OPTION PRICING WITH MARKET IMPACT AND NON-LINEAR BLACK AND SCHOLES PDES'S

GRÉGOIRE LOEPER

Abstract. We propose a few variations around a simple model in order to take into account the market impact of the option seller when hedging an option. This "retro-action" mechanism turns the linear Black and Scholes PDE into a non-linear one. This model allows also to retrieve some earlier results of [9]. Numerical simulations are then performed.

1. Introduction

We are interested in the derivation of a pricing model for options that takes into account the market impact of the option issuer, i.e. the feedback mechanism between the option hedging induced stock trading activity and the price dynamics. For a start, we will assume a linear market impact: each order to buy N stocks impacts the stock price by λNS^2 ($\lambda \geq 0$). The scaling in S^2 means that the impact in relative price move depends on the amount of stock traded expressed in currency, hence λ is homogeneous to the inverse of the currency *i.e. percents per dollar*.

$$\text{Order to buy N stocks} \implies S \to S(1 + \lambda NS). \tag{1}$$

In the literature devoted to the study of the market impact, it is more frequent to find an impact that varies as a power law of the size of the trade (see for example [2] for a detailed analysis of the subject). Although our linear approach is clearly not the most realistic, it has the advantage of avoiding arbitrage opportunities, as well as not being sensitive to the hedging frequency: in a non-linear model, splitting an order in half and repeating it twice would not yield the same result, a situation that we want to avoid here. Later on, we will show that this can be refined to other "non-linear" situations, by allowing λ to depend on the *trading intensity*.

In addition we assume that there is no lasting effect of trade orders. This seems in contradiction with other studies (see Bouchaud [3] for example) that give a long term correlation between an order given at time t, and returns at times $t+\tau, \tau > 0$. However this impact is due to large orders that are split into several small pieces, hence to correlation between market orders. This point of view is not relevant here, as we are concerned with our *own* market impact.

In terms of concrete applications, this problem arises notably when equity derivatives houses trade corporate deals, (i.e. sell options to a company on its own stock, for example when hedging a stock option's plan), as the amount engaged are often large compared to the average daily volume traded. In such situations, the option hedger can not immediately unwind his position in the market, and he has to anticipate the feedback mechanism between its own hedging activity and the price dynamic. This is also observed when a large long call position is hedged near the maturity. The hedging activity has the effect to stick the price around the strike, an effect known as *stock pinning*.

BNP Paribas Equities and Derivatives.
Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie.
Chaire de Finance Quantitative, Ecole Centrale Paris.

There is already a substantial amount of literature on the subject of option hedging with transaction costs, market impact, or liquidity constraints. We mention the related works by Soner and Touzi [11] and Cheridito Soner Touzi [9] that deal with the problem of liquidity constraints, as well as the work [4] by Cetin, Jarrow and Protter and [5] by Cetin, Soner and Touzi that deal with the problem of liquidity (or transaction) costs. Recently Almgren et al. [1] have also looked at the subject of hedging with market impact. Our approach through the "physical" incorporation of the market impact is clearly related to those works, and a more thorough discussion of the consistency between our results and theirs will be done hereafter.

The aim of the paper is to propose a simple and direct approach to derive a pricing equation that is adaptable to a wide range of situations. This can be also seen as related to the approach of [8] that derive non-linear heat equations starting from a stationary game approach. We will also provide a numerical implementation of the non-linear Black and Scholes pde that appear. One of the key points in the paper is the incorporation of the market impact in the Black and Scholes pde that is turned into the original form (when using rescaled variables)

$$\frac{1}{\partial_t u} + \frac{1}{s^2 \partial_{ss} u} = \lambda(\partial_t u),$$

where the function $\partial_t u \to \lambda(\partial_t u)$ accounts for the *trading intensity* dependent market impact.

The rest of the paper is organized as follows: through P&L calculations, we first derive in the simple linear case the pricing pde. We then extend this approach to more general non-linear market impact model which in turn can be extended to the known *gamma max* model (treated in [11]). We then show rigorously how the the gamma max model can be obtained as a limit of well chosen market impact models. We then discuss the consistency of our approach with the above mentioned papers. Finally we present the results of a numerical implementation.

