $\exists\, \tau > 0$ s.t. for $t \in (0, \tau)$

[i)]

1. *$g_t(z)$ is Lipschitz continuous with bounded Lipschitz constant,*
2. *$|g_t(z) - g(z)| < h(z)$ with $\lim_{z\to\infty} h(z) = 0$,*

then

$$\lim_{t\to 0} \mathbb{E}[g_t(X_t)] = \mathbb{E}[g(X)] \ .$$

Proof. It is possible to apply the Ascoli-Arzelá theorem (see, e.g.,, Rudin 1976, Th.7.25, p.158) on every compact set $[0, K]$, $K > 0$, because a sequence of Lipschitz continuous functions with bounded Lipschitz constant is equicontinous on any compact set. Thus, a sub-sequence of $g_t(z)$ converges uniformly to $g(z)$ in any $[0, K]$. For every $\epsilon > 0$, $\exists\, K$ s.t.

$$\lim_{t\to 0} \mathbb{E}[|g_t(X_t) - g(X_t)|] = \lim_{t\to 0} \mathbb{E}\left[|g_t(X_t) - g(X_t)| \mathbb{1}_{X_t < K}\right] + \lim_{t\to 0} \mathbb{E}\left[|g_t(X_t) - g(X_t)| \mathbb{1}_{X_t > K}\right] < \epsilon \ .$$

The first expected value goes to zero because $g_t(z)$ converges uniformly to $g(z)$ on $[0, K]$, as proven above via Ascoli-Arzelá theorem. There exists K s.t. it is possible to bound the second with ϵ because $h(z)$ goes to zero as z goes to infinity.
Moreover, $g(z)$ is bounded because it is the limit of a uniformly bounded sequence and

$$\lim_{t\to 0} \mathbb{E}\left[|g(X_t) - g(X)|\right] = 0 \ .$$

by definition of convergence in distribution, because $g(z)$ is bounded. We have that

$$0 \leq \lim_{t\to 0} \mathbb{E}[|g_t(X_t) - g(X)|] \leq \lim_{t\to 0} \left\{\mathbb{E}[|g_t(X_t) - g(X_t)|] + \mathbb{E}\left[|g(X_t) - g(X)|\right]\right\} = 0 \ ,$$

this proves the thesis

□

Lemma .5.2. *For $\delta = -1/2$, let X_t be a sequence of positive random variable s.t. $X_t \to X$ in distribution for t that goes to zero.*
Then,

$$\lim_{t\to 0} \mathbb{E}\left[N\left(\bar{\sigma}\bar{\eta}\left(-\sqrt{X_t} + \varphi_t/(\bar{\sigma}^2\sqrt{X_t}\eta_t)\right) - \bar{\sigma}\sqrt{tX_t}/2\right)\right] = \mathbb{E}\left[N(\bar{\sigma}\bar{\eta}(-\sqrt{X} + 1/\sqrt{X}))\right] \ .$$

Proof. Define

$$g_t(z) := N\left(\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\sqrt{z}\eta_t)\right)-\bar\sigma\sqrt{t\,z}/2\right) \text{ and } g(z) := N(\bar\sigma\bar\eta(-\sqrt{z}+1/\sqrt{z})) \ .$$

We emphasize that $g_t(z)$ is uniformly bounded by one and $g_t(z)$ converges point-wise to $g(z)$ because, thanks to **Lemma .4.6** point 3, $\lim_{t\to 0}\varphi_t/(\bar\sigma^2\eta_t)=1$.
We prove that $\exists\,\tau\in(0,1)$ s.t. the derivative of $g_t(z)$ is uniformly bounded, if $t\in(0,\tau)$. Fix $\tau\in(0,1)$ s.t.

$$\begin{cases} \dfrac{\varphi_\tau}{\bar\sigma^2\,\eta_\tau} & > \frac{2}{3} \\ \dfrac{\bar\sigma\bar\eta}{\bar\sigma\bar\eta+\bar\sigma\sqrt{\tau}/2} & > \frac{3}{4} \end{cases} \ .$$

The following hold for $t<\tau$,

$$\begin{aligned}
&\left|\frac{\partial g_t}{\partial z}\right| \\
&= N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{zt}/2\right)\left|\bar\sigma\bar\eta\left(-1/(2\sqrt{z})-\varphi_t/(2\bar\sigma^2\eta_t\,z^{3/2})\right)-\bar\sigma\sqrt{t}/(4\sqrt{z})\right| \\
&= N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{zt}/2\right)\left(1+\varphi_t/(\bar\sigma^2\eta_t z)+\sqrt{t}/(2\bar\eta)\right)\bar\sigma\bar\eta/\left(2\sqrt{z}\right) \\
&\le N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{zt}/2\right)(1+1/z+1/(2\bar\eta))\,\bar\sigma\bar\eta/\left(2\sqrt{z}\right) \qquad (.5.1) \\
&\le \left[\frac{1}{\sqrt{2\pi}}\mathbb{1}_{D_2}+N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+2/(3\sqrt{z})\right)-\tau\bar\sigma\sqrt{z}/2\right)\mathbb{1}_{D_1}+N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+1/\sqrt{z}\right)\right)\mathbb{1}_{D_3}\right]\cdot \\
&\quad (1+1/z+1/(2\bar\eta))\,\bar\sigma\bar\eta/\left(2\sqrt{z}\right) := M(z) \ . \qquad (.5.2)
\end{aligned}$$

Inequality (.5.1) holds because, by **Lemma .4.6** point 2, $\varphi_t/(\bar\sigma^2\eta_t)<1$ and $\tau\in(0,1)$. Let us observe that (.5.1) is the product of positive quantities. In (.5.2) we bound from above only the first factor, the only one that still depends from t. Inequality (.5.2) is deduced by dividing the domain of $z\in\mathbb{R}^+$ in the three sets $D_1\equiv(0,1/2]$, $D_2\equiv(1/2,3/2]$ and $D_3\equiv(3/2,\infty)$.
For $z\in D_2$, we bound the first factor with its maximum $\frac{1}{\sqrt{2\pi}}$.
For $z\in D_1$, we observe that for $t<\tau$

$$\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{zt}/2>\bar\sigma\bar\eta\left(-\sqrt{z}+2/(3\sqrt{z})\right)-\tau\bar\sigma\sqrt{z}/2>0 \ .$$

