Option pricing in exponential Lévy models with transaction costs.

Nicola Cantarutti*, João Guerra, Manuel Guerra,
Maria do Rosário Grossinho

CEMAPRE - Centre for Applied Mathematics and Economics
ISEG - University of Lisbon

November 3, 2016

Abstract

We present an approach for pricing a European call option in presence of proportional transaction costs, when the stock price follows a general exponential Lévy process. The model is a generalization of the celebrated work of Davis, Panas and Zariphopoulou (1993), where the value of the option is found using the concept of utility indifference price. This requires to solve two stochastic singular control problems in finite time, satisfying the same Hamilton-Jacobi-Bellman equation and with different terminal conditions. We solve numerically the continuous time optimization problem using the Markov chain approximation method, and consider the underlying stock following an exponential Merton jump-diffusion process. This model takes into account the possibility of portfolio bankruptcy. We show numerical results for the simpler case of an infinitely rich investor, whose probability of default can be ignored. Option prices are obtained for both the writer and the buyer.

Key words: option pricing, transaction costs, Lévy processes, indifference price, singular stochastic control, variational inequality, Markov chain approximation.

1 Introduction

The problem of pricing a European call option was first solved mathematically in the paper of Black and Scholes (1973). Even if it is quite evident that this model is too simplistic to represent the real features of the market, it is still nowadays one of the most used model to price and hedge options. The reason for its success is that it gives a closed form solution for the option price, and that the hedging strategy is easly implementable. The Black-Scholes model considers a *complete* market, i.e. a market where it is possible to create a portfolio containing cash and shares of the underlying stocks, such that following a particular trading strategy it is always possible to replicate the payoff of the

*Corresponding author. Email: nicolacantarutti@gmail.com

option. In this framework, this particular portfolio is called *replicating portfolio* and the trading strategy to hedge the option is called *delta hedge*. However, this model does not consider many features that characterize the real market.

In the Black-Scholes model, the stock price follows a geometric Brownian motion. This is equivalent to assume that the log-returns are normally distributed. However, a deep statistical analysis of financial data, reveals that the normality assumption is not a very good approximation of reality (see Cont (2001)). Indeed, it is easy to see that empirical log-return distributions have substantially more mass around the origin and along the tails (*heavy tails*). This means that normal distribution underestimates the probability of large log-returns, and considers them just as rare events. In the real market instead, log-returns manifest frequently high peaks, that come more and more evident when looking at short time scales. The log-returns peaks correspond to sudden large changes in the price, which are called *jumps*. There is a huge literature of option pricing models that consider an underlying process with a discontinuous path. Most of these models consider the log-prices dynamics following a *Lévy process*. These are stochastic processes with independent and stationary increments, that satisfy the additional property of stochastic continuity. Good references on the theory of Lévy processes are the books of Sato (1999) and Applebaum (2009). Financial applications are discussed in the book of Cont and Tankov (2003).

A second issue of the Black-Scholes model is that it does not consider the presence of market frictions such as bid/ask spread, transaction fees or budget constraints. The securities in the market are traded with a bid-ask spread, and this means that there are two prices for the same security. But the Black-Scholes formula just gives one price. Moreover, the replicating portfolio cannot be perfectly implemented, since the delta-hedging strategy involves continuous time trading. This is impractical because the presence of transaction costs makes it infinitely costly. Another kind of market friction that needs to be considered are the budget constraints. A bound in the budget or a restriction in the possibility of selling short, clearly restricts the set of possible trading strategies.

Many authors attempted to include the presence of proportional transaction costs in option pricing models. In Leland (1985), in order to avoid continuous trading, the author specifies a finite number of trading dates. He obtains a Black-Scholes type partial differential equation (PDE) with an adjusted volatility term, that takes into account the transaction costs. However, trading at fixed dates is not optimal, and the option price goes to infinity as the number of dates grows. Further work in this direction has been done by Boyle and Vorst (1992), which considers a multiperiod binomial model as in Cox et al. (1979), with transaction costs. Here again, the cost of the replicating portfolio depends on the number of time periods. A different approach has been introduced by Hodges and Neuberger (1989). In their work, they used an alternative definition of the option price called *indifference price*, based on the concepts of *expected utility* and *certainty equivalent*. An overview of these concepts applied to several incomplete market model can be found in Carmona (2009).

As long as the perfect replicating portfolio is no longer implementable in presence of transaction costs, the hedging strategy cannot be anymore riskless. The model has to take into account the risk profile of the writer/buyer to describe his trading preferences. Hodges and Neuberger (1989) define the option price as the value that makes an investor indifferent between holding a porfolio with

an option and without, in terms of expected utility of the final wealth. They showed that it is impossible to hedge perfectly the option. The optimal strategy is to keep the portfolio's value within a band called *no transaction region*. Using numerical experiments, they showed that this strategy outperform the one proposed in Leland (1985). This approach has been further developed in Davis et al. (1993), where the problem is formulated rigorously as a singular stochastic optimal control problem. The authors proved that the value functions of the two optimization problems can be interpreted as the solutions of the associated Hamilton-Jacobi-Bellman (HJB) equation in the viscosity sense. They prove also that the numerical scheme, based on the *Markov chain approximation*, converges to the viscosity solution. Numerical methods for this model are presented in Davis and Panas (1994), Clewlow and Hodges (1997) and Monoyios (2003), Monoyios (2004).
In Whalley and Wilmott (1997) and Barles and Soner (1998) are discussed two different asymptotic analysis of the problem for small levels of transaction costs. Essentially the authors reduce the HJB equation to a simpler non-linear PDE. Further studies are presented in the thesis work of Damgaard (1998), where the author studied the robustness of the model from a theoretical and numerical point of view. He found that under particular conditions the model is quite robust with respect to the choice of the utility function.

In this work we want to develop a model for pricing options using the concept of indifference price, as done in Davis et al. (1993). In our model we consider proportional transaction costs, and a stock dynamics following an exponential Lévy process. In contrast with the Black-Scholes model, we obtain two prices: the price for the buyer and the price for the writer of the option, which is a more realistic property of the model. The presence of jumps means we have to consider the possibility of insolvency. A sudden jump in the price can have dangerous consequences and cause the bankruptcy of the investor. It turns out that it is very difficult to solve the general maximization problem numerically. In order to simplify the problem, we consider the special case of an investor with infinite wealth (always solvent) and an exponential utility function to describe his risk profile. Under these assumptions it is possible to reduce by one the number of variables of the HJB equation.

In Section 2, after a short review of Lévy processes theory, we introduce the model's equations and definitions. We derive the HJB equation associated to the maximization problem. In Section 3 we obtain a general Markov chain approximation of the continuous time problem. We introduce the Merton jump-diffusion process and refer to Appendix B for its specific Markov chain approximation. The numerical results are presented in Section 4 and a complete summary of all the outcomes is presented in the conclusive Section 5.

2 The model

2.1 Exponential Lévy models

Let X_t be a Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where $\mathcal{F}_t$ is the natural filtration and $t \in [t_0, T]$. We assume that X_t has the characteristic Lévy triplet (b, σ, ν), where $b \in \mathbb{R}$, $\sigma \geq 0$ and ν is a positive measure on $\mathbb{R}$, called *Lévy measure* which satisfies:

$$\nu(\{0\}) = 0, \qquad \int_{\mathbb{R}} (1 \wedge z^2)\nu(dz) < \infty. \tag{1}$$

We model the log-prices dynamics with the Lévy process X_t. Thus the price of a stock S_t follows an *exponential Lévy process*:

$$S_t = S_0 e^{X_t}. \tag{2}$$

Motivated by practical reasons, we only consider processes with finite mean and variance. The conditions for finite second moments: $\mathbb{E}[|X_t|^2] < \infty$ and $\mathbb{E}[\exp(2X_t)] < \infty$, are directly related to the integrability conditions of the Lévy measure:

$$\int_{|z|\geq 1} |z|^2 \nu(dz) < \infty \qquad \int_{|z|\geq 1} e^{2z} \nu(dz) < \infty. \tag{3}$$

Considering the above conditions, the processes follow the SDEs:

$$dX_t = \left(b + \int_{|z|\geq 1} z\nu(dz) \right) dt + \sigma dW_t + \int_{\mathbb{R}} z\tilde{N}(dt, dz), \tag{4}$$

and

$$\begin{aligned} \frac{dS_t}{S_t} = & \left(b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} \left(e^z - 1 - z\mathbb{1}_{\{|z|<1\}}\right)\nu(dz) \right) dt \\ & + \sigma dW_t + \int_{\mathbb{R}} (e^z - 1)\tilde{N}(dt, dz), \end{aligned} \tag{5}$$

where the term $\tilde{N}(dt, dz)$ is the compensated Poisson martingale measure. It is defined as

$$\tilde{N}(dt, dz) = N(dt, dz) - dt\, \nu(dz), \tag{6}$$

where $N(dt, dz)$ is the Poisson random measure with intensity $dt\, \nu(dz)$.
The Eq. (4) is also called the *Lévy-Itô decomposition*, and Eq. (5) can be derived applying the Itô lemma to (2).
In the following, we indicate the drift term simply as

$$\mu = b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} \left(e^z - 1 - z\mathbb{1}_{\{|z|<1\}}\right)\nu(dz). \tag{7}$$

