Lecture Notes Computational Finance

Michele Azzone[‡]

November 28, 2025

(‡) Politecnico di Milano, Department of Mathematics, 32 p.zza L. da Vinci, Milano

Address for correspondence:
Michele Azzone
Department of Mathematics
Politecnico di Milano
32 p.zza Leonardo da Vinci
I-20133 Milano, Italy
michele.azzone@polimi.it

Notes

1 Lévy and Additive processes

1.1 Basic concepts and notation

Definition 1.1. A function $f : [0, T] \to \mathbb{R}$ is **càdlàg** (a French acronym for *continue à droite, limite à gauche*, meaning "right-continuous with left limits") if it satisfies the following two conditions:

1. **(Left limits exist)** For all $t \in (0, T]$, the left-hand limit
$$f(t_-) := \lim_{s \to t, s<t} f(s)$$
exists and is finite.

2. **(Right-continuous)** For all $t \in [0, T)$, the right-hand limit $f(t_+) := \lim_{s \to t, s>t} f(s)$ exists and is equal to the function's value at t:
$$f(t) = f(t_+)$$

Definition 1.2. A **stochastic process** $(Z_t)_{t \geq 0}$ is a collection of random variables $\{Z_t\}_{t \geq 0}$, indexed by time $t \in [0, \infty)$, all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

We indicate the process by (Z_t) and the random variable at time t as Z_t.

Definition 1.3. A **càdlàg Process** is a stochastic process with **càdlàg** paths.

Definition 1.4. $\hat{\mu}(u) = \mathbb{E}[e^{iuX}]$ is the characteristic function of the random variable X.

Definition 1.5. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space. (X_t) is a **martingale** with respect to the filtration $(\mathcal{F}_t)$ if it satisfies the following three conditions:

1. **(Adapted)** X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.

2. **(Integrable)** $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$.

3. **(Martingale Property)** For all $s \leq t$,
$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$$

Proposition 1.6. *Let X be an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$*
Let $\mathcal{G}_1$ be a sub-σ-algebras of $\mathcal{F}$ s $\mathcal{G}_1 \subseteq \mathcal{F}$.
Then, the following identity holds almost surely:

$$\mathbb{E}\left[\mathbb{E}[X \mid \mathcal{G}_1]\right] = \mathbb{E}[X]$$

Definition 1.7 (Measure)**.** Let $(X, \mathcal{F})$ be a measurable space, where X is a set and $\mathcal{F}$ is a σ-algebra on X.
A **measure** is a function $\mu : \mathcal{F} \to [0, \infty]$ that satisfies the following two properties:

1. **(Null Empty Set)** $\mu(\emptyset) = 0$.

2. **(Countable Additivity / σ-additivity)** For any sequence $\{E_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{F}$ (i.e., $E_i \cap E_j = \emptyset$ for $i \neq j$),

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) \ .$$

Definition 1.8 (Measure with a Density)**.** Let $\nu : \mathbb{R} \to [0, \infty)$ be a Lebesgue measurable function. The integral measure ν on $\mathcal{B}(\mathbb{R})$ is

$$\nu(B) = \int_B \nu(x)\, dx \quad \text{for every } B \in \mathcal{B}(\mathbb{R}) \ .$$

1.2 Poisson Process

Definition 1.9 (Exponential Distribution)**.** A continuous random variable Y follows an Exponential distribution with rate parameter $\lambda > 0$, denoted $Y \sim \text{Exp}(\lambda)$, if its probability density function (PDF) $f(x)$ is:

$$f(y) = \lambda e^{-\lambda y} \mathbf{1}_{[0,\infty)}(y)$$

The CDF is $F(y) = \int_0^y f(x)dx = 1 - e^{-\lambda y}$ and the characteristic function is

$$\mathbb{E}[e^{iuY}] = \int_0^{\infty} e^{iux} \lambda e^{-\lambda x} dx = \frac{\lambda}{iu - \lambda} e^{iux} \lambda e^{-\lambda x} \ |_0^{\infty} = \frac{\lambda}{iu - \lambda} \ .$$

Definition 1.10 (Poisson Distribution)**.** A discrete random variable N follows a Poisson distribution with parameter $\lambda > 0$ (the average rate or mean), denoted $X \sim \text{Pois}(\lambda)$, if its probability mass function (PMF) $p(k)$ is given by:

$$p(k) = \mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$

$\mathbb{E}[N] = \lambda$ and CF of $\text{Pois}(\lambda)$ is $\exp(\lambda(e^{iu} - 1))$.

Definition 1.11 (Gamma Distribution)**.** A continuous random variable X follows a Gamma distribution with shape parameter $k > 0$ and scale parameter $1/\lambda$, denoted $X \sim \text{Gamma}(1/\lambda, k)$, if its PDF is given by:

$$f(x) = \frac{1}{\Gamma(k)} \lambda^k x^{k-1} e^{-\lambda x} \quad \text{for } x \geq 0$$

where $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t}\, dt$ is the **Gamma function**.

The characteristic function is $(1 - iu/\lambda)^{-k}$.

Proposition 1.12. *τ_j i.i.d., $\tau_j \sim \exp(\lambda)$. $T_k = \sum_{j=1}^{k} \tau_j \sim \Gamma(1/\lambda, k)$.*

Proof.

$$\mathbb{E}[e^{iuT_k}] = \prod_{j=1}^{k} \mathbb{E}[e^{iu\tau_j}] = \left(\frac{\lambda}{iu - \lambda}\right)^k = (1 - iu/\lambda)^{-k} \ .$$

The first equality is due to the independence of the τ_j the second to the fact that they are all $\exp(\lambda)$. The thesis follow because the characteristic function of T_k is the one of a Gamma RV. $\square$

Proposition 1.13.

$$N_t = \sum_{n\geq 1} \mathbf{1}_{\{t\geq T_n\}} \sim \text{Poiss}(\lambda t)$$

Proof.

$$T_n = \sum_{j=1}^{n} \tau_j \quad \text{(sum of first } n \text{ interarrival times, each } \tau_j \sim \text{Exp}(\lambda))$$

In particular

$$P(N_t = n) = P(T_n \leq t < T_{n+1})$$

because it is the probability that the sum of the first n τ_j is below or equal to t while the sum of the first $n+1$ τ_j is above. Define $A = \{T_{n+1} \geq t\}$, $B = \{T_n \geq t\}$. Notice that $T_n \geq t$ implies $T_{n+1} \geq t$, then $B \subseteq A$.

$$P(T_n \leq t < T_{n+1}) = P(A \setminus B) = P(A) - P(B)\ .$$

$$\begin{aligned} P(B) = P(T_n \geq t) &= \int_t^\infty \frac{(\lambda x)^{n-1}\lambda e^{-\lambda x}}{(n-1)!}\,\mathrm{d}x \\ &= \int_t^\infty \frac{e^{-\lambda x}}{(n)!}\,\mathrm{d}\left[(\lambda x)^n\right] \\ &= \frac{e^{-\lambda x}(\lambda x)^n}{n!}\bigg|_t^\infty + \int_t^\infty \frac{\lambda e^{-\lambda x}(\lambda x)^n}{n!}\,\mathrm{d}x \\ &= P(A) - \frac{e^{-\lambda t}(\lambda t)^n}{n!} \end{aligned}$$

Then,

$$P(N_t = n) = P(A) - P(B) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \sim \text{Poiss}(\lambda t)\ .$$

□

Remark 1.14. The process $N_t = \sum_{n\geq 1} \mathbf{1}_{\{t\geq T_n\}}$ is called a counting process because it counts the number of events (or "jumps") that have occurred up to (and including) time t. The variables T_n represent the jump times (i.e., the time of the n-th event), while the τ_j (where $T_n = \sum_{j=1}^n \tau_j$) are the interarrival times between consecutive jumps. The parameter λ is called the jump intensity or rate. This terminology is used because the expected number of jumps by time t is $\mathbb{E}[N_t] = \lambda t$, meaning λ represents the expected number of jumps per unit time.

Definition 1.15. The counting process $(N_t) = \left(\sum_{n\geq 1} \mathbf{1}_{\{t\geq T_n\}}\right)$ is called a Poisson process.

Proposition 1.16. *1. For any $t > 0$, N_t is finite a.s..*

2. *The paths of (N_t) are piecewise constant and increase by jumps of size one. The paths are càdlàg .*

3. *(N_t) is stochastically continuous, i.e., $\forall t > 0$ and $\forall \varepsilon > 0$,*

$$\lim_{h\to 0} P(|N_{t+h} - N_t| > \varepsilon) = 0.$$

4. *The characteristic function is given by:*

$$\mathbb{E}[e^{iuN_t}] = \exp(\lambda t(e^{iu} - 1)).$$

5. (N_t) *has* ***independent increments****: for any sequence of times* $0 \leq t_0 < t_1 < \cdots < t_n$*, the random variables*

$$(N_{t_1} - N_{t_0}), (N_{t_2} - N_{t_1}), \ldots, (N_{t_n} - N_{t_{n-1}})$$

are mutually independent.

6. (N_t) *has* ***stationary increments****: for any* $s, t \geq 0$*, the distribution of the increment* $N_{t+s} - N_s$ *depends only on* $t - s$*. Specifically,*

$$N_{t+s} - N_s \sim N_t.$$

Proof. 1. T_k is a sum of i.i.d. exponential random variables with mean $1/\lambda$. $\frac{T_k}{k} \to 1/\lambda$ a.s. as $k \to \infty$ by the law of large numbers. Then $T_k \to \infty$ a.s. and then $\forall t > 0, \exists n_0(t)$ s.t. $T_n > t, \forall n > n_0$ a.s. Then N_t is finite a.s.

2. (N_t) is constant in each interval (T_n, T_{n+1}) and increases by one at each T_n. Moreover, the path are right continous by definition.

3. In the interval $[0, T]$ is exists n_0 s.t. there are at most n_0 jumps in the path. Then, $P(N_{t^-} \neq N_t) = 0$ because the probability that the jump is exactly in t is zero and then $N_{t^-} = N_t$ a.s. This jointly with $N_t = N_{t^+}$ by definition ensures convergence in probability for any t that ensures stochastic continuity.

4. From definition of Poisson RV.

□

Proposition 1.17. (N_t^1), (N_t^2) *two independent Poisson process with intensity* λ_1 λ_2*. Then* $(N_t^1 + N_t^2)$ *is a Poisson process with intensity* $\lambda_1 + \lambda_2$*.*

Definition 1.18. A compensated Poisson process is $(\tilde{N}_t) := (N_t - \lambda t)$.

Let us notice that

$$\mathbb{E}[e^{iu\tilde{N}_t}] = \exp(\lambda t(e^{iu} - 1 - iu)$$

$\tilde{N}_t$ has independent and stationary increments because its increments are $N_t - N_s - (t - s)\lambda$ and

$$\mathbb{E}[\tilde{N}_t|\mathcal{F}_s] = \mathbb{E}[N_t - \lambda t|\mathcal{F}_s] = \mathbb{E}[N_t - N_s + N_s - \lambda t|\mathcal{F}_s] = \lambda(t - s) + N_s - \lambda t = \tilde{N}_s \ ,$$

which implies that $\tilde{N}_t$ is a martingale.

Theorem 1.19. *Let* (X_t) *be a counting process with stationary and independent increments. Then* (X_t) *is a Poisson process.*

1.3 Lévy processes

Definition 1.20. A probability distribution μ on $\mathbb{R}$ is **infinitely divisible** if, for all $n \in \mathbb{N}$, there exists a probability distribution μ_n on $\mathbb{R}$ such that

$$\hat{\mu}(u) = (\hat{\mu}_n(u))^n, \quad \forall u \in \mathbb{R}$$

his is equivalent to saying that for a random variable X with distribution μ, for all $n \in \mathbb{N}$, there exist independent and identically distributed (i.i.d.) random variables $X_1, \ldots, X_n$ such that X has the same distribution as their sum:

$$X \stackrel{d}{=} \sum_{j=1}^{n} X_j$$

Consider X with characteristic function $\hat{\mu}$ and X_j with characteristic function $\hat{\mu}_n$. By the properties of characteristic functions and the independence of the X_j:

$$\hat{\mu}(u) = \mathbb{E}[e^{iuX}] = \mathbb{E}\left[e^{iu\sum_{j=1}^n X_j}\right] = \mathbb{E}\left[\prod_{j=1}^n e^{iuX_j}\right] = \prod_{j=1}^n \mathbb{E}[e^{iuX_j}] = \prod_{j=1}^n \hat{\mu}_n(u) = (\hat{\mu}_n(u))^n$$

Definition 1.21. A stochastic process (X_t) defined on $\mathbb{R}$, $X_0 = 0$ almost surely, is a **Lévy process** if it satisfies the following properties:

1. Independent increments
2. Stationary increments
3. Stochastic continuity

Example 1.22. A Poisson process is a Lévy process. By proposition 1.16 we verify the 3 conditions above moreover $N_0 = 0$ a.s. by construction of Poisson process.

Example 1.23. (W_t) is a Lévy process. $W_t - W_s$ independent from W_s by definition of BM. $W_{t+s} - W_t \sim N(0, t-s)$ stationary incrementets. The process is continous and thas stochastic continuity follow and $W_0 = 0$ a.s. The unique contionuous (pure diffusion) Lévy process is the Brownian motion.

Notice that stochastic continuity does not imply continuity (e.g. Poisson process path have jumps). It means that the probability of observing a finite jump ($> \epsilon$) at a fixed time t is zero.

Lemma 1.24. *If (X_t) is a Lévy process than the distribution of X_t is infinitely divisible.*

Proof. For each n define a grid of step size t/n: $t_0 < t_1 < ....t_n = t$. We have that

$$X_t = \sum_{j=1}^n (X_{t_j} - X_{t_{j-1}})$$

by definition of telescopic sum.

$$\hat{\mu}(u) := \mathbb{E}[e^{iuX_t}] = \mathbb{E}\left[e^{iu\sum_{j=1}^n (X_{t_j} - X_{t_{j-1}})}\right] = \mathbb{E}\left[\prod_{j=1}^n e^{iu(X_{t_j} - X_{t_{j-1}})}\right] = \prod_{j=1}^n \mathbb{E}\left[e^{iu(X_{t_j} - X_{t_{j-1}})}\right] \quad ,$$

where the last equality is due to independence of increments and

$$\prod_{j=1}^n \mathbb{E}\left[e^{iu(X_{t_j} - X_{t_{j-1}})}\right] = (\hat{\mu}^{(n)}(u))^n$$

due to stationarity of increments, where $\hat{\mu}^{(n)}(u) = \mathbb{E}\left[e^{iu(X_{t_j} - X_{t_{j-1}})}\right]$. This proves the thesis by definition of infinitely divisible distribution. □

Theorem 1.25. *If μ is an infinitely divisible distribution on $\mathbb{R}$ then*

$$\hat{\mu}(u) = exp\left(-1/2u^2 A + iu\gamma + \int_{\mathbb{R}} e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1}\nu(dx)\right) \quad ,$$

where A positive scalar, $\gamma \in \mathbb{R}$ and ν is a measure on $\mathbb{R}$ such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}} x^2 \wedge 1 \nu(dx) < \infty$

Proof. See Theorem 8.1 Sato (1999) □

A diffusion term, γ drift and ν Lévy measure or jump measure. (A, ν, γ) is called generating triplet of the distribution

Example 1.26. W_t BM. $\mathbb{E}[e^{iuW_t}] = e^{-u^2 t}$. $A = t$, $\gamma = 0$, $\nu = 0$.

Example 1.27. Compensated Poisson, CF=$exp(\lambda t(e^{iu-1-iu})) = exp(\int_{\mathbb{R}}(e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1})\lambda t \delta_1(dx))$ where δ_1 is the Dirac's delta in 1. Then $\nu = \lambda t \delta_1$.

1.4 Additive processes and infinitely divisible distributions

Definition 1.28. A stochastic process (X_t) defined on $\mathbb{R}$, $X_0 = 0$ almost surely, is an **Additive process** if it satisfies the following properties:

1. Independent increments
2. Stochastic continuity

Additive processes are an extension of Lévy processes that relax the assumption of stationary increments.

Theorem 1.29 (Theorem 9.8 Sato)**.** *1. If (X_t) is an additive process on $\mathbb{R}$ the distribution of X_t is infinitely divisible with generating triplet (A_t, ν_t, γ_t) such that*

a) ***Initialization:***

$$A_0 = 0, \quad \nu_0 = 0, \quad \gamma_0 = 0.$$

b) ***Monotonicity:*** *For any $0 \leq s \leq t$:*

- $A_t - A_s \geq 0$
- $\nu_t(B) \geq \nu_s(B)$ *for all Borel sets $B \in \mathcal{B}(\mathbb{R})$.*

c) ***Continuity:***

- A_t *and* γ_t *are continuous in t;*
- $\nu_t(B)$ *is continuous in t for any Borel set B bounded away from the origin (i.e., $B \subset \{x : |x| > \varepsilon\}$ for some $\varepsilon > 0$).*

2. Given a family of generating triplets (A_t, ν_t, γ_t) that satifies (a,b,c) then, there exist a unique in law additive process (X_t) whit probability measure (μ_t). Such that μ_t is infinitely divisible with triplet (A_t, ν_t, γ_t).

Corollary 1.30. *1) If (X_t) is a Lévy process the generating triplet of X_t (A_t, ν_t, γ_t) is $(tA, t\nu, t\gamma)$ and 2) given an infinitely divisible distribution μ with triplet (A, ν, γ) it exists a Lévy process with generating triplet $(tA, t\nu, t\gamma)$ that has law μ at time $t = 1$.*

Proof. We prove first point one. By independence of increments for any additive process

$$\mathbb{E}\left[e^{iuX_t}\right] = \mathbb{E}\left[e^{iu(X_t - X_s + X_s)}\right] = \mathbb{E}\left[e^{iu(X_t - X_s)}\right]\mathbb{E}\left[e^{iuX_s}\right] \quad .$$

Then,

$$\mathbb{E}\left[e^{iu(X_t - X_s)}\right] = \mathbb{E}\left[e^{iuX_t}\right] / \mathbb{E}\left[e^{iuX_s}\right] = \exp\left(iu(\gamma_t - \gamma_s) - 1/2u^2(A_t - A_s) + \int_{\mathbb{R}} ...(\nu_t - \nu_s)(dx)\right) \quad .$$

By stationarity of increments of the Lévy process we need that $\gamma_{t+h} - \gamma_t = \gamma_h$ $A_{t+h} - A_t = A_h$ and $\nu_{t+h}(B) - \nu_t(B) = \nu_h(B)$ $\forall B \in \mathcal{B}(\mathbb{R})$ and does not depend on t. You can check that the unique function of t that satisfy such property and is continous is linear in t and then $\gamma_t = t\gamma$, $A_t = tA$ and $\nu_t(B) = t\nu(B)$.
We prove now point 2. We use Therem 9.8 of Sato. We check the conditions

1. $A_0 = 0$, $\nu_0(B) = 0 * \nu(B) = 0$, $\gamma_0 = 0$.
2. $A_t = tA$, $\nu_t(B) = t\nu(B)$ are linear in t and then increasing (because A and $\nu(B)$) are positive.

3. $A_t = tA$, $\gamma_t = t\gamma$ and $\nu_t(B) = t\nu(B)$ are continuous in t.