2. Derivation of the pricing equation

2.1. **The linear case.** We assume that we have **sold** an option whose value is $u(S, t)$, and greeks are as usual

$$\begin{aligned} \Delta &= \partial_s u, \\ \Gamma &= \partial_s \Delta, \\ \Theta &= \partial_t u. \end{aligned}$$

We also introduce the Gamma *in currency*, i.e.

$$\Gamma_c = \Gamma S^2. \tag{2}$$

Assume the stock price S moves by dS. A naive trader would want to buy ΓdS stocks, but our smart trader anticipates that his order will impact the market (if $\Gamma > 0$ the hedging trade will amplify the initial move and conversely otherwise), hence will change the spot move of dS to μdS for some μ to be found. He then buys instead $\mu \Gamma dS$. Let us now compute the value of μ: we want to be delta-hedged when the spot reaches its final value, i.e. $S + \mu dS$. Hence using (1), we write that

$$S_{\text{after re-hedging}} - S_{\text{before re-hedging}} = \lambda S^2 \text{ Number of stocks bought to re-hedge},$$

and this yields

$$\mu dS - dS = \lambda[\Delta(S + \mu dS) - \Delta(S)]S^2.$$

This identity expresses the fact that the number of titles we have bought is $\Delta(S+\mu dS)-\Delta(S)$, hence that we are Delta-hedged at the end of the day. Performing a Taylor expansion, we get that

$$\Gamma\mu dS\lambda S^2=(\mu-1)dS, \tag{3}$$

which yields

$$\mu=\frac{1}{1-\lambda\Gamma_c}. \tag{4}$$

Remember that Γ is computed with respect to the option we have sold, hence $\Gamma>0$ when we sell the call, and we see that in this case, assuming that $\lambda\Gamma_c$ is not too big (we'll discuss that later), we get $\mu>1$, hence the hedger increases the volatility by buying when the spot rises, and selling when it goes down. Equation (4) can be interpreted as follows: after the initial move $S\to S+dS$, the hedge is adjusted "naively" by ΓdS stocks, which then impacts the price by $\lambda S^2\Gamma dS$, which in turn impacts the hedge by $\Gamma(\lambda S^2\Gamma dS)$, etc ... The final spot move is thus the sum of the geometric sequence

$$dS(1+\lambda S^2\Gamma+(\lambda S^2\Gamma)^2+...)=\frac{1}{1-\lambda\Gamma_c^2}.$$

One sees right away that in this sequence, the critical point is reached when the "first" rehedge (i.e. buying ΓdS stocks after the initial move of dS) doubles the initial move: the sum will not converge. This situation will discussed hereafter.

Then the value of our portfolio Q, containing -1 option + Δ stocks at the beginning of the day, and $\Delta(S+\mu dS)$ stocks at the end of the day evolves as

$$dQ=-du+\Delta dS+\mathcal{R}$$

and $\mathcal{R}$ is the P&L realized during the re-hedging (in the usual Black and Scholes derivation, this term is zero). Our approach is thus the following: we first observe a spot move of dS, and then we re-hedge, and during this operation, the only spot moves are due to our market impact.

We assume that during the re-hedging S evolves along

$$S_\theta:=\theta(S+\mu dS)+(1-\theta)(S+dS),\theta\in[0,1],$$

and similarly, Δ (the number of stocks that we hold) evolves along

$$\begin{aligned}\Delta_\theta &:= \Delta(S)+\theta(\Delta(S+\mu dS)-\Delta(S))\\ &= \Delta(S)+\theta\Gamma\mu dS,\theta\in[0,1].\end{aligned}$$

The computation of $\mathcal{R}$ then yields

$$\begin{aligned}\mathcal{R} &= \int_0^1\Delta_\theta dS_\theta\\ &= \int_0^1(\Delta_\theta-\Delta_0+\Delta_0)dS_\theta\\ &= \int_0^1(\Delta_\theta-\Delta_0)dS_\theta+\Delta_0(\mu-1)dS\\ &= \int_0^1\Gamma\mu dS\theta(\mu-1)dSd\theta+\Delta_0(\mu-1)dS\\ &= \frac{1}{2}\Gamma\mu(\mu-1)(dS)^2+\Delta_0(\mu-1)dS.\end{aligned}$$

The result is expected: the first term amounts to say that we have bought (resp. sold) $\Gamma\mu dS$ stocks at a price lower (resp. higher) by $(\mu-1)dS/2$ than the "final" price. This is similar to a transaction cost, but *in favour* of the hedge provider, which is somehow surprising (this could not be turned into real money, as reverting the operation would have the inverse market impact). The second term is the usual profit obtained by holding Δ_0 stocks. We assume that the option is sold at its fair price, hence $dQ = 0$ (we neglect the interest rates, but they can be added to the study without modifying its conclusions). Then we have as S moves to $S + \mu dS$,