2.1.1 Lévy processes as Markov processes

Lévy processes are Markov processes. The infinitesimal generator associated to the price process (5) is given by:

$$\begin{aligned} \mathcal{L}^S f(s) &= \lim_{t\to 0} \frac{\mathbb{E}[f(S_t)|S_0 = s] - f(s)}{t} \\ &= \mu s \frac{\partial f(s)}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 f(s)}{\partial s^2} \\ &+ \int_{\mathbb{R}} \left[f(se^z) - f(s) - s(e^z - 1)\frac{\partial f(s)}{\partial s} \right] \nu(dz). \end{aligned} \tag{8}$$

with $f \in C^2(\mathbb{R}) \bigcap C_2(\mathbb{R})$, where $C^2(\mathbb{R})$ is the space of twice differenciable functions and $C_2(\mathbb{R})$ is the space of continuous functions with polinomial growth of second degree at infinity.
We can define the *transition probabilities* associated to the Lévy process. For any Borel set $B \in \mathcal{B}(\mathbb{R})$:

$$p_{s,t}(x, B) = P(X(t) \in B | X(s) = x). \tag{9}$$

Transition probabilities are connected with conditional expectation by the simple relation:

$$\mathbb{E}[f(X(t))|X(s) = x] = \int_{\mathbb{R}} f(y) p_{s,t}(x, dy) \tag{10}$$

2.2 Portfolio dynamics with transaction costs

In this section we introduce the market model with proportional transaction costs that generalizes the model of Davis et al. (1993). Let us consider a portfolio composed by one risk-free asset B (bank account) paying a fixed interest rate $r > 0$ and a stock S. We denote with Y the number of shares of the stock S that the investor holds. The state of the investor at time $t \in [t_0, T]$ is (B_t^π, Y_t^π, S_t). The portfolio evolves following the SDE:

$$\begin{cases} dB_t^\pi & = rB_t dt - (1+\theta_b) S_t dL_t + (1-\theta_s) S_t dM_t \\ dY_t^\pi & = dL_t - dM_t \\ dS_t & = S_t \left(\mu dt + \sigma dW_t + \int_{\mathbb{R}} (e^z - 1) \tilde{N}(dt, dz) \right). \end{cases} \tag{11}$$

The parameters θ_b, $\theta_s \geq 0$ are the proportional transaction costs when buying and selling, respectively. The process $\pi(t) = (L(t), M(t))$ is the trading strategy and represents the cumulative number of shares bought and sold respectively in $[t_0, T]$.
These processes are right-continuous, $\mathcal{F}_t$-adapted and nondecreasing. By convention $L(t_0^-) = M(t_0^-) = 0$ and we allow a possible initial transaction at t_0. Furthermore, $\pi(t)$ is progressively measurable:

$$\mathbb{E}[L(t)^n] < \infty \text{ , } \mathbb{E}[M(t)^n] < \infty \qquad \text{for} \quad n = 1, 2 \quad \text{and} \quad \forall t \in [t_0, T].$$

Define the *cash value* as the value in cash when the shares in the portfolio are liquidated: long positions are sold and short positions are covered.

$$c(y, s) = \begin{cases} (1+\theta_b) ys, & \text{if } y \leq 0 \\ (1-\theta_s) ys, & \text{if } y > 0. \end{cases} \tag{12}$$

For $t \in [t_0, T]$, we define the *total wealth* process:

$$W_t^\pi = B_t^\pi + c(Y_t^\pi, S_t). \tag{13}$$

In this model we require that the portfolio's wealth W_t is greater than a fixed constant $-C$, with $C \geq 0$ for all $t \in [t_0, T]$, as a condition for the investor to be solvent. We define the *solvency region*:

$$\mathcal{S} = \Big\{ (B_t, Y_t, S_t) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ : B_t + c(Y_t, S_t) \geq -C \Big\}. \tag{14}$$

As long as we describe the underlying stock as a process with jumps, we cannot guarantee that the portfolio stays solvent for all $t \in [t_0, T]$. Holding short positions, it is possible that a sudden increase in the value of the stock can cause the total wealth to jump instantaneously out of the solvency region. The same happens with a downward jump when the investor is long in stocks and negative in cash. The sudden decrease of the stock's price makes him unable to pay his debts. If the investor goes bankrupt, there are no trading strategies to save him.
Define the first exit time from the solvency region as

$$\tau_{\mathcal{S}} = \inf\big\{ t \in [t_0, T] : W_t \notin \mathcal{S} \big\}. \tag{15}$$

Define the set of *admissible trading strategies* $\Pi(B_t, Y_t, S_t)$, as the set of all right-continuous, nondecreasing, $\mathcal{F}_t$-measurable processes $L(t)$ and $M(t)$, such that $(B^\pi(t), Y^\pi(t), S(t); \pi(t))$ is a solution of (11) for $t \in [t_0, \tau_{\mathcal{S}} \wedge T]$, and with initial values (B_0, Y_0, S_0) .

Remark: The original model, as formulated by Hodges and Neuberger (1989) and Davis et al. (1993), considers a portfolio starting with zero total wealth. But it does not consider the possibility of insolvency.
The writer (or buyer) of the option creates a portfolio at time t_0 in order to hedge the option. Therefore it is reasonable to assume that it does not own any shares in the underlying stock before t_0. We consider, following the previous works, an initial portfolio with zero shares and the initial value B_0 in the cash account. This assumption can be easly relaxed to include an initial amount Y_0 of shares if needed.

2.3 Utility maximization

2.3.1 No option

The objective of the investor is to maximize the expected utility of the wealth at $\tau_{\mathcal{S}} \wedge T$ over all the admissible strategies. This expectation is conditioned on the initial value of cash B_0, number of shares Y_0 and value of the stock S_0. We can write the value function of the maximization problem:

$$V^0(t_0, B_0, Y_0, S_0) = \sup_{\pi \in \Pi(B_t, Y_t, S_t)} \mathbb{E}_{B_0, Y_0, S_0} \Big[\mathcal{U}(W_T) \, \mathbb{1}_{\{\tau_{\mathcal{S}} > T\}} + \mathcal{U}(-C) \, \mathbb{1}_{\{\tau_{\mathcal{S}} \leq T\}} \Big] \tag{16}$$

where $\mathcal{U} : \mathcal{S} \to \mathbb{R}$ is a concave and increasing utility function, such that $\mathcal{U}(0) = 0$. It is important to note that the utility function has to be defined also for negative numbers (this condition excludes the use of a logarithmic utility).
We use indicator functions to consider separately the cases of solvency and insolvency. The second term inside the expectation considers the case of bankruptcy, where the value function assumes the least possible value attainable by the utility function.

2.3.2 Writer/Buyer of the option

Assume that the investor builds a portfolio with a cash account, shares of a stock and in addition sells or purchases a European call option written on the

same stock, with strike price K and expiry date T. This means that at time t_0 the initial amount in the cash account increases by the option's value p^w (in the writer case), or decrease by p^b (in the buyer case).
Define the wealth process for the writer and buyer respectively by:

- $W_t^w = B_t + c(Y_t, S_t)\mathbb{1}_{\{t<T, c(1,S_T)\leq K\}} + \Big(c\big(Y_t - 1, S_t\big) + K\Big)\mathbb{1}_{\{t=T, c(1,S_T)>K\}}$
- $W_t^b = B_t + c(Y_t, S_t)\mathbb{1}_{\{t<T, c(1,S_T)\leq K\}} + \Big(c\big(Y_t + 1, S_t\big) - K\Big)\mathbb{1}_{\{t=T, c(1,S_T)>K\}}$.

In the case the option is exercised, $c(1, S_T) > K$, the buyer pays the writer the strike K in cash, and the writer delivers one share to the buyer.
Remember that in a market with transaction costs the real value of a share is given by the bilinear cash value function (12). So, the buyer of the option does not exercise when $S_T > K$, but when $c(1, S_T) = S_T(1 - \theta_s) > K$.
The *solvency regions* are:

- $\mathcal{S}^w = \Big\{(B_t, Y_t, S_t) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ : W_t^w > -C\Big\}$
- $\mathcal{S}^b = \Big\{(B_t, Y_t, S_t) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ : W_t^b > -C\Big\}$.