Hence, because N' is a decreasing function of its argument in $\mathbb{R}^+$,

$$N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{zt}/2\right)\le N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+2/(3\sqrt{z})\right)-\tau\bar\sigma\sqrt{z}/2\right) \ , \ z\in D_1 \ .$$

Finally, for $z\in D_3$

$$\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{tz}/2<\bar\sigma\bar\eta\left(-\sqrt{z}+1/\sqrt{z}\right)<0 \ . \qquad (.5.3)$$

Thus, because N' is an increasing function of its argument in $\mathbb{R}^-$

$$N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+\varphi_t/(\bar\sigma^2\eta_t\sqrt{z})\right)-\bar\sigma\sqrt{tz}/2\right)<N'\left(\bar\sigma\bar\eta\left(-\sqrt{z}+1/\sqrt{z}\right)\right) \ , \ z\in D_3 \ .$$

Notice that $M(z)$ is positive and bounded on $\mathbb{R}^+$; this implies that the derivatives of $g_t(z)$ is uniformly bounded. Thus, the sequence $g_t(z)$ is Lipschitz continuous in z with bounded Lipschitz

constant on $(0,\tau)$.
Moreover, for $t<\tau<1$ we have that

$$\begin{aligned}|g_t(z)-g(z)| &\leq \mathbb{1}_{z\in(0,1]}+N'\left(\bar{\sigma}\bar{\eta}\left(-\sqrt{z}+1/\sqrt{z}\right)\right)\left(\bar{\sigma}\sqrt{zt}/2+\bar{\sigma}\bar{\eta}(1-\varphi_t/(\bar{\sigma}^2\eta_t))/\sqrt{z}\right)\mathbb{1}_{z\in(1,\infty)}\\ &\leq \mathbb{1}_{z\in(0,1]}+N'\left(\bar{\sigma}\bar{\eta}\left(-\sqrt{z}+1/\sqrt{z}\right)\right)\left(\bar{\sigma}\sqrt{z}/2+\bar{\sigma}\bar{\eta}/\sqrt{z}\right)\mathbb{1}_{z\in(1,\infty)}:=h(z)\ .\end{aligned}$$

In the first inequality we divide the domain of $z\in\mathbb{R}^+$ in two sets, $D_1\equiv(0,1]$ and $D_2\equiv(1,\infty)$. In the first domain the difference is bounded by one. In the second set, notice that (.5.3) is still valid for $z>1$; then, the difference is lower than N' computed on the max of the arguments of N multiplied by the positive difference of the arguments of N. The second inequality holds because $\varphi_t/(\bar{\sigma}^2\eta_t)$ is positive and $t<1$. We observe that $h(z)$ goes to zero as z goes to infinity. Notice that X_t converges to X in distribution, $g_t(z)$ is a sequence of positive function uniformly bounded, Lipschitz continuous with bounded Lipschitz constant on $(0,\tau)$, and $\lim_{z\to\infty}h(z)=0$. Thus, we prove the thesis via **Lemma .5.1** □

Lemma .5.3. *For $t>0$,*

$$\sup_z\left|\mathbb{P}(S_t<z)-N\left((z-1)\sqrt{\frac{t}{k_t}}\right)\right|\leq\frac{2-\alpha}{1-\alpha}\sqrt{\frac{k_t}{t}}\ , \tag{.5.4}$$

*where S_t is the random variable of **Definition 2.2.1** with Laplace transform $\mathcal{L}_t(u;\,k_t,\,\alpha)$. Moreover, if $\beta>1$,*

1. $$\lim_{t\to0}\mathbb{P}(S_t<1)=\lim_{t\to0}\mathbb{P}(S_t\geq1)=\lim_{t\to0}\mathbb{P}\left(S_t\leq\frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right)=\frac{1}{2}\ .$$

2. $$\lim_{t\to0}\mathbb{P}(S_t\leq1-t^q)=\begin{cases}1/2 & \text{if } q>\frac{\beta-1}{2}\\ N(-1/\sqrt{k}) & \text{if } q=\frac{\beta-1}{2}\\ 0 & \text{if } q<\frac{\beta-1}{2}\end{cases}\ .$$

3. $$\lim_{t\to0}\mathbb{P}(S_t\leq1+t^q)=\begin{cases}1/2 & \text{if } q>\frac{\beta-1}{2}\\ N(1/\sqrt{k}) & \text{if } q=\frac{\beta-1}{2}\\ 1 & \text{if } q<\frac{\beta-1}{2}\end{cases}\ .$$

Proof. We use an approach similar to Küchler and Tappe (2013, Th.4.7, p.4271). Given $t>0$, $n\in\mathbb{N}$ we define $X_t^i:=S_t^i-1$ for $i=1,2,...,n$ with S_t^i independent positive random variables with Laplace transform $\mathcal{L}_t(u;\,k_t\,n,\,\alpha)$. The standard deviation of S_t^i is $\Sigma_t^n:=\sqrt{k_tn/t}$. We define
$Q_t^n:=\sum_{i=1}^n X_t^i/(\sqrt{n}\Sigma_t^n)$. Notice that $Q_t^n+\sqrt{n}/\Sigma_t^n$ has the same law of $\sqrt{t/k_t}S_t$ by identity in Laplace transform because

$$\mathbb{E}\left[e^{-u\left(P_t^n+\sqrt{n}/\Sigma_t^n\right)}\right]=\mathbb{E}\left[e^{-u\sum_{i=1}^n S_t^i/(\sqrt{n}\Sigma_t^n)}\right]=e^{\sum_{i=1}^n\frac{t(1-\alpha)}{k_t\alpha n}\left(1-\left(1+\frac{u\sqrt{k_t}}{\sqrt{t}(1-\alpha)}\right)^\alpha\right)}=\mathbb{E}\left[e^{-u\sqrt{\frac{t}{k_t}}S_t}\right]\ .$$