Then it exist an additive process (X_t) with law characterized by this triplet. We have to check that it is stationary

$$\mathbb{E}\left[e^{iu(X_{t+h}-X_t)}\right] = \exp\left(iu\gamma h - 1/2u^2Ah + \int_{\mathbb{R}} ...h\nu(dx)\right) \quad .$$

which does not depend from t. Then, X_t is a Lévy process and this proves the thesis □

Example 1.31. The BM $\mathbb{E}[e^{iuW_t}] = e^{-u^2/2t}$. $A = 1$.Compensated Poisson, CF=$exp(\lambda t(e^{iu-1-iu}))$ = $exp(\int_{\mathbb{R}}(e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1})\lambda t\delta_1(dx))$ where δ_1 is the Dirac's delta in 1. Then $\nu = \lambda\delta_1$.

Corollary 1.32. *If (X_t) is a Lévy process then it exists $\psi(u)$ s.t. $\hat{\mu}_t(u) = exp(t\psi(u))$.*

Proof. By corollary 1.30 the characteristic function of of (X_t) is

$$\hat{\mu}_t(u) = exp\left(-t/2u^2A + tiu\gamma + t\int_{\mathbb{R}} e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1}\nu(dx)\right) \quad .$$

The thesis follows by identifying

$$\psi(u) := -1/2u^2A + iu\gamma + \int_{\mathbb{R}} e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1}\nu(dx) \quad .$$

□

Proposition 1.33. *Let (X_t) be a Lévy process with characteristic triplet (γ, A, ν).*

1. *The n-th absolute moment $\mathbb{E}[|X_t|^n]$ is finite for all t if and only if $\int_{|x|\geq 1} |x|^n\nu(dx) < \infty$.*
2. *If this condition is satisfied, the n-th cumulant $c_n(X_t)$ of X_t is given by:*

$$c_1(X_t) = t\left(\gamma + \int_{|x|\geq 1} x\nu(dx)\right), \tag{1}$$

$$c_2(X_t) = t\left(A + \int_{-\infty}^{\infty} x^2\nu(dx)\right), \tag{2}$$

$$c_n(X_t) = t\int_{-\infty}^{\infty} x^n\nu(dx) \quad for\ n \geq 3. \tag{3}$$

Notice that variance ($c_2(X_t)$) is controlled both by the jumps and the diffusion term (over-parametrization). Skewness is $c_3(X_t)/c_2(X_t)^{3/2}$ and excess kurtosis $c_4(X_t)/c_2(X_t)^2$.

1.5 The compound Poisson

Definition 1.34. A stochastic process (X_t) is said to be a *Compound Poisson Process* if it can be represented as:

$$X(t) = \sum_{j=1}^{N_t} Y_j, \quad t \geq 0$$

where:

- (N_t) is a Poisson process with intensity rate $\lambda > 0$.

- Y_j i.i.d. random variables.
- The process (N_t) and the Y_j are independent.

Proposition 1.35. *The compound Poisson $X_t = \sum_{j=1}^{N_t} y_j$; $y_j \perp\!\!\!\perp N_t$; $y_j \sim f$ i.i.d. is a Lévy process.*

Proof. We prove that a Compound Poisson is a Lévy process by verifying directly the definition of Lévy. For every increment

$$X_t - X_s = \sum_{j=1}^{N_t - N_s} y_j^* \perp\!\!\!\perp \sum_{j=1}^{N_s} y_j = X_s \ ,$$

where $y_j^* = y_{j+N_s}$ and the independence holds thanks to the fact that all y_j are independent. Moreover

$$X_{t+h} - X_t = \sum_{j=1}^{N_{t+h} - N_t} \tilde{y}_j \overset{L}{=} \sum_{j=1}^{N_h} y_j = X_h \ ,$$

where $\tilde{y}_j = y_{j+N_t}$. N_t is constant in (T_n, T_{n+1}) and thus also X_t is constant in (T_n, T_{n+1}). Moreover, $\Delta X_{T_n} = X_{T_n} - X_{T_n^-} = y_j$, which is finite a.s. (because y_j has a distribution on $\mathbb{R}$) and $X_{t^+} = X_t$ because $N_{t^+} = N_t$. Then, (X_t) is stochastically continuous (similar argument as the Poisson process) and Cadlag. $X_0 = 0$ a.s because $N_0 = 0$ a.s. and then X_t is Lévy.
We prove now that a Lévy process with the same distribution of the Compound Poisson exists. We compute the characteristic function of the compound Poisson proces

$$\begin{aligned}\mathbb{E}\left[e^{iuX_t}\right] &= \mathbb{E}\left[e^{iu\sum_{j=1}^{N_t} y_j}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{iu\sum_{j=1}^{N_t} y_j}|N_t\right]\right] = \mathbb{E}\left[\hat{f}(u)^{N_t}\right] = \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n (\hat{f}(u))^n}{n!} \\ &= e^{-\lambda t} e^{\lambda t \hat{f}(u)} = e^{\lambda t(\int_{\mathbb{R}} e^{iux} f(x)dx - 1)} = e^{\lambda t(iu\int_{|x|\le 1} xf(x)dx + \int_{\mathbb{R}} e^{iux} - 1 - \mathcal{I}_{|x|\le 1} f(x)dx)} \ ,\end{aligned}$$

where $\hat{f}(u)$ is the characteristic function of f. The second equality is due to the tower property, the third to the independence of the y_j, the fourth to the Poisson process distribution, the fifth to the power series of the exponential, the sixth to the definition of characteristic function. We can now identify that X_t has infinitely divisible distribution with $A_t = 0$, $\nu_t = \lambda t f$ and $\gamma_t = \int_{|x|\le 1} xf(x)dx$. The triplet satisfies the conditions of Sato so an additive process exists and we can check that increments are stationary (or directly that t factorize out of the characteristic function) so X_t is a Lévy process. □

Notice that in the case of the standard Poisson process f is δ_1. You just have unitary jumps.
Given the jump size process $\Delta X_s = X_s - X_{s-}$, the Lévy measure $\nu_t(B)$ of an additive process can be identified with the expected number of jumps with size in the Borel set B occurring between time 0 and t:

$$\nu_t(B) = \mathbb{E}\left[\sum_{0<s\le t} \mathbf{1}_{\{\Delta X_s \in B\}}\right].$$

We verify this interpretation for the Compound Poisson process. Recall that jumps occur at times

determined by a Poisson process N_t with magnitudes y_j.

$$\begin{aligned}\nu_t(B) &= \mathbb{E}\left[\sum_{j=1}^{N_t} \mathbf{1}_{\{y_j \in B\}}\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{j=1}^{N_t} \mathbf{1}_{\{y_j \in B\}} \;\middle|\; N_t\right]\right] \\ &= \mathbb{E}\left[\sum_{j=1}^{N_t} P(y_j \in B)\right] \quad \text{(by i.i.d. property of } y_j\text{)} \\ &= \mathbb{E}[N_t] P(y_1 \in B) \\ &= \lambda t \int_B f(x)dx.\end{aligned}$$

This matches the result derived from the characteristic function, where $\nu_t(dx) = t\lambda f(x)dx$.

Proposition 1.36. *A piecewhise constant process is a Lévy process if and only if it is a Compound Poisson.*

1.6 Finite activity, finite variation and infinite variation

For a compound Poisson process $\nu_t(B)$ is always finite for all $B \in \mathbb{B}(\mathbb{R})$ because $\lambda t \int_B f(x)dx \leq \lambda t$. In general, for a cadlag process (X_t), (X_t) is bounded a.s. on $[0,T]$ and then $\nu_t(B) < \infty \forall B \subset \{x : |x| > \varepsilon\}$ for some $\varepsilon > 0$.
One might ask whether a process could have infinitely many jumps of magnitude greater than ε on a finite interval $[0,T]$ if the positive and negative jumps "compensate" each other (e.g., an infinite sequence of $+C$ and $-C$).
This is impossible for a càdlàg process. Let $D_\varepsilon = \{t \in [0,T] : |\Delta X_t| > \varepsilon\}$ be the set of times with large jumps. If $|D_\varepsilon| = \infty$, by the Bolzano-Weierstrass theorem, these jump times must have an accumulation point $\tau \in [0,T]$.
As $t \to \tau$, the process X_t would oscillate by an amount of at least ε infinitely often. Consequently, the left limit:

$$X_{\tau-} = \lim_{s \uparrow \tau} X_s$$

would not exist. Since X is defined as càdlàg (right-continuous with left limits), this situation is forbidden.
Therefore, for any $\varepsilon > 0$, the number of jumps is almost surely finite if the process is càdlàg.
We distinguish between two regimes for the jump component:

1. **Finite Activity (Compound Poisson):** The process has finitely many jumps in any bounded interval.
$$X_t = \sum_{j=1}^{N_t} \Delta X_{T_j} = \sum_{j=1}^{N_t} Y_j,$$
where N_t is a Poisson process and Y_j are the jump sizes.

2. **Infinite Activity:** The process exhibits infinitely many jumps on $[0,t]$.
$$X_t = \sum_{s \leq t} \Delta X_s.$$

In the infinite activity case, although the number of jumps is infinite, the sum of jumps remains finite because the vast majority of them are arbitrarily small (infinitesimal).

Consider the case of an integral measure. The density $\nu(x)$ must be integrable on any set bounded away from the origin, meaning:

$$\int_{|x|>\varepsilon} \nu(x)\,dx < \infty \quad \text{for all } \varepsilon > 0.$$

However, as $x \to 0$, the density $\nu(x)$ may explode. In the case of **infinite activity** processes, the density is *not* integrable at the origin:
$\int_{-\varepsilon}^{\varepsilon} \nu(x)\,dx = \infty$.
This divergence is valid because the Lévy condition:
$\int_{\mathbb{R}}(1 \wedge x^2)\nu(x)\,dx < \infty$
does not require $\nu(x)$ to be integrable near zero; it only requires that the singularity is strictly weaker than $1/x^3$. Specifically, $x^2\nu(x)$ must be integrable near the origin. This condition is relevant because it ensures that

$$\hat{\mu}(u) = exp\left(-1/2u^2A + iu\gamma + \int_{\mathbb{R}} e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1}\nu(dx)\right)$$

is integrable because near the origin

$$(e^{iux} - 1 - iux\mathcal{I}_{|x|\leq 1})\nu(x) \approx -u^2x^2v(x) \ .$$

Example: Consider the density $\nu(x) = \frac{1}{x^2}$ (associated with the Cauchy process).

- **Activity:** $\int_0^{\varepsilon} \frac{1}{x^2}\,dx = \left[-\frac{1}{x}\right]_0^{\varepsilon} = \infty$ (Infinite Activity).
- **Lévy Condition:** $\int_0^{\varepsilon} x^2 \cdot \frac{1}{x^2}\,dx = \int_0^{\varepsilon} 1\,dx = \varepsilon < \infty$ (Valid Lévy process).

Whitin the infinite activity case we distinguish between the finite variation and the infinite variation case.

Definition 1.37. Let $f : [a, b] \to \mathbb{R}$ be a real-valued function. The total variation of f on the interval $[a, b]$, denoted by $V_a^b(f)$, is defined as:

$$V_a^b(f) = \sup_{\mathcal{P}} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|\,,$$

where the supremum is taken over all finite partitions $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ of the interval $[a, b]$.

1. Infinite activity finite variation processes satisfy $A = 0$ and $\int_{|x|\leq 1} |x|\nu(dx) < \infty$ and $\int_{|x|\leq 1} \nu(dx) = \infty$. Notice that the variation coincide with the sum of absolute values jumps because we have no diffusion part and then $\sum_{s\leq t} |\Delta X_s| < \infty$.
2. Infinite variation, $A \neq 0$ or $\int_{|x|\leq 1} \nu(dx) = \infty$. In case $\int_{|x|\leq 1} |x|\nu(dx) = \infty$ $\sum_{s\leq t} |\Delta X_s| = \infty$.

Corollary 1.38. *The characteristic function of a Lévy process (X_t) with finite variation is given by:*

$$\mathbb{E}\left[e^{iuX_t}\right] = \exp\left(t\left(iub + \int_{\mathbb{R}}(e^{iux} - 1)\nu(dx)\right)\right),$$

where the drift is defined as $b = \gamma - \int_{|x|\leq 1} x\nu(dx)$.

Proof. Recall the general Lévy-Khintchine formula. If (X_t) has finite variation, then the diffusion coefficient $A = 0$ and the Lévy measure satisfies the condition $\int_{|x|\leq 1} |x|\nu(dx) < \infty$.
We start with the standard representation:

$$\mathbb{E}\left[e^{iuX_t}\right] = \exp\left(t\left(iu\gamma + \int_{\mathbb{R}}(e^{iux} - 1 - iux\mathbb{1}_{|x|\leq 1})\nu(dx)\right)\right).$$

Since $\int_{|x|\leq 1} |x|\nu(dx) < \infty$, the term x is integrable near the origin. We can therefore split the integral and group the linear terms:

$$\begin{aligned}\Psi_t(u) &= \exp\left(t\left(iu\left(\gamma - \int_{|x|\leq 1} x\nu(dx)\right) + \int_{\mathbb{R}}(e^{iux} - 1)\nu(dx)\right)\right) \\ &= \exp\left(t\left(iub + \int_{\mathbb{R}}(e^{iux} - 1)\nu(dx)\right)\right).\end{aligned}$$

This concludes the proof. □

Corollary 1.39. *For an infinite activity finite variation Lévy process $X_t = bt + \sum_{s\in 0,t} \Delta X_s$.*

Theorem 1.40. *Let $(X_t)_{t\geq 0}$ be a Lévy process with generating triplet (A, ν, γ). The following conditions are equivalent:*

1. *$X_t \geq 0$ almost surely for a fixed $t > 0$.*
2. *$X_t \geq 0$ almost surely for all $t > 0$.*
3. *The sample paths of X are almost surely non-decreasing (i.e., for all $t \geq s$, $X_t \geq X_s$ a.s.).*
4. *The process is of finite variation ($A = 0$, $\int_{|x|\leq 1} |x|\nu(dx) < \infty$) with:*

$$\begin{aligned}\nu((-\infty, 0]) &= 0, \\ b &\geq 0,\end{aligned}$$

where $b = \gamma - \int_0^1 x\nu(dx)$ is the drift of the finite variation representation.

Proof. We proceed by proving the implications in a logical sequence.
(1) $\implies$ (3): Fix $t > 0$ and $N \in \mathbb{N}$. Consider the increments $Y_k = X_{kt/N} - X_{(k-1)t/N}$ for $k = 1, \ldots, N$. Since X is a Lévy process, these are independent and identically distributed (i.i.d.) random variables, and $X_t = \sum_{k=1}^N Y_k$. Since $X_t \geq 0$ almost surely for any chosen t, the support of the distribution of the increments must be contained in $[0, \infty)$. Consequently, $X_s - X_r \geq 0$ almost surely for any rational time points $s \geq r$. By the right-continuity of paths and the density of $\mathbb{Q}$ in $\mathbb{R}$, we conclude that $X_t \geq X_s$ almost surely for all real $t \geq s$.
(3) $\implies$ (2) $\implies$ (1): If the paths are non-decreasing and start at $X_0 = 0$, then clearly $X_t \geq 0$ for all $t > 0$. This immediately implies condition (1).
(3) $\iff$ (4): We utilize the finite variation decomposition (Corollary 1.39).
Step 4 $\implies$ 3*:* Assume condition (4) holds. Since the process is of finite variation, we can write:

$$X_t = bt + \sum_{0<s\leq t} \Delta X_s.$$

We are given that $b \geq 0$ and $\nu((-\infty, 0]) = 0$. Since the Lévy measure does not charge negative values (and by definition $\nu(\{0\}) = 0$), all jumps must be strictly positive ($\Delta X_s > 0$ a.s.). Thus, for any $t \geq s$:

$$X_t - X_s = b(t - s) + \sum_{s<r\leq t} \Delta X_r.$$

Since both terms on the right-hand side are non-negative, $X_t - X_s \geq 0$ almost surely.
Step 3 $\implies$ 4: If the trajectories are almost surely non-decreasing, the process must be of finite variation (monotonic functions have finite variation). Furthermore, a non-decreasing process cannot have negative jumps, so the measure must satisfy $\nu((-\infty, 0]) = 0$. Finally, consider the small-increment behavior. If $b < 0$, then for sufficiently small $t - s$, $X_t - X_s$ would be negative. Therefore, we must have $b \geq 0$. □

1.7 Jump Diffusion Processes

Proposition 1.41. *Let $(X_t)_{t\geq 0}$ be a stochastic process with independent increments. If $\mathbb{E}[e^{X_t}] < \infty$ for all $t \geq 0$, then the process defined by*

$$Y_t = \frac{e^{X_t}}{\mathbb{E}[e^{X_t}]}$$

is a martingale with respect to its natural filtration $\mathcal{F}_t$.

Proof. First, observe that Y_t is $\mathcal{F}_t$-measurable by definition and integrable by hypothesis. To prove the martingale property, let $s < t$. We compute the conditional expectation:

$$\mathbb{E}[Y_t \mid \mathcal{F}_s] = \mathbb{E}\left[\frac{e^{X_t}}{\mathbb{E}[e^{X_t}]}\middle| \mathcal{F}_s\right] = \frac{1}{\mathbb{E}[e^{X_t}]}\mathbb{E}\left[e^{X_t - X_s}e^{X_s} \mid \mathcal{F}_s\right].$$

Since e^{X_s} is $\mathcal{F}_s$-measurable, we can pull it out of the conditional expectation. Furthermore, since the increment $X_t - X_s$ is independent of $\mathcal{F}_s$, the conditional expectation becomes the unconditional expectation:

$$\mathbb{E}[Y_t \mid \mathcal{F}_s] = \frac{e^{X_s}}{\mathbb{E}[e^{X_t}]}\mathbb{E}\left[e^{X_t - X_s}\right].$$

Now, using the property of independent increments, we can factorize the expectation in the denominator as $\mathbb{E}[e^{X_t}] = \mathbb{E}[e^{X_t - X_s + X_s}] = \mathbb{E}[e^{X_t - X_s}]\mathbb{E}[e^{X_s}]$. Substituting this back into the equation gives:

$$\mathbb{E}[Y_t \mid \mathcal{F}_s] = \frac{e^{X_s}\mathbb{E}[e^{X_t - X_s}]}{\mathbb{E}[e^{X_t - X_s}]\mathbb{E}[e^{X_s}]} = \frac{e^{X_s}}{\mathbb{E}[e^{X_s}]} = Y_s.$$

Thus, (Y_t) is a martingale. □

Remark 1.42. Let $\phi_t(u) = \mathbb{E}[e^{iuX_t}]$ be the characteristic function of X_t. Notice that $\mathbb{E}[e^{X_t}] = \phi_t(-i)$. Consequently, if the characteristic function satisfies condition $\phi_t(-i) = 1$ for all t (i.e., the exponential moment is normalized), then the process e^{X_t} itself is an exponential martingale.

We model the forward price with maturity T as

$$F(t,T) = F(0,T)\frac{e^{X_t}}{\mathbb{E}[e^{X_t}]} \quad .$$

we can use this to price European options because $F(T,T) = S_T$

Definition 1.43. (X_t) is called a **jump-diffusion process** :

$$X_t = bt + \sigma W_t + \sum_{j=1}^{N_t} y_j,$$

where:

- b and σ are the drift and diffusion coefficients.
- (W_t) is a standard Brownian motion.
- (N_t) Poisson process and Y_j r.v.