The investor wishes to maximize the expected utility of the wealth of his portfolio.

$$V^j(t_0, B_0, Y_0, S_0) = \sup_{\pi \in \Pi(B_t, Y_t, S_t)} \mathbb{E}_{B_0, Y_0, S_0}\Big[\mathcal{U}(W_T^j)\ \mathbb{1}_{\{\tau_S > T\}} + \mathcal{U}(-C)\ \mathbb{1}_{\{\tau_S \leq T\}}\Big], \tag{17}$$

for $j = w, b$.

2.3.3 Indifference pricing

With this model we can compute two option prices: the price for the writer and the price for the buyer. These prices are defined, respectively, as the amount required to get the same maximal expected utility of the wealth of the portfolio without the option. To compute the option price, it is necessary to solve two portfolio optimization problems: the problem without the option and the problem with the option. We define the

- Writer price:

$$V^0(t_0, B_0, Y_0, S_0) = V^w(t_0, B_0 + p^w, Y_0, S_0) \tag{18}$$

- Buyer price:

$$V^0(t_0, B_0, Y_0, S_0) = V^b(t_0, B_0 - p^b, Y_0, S_0) \tag{19}$$

The prices p^w and p^b can be obtained implicitly by these conditions.

2.3.4 Hamilton-Jacobi-Bellman Equation

We present the HJB equation associated to the singular stochastic optimal control problems described before. These problems are called singular because the controls $(dL(t), dM(t))$ are allowed to be singular with respect to the Lebesgue measure dt. A rigorous derivation of the following equation can be found in Fleming and Soner (2005). It turns out that the HJB equation of the singular control problem is a variational inequality:

$$\max \Bigg\{ \frac{\partial V^j}{\partial t} + rb\frac{\partial V^j}{\partial b} + \mu s \frac{\partial V^j}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V^j}{\partial s^2} \tag{20}$$

$$+ \int_{\mathbb{R}} \left[V^j(t,b,y,se^z) - V^j(t,b,y,s) - s(e^z - 1)\frac{\partial V^j}{\partial s} \right] \nu(dz) \; ,$$

$$\frac{\partial V^j}{\partial y} - (1+\theta_b)s\frac{\partial V^j}{\partial b} \; , \; -\left(\frac{\partial V^j}{\partial y} - (1-\theta_s)s\frac{\partial V^j}{\partial b} \right) \Bigg\} = 0,$$

for $(t,b,y,s) \in [t_0,T] \times \mathcal{S}^j$ and $j = 0, w, b$. The boundary conditions are given by Eqs. (16) and (17). The main difference between our model and the previous models in the literature, is that this HJB equation is a partial integro-differential equation (PIDE), which involves an additional integral operator. The presence of this non-local operator implies that we need to define the lateral conditions not only on the boundary of the solvency region, but also beyond.
This is given by the condition

$$V^j(t, W_t) = \mathcal{U}(-C) \qquad \text{for} \qquad W_t \notin \mathcal{S}^j \quad j = 0, w, b. \tag{21}$$

The variational inequality (20) says that the maximum of three operators is equal to zero. This feature can be interpreted better if we consider the state space divided into three different regions: the Buy, the Sell and the No Transaction (NT) regions.
The optimization problem is a free boundary problem, and its solution consist of finding the value function V and the optimal boundaries of these three regions. These boundaries completely characterize the investor's trading strategy. The optimal strategy consists in keeping the portfolio process inside the NT region. If the portfolio jumps outside the NT region, the optimal strategy is to trade in order to bring back the portfolio on the boundary of the NT region. We can argue that Buy and Sell regions are separated by the NT region, since it is clearly not optimal to buy and sell a stock at the same time.
In the Buy and Sell regions the value functions remain constant along the directions of the trades. Respectively, we have:

- Buy: $V(t,b,y,s) = V(t, b - s(1+\sigma_b)\Delta L_t, y + \Delta L_t, s).$
- Sell: $V(t,b,y,s) = V(t, b + s(1-\sigma_s)\Delta M_t, y - \Delta M_t, s).$

where $\Delta L_t = L(t) - L(t^-)$ and $\Delta M_t = M(t) - M(t^-)$ are the number of shares respectively bought or sold in the trade.
The second and third terms in the HJB equation (20) are the gradient of the value function along the optimal trading direction from the Buy and Sell regions to the NT boundaries. In the NT region the portfolio evolves according to the portfolio equation (11), with $dL = dM = 0$. The number of shares remains

constant as long as the portfolio stays in the NT region. By the dynamic programming principle, we can write the value function as

$$V(t,b,y,s) = \mathbb{E}_{b,y,s}\left[V(t+\Delta t, b+\Delta B, y, s+\Delta S) \right]$$

where Δt is a (small) finite interval of time. We indicate with ΔB and ΔS the change in the cash account and in the stock price during the time interval Δt. For $\Delta t \to 0$ we can use Itô's lemma to obtain the infinitesimal generator of the process, that corresponds to the first term in Eq. (20).

2.4 Variable reduction

In this model we introduce the possibility of insolvency, which is directly reflected in the definition of the set of admissible trading strategies. In the literature, all the models that consider diffusion processes use to define the set of admissible trading strategies $\Pi(B_0, Y_0, S_0)$ as the set of all right-continuous, measurable processes $(B^\pi(t), Y^\pi(t), S(t); \pi(t))$ solution of (11) with initial values (B_0, Y_0, S_0), such that:

$$W^\pi_t \in \mathcal{S} \qquad \forall t \in [t_0, T]. \tag{22}$$

This set is completely determined by the initial values of the portfolio.
In this model, the stock process can jump and the portfolio can go bankruptcy. This means that the set of admissible strategies $\Pi(B_t, Y_t, S_t)$ is a dynamic set, and at every time $t \in [t_0, \tau_{\mathcal{S}} \wedge T]$ it depends on the current state.

Assumption: We assume that the portfolio is always solvent. Therefore the set of trading strategies only depends on the initial wealth of the investor:

$$\Pi(B_t, Y_t, S_t) \approx \Pi(B_0, Y_0, S_0). \tag{23}$$

This can be the case of a large investor, with a very small probability of default that can be ignored. To express this concept we can say that the investor has infinite initial value in the cash account $B_0 = \infty$. Consequently, the solvency constraint (14) lose meaning, and we can set for convenience $C = 0$. The lateral boundary conditions (21) lose importance as well, and are ignored.

The reason for this strong assumption is that the HJB Eq. (20) is difficult to solve numerically. With a solvent portfolio we can use the properties of the exponential utility function to reduce by one the number of variables of the problem. In the conclusive Section 5 we give a further comment on this assumption.
The exponential utility is defined as:

$$\mathcal{U}(w) = 1 - e^{-\gamma w}. \tag{24}$$

It has already been used to reduce the number of variables in the work of Hodges and Neuberger (1989), Davis et al. (1993) and related works. The exponential utility has the property that the coefficient of risk aversion $-\mathcal{U}''(x)/\mathcal{U}'(x) = \gamma$ is constant, and does not depend on the wealth w. This means that the amount invested in the risky asset, at time $t \in [t_0, T]$, is independent of the total wealth at time t. This choiche of utility function let us simplify the problem by reducing

the number of variables from four to three. As long as the amount in the risky asset is independent of the total wealth, the amount in the cash account is irrelevant to the trading strategy. We can thus remove B_t from the state dynamics.
The integral representation of the evolution of the cash account B_t in (11) is:

$$B^\pi(T) = \frac{B_0}{\delta(t_0,T)} - \int_{t_0}^{T} (1+\theta_b)\frac{S(t)}{\delta(t,T)} dL(t) + \int_{t_0}^{T} (1-\theta_s)\frac{S(t)}{\delta(t,T)} dM(t) \quad (25)$$

where $\delta(t,T) = e^{-r(T-t)}$. The explicit expression of the value function (16) for the maximization problem without option, using the exponential utility (24) is:

$$\begin{aligned} V^0(t_0,B_0,Y_0,S_0) = &\sup_{\pi\in\Pi(B_0,Y_0,S_0)} \mathbb{E}_{B_0,Y_0,S_0}\Bigg[\left(1 - e^{-\gamma\left(B^\pi(T)+c(Y^\pi(T),S(T))\right)}\right) \\ &\mathbb{1}_{\{\tau_S>T\}}\Bigg] + \mathbb{E}_{b,y,s}\left[\left(1-e^{-\gamma(-C)}\right)\mathbb{1}_{\{\tau_S\le T\}}\right] \\ = 1 - &\inf_{\pi\in\Pi(B_0,Y_0,S_0)} \mathbb{E}_{B_0,Y_0,S_0}\Bigg[e^{-\gamma\left(B^\pi(T)+c(Y^\pi(T),S(T))\right)}\mathbb{1}_{\{\tau_S>T\}} \\ &+ e^{\gamma C}\,\mathbb{1}_{\{\tau_S\le T\}}\Bigg] \\ = 1 - &e^{-\gamma\frac{B_0}{\delta(t_0,T)}} Q^0(t_0,Y_0,S_0) \end{aligned}$$

where we have used the expression (25). The new minimization problem is:

$$\begin{aligned} Q^0(t_0,Y_0,S_0) = &\inf_{\pi\in\Pi(B_0,Y_0,S_0)} \mathbb{E}_{Y_0,S_0}\Bigg[e^{-\gamma\left[-\int_{t_0}^T(1+\theta_b)\frac{S_t}{\delta(t,T)}dL_t + \int_{t_0}^T(1-\theta_s)\frac{S_t}{\delta(t,T)}dM_t\right]} \\ &e^{-\gamma c(Y^\pi(T),S(T))}\mathbb{1}_{\{\tau_S>T\}} + e^{\gamma\left(C+\frac{B_0}{\delta(t_0,\tau_S)}\right)}\,\mathbb{1}_{\{\tau_S\le T\}}\Bigg]. \end{aligned} \quad (26)$$