Thus, $\sqrt{k_t/t}\,Q_t^n+1$ is equal in distribution to S_t. Moreover, for any $t>0$

$$\sup_z|\mathbb{P}(Q_t^n<z)-N(z)|\leq\frac{\mathbb{E}\left[\left|X_t^i\right|^3\right]}{(\Sigma_t^n)^3\sqrt{n}}<\frac{\mathbb{E}\left[(S_t^i)^3\right]+1}{(\Sigma_t^n)^3\sqrt{n}}=t^{3/2}\frac{2+3\frac{k_tn}{t}+\frac{(2-\alpha)k_t^2n^2}{(1-\alpha)t^2}}{k_t^{3/2}n^2}=\frac{2-\alpha}{1-\alpha}\sqrt{\frac{k_t}{t}}+O\left(\frac{1}{n}\right)\ .$$

The first inequality holds thanks to the Berry-Esseen theorem (see, e.g., Durrett 2019, Th.3.4.17, p.136). The first equality is obtained by substituting the third moment of S_t^i and in the last equality we emphasize the leading term in $1/n$. Thus, $\forall\, \epsilon > 0$ it exists n such that

$$\sup_z |\mathbb{P}(Q_t^n < z) - N(z)| < \frac{2-\alpha}{1-\alpha}\sqrt{\frac{k_t}{t}} + \epsilon \ .$$

By definition of cumulative distribution function, we get (.5.4).
Equation (.5.4) allow us to prove the limits of the probability.

1. $$\mathbb{P}(S_t \leq 1) = N(0) + O\left(t^{(\beta-1)/2}\right) = 1/2 + O\left(t^{(\beta-1)/2}\right) \ ,$$
   where the first equality is due to (.5.4) and second term goes to zero because $\beta > 1$. Moreover,
   $$P\left(S_t \leq \frac{\varphi_t}{\bar{\sigma}^2\eta_t}\right) = N\left(\frac{\varphi_t - \bar{\sigma}^2\eta_t}{\bar{\sigma}^2\eta_t}\sqrt{\frac{t}{k_t}}\right) + O\left(t^{(\beta-1)/2}\right) \ ,$$
   where $\lim_{t\to 0} N\left(\frac{\varphi_t - \bar{\sigma}^2\eta_t}{\bar{\sigma}^2\eta_t}\sqrt{\frac{t}{k_t}}\right) = \frac{1}{2}$ thanks to **Lemma .4.6** point 3 observing that
   $$\left(\frac{\varphi_t}{\bar{\sigma}^2\eta_t} - 1\right)\sqrt{\frac{t}{k_t}} = \sqrt{t}\bar{\sigma}^2\eta_t\sqrt{k_t} + O\left(t^{2\delta+(3\beta+1)/2}\right) = o(1) \ ,$$
   because, in the additive process boundaries, for $\beta > 1$, $\delta + (\beta+1)/2 > \delta + 1 > 0$.

2. $$\mathbb{P}(S_t \leq 1 - t^q) = N\left(-t^q\sqrt{\frac{t}{k_t}}\right) + O\left(t^{(\beta-1)/2}\right) \ ,$$
   where $N\left(-t^q\sqrt{\frac{t}{k_t}}\right)$ goes to $1/2$ if $q > (\beta-1)/2$, to $N(-1/(\sqrt{k})$ if $q = (\beta-1)/2$, and to 0 if $q < (\beta-1)/2$. We emphasize that the second term goes to zero as $O\left(t^{(\beta-1)/2}\right)$.

3. $$\mathbb{P}(S_t \leq 1 + t^q) = N\left(t^q\sqrt{\frac{t}{k_t}}\right) + O\left(t^{(\beta-1)/2}\right) \ ,$$
   where $N\left(t^q\sqrt{\frac{t}{k_t}}\right)$ goes to $1/2$ if $q > (\beta-1)/2$, to $N(1/(\sqrt{k})$ if $q = (\beta-1)/2$, and to 1 if $q < (\beta-1)/2$. We emphasize that the second term goes to zero as $O\left(t^{(\beta-1)/2}\right)$.

□

Lemma .5.4. *If $\delta = -1/2$, $\exists H > 1$ s.t.*

$$m(z) := N'\left(-l_t^z + \frac{\bar{\sigma}\sqrt{zt}}{2}\right)\bar{\sigma}\sqrt{z} \quad , \quad z > 0 \ ,$$

is increasing for $z \in [1, H]$ for sufficiently small t, where l_t^z is the quantity defined in equation (2.2.6).

Proof. We compute the derivative w.r.t. z of $m(z)$ and study its sign at short-time.

$$\begin{aligned}\frac{\partial m(z)}{\partial z} &= \frac{\frac{\partial N'\left(-l_t^z+\frac{\bar{\sigma}\sqrt{zt}}{2}\right)\bar{\sigma}\sqrt{z}}{\partial z}}{N'\left(\bar{\sigma}\bar{\eta}\left(\sqrt{z}-\frac{\varphi_t}{\bar{\sigma}^2\eta_t\sqrt{z}}\right)+\frac{\bar{\sigma}\sqrt{tz}}{2}\right)\bar{\sigma}} \\ &= \frac{1}{2\sqrt{z}} - 2\left(\bar{\sigma}\bar{\eta}\left(\sqrt{z}-\frac{\varphi_t}{\bar{\sigma}^2\sqrt{z}\eta_t}\right)+\frac{\bar{\sigma}\sqrt{tz}}{2}\right)\left(\bar{\sigma}\bar{\eta}\left(\frac{1}{2}+\frac{\varphi_t}{2\bar{\sigma}^2\eta_t z}\right)+\frac{\bar{\sigma}\sqrt{t}}{4}\right) \\ &= \frac{1}{z^{3/2}}\left(\frac{z}{2}-z^2\left(\bar{\sigma}\bar{\eta}+\bar{\sigma}\frac{\sqrt{t}}{2}\right)^2+\bar{\sigma}^2\bar{\eta}^2\frac{\varphi_t^2}{\bar{\sigma}^4\eta_t^2}\right) \ .\end{aligned}$$