All r.v. are independents among each others. This is a Lévy process because it is a sum of independent Lévy processes.

Merton Model: The jump sizes are distributed as $y_j \sim \mathcal{N}(\mu, \delta^2)$, where μ and δ are the mean and standard deviation of the jumps. In total, there are 4 parameters, as the drift parameter b is constrained so that the forward price is a martingale ($\mathbb{E}[e^{X_t}] = \hat{\mu}(-i) = 1$).
The Lévy density is $\nu(x) = \lambda f(x)$, where $f(x)$ is the Gaussian density:

$$f(x) = \frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(x-\mu)^2}{2\delta^2}} .$$

The process comprises three independent parts (drift, diffusion, and jumps), so the characteristic function is:

$$\hat{\mu}(u) = e^{iubt}\mathbb{E}[e^{iu\sigma W_t}]\mathbb{E}\left[e^{iu\sum_{j=1}^{N_t} y_j}\right] = \exp\left\{t\left(iub - \frac{\sigma^2 u^2}{2} + \lambda\int_{\mathbb{R}}(e^{iux} - 1)f(x)dx\right)\right\} .$$

By computing the integral (identifying the characteristic function of the Normal distribution), we get:

$$\hat{\mu}(u) = \exp\left\{t\left(iub - \frac{\sigma^2 u^2}{2} + \lambda\left(e^{iu\mu - \delta^2 u^2/2} - 1\right)\right)\right\} .$$

Finally, setting the martingale condition $\hat{\mu}(-i) = 1$, we obtain:

$$b = -\frac{1}{2}\sigma^2 - \lambda\left(e^{\mu + \delta^2/2} - 1\right) .$$

Kou Model: The jump measure is given by:

$$\nu(x) = \lambda f(x) = \lambda\left(p\lambda_+ e^{-\lambda_+ x}\mathcal{I}_{x>0} + (1-p)\lambda_- e^{\lambda_- x}\mathcal{I}_{x<0}\right) ,$$

where $p \in [0,1]$ and $\lambda_+ > 0, \lambda_- > 0$. Notice that if $\lambda^+ <= 1$ the expecxted value of e^{X_t} is infinite. The Poisson intensity is λ, and the law of jumps is a mixture of positive and negative exponential distributions.
Computing the positive and negative integrals at the exponent of the characteristic function separately:

$$\int_0^\infty (e^{iux} - 1)\lambda_+ e^{-\lambda_+ x}dx = \frac{\lambda_+}{\lambda_+ - iu} - 1 = \frac{iu}{\lambda_+ - iu} ,$$
$$\int_{-\infty}^0 (e^{iux} - 1)\lambda_- e^{\lambda_- x}dx = \frac{\lambda_-}{\lambda_- + iu} - 1 = \frac{-iu}{\lambda_- + iu} .$$

Substituting these into the characteristic function , we get:

$$\hat{\mu}(u) = \exp\left\{t\left(iub - \frac{\sigma^2 u^2}{2} + \lambda\left[\frac{piu}{\lambda_+ - iu} - \frac{(1-p)iu}{\lambda_- + iu}\right]\right)\right\} .$$

Thus, the drift parameter is:

$$b = -\frac{1}{2}\sigma^2 + \lambda\left(\frac{p}{\lambda_+ - 1} - \frac{1-p}{\lambda_- + 1}\right) .$$

MC Simulation of Jump Diffusions: The simulation method is exact because we sample directly from the distribution of the process, rather than approximating the stochastic differential equation (there is no discretization bias). The simulation is divided into 3 independent steps:

1. Simulate the Brownian Motion (Diffusion component).
2. Simulate the jump times (Poisson process).
3. Simulate the jump sizes (conditional on a jump occurring) with law $f(x) = \nu(x)/\lambda$.

Recall that the counting process is defined as:

$$N_t = \sum_{n \geq 1} \mathcal{I}_{\{t \geq T_n\}} \sim \text{Poisson}(\lambda t)$$

where the arrival times are $T_n = \sum_{j=1}^{n} \tau_j$, and the inter-arrival times are i.i.d. with $\tau_j \sim \text{Exp}(\lambda)$. We can simulate the jump times using one of two methods:
Method 1 This method generates jumps sequentially until the time horizon t is exceeded.

- Initialize $T_0 = 0$, $j = 0$.
- Loop:
  - Generate $\tau_{j+1} \sim \text{Exp}(\lambda)$.
  - Set $T_{j+1} = T_j + \tau_{j+1}$.
  - If $T_{j+1} > t$, stop.
  - Else, increment $j = j + 1$ and repeat.
- The jump times are $\{T_1, \ldots, T_j\}$ and the total number of jumps is $N_t = j$.

Method 2 This method utilizes the property that, conditioned on $N_t = n$, the arrival times are uniformly distributed.

- Simulate the total number of jumps: $n \sim \text{Poisson}(\lambda t)$.
- Generate n independent uniform random variables: $U_1, \ldots, U_n \sim \mathcal{U}[0, 1]$.
- The jump times are obtained by sorting and scaling:

$$\{T_1, \ldots, T_n\} = t \cdot \text{sort}(\{U_1, \ldots, U_n\}) \ .$$

Once the jump times are determined, we simulate the jump sizes y_j from the density $f(x)$. For computational efficiency, it is preferable to simulate all jump sizes y_j simultaneously (vectorization). However, a technical challenge arises because the number of jumps N_T is random; consequently, each simulation path contains a different number of jumps, resulting in "jagged" data structures that are difficult to handle in standard matrix operations.
To simulate the jump-diffusion process on a discrete time grid $0 = t_0 < t_1 < \cdots < t_N$ with step size Δt:

1. Diffusion Component: Simulate the independent Brownian increments for each step i:

$$G_i = W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, \Delta t) \ .$$

2. Jump Component: Simulate the jump times T_j and the corresponding jump sizes y_j using Method 1 or 2 described previously.

3. Path Construction: The value of the log-process X at time t_i is the sum of the drift, the accumulated diffusion, and the accumulated jumps up to that time:

$$X_{t_i} = bt_i + \sigma \sum_{k=1}^{i} G_k + \sum_{j=1}^{N_{t_i}} y_j$$

where $N_{t_i} = \sum_{n\geq 1} \mathbf{1}_{\{T_n \leq t_i\}}$ is the number of jumps that have occurred by time t_i.

This formula can be vectorialized.

1.8 Levy subordination and the LTS

Theorem 1.44. *Let (X_t) be a Lévy process on $\mathbb{R}$ with triplet (A, ν, γ) and S_t be a Lévy subordinator with jump measure ρ and drift b. Then,*

1. *$(y_t) = (X_{S_t})$ is a Lévy process with characteristic function $L_t(-\psi(u))$ where $L_t(s) = \mathbb{E}[e^{-sS_t}]$ is the Laplace transform fo S_t and $\psi(u)$ is the characteristic exponent of X_t.*

2. *The triplet of y_t is*

$$\begin{aligned} A^y &= bA \\ \nu^y(B) &= b\nu(B) + \int_0^\infty P_{X_s}(B)\rho(ds) \\ \gamma^y &= b\gamma + \int_0^\infty \rho(ds) \int_{|x|\leq 1} x P_{X_s}(dx) \quad , \end{aligned}$$

where $B \in \mathcal{B}(\mathbb{R})$ and $P_{X_s}(B) = P(X_s \in B)$.

Proof. We prove only point 1. We prove derive the characteristic function of y_t

$$\mathbb{E}[e^{iuy_t}] = \mathbb{E}[e^{iuX_{S_t}}] = \mathbb{E}[\mathbb{E}[e^{iuX_{S_t}}|\mathcal{F}_t^S]] = \mathbb{E}[e^{S_t\psi(u)}] = L_t[-\psi(u)] \quad ,$$

where $\mathcal{F}_t^S$ is the filtration of S_t. The second equality is due to the tower property the third to the characteristic function of a Lévy process and the last to the definition of Laplace transform. We verify the definition of Lévy process. $X_0 = 0$ a.s. and $S_0 = 0$ a.s. ensures that $X_{S_0} = 0$ a.s and composition of cadlag processes is cadlag. We prove independence of increments by showing that the characteristic function of a random vector of increments factorizes:

$$\begin{aligned} \mathbb{E}[e^{iu_1(y_t - y_r) + iu_2 y_r}] &= \mathbb{E}[e^{iu_1(X_{S_t} - X_{S_r}) + iu_2 X_{S_r}}] = \mathbb{E}[\mathbb{E}[e^{iu_1(X_{S_t} - X_{S_r}) + iu_2 X_{S_r}}|\mathcal{F}_t^S]] \\ &= \mathbb{E}[\mathbb{E}[e^{iu_1(X_{S_t} - X_{S_r})}|\mathcal{F}_t^S]\mathbb{E}[e^{iu_2 X_{S_r}}|\mathcal{F}_t^S]] = \mathbb{E}[e^{(S_t - S_r)\psi(u_1)} e^{S_r\psi(u_2)}] \\ &= \mathbb{E}[e^{(S_t - S_r)\psi(u_1)}]\mathbb{E}[e^{S_r\psi(u_2)}] = L_{t-r}(-\psi(u_1))L_r(-\psi(u_2)) = \mathbb{E}[e^{iu_1(y_t - y_r)}]\mathbb{E}[e^{iu_2 y_r}] \quad , \end{aligned}$$

where the second equality is due to the towe property the third to the independence of increments of X_t the fourth to the characteristic function of a Lévy process the fifth to the independence of increments of S_t the sixth to the stationarity of increments of S_t and the definition of Laplace transform. The last equality that proves the factorizazion is due to similar steps to the single increments.
We prove stationarity of increments.

$$\mathbb{E}[e^{iu(y_{t+h} - y_t)}] = \mathbb{E}[\mathbb{E}[e^{iu(X_{S_{t+h}} - y_{X_t})}|\mathcal{F}_{t+h}^S] = \mathbb{E}[e^{(S_{t+h} - S_t)\psi(u)}] = L_h(-\psi(u)) \quad ,$$

where the third equatility is due to the characteristic function of a Lévy process increment and the last to the stationarity of S_t and the definition of Laplace transform. We prove stochastic continuity. We use the fact that X_t is uniformly stochastically continuous in probability i.e. fix $\epsilon > 0$ and $q > 0$. $\forall t$ it exists $\bar{h}$ (the same $\bar{h}$ for all t) such that $P(|X_{t+h} - X_t| > \epsilon) < q \; \forall h < \bar{h}$. This comes from the fact that the increments of X_t are stationary and X_t is stochastically continous. Them by the total probability theorem fixing $\delta > 0$

$$\begin{aligned}
&P(|X_{S_t} - X_{S_s}| > \epsilon) \\
&= P(|X_{S_t} - X_{S_s}| > \epsilon || S_t - S_s| < \delta) P(|S_t - S_s| < \delta) \\
&+ P(|X_{S_t} - X_{S_s}| > \epsilon || S_t - S_s| \geq \delta) P(|S_t - S_s| \geq \delta) \\
&\leq P(|X_{S_t} - X_{S_s}| > \epsilon || S_t - S_s| < \delta) + P(|S_t - S_s| \geq \delta) \;\; .
\end{aligned}$$

To prove stochastic continuity we have to prove that $P(|X_{S_t} - X_{S_s}| < \epsilon)$ goes to zero for s that goes to t. We show it by proving that for any $\epsilon > 0$ and $q > 0$ it exists $\bar{h}$ such that $P(|X_{S_t} - X_{S_s}| > \epsilon) < q$ for all $|t - s| < \bar{h}$. In particular, $P(|X_{S_t} - X_{S_s}| > \epsilon || S_t - S_s| < \delta)$ for any s, t it exists by uniformly stochastic continuity $\bar{\delta}$ such that $P(|X_{S_t} - X_{S_s}| < \epsilon || S_t - S_s| < \delta) < q/2$ for all $\delta \leq \bar{\delta}$. Now, by stochastic continuity of S_t it exists $\bar{h}$ such that $P(|S_t - X_s| \leq \bar{\delta}) < q/2$ for every $t - s \leq \bar{h}$ and then we have that

$$P(|X_{S_t} - X_{S_s}| > \epsilon) < q$$

for every $t - s \leq \bar{h}$ which proves stochastic continuity. □

Stable Subordinators: These are subordinators characterized by the Lévy measure:

$$\rho(x) = \frac{C}{x^{1+\alpha}} \mathcal{I}_{\{x>0\}} \; ,$$

where $C > 0$ and $\alpha \in (0, 1)$.
A significant drawback of strictly stable subordinators is that their moments are infinite for orders $p \geq \alpha$. Consequently, the process has infinite mean and infinite variance. To overcome this limitation while preserving the local behavior of stable processes (explosion of the measure for small jumps), Tempered Stable Subordinators (TSS) are introduced.

Tempered Stable Subordinators (TSS): We consider a process with drift $b = 0$. The Lévy measure is given by:

$$\rho(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{\kappa} \right)^{1-\alpha} \frac{e^{-(1-\alpha)x/\kappa}}{x^{1+\alpha}} \mathcal{I}_{\{x>0\}} \; ,$$

where $\kappa > 0$ and $\alpha \in [0, 1)$.
A subordinator with these characteristics is well-defined because there are no negative jumps, the diffusion component is zero, and the Lévy measure satisfies the integrability condition for subordinators. Specifically, we verify that $\int_{\mathbb{R}^+} (x \wedge 1) \rho(dx) < \infty$:

$$\int_{\mathbb{R}^+} (x \wedge 1) \rho(dx) \leq C_1 \int_0^1 \frac{x}{x^{1+\alpha}} dx + C_2 \int_1^\infty e^{-(1-\alpha)x/\kappa} dx < \infty \; ,$$

where $C_1, C_2 > 0$. Since $\alpha < 1$, the integral near zero converges (since $x^{-\alpha}$ is integrable), and the exponential term ensures convergence at infinity.
Thus, there exists an infinitely divisible distribution μ with Lévy density $\rho(x)$, no diffusion, and drift $b = 0$. Consequently, there exists a Lévy process associated with this triplet. Note that this process is a *subordinator* because it is of finite variation (since $\alpha < 1$) and has no negative jumps. All TSS are of infinite activity and finite variation.
The two most famous cases are:

- Gamma process ($\alpha = 0$):

$$\rho(x) = \frac{1}{\kappa}\frac{e^{-x/\kappa}}{x}\mathcal{I}_{\{x>0\}}$$

  associated with the Gamma distribution.

- Inverse Gaussian (IG) process ($\alpha = 1/2$):

$$\rho(x) = \frac{1}{\sqrt{2\pi}\kappa}\frac{e^{-x/(2\kappa)}}{x^{3/2}}\mathcal{I}_{\{x>0\}}$$

  associated with the Inverse Gaussian distribution.

The Laplace transform of the TSS, defined as $L_t(s) = \mathbb{E}[e^{-sS_t}]$, is given by:

$$L_t(s) = \exp\left(t\frac{1-\alpha}{\kappa\alpha}\left(1 - \left(1 + \frac{\kappa s}{1-\alpha}\right)^{\alpha}\right)\right)$$

if $\alpha \in (0, 1)$, and

$$L_t(s) = (1 + \kappa s)^{-t/\kappa}$$

if $\alpha = 0$. Note that for $\alpha = 0$, this corresponds to the Laplace transform of a $\Gamma(t/\kappa, \kappa)$ random variable. This parameterization is chosen specifically to ensure that $\mathbb{E}[S_t] = t$ and $\mathrm{Var}[S_t] = \kappa t$. We can verify this for the Gamma case ($\alpha = 0$): recalling that $b = \gamma - \int_{|x|\geq 1}\rho(dx)$

$$\begin{aligned}
\mathbb{E}[S_t] &= t\left(\gamma + \int_{|x|\geq 1}^{\infty} x\rho(dx)\right) = t\int_0^{\infty} x\rho(x)dx \\
&= t\int_0^{\infty} x\left(\frac{1}{\kappa}\frac{e^{-x/\kappa}}{x}\right)dx = \frac{t}{\kappa}\int_0^{\infty} e^{-x/\kappa}dx = t\ , \\
\mathrm{Var}[S_t] &= t\int_0^{\infty} x^2\rho(x)dx \\
&= t\int_0^{\infty} x^2\left(\frac{1}{\kappa}\frac{e^{-x/\kappa}}{x}\right)dx = \frac{t}{\kappa}\int_0^{\infty} xe^{-x/\kappa}dx \\
&= \frac{t}{\kappa}\left(\left[-\kappa x e^{-x/\kappa}\right]_0^{\infty} + \kappa\int_0^{\infty} e^{-x/\kappa}dx\right) = \kappa t\ .
\end{aligned}$$

The Normal Tempered Stable (NTS or LTS) Process: Let $X_t = \mu t + \sigma W_t$ be a Brownian motion with drift, where W_t is a standard Brownian motion, $\sigma \in \mathbb{R}^+$, and $\mu \in \mathbb{R}$. Let (S_t) be a Tempered Stable Subordinator (TSS). The Normal Tempered Stable (LTS) process, is defined by the subordination $Y_t = X_{S_t}$.
By the Theorem of Subordination, Y_t is a Lévy process with characteristic function :

$$\hat{\mu}_{Y_t}(u) = L_t\left(-iu\mu + \frac{u^2\sigma^2}{2}\right) = \begin{cases} \exp\left(t\frac{1-\alpha}{\kappa\alpha}\left(1 - \left(1 + \frac{\kappa\left(-iu\mu + \frac{u^2\sigma^2}{2}\right)}{1-\alpha}\right)^{\alpha}\right)\right) & \text{if } \alpha \in (0,1) \\ \left(1 + \kappa\left(-iu\mu + \frac{u^2\sigma^2}{2}\right)\right)^{-t/\kappa} & \text{if } \alpha = 0 \end{cases}.$$

We introduce an alternative parameterization: $\mu = -\sigma^2(\eta + 1/2)$. If $\eta = 0$, the smile has no skew and is symmetric.
For pricing purposes, the forward price process is a martingale if we consider the process $e^{Y_t+\varphi t}$, where the drift compensator φ is chosen such that:

$$\mathbb{E}[e^{Y_t+\varphi t}] = \hat{\mu}_{Y_t}(-i)e^{\varphi t} = L_t\left(-\mu - \frac{\sigma^2}{2}\right)e^{\varphi t} = 1\ .$$

Substituting the parameterization for μ, the argument of the Laplace transform simplifies to $\sigma^2\eta$. Thus, we require $L_t(\sigma^2\eta)e^{\varphi t} = 1$. Solving for φ:

$$\varphi = -\frac{1}{t}\log\left(L_t(\sigma^2\eta)\right) \ .$$

The Lévy Triplet and Subcases: The Lévy triplet of the subordinated process Y can be computed using the Theorem of Subordination. Notice that the diffusion component $A_Y = 0$, meaning this is a pure jump process. The Lévy density $\nu_Y(x)$ is given by mixing the Gaussian density with the Lévy measure of the subordinator:

$$\nu_Y(x) = \int_0^\infty P_{X_s}(x)\rho(ds) = \int_0^\infty \frac{1}{\sqrt{2\pi s}\sigma}\exp\left(-\frac{(x-\mu s)^2}{2\sigma^2 s}\right)\left(\frac{1-\alpha}{\kappa}\right)^{1-\alpha}\frac{e^{-(1-\alpha)s/\kappa}}{\Gamma(1-\alpha)s^{1+\alpha}}ds \ ,$$

where $P_{X_s}(x)$ denotes the density of the Brownian motion X_s. This simplifies to:

$$\nu_Y(x) = C_1\int_0^\infty \frac{\exp\left(-\frac{(x-\mu s)^2}{2\sigma^2 s} - \frac{(1-\alpha)s}{\kappa}\right)}{s^{3/2+\alpha}}ds \ .$$

The Lévy density $\nu_Y(x)$ is integrable for $x \neq 0$. Its asymptotic behavior is characterized as follows:

1. As $x \to 0$: $\nu_Y(x) \approx \frac{1}{|x|^{2\alpha+1}}$
2. As $x \to \infty$: $\nu_Y(x) \approx \frac{e^{-\lambda_+ x}}{x^{\alpha+1}}$
3. As $x \to -\infty$: $\nu_Y(x) \approx \frac{e^{\lambda_- x}}{|x|^{\alpha+1}}$

where the decay rates are given by $\lambda_\pm = \sqrt{\frac{\mu^2}{\sigma^4} + \frac{2}{\kappa\sigma^2(1-\alpha)}} \mp \frac{\mu}{\sigma^2}$.
Consequently, if $\eta \gg 0$ (implying negative skew), then $\mu \ll 0$. This results in $\lambda_+ \gg \lambda_-$, meaning negative jumps decay slower than positive jumps, justifying the negative skew (fatter left tail).
Regarding the path properties:

- For $\alpha < 1/2$, the process is of finite variation.
- For $\alpha \geq 1/2$, the process is of infinite variation.
- All Normal Tempered Stable (LTS) processes are of infinite activity.