In order to simplify further the equation, it is convenient to pass to log-variables: $x = \log(s)$. Note that the derivative operators change as:

$$s\frac{\partial}{\partial s} = \frac{\partial}{\partial x}, \qquad s^2\frac{\partial^2}{\partial s^2} = \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}.$$

The HJB Eq. (20) now becomes:

$$\begin{aligned} &\min\Bigg\{ \frac{\partial Q^j}{\partial t} + (\mu - \frac{1}{2}\sigma^2)\frac{\partial Q^j}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 Q^j}{\partial x^2} \\ &+ \int_{\mathbb{R}} \left[Q^j(t,y,x+z) - Q^j(t,y,x) - (e^z-1)\frac{\partial Q^j}{\partial x}\right]\nu(dz)\;, \\ &\frac{\partial Q^j}{\partial y} + (1+\theta_b)e^x\frac{\gamma}{\delta(t,T)}Q^j\;,\; -\left(\frac{\partial Q^j}{\partial y} + (1-\theta_s)e^x\frac{\gamma}{\delta(t,T)}Q^j\right)\Bigg\} = 0 \end{aligned} \quad (27)$$

with $j = 0, w, b$.
For the case of no-option, $j = 0$, we have:

- Terminal conditions:

$$Q^0(T,y,s) = e^{-\gamma c(y,s)}, \quad (28)$$

- Lateral conditions:

$$Q^0(t, y, s) = e^{\gamma\left(C+\frac{B_0}{\delta(t_0,t)}\right)} \tag{29}$$

for all $y, s \in \mathbb{R} \times \mathbb{R}^+$ such that $c(y, s) \leq -\left(C + \frac{B_0}{\delta(t_0,t)}\right)$. Under the assumption of infinite initial capital, $B_0 = \infty$, the lateral conditions are ignored.

Analogous computations can be done for the portfolio of the writer and the buyer of the option. What changes are the terminal conditions. We obtain the terminal conditions for the writer and buyer respectively:

$$Q^w(T, y, s) = e^{-\gamma\left[c(y,s)\mathbb{1}_{\{c(1,s)\leq K\}}+\left(c(y-1,s)+K\right)\mathbb{1}_{\{c(1,s)>K\}}\right]} \tag{30}$$

$$Q^b(T, y, s) = e^{-\gamma\left[c(y,s)\mathbb{1}_{\{c(1,s)\leq K\}}+\left(c(y+1,s)-K\right)\mathbb{1}_{\{c(1,s)>K\}}\right]} \tag{31}$$

Using conditions (18), (19), we obtain the price of the option as:

$$p^w(t_0, y, s) = \frac{\delta(t_0, T)}{\gamma} \log\left(\frac{Q^w(t_0, y, s)}{Q^0(t_0, y, s)}\right) \tag{32}$$

$$p^b(t_0, y, s) = \frac{\delta(t_0, T)}{\gamma} \log\left(\frac{Q^0(t_0, y, s)}{Q^b(t_0, y, s)}\right) \tag{33}$$

2.5 The dynamic programming equation

The HJB Eq. (27) has the following integral representation:

$$\begin{aligned} Q^j(t, y, x) = \min \Bigg\{ & \mathbb{E}_{y,x}\left[Q^j\big(t + \Delta t, y, x + \Delta X\big)\right], \\ & \exp\left(\frac{\gamma}{\delta(t,T)}(1 + \sigma_b)e^x \Delta L_t^*\right) Q^j\big(t, y + \Delta L_t^*, x\big), \\ & \exp\left(-\frac{\gamma}{\delta(t,T)}(1 - \sigma_s)e^x \Delta M_t^*\right) Q^j\big(t, y - \Delta M_t^*, x\big) \Bigg\}. \end{aligned} \tag{34}$$

Each term inside the "min" is the integral form of the corresponding term in the differential equation (27). The values ΔL_t^* and ΔM_t^* are the optimal number of shares needed to trade the portfolio to the boundary of the NT region.

3 Markov chain approximation

To solve the variational inequality (27) we use the Markov chain approximation method developed by Kushner and Dupuis (2001). The numerical tecnique for singular controls has been developed in the work of Kushner and Martins (1991). The portfolio dynamics (11) is approximated by a discrete state controlled Markov chain in discrete time. The method consist in creating a backward recursive dynamic programming algorithm, in order to compute the value function at time t, given its value at time $t + \Delta t$.
Kushner and Dupuis (2001) proved that the value function obtained from the discrete dynamic programming algorithm, converges to the value function of the original continuous time problem as $\Delta t \to 0$. Their proof uses a weak

convergence in probability argument. Another proof of convergence has been introduced by Barles and Souganidis (1991), where they prove convergence to the viscosity solution of the original HJB equation. In the work of Davis et al. (1993), they prove the existence and uniqueness of the viscosity solution of the HJB Eq. (20) for the diffusion case, and using the method developed by Barles and Souganidis (1991) proved that the value function obtained with the Markov chain approximation converges to it.

In this work we model the stock as a general exponential Lévy process. For practical computations we need to specify which Lévy process we intend to use, and this is equivalent to specify its Lévy triplet. In this work we consider a Merton jump diffusion process for the log-prices. The Markov chain approximation can be applied to all the Lévy processes with finite activity of jumps, following a similar discretization [1].

3.1 The Merton jump-diffusion model

One of the firsts jump-diffusion models applied to finance is the *Merton model*, presented in Merton (1976). In the paper, the author obtained a closed form solution (as a series expansion) for the price of a European vanilla call option. The Merton model describes the log-prices evolution as a Lévy process X_t with a characteristic Lévy triplet (b, σ, ν) with $b \in \mathbb{R}$, $\sigma > 0$ and *Lévy measure* (1):

$$\nu(dz) = \frac{\lambda}{\xi\sqrt{2\pi}} e^{-\frac{(z-\alpha)^2}{2\xi^2}} dx. \tag{35}$$

The process can be represented as the superposition of a drift component, a diffusion component and a finite number of jumps. The number of jumps is a Poisson process N_t with intensity $\lambda > 0$, and the size of the jumps are normal distributed $\sim \mathcal{N}(\alpha, \xi^2)$. Recall that for processes with finite number of jumps in a time interval, the Lévy measure is the product of the jump intensity λ with the density of the jump sizes. Using the Eqs. (4) and (7), we obtain the SDE for X_t:

$$\begin{aligned} dX_t &= \left(\mu - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}} (e^z - 1 - z)\nu(dz) \right) dt + \sigma dW_t + \int_{\mathbb{R}} z\tilde{N}(dt, dz) \qquad (36) \\ &= \left(\mu - \frac{1}{2}\sigma^2 - m \right) dt + \sigma dW_t + \int_{\mathbb{R}} zN(dt, dz), \end{aligned}$$

where we used (6), and defined:

$$m = \int_{\mathbb{R}} (e^z - 1)\nu(dz) = \lambda\left(e^{\alpha + \frac{\xi^2}{2}} - 1 \right).$$

For any $f \in C^2(\mathbb{R}) \bigcap C_2(\mathbb{R})$, the associated infinitesimal generator is:

$$\begin{aligned} \mathcal{L}^X f(x) = & \left(\mu - \frac{1}{2}\sigma^2 - m \right) \frac{\partial f(x)}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f(x)}{\partial x^2} \qquad (37) \\ & - \lambda f(x) + \int_{\mathbb{R}} f(x+z)\, \nu(dz). \end{aligned}$$

[1] For Lévy processes with infinite activity, a common procedure is to approximate the small jumps with a Brownian motion, as explained in Cont and Voltchkova (2005). This helps to remove the singularity of the Lévy measure near zero.

