Abstract

The goal of this paper is to show that the jump-diffusion models are an essential and easy-to-learn tool for option pricing and risk management, and that they provide an adequate description of stock price fluctuations and market risks. We try to give an overview of the field without focusing on technical details. After introducing several widely used jump-diffusion models, we discuss Fourier transform based methods for European option pricing, partial differential equations for barrier and American options, and the existing approaches to calibration and hedging.

1 Introduction

Starting with Merton’s seminal paper [21] and up to the present date, various aspects of jump-diffusion models have been studied in the academic finance community (see [8] for a list of almost 400 references on the subject). In the last decade, also the research departments of major banks started to accept jump-diffusions as a valuable tool in their day-to-day modeling. This increasing interest to jump models in finance is mainly due to the following reasons.

First, in a model with continuous paths like a diffusion model, the price process behaves locally like a Brownian motion and the probability that the stock moves by a large amount over a short period of time is very small, unless one fixes an unrealistically high value of volatility. Therefore, in such models the prices of short term out of the money options should be much lower than what one observes in real markets. On the other hand, if stock prices are allowed to jump, even when the time to maturity is very short, there is a non-negligible probability that after a sudden change in the stock price the option will move in the money.

Second, from the point of view of hedging, continuous models of stock price behavior generally lead to a complete market or to a market, which can be made complete by adding one or two additional instruments, like in stochastic volatility models. Since in such a market every terminal payoff can be exactly replicated, options are redundant assets, and the very existence of traded options becomes a puzzle. The mystery is easily solved by allowing for discontinuities:
in real markets, due to the presence of jumps in the prices, perfect hedging is impossible and options enable the market participants to hedge risks that cannot be hedged by using the underlying only.

From a risk management perspective, jumps allow to quantify and take into account the risk of strong stock price movements over short time intervals, which appears non-existent in the diffusion framework. To be more specific, let us give an example from the domain of portfolio management. The constant proportion portfolio insurance strategy consists in holding a proportion \( x_t \) of the risky asset in the portfolio, where \( x_t \) is given by

\[
x_t = m \frac{V_t - F_t}{V_t},
\]

where \( V_t \) is the portfolio value, \( F_t \) is the ‘floor’, i.e., the ‘insured’ lower bound on the portfolio value, and \( m \) is a constant multiplier. When the portfolio value approaches the lower bound, the proportion of risky asset tends to zero. In a continuous-path model with frequent trading, the portfolio will therefore never go below the barrier \( F_t \). Taking a large multiplier, one can then construct a portfolio with a very important upside potential and almost no downside risk. However, this illusion breaks down as soon as one takes into account the jump risk: there is always a non-zero probability that due to a sudden downward jump in the risky asset price, the investor will not have a chance to withdraw before the portfolio value drops below \( F_t \).

The last and probably the strongest argument for using discontinuous models is simply the presence of jumps in observed prices. Figure 1 depicts the evolution of the DM/USD exchange rate over a two-week period in 1992, and one can see at least three points where the rate moved by over 100 bp within a 5-minute period. Price moves like these ones clearly cannot be accounted for in a diffusion model with continuous paths, but they must be dealt with if the market risk is to be measured and managed correctly.
In this paper we give a brief introduction to jump-diffusion models and review various mathematical and numerical tools needed to use these models for option pricing and hedging. Since we are focusing on explanations rather than technical details, no proofs are given, but the reader will always be able to find complete proofs in the references we provide.

The rest of this paper is structured as follows. In section 2 we provide a brief mathematical introduction to jump diffusions and define several important parametric and non-parametric classes. Section 3 discusses the Fourier-transform methods for European option pricing, based on the explicit knowledge of the characteristic function in many jump-diffusion models. Section 4 discusses the partial integro-differential equations which play the role of the Black-Scholes equation in jump-diffusion models and can be used to value American and barrier options. Finally, section 5 discusses hedging in presence of jumps and section 6 explains how jump-diffusion models can be calibrated to market data.

2 A primer on jump-diffusion models

The two basic building blocks of every jump-diffusion model are the Brownian motion (the diffusion part) and the Poisson process (the jump part). The Brownian motion is a familiar object to every option trader since the appearance of the Black-Scholes model, but a few words about the Poisson process are in order. The proofs of the statements below can be found in [8, chapter 2].

The Poisson process Take a sequence \( \{\tau_i\}_{i \geq 1} \) of independent exponential random variables with parameter \( \lambda \), that is, with cumulative distribution function \( P[\tau_i \geq y] = e^{-\lambda y} \) and let \( T_n = \sum_{i=1}^n \tau_i \). The process

\[
N_t = \sum_{n \geq 1} 1_{t \geq T_n}
\]

is called the Poisson process with parameter \( \lambda \). For example, if the waiting times between buses at a bus stop are exponentially distributed, the total number of buses arrived up to time \( t \) is a Poisson process. The trajectories of a Poisson process are piecewise constant (right-continuous with left limits or RCLL), with jumps of size 1 only. The jumps occur at times \( T_i \) and the intervals between jumps (the waiting times) are exponentially distributed. At every date \( t > 0 \), \( N_t \) has the Poisson distribution with parameter \( \lambda t \), that is, it is integer-valued and

\[
P[N_t = n] = e^{-\lambda t} \frac{\lambda^t}{n!}.
\]

(1)

The Poisson process shares with the Brownian motion the very important property of independence and stationarity of increments, that is, for every \( t > s \) the increment \( N_t - N_s \) is independent from the history of the process up to time \( s \) and has the same law as \( N_{t-s} \). The processes with independent and stationary increments are called Lévy processes after the French mathematician Paul Lévy.
Characteristic function  The notion of characteristic function of a random variable plays an essential role in the study of jump-diffusion processes: often we do not know the distribution function of such a process in closed form but the characteristic function is known explicitly. The characteristic function of a random variable $X$ is defined by

$$\phi_X(u) \equiv E[e^{iuX}].$$

For the Poisson process, this gives

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

Here, the computation can be done directly using equation (1).