Two relevant subcases are:

1. The Variance Gamma (VG) ($\alpha = 0$):

$$\hat{\mu}_Y(u) = \left(1 + \kappa\left(-iu\mu + \frac{u^2\sigma^2}{2}\right)\right)^{-t/\kappa} \ .$$

   The PDF is available in closed form (in terms of Bessel functions), while the CDF is available only in integral form.

2. The Normal Inverse Gaussian (NIG) ($\alpha = 1/2$): Substituting $\alpha = 1/2$ into the general NTS characteristic function, we obtain:

$$\hat{\mu}_Y(u) = \exp\left(\frac{t}{\kappa}\left(1 - \sqrt{1 + 2\kappa\left(-iu\mu + \frac{u^2\sigma^2}{2}\right)}\right)\right) \ .$$

   The PDF is available in closed form, while the CDF is available only in integral form.

1.9 The CGMY Process (Tempered Stable Processes)

The CGMY process is a pure jump process defined by the Lévy density:

$$\nu(x) = \frac{C}{x^{1+\alpha}} e^{-\lambda_+ x} \mathcal{I}_{\{x>0\}} + \frac{C}{|x|^{1+\alpha}} e^{\lambda_- x} \mathcal{I}_{\{x<0\}} \ ,$$

where $C \in \mathbb{R}^+$ is the jump intensity, and $\lambda_+ > 0$ and $\lambda_- > 0$ govern the decay of the tails for positive and negative jumps, respectively. The parameter $\alpha < 2$ determines the path properties (activity and variation) of the process.
The acronym CGMY is derived from the names of the authors: Carr, Geman, Madan, and Yor.
The Characteristic Function, $\phi_t(u) = \mathbb{E}[e^{iuX_t}]$, for the case where $\alpha \neq 0$ and $\alpha \neq 1$, is given by:

$$\hat{\mu}_t(u) = \exp\left(tC\Gamma(-\alpha)\left[(\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha\right]\right) \ .$$

The probability density function of the process is not available in closed form.
Constraints on α:

- If $\alpha < 0$: Finite activity (Compound Poisson-like).
- If $0 \leq \alpha < 1$: Infinite activity, finite variation.
- If $1 \leq \alpha < 2$: Infinite activity, infinite variation.
- If $\alpha \geq 2$: The condition for an infinitely divisible distribution is violated because the singularity at zero becomes non-integrable against x^2. Specifically, $\int_{-1}^{1} x^2 \nu(dx) = \infty$.

Note that a more general parameterization exists allowing for asymmetric intensities (C_+ and C_-) and stability indices (α_+ and α_-). However, these are less relevant in practice since the implied volatility smile is primarily influenced by the large jumps (tails) rather than the fine structure of small jumps.

1.10 Exact simulation of VG and NIG

To simulate the Variance Gamma (VG) and Normal Inverse Gaussian (NIG) processes exactly, we utilize the property of subordination. Since the processes are defined as Brownian Motion with drift evaluated at a random time S_t (the subordinator), we can simulate Y_t by first simulating the increment of the subordinator S_t.
We focus here on the **Gamma process** (the subordinator for VG).
The Gamma subordinator S_t follows a Gamma distribution $\Gamma(\kappa, t/\kappa)$, where t/κ is the shape parameter and κ is the scale parameter. This is derived from the Laplace transform $L_t(s) = (1 + \kappa s)^{-t/\kappa}$ and the characteristic function $\hat{\mu}(u) = (1 - iu\kappa)^{-t/\kappa}$.
By the scaling property of the Gamma distribution, if $Z \sim \Gamma(1, \alpha)$, then $\kappa Z \sim \Gamma(\kappa, \alpha)$. Therefore, to simulate S_t, we only need to generate a random variable Z from a standard Gamma distribution with shape parameter $c = t/\kappa$ and scale 1, and then return κZ. The density of Z is:

$$f_Z(x) = \frac{x^{c-1} e^{-x}}{\Gamma(c)} \mathcal{I}_{\{x>0\}} \ .$$

Depending on the value of the shape parameter $c = t/\kappa$, we use different algorithms.

Case 1: $c < 1$ (Johnk's Algorithm) For a shape parameter less than 1, we use Johnk's rejection sampler.

1. **DO**
2. Generate U and V independently from $\mathcal{U}(0,1)$.
3. Set $X = U^{1/c}$ and $Y = V^{1/(1-c)}$.
4. **WHILE** $X + Y > 1$.
5. Generate $E \sim \text{Exp}(1)$ (Standard Exponential).
6. Return $Z = \frac{X}{X+Y}E$.

Case 2: $c > 1$ (Best's Algorithm) For a shape parameter greater than 1, we use Best's Algorithm (1978). Let $b = c - 1$.

1. Set constants: $b = c - 1$ and $H = \sqrt{3c - 0.75}$.
2. **DO**
3. Generate U and V independently from $\mathcal{U}(0,1)$.
4. Set $W = U(1-U)$ and $Y = H\frac{U-0.5}{W}$.
5. Set $X = b + Y$.
6. **IF** $X < 0$, restart loop (Go to step 2).
7. Set $Z = 64W^3V^2$.
8. **WHILE** $\ln(Z) > 2(b\ln(X/b) - Y)$.
9. Return X.

Note: If $c = 1$, the distribution is simply Exponential with mean 1, which can be simulated as $-\ln(U)$.
Thus, we can simulate Z using one of the two algorithms described above; the realization of the Gamma subordinator is then $S_t = \kappa Z$.
For the **Inverse Gaussian (IG) subordinator** (used in the NIG process), we can leverage an exact simulation algorithm developed for the IG law. The variable S_t follows an Inverse Gaussian distribution $IG(\mu, \lambda)$ with density:

$$f_{IG}(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right) \mathcal{I}_{\{x>0\}} .$$

To ensure the subordinator satisfies $\mathbb{E}[S_t] = t$ and $\text{Var}[S_t] = \kappa t$, the parameters are defined as:

$$\mu = t \quad \text{and} \quad \lambda = \frac{t^2}{\kappa} .$$

We use the **Michael-Schucany-Haas (1976)** algorithm to simulate $X \sim IG(\mu, \lambda)$:

1. Generate $\nu \sim \mathcal{N}(0,1)$ (Standard Normal).
2. Set $Y = \nu^2$ (Chi-squared with 1 degree of freedom).

3. Calculate the first root:

$$X_1 = \mu + \frac{\mu^2 Y}{2\lambda} - \frac{\mu}{2\lambda}\sqrt{4\mu\lambda Y + \mu^2 Y^2} \ .$$

4. Generate $U \sim \mathcal{U}(0,1)$ (Uniform).

5. **IF** $U \leq \frac{\mu}{\mu + X_1}$:
   - Return $S_t = X_1$.

6. **ELSE**:
   - Return $S_t = \frac{\mu^2}{X_1}$.

1.11 Simulation of NIG/VG Processes

We can simulate Normal Inverse Gaussian (NIG) or Variance Gamma (VG) processes using the structure of subordination.

1. Exact Simulation at a fixed time t

1. Simulate the subordinator S_t (using the IG or Gamma algorithms described previously).
2. Simulate a standard normal variable $Z \sim \mathcal{N}(0,1)$.
3. Compute the process value:
$$Y_t = \mu S_t + \sigma\sqrt{S_t} Z \ .$$

2. Path Simulation on a Time Grid To simulate the process on a grid $t_0, t_1, \ldots, t_n$ with step size Δt:

1. For each step j, simulate the independent increment of the subordinator ΔS_j, which follows the law of the subordinator for time interval Δt (IG or Gamma).
2. Simulate the increment of the process:
$$\Delta Y_j = \mu \Delta S_j + \sigma\sqrt{\Delta S_j} Z_j \ ,$$
where $Z_j \sim \mathcal{N}(0,1)$ are independent.
3. Accumulate the increments: $Y_{t_i} = \sum_{j=1}^{i} \Delta Y_j$.

The Acceptance-Rejection Principle The algorithms used to simulate the Gamma subordinator (Johnk's and Best's) and partialli also the IG are based on the Acceptance-Rejection (A-R) method.
In general, we want to simulate a random variable X with a PDF $f(x)$ and CDF $F(x)$ (usually $F(x)$ is difficult to compute otherwhise we would simulated by CDF inversion) using a simpler auxiliary random variable Y with PDF $g(x)$ and CDF $G(x)$. Usually, we are able to simulate fast Y but are unable to do the same for X. We require a constant $c \geq 1$ such that $f(x) \leq cg(x)$ for all x (the "envelope" condition). The algorithm is:

1. Generate $Y \sim g(x)$.
2. Generate $U \sim \mathcal{U}[0,1]$ independent of Y.

3. If $U \leq \frac{f(Y)}{cg(Y)}$, set $X = Y$. Otherwise, reject Y and repeat from Step 1.

Link to previous algorithms:

- Johnk's Algorithm: This is a classic A-R implementation. The condition `WHILE` $X + Y > 1$ corresponds to the rejection step. The pair is accepted only when it falls inside the required region.
- Best's Algorithm: This uses a proposal density $g(x)$ with heavy tails (derived from a t-distribution) to efficiently envelope the Gamma density. The inequality $\ln(Z) > 2(b\ln(X/b) - Y)$, is simply the computational optimization of the condition $U \leq f(Y)/cg(Y)$.

Proposition 1.45. *The acceptance rejection method is exact.*

Proof. We prove that $P(Y \leq s | U \leq f(Y)/(cg(Y))) = F(s)$. Let us call $A := \{Y \leq s\}$, $B := \{U \leq f(Y)/(cg(Y))\}$.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad .$$

$P(A) = P(Y \leq s) = G(s)$, using the definition of conditional probability

$$P(B) = \int_{\mathbb{R}} P(B|Y = w)g(w)dw = \int_{\mathbb{R}} P(U \leq f(w)/(cg(w))g(w)dw = 1/c \ ,$$

and

$$P(B|A) = \frac{P(B,A)}{P(A)} = \frac{\int_{-\infty}^{s} P(U \leq f(w)/(cg(w)))g(w)dw}{G(s)} = \frac{F(s)}{cG(s)} \quad .$$

By combining the elements we get the thesis. □

1.12 Simulating infinite activity Lévy processes

Theorem 1.46 (Lévy Ito Decomposition)**.** *Let (X_t) be a Lévy process on $\mathbb{R}$. The process can be decomposed as a sum of independent Lévy processes:*

$$X_t = \gamma t + \sqrt{A}W_t + \bar{N}_t + \lim_{\epsilon \downarrow 0} \tilde{N}_t^{\epsilon},$$

where:

- *γ is the drift parameter.*
- *W_t is a standard Brownian motion.*
- *$\bar{N}_t$ is a compound Poisson process representing large jumps ($|x| \geq 1$), with intensity $\bar{\lambda} = \nu(\{x : |x| \geq 1\})$ and the jumps distribution is $\frac{1}{\bar{\lambda}}\mathbf{1}_{\{|x|\geq 1\}}\nu(dx)$.*
- *$\tilde{N}_t^{\epsilon}$ is a compensated compound Poisson process representing small jumps, defined as:*

$$\tilde{N}_t^{\epsilon} = \sum_{j=1}^{M_t^{\epsilon}} y_j^{\epsilon} - t \int_{\epsilon<|x|<1} x\, \nu(dx).$$

Here, M_t^{ϵ} is a Poisson process with intensity $\lambda^{\epsilon} = \nu(\{x : \epsilon < |x| < 1\})$ and the jumps distribution is $\frac{1}{\lambda^{\epsilon}}\mathbf{1}_{\{\epsilon<|x|<1\}}\nu(dx)$.

Idea of Proof. The terms in the decomposition are independent by construction. Therefore, the characteristic function of X_t is the product of the characteristic functions of its components. Since the characteristic functions for Brownian motion and Compound Poisson processes are known, we sum their exponents directly:

$$\ln \mathbb{E}[e^{iuX_t}] = \underbrace{i\gamma u t}_{\text{Drift}} \underbrace{-\frac{1}{2}Au^2 t}_{\text{Diffusion}} + \underbrace{t\int_{|x|\geq 1}(e^{iux}-1)\nu(dx)}_{\text{Large Jumps}} + \underbrace{t\lim_{\epsilon\downarrow 0}\int_{\epsilon<|x|<1}(e^{iux}-1-iux)\nu(dx)}_{\text{Compensated Small Jumps}}.$$

Note that the compensator term $-iux$ in the small jumps integral arises from the centering of $\tilde{N}_t^\epsilon$. Combining the integrals over the domains $\{|x| \geq 1\}$ and $\{|x| < 1\}$, we recover the characteristic exponent from the Lévy-Khintchine formula:

$$\Psi_t(u) = t\left(i\gamma u - \frac{1}{2}Au^2 + \int_{\mathbb{R}}\left(e^{iux} - 1 - iux\mathbf{1}_{|x|<1}\right)\nu(dx)\right).$$

Since the characteristic function uniquely determines the distribution, the decomposition holds. □

For the small jumps (where $|x| < 1$), the Lévy measure ν can have a singularity at zero such that $\int_{|x|<1}|x|\nu(dx) = \infty$. Consequently, the direct sum of small jumps $\sum_{s\leq t}\Delta X_s \mathbf{1}_{|\Delta X_s|<1}$ may diverge (the path has infinite variation) and so the sum $\mathbf{1}_{\Delta X_s<1}$ may oscillates and be not defined. However by subtracting the mean to the sum we ensure the convergence in absolute value.
However, the definition of a Lévy process guarantees that $\int_{|x|<1} x^2\nu(dx) < \infty$. By compensating the process (subtracting the expected drift), we modify the integrand in the characteristic exponent:

$$\text{From } (e^{iux}-1) \sim iux \quad \text{to} \quad (e^{iux}-1-iux) \sim -\frac{1}{2}u^2x^2 \quad \text{as } x \to 0.$$

The quadratic behavior of the compensated term ensures the integral converges near zero, making the limit $\epsilon \to 0$ well-defined in L^2.
We can exploit the Lévy-Itô Decomposition to simulate a Lévy process $(X_t)_{t\geq 0}$. For simplicity, we neglect the diffusion term (W_t) in this discussion.
We approximate X_t by truncating the small jumps. Let $\epsilon > 0$ be a cutoff level. We define the approximated process X_t^ϵ as the sum of the drift, large jumps, and compensated jumps down to size ϵ:

$$X_t^\epsilon = \gamma t + \bar{N}_t + \tilde{N}_t^\epsilon,$$

where $\tilde{N}_t^\epsilon$ represents the compensated compound Poisson process of jumps with magnitude in $(\epsilon, 1)$. This approximation X_t^ϵ can be simulated exactly as a sum of a drift and independent compound Poisson processes.
The residual error of this approximation is defined as $R_t^\epsilon := X_t - X_t^\epsilon$. Formally, this is the limit of the compensated small jumps:

$$R_t^\epsilon = \lim_{\delta\to 0}(\tilde{N}_t^\delta - \tilde{N}_t^\epsilon).$$

Since $\tilde{N}_t^\epsilon$ is a martingale, the error has zero mean. Explicitly:

$$\begin{aligned}\mathbb{E}[\tilde{N}_t^\epsilon] &= \mathbb{E}[M_t^\epsilon]\mathbb{E}[Y_1] - t\int_{\epsilon<|x|<1} x\,\nu(dx) \\ &= (t\lambda^\epsilon)\left(\frac{1}{\lambda^\epsilon}\int_{\epsilon<|x|<1} x\,\nu(dx)\right) - t\int_{\epsilon<|x|<1} x\,\nu(dx) \\ &= 0.\end{aligned}$$

Consequently, $\mathbb{E}[R_t^\epsilon] = 0$. The variance of the error is determined by :

$$\mathrm{Var}(R_t^\epsilon) = \lim_{\delta\to 0} t\int_{\delta<|x|<\epsilon} x^2\,\nu(dx) := t\sigma^2(\epsilon) = t\int_{|x|<\epsilon} x^2\,\nu(dx) := t\sigma^2(\epsilon) \ \ ,$$

where we can take the limit because of the Lévy measure condition. This variance $\sigma^2(\epsilon)$ quantifies the quality of the approximation.
If X_t is a finite variation process, the decomposition is simpler because the integral of the small jumps, $\int_{|x|<1}|x|\nu(dx)$, is finite (it does not explode). Therefore, we do not strictly need the compensator for convergence.
However, to minimize simulation error when neglecting jumps smaller than ϵ, we replace their aggregate effect with their expected value (drift). The approximation by just reorganizing the terms in X_t^ϵ becomes:

$$X_t^\epsilon = \left(b + \int_{|x|<\epsilon} x\,\nu(dx)\right)t + \bar{N}_t^\epsilon,$$

where $\bar{N}_t^\epsilon$ is a compound Poisson process with intensity $\bar{\lambda}_\epsilon = \nu(\{x : |x| \geq \epsilon\})$ and jump distribution $\frac{1}{\bar{\lambda}_\epsilon}\mathbf{1}_{\{|x|\geq\epsilon\}}\nu(dx)$.
In this scheme, we are effectively replacing the stochastic sum of small jumps with their deterministic mean.