The derivative is positive if

$$0 < z < \frac{1/2+\sqrt{1/4+4\left(\bar{\sigma}\bar{\eta}+\bar{\sigma}\frac{\sqrt{t}}{2}\right)^2\bar{\sigma}^2\bar{\eta}^2\frac{\varphi_t}{\bar{\sigma}^4\eta_t^2}}}{2\left(\bar{\sigma}\bar{\eta}+\bar{\sigma}\frac{\sqrt{t}}{2}\right)^2} \ .$$

Notice that $\exists\ \tau$ and $H > 1$ such that for every $t < \tau$ the derivative is positive if $z < H$ because for sufficiently small time

$$\frac{1/2+\sqrt{1/4+4\left(\bar{\sigma}\bar{\eta}+\bar{\sigma}\frac{\sqrt{t}}{2}\right)^2\bar{\sigma}^2\bar{\eta}^2\frac{\varphi_t^2}{\bar{\sigma}^4\eta_t^2}}}{2\left(\bar{\sigma}\bar{\eta}+\bar{\sigma}\frac{\sqrt{t}}{2}\right)^2} > \frac{1/2}{2(\bar{\sigma}\bar{\eta}+\bar{\sigma}\sqrt{t}/2)}+\frac{\bar{\eta}}{\bar{\eta}+\sqrt{t}/2}\,\frac{\varphi_t}{\bar{\sigma}^2\eta_t} > H > 1 \ ,$$

where the first inequality is obtained by bounding from below 1/4 with 0 inside the square root and the second holds because, by **Lemma .4.6** point 3,

$$\lim_{t\to 0}\frac{1/2}{2(\bar{\sigma}\bar{\eta}+\bar{\sigma}\sqrt{t}/2)}+\frac{\bar{\eta}}{\bar{\eta}+\sqrt{t}/2}\,\frac{\varphi_t}{\bar{\sigma}^2\eta_t} = \frac{1}{4\bar{\sigma}^4\bar{\eta}^2}+1 \ .$$

Thus, $m(z)$ is increasing in [1,H] for sufficiently small t □

.6 Proofs of Chapter 3

Proof of Proposition 3.2.1
We bound the range and the discretization error separately.
First, we bound the CDF range error, i.e. the error we introduce considering the integral (3.2.1) in the range $(0, Nh)$. Fix h, it exists $N \in \mathbb{N}$ s.t.

$$\begin{aligned}
&\left| P(x) - \left(1 - \frac{e^{-ax}}{\pi} \int_0^{N\,h} du\; Re\left[e^{-iux} \frac{\phi_{s,t}(u - ia)}{iu + a} \right) \right] \right| \\
&\leq \frac{e^{-ax}}{\pi} \int_{N\,h}^{\infty} du\; \frac{B\, e^{-b\, u^{\omega}}}{u} < \frac{e^{-ax}}{N\,h} \frac{1}{\pi} \int_{N\,h}^{\infty} du\; B e^{-b\, u^{\omega}} \qquad (.6.1) \\
&= \frac{e^{-ax}\, B}{N\,h} \frac{1}{\omega\; b^{1/\omega}} \Gamma\left[\frac{1}{\omega}, b\,(N\,h)^{\omega} \right] = O\left((N\,h)^{-\omega} e^{-b\,(N\,h)^{\omega}} \right) \quad , \qquad (.6.2)
\end{aligned}$$

where $\Gamma(\epsilon, u)$ is the upper incomplete gamma function.
The first inequality is due to $|iu + a| > u$ and to the fact that $|\phi_{s,t}(u - ia)| \leq B e^{-b\, u^{\omega}}$ for sufficiently large values of u, thanks to Assumption 2.

Second, we bound the discretization error.
By theorem 6.2 of Lee (2004), we have that

$$\begin{aligned}
&\left| \frac{e^{-ax}}{2\pi} \int_{\mathbb{R}} du\; e^{-iux} \frac{\phi_{s,t}(u - ia)}{iu + a} - \frac{e^{-ax}}{2\pi} \sum_{j=1}^{M} \frac{e^{-iu_j x} \phi_{s,t}(u_j - ia)}{iu_j + a} \right| \\
&\leq \frac{e^{-2\pi a/h}}{1 - e^{-4\pi a/h}} + \frac{e^{-2\pi(p-a)/h - p\,x}}{1 - e^{-4\pi(p-a)/h}} \phi_{s,t}(-ip) \qquad \forall p < p_t^+ \quad . \qquad (.6.3)
\end{aligned}$$

Notice that $\phi_{s,t}(-i\,p)$ is well defined because $p < p_t^+$.
We select a and p to minimize the discretization error. Notice that, for a sufficiently large M, the leading terms in (.6.3) are $e^{-2\pi a/h}$ and $e^{-2\pi(p-a)/h - p\,x}$. Hence, for a given p the best choice of a is

$$\hat{a} = \frac{p}{2} \left(1 - \frac{x}{\pi}\, h \right) \quad .$$

This last quantity, for a sufficiently small h, is close to $p/2$, whatever x we consider. Thus, to minimize the discretization error, we select $a = p/2$. Then p can be chosen to its maximum value p_t^+ and obtain

$$\left| \frac{e^{-ax}}{2\pi} \int_{\mathbb{R}} du\; e^{-iux} \frac{\phi_{s,t}(u - ia)}{iu + a} - \frac{e^{-ax}}{2\pi} \sum_{j=1}^{M} \frac{e^{-iu_j x} \phi_{s,t}(u_j - ia)}{iu_j + a} \right| \leq \frac{e^{-\pi p_t^+/h} + e^{-\pi p_t^+/h - p_t^+\, x} \phi_{s,t}(-ip_t^+)}{1 - e^{-2\pi p_t^+/h}} \quad . \qquad (.6.4)$$