Compound Poisson process  For financial applications, it is of little interest to have a process with a single possible jump size. The compound Poisson process is a generalization where the waiting times between jumps are exponential but the jump sizes can have an arbitrary distribution. More precisely, let $N$ be a Poisson process with parameter $\lambda$ and $\{Y_i\}_{i \geq 1}$ be a sequence of independent random variables with law $f$. The process

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

is called compound Poisson process. Its trajectories are RCLL and piecewise constant but the jump sizes are now random with law $f$. The compound Poisson process has independent and stationary increments. Its law at a given time $t$ is not known explicitly but the characteristic function is known and has the form

$$E[e^{iuX_t}] = \exp\{t\lambda \int_{\mathbb{R}} (e^{iux} - 1)f(dx)\}.$$ 

Simulation of compound Poisson process  Contrary to more complex jump processes, the compound Poisson process is easy to simulate. The algorithm is based on the following fact [8, chapter 2]:

**Fact 1** Conditionally on $N_T = n$, the jump times $T_1, \ldots, T_n$ of a Poisson process on the interval $[0, T]$ are distributed as $n$ independent ordered uniforms on $[0, T]$.

This leads to the following algorithm:

- Simulate $N_T$ from the Poisson distribution\(^1\) with parameter $\lambda T$.

\(^1\)The random number generator from the Poisson distribution is available in most MATLAB-like scientific computing environments. If you need to implement it, see [13].
• Simulate $N_T$ uniform random variables \( \{U_i\}_{i=1}^{N_T} \) on \([0,T]\).

• Simulate $N_T$ independent variables \( \{Y_i\}_{i=1}^{N_T} \) with law $f$.

• The process is given by

\[
X_t = \sum_{i=1}^{N_T} Y_i U_i \mathbf{1}_{U_i \leq t}.
\]

### Jump-diffusions and Lévy processes

Combining a Brownian motion with drift and a compound Poisson process, we obtain the simplest case of a jump-diffusion — a process which sometimes jumps and has a continuous but random evolution between the jump times (cf. Fig. 2):

\[
X_t = \mu t + \sigma B_t + \sum_{i=1}^{N_t} Y_i.
\]  

(2)

The best known model of this type in finance is the Merton model [21], where the stock price is $S_t = S_0 e^{X_t}$ with $X_t$ as above and the jumps \( \{Y_i\} \) have Gaussian distribution.

The process (2) is again a Lévy process and its characteristic function can be computed by multiplying the CF of the Brownian motion and that of the compound Poisson process (since the two parts are independent):

\[
E[e^{iux_t}] = \exp \left\{ t \left( i\mu u - \frac{\sigma^2 u^2}{2} + \lambda \int_{\mathbb{R}} (e^{iux} - 1) f(dx) \right) \right\}.
\]

The class of Lévy processes is not limited to jump-diffusions of the form (2): there exist Lévy processes with infinitely many jumps in every interval. Most
of such jumps are very small and there is only a finite number of jumps with absolute value greater than any given positive number. One of the simplest examples of this kind is the gamma process, a process with independent and stationary increments such that for all $t$, the law $p_t$ of $X_t$ is the gamma law with parameters $\lambda$ and $ct$:

$$
p_t(x) = \frac{\lambda^{ct}}{\Gamma(ct)} e^{-\lambda x} x^{ct-1}.
$$

The gamma process is an increasing Lévy process (also called subordinator). Its characteristic function has a very simple form:

$$
E[e^{iuX_t}] = (1 - iu/\lambda)^{-ct}.
$$

The gamma process is the building block for a very popular jump model, the variance gamma process [20, 19], which is constructed by taking a Brownian motion with drift and changing its time scale with a gamma process:

$$
Y_t = \mu X_t + \sigma B_t.
$$

Using $Y_t$ to model the logarithm of stock prices can be justified by saying that the price is a geometric Brownian motion if viewed on a stochastic time scale given by the gamma process [16]. The variance gamma process is another example of a Lévy process with infinitely many jumps and has characteristic function

$$
E[e^{iuY_t}] = \left(1 + \frac{\sigma^2 u^2}{2} - i\mu\kappa u\right)^{-\kappa t}.
$$

The parameters have the following (approximate) interpretation: $\sigma$ is the variance parameter, $\mu$ is the skewness parameter and $\kappa$ is responsible for the kurtosis of the process.

In general, every Lévy process can be represented in the form

$$
X_t = \gamma t + \sigma B_t + Z_t,
$$

where $Z_t$ is a jump process with (possibly) infinitely many jumps. A detailed description of this component is given by the Lévy-Itô decomposition which is beyond the scope of this introductory paper. The characteristic function of a Lévy process is given by the Lévy-Khintchine formula:

$$
E[e^{iuX_t}] = \exp \left\{ t \left( i\gamma u - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}} \left( e^{iux} - 1 - iux 1_{|x|\leq 1}\right) \nu(dx) \right) \right\},
$$

where $\nu$ is a positive measure on $\mathbb{R}$ describing the jumps of the process: the Lévy measure. If $X$ is compound Poisson, then $\nu(\mathbb{R}) < \infty$ and $\nu(dx) = \lambda f(dx)$ but in the general case $\nu$ need not be a finite measure. It must satisfy the constraint

$$
\int_{\mathbb{R}} (1 \land x^2) \nu(dx) < \infty
$$
and describes the jumps of $X$ in the following sense: for every closed set $A \subset \mathbb{R}$ with $0 \not\in A$, $\nu(A)$ is the average number of jumps of $X$ in the time interval $[0, 1]$, whose sizes fall in $A$.