Proposition 1.47. *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function a.e. s.t. $|f'(x)| \leq C$. Then,*

$$|\mathbb{E}[f(X_T)] - \mathbb{E}[f(X_T^\epsilon)]| \leq C\sigma(\epsilon)\sqrt{T}.$$

Proof. Define the error term $R_T^\epsilon = X_T - X_T^\epsilon$. By the Fundamental Theorem of Calculus, we notice that:

$$f(X_T) - f(X_T^\epsilon) = f(X_T^\epsilon + R_T^\epsilon) - f(X_T^\epsilon) = R_T^\epsilon\int_0^1 f'(X_T^\epsilon + uR_T^\epsilon)\,du.$$

Then, by applying Jensen's inequality and the bound on f', we obtain:

$$\begin{aligned}|\mathbb{E}[f(X_T)] - \mathbb{E}[f(X_T^\epsilon)]| &\leq \mathbb{E}\left[\left|R_T^\epsilon\int_0^1 f'(X_T^\epsilon + uR_T^\epsilon)\,du\right|\right] \\ &\leq C\mathbb{E}[|R_T^\epsilon|] \\ &\leq C\sqrt{\mathbb{E}[(R_T^\epsilon)^2]} \\ &= C\sigma(\epsilon)\sqrt{T}.\end{aligned}$$

□

Notice that the approximation error increases with time and with the variance of small jumps. In case X_t is of finite variation the $\sigma^2(\epsilon) = \int_{|x|<\epsilon} x^2\nu(dx) \leq \epsilon\int_{|x|<\epsilon}|x|\nu(dx)$ and then goes to zero as ϵ. For other cases $\sigma^2(\epsilon)$ will still go to zero but it can go slower than ϵ.

Remark 1.48. This result applies to the payoff of a Put option where $f(x) = (K - F_0 e^x)^+$. The derivative is given by $f'(x) = -F_0 e^x \mathbb{I}_{\{x<\ln(K/F_0)\}}$. Observe that whenever the option is in-the-money (i.e., $x < \ln(K/F_0)$), we must have $F_0 e^x < K$. Consequently, the derivative is bounded by the strike price:

$$|f'(x)| = F_0 e^x \mathbb{I}_{\{x<\ln(K/F_0)\}} < K.$$

Thus, the condition holds with $C = K$.

Example 1.49. Let the tempered stable subordinator be defined by the Lévy density $\nu(x) = Cx^{-1-\alpha}e^{-(1-\alpha)x/k}$. The approximation X_T^ϵ is simulated as follows:

1. Calculate the large jump intensity Λ_ϵ and the small jump drift compensation b_ϵ:
$$\Lambda_\epsilon = \int_\epsilon^\infty \nu(x)\,dx, \qquad T\int_0^\epsilon x\nu(x)\,dx.$$
2. Draw the number of jumps $N \sim \text{Poisson}(\Lambda_\epsilon T)$.
Then, simulate the jump sizes y_k by inverting the cumulative distribution function (CDF) associated with the conditional density $\frac{\nu(x)}{\Lambda_\epsilon}\mathbb{I}_{\{x>\epsilon\}}$.
3. Compute the final value:
$$X_T^\epsilon = \sum_{k=1}^N y_k + b_\epsilon.$$

To improve the cases of infinite variation we can consider an additional improvement proposed by Asmussen and Rosinsky.

$$\tilde{X_t}^\epsilon = X_t^\epsilon + \sigma(\epsilon)W_t \ ,$$

which correspond with removing small jumps R_t^ϵ and replacing them with a diffusion with the same variance. This is based on the asymptotic result that follows.

Theorem 1.50. *$R_t^\epsilon/\sigma(\epsilon) \to W_t$ in distribution if and only if $\forall k > 0$ $\frac{\sigma(k\sigma(\epsilon)\wedge\epsilon)}{\sigma(\epsilon)} \to 1$ as $\epsilon \to 0$.*

Proof. We give an idea of the proof only the if part and in the subcase $\sigma(\epsilon)/\epsilon \to \infty$. Notice that this implies $k\sigma(\epsilon) \wedge \epsilon \to \epsilon$ and $\frac{\sigma(k\sigma(\epsilon)\wedge\epsilon)}{\sigma(\epsilon)} \to \frac{\sigma(\epsilon)}{\sigma(\epsilon)} \to 1$.

$$\mathbb{E}\left[\frac{R_t^\epsilon}{\sigma(\epsilon)}\right] = 0 \qquad Var\left[\frac{R_t^\epsilon}{\sigma(\epsilon)}\right] = t \ .$$

Jumps of R_t^ϵ by definition are below ϵ in absolute values and jumps of $R_t^\epsilon/\sigma(\epsilon)$ are below $\epsilon/\sigma(\epsilon)$ than jumps sizes go to zero for $R_t^\epsilon/\sigma(\epsilon)$ and we can conclude that it converges to a continuous Lévy process with mean zero and variance t. Moreover, the only continuous Lévy process is the Brownian motion. □

Proposition 1.51. *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function a.e. s.t. $|f'(x)| \le C$. Then,*

$$|\mathbb{E}[f(X_T)] - \mathbb{E}[f(X_T^\epsilon + \sigma(\epsilon)W_T)]| \le A\rho(\epsilon)\sigma(\epsilon) \ ,$$

where

$$\rho(\epsilon) = \frac{\int_{|x|\le\epsilon}|x|^3\nu(dx)}{\sigma(\epsilon)^3} \le \epsilon\frac{\int_{|x|\le\epsilon}|x|^2\nu(dx)}{\sigma(\epsilon)^3} \ .$$

Remark 1.52. If $\sigma(\epsilon)/\epsilon \to \infty$

$$\rho(\epsilon) = \frac{\int_{|x|\le\epsilon}|x|^3\nu(dx)}{\sigma(\epsilon)^3} \le \epsilon\frac{\int_{|x|\le\epsilon}|x|^2\nu(dx)}{\sigma(\epsilon)^3} = \frac{\epsilon}{\sigma(\epsilon)} \to 0 \ .$$

Remark 1.53. Approximating or removing small jumps in the Monte Carlo simulation introduces a bias in the estimator $\hat{\theta}$ of the true value θ. The Mean Squared Error (MSE) of the MC estimator decomposes into variance and squared bias:

$$\mathrm{MSE}[\hat{\theta}] = \mathbb{E}[(\hat{\theta} - \theta)^2] = \mathbb{V}\mathrm{ar}(\hat{\theta}) + \Big(\mathbb{E}[\hat{\theta}] - \theta\Big)^2.$$

If the simulation method is exact, the bias is zero, and one need only increase the number of simulations until the variance is sufficiently low. However, in the presence of bias, it is optimal to reduce the truncation error only until the squared bias is of the same order of magnitude as the variance (once I consider the variance small enough).

2 Pricing and simulation with characteristic functions

2.1 Pricing Formulas

Let X be a random variable on $\mathbb{R}$ such that $\mathbb{E}[e^x] = 1$ (i.e. $\hat{\mu}(-i) = 1$) and $\hat{\mu}(0) = 1$ and analytic in a neighborhood of the origin. Let A_X denote the interior set of

$$\left\{v \in \mathbb{R} : \mathbb{E}[e^{vX}] < \infty\right\}.$$

and the strip

$$\Lambda_X := \{\zeta \in \mathbb{C} : -\mathrm{Im}(\zeta) \in A_X\}.$$

Theorem 2.1. *The characteristic function of X is well defined and analytic in Λ_X.*

By Lucacs theorem the strip Λ_X is an open set and we identify its boundaries as $-p^-, p^+ + 1$. Assuming that $p^+, p^- > 0$ corresponds to assuming that $E[e^X]$ and $E[e^{0X}]$ are finite (necessary for martingale condition). In our setting, we already have asked for $p^+, p^- > 0$ because of $\mathbb{E}[e^x] = 1$ (i.e. $\hat{\mu}(-i) = 1$) and $\hat{\mu}(0) = 1$.

Remark 2.2. For model utilized in the equity market usually $p^+ > p^-$ due to the observed equity market skew.

We are interested in pricing two pay-offs

$$G_1(x, k) := (exp(x) - exp(k))^+$$

the call option and

$$G_2(x, k) := \mathcal{I}_{x>k} \ ,$$

the digital call option. We also define the expected value of the aformentioned pay-off for the random variable X as

$$C(k) = \mathbb{E}[G_1(X, k)] \qquad D(k) = \mathbb{E}[G_2(X, k)] \ .$$

Notice that from the expected value of the digital call option at different strikes we are able to obtain the CDF of X ($F_X(k) = 1 - \mathbb{E}[G_2(X, k)]$).
We give some bounds on the option prices. They hold for all k but we will use one for large positive k, and one for large negative k.

Theorem 2.3. *For any $p > 0$*

$$C(k) \leq \frac{\mathbb{E}[e^{(p+1)X}]}{(p+1)e^{pk}} \left(\frac{p}{p+1}\right)^p \quad \textit{and} \quad C(k) \leq 1\,\mathbb{E}[e^X] \ .$$

Moreover,For any $p > 0$,

$$D(k) \leq \frac{\mathbb{E}[e^{p\cdot X}]}{e^{pk}}, \quad \textit{and} \quad D(k) \leq \ .$$

Proof. See Lee *et al.* (2004), Th.3.1, 3.3. □

The Fourier transform of $C(k)$ and D_k do not exists because the $C(k)$ and $D(k)$ are not $L^1(\mathbb{R})$ since they tend to a positive constant for $k \to -\infty$ (by no arbitagrage argument $C(k)$ goes to $BF_0(T)$ and $D(k)$ goes to 1). However, the Fourier transform of the *damped* option price

$C_\alpha(k) = exp(\alpha k)C(k)$ and $D_\alpha(k) = exp(\alpha k)D(k)$ do have a Fourier transform provided that α is chosen correctly so that $C_\alpha(k), D_\alpha(k)$ is $L^1(\mathbb{R})$. Let us define

$$\hat{C}_\alpha(u) := \int_{\mathbb{R}} e^{iuk} C_\alpha(k) dk$$

and

$$\hat{D}_\alpha(u) := \int_{\mathbb{R}} e^{iuk} D_\alpha(k) dk \quad .$$

Theorem 2.4. *Consider $\alpha > 0$ with $\alpha < p^+$ for the European call and $\alpha < p^+ + 1$ for the digital. For any such α, the Fourier transform $\hat{C}_\alpha(u)$ and $\hat{D}_\alpha(u)$ exist and*

$$\hat{C}_\alpha(u) = \frac{\hat{\mu}(u - (\alpha+1)i)}{\alpha^2 + \alpha - u^2 + i(2\alpha+1)u},$$
$$\hat{D}_\alpha(u) = \frac{\hat{\mu}(u - \alpha i)}{\alpha + iu} \quad ,$$

where $\hat{\mu}$ is the characteristic function of X.

Proof. It exists $p > \alpha$ because A_X is open. By theorem 2.3, $C_\alpha(k)$ and $D_\alpha(k)$ decays exponentially for $|k| \to \infty$ and are bounded. E.g. for positive k consider the bound $C_\alpha(k) \leq \frac{Ee^{\alpha k}}{e^{pk}}$ where E is a positive constant and for negative k $C_\alpha(k) \leq e^{\alpha k}$. Therefore, $C_\alpha(k)$ and $D_\alpha(k)$ are in L^1. For the call

$$\hat{C}_\alpha(u) = \int_{\mathbb{R}} e^{iuk} C_\alpha(k) du = \int_{\mathbb{R}} e^{iuk} e^{\alpha k} \mathbb{E}[G_1(X,k)] du = \mathbb{E}\left[\int_{\mathbb{R}} e^{iuk} e^{\alpha k} G_1(X,k) dk\right] \quad ,$$

where the last equality is due to Fubini. Now we can compute the integral and get

$$\mathbb{E}\left[\int_{\mathbb{R}} e^{iuk} e^{\alpha k} G_1(X,k) dk\right] = \mathbb{E}\left[\int_{-\infty}^{x} e^{iuk} e^{\alpha k} (e^X - e^k) dk\right] = \frac{\mathbb{E}\left[e^{(1+\alpha+iu)X}\right]}{(\alpha + iu)(1 + \alpha + iu)}$$

which implies the thesis because $\alpha \in A_X$. Similar (and easier computations) can be done for $D(k)$. □

Now we can recover option prices by Fourier inversion. $f(k) \in L^1$ and $\int_{\mathbb{R}} \hat{f}(u) \in L^1$ then $f(k) = \frac{1}{2\pi} e^{-iuk} \hat{f}(u) du$.

Theorem 2.5. *Consider $\alpha > 0$ with $\alpha < p^+$ for the European call and $\alpha < p^+ + 1$ for the digital. Then*

$$C(k) = \frac{e^{-\alpha k}}{2\pi} \int_{\mathbb{R}} e^{-iuk} \hat{C}_\alpha(u) du = \frac{e^{-\alpha k}}{\pi} \int_{\mathbb{R}^+} Re[e^{-iuk} \hat{C}_\alpha(u)] du$$

and if $\hat{\mu}(u - \alpha_i)/(\alpha + iu) \in L^1$

$$D(k) = \frac{e^{-\alpha k}}{2\pi} \int_{\mathbb{R}} e^{-iuk} \hat{D}_\alpha(u) du = \frac{e^{-\alpha k}}{\pi} \int_{\mathbb{R}^+} Re[e^{-iuk} \hat{D}_\alpha(u)] du \quad .$$

Proof. For the call it is enough to notice that $C_\alpha(k)$ is L^1 and $\hat{C}_\alpha(u)$ is also L^1 because $|\hat{C}_\alpha(u)| \leq C/(u^2 + \alpha^2)$ for all $|u| > \bar{u}$ for a certain $u > 0$ and is bounded for for all $|u| \leq \bar{u}$ and the Fourier inversion hold. We can consider the integral on the positive half-line because $I(u) := e^{-iuk} \hat{C}_\alpha(u)$ is such that $I(-u) = \overline{I(u)}$ (it has conjugate symmetry $I(\bar{u}) = R(I(u)) - iIm(I(u))$) which means that the real part is even and the imaginary part is odd. This is true because Fourier transform

of a real-valued function has conjugate symmetry and then e^{-iuk} has conjugate symmetry ($e^{iuk} = cos(uk) + isin(uk) = cos(-uk) - isin(-uk) = \overline{e^{-iuk}}$) and

$$\overline{I(u)} = \overline{e^{-iuk}}\overline{\hat{C}_\alpha(u)} = e^{iuk}\hat{C}_\alpha(-u) = I(-u) \ ,$$

using that the product of complex conjugates is the complex cojugate of the product.
The statement on the digital call is due to the fact that, $D_\alpha(k)$ is L^1 and $\hat{D}_\alpha(u)$ is also L^1 by hypothesis. The rest is proven as if for the European call. □

Theorem 2.6. *Consider $\alpha \in A_X$ with $\alpha < p^+$ for the European call and $\alpha < p^+ + 1$ for the digital. Then*

$$C(k) = R_{\alpha,C} + \frac{e^{-\alpha k}}{2\pi}\int_{\mathbb{R}} e^{-iuk}\hat{C}_\alpha(u)du = R_{\alpha,C} + \frac{e^{-\alpha k}}{\pi}\int_{\mathbb{R}^+} Re[e^{-iuk}\hat{C}_\alpha(u)] \ ,$$

where

$$R_{\alpha,C} = \begin{cases} 1 & -1 < \alpha < 0 \\ 1/2 & \alpha = 0 \\ 0 & \alpha > 0 \end{cases}$$

and

$$D(k) = R_{\alpha,D} + \frac{e^{-\alpha k}}{2\pi}\int_{\mathbb{R}} e^{-iuk}\hat{D}_\alpha(u)du = R_{\alpha,D} + \frac{e^{-\alpha k}}{\pi}\int_{\mathbb{R}^+} Re[e^{-iuk}\hat{D}_\alpha(u)]$$

where

$$R_{\alpha,D} = \begin{cases} 1 & \alpha < 0 \\ 1/2 & \alpha = 0 \\ 0 & \alpha > 0 \end{cases} \ .$$

Proof. We prove it for the call. The proof for the digital call is exactly the same. From theorem 2.5 the thesis follow for $\alpha > 0$. For $-1 < \alpha < 0$ we know that for $\alpha_1 > 0$ the following holds

$$C(k) = \frac{e^{-\alpha_1 k}}{2\pi}\int_{\mathbb{R}} e^{-iuk}\hat{C_{\alpha_1}}(u)du \ = \frac{e^{-\alpha_1 k}}{2\pi}\int_{\mathbb{R}} e^{-iuk}\frac{\hat{\mu}(u-(\alpha_1+1)i)}{\alpha_1^2+\alpha_1-u^2+i(2\alpha_1+1)u}du \ .$$

By doing a change of variable $z = u - i\alpha_1$ we get

$$C(k) = \frac{1}{2\pi}\int_{-\infty-\alpha_1 i}^{\infty-\alpha_1 i} e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2}dz \ ,$$

which has two poles in $z = 0$, $z = i$. Consider a rectangular path Γ on $\mathbb{C}$ with horizontal segment on $Im(z) = -\alpha_1$ and $Im(z) = -\alpha$ and vertical segments on $Re(z) = \pm R$ that includes the pole in zero. The integrand is holomorphic outside of the poles. We can compute the residual of the pole as

$$Res(0) = \lim_{z\to 0}\frac{1}{2\pi}e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2} * z = \frac{1}{2\pi i}$$

Then, by Cauchy integration theorem $\int_\Gamma \frac{1}{2\pi}e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2}du = 2\pi i Res(0) = 1$ proceding counterclockwise

$$\frac{1}{2\pi}\int_{-\infty-\alpha_1 i}^{\infty-\alpha_1 i} e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2}dz - \frac{1}{2\pi}\int_{-\infty-\alpha i}^{\infty-\alpha i} e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2}dz = 1 \ ,$$

because $\lim_{R\to\infty}\frac{1}{2\pi}\int_{R-\alpha i}^{R-\alpha_1 i} e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2} = \lim_{R\to\infty}\frac{1}{2\pi}\int_{-R-\alpha i}^{-R-\alpha_1 i} e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2} = 0$. We can exchange limit and integrals by dominated convergence theorem because the integrand is bounded and the integration set is finite. For the $\alpha = 0$ case refer to Lee *et al.* (2004). We finish the proof by doing another change of variables $u = z + i\alpha$

$$C(k) = 1 + \frac{1}{2\pi}\int_{-\infty-\alpha i}^{\infty-\alpha i} e^{-ikz}\frac{\hat{\mu}(z-i)}{iz-z^2}dz = 1 + \frac{1}{2\pi}e^{-k\alpha}\int_{-\infty}^{\infty} e^{-iku}\frac{\hat{\mu}(u-i-i\alpha)}{(iu+\alpha)(iu+1+\alpha)}du$$

□

The European and Digital option prices are called the Lewis formula with a generic α. The formula in the original paper of Lewis is for $\alpha = -1/2$.