By combining equations (.6.2,.6.4), we obtain the thesis □

To prove proposition 3.4.1, we report the ATS characteristic function in (1.2.4) and the ATS jump measure.

$$\phi_t(u) = \mathcal{L}_t \left(iu \left(\frac{1}{2} + \eta_t \right) \sigma_t^2 + \frac{u^2 \sigma_t^2}{2};\; k_t,\; \alpha \right) e^{-iu \log \mathcal{L}_t\left(\eta_t \sigma_t^2;\, k_t,\, \alpha \right)} \quad , \qquad \alpha \in (0, 1) \quad .$$

σ_t, k_t are continuous on $[0,\infty)$ and η_t, φ_t are continuous on $(0,\infty)$ with $\sigma_t > 0$, $k_t \geq 0$ and

$$\ln \mathcal{L}_t\left(u;\ k,\ \alpha\right) := \frac{t}{k}\frac{1-\alpha}{\alpha}\left\{1-\left(1+\frac{u\,k}{1-\alpha}\right)^{\alpha}\right\} \ .$$

Moreover, we recall a sufficient condition for the existence of ATS is provided in **Theorem 1.2.1**. The ATS jump measure is

$$\nu_t(x) = \frac{tC\left(\alpha,k_t,\sigma_t,\eta_t\right)}{|x|^{1/2+\alpha}} e^{-(1/2+\eta_t)x} K_{\alpha+1/2}\left(|x|\sqrt{\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(k_t\,\sigma_t^2)}\right) \ ,$$

with

$$C\left(\alpha,k_t,\sigma_t,\eta_t\right) := \frac{2}{\Gamma(1-\alpha)\sqrt{2\pi}}\left(\frac{1-\alpha}{k_t}\right)^{1-\alpha}\sigma_t^{2\alpha}\left(\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(k_t\,\sigma_t^2)\right)^{\alpha/2+1/4} \ ,$$

$K_\nu(z)$ the modified Bessel function of the second kind (see e.g., Abramowitz and Stegun 1948, Ch.9 p.376).

Theorem .6.1. Sufficient conditions for existence of ATS
There exists an additive process $\{f_t\}_{t\geq 0}$ with the characteristic function (1.2.4) if the following two conditions hold.

1. *$g_1(t)$, $g_2(t)$, and $g_3(t)$ are non decreasing, where*

$$\begin{aligned} g_1(t) &:= \left(1/2+\eta_t\right) - \sqrt{\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)} \\ g_2(t) &:= -\left(1/2+\eta_t\right) - \sqrt{\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)} \\ g_3(t) &:= \frac{t^{1/\alpha}\sigma_t^2}{k_t^{(1-\alpha)/\alpha}}\sqrt{\left(1/2+\eta_t\right)^2 + 2(1-\alpha)/(\sigma_t^2 k_t)} \ ; \end{aligned}$$

2. *Both $t\,\sigma_t^2\,\eta_t$ and $t\,\sigma_t^{2\alpha}\,\eta_t^{\alpha}\,/k_t^{1-\alpha}$ go to zero as t goes to zero.*

Proof of Proposition 3.4.1 We prove that Assumption 1 holds.
This entails showing that p_t^+ is non increasing for the ATS.
At time t, the ATS characteristic function in equation (1.2.4) is analytic on the imaginary axis $u = i\,a$, $a \in \mathbb{R}$ iff

$$1 + \frac{k_t}{1-\alpha}\left(a\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 - \frac{a^2\sigma_t^2}{2}\right) > 0 \ .$$

By solving the second order inequality, we get

$$g_1(t) < a < -g_2(t) - 1 \ .$$

Hence, $p_t^+ = -g_2(t) - 1$ and $p_t^- = -g_1(t)$. Notice that $-g_2(t)$ and $-g_1(t)$ are non increasing w.r.t. t because of condition 1 in **Theorem 1.2.1** and p_t^- and p_t^+ are positive through direct inspection. Moreover, $p_t^+ \geq p_t^-$ because

$$p_t^- \geq p_t^- = \eta_t \geq 0 \ .$$

We prove that Assumption 2 holds. We have to show that, given s and $t > s$, there exists $B > 0$, $b > 0$ and $\omega > 0$ such that, for sufficiently large u, Assumption 2 holds for the characteristic

function of ATS.

We choose $\log(B) > \frac{1-\alpha}{\alpha}\left(\frac{t}{k_t} - \frac{s}{k_s}\right)$, $0 < b < \frac{1-\alpha}{\alpha}\left(\frac{t}{k_t^{1-\alpha}}\left(\frac{\sigma_t^2}{2(1-\alpha)}\right)^{\alpha} - \frac{s}{k_s^{1-\alpha}}\left(\frac{\sigma_s^2}{2(1-\alpha)}\right)^{\alpha}\right)$, $\omega < 2\alpha$. Notice that it is possible to fix $b > 0$ because if $g_3(t) > g_3(s)$ by condition 1 of theorem 1.2.1 $\frac{t}{k_t^{1-\alpha}}\left(\frac{\sigma_t^2}{2(1-\alpha)}\right)^{\alpha}$ is increasing in t. For sufficiently large u, $|\phi_{t,s}(u - i\,a)|$ goes to zero faster than $Be^{-b\,u^{\omega}}$ because $log\phi_{t,s}(u - i\,a)$ is asymptotic to

$$-\frac{1-\alpha}{\alpha}\left(\frac{t}{k_t^{1-\alpha}}\left(\frac{\sigma_t^2}{2(1-\alpha)}\right)^{\alpha} - \frac{s}{k_s^{1-\alpha}}\left(\frac{\sigma_s^2}{2(1-\alpha)}\right)^{\alpha}\right)u^{2\alpha} \quad .$$