To keep the discussion simple, in the rest of this paper we will only consider Lévy jump-diffusions, that is, Lévy processes with finite jump intensity of the form (2), but with the new notation $\nu(dx) = \lambda f(dx)$ for the Lévy measure. The characteristic function of such a process therefore takes the form

$$E[e^{iuX_t}] = \exp \left\{ t \left( i\mu u - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}} (e^{iux} - 1)\nu(dx) \right) \right\}. \quad (3)$$

**Exponential Lévy models** To ensure positivity as well as the independence and stationarity of log-returns, stock prices are usually modeled as exponentials of Lévy processes:

$$S_t = S_0 e^{X_t}. \quad (4)$$

In the jump-diffusion case this gives

$$S_t = S_0 \exp \left( \mu t + \sigma B_t + \sum_{i=1}^{N_t} Y_i \right).$$

Between the jumps, the process evolves like a geometric Brownian motion, and after each jump, the value of $S_t$ is multiplied by $e^{Y_i}$. This model can therefore be seen as a generalization of the Black-Scholes model:

$$dS_t = S_t \left( \tilde{\mu} dt + \sigma dB_t + dJ_t \right). \quad (5)$$

Here, $J_t$ is a compound Poisson process such that the $i$-th jump of $J$ is equal to $e^{Y_i} - 1$. For instance, if $Y_i$ have Gaussian distribution, $S$ will have lognormally distributed jumps. The notation $S_{t-}$ means that whenever there is a jump, the value of the process before the jump is used on the left-hand side of the formula. The forms (4) and (5) are equivalent: for a model of the first kind one can always find a model of the second kind with the same law. In the rest of the paper, unless explicitly stated otherwise, we will use the exponential form (4).

For option pricing, we will explicitly include the interest rate into the definition of the exponential Lévy model:

$$S_t = S_0 e^{rt + X_t}. \quad (6)$$

While the forms (4) and (6) are equivalent, the second one leads to a slightly simpler notation. In this case, under the risk-neutral probability, $e^{X_t}$ must be a martingale and from the Lévy-Khintchine formula (3) combined with the independent increments property we conclude that this is the case iff

$$b + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^x - 1)\nu(dx) = 0. \quad (7)$$
The model (6) admits no arbitrage opportunity if there exists an equivalent probability under which $e^{X_t}$ is a martingale. For Lévy processes it can be shown that this is almost always the case, namely an exponential Lévy model is arbitrage-free if and only if the trajectories of $X$ are not almost surely increasing nor almost surely decreasing.

If a Brownian component is present, the martingale probability can be obtained by changing the drift as in the Black-Scholes setting. Otherwise, in finite-intensity models, the drift must remain fixed under all equivalent probabilities since it can be observed from a single stock price trajectory. To satisfy the martingale constraint (7), one must therefore change the Lévy measure, i.e. the intensity of jumps. To understand how this works, suppose that $X$ is a Poisson process with drift:

$$X_t = N_t - at, \quad a > 0.$$  

We can obtain a martingale probability by changing the intensity of $N$ to $\lambda_{\text{mart}} = \frac{a}{e-1}$. If, however, $X$ is a Poisson process without drift (increasing trajectories), one cannot find a value of $\lambda > 0$ for which $e^{X_t}$ is a martingale.

**Beyond Lévy processes** Although the class of Lévy processes is quite rich, it is sometimes insufficient for multi-period financial modeling for the following reasons:

- Due to the stationarity of increments, the stock price returns for a fixed time horizon always have the same law. It is therefore impossible to incorporate any kind of new market information into the return distribution.

- For a Lévy process, the law of $X_t$ for any given time horizon $t$ is completely determined by the law of $X_1$. Therefore, moments and cumulants depend on time in a well-defined manner which does not always coincide with the empirically observed time dependence of these quantities [3].

For these reasons, several models combining jumps and stochastic volatility appeared in the literature. In the Bates [2] model, one of the most popular examples of the class, an independent jump component is added to the Heston stochastic volatility model:

$$dX_t = \mu dt + \sqrt{V_t}dW^X_t + dZ_t, \quad S_t = S_0 e^{X_t},$$  

$$dV_t = \xi(\eta - V_t)dt + \theta \sqrt{V_t}dW^V_t, \quad d\langle W^V, W^X \rangle_t = \rho dt,$$  

where $Z$ is a compound Poisson process with Gaussian jumps. Although $X_t$ is no longer a Lévy process, its characteristic function is known in closed form [8, chapter 15] and the pricing and calibration procedures are similar to those used for Lévy processes.
3 Pricing European options via Fourier transform

In the Black-Scholes setting, the prices of European calls and puts are given explicitly by the Black-Scholes formula. In the case of Lévy jump-diffusions, closed formulas are no longer available but a fast deterministic algorithm, based on Fourier transform, was proposed by Carr and Madan [7]. Here we present a slightly improved version of their method, due to [22, 8].

Let \( \{X_t\}_{t \geq 0} \) be a Lévy process and, for simplicity, take \( S_0 = 1 \). We would like to compute the price of a European call with strike \( K \) and maturity \( T \) in the exponential Lévy model (6). Denote \( k = \log K \) the logarithm of the strike.

To compute the price of a call option

\[
C(k) = e^{-rT}E[(e^{rT}+X_T - e^k)^+],
\]

we would like to express its Fourier transform in log strike in terms of the characteristic function \( \Phi_T(v) \) of \( X_T \) and then find the prices for a range of strikes by Fourier inversion. However we cannot do this directly because \( C(k) \) is not integrable (it tends to 1 as \( k \) goes to \(-\infty\)). The idea is to subtract the Black-Scholes call price with non-zero volatility and compute the Fourier transform of the resulting function which is integrable and smooth:\(^2\)

\[
z_T(k) = C(k) - C^\Sigma_{BS}(k),
\]

where \( C^\Sigma_{BS}(k) \) is the Black-Scholes price of a call option with volatility \( \Sigma \) and log-strike \( k \) for the same underlying value and the same interest rate.