2.2 Numerical errors

When evaluating European and digital options with the Lewis formula we are performing a numerical integration and the total error is defined as the difference between the true value

$$C(k) = R_{\alpha,C} + \frac{e^{-\alpha k}}{\pi}\int_{\mathbb{R}^+} Re[e^{-iuk}\hat{C}_\alpha(u)]$$

and the discrete approximation given by the sum

$$\bar{C}^N_{\Delta u}(k) = R_{\alpha,C} + \frac{e^{-\alpha k}\Delta u}{\pi} Re\left[\sum_{n=0}^{N-1} e^{-i(n+0.5)\Delta uk}\hat{C}_\alpha((n+0.5)\Delta u)\right] \quad , \tag{4}$$

where N is the number of point that we use to discretize, Δu is the discretization step in the Fourier space. The approximation of the sum with the integral is done with the midpoint rule (rectangle rule). When discretizing the integral we get numerical errors due to truncation (the choice of $N\Delta u$: $|\bar{C}^N_{\Delta u}(k) - \bar{C}^\infty_{\Delta u}(k)|$) and discretization ($\Delta u$: $|C(k) - \bar{C}^\infty_{\Delta u}(k)|$). In particular, the two error can be identified as

$$|C(k) - \bar{C}^N_{\Delta u}(k)| \le |C(k) - \bar{C}^\infty_{\Delta u}(k)| + |\bar{C}^N_{\Delta u}(k) - \bar{C}^\infty_{\Delta u}(k)|$$

We deal first with the truncation error

Theorem 2.7. *If $\hat{\mu}(u)$ decays as a power law i.e. $\hat{\mu}(u) \le \frac{l(u)}{u^\gamma}$ for $\gamma + 1 > 0$ and for all $u > u_0$, where $l(u)$ is non increasing in u.*

$$|\bar{C}^N_{\Delta u}(k) - \bar{C}^\infty_{\Delta u}(k)| \le \frac{e^{-\alpha k} l((N-1)\Delta u)}{\pi(\gamma+1)((N-1)\Delta u)^{\gamma+1}}.$$

Moreover, if $\gamma > 0$

$$|\bar{D}^N_{\Delta u}(k) - \bar{D}^\infty_{\Delta u}(k)| \le \frac{e^{-\alpha k} l((N-1)\Delta u)}{\pi\gamma((N-1)\Delta u)^{\gamma}},$$

for all N s.t. $N\Delta u > u_0$.

Proof.

$$|\bar{C}^N_{\Delta u}(k) - \bar{C}^\infty_{\Delta u}(k)| \le \frac{\Delta u e^{-\alpha k}}{\pi}\sum_{n=N}^{\infty}|\hat{C}_\alpha((n+0.5)\Delta u)| \le \frac{e^{-\alpha k} l(N\Delta u)}{\pi}\sum_{n=N}^{\infty}\frac{\Delta u}{((n+1/2)\Delta u)^{\gamma+2}}$$

The second inequality is because $|e^{ic}| = 1$ the second because $|(\alpha+iu)(1+\alpha+iu)| = \sqrt{(u^2+\alpha^2)((u^2+(1+\alpha)^2)} \geq u^2$ and $l(u)$ non increasing in u. To conclude the proof for the call we use that, since $1/x^{\gamma+2}$ is decreasing

$$\sum_{n=N}^{\infty} \frac{\Delta u}{((n+1/2)\Delta u)^{\gamma+2}} \leq \int_{(N-1)\Delta u}^{\infty} 1/u^{2+\gamma} du = \frac{1}{(\gamma+1)((N-1)\Delta u)^{\gamma+1}} \ .$$

Similar results hold for the digital put just noticing that $|\alpha+iu| \leq u$ and that $1/x^{\gamma+1}$ is convex. □

Theorem 2.8. *If $\hat{\mu}(u)$ decays as an exponential i.e. $\hat{\mu}(u) \leq l(u)e^{-\gamma u^w}$ for $\gamma > 0$, $w > 0$ and for all $u > u_0$, where $l(u)$ is non increasing in u.*

$$|\bar{C}^N_{\Delta u}(k) - \bar{C}^{\infty}_{\Delta u}(k)| \leq \frac{e^{-\alpha k} l((N-1)\Delta u)\Gamma^*(1, \gamma((N-1)\Delta u)^w)}{\pi\omega}$$
$$|\bar{D}^N_{\Delta u}(k) - \bar{D}^{\infty}_{\Delta u}(k)| \leq \frac{e^{-\alpha k} l((N-1)\Delta u)\Gamma^*(0, \gamma((N-1)\Delta u)^w)}{\pi\omega} \ ,$$

where $\Gamma^*(s,u)$ is the upper incomplete gamma function.

Proof. We prove it for the digital call

$$|\bar{D}^N_{\Delta u}(k) - \bar{D}^{\infty}_{\Delta u}(k)| \leq \frac{\Delta u e^{-\alpha k}}{\pi} \sum_{n=N}^{\infty} |\hat{C}_{\alpha}((n+0.5)\Delta u)| \leq \frac{e^{-\alpha k} l(N\Delta u)}{\pi} \sum_{n=N}^{\infty} \frac{\Delta u e^{-\gamma((n+1/2)\Delta u)^w}}{\Delta u(n+1/2)} \ .$$

Moreover,

$$\sum_{n=N}^{\infty} \frac{\Delta u e^{-\gamma((n+1/2)\Delta u)^w}}{\Delta u(n+1/2)} \leq \int_{(N-1)\Delta u}^{\infty} \frac{e^{-\gamma u^w}}{u} du = \Gamma^*(0, \gamma((N-1)\Delta u)^w)/w \ .$$

□

We add that $\Gamma^*(s,x) \sim e^{-x}/x^{1+s}$ for large x.

Theorem 2.9. *The discretization error for $\alpha > 0$ is bounded by*

$$|C(k) - \bar{C}^{\infty}_{\Delta u}(k)| \leq e^{-2\pi\alpha/\Delta u} + e^{-2\pi(p^+-\alpha)/\Delta u + p^+ k} \hat{\mu}(-i(p^+ + 1)) \ ,$$

and

$$|D(k) - \bar{D}^{\infty}_{\Delta u}(k)| \leq e^{-2\pi\alpha/\Delta u} + e^{-2\pi(p^+-\alpha)/\Delta u + p^+ k} \hat{\mu}(-i(p^+)) \ .$$

Remark 2.10. The α that minimizes the discretization error is $p^+/2$. Notice that the truncation error does not depend explicitly from α. Hence, fixed α we can select Δu such that the discretization error is below a certain boundary and then N so that the truncation error is also below a certain boundary.

2.3 Pricing with Lewis formula

To evaluate (4) on a set of strikes one can compute the summation by hand but the Fast Fourier Transform is by far faster if we consider a set of strikes.
Given a vector x the function FFT (implemented in MATLAB) is very efficient in computing

$$X_m = \sum_{n=1}^{N} e^{-i(2\pi/N)(n-1)(m-1)} x_n \qquad m = 1:N \ .$$

In particular, the algorithm is applied to the vector x (so we compute $FFT(x_m)$ to get the vector X_m above).
To utilize this algorithm we need to rewrite the sum in (4) in this formulation. In particular we consider a grid $u_1, ...u_N$ with step Δu in the Fourier space and a grid $k_1, ...k_N$ with step $\Delta k = 2\pi/(N\Delta u)$ in the "moneyness" space. Then $k_m = k_1 + \Delta k(m-1)$ and $u_n = u_1 + (n-1)\Delta_u$. Then, for the strike k_m the sum in (4) is

$$\sum_{n=0}^{N-1} \Delta u e^{-i(n+0.5)\Delta u k_m} \hat{C}_\alpha((n+0.5)\Delta u) = e^{-i\Delta u k_m/2}\Delta u \sum_{n=1}^{N} e^{-i(n-1)\Delta u(k_1+\Delta k(m-1))}\hat{C}_\alpha((n-0.5)\Delta u)$$
$$= e^{-i\Delta u k_m/2}\Delta u \sum_{n=1}^{N} e^{-i(n-1)\Delta u \Delta k(m-1))} e^{-i(n-1)\Delta u k_1}\hat{C}_\alpha((n-0.5)\Delta u)$$
$$= e^{-i\Delta u k_m/2}\Delta u \sum_{n=1}^{N} e^{-i(2\pi/N)(n-1)(m-1))} e^{-i(n-1)\Delta u k_1}\hat{C}_\alpha((n-0.5)\Delta u)$$
$$=: e^{-i\Delta u k_m/2}\Delta u \sum_{n=1}^{N} e^{-i(2\pi/N)(n-1)(m-1))} x_n \ ,$$

and we are able to find our x_n. At time $t = 0$, we aim to price a set of European option with maturity T with different strikes. We model the forward as $F_T(T) = F_0(T)exp(X_T)$, where X_T has characteristic function $\hat{\mu}_T(u)$. In this setting the call price with strike K

$$C(F(0,T), B_0(T), K) = B_0(T)\mathbb{E}[(F_T(T) - K)^+] = B_0(T)\mathbb{E}[(F(0,T)e^{X_T} - K)^+]$$
$$= B_0(T)F_0(T)\mathbb{E}[(e^{X_T} - e^k)^+] = B_0(T)F_0(T)C(k) \ ,$$

where $k = log(K/F_0(T))$ is the log-moneyness and $B_0(T)$ is the discount factor.

1. From the characteristic function $\hat{\mu}_T(u)$ we identify p_T^+ and p_T^- by finding where $\hat{\mu}_T(-i\zeta)$ is not defined.

2. We select $\alpha = p_t^+/2$ (or $p_t^-/2$ if $p_t^+ < p_t^-$).

3. Consider a grid of moneyness k and a grid in the Fourier space u. Notice that to use FFT $\Delta u \Delta k = 2\pi/N$ and then fixing N i will have large Δk if Δu is small. Suggestion, find Δk sufficiently small for your grid of log-moneyness (e.g. 0.01). Then find the truncation N such that both errors are below a reasonable threshold (e.g. 10^{-8}.)

4. We compute the $x_n = e^{-i(n-1)\Delta u k_1}\hat{C}_\alpha((n-0.5)\Delta u)$ and the call prices as

$$C(k) = \left(R_{\alpha,C} + \frac{e^{-\alpha k}e^{-i\Delta u k_m/2}\Delta u}{\pi} FFT(x)\right) \ ,$$

   where x and k are n-dimensional vectors.

5. If I am interest only in a subset of log-moneyness k^* I can obtain $c(k^*) = spline(k, C(k), k^*)$.

An analogous algorithm can be applied also for the digital call (clearly in this case we do not need to multiply with the Forward price).

2.4 Simulating with characteristic function

We aim to simulate the aformentioned random variable X. The main idea is to estimate the CDF on a grid exploiting FFT and simulate from it by generating uniform random variable. Given the CDF of X P and a uniform random variable Z we can simulate X as $P^{-1}(Z)$. The proof is left to the reader.
We recall that the CDF of X estimated with the Lewis formula is

$$\bar{P}^N_{\Delta u}(k) = 1 - \bar{D}^N_{\Delta u}(k) \ .$$

- We find p^+ and p^- for X. We select parameters α, Δu, and N (as discussed for the call) and compute $\bar{P}^N_{\Delta u}(k)$ on a grid of k values using the FFT algorithm.
- Ensure that $\bar{P}^N_{\Delta u}(k)$ is monotonically increasing and bounded in $[0, 1]$.
  - i.e., cap and floor the values and consider only the interval where it is monotonically increasing (discarding all other log-moneyness).
  - We end up with a new grid $\bar{k}$ in which we identify $\bar{k}_{\min}$ and $\bar{k}_{\max}$.
  - We call the values at these boundaries $\bar{P}_{\min} = \bar{P}^N_{\Delta u}(\bar{k}_{\min})$ and $\bar{P}_{\max} = \bar{P}^N_{\Delta u}(\bar{k}_{\max})$.
- Simulate N_{sim} independent uniform random variables $Z_j \sim U[0, 1]$.
- Simulate X inverting the CDF, $X_j = (\bar{P}^N_{\Delta u})^{-1}(Z_j)$, for each Z_j:

  a) **Interpolation:** If $Z_j \in [\bar{P}_{\min}, \bar{P}_{\max}]$, we find $\bar{X}_j$ by inverting the gridded function:

  $$\bar{X}_j = \text{interp}(\bar{P}^N_{\Delta u}(\bar{k}), \bar{k}, Z_j)$$

  b) **Extrapolation (Clamping):** If Z_j is outside the interpolation range:
  - If $Z_j > \bar{P}_{\max}$, set $\bar{X}_j = \bar{k}_{\max}$.
  - If $Z_j < \bar{P}_{\min}$, set $\bar{X}_j = \bar{k}_{\min}$.

We can consider linear or spline interpolation. We can further refine the algorithm by simulating the tails of X it is possible to derive an exponential extrapolation for $Z_j \notin [\bar{P}_{\min}, \bar{P}_{\max}]$.

Simulate the increments of a Lévy process If we are simulating from zero to T the procedure above applies. Otherwhise for $0 < s < t$

$$\mathbb{E}[e^{iu(X_t - X_s)}] = \mathbb{E}[e^{iu(X_{t-s})}] = e^{(t-s)\psi(u)} \ ,$$

where the equality is due to the stationarity of increments and $\psi(u)$ is the characteristic exponent. At this point, I can find p^+ and p^- as the point in which $\psi(u)$ is not analytic on the imaginary axis. $t - s$ does not interfer with the analiticity hence the p^+ and p^- are the same for all t and s. At this point I can simulate the increments we have introduced for a generic random variable.

Simulate the increments of an Additive process . For an additive process the non stationarity of increments need to be dealt with. In particular, for any time t the characteristic function of the process $\hat{\mu}_t(u)$ has potentially different analyticity strip p^+_t and p^-_t.

Theorem 2.11. *For an additive process p^+_t and p^-_t are non increasing in t.*

Proof. See Azzone and Baviera (2023). □

Corollary 2.12. *For any $0 < s < t$, the analyticity strip of an additive process increments $X_t - X_s$ is identified by p_t^+ and p_t^-.*

Proof. For an additive process

$$\hat{\mu}_t(u) = \mathbb{E}[e^{iu(X_t)}] = \mathbb{E}[e^{iu(X_t - X_s)}]\mathbb{E}[e^{iuX_s}] =: \hat{\mu}_{s,t}(u)\hat{\mu}_s(u) \ ,$$

for the independece of increments. Hence, the characteristic function of the increment $\hat{\mu}_{s,t}(u) = \hat{\mu}_t(u)/\hat{\mu}_s(u)$ and then the analiticity strip is the intersection of the analiticity strips of $\hat{\mu}_t(u)$ and $\hat{\mu}_s(u)$. But since p_t^+ and p_t^- are non increasing $[-p_t^-, p_t^+ + 1] \cap [-p_s^-, p_s^+ + 1] = [-p_t^-, p_t^+ + 1]$. □

Them, to simulate the increments of an additive process we need to utilize the analiticity strip at time t_1 and the algorithm for X applies.

3 How to calibrate a volatility surface

3.1 Discount factors and Forward from option prices

On a given day $t_0 = 0$, we are faced with the problem of obtaining model parameters from quoted option prices on a set of maturity and strikes. We call quoted option prices as $c_{K,T}$. In this course, we consider the

$$F_t(T) := F_0(T) \ \exp(f_t) \ , \tag{5}$$

The call option price for the model is $C_{k,t}^f$

$$C_{k,T}^f = B_0(T)\mathbb{E}[(F_0(T)e^{f_T} - K)^+] \ ,$$

Notice that the call prices depends also from the forward price and the discount factor. To calibrate we follow the concept of a calibration cascade approach: i.e. we calibrate first the most liquid instruments (discounts and forwards) and then focus on the model parameters. The standard approach would be to bootstrap interest rate from market instruments and obtain prices for forward or futures on the underlying. However, this is not the best approach because you can have futures not available at all maturity and the interest rate, the futures and the option market can be not perfectly synchronized. We introduce here a technique to get interest rates and forwards directly from call and put prices without adding any additional data.
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 option w.r.t. the forward price F and the strike price K

$$C(K) - P(K) = B_0(T)(F - K) \ , \tag{6}$$

where $C(K)$ and $P(K)$ are respectively the European call and put option prices and $\overline{B}_0(T)$ is the market discount factor between 0 and T. We drop the time variable in calls puts and forwards for simplicity but what follow can be applied to option expiring on any maturity.
Instead of considering a standard forward contract, a trader 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 $\mathcal{G}(K)$ 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.[1] The market implied discount factor $\overline{B}_0(T)$ is the (unique) factor such that the forward price

$$F = \frac{\mathcal{G}(K)}{\overline{B}(t_0, T)} + K \tag{7}$$

does not depend on the strike K:
The discount factor used in the market $\overline{B}_0(T)$ can be obtained as the angular coefficient in the linear regression

$$\mathcal{G}_i = -\overline{B}_0(T)\, K_i + \overline{B}(t_0, T)\, F + \epsilon_i \qquad i = 1, .., N \tag{8}$$

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

$$\overline{B}_0(T) = -\frac{\sum_{i=1}^{N}(K_i - \hat{K})(\mathcal{G}_i - \hat{\mathcal{G}})}{\sum_{i=1}^{N}(K_i - \hat{K})^2} \tag{9}$$

where

$$\hat{\mathcal{G}} := \frac{1}{N}\sum_{i=1}^{N} \mathcal{G}_i \quad , \quad \hat{K} := \frac{1}{N}\sum_{i=1}^{N} K_i \, . \tag{10}$$

Let us underline that this methodology allows us to determine also the forward price obtained via synthetic forwards.
Empirical observation in Azzone and Baviera (2021) shows that in the US market the estimated interest rate is above the OIS one by around 30 bp which indicate the presence of a cost of funding in the equity market.
This forward price is obtained from the put-call parity relation (7) using the $\overline{B}(t_0, T)$ that includes the cost of funding: at a given maturity T, the forward ask price is the lowest forward ask $F^{ASK} = min_K \frac{C^{ASK}(K) - P^{BID}(K)}{\overline{B}(t_0,T)} + K$ and the forward bid price is the highest forward bid $F^{BID} = max_K \frac{C^{BID}(K) - P^{ASK}(K)}{\overline{B}(t_0,T)} + K$. Further improvement: select the forward price such that the implied volatility smile is smooth in the point in which you pass from put to call in case you consider OTM put and calls in the calibration (i.e. the strikes close to the ATM can be fitted with a straight line). You have to ensure that the straitg lines intersecting the IV of the first two call and the first two puts intersect exactly in for zero log-moneyness.

3.2 Calibration cascade

Consider a volatility surfaces that comprise of bid and ask prices of call options on several strikes and maturities. You need to calibrate first discount factors and forwards that are much more liquid than options.

- Data pre-processing: 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 a reasonable threshold (e.g. 0.1 index point for S&P 500 and EURO STOXX) , 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.

[1] 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.

- Forward and discounts: obtain interest rates and forward prices synchronized with option prices following the methodology described above.
- Consider and additional filtering based on liquidity. Usually, you can consider just options in the 10%-90% delta range, i.e. whose B&S delta falls in the (10%,90%) range in absolute value. To compute the delta you need to compute the option B&S implied volatility using Forward and discounts.
- Select a model and minimize the error: minimize the distance between market and model prices. You can calibrate on OTM call and put or on implied volatility. E.g.

$$\sum_{K,T}(C^{mkt}(K,T)-C^{mdl}(K,T))^2\mathcal{I}_{K>F_0(T)}+(P^{mkt}(K,T)-P^{mdl}(K,T))^2\mathcal{I}_{K\leq F_0(T)} \quad .$$

3.3 Hedging

This section will be based on the detailed textbook Taleb (1997). We will detail the case of a market maker who has to hedge a portfolio of derivatives P. We will consider the baseline case in which we just have one asset S and a set of plain vanilla options $\mathbf{C}$ of which he/she computes the implied volatility surface $\mathbf{\Sigma}_t$ for different strikes and maturities. We assume the market maker is calibrating a certain parametric model (with parameters θ). We assume the market maker is calibrating the model on market data and is getting an estimate for the model parameters $\hat{\theta}(S_t,\mathbf{\Sigma}_t)$ at time t. The value of the portfolio computed with the calibrated model will be $\hat{P}(S_t,\hat{\theta}(S_t,\mathbf{\Sigma}_t))$.