We prove that Assumption 3 holds. We have to demonstrate that $m^+_{s,t}(x)$ is non increasing in x and $m^-_{s,t}(x)$ is non decreasing.
We prove the thesis by showing that the derivative of $\nu_t(x)$ w.r.t. x is non increasing w.r.t. t for any $x > 0$. Notice that if this holds then

$$m^+_{s,t}(x) = \mathbb{I}_{x>\epsilon}\frac{\nu_t(x) - \nu_s(x)}{\int_{\epsilon}^{\infty} dz(\nu_t(z) - \nu_s(z))}$$

is non increasing in x.
We observe that by the condition on $g_1(t)$ and $g_2(t)$ of theorem 1.2.1 we have that that

$$g_t := \sqrt{(1/2 + \eta_t)^2 + 2(1-\alpha)/(k_t\,\sigma_t^2)}$$

is non increasing w.r.t. t.

$$\frac{\partial \nu_t(x)}{\partial x} = -C_2\int_0^{\infty} dz\frac{e^{-z}z^{\alpha}g_3(t)\,e^{x g_2(t)}}{x^{2+\alpha}}\left(\alpha + \frac{z}{z/2 + x\,h_t} + 1 - x\,g_2(t)\right) \quad ,$$

where C_2 is a positive constant. The derivative of $\nu_t(x)$ is non increasing for any $x > 0$ because

1. $g_3(t)$ is positive and non decreasing in t by condition 1 of **Theorem 1.2.1**.
2. $e^{x\,g_2(t)}\left(\alpha + \frac{s}{s/2 + x\,g_t}\right)$ is the combination of two non decreasing function of t for any $x > 0$.
3. $g_2(t)$ is non decreasing and $(1 - cx)e^{c\,x}$ is non decreasing for $c < 0$.

This proves the thesis for $x > 0$.
The same holds true for $x < 0$. *Mutatis mutandis*, by substituting $g_2(t)$ with $g_1(t)$, the proof is the same □

.7 An application to commodities

We analyze all quoted NYMEX WTI future and american option prices observed at 11:00 am NT on the 23^{rd} of June 2021. The dataset is composed of real market quotes (no smoothing or interpolation). Let us recall that volume of crude oil futures and options is more of 50% of the total energy contracts traded on NYMEX. Option and futures data is provided by Eikon Reuters (chain $0\#CL+$). Futures stop trading three business days prior to the 25^{th} of each delivery month (if the 25^{th} is not a business day trading ceases three business days prior to the first business day before the 25^{th}). Options expire three business days before the futures termination dates. We follow Kyriakou *et al.* (2016) and consider only options expiring on the first six months, on March, June, September and December of the front year and on December of the next year.
The dataset contains all bid/ask prices for both call and put. The strikes are in a regular grid for each available maturity. We exclude options that do not satisfy two simple liquidity thresholds. We discard options whose price is less than 10% the minimum difference in the grid of market strikes (the so-called *penny options*) and options with bid-ask over bid bigger than 60%. The last condition filters out strikes for which either a bid or an ask price is missing.

We calibrate the ATS following the procedure discussed in chapter 1. We cut the volatility surface into slices, each one containing options with the same maturity, and calibrate each slice separately. Hereinafter, we focus on $\alpha = 1/2$ (NIG) and $\alpha = 0$ (VG), which are the two (ATS and Sato) generalizations of the two most frequently used LTS processes. For every fixed maturity T, it is possible to define a new Lévy normal tempered stable process such that, at time T, its marginal distribution is equal to the marginal distribution of an ATS. A different Lévy NIG and VG is calibrated for every different maturity and the three time-dependent parameters k_T, η_T, σ_T are obtained. The calibration is performed imposing the ATS conditions of monotonicity.
Beneath the ATS processes, we consider the calibration of the standard Lévy processes and of the (four parameters) Sato processes proposed by Carr *et al.* (2007).[1] We remind that the latter are additive and self-similar processes (see, e.g., Sato 1991). Call option prices, with strike K and maturity T, are computed using the Lewis (2001) formula.
The calibration is performed by minimizing the Euclidean distance between model and market prices. The simplex method is used to calibrate every maturity of the ATS process. For Lévy processes and Sato processes, because standard routines for global minimum algorithms are not satisfactory, we consider a differential evolution algorithm together with a multi-start simplex method.
The calibration performance is reported in Table .4 in terms of Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE).[2] It is possible to observe that Sato processes slightly improve Lévy performance, as reported in the literature (see, e.g Carr *et al.* 2007), while the ATS processes improvement is, on average, above two orders of magnitude. Although we present the results for VG and NIG, similar results can be obtained for all ATS processes with $\alpha \in [0, 1)$. The worst results are observed in the VG case.

	MSE			MAPE		
Model	Lévy	Sato	ATS	Lévy	Sato	ATS
NIG	0.02	$5 * 10^{-3}$	$\mathbf{7 * 10^{-5}}$	3.16%	1.78%	**0.19**%
VG	0.02	$5 * 10^{-3}$	$\mathbf{10^{-3}}$	3.23%	1.77%	**0.84**%

[1] We underline that, in both cases (LTS and Sato), model parameters are obtained through a global calibration of the whole volatility surface.
[2] Calibrated model parameters are available upon request.

Table .4: Calibration performance for the WTI options in terms of MSE and MAPE. In the NIG and VG cases, we consider the standard Lévy process, the Sato process, and the corresponding ATS process. Sato processes perform better than Lévy processes but ATS improvement is far more significant: two orders of magnitude for MSE and one order of magnitude for MAPE.

Figure .1 shows the differences of MSE w.r.t. the different times-to-maturity for WTI volatility surface calibrated with a NIG process. Sato and Lévy LTS have a MSE of the same order of magnitude, while the improvement of ATS is of two orders of magnitude and particularly significant at short-time. The short time improvement in implied volatility calibration is particularly evident, as shown in Figure .2.