Proposition 1 Let \( \{X_t\}_{t \geq 0} \) be a real-valued Lévy process such that \((e^{X_t})\) is a martingale, and

\[
\int_{x>1} e^{(1+\alpha)x}\nu(dx) < \infty
\]

for some \( \alpha > 0 \). Then the Fourier transform in log-strike \( k \) of \( z_T(k) \) is given by:

\[
\zeta_T(v) = e^{ivrT}\Phi_T(v-i) - \Phi^\Sigma_T(v-i)
\]

where \( \Phi^\Sigma_T(v) = \exp(-\frac{\Sigma^2T}{2}(v^2 + iv)) \) is the characteristic function of log-stock price in the Black-Scholes model.

The optimal value of \( \Sigma \) is the value for which \( \zeta_T(0) = 0 \). However, the convergence is good for any \( \Sigma > 0 \). One can take for example \( \Sigma = 0.2 \) for practical calculations.

\(^2\)Carr and Madan proposed to subtract the (non-differentiable) intrinsic value of the price \((1 - e^{k-rT})^+\) but this leads to a slower convergence.
Numerical Fourier inversion  Option prices can be computed by evaluating numerically the inverse Fourier transform of $\zeta_T$:

$$z_T(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ivk} \zeta_T(v) dv.$$  \hfill (9)

This integral can be efficiently computed for a range of strikes using the Fast Fourier Transform. Recall that this algorithm allows to calculate the discrete Fourier transform $\text{DFT}[f][n]$, defined by,

$$\text{DFT}[f][n] := \sum_{k=0}^{N-1} f_k e^{-2\pi ink/N}, \quad n = 0 \ldots N - 1,$$

using only $O(N \log N)$ operations.

To approximate option prices, we truncate and discretize the integral (9) as follows:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \zeta_T(v) dv = \frac{1}{2\pi} \int_{-L/2}^{L/2} e^{-ivk} \zeta_T(v) dv + \varepsilon_{\text{trunc}}$$

$$= \frac{L}{2\pi(N-1)} \sum_{m=0}^{N-1} w_m \zeta_T(v_m) e^{-iv_m} + \varepsilon_{\text{trunc}} + \varepsilon_{\text{discr}},$$

where $\varepsilon_{\text{trunc}}$ is the truncation error, $\varepsilon_{\text{discr}}$ is the discretization error, $v_m = -L/2 + m\Delta$, $\Delta = L/(N - 1)$ is the discretization step and $w_m$ are weights, corresponding to the chosen integration rule (for instance, for the Simpson’s rule $w_0 = 1/3$, and for $k = 1, \ldots, N/2$, $w_{2k-1} = 4/3$ and $w_{2k} = 2/3$).\footnote{We use the FFT with $N = 2^p$, so $N$ is even.}


\[
\Phi_T(u) = \mathbb{E}[e^{iuX_T}]
\]

Table 1: Examples of characteristic functions of jump-diffusion processes used in financial modeling. For further examples, see [8, chapter 4].

Choosing \( k_n = k_0 + \frac{2\pi n}{N} \) we see that the sum in the last term becomes a discrete Fourier transform:

\[
\frac{L}{2\pi(N-1)} e^{ik_nL/2} \sum_{m=0}^{N-1} w_m \zeta_T(k_m) e^{-ik_0m\Delta} e^{-2\pi inm/N} = \frac{L}{2\pi(N-1)} e^{ik_nL/2} \text{DFT}_n[w_m \zeta_T(k_m) e^{-ik_0m\Delta}]
\]

Therefore, the FFT algorithm allows to compute \( z_T \) and option prices for the log strikes \( k_n = k_0 + \frac{2\pi n}{N} \). The log strikes are thus equidistant with the step \( d \) satisfying

\[
d\Delta = \frac{2\pi}{N}.
\]

This relationship implies that if we want to compute option prices on a fine grid of strikes, and at the same time keep the discretization error low, we must use a large number of points.

This method applies to all models where the characteristic function of log-stock price is known or easy to compute. This is the case for exponential Lévy models (see, e.g., Table 1) but also holds for a more general class of affine processes [14, 15], which includes in particular the Bates model mentioned in section 2.

4 Integro-differential equations for barriers and American options

The Fourier-transform based algorithm of the preceding section is very efficient for European vanilla options, but does not apply to more complicated contracts with barriers or American-style exercise\(^4\). In diffusion models their prices are usually expressed as solutions of the Black-Scholes partial differential equation

\[
\frac{\partial P}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC - rS \frac{\partial C}{\partial S}
\]

\(^4\)Fourier-transform based methods for pricing single-barrier options can be found in the literature [24, 5, 18] but except for some particular models [17], the numerical complexity of the resulting formulae is prohibitive.
with appropriate boundary conditions. In this section, we show how this method can be generalized to models with jumps by introducing partial integro-differential equations (PIDEs). A complete presentation with proofs, as well as the general case of possibly infinite Lévy measure, can be found in [23, 12].

Barrier “out” options We start with up-and-out, down-and-out, and double barrier options which have, respectively, an upper barrier \( U > S_0 \), a lower barrier \( L < S_0 \), or both of them. If the stock price \( S_t \) has not crossed any of the barriers before maturity \( T \), then the payoff of the option is \( H(S_T) \); otherwise, the option expires worthless or pays out a rebate \( G(\tau^*, S_{\tau^*}) \) where \( \tau^* \) is the moment when the stock price first touches the barrier (usually, the rebate is simply a constant amount).

The barrier options are said to be weakly path dependent, because at any given time \( t \), their price does not depend on the entire trajectory of the stock price prior to \( t \) but only on the current value \( S_t \) and on the event \( \{ t < \tau^* \} \), that is, on the information, whether the barrier has already been crossed. If the price of a barrier option is denoted by \( C_t \) then \( C_t = C_b(t, S_t) \) where \( C_b \) is a deterministic function, which satisfies a generalized Black-Scholes equation given below.