3.2 The discrete model

We now discretize the time and space in order to construct a Markov chain approximation of the log-price dynamics. The desired Markov chain approximation for a jump-diffusion model has to satisfy two conditions to be admissible:

1. the transition probabilities are represented in a specific form (Eq. (48) in Appendix B)).
2. the transition probabilities have to be *locally consistent* with the SDE.

The concept of locally consistence can be thought intuitively as a moment matching condition between the Markov chain and the continuous process at each time step. In the Appendix B we breafly describe the two properties and check that are satisfied by our Markov chain. For the general theory of Markov chain approximations we refer to Kushner and Dupuis (2001). The authors prove that under the two conditions, the discrete Markov chain converges in probability to the corresponding continuous SDE as the time step goes to zero.
For $n = 0, 1, ...N \in \mathbb{N}$, define the discrete time step $\Delta t = \frac{T-t_0}{N}$ such that $t_n = t_0 + n\Delta t$.
Define the space $\Sigma_x = \{-K_1 h_x, ... - h_x, 0, h_x, ... + K_2 h_x\}$ where we consider the discrete space step h_x in the x direction, and $K_1, K_2 \in \mathbb{N}$. The values of K_1 and K_2 can be different to capture the possible asymmetry of the jump sizes.
The discretized version of the SDE (36) is:

$$\Delta X_n = \left(\mu - \frac{1}{2}\sigma^2 - m\right)\Delta t + \sigma \Delta W_n + \Delta J_n, \tag{38}$$

where $\Delta X_n = X(t_n + \Delta t) - X(t_n)$ and the term $\Delta W_n = W(t_n + \Delta t) - W(t_n) \in \Sigma_x$ assumes only three possible values:

$$\Delta W_n = \begin{cases} h_x \quad \text{with} \quad \mathbb{P}(W_{n+1} = W_n + h_x | W_n) \\ 0 \quad \text{with} \quad \mathbb{P}(W_{n+1} = W_n | W_n) \\ -h_x \quad \text{with} \quad \mathbb{P}(W_{n+1} = W_n - h_x | W_n) \end{cases}$$

and ΔJ_n is the compound Poisson jump term, that assumes all the values in Σ_x with specific transition probabilities.
In the construction of ΔX_n the transition probabilities and the parameters h_x and Δt have to be chosen such that the Markov chain is locally consistent with the first two moments of the SDE (36) :

$$\mathbb{E}\big[\Delta X_n\big] = \big(\mu - \frac{1}{2}\sigma^2 - m + \lambda\alpha\big)\Delta t, \tag{39}$$

$$\mathbb{E}\Big[\big[\Delta X_n - \mathbb{E}[\Delta X_n]\big]^2\Big] = \big(\sigma^2 + \lambda\xi^2 + \lambda\alpha^2\big)\Delta t. \tag{40}$$

(see Appendix B.2).

Define the space $\Sigma_y = \{-K_3 h_y, ...-h_y, 0, h_y, ...+K_4 h\}$, where h_y is a discrete step in y direction and $K_3, K_4 \in \mathbb{N}$. The number of shares y takes values in Σ_y. The two increments ΔL_n, ΔM_n which describe the change in the number of shares bought or sold are positive multiples of h_y. The discretized SDE for the number of shares is:

$$\Delta Y_n = \Delta L_n - \Delta M_n \tag{41}$$

The action of the control is supposed to happen instantaneously: $\Delta L_n = L(t_n) - L(t_n^-)$ and $\Delta M_n = M(t_n) - M(t_n^-)$ happen at the same time t_n. We indicate with $L(t_n^-)$ and $M(t_n^-)$ the number of shares just before the transaction.
The backward algorithm for computing the value function using the dynamic programming equation, considers thus two different steps: a jump-diffusion step and a control step. However, for the numerical implementation, we cannot use the equation written as in (34) because the value function at time t on the right hand side is still unknown. We can represent the value function at time t, as an expectation of its values at time $t + \Delta t$. The final equation is therefore:

$$\begin{aligned} Q(t_n, Y_n, X_n) = \min \bigg\{ & \mathbb{E}\Big[Q\big(t_{n+1}, Y_n, X_n + \Delta X_n\big)\Big], \qquad (42) \\ & \min_{\Delta L} \exp\bigg(\frac{\gamma}{\delta(t_n, T)} (1 + \sigma_b) e^{X_n} \Delta L_n \bigg) \mathbb{E}\Big[Q\big(t_{n+1}, Y_n + \Delta L_n, X_n + \Delta X_n\big)\Big], \\ & \min_{\Delta M} \exp\bigg(-\frac{\gamma}{\delta(t_n, T)} (1 - \sigma_s) e^{X_n} \Delta M_n \bigg) \mathbb{E}\Big[Q\big(t_{n+1}, Y_n - \Delta M_n, X_n + \Delta X_n\big)\Big] \bigg\}. \end{aligned}$$

where all the expectations are conditioned on the current state (Y_n, X_n).
The algorithm runs as follows:

1. Create the log-price lattice for the Markov chain (38). This usually has the shape of a recombining multinomial tree with $\bar{L} = K_1 + K_2 + 1$ branches and step size $(\Delta t, h_x)$. The transition probabilities are evaluated from the explicit finite differences discretization of the infinitesimal generator (see Appendix B). The reason for using a multinomial tree instead of a rectangular lattice (like the finite differences methods) is that we do not need to introduce artificial lateral boundary conditions.

2. Create the shares vector y with discretization step h_y. Its dimension is $\bar{M} = K_3 + K_4 + 1$

3. Create the value functions at terminal time using the terminal conditions (28),(30),(31). The result are two-dimensional grids with dimensions $\big(N(\bar{L} - 1) + 1\big) \times \bar{M}$.

4. Create a backward loop over time, with index n that starts at N and finishes at 1.

5. Given the value functions at time n, compute the value functions at time $n - 1$ on a new grid of size $\big((n-1)(\bar{L} - 1) + 1\big) \times \bar{M}$. This is done in two steps:

    - *Time step*: For each node of the tree at time $n - 1$ and for each y, take a weighted average of the value function at time n over all the possible nodes at time n connected with the starting node. We use the discretization presented in Appendix B.2. (see Eq. (54))
    - *Control step*: Compute the minimum of the second term in Eq. (42) for all possible increments of ΔL (of size h_y). Do the same for the

third term for all size ΔM, and then compute the minimum between the three terms.

6. Use formulas (32) or (33) to find the option price.

4 Numerical results

In this section we implement the algorithm and calculate the price of a European call option for the writer and the buyer. The fixed parameters of the option and the Merton model (36) are as follows:

K	T	r	μ	σ	α	ξ	λ
15	1	0.1	0.1	0.25	0	0.5	0.8

With these parameters it is possible to compute the option price with no transaction costs using the standard *martingale pricing theory* (see Appendix A). We computed the price using the series approximation formula of Merton (1976), using Monte Carlo simulation [2] and solving the Merton PIDE [3] (46).
For $S_0 = 15$, the "at the money" price is:

Merton formula	Monte Carlo	Merton PIDE
3.4776	3.4779 ± 0.0903	3.4749

The PIDE price is our reference price. Of course, the parameter μ has not been used to compute these prices. We will see that even in our model the drift parameter μ does not play an important role.

We used a stock log-price tree with $N = 100$ time steps, and number of branches $\bar{L} = 61$. These two values are not independent. Every Lévy process that satisfies the condition (3), satisfies the square root rule. Therefore the size of a space step is $h_x = \sqrt{\mathbb{E}[\Delta X^2]}$ which corresponds to $h_x = \sqrt{(\sigma^2 + \lambda\xi^2 + \lambda\alpha^2)\,\Delta t} = \sigma_X \sqrt{\Delta t}$ (see Eq. (40)). However, the Poisson jump sizes do not scale with Δt. So the number $\bar{L}$ has to be chosen big enough in order to cover the domain of the jump sizes with the discrete tree. We consider the jump domain at least bigger than three standard deviations of the jump density. This is an indicative value, and a bigger domain gives a more precise result. But the computational time can increase a lot.
The constraint is $\bar{L}\, h_x \geq 3\xi$. The time increment is defined as $\Delta t = T/N$, so $h_x = \sigma_X \sqrt{T/N}$. Using our parameters and putting together the formulas we obtain the relation $\bar{L} \geq 5.7\sqrt{N}$.

The prices are computed with a small risk aversion coefficient γ. We illustrate in Figure 1 that our model replicates the Merton model prices in the case of zero transaction costs, $\mu = r$ and γ small close to zero. An intuitive argument to explain this feature is that for small values of γ, the utility function looks like a linear utility function $\mathcal{U}(w) = 1 - e^{-\gamma w} \approx \gamma w$. Therefore the investor profile

[2] The price is the average of 10000 Monte Carlo runs, each with 10000 paths. The error is the standard deviation.