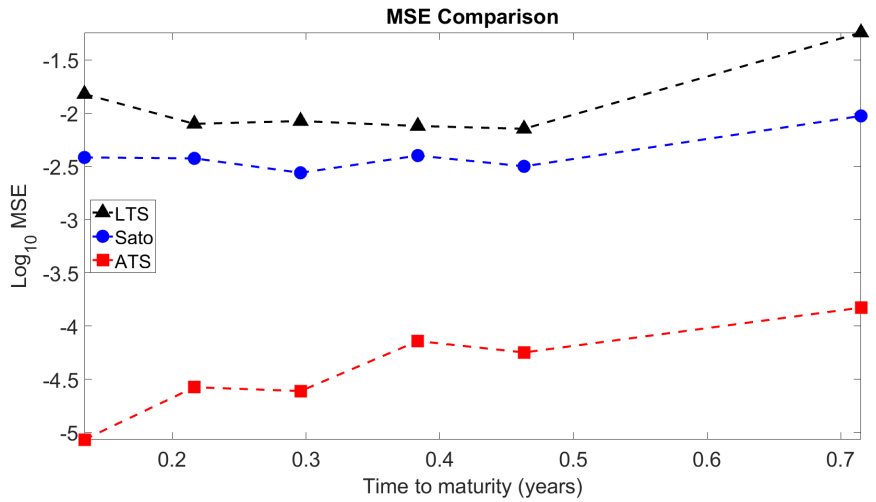

Figure .1: MSE w.r.t. the different times-to-maturity (in years) for S&P 500 volatility surface calibrated with a NIG process. Sato (circles) and Lévy (triangles) have a MSE of the same order of magnitude, while the improvement of ATS (squares) is of two orders of magnitude and particularly significant at short-time.

In Figure 1.5, we plot the market implied volatility and the volatility replicated via ATS, LTS, and Sato processes at the 1 month and 24 days (on the left) and 9 months and 24 days (on the right) maturities. We observe that the ATS implied volatility is the closest to the market implied volatility in any case and it significantly improves both LTS and Sato processes, particularly for small maturities. Similar results hold for all other ATS.

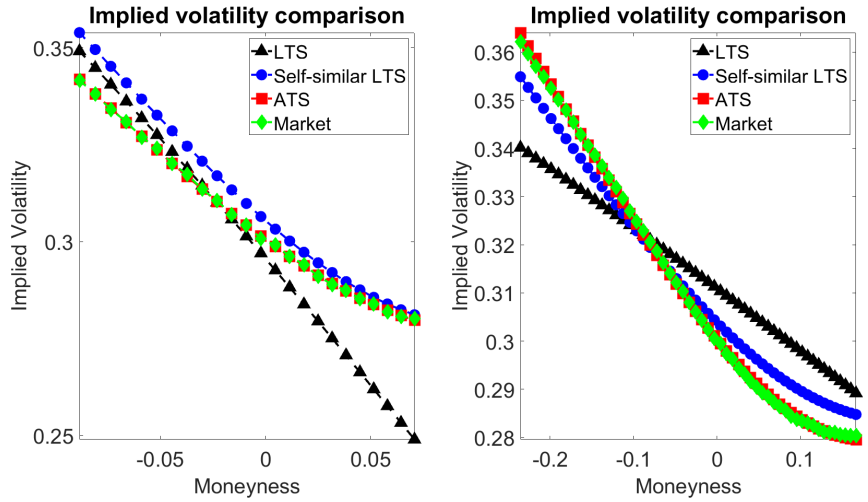

Figure .2: Implied volatility smile for WTI options for a given time-to-maturity: 1 month and 24 days (on the left) and 9 months and 24 days (on the right). The NIG ATS process, Sato process, and LTS process implied volatilities are plotted together with the market-implied volatility. ATS reproduces the smile significantly better then the alternatives, the improvement is particularly evident for small maturities.

In Figure .3, we plot the market and the ATS implied volatility *skew* w.r.t. the times-to-maturity. We observe that the calibrated ATS replicates accurately the market implied volatility *skew*.

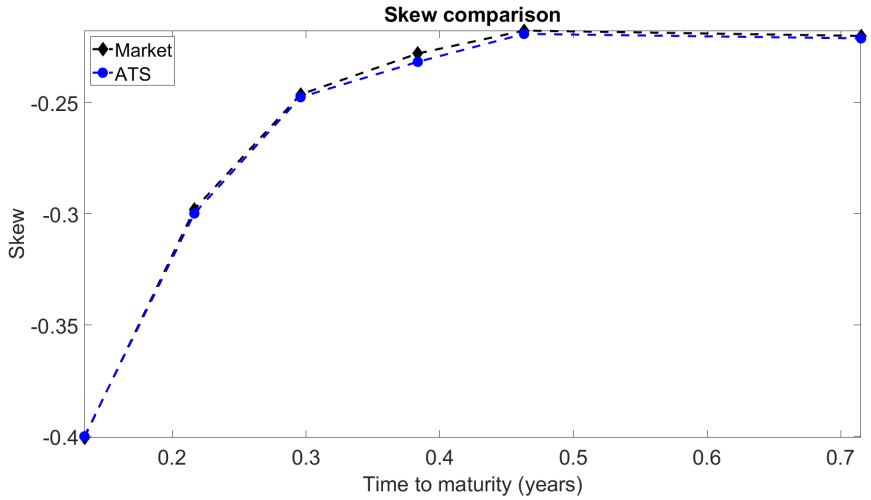

Figure .3: The market and the NIG ATS implied volatility *skew* w.r.t. the times-to-maturity. ATS replicates the market implied volatility *skew* behavior.

Hereinafter, we show that power-law scaling parameters are observed in market data. This stylized fact is extremely relevant: we observe statistical evidence that the market-implied volatility surface is compatible with a power-law scaling ATS.

We introduce a new ATS process, w.r.t. the time $\theta := T\sigma_T^2$.
We define $\hat{k}_\theta := k_T\sigma_T^2$ and $\hat{\eta}_\theta := \eta_T$. We call $\hat{f}_\theta$ the ATS with parameters $\hat{k}_\theta$, $\hat{\eta}_\theta$ and $\sigma_\theta := 1$.