To obtain an equation with constant coefficients we switch to log-prices and denote:

- \( \tau = T - t \) (time to maturity), \( x = \log(S/S_0) \) (log-price),
- \( l = \log(L/S_0) \), \( u = \log(U/S_0) \) (barriers in terms of log-price),
- \( h(x) = H(S_0 e^x) \) (payoff function after the change of variables),
- \( g(\tau, x) = e^{\tau r} G(T - \tau, S_0 e^x) \) (rebate after the change of variables),
- \( v(\tau, x) = e^{\tau r} C_b(T - \tau, S_0 e^x) \) (option’s forward price).

Then the transformed option price \( v(\tau, x) \) satisfies

\[
\frac{\partial v}{\partial \tau}(\tau, x) = L v(\tau, x), \quad (\tau, x) \in (0, T] \times (l, u),
\]

\[
v(0, x) = h(x), \quad x \in (l, u),
\]

\[
v(\tau, x) = g(\tau, x), \quad \tau \in [0, T], \quad x \notin (l, u),
\]

where \( L \) is an integro-differential operator:

\[
Lf(x) = \frac{\sigma^2}{2} f''(x) - \left( \frac{\sigma^2}{2} - r \right) f'(x) + \int_{\mathbb{R}} \nu(dy) f(x + y) - \lambda f(x) - \alpha f'(x),
\]

with \( \lambda = \int_{\mathbb{R}} \nu(dy), \ \alpha = \int_{\mathbb{R}} (e^y - 1) \nu(dy) \). By convention, we set \( l = -\infty \) if there is no lower barrier and \( u = \infty \) if there is no upper barrier. So, (11)-(13) covers all types of barrier options above, as well as the European vanilla case.
In the case of the Black-Scholes model \( \nu \equiv 0 \), equation (11)–(13) is nothing more than the standard heat equation

\[
\frac{\partial v}{\partial \tau} = \frac{\sigma^2}{2} \frac{\partial^2 v}{\partial x^2} - \left( \frac{\sigma^2}{2} - r \right) \frac{\partial v}{\partial x},
\]

which can be obtained from the Black-Scholes equation (10) by an exponential change of variable.

Note, that (13) is different from usual boundary conditions for differential equations: it gives the values of the solution not only at the barriers but also beyond the barriers. It is an important consequence of the non-local character of the operator \( L \) due to the integral part.

**Numerical solution of the integro-differential equation** To solve numerically the problem (11)–(13), we proceed with the following steps:

- **Truncation of large jumps.** This corresponds to truncating the integration domain in (14).
- **Localization.** If the problem was initially stated on an unbounded interval (as in the European or one-barrier cases), we must choose a bounded computational domain and, consequently, impose artificial boundary conditions.
- **Discretization.** The derivatives of the solution are replaced by usual finite differences and the integral terms are approximated using the trapezoidal rule. The problem is then solved using an explicit-implicit scheme.

Let us now consider these steps in detail.

**Truncation of large jumps** Since we cannot calculate numerically an integral on the infinite range \((-\infty, \infty)\), the domain is truncated to a bounded interval \((B_l, B_r)\). In terms of the process, this corresponds to removing the large jumps. Usually, the tails of \( \nu \) decrease exponentially, so the probability of large jumps is very small. Therefore, we don’t change much the solution by truncating the tails of \( \nu \).

**Localization** Similarly, for the computational purposes, the domain of definition of the equation has to be bounded. For barrier options, the barriers are the natural limits for this domain and the rebate is the natural boundary condition. In absence of barriers, we have to choose artificial bounds \((-A_l, A_r)\) and impose artificial boundary conditions. Recall that “boundary” conditions in this case must extend the solution beyond the bounds as well: \( v(\tau, x) = g(\tau, x) \) for all \( x \notin (-A_l, A_r), \tau \in [0, T] \).

In [23], it is shown that a good choice for the boundary conditions is \( g(\tau, x) = h(x + \tau r) \) where \( h \) is the payoff function. For example, for a put option, we have \( h(x) = (K - S_0 e^x)^+ \) and thus \( g(\tau, x) = (K - S_0 e^{x + \tau r})^+ \).

In the case of one barrier, we need this boundary condition only on one side of the domain: the other is zero or given by the rebate.
Discretization: We consider now the localized problem on \((-A_l, A_r)\):

\[
\begin{align*}
\frac{\partial v}{\partial \tau} &= Lv, \\
v(0, x) &= h(x), \\
v(\tau, x) &= g(\tau, x), \\
\end{align*}
\]

on \((0, T] \times (-A_l, A_r)\) (15)

\[
\begin{align*}
v_n i &= v(\tau_n, x_i), \\
v_n i &= g(\tau_n, x_i), \\
\end{align*}
\]

where \(L\) is the following integro-differential operator:

\[
Lv = \frac{\sigma^2}{2} \frac{\partial^2 v}{\partial x^2} - \left( \frac{\sigma^2}{2} - r \right) \frac{\partial v}{\partial x} + \int_{B_l}^{B_r} \nu(dy) v(\tau, x + y) - \lambda v - \alpha \frac{\partial v}{\partial x},
\]

with \(\lambda = \int_{B_l}^{B_r} \nu(dy), \quad \alpha = \int_{B_l}^{B_r} (e^y - 1) \nu(dy)\). Let us introduce a uniform grid on \([0, T] \times \mathbb{R}\):

\[
\tau_n = n\Delta t, \quad n = 0 \ldots M, \quad x_i = -A_l + i\Delta x, \quad i \in \mathbb{Z},
\]

with \(\Delta t = T/M, \Delta x = (A_r + A_l)/N\). The values of \(v\) on this grid are denoted by \(\{v_n^i\}\). The space derivatives of \(v\) are approximated by finite differences:

\[
\begin{align*}
\left( \frac{\partial^2 v}{\partial x^2} \right)_i &\approx \frac{v_{i+1} - 2v_i + v_{i-1}}{(\Delta x)^2}, \\
\left( \frac{\partial v}{\partial x} \right)_i &\approx \frac{v_{i+1} - v_i}{\Delta x}, \quad \text{or} \quad \left( \frac{\partial v}{\partial x} \right)_i \approx \frac{v_{i-1} - v_i}{\Delta x}. 
\end{align*}
\]

The choice of the approximation of the first order derivative — forward or backward difference — depends on the parameters \(\sigma, r, \alpha\) (see below).