[3] We solved the PIDE (46) with boundary conditions as in Appendix A using the Implicit-Explicit method proposed in Cont and Voltchkova (2005). We used a discretization as in Appendix B.2 for the continuous differential and integral operators.

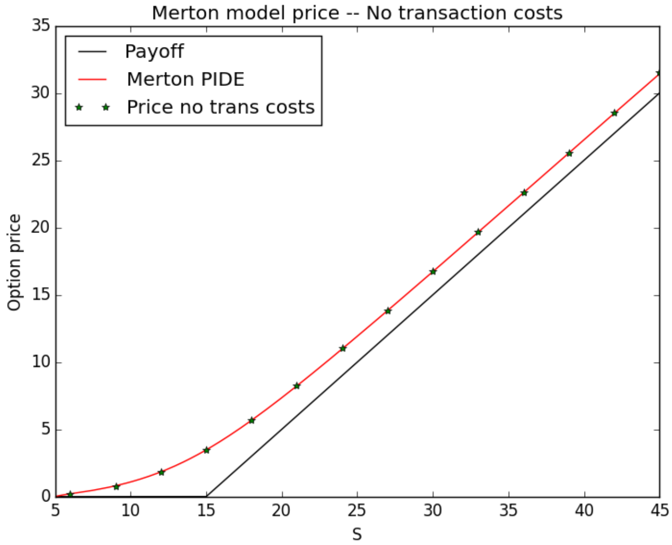

Figure 1: The model prices correspond to the Merton prices with zero transaction costs, $\mu = r = 0.1$ and $\gamma = 0.04$. The other paramethers are in the tables.

is risk neutral. For a formal proof we refer to Carmona (2009) and references therein.
The following table shows how the writer prices change when introducing the transaction costs. We consider same transaction costs for buying and selling $\sigma_b = \sigma_s$, and $\gamma = 0.04$.

cost = 0	cost = 0.01	cost = 0.02	cost = 0.03	cost = 0.04
3.4771	3.6400	3.8212	4.0054	4.1864

In Figure 2 and 3 we present the option prices for the writer and buyer respectively, with different values of transaction costs. We can see that the prices for the writer are an increasing function of the transaction costs, while the prices for the buyer a are decreasing function. This property is verified in Clewlow and Hodges (1997) for the case of the diffusion process.
In Figure 4 we can see how the price for the writer is affected by the change of the transaction costs. The picture shows prices for different values of the risk coefficient.

The risk profile of the investor also plays an important role. As already shown in Davis et al. (1993), the option price is an increasing function of the risk aversion. The Figure 5 confirms their results.
In our computations, we always used the drift term μ equal to the risk free interest rate r. This is the same choice of Hodges and Neuberger (1989). They do not explain the reasons for this choiche, but follow the common knowledge that the option price is independent of the expected return of the underlying asset. However, it turns out that this empirical fact is still true in this model, as we can see in Figure 6. This feature of the model has been analized in Damgaard

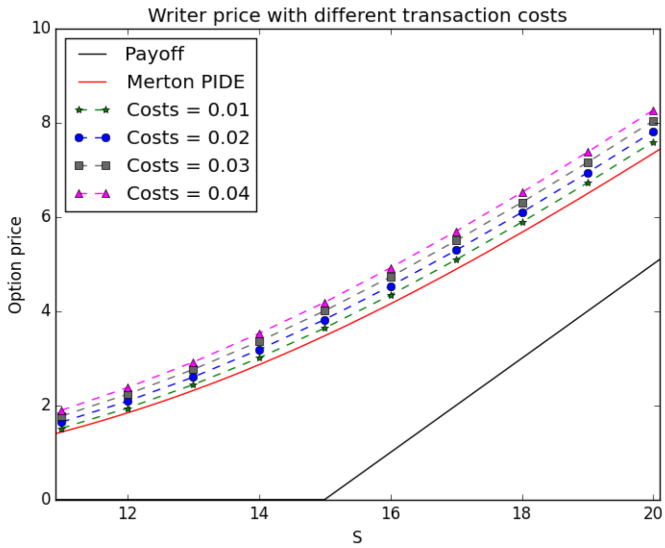

Figure 2: Option prices for the writer with different transaction costs. $\gamma = 0.04$, $\sigma_b = \sigma_s$. The other paramethers are in the tables.

Figure 3: Option prices for the buyer with different transaction costs. $\gamma = 0.04$, $\sigma_b = \sigma_s$. The other paramethers are in the tables.

(1998) for the diffusion case. This numerical experiment confirm that the option prices are not very sensitive to the change of the drift μ.

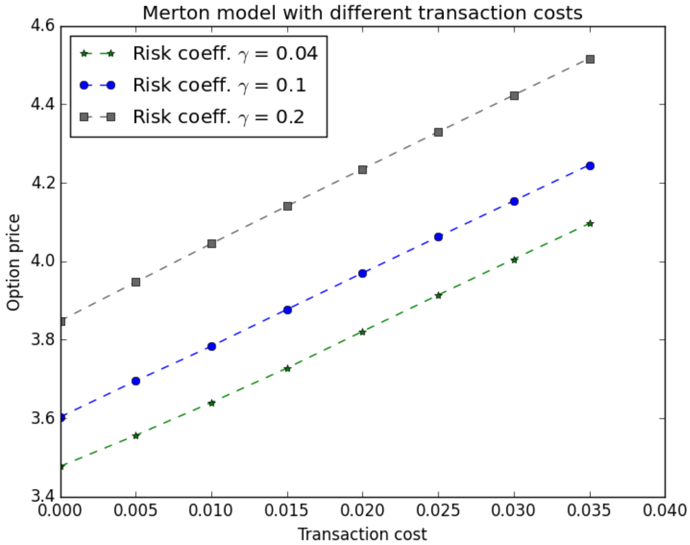

Figure 4: Option prices for the writer as a function of the transaction costs, $\sigma_b = \sigma_s$. Different values of γ.

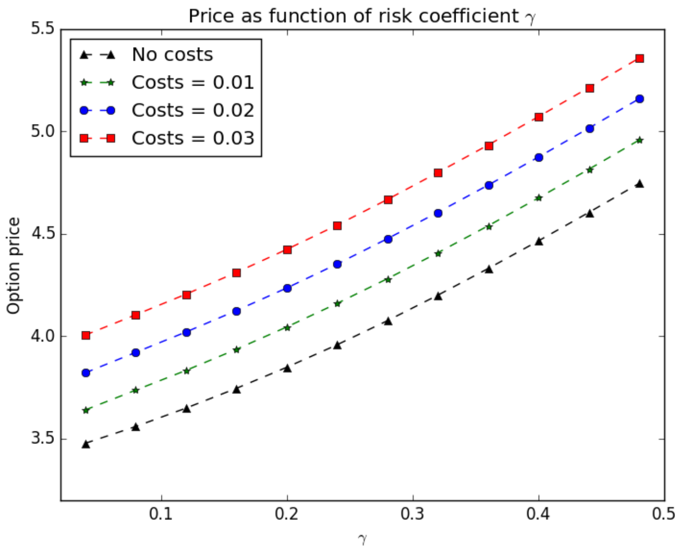

Figure 5: Option prices for the writer as a function of γ. Different values of transaction costs. $\sigma_b = \sigma_s$.

5 Conclusions

We presented a model for pricing options in presence of proportional transaction. This is an extension of the model first introduced by Hodges and Neuberger

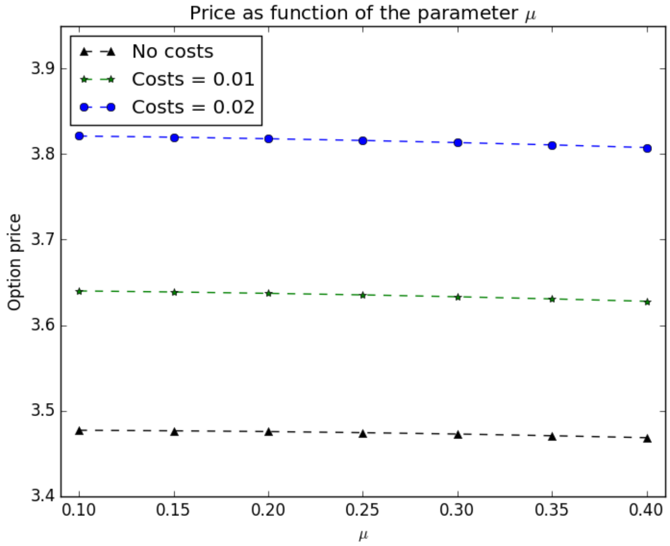

Figure 6: Option prices for the writer as a function of μ. $\gamma = 0.04$ Different values of transaction costs $\sigma_b = \sigma_s$.