Notice that for every maturity $\hat{f}_\theta$ has the same characteristic function of the calibrated ATS f_T. We analyze the re-scaled parameters. We observe a self-similar behavior of $\hat{k}_\theta$ and $\hat{\eta}_\theta$; that is reported in equation equations (1.3.2) ,

$$\begin{cases} \hat{k}_\theta &= \bar{k}\theta^\beta \\ \hat{\eta}_\theta &= \bar{\eta}\theta^\delta \end{cases},$$

where $\bar{k}$, $\bar{\eta}$ are positive constants and β, δ are real constant parameters. To investigate this behavior and to infer the value of the scaling parameters we consider equations (1.3.2) in the log-log scale.
In Figures .4 and .5 we plot the weighted regression lines and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ with their confidence intervals . The confidence intervals are two times the standard deviations of $\ln \hat{k}_\theta$, of $\ln \hat{\eta}_\theta$ and of $\ln \theta$.

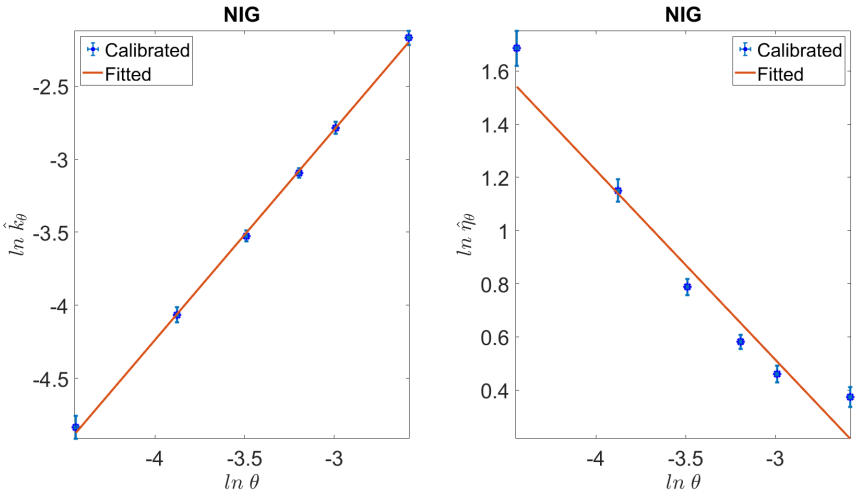

Figure .4: Weighted regression line and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ w.r.t. $\ln \theta$ for the NIG calibrated model for WTI. We plot a confidence interval equal to two times the corresponding standard deviation. Notice that confidence intervals on $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ are one order of magnitude wider than confidence intervals on $\ln \theta$.

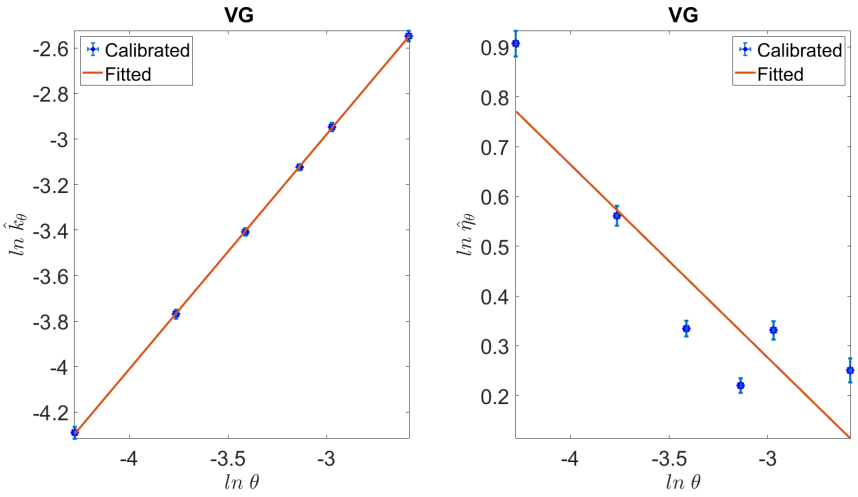

Figure .5: Weighted regression line and the observed time-dependent parameters $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ w.r.t. $\ln \theta$ for the VG model calibrated on WTI. We plot a confidence interval equal to two times the corresponding standard deviation. Notice that confidence intervals on $\ln \hat{k}_\theta$ and $\ln \hat{\eta}_\theta$ are one order of magnitude wider than confidence intervals on $\ln \theta$.

Moreover, let us emphasize that the scaling parameters appear qualitatively compatible to $\beta = 1$ and $\delta = -\frac{1}{2}$ in all observed cases but the tests p-values are always below 10%. It seems that in the commodity case a different scaling is observed in market data. This deserves further investigations.

Model	Parameter	Parameter's Value	p-value
NIG	β	1.44	0.00
NIG	δ	-0.68	0.00
VG	β	1.03	0.10
VG	δ	-0.38	0.00

Table .5: Scaling parameters calibrated from WTI volatility surfaces. Parameter estimates are provided together with the p-values of the statistical tests that verify whether it is possible to accept the null hypotheses $\beta = 1$ and $\delta = -\frac{1}{2}$.

In Table .6 we report an estimation of the parameter $\bar{k}$ and $\bar{\eta}$.

Model	Parameter	Parameter's Value	p-value
NIG	$\bar{k}$	4.65	0.002
NIG	$\bar{\eta}$	0.21	0.000
VG	$\bar{k}$	1.12	0.000
VG	$\bar{\eta}$	0.41	0.001

Table .6: $\bar{k}$ and $\bar{\eta}$ calibrated from WTI volatility surfaces. Parameter estimates are provided together with the p-values of the statistical tests that verify whether it is possible to accept the null hypothesis $\bar{k} = 0$ and $\bar{\eta} = 0$.

We have statistical evidence that in all cases $\bar{k}$ and $\bar{\eta}$ are positive (we reject the null hypotheses of $\bar{k} = 0$ and $\bar{\eta} = 0$ with a 5% confidence level). From these results and from Figure .3 it is possible to infer a connection between a positive $\bar{\eta}$ and the observed negative *skew*.