To approximate the integral term, we use the trapezoidal rule with the same discretization step \(\Delta x\). Choose integers \(K_l, K_r\) such that \([B_l, B_r]\) is contained in \([K_l - 1/2] \Delta x, (K_r + 1/2) \Delta x\) (Fig. 4). Then,

\[
\int_{B_l}^{B_r} \nu(dy) v(\tau, x_i + y) \approx \sum_{j=K_l}^{K_r} \nu_j v_{i+j}, \quad \text{where} \quad \nu_j = \int_{(j-1/2)\Delta x}^{(j+1/2)\Delta x} \nu(dy). \quad (20)
\]

Using (18)–(20) we obtain an approximation for \(Lv \approx D_{\Delta v} + J_{\Delta v}\), where \(D_{\Delta v}\) and \(J_{\Delta v}\) are chosen as follows.
Explicit-Implicit Scheme  Without loss of generality, suppose that $\sigma^2/2 - r < 0$. Then

\[
(D\Delta v)_i = \frac{\sigma^2}{2} \frac{v_{i+1} - 2v_i + v_{i-1}}{(\Delta x)^2} - \left(\frac{\sigma^2}{2} - r\right) \frac{v_{i+1} - v_i}{\Delta x}.
\]

If $\sigma^2/2 - r > 0$, to ensure the stability of the algorithm, we must change the discretization of $\partial v/\partial x$ by choosing the backward difference instead of the forward one. Similarly, if $\alpha < 0$ we discretize $J$ as follows:

\[
(J\Delta v)_i = \sum_{j=K}^{K_r} \nu_j v_{i+j} - \lambda v_i - \alpha \frac{v_{i+1} - v_i}{\Delta x}.
\] (21)

Otherwise, we change the approximation of the first derivative. Finally, we replace the problem (15)--(17) with the following explicit-implicit scheme:

Initialization:

\[
v_0^i = h(x_i), \quad \text{if } i \in \{0, \ldots, N-1\},
\]

\[
v_0^i = g(0, x_i), \quad \text{otherwise.}
\] (22) (23)

For $n = 0, \ldots, M-1$:

\[
\frac{v^{n+1}_i - v^n_i}{\Delta t} = (D\Delta v^{n+1})_i + (J\Delta v^n)_i, \quad \text{if } i \in \{0, \ldots, N-1\}
\]

\[
v^{n+1}_i = g((n+1)\Delta t, x_i), \quad \text{otherwise.}
\] (24) (25)

Here, the non-local operator $J$ is treated explicitly to avoid the inversion of the dense matrix $J\Delta$, while the differential part $D$ is treated implicitly. At each time step, we first evaluate vector $J\Delta v^n$ where $v^n$ is known from the previous iteration\(^5\), and then solve the tridiagonal system (24) for $v^{n+1} = (v^{n+1}_0, \ldots, v^{n+1}_{N-1})$.

This scheme is stable if

\[
\Delta t < \frac{\Delta x}{|\alpha| + \lambda \Delta x}.
\]

Pricing American options  The simplest way to adapt the above method to pricing American options is to use the dynamic programming. If we approximate continuous time by a discrete grid of exercise dates $t_n = n\Delta t$, the value of the American option at $t_n$ is the maximum between profits from exercising immediately and holding the option until $t_{n+1}$:

\[
V_n = \max\{H(S_{t_n}), V_{n+1}^c\},
\] (26)

where $V_n^c = e^{-r\Delta t} \mathbb{E}[V_{n+1}^c | F_{t_n}]$ may be interpreted as the value of a European option with payoff $V_{n+1}$ and maturity $t_{n+1}$. Therefore, at each time step, we can compute $V_n^c$ as above and then adjust the result by taking the maximum as in (26).

---

\(^5\)The particular form of the sum in (21) (discrete convolution of two vectors) allows to compute it efficiently and simultaneously for all $i$ using Fast Fourier transform.
More precisely, after the same change of variables, localization and discretization procedures, we end up with the following scheme:

Initialization:
\[ v_0^i = h(x_i), \quad \text{for all } i, \]

For \( n = 0, \ldots, M - 1 \):
\[ \frac{\tilde{v}^{n+1} - v_n^i}{\Delta t} = (D_{\Delta} \tilde{v}^{n+1})_i + (J_{\Delta} v^n)_i, \quad \text{if } i \in \{0, \ldots, N-1\} \]
\[ \tilde{v}^{n+1}_i = g((n+1)\Delta t, x_i), \quad \text{otherwise;} \]
\[ v^{n+1}_i = \max\{h(x_i), \tilde{v}^{n+1}_i\}, \quad \text{for all } i. \]

A slightly different approach to pricing American puts in jump-diffusion models, also based on the explicit-implicit scheme (22)–(25), is described in [1].

5 Hedging the jump risk

In the Black-Scholes model, the delta-hedging strategy is known to completely eliminate the risk of an option position. This strategy consists in holding the amount of stock equal to \( \frac{\partial C}{\partial S} \), the sensitivity of the option price with respect to the underlying. However, in presence of jumps, delta-hedging is no longer optimal. Suppose that a portfolio contains \( \phi_t \) stock, with price \( S_t \), and a short option position. After a jump \( \Delta S_t \), the change in the stock position is \( \phi_t \Delta S_t \), and the option changes by \( C(t, S_t + \Delta S_t) - C(t, S_t) \). The jump will be completely hedged if and only if
\[ \phi_t = \frac{C(t, S_t + \Delta S_t) - C(t, S_t)}{\Delta S_t}. \]

Since the option price is a nonlinear function of \( S \), \( \phi_t \neq \frac{\partial C}{\partial S} \) and delta-hedging does not offset the jump risk completely. This is illustrated in figure 5 where a single 7% jump in the stock price leads to an important residual hedging error.