(1989) and then formalized by Davis et al. (1993). The main difference between this work and the former works in the literature is that we considered a stock dynamics that follows a general exponential Lévy process. The presence of jumps in the stock dynamics is the cause of portfolio bankruptcy. In this model we consider the theoretical possibility of default. It is thus possible to compute option prices for investors with different risk profile and different wealth. A small firm can have a thin solvency region, therefore its set of trading strategies may be quite smaller than the respective set of a big firm. The option price changes depending on the size of the company. We can give an intuitive explanation: if a company with a high default probability wants to sell an option, it has to sell it at a smaller price respect to a company with smaller default probability, in order to include a reward for the buyer who is taking a higher risk.

In this paper we don't consider the possibility of bankruptcy in the numerical computations. The general equation (20) is a complicate equation with three state variables and a time variable, so we opted to consider only investors "too big to fail". Thanks to this assumption and the choice of an exponential utility, it was possible to reduce the number of variables of the problem, obtaining the simpler HJB equation (34). The solution of the general equation (20) and the numerical investigation of how the option price changes under different initial wealth, are interesting ideas for a future research project.

The optimization problem has been solved with the Markov chain approximation method. The same method has been used frequently in the literature in the case of diffusion processes by Davis et al. (1993), Davis and Panas (1994), Clewlow and Hodges (1997), Damgaard (1998) and Monoyios (2003).

We considered the case of the Merton jump-diffusion model, although any Lévy process satisfying the conditions (3) can be used. A future work can be to

consider a process with infinite activity of jumps (such as a Variance Gamma or a NIG process).
Using numerical experiment, we confirm some features of the model already proven for the diffusion case:

1. The writer prices in presence of transaction costs are always greater than the prices computed with zero transaction costs. The option value is an increasing function of the transaction cost parameters. Respectively, the buyer prices are always smaller that the prices with no transaction costs, and are a decreasing function of the transaction costs.

2. The option price is an increasing function of the risk aversion coefficient.

3. The underlying drift is not a relevant parameter for the computation of the option.

In this work we concentrated on the evaluation of the option price and do not consider the problem of hedging. A future improvement can be to find the free boundaries that divide the No Transaction region from the Buy and Sell regions. These boudaries serve as an indicator that says when it is optimal to trade, and are needed in the hedging practice.
Another direction for future improvements, is the development a more efficient numerical method to solve the HJB equation. There are different approachs in the literature to solve variational inequalities, such as the penalty method. The multinomial method we considered corresponds to an explicit scheme. We argue that an implicit scheme (or implicit/explicit if considering jumps) can strongly increase the efficiency of the numerical method.

Acknowledgements

Our sincere thanks go to the Department of Mathematics, ISEG and CEMAPRE, University of Lisbon, `http://cemapre.iseg.ulisboa.pt/`.
This research was supported by the European Union in the FP7-PEOPLE-2012-ITN project STRIKE - Novel Methods in Computational Finance (304617), and by CEMAPRE MULTI/00491 financed by FCT/MEC through Portuguese national funds. We wish also to acknowledge all the members of the STRIKE network, `http://www.itn-strike.eu/`.

A Martingale option pricing

Under a risk neutral measure $\mathbb{Q}$, the dynamics of the stock price is described by the following *exponential Lévy model*:

$$S_t = S_0 e^{rt+X_t} \tag{43}$$

where r is the risk free interest rate, and X_t is a Lévy process with Lévy triplet (b, σ, ν). Under $\mathbb{Q}$ the discounted price is a $\mathbb{Q}$-martingale:

$$\mathbb{E}^{\mathbb{Q}}\big[e^{-rt} S_t \big| S_0\big] = \mathbb{E}^{\mathbb{Q}}\big[S_0 e^{X_t} \big| S_0\big] = S_0, \tag{44}$$

such that $\mathbb{E}^{\mathbb{Q}}[e^{X_t}|X_0 = 0] = 1$. This condition is equivalent to the condition for the triplet:

$$b = -\frac{1}{2}\sigma^2 - \int_{\mathbb{R}} (e^z - 1 - z1_{|y|<1})\nu(dz).$$

Let $C(t, S_t)$ be the value of a European call option at time t. By the *martingale pricing theory* the discounted price of the option is a martingale. It is possible to derive the partial integro-differential equation (PIDE) for the price of the option.

$$\mathbb{E}^{\mathbb{Q}}\Big[d\big(e^{-rt}C(t, X_t)\big)\Big] = \frac{\partial C(t, x)}{\partial t} + \mathcal{L}^{X_t}C(t, x) - rC(t, x) = 0 \tag{45}$$

where $\mathcal{L}^{X_t}$ is the infinitesimal generator of X_t. For a Merton model, with generator as in (37), the resulting equation is called *Merton PIDE*:

$$\begin{aligned} &\frac{\partial C(t, x)}{\partial t} + (r - \frac{1}{2}\sigma^2 - m)\frac{\partial C(t, x)}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 C(t, x)}{\partial x^2} \\ &+ \int_{\mathbb{R}} C(t, x + y)\nu(dy) = (r + \lambda)C(t, x) \end{aligned} \tag{46}$$

The value of the call option is given by the solution of the PIDE (46) with the usual boundary conditions:

- Payoff:
$$C(T, x) = \max\{e^x - K, 0\}.$$
- Lateral conditions:
$$C(t, 0) = 0 \qquad \text{and} \qquad C(t, x) \overset{x \to \infty}{=} e^x - Ke^{-r(T-t)}.$$

B Properties of the Markov chain

We explained in the Section 3.2 that the Markov chain approximation of a continuous time jump-diffusion process has to satisfy two properties. This section makes a summary of the key concepts and refer to Kushner and Dupuis (2001) for detailed definitions and proofs of convergence.

B.1 Transition probabilities

The random components of the SDE (36) for the log-price are two independent stochastic processes: a Brownian motion and a compound Poisson process. The number of jumps is Poisson distributed:

$$\mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \tag{47}$$

For a small Δt, we can compute the first order approximated probability jumps:

- $\mathbb{P}(N_{t+\Delta t} - N_t = 0) \overset{d}{=} \mathbb{P}(N_{\Delta t} = 0) = e^{-\lambda \Delta t} \approx 1 - \lambda \Delta t$,
- $\mathbb{P}(N_{\Delta t} = 1) = e^{-\lambda \Delta t}(\lambda \Delta t) \approx \lambda \Delta t$,

- $\mathbb{P}(N_{\Delta t} > 0) = 1 - \mathbb{P}(N_{\Delta t} = 0) \approx \lambda \Delta t$.

At first order in Δt the probability of one jump is the same as the probability of any positive number of jumps.
In the discretized SDE (38), we can assume that in a small time step Δt the process or jumps exactly once, or does not jump at all. At every time step Δt, there are two possible independent moves:

1. *diffusion*: with transition probability $p^D(x,z)$ with $z = x + Z^D$ and the random variable $Z^D \in \{-h_x, 0, h_x\}$,
2. *jumps*: with transition probability $p^J(x,z)$ with $z = x + Z^J$. The random variable Z^J takes values is Σ_x.

The total transition probability for the process X_n is

$$p(x,z) = \mathbb{P}(X_{n+1} = z | X_n = x).$$

We can compute by conditioning on the values of the Poisson process:

$$\begin{aligned} p(x,z) &= p^D(x,z)\,\mathbb{P}(N_{\Delta t} = 0) + p^J(x,z)\,\mathbb{P}(N_{\Delta t} = 1) \\ &= (1 - \lambda \Delta t)\, p^D(x,z) + (\lambda \Delta t)\, p^J(x,z) \end{aligned} \tag{48}$$

The first property of the Markov chain approximation of a jump-diffusion process is that the one step transition probability can be represented in the form (48), at first order in Δt.

B.2 Local consistency

The second property says that the first two moment of the increment ΔX_n of the Markov chain in a time step Δt, are the same of corresponding increment ΔX_t of the continuous time process (at first order in Δt).

In order to check that the Markov chain (38) satisfies this property, together with the first property, it is necessary to find the explicit form of the transition probabilities. This is achieved by discretizing the infinitesimal generator of the uncontrolled process, that corresponds to the first term inside the "max" in the HJB equation (27), using an explicit finite different scheme. For the Merton model we derived the form of the infinitesimal generator in Eq. (37).
We use the short notation: $Q(t_n, y_j, x_i) = Q_i^n$ where we drop the variable y, because we are considering the uncontrolled dynamics. The derivatives are discretized by the finite differences:

- Backward approximation in time: $\frac{\partial Q}{\partial t} \approx \frac{Q_i^{n+1} - Q_i^n}{\Delta t}$.
- Central approximation in space: $\frac{\partial Q}{\partial x} \approx \frac{Q_{i+1}^{n+1} - Q_{i-1}^{n+1}}{2h_x}$.
- Second order: $\frac{\partial^2 Q}{\partial x^2} \approx \frac{Q_{i+1}^{n+1} + Q_{i-1}^{n+1} - 2Q_i^{n+1}}{h_x^2}$.