The Portfolio Δ :

$$\Delta_t=\frac{\partial\hat{P}(S_t,\hat{\theta}(S_t,\mathbf{\Sigma}_t))}{\partial S_t} \quad .$$

The sensitivity of a derivative price to the movement in the underlying asset. For an exotic derivatives portfolio it is usually computed numerically as

$$\hat{\Delta}_t=\frac{\hat{P}(S_t+\epsilon,\hat{\theta}(S_t,\mathbf{\Sigma}_t)-\hat{P}(S_t,\hat{\theta}(S_t,\mathbf{\Sigma}_t))}{\epsilon} \quad ,$$

or more robustly as

$$\hat{\Delta}_t=\frac{\hat{P}(S_t+\epsilon,\hat{\theta}(S_t,\mathbf{\Sigma}_t)-\hat{P}(S_t-\epsilon,\hat{\theta}(S_t,\mathbf{\Sigma}_t))}{2\epsilon} \quad ,$$

where ϵ is a reasonable monetary quantity for the operator (e.g. a bp.). When computing the Δ notice that you are not changing the model calibrated parameters. The Δ is usually hedged by considering a long or short position in the spot or in the forward. Notice that in the equity market the Delta of a forward is not exactly one but is, at time t $S_t/F_t(T)$.

The Portfolio Vega:

$$V_t=\frac{\partial\hat{P}(S_t,\hat{\theta}(S_t,\mathbf{\Sigma}_t))}{\partial\mathbf{\Sigma}_t} \quad .$$

This is the standard definition. However, in principle one should hedge against all possible $\sigma_{i,j}$ of the implied volatility surface, which is clearly impractical due to the exponential hedging costs. When computing the numerical derivative, we consider different possible movements of the volatility surface for hedging the Vega:

1. **Parallel shift:**
$$\hat{V}_t = \frac{\hat{P}(S_t, \hat{\theta}(S_t, \mathbf{\Sigma}_t + \epsilon)) - \hat{P}(S_t, \hat{\theta}(S_t, \mathbf{\Sigma}_t - \epsilon))}{2\epsilon} \ .$$
This approach considers a uniform parallel shift of the entire volatility surface (typically with $\epsilon \approx 1\%$). In this case, the model parameters are recalibrated under the shifted implied volatility.

2. **Bucketed parallel shift:** For each time-to-maturity (bucket) one computes the Vega by assuming a parallel shift of the implied volatility for that maturity. The hedge can then be constructed either at the bucket level or by grouping buckets into coarser aggregates.

   Coarse-grained Bucket: In this setting, the focus is on the impact of selected macro-buckets of volatility.

   (a) Compute the individual bucket Vegas V^{t_i}.

   (b) Compute the aggregate Vega for the macro-bucket as a weighted average of the individual bucket Vegas: We call t_i the time to maturites and k_j the time to maturity corresponding to the m course grained buckets (less k_j than t_i)
$$w_i^{(j)} = \begin{cases} 1, & t_i \leq k_1 \\ \frac{t_i - k_{j-1}}{k_j - k_{j-1}}, & k_{j-1} \leq i \leq k_{j+1},\ i \neq k_j, \\ 1, & t_i = k_j \\ 1, & t_i \geq k_m \end{cases}$$
   where k_j is the maturity corresponding to the j-th macro-bucket.

   (c) $\mathcal{V}^{k_j} = \sum w_i^{(j)} V^{t_i}$

   The macro-buckets are typically established by the risk committee (in agreement with the bank's risk management). An example could be $k_1 = 6\text{m}, k_2 = 1\text{y}, k_3 = 2\text{y}, k_4 = 5\text{y}$.

   For instance, consider three macro-buckets. One hedges the k_3 Vega with an option expiring at two years C_3, the k_2 Vega with an option expiring at one year C_2, and the k_1 Vega with an option expiring at six months C_1. The following linear system must then be solved:
$$\begin{aligned} \mathcal{V}_P^{k3} + x_3 \mathcal{V}_{C3}^{k3} &= 0, \\ \mathcal{V}_P^{k2} + x_2 \mathcal{V}_{C2}^{k2} + x_3 \mathcal{V}_{C3}^{k2} &= 0, \\ \mathcal{V}_P^{k1} + x_1 \mathcal{V}_{C1}^{k1} + x_2 \mathcal{V}_{C2}^{k1} + x_3 \mathcal{V}_{C3}^{k1} &= 0 \ , \end{aligned}$$
   where $\mathcal{V}_j^{k_i}$ denotes the Vega with respect to bucket i of instrument j.

3. **Smile rotation (skew):** To hedge against skew movements, one may consider a rotation of the implied volatility smile (typically centered around the ATM). For a single smile, an example rotation vector is
$$[-1\%, -1\%, -1\%, -1\%, -1\%, -1\%, \ 1\%, 1\%, 1\%, 1\%, 1\%, 1\%].$$
   A similar construction can be extended to the entire volatility surface. One may hedge skew risk either for a specific smile, for selected smiles of interest, or for the entire surface. This is typically achieved using a call spread that is neutral to parallel shifts (thus preserving the hedge enforced in the previous step) but remains sensitive to rotation shifts.

Vega Hedging with model parameters: Instead of hedging against arbitrary movements of the volatility surface, one can hedge directly with respect to the model parameters. The Greeks we consider are of the form

$$\frac{\partial P(S_t, \hat{\theta})}{\partial \theta_i}.$$

Consider the simple case of the NIG model, which is characterized by three parameters: σ, κ, and η. Specifically, σ controls the volatility level, η the skew, and κ the convexity. One can therefore hedge

$$\frac{\partial P(S_t, \hat{\theta})}{\partial \sigma}, \quad \frac{\partial P(S_t, \hat{\theta})}{\partial \eta}, \quad \frac{\partial P(S_t, \hat{\theta})}{\partial \kappa},$$

thus achieving immunization against the most common movements of the volatility smile.
For a full volatility surface this becomes more challenging, but we will see that additive processes still allow for partial protection. Ultimately, one must decide which parameters are relevant to hedge. A practical rule of thumb is the following: compute the average standard deviation of a parameter, say $\mathrm{std}(\sigma)$, and multiply it by the corresponding Greek. The product

$$\mathrm{std}(\sigma) \cdot \frac{\partial \hat{P}(S_t, \hat{\theta})}{\partial \sigma}$$

provides a measure of how much fluctuations in σ affect the derivative price.
Moreover, if a parameter oscillates significantly due to calibration issues (e.g., over-parameterization), hedging it may be both costly and risky, since the hedging costs can become prohibitively high. For these reasons hedging with model parameters is not very common in practitioner desks.

4 An example of additive process: the Additive normal tempered stable process (ATS).

We recall the Lévy normal tempered stable processes (LTS) a According to this modeling approach, the forward with expiry T is an exponential Lévy; i.e.

$$F_t(T) := F_0(T)\ \exp(f_t)\ , \tag{11}$$

with f_t a LTS

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

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 $F_t(T)$.[2] W_t is a Brownian motion and S_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

$$\mathbb{E}\left[e^{iuf_t}\right] = \mathcal{L}_t\left(iu\left(\frac{1}{2} + \eta\right)\sigma^2 + \frac{u^2\sigma^2}{2};\ k,\ \alpha\right) e^{iu\varphi\, t}\ , \tag{12}$$

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

$$\ln \mathcal{L}_t\left(u;\ k,\ \alpha\right) := \begin{cases} \dfrac{t}{k}\dfrac{1-\alpha}{\alpha}\left\{1 - \left(1 + \dfrac{u\ k}{1-\alpha}\right)^{\alpha}\right\} & \text{if } \ 0 < \alpha < 1 \\ -\dfrac{t}{k}\ln\left(1 + u\ k\right) & \text{if } \ \alpha = 0 \end{cases}. \tag{13}$$

[2] We utilize here $\mu = -\left(\frac{1}{2} + \eta\right)\ \sigma^2$. A parametrization scheme of the drift in terms of η can be suitable in 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. This is the parametrization used in the FE course.

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 (12) but with time-dependent parameters

$$\mathbb{E}\left[e^{iuf_t}\right] = \mathcal{L}_t\left(iu\left(\frac{1}{2}+\eta_t\right)\sigma_t^2 + \frac{u^2\sigma_t^2}{2};\ k_t,\ \alpha\right)e^{iu\varphi_t t}\ , \tag{14}$$

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

Theorem 4.1. Sufficient conditions for existence of ATS
There exists an additive process $\{f_t\}_{t\geq 0}$ with the characteristic function (14) 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.*

For the full proof see Azzone and Baviera (2022). The idea of this proof is to show that there exists an additive process with the characteristic function in (14) using the result in Sato (1999, Th.9.8, p.52).
At any given time $t > 0$ the characteristic function in (14) is the characteristic function of a LTS (12), 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 (14) (see, e.g. Cont and Tankov 2003, eq. 4.24, p.130), we just need to replace $k = k_t$, $\eta = \eta_t$, $\sigma = \sigma_t$ and $\varphi = \varphi_t$ and then check that if the hypotheses of the theorem hold the condition of theorem 9.8 of Sato are satisfied.
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. It is possible to calibrate the model starting from the first, the last or a maturity in the middle. We outline the procedure starting for the first maturity.

1. For the first maturity T_1 calibrate the three parameters $k_{T_1}, \eta_{T_1}, \sigma_{T_1}$ minimizing the mean squared errors between market prices and model prices.

2. For maturity T_i, calibrate the ATS parameters $k_{T_i}, \eta_{T_i}, \sigma_{T_i}$ minimizing the mean squared errors between market prices and model prices and ensuring that $g_1(T_i) \geq g_1(T_{i-1})$, $g_2(T_i) \geq g_2(T_{i-1})$ and $g_3(T_i) \geq g_2(T_{i-1})$.

Notice that in this case we are considering 3 parameters per maturity (i.e. 10 maturities 30 parameters). However, this parameters are interpretable bevcause they are connected to level skew and convexity of every single smile.
We present a calibration example on S&P 500 and EURO STOXX 50 option prices on 11 am New York Time on the 30 of May 2013.

The calibration performance is reported in Table 1 in terms of Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). The ATS processes improvement is, on average, above two orders of magnitude. Although we present the results for VG and NIG ($\alpha = 0$ and $\alpha = 1/2$), 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	ATS	Lévy	ATS
S&P 500	NIG	4.56	**0.02**	3.13%	**0.23%**
S&P 500	VG	8.49	**0.24**	4.31%	**0.79%**
Euro Stoxx 50	NIG	22.15	**0.10**	1.75%	**0.09%**
Euro Stoxx 50	VG	55.81	**0.35**	2.85%	**0.21%**

Table 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.

In Figure 1, 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*.

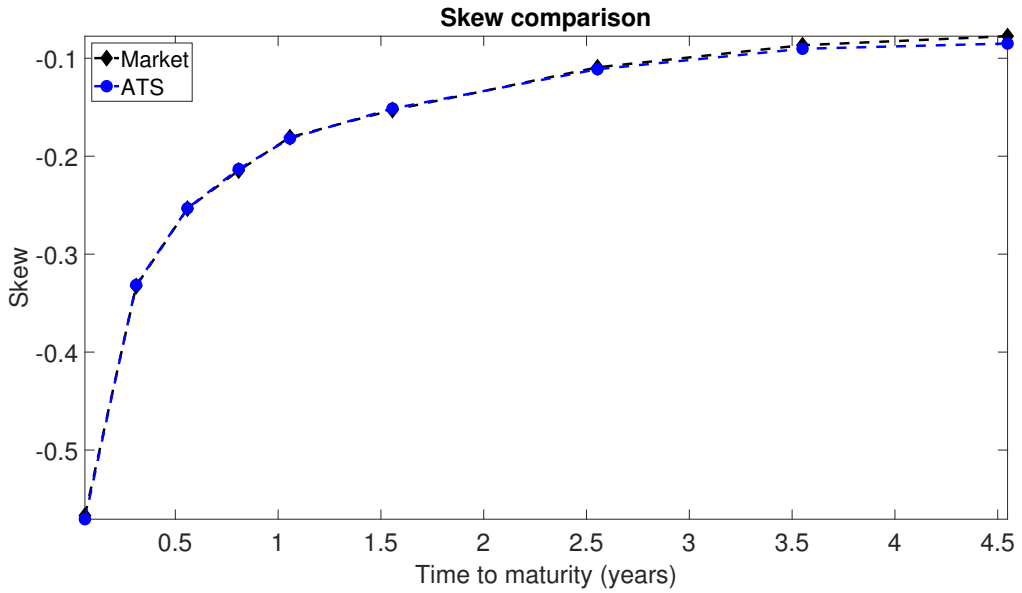

Figure 1: The market and the NIG ATS implied volatility *skew* for EURO STOXX 50 w.r.t. the times to maturity. ATS replicates the market implied volatility *skew* behavior.

We introduce a sub-case of ATS, determined by self-similar functions of time.
This family of processes describes accurately market implied volatility surfaces. Power-law scaling functions of time allow us to rewrite **Theorem 4.1** conditions as simple inequalities on the scaling parameters.

Theorem 4.2. 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. We check that the power-law sub-case of ATS satisfies the two conditions of **Theorem 4.1**.
First, we verify that $g_1(t)$, $g_2(t)$, and $g_3(t)$ are non decreasing.

$$g_1(t) = (1/2 + \bar{\eta} t^{\delta}) - \sqrt{(1/2 + \bar{\eta} t^{\delta})^2 + 2(1-\alpha)/(\bar{\sigma}^2 \bar{k} t^{\beta})}$$

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 $\beta > -\delta$.

$$g_2(t) = -(1/2 + \bar{\eta} t^{\delta}) - \sqrt{(1/2 + \bar{\eta} t^{\delta})^2 + 2(1-\alpha)/(\bar{\sigma}^2 \bar{k} t^{\beta})}$$

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 $\delta \leq 0$.

$$g_3(t) = \frac{\sqrt{\bar{\sigma}^4 t^{2/\alpha - 2\beta(1-\alpha)/\alpha}\,(1/2 + \bar{\eta} t^{\delta})^2 + 2t^{-\beta+2/\alpha-2\beta(1-\alpha)/\alpha}\bar{\sigma}^2(1-\alpha)/(\bar{k})}}{\bar{k}^{(1-\alpha)/\alpha}}$$

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 $\beta \leq \frac{1}{1-\alpha/2}$ □

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. $x =: y\sqrt{t}$. 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. 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

$$\mathcal{I}_t(y) = \mathcal{I}_t(0) + y\ \left.\frac{d\mathcal{I}_t(y)}{dy}\right|_{y=0} + o(y) =: \hat{\sigma}_t + y\,\hat{\xi}_t + o(y) \ . \tag{15}$$

We call $\hat{\xi}_t$ the skew term. We define $\hat{\sigma}_0$ and $\hat{\xi}_0$ as 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

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

In Figure 4, 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 bounded 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 $O\left(\sqrt{\frac{1}{t}}\right)$.

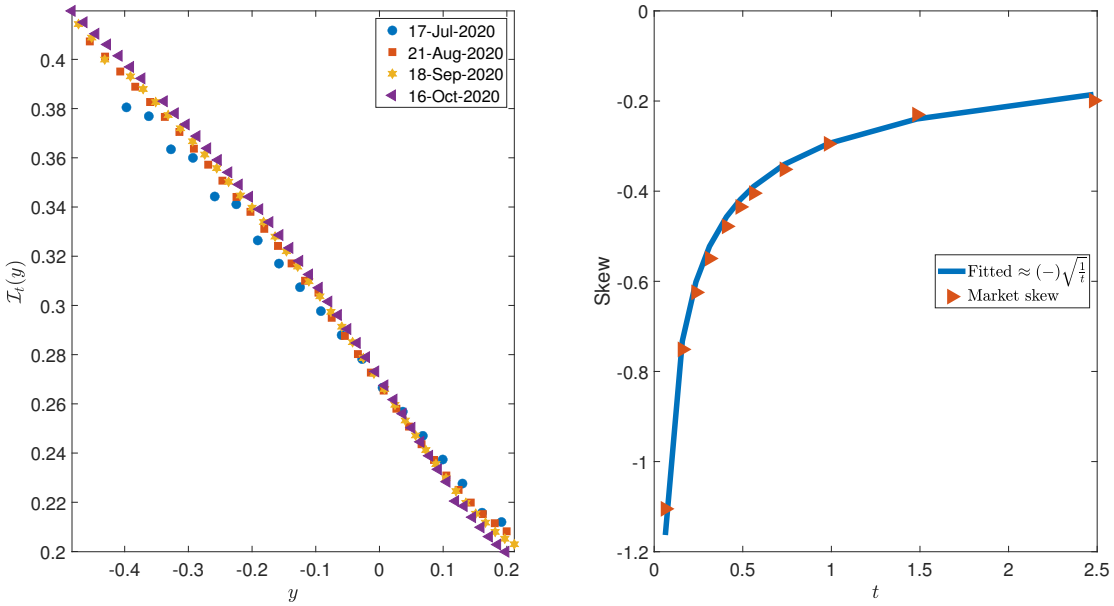

Figure 2: Example of the S&P 500 short-time implied volatility and skew on the 22^{nd} of June 2020. On the left, we plot the one month (blue circles), two months (red squares), three months (orange stars), and four months (purple triangles) market implied volatility w.r.t. the *moneyness degree* y. 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 $\approx (-)\sqrt{\frac{1}{t}}$.

Equity market data are compatible with a positive and bounded $\hat{\sigma}_0$ and a negative and bounded $\hat{\xi}_0$, that leads to a skew proportionally inverse to the square root of the time-to-maturity.

Theorem 4.3.
The ATS has a positive and constant short-time implied volatility $\hat{\sigma}_0$ and a negative and constant short-time skew term $\hat{\xi}_0$, if and only if $\beta = 1$ and $\delta = -1/2$.

Proof. See Azzone and Baviera (2024). □

[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.

This process has just 3 free parameters and replicates the short time volatility and skew. It is possible to consider another variation with more parameters but still less than 3 parameters for maturity. We can do so by constructing a deterministic time-change for the ATS.

Proposition 4.4. 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. 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 □

□

Consider a power scaling ATS with $\bar{\sigma} = 1$ if we time-change it with a non decreasing function of time θ_t we get a new additive process that has 4 parameters plus one parameters per maturity or 2 paramaters +1 one per maturity if we fix β and δ.
In this case the charactersitic function is

$$\mathbb{E}\left[e^{iuf_t}\right] = \mathcal{L}_{\theta_t}\left(iu\left(\frac{1}{2} + \bar{\eta}\theta_t^{\delta}\right) + \frac{u^2}{2};\ \bar{k}\theta_t^{\beta},\ \alpha\right) e^{\hat{\varphi}_t t} \quad , \tag{16}$$

where $\hat{\varphi}_t$ is selected to satisfy the martingale condition. The calibration of this model can only be done globally (i.e. all parameters should be fitted at the same time). During the calibration you should ensure the monotonicity of θ_t and the admissible range for β and δ if they are not fixed.