Thus, to hedge a jump of a given size, one should use the sensitivity to movements of the underlying of this size rather than the sensitivity to infinitesimal movements. Since typically the jump size is not known in advance, the risk associated to jumps cannot be hedged away completely: we are in an incomplete market. In this setting, the hedging becomes an approximation problem: instead of replicating an option, one tries to minimize the residual hedging error. Empirical studies show that strategies using only stock lead to high levels of residual risk, and to obtain realistic hedges, liquid options should be added to the hedging portfolio (gamma-hedging).

In this section we show how to compute optimal hedging strategies in presence of jumps. First, we treat the case when the hedging portfolio contains only stock and the risk-free asset. Let \( S_t \) denote the stock price and \( \phi \) the quantity of stock in the hedging portfolio, and suppose that \( S \) satisfies (5) with the Lévy
figure 5: evolution of an option position and the corresponding delta-hedging portfolio in presence of stock jumps.

measure of the jump part denoted by $\nu$. Then the (self-financing) portfolio evolves as

$$dV_t = (V_t - \phi_t S_t) rd_t + \phi_t dS_t.$$ 

The ‘forward’ values of the stock and the portfolio $S_t^* = e^{r(T-t)} S_t$ and $V_t^* = e^{r(T-t)} V_t$
satisfy

$$V_T^* = e^{rT} V_0 + \int_0^T \phi_t dS_t^*.$$ 

We would like to compute the strategy which minimizes the expected squared residual hedging error under the martingale probability:

$$\phi^* = \arg \inf E[(V_T - H_T)^2] = \arg \inf E \left[ \left( V_0 + \int_0^T \phi_t dS_t^* - H_T \right)^2 \right]$$

with $H_T$ the option’s payoff. Using the Itô formula for jump processes and the isometry relation for stochastic integrals (both are out of scope of the present paper but see [11] for details), the residual hedging error can be expressed as

$$E[(V_T - H_T)^2] = \left( e^{rT} V_0 - E[H_T] \right)^2 + E \int_0^T dt (S_t^*)^2 \sigma^2 \left\{ \phi_t - \frac{\partial C}{\partial S} \right\}^2$$

$$+ E \int_0^T \int_{\mathbb{R}} \nu(dz) e^{2r(T-t)} \left\{ C(t, S_t(1 + z)) - C(t, S_t) - S_t \phi_t z \right\}^2.$$ 

From this formula, three immediate conclusions can be made:

• The initial capital minimizing the hedging error is

$$V_0 = e^{-rT} E[H_T].$$

(27)
If the initial capital is given by (27), the residual hedging error is zero (and the market is complete) only in the following two cases:

- No jumps in the stock price ($\nu \equiv 0$). This case correspond to the Black-Scholes model and the optimal hedging strategy is

$$\phi_t = \frac{\partial C}{\partial S}.$$ 

- No diffusion component ($\sigma = 0$) and only one possible jump size ($\nu = \delta z_0(z)$). In this case, the optimal hedging strategy is

$$\phi_t = \frac{C(S_t(1 + z_0)) - C(S_t)}{S_t z_0}.$$ 

In all other cases, the residual hedging error is non-zero (and the market is incomplete) and is minimized by

$$\phi^*(t, S_t) = \frac{\sigma^2 \frac{\partial C}{\partial S} + \frac{1}{S_t} \int \nu(dz) z(C(t, S_t(1 + z)) - C(t, S_t))}{\sigma^2 + \int z^2 \nu(dz)}.$$ 

The optimal quadratic hedging strategy is a weighted sum of two terms: the sensitivity of option price to infinitesimal stock movements, and the average sensitivity to finitely-sized jumps. Figure 6 shows the difference between the optimal strategy and the delta $\frac{\partial C}{\partial S}$. These data were obtained in Merton’s jump diffusion model (2) with parameters $\mu = 0.1$, $r = 0$, $\sigma = 0.2$, $\lambda = 1$, mean jump of $-0.1$, jump standard deviation of 0.05, for a European put option with strike $K = 1.2$ and maturity $T = 1$ month. As we see, the two strategies are not so different after all. The residual hedging errors are also similar: for delta-hedging it has a standard deviation of 1.7% (of the initial stock price) and for the optimal
strategy 1.6%. For comparison, in absence of jump risk, the residual hedging error (due to discrete rebalancing) has a standard deviation of 0.7% and if we do not hedge at all, the error is of order of 16%. In conclusion,

- Hedging with stock only in presence of jumps eliminates a large part risk but still leads to an important residual hedging error.
- Performances of delta hedging and of the optimal quadratic hedging with stock only are very similar.

To eliminate the remaining hedging error, a possible solution is to introduce liquid options into the hedging portfolio. In the above example, if, in addition to the stock, the hedging portfolio contains a European option with strike \( K = 1 \), the standard deviation is 0.76%, that is, the risk due to jumps becomes negligible compared to the one associated to discrete rebalancing. Optimal quadratic hedge ratios in the case when the hedging portfolio may contain options can be found in [11].

6 Model calibration

In the Black-Scholes setting, the only model parameter to choose is the volatility \( \sigma \), originally defined as the annualized standard deviation of logarithmic stock returns. The notion of model calibration does not exist, since after observing a trajectory of the stock price, the pricing model is completely defined. On the other hand, since the pricing model is defined by a single volatility parameter, this parameter can be reconstructed from a single option price (by inverting the Black-Scholes formula). This value is known as the implied volatility of this option.

If the real markets obeyed the Black-Scholes model, the implied volatility of all options written on the same underlying would be the same and equal to the standard deviation of returns of this underlying. However, empirical studies show that this is not the case: implied volatilities of options on the same underlying depend on their strikes and maturities (figure 7, left graph).