The integral term in (37) is truncated and restricted to the domain $[B_1, B_2] = [(-K_1 - 1/2)h_x, (K_2 + 1/2)h_x]$. The discretization is obtainded by trapezoidal quadrature (see Cont and Tankov (2003)):

$$\int_{B_1}^{B_2} Q(t_{n+1}, y_j, x_i + z)\nu(dz) \approx \sum_{k=-K_1}^{K_2} \nu_k Q_{i+k}^{n+1} \tag{49}$$

where

$$\nu_k = \int_{(k-\frac{1}{2})h_x}^{(k+\frac{1}{2})h_x} \nu(z)dz, \quad \text{for} \quad -K_1 < k < K_2. \tag{50}$$

and for $k = -K_1$ and $k = K_2$:

$$\nu_{-K_1} = \int_{-\infty}^{(-K_1+\frac{1}{2})h_x} \nu(z)dz, \qquad \nu_{K_2} = \int_{(K_2-\frac{1}{2})h_x}^{\infty} \nu(z)dz. \tag{51}$$

As long as $\lambda = \sum_{k=-K_1}^{K_2} \nu_k$, the jump transition probabilities can be defined as:

$$p^J(x_i, x_{i+k}) = \frac{\nu_k}{\lambda}. \tag{52}$$

The discretized equation becomes:

$$\begin{aligned} &\frac{Q_i^{n+1} - Q_i^n}{\Delta t} + (\mu - \frac{1}{2}\sigma^2 - m)\frac{Q_{i+1}^{n+1} - Q_{i-1}^{n+1}}{2h_x} \\ &+ \frac{1}{2}\sigma^2 \frac{Q_{i+1}^{n+1} + Q_{i-1}^{n+1} - 2Q_i^{n+1}}{h_x^2} + \sum_{k=-K_1}^{K_2} \nu_k Q_{i+k}^{n+1} - \lambda Q_i^n = 0 \end{aligned} \tag{53}$$

Rearranging the terms:

$$\begin{aligned} \Big(1 + \lambda\Delta t\Big) Q_i^n = \; & Q_i^{n+1}\Big(1 - \sigma^2 \frac{\Delta t}{h_x^2}\Big) \\ & + Q_{i+1}^{n+1}\Big((\mu - \frac{1}{2}\sigma^2 - m)\frac{\Delta t}{2h_x} + \frac{1}{2}\sigma^2 \frac{\Delta t}{h_x^2}\Big) \\ & + Q_{i-1}^{n+1}\Big(-(\mu - \frac{1}{2}\sigma^2 - m)\frac{\Delta t}{2h_x} + \frac{1}{2}\sigma^2 \frac{\Delta t}{h_x^2}\Big) \\ & + (\lambda\Delta t) \sum_{k=-K_1}^{K_2} p^J(x_i, x_{i+k}) Q_{i+k}^{n+1} \end{aligned}$$

The diffusion transition probabilities $p^D(x_i, x_{i+k})$ are the coefficients of the terms Q_{i-1}^n, Q_i^n, Q_{i+1}^n, for $k \in \{-1, 0, 1\}$ and identically equal to zero for other values of k.

We bring the term $\big(1 + \lambda\Delta t\big)$ on the right hand side and use the first order Taylor approximation $\big(1 + \lambda\Delta t\big)^{-1} \approx 1 - \lambda\Delta t$ for small time step Δt, in order to obtain:

$$\begin{aligned} Q_i^n = & \sum_{k=-K_1}^{K_2} p(x_i, x_{i+k})\, Q_{i+k}^{n+1} \\ = & \big(1 - \lambda\Delta t\big) \sum_{k=-1}^{1} p^D(x_i, x_{i+k})\, Q_{i+k}^{n+1} + \big(\lambda\Delta t\big) \sum_{k=-K_1}^{K_2} p^J(x_i, x_{i+k})\, Q_{i+k}^{n+1}, \end{aligned} \tag{54}$$

where $p(x_i, x_{i+k}) = (1 - \lambda\Delta t)p^D(x_i, x_{i+k}) + (\lambda\Delta t)p^J(x_i, x_{i+k})$ is the total transition probability. It is straightforward to check that all the terms sums to one.

The transition probabilities have the form (48), so this Markov chain approximation satisfies the first property condition.
Lets check that also the *local consistency* conditions (39),(40) are satisfied:

$$\begin{aligned}\mathbb{E}\big[\Delta X_n\big] =&\big(1-\lambda\Delta t\big)\sum_{k=-1}^{1} p^D(x_i,x_{i+k})\,kh_x + \big(\lambda\Delta t\big)\sum_{k=-K_1}^{K_2} p^J(x_i,x_{i+k})\,kh_x \\ =&\big(1-\lambda\Delta t\big)\big(\mu-\frac{1}{2}\sigma^2-m\big)\,\Delta t+\big(\lambda\Delta t\big)\,\alpha \\ \approx&\big(\mu-\frac{1}{2}\sigma^2-m+\lambda\alpha\big)\,\Delta t,\end{aligned}$$

where we considered only terms at first order in Δt. The sum that involves the jump probabilities converges to the expected value of the size of the jumps ($\sim\mathcal{N}(\alpha,\xi)$). Use (50),(51),(52) and $K_1,K_2\to\infty$.
For the second moment:

$$\begin{aligned}\mathbb{E}\Big[\big[\Delta X_n\big]^2\Big] =&\big(1-\lambda\Delta t\big)\sum_{k=-1}^{1} p^D(x_i,x_{i+k})\,(kh_x)^2 \\ &+\big(\lambda\Delta t\big)\sum_{k=-K_1}^{K_2} p^J(x_i,x_{i+k})\,(kh_x)^2 \\ =&\big(1-\lambda\Delta t\big)\sigma^2\,\Delta t+\big(\lambda\Delta t\big)\,\big(\xi^2+\alpha^2\big) \\ \approx&\big(\sigma^2+\lambda\xi^2+\lambda\alpha^2\big)\,\Delta t.\end{aligned}$$

where again we do not consider any second order term in Δt. The local consistency property is thus satisfied.

References

Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*. Cambridge University Press; 2nd edition.

Barles, G. and Soner, H. M. (1998). Option pricing with transaction costs and a nonlinear black scholes equation. *Finance and Stochastics*, 2(4):369–397.

Barles, G. and Souganidis, P. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *J. Asymptotic Analysis*, 4:271–283.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654.

Boyle, P. and Vorst, T. (1992). Option replication in discrete time with transaction costs. *The Journal of Finance*, 47(1):271–293.

Carmona, R. (2009). *Indifference Pricing*. Princeton University Press.

Clewlow, L. and Hodges, S. (1997). Optimal delta hedging under transaction costs. *Jornal of Economic Dynamics and Control*, 21:1353–1376.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236.

Cont, R. and Tankov, P. (2003). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC; 1 edition.

Cont, R. and Voltchkova, E. (2005). A finite difference scheme for option pricing in jump diffusion and exponential lévy models. *SIAM Journal of numerical analysis*, 43(4):1596–1626.

Cox, J., Ross, S., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial economics*, 7:229–263.

Damgaard, A. (1998). *Optimal Portfolio Choice and Utility Based Option Pricing in Markets with Transaction Costs*. PhD thesis, School of Business and Economics, Odense University.

Davis, M. H. A. and Panas, V. G. (1994). The writing price of a european contingent claim under proportional transaction costs. *Computational and Applied Mathematics*, 13(2):0101–8205.

Davis, M. H. A., Panas, V. G., and Zariphopoulou, T. (1993). European option pricing with transaction costs. *SIAM J. Control Optim.*, 31(2):470–493.

Fleming, W. H. and Soner, M. H. (2005). *Controlled Markov Processes and Viscosity Solutions*. Springer; 2nd edition.

Hodges, S. D. and Neuberger, A. (1989). Optimal replication of contingent claims under transaction costs. *The Review of Future Markets*, 8(2):222–239.

Kushner, H. and Dupuis, P. G. (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer; 2nd ed.

Kushner, H. J. and Martins, F. L. (1991). Numerical methods for stochastic singular control problems. *SIAM Journal of Control and Optimization*, 29(6):1443–1475.

Leland, H. (1985). Option pricing and replication with transaction costs. *The Journal of Finance*, 40(5):1283–1301.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

Monoyios, M. (2003). Efficient option pricing with transaction costs. *Journal of Computational Finance*, 7:107–128.

Monoyios, M. (2004). Option pricing with transaction costs using a markov chain approximation. *Jornal of Economic Dynamics and Control*, 28:889–913.

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

Whalley, A. E. and Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7(3):307–324.