Proposition 4.5. *For ATS processes with* $\alpha \in (0, 1)$ *we have that* $p_t^+ \geq p_t^-$.

Proof. To identify p_t^+ and p_t^-, we apply the Lukacs theorem (cf. **?**, th.3.1, p.12). At time t, the ATS characteristic function in equation (14) is analytic on the imaginary axis $u = -i\,a$, $a \in \mathbb{R}$ iff

$$1 + \frac{k_t}{1-\alpha}\left(a\left(\frac{1}{2} + \eta_t\right)\sigma_t^2 - \frac{a^2\sigma_t^2}{2}\right) > 0 \quad .$$

By solving the second order inequality, we get

$$g_1(t) < a < -g_2(t) \quad ,$$

with $g_1(t)$ and $g_2(t)$ defined in (4). Hence, $p_t^+ := -g_2(t) - 1$ and $p_t^- := -g_1(t)$.
It holds that $p_t^+ \geq p_t^-$ because

$$p_t^+ - p_t^- = 2\eta_t \geq 0$$

□

5 Modelling Volatility

5.1 Stochastic volatility and Heston model

A widely studied class of models satisfy the following SDE

$$dS_t = \mu S_t dt + \sigma_t S_t dW_t \ ,$$

where σ_t is called the istantaneous volatility process. This process is not Markov but the bivariate process (S_t, σ_t) is Markov. This is relevant to simulate the model you need to simulate the two dimensional process otherwhise you miss the information.
Usually, we want the stochastic volatility process to be positive and mean reverting (volatility is stationary in market data). It is standard to define $\sigma_t := f(y_t)$ with f a positive function and y_t a process that satisfies

$$dy_t = \lambda(\eta - y_t)dt + ...d\hat{Z}_t \ ,$$

where λ is the speed of mean reversion η the long-term mean and $\hat{Z}_t$ a Brownian motion correlated with W_t with correlation parameter ρ. Some common driving process are

1. Geometric BM (no mean reversion)
2. Gaussian Ornstein-Uhlenbeck:

$$dy_t = \alpha(\eta - y_t)dt + \beta d\hat{Z}_t$$

3. Cox-Ingersoll-Ross

$$dy_t = \kappa(\eta - y_t)dt + v\sqrt{y_t}d\hat{Z}_t$$

In all this cases we also have an initial condition for the volatility. I.e. $y_0 = v$.
To keep a consistent notation we model directly the forward as $F_t(T) = F_0(t)\exp(X_t)$. The Heston model is a stochastic volatility model with a CIR volatility process and $f = \sqrt{}$. We model directly the standardized forward as a martingale.

$$\begin{cases} d\exp(X_t) & = \exp(X_t)\sqrt{V_t}dW_t \\ dV_t & = \kappa(\eta - y_t)dt + \theta\sqrt{V_t}d\hat{Z}_t \end{cases}$$

with $y_0 = v, \exp(X_0) = 1$ or equivalently for the exponent X_t

$$\begin{cases} dX_t & = (-1/2V_t)dt + \sqrt{V_t}dW_t \\ dV_t & = \kappa(\eta - V_t)dt + \theta\sqrt{V_t}d\hat{Z}_t \ , \end{cases}$$

with $V_0 = v, X_0 = 0$. The Heston model has 5 parameters, v the initial value of volatility (level of the smile), η the long-term mean of volatility, κ the speed of mean reversion, θ the vol-of-vol (linked to the convexity) and ρ the correlation (negative rho implies negative skew). It is difficult to disentangle the effects of v, κ, η.

Heston characteristic function: Good property of Heston model: you can get characteristic function of the log-return in closed form. Let us define

$$f(x, v, t) = \mathbb{E}[e^{iuX_T}|X_t = x, V_t = v]$$

which is the characteristic function of X_T given some initial conditions (notice we can start from any time $t < T$). Notice that $f(X_t, V_t, t) = \mathbb{E}[e^{iuX_T}|\mathcal{F}_t]$ is a martingale with respect to the filtration

$\mathcal{F}_t$ due to the tower property. We write its SDE and set the drift equal to zero to get an expression for f. By two-dimensional Ito

$$\begin{aligned} df(X_t, V_t, t) &= \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial v} dV_t + \frac{\partial f}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} < X_t > dt + \frac{1}{2}\frac{\partial^2 f}{\partial v^2} < V_t > dt + \frac{\partial^2 f}{\partial x v} < X_t, V_t > dt \\ &= \frac{\partial f}{\partial x}(-1/2 V_t dt + \sqrt{y_t} dW_t) + \frac{\partial f}{\partial v}(\kappa(\eta - V_t) + \theta\sqrt{V_t} d\hat{Z}_t) + \frac{\partial f}{\partial l} dt \\ &+ \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2} V_t + \frac{\partial^2 f}{\partial v^2}\theta^2 V_t + 2\frac{\partial^2 f}{\partial x v}\rho\theta V_t\right) dt \end{aligned}$$

By setting the drift equal to zero (for any x and v) we get the PDE

$$-\frac{\partial f}{\partial x} 1/2 v + \frac{\partial f}{\partial v}\kappa(\eta - v) + \frac{\partial f}{\partial t} + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2} v + \frac{\partial^2 f}{\partial v^2}\theta^2 v + 2\frac{\partial^2 f}{\partial x v}\rho\theta v\right) = 0 \ .$$

We also have a terminal condition by setting $t = T$ $f(x, v, T) = e^{iux}$. We solve the PDE by taking a guess on form of f:

$$f(x, u, t) = \exp(C(T - t) + vD(T - t) + iux)$$

By substiting in the PDE we get

$$\begin{aligned} 0 = &-f(x,u,t) iu \frac{1}{2} v + f(x,u,t) D(T-t)\kappa(\eta - v) - f(x,u,t)(C'(T-t) + vD'(T-t)) \\ &+ \frac{f(x,u,t)}{2}\left(-u^2 v + vD^2(T-t)\theta^2 + 2iuD(T-t)\rho\theta v\right) \end{aligned}$$

By simplifying for f and collecting the terms without v and linear in v we get the system of ODE

$$\begin{cases} D'(T-t) &= \frac{\theta^2}{2} D(T-t)^2 + (iu\rho\theta - k) D(T-t) - \frac{iu+u^2}{2} \\ C'(T-t) &= D(T-t) k\eta \end{cases}$$

and $D(0) = 0$, $C(0) = 0$. By calling $T - t =: s$ we get

$$\begin{cases} D'(s) &= \frac{\theta^2}{2} D(s)^2 + (iu\rho\theta - k\kappa) D(s) - \frac{iu+u^2}{2} \\ C'(s) &= D(s) k\eta \end{cases} ,$$

which is a two-dimensional system of Riccati equations. We can solve the first that does not depend from C.

$$D(s) = -\frac{u^2 + iu}{\gamma coth(\gamma s/2) + \kappa - i\rho\theta u} \ ,$$

where $\gamma := \sqrt{\theta^2(u^2 + iu) + (k - i\rho\theta u)^2}$. Now we can solve for C by integrating on both sides $C'(s) = D(s) k\eta$ and get

$$C(s) = \frac{k\eta s(k - i\rho\theta u)}{\theta^2} - \frac{2k\eta}{\theta^2} log\left(cosh(\gamma s/2) + \frac{k - i\rho\theta u}{\gamma} sinh(\gamma s/2)\right) \ .$$

The non analyticity strip in the characteristic function is due to the logarithm.

5.2 Bates model

Stochastic volatility model are unable to get the short time equity skew. Some improvement can be obtained using jumps. The Bates models

$$\begin{cases} dX_t &= (-1/2 V_t - \lambda) dt + \sqrt{V_t} dW_t + dJ_t \\ dV_t &= \kappa(\eta - V_t) dt + \theta\sqrt{V_t} d\hat{Z}_t \ , \end{cases}$$

with $V_0 = v, X_0 = 0$ and J_t a compound Poisson process independent from the BMs that is an exponential martingale and λ ensures the martingality.
To price with the Bates model we need to compute its characteristic function. To do so we separate the process in two parts the continuous one X_t^c such that $dX_t^C = (-1/2V_t)dt + \sqrt{V_t}dW_t$ and the discontinous part J_t. One can check that $X_t^C + J_t - \lambda t$ satisfies the Bates SDE. Then,

$$\mathbb{E}[e^{iuX_T}|\mathcal{F}_t] = \mathbb{E}[e^{iuX_T^C}|X_t^C = x, V_t = v]\mathbb{E}[e^{iuJ_T}|J_t = j]e^{-iu\lambda t} = f(x,v,t)e^{iuj-iu\lambda t}\hat{\mu}_{T-t}(u)$$

where $\hat{\mu}_{T-t}(u)$ is the characteristic function of the jump process. In general the ρ determine the skew at large maturities and the jumps the skew at short maturity. The main issue with Bates is that when calibrating on market data it is easy to find a lot of local minima.

5.3 Non-Gaussian OU

Definition 5.1. A non-Gaussian OU (or RCLL process in the defition of Benth *et al.* (2008)) $X(t)$, $t \leq s \leq T$ is the strong solution of the SDE

$$dX_t = (\mu_t - a_t X_t)dt + \sigma_t dI_t \ , \qquad X_s = x \ ,$$

where I_t is an additive process with triplet (A_t, γ_t, ν_t) and all function are real-valued on $[0,T]$.

Notice that a Gaussian OU is obtained when $\nu_t = 0$.

Proposition 5.2. *The unique strong solution of the non-Gaussian Ornstein-Uhlenbeck (OU) SDE is given by:*

$$X_t = x\exp\left(-\int_s^t a_r\,dr\right) + \int_s^t \mu_r \exp\left(-\int_r^t a_q\,dq\right)dr + \int_s^t \sigma_r \exp\left(-\int_r^t a_q\,dq\right)dI_r.$$

Proof. We restrict our proof to verifying that the expression above satisfies the OU SDE. For a complete proof of existence and uniqueness, see Benth *et al.* (2008), Proposition 3.1.
Let us define the process Z by:

$$Z_t := \exp\left(\int_s^t a_r\,dr\right)X_t.$$

Applying Itô's product rule to $f(t, X_t) = Z_t$, we obtain:

$$dZ_r = a_r Z_r\,dr + \exp\left(\int_s^r a_q\,dq\right)dX_r.$$

Substituting the dynamics of dX_r, the drift terms involving a_r cancel out, yielding:

$$dZ_r = \mu_r \exp\left(\int_s^r a_q\,dq\right)dt + \sigma_r \exp\left(\int_s^r a_q\,dq\right)dI_r.$$

Integrating both sides from s to t, we have:

$$Z_t - Z_s = \int_s^t \mu_r \exp\left(\int_s^r a_q\,dq\right)dr + \int_s^t \sigma_r \exp\left(\int_s^r a_q\,dq\right)dI_r.$$

Finally, multiplying by $\exp\left(-\int_s^t a_r\,dr\right)$ and noting that $Z_s = X_s = x$, we recover the expression for X_t:

$$X_t = x\exp\left(-\int_s^t a_r\,dr\right) + \int_s^t \mu_r \exp\left(-\int_r^t a_q\,dq\right)dr + \int_s^t \sigma_r \exp\left(-\int_r^t a_q\,dq\right)dI_r.$$

This completes the proof. □

Proposition 5.3. *The characteristic function of an OU process is*

$$\mathbb{E}[e^{iuX_t}|X_s = x] = \exp\left(iu\left\{xe^{-\int_s^t a_r dr} + \int_s^t \mu_r e^{-\int_r^t a_q dq} dr\right\} + \psi(s,t,u\sigma(.)e^{-\int_.^t a(r)dr})\right) \quad ,$$

where

$$\psi(s,t,g(.)) = i\int_s^t g(r)d\gamma_t - 1/2\int_s^t g(r)^2 dA_t + \int_s^t\int_{\mathbb{R}}(e^{ig(r)z} - 1 - ig(r)z\mathcal{I}_{|z|\leq 1})l(dz,dr) \quad ,$$

where $l(z,r) = \frac{\partial \nu_t(x)}{\partial(t)}$.

Proof. Given the strong solution of the OU SDE the deterministic parts comes out of the expected value. We focus on the stochastic part

$$\mathbb{E}\left[e^{iu\int_s^t \exp\left(-\int_r^t a_q dq\right)\sigma_r dI_r}\right] \quad .$$

We prove the equality for a generic continuous integrand function g

$$\mathbb{E}\left[e^{iu\int_s^t g(r)dI_r}\right] = \mathbb{E}\left[e^{\psi(s,t,ug(r))}\right] \quad .$$

Any continuous function on [s,t] can be approximated by a sequence of step function on a grid $s = t_0 < t_1 < ... < t_k = t$ $g^{(k)}(r) := \sum_{j=1}^k a_j \mathcal{I}_{r\in[t_{j-1},t_j)}$,

$$\mathbb{E}\left[e^{iu\int_s^t g^{(k)}(r)dI_r}\right] = \prod_{j=1}^k \mathbb{E}[e^{iua_{j-1}(I(t_j)-I(t_{j-1}))}] = e^{\sum_{j=1}^k \psi(t_{j-1},t_j,a_{j-1},u)} = e^{\psi(s,t,ug^{(k)}(.))}$$

where the first equality is due to the independence of increments, the second to the characteristic function of an increment of an additive process and the third to the definition of integral of a step function. By taking the limit for k we get the thesis for any function g and in particular, for $u\exp\left(-\int_r^t a_q dq\right)\sigma_r$.

□

5.4 The BNS model

The key point in this approach is the presence of jumps in the volatility correlated with jumps in the log-price.

$$\begin{cases} dX_t &= (\mu - 1/2\sigma_t^2)dt + \sigma_t dW_t + \rho dZ_t \\ d\sigma_t^2 &= -\lambda\sigma_t^2 dt + dZ_t \quad , \end{cases} \qquad \sigma_0^2 > 0; X_0 = 0 \quad ,$$

where Z_t is a Lévy subordinator without drift (b=0) independent from the BM. e variance process σ_t^2 follows a non-Gaussian Ornstein-Uhlenbeck (OU) equation. Its characteristic function is given by:

$$\mathbb{E}\left[e^{iu\sigma_t^2}\right] = \exp\left(iu\sigma_0^2 e^{-\lambda t} + \int_0^t\int_{\mathbb{R}^+}\left(e^{iuxe^{-\lambda r}} - 1\right)\nu(dx)\,dr\right) = \exp\left(iu\sigma_0^2 e^{-\lambda t} + \int_0^t \psi\left(ue^{-\lambda r}\right)dr\right),$$

where $\psi(\xi) = \int_{\mathbb{R}^+}(e^{i\xi x} - 1)\nu(dx)$ is the characteristic exponent of the driving Lévy subordinator Z_t. The explicit solution to the variance SDE is:

$$\sigma_t^2 = \sigma_0^2 e^{-\lambda t} + \int_0^t e^{-\lambda(t-r)}\,dZ_r.$$

To find the characteristic function of the log-price X_t, we condition on the path of the variance (and thus on Z).

$$\begin{aligned}\mathbb{E}[e^{iuX_t}] &= \mathbb{E}\left[\exp\left(iu\left(\mu t - \frac{1}{2}\int_0^t \sigma_r^2\,dr + \int_0^t \sigma_r\,dW_r + \rho Z_t\right)\right)\right] \\ &= e^{iu\mu t}\mathbb{E}\left[\exp\left(iu\rho Z_t - \frac{iu}{2}\int_0^t \sigma_r^2\,dr\right)\mathbb{E}\left[\exp\left(iu\int_0^t \sigma_r\,dW_r\right)\middle|\sigma, Z\right]\right].\end{aligned}$$

Using the standard result for the Itô integral of a deterministic function (conditional on σ), the inner expectation is $e^{-\frac{1}{2}u^2\int_0^t \sigma_r^2 dr}$:

$$\mathbb{E}[e^{iuX_t}] = e^{iu\mu t}\mathbb{E}\left[\exp\left(iu\rho Z_t - \frac{1}{2}(iu+u^2)\int_0^t \sigma_r^2\,dr\right)\right].$$

Lemma (Stochastic Fubini) To compute the integral of the variance, we substitute the solution for σ_r^2 and swap the order of integration:

$$\begin{aligned}\int_0^t \sigma_r^2\,dr &= \int_0^t \left(\sigma_0^2 e^{-\lambda r} + \int_0^r e^{-\lambda(r-q)}\,dZ_q\right)dr \\ &= \sigma_0^2\frac{1-e^{-\lambda t}}{\lambda} + \int_0^t\int_q^t e^{-\lambda(r-q)}\,dr\,dZ_q \\ &= \sigma_0^2\frac{1-e^{-\lambda t}}{\lambda} + \int_0^t \frac{1-e^{-\lambda(t-q)}}{\lambda}\,dZ_q.\end{aligned}$$

Substituting this back into the expectation:

$$\begin{aligned}\mathbb{E}[e^{iuX_t}] &= \exp\left(iu\mu t - \frac{1}{2}(iu+u^2)\sigma_0^2\frac{1-e^{-\lambda t}}{\lambda}\right) \\ &\quad\times\mathbb{E}\left[\exp\left(iu\rho\int_0^t dZ_q - \frac{1}{2}(iu+u^2)\int_0^t \frac{1-e^{-\lambda(t-q)}}{\lambda}\,dZ_q\right)\right] \\ &= \exp\left(iu\mu t - \frac{iu+u^2}{2\lambda}\sigma_0^2(1-e^{-\lambda t})\right) \\ &\quad\times\mathbb{E}\left[\exp\left(i\int_0^t\left[u\rho + \frac{i}{2}(iu+u^2)\frac{1-e^{-\lambda(t-q)}}{\lambda}\right]dZ_q\right)\right].\end{aligned}$$

Proceeding as in the proof of the Non Gaussian OU characteristic function $\mathbb{E}[\exp(i\int f(q)dZ_q)] = \exp(\int \psi(f(q))dq)$, and simplifying the complex term $\frac{i}{2}(iu+u^2) = -\frac{1}{2}(u-iu^2)$, we obtain the final result:

$$\mathbb{E}[e^{iuX_t}] = \exp\left(iu\mu t - \frac{iu+u^2}{2\lambda}\sigma_0^2(1-e^{-\lambda t}) + \int_0^t \psi\left(u\rho - \frac{1}{2}(u-iu^2)\frac{1-e^{-\lambda r}}{\lambda}\right)dr\right).$$

Problem, ρ controls both the jumps in the smile and the correlation spot/volatility. This is a problem because the jumps determine the short time skew and the correlation the skew for all maturity.

References

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

Azzone, M. and Baviera, R., 2022. Additive normal tempered stable processes for equity derivatives and power-law scaling, *Quantitative Finance*, 22 (3), 501–518.

Azzone, M. and Baviera, R., 2023. A fast monte carlo scheme for additive processes and option pricing, *Computational Management Science*, 20 (1), 31.

Azzone, M. and Baviera, R., 2024. Short-time implied volatility of additive normal tempered stable processes, *Annals of Operations Research*, 336 (1), 93–126.

Benth, F.E., Benth, J.S., and Koekebakker, S., 2008. *Stochastic modelling of electricity and related markets*, vol. 11, World Scientific.

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

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

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

Taleb, N.N., 1997. *Dynamic hedging: managing vanilla and exotic options*, John Wiley & Sons.