Jump-diffusion models provide an explanation of the implied volatility smile phenomenon since in these models the implied volatility is both different from the historical volatility and changes as a function of strike and maturity. Figure 7, right graph shows possible implied volatility patterns (as a function of strike) in the Merton jump-diffusion model.

The results of calibration of the Merton model to S&P index options are presented in figure 8. The calibration was carried out separately for each maturity using the routine [4] from Premia software. In this program, the vector of unknown parameters \( \theta \) is found by minimizing numerically the squared norm of the difference between market and model prices:

\[
\theta^* = \arg \inf \| P_{\text{obs}} - P^\theta \|^2 \equiv \arg \inf \sum_{i=1}^{N} \mathit{w}_i (P_{\text{obs}}^i - P^\theta (T_i, K_i))^2, \quad (28)
\]
Figure 7: Left: implied volatilities of options on S&P 500 index as a function of their strikes and maturities. Right: implied volatilities as a function of strike for different values of the mean jump size in Merton jump diffusion model. Other parameters: volatility $\sigma = 0.2$, jump intensity $\lambda = 1$, jump standard deviation $\delta = 0.05$, option maturity $T = 1$ month.

where $P_{\text{obs}}$ denotes the prices observed in the market and $P^\theta(T_i, K_i)$ is the Merton model price computed for parameter vector $\theta$, maturity $T_i$ and strike $K_i$. Here, the weights $w_i := \frac{1}{(P_{\text{obs}}^i)^2}$ were chosen to ensure that all terms in the minimization functional are of the same order of magnitude. The model prices were computed simultaneously for all strikes present in the data using the FFT-based algorithm described in section 3. The functional in (28) was then minimized using a quasi-newton method (LBFGS-B described in [6]). In the case of Merton model, the calibration functional is sufficiently well behaved, and can be minimized using this convex optimization algorithm. In more complex jump-diffusion models, in particular, when no parametric shape of the Lévy measure is assumed, a penalty term must be added to the distance functional in (28) to ensure convergence and stability. This procedure is described in detail in [9, 10, 22].

The calibration for each individual maturity is quite good, however, although the options of different maturities correspond to the same trading day and the same underlying, the parameter values for each maturity are different, as seen from table 2. In particular, the behavior for short (1 to 5 months) and long (1 to 3 years) maturities is qualitatively different, and for longer maturities the mean jump size tends to increase while the jump intensity decreases with the length of the holding period.

Figure 9 shows the result of simultaneous calibration of Merton model to options of 4 different maturities, ranging from 1 month to 3 years. As we see, the calibration error is much bigger than in figure 8. This happens because, as already observed in section 2, for processes with independent and stationary increments (and the log-price in Merton model is an example of such process), the law of the entire process is completely determined by its law at any given
Figure 8: Calibration of Merton jump-diffusion model to market data separately for each maturity. Top left: maturity 1 month. Bottom left: maturity 5 months. Top right: maturity 1.5 years. Bottom right: maturity 3 years.

Figure 9: Calibration of Merton jump-diffusion model simultaneously to 4 maturities. Calibrated parameter values: $\sigma = 9.0\%$, $\lambda = 0.39$, jump mean $-0.12$ and jump standard deviation $0.15$. Top left: maturity 1 month. Bottom left: maturity 5 months. Top right: maturity 1.5 years. Bottom right: maturity 3 years.
### Table 2: Calibrated Merton model parameters for different times to maturity.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>( \sigma )</th>
<th>( \lambda )</th>
<th>jump mean</th>
<th>jump std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>9.5%</td>
<td>0.097</td>
<td>-1.00</td>
<td>0.71</td>
</tr>
<tr>
<td>2 months</td>
<td>9.3%</td>
<td>0.086</td>
<td>-0.99</td>
<td>0.63</td>
</tr>
<tr>
<td>5 months</td>
<td>10.8%</td>
<td>0.050</td>
<td>-0.59</td>
<td>0.41</td>
</tr>
<tr>
<td>11 months</td>
<td>7.1%</td>
<td>0.70</td>
<td>-0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>17 months</td>
<td>8.2%</td>
<td>0.29</td>
<td>-0.25</td>
<td>0.12</td>
</tr>
<tr>
<td>23 months</td>
<td>8.2%</td>
<td>0.26</td>
<td>-0.27</td>
<td>0.15</td>
</tr>
<tr>
<td>35 months</td>
<td>8.8%</td>
<td>0.16</td>
<td>-0.38</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Figure 10: Calibration of the Bates stochastic volatility jump-diffusion model simultaneously to 4 maturities. Top left: maturity 1 month. Bottom left: maturity 5 months. Top right: maturity 1.5 years. Bottom right: maturity 3 years. Calibrated parameters (see equation (8)): initial volatility \( \sqrt{\nu_0} = 12.4\% \), rate of volatility mean reversion \( \xi = 3.72 \), long-run volatility \( \sqrt{\eta} = 11.8\% \), volatility of volatility \( \theta = 0.501 \), correlation \( \rho = -48.8\% \), jump intensity \( \lambda = 0.038 \), mean jump size \(-1.14\), jump standard deviation 0.73.

time \( t \) (cf. equation 3). If we have calibrated the model parameters for a single maturity \( T \), this fixes completely the risk-neutral stock price distribution for all other maturities. A special kind of maturity dependence is therefore hard-wired into every Lévy jump diffusion model, and table 2 shows that it does not always correspond to the term structures of market option prices.

To calibrate a jump-diffusion model to options of several maturities at the same time, the model must have a sufficient number of degrees of freedom to reproduce different term structures. This is possible for example in the Bates model (8), where the smile for short maturities is explained by the presence of jumps whereas the smile for longer maturities and the term structure of implied volatility is taken into account using the stochastic volatility process. Figure 10 shows the calibration of the Bates model to the same data set as above. As we see, the calibration quality has improved and is now almost as good as when each maturity was calibrated separately. The calibration was once again carried out using the tool [4] from Premia.
References