$$du = \Theta dt + \Delta \mu dS + \frac{1}{2}\Gamma(\mu dS)^2,$$

and we thus get the following Black&Scholes identity

$$-\Theta dt + \Gamma(dS)^2/2\left[-\mu^2 + \mu(\mu - 1)\right] = 0,$$

hence writing

$$\begin{aligned} -\mu^2 + \mu(\mu-1) &= -\mu \\ &= -\frac{1}{1-\lambda\Gamma_c}, \end{aligned}$$

we obtain the following pde:

$$\partial_t u + \frac{1}{2}\sigma^2 F(S^2\partial_{ss}u) = 0, \tag{5}$$

$$F(\Gamma_c) = \frac{\Gamma_c}{1-\lambda\Gamma_c}. \tag{6}$$

(Remember that $\Gamma_c = S^2\partial_{ss}u$.) This can be formally rewritten as

$$\frac{2}{\sigma^2}\frac{1}{\partial_t u} + \frac{1}{s^2\partial_{ss}u} = \lambda. \tag{7}$$

When $\lambda \to 0$, we note that we recover the usual B&S equation. This equation clearly poses problems when $1 - \lambda\Gamma_c$ goes to 0, and even becomes negative. This case arises when one has sold a convex payoff ($\Gamma > 0$) and an initial move of dS would be more than doubled during a naive re-hedging (see above) due to our market impact. Intuitively, we run after the spot, and each time we re-hedge, it goes away further. Thus if we hedge naively, the spot will clearly go away very quickly. This arises near the maturity of the product, if the spot is close to the strike. Hence, following the model, when the spot moves up, one should **sell** stocks instead of buying, because the market impact will make the spot go down. We rely on the market impact to take back the spot at a level where we are hedged, whereas usally we change our hedge to adapt it to the spot. This situation is clearly not realistic (or at least no trader would accept to do that ...), and is a limit of our linear model.

On the other hand, if we consider a function that satisfies for all time $t \in [0, T]$ the constraint

$$\partial_{ss}u(S) \leq \frac{1}{\lambda S^2}, \tag{8}$$

solving (5, 6) with terminal payoff $u(T, S_T) = \Phi(S_T)$ then our approach is valid, and the pde system (5, 6) yields the exact replication strategy (we shall prove this in the verification theorem hereafter). Note that our approach has been presented in the case of a constant volatility, but would adapt with no modification to another volatility process, local or stochastic. The function F is increasing, which guarantees that the time independent problem is elliptic, and thus the evolution problem is well posed.

Another (informal) way to see the constraint (8), is to consider instead of F

$$\begin{aligned} F(\Gamma_c) &= \frac{\Gamma_c}{1-\lambda\Gamma_c} \text{ if } \lambda\Gamma_c < 1, \\ &= +\infty \text{ otherwise.} \end{aligned}$$

The physical interpretation of the singular part is that areas with large positive Γ (i.e such as $\Gamma_c > \lambda^{-1}$) will be quickly smoothed out and will instantaneously disappear, as if the final payoff was smoothed (again this argument will be made rigorous later on). This amounts to replace the solution u by the smallest function greater than u and satisfying the constraint (8) (a semi-concave envelope, the so-called "face-lifting" in [11]). That could be compared to what is done by practitioners: one replaces then a single call by a strip of calls, in order to cap the Γ. The constraint can be rigorously included expressed by turning the system (5, 6) into

$$\max\{\partial_t u + \frac{1}{2}\sigma^2 S^2 F(S^2\partial_{ss}u), \partial_{ss}u - \frac{1}{\lambda S^2}\} = 0, \tag{9}$$

$$F(\Gamma_c) = \frac{\Gamma_c}{1-\lambda\Gamma_c}. \tag{10}$$

Under this formulation the problem enters into the framework of viscosity solutions, see [6].

Note that on the other hand, areas with large negative Γ (when we buy the call) would have very little diffusion, but pose no theoretical difficulties.

2.2. **The intensity dependent impact.** Let us now turn to a slightly more general heuristic. In the previous approach we have assumed that the market impact of buying N stocks is, in terms of price, $\lambda N S^2$, regardless of the time on which the order is spread. We now define the intensity as follows: let N_t be the number of stocks detained in the portfolio, and assume that N_t is a stochastic process with finite quadratic variation. Let I_t be the square root of the quadratic variation of N_t i.e. $I_t^2 dt = (dN_t, dN_t)$. In the usual Black and Scholes case, $dN_t = \Gamma dS$ so $I_t = \Gamma S\sigma$. To express the intensity in currency so we introduce $I_t^c(t) = S_t I_t$. Note that we have the representation

$$dN_t = I_t dB_t + \nu dt, \tag{11}$$

for some Brownian motion B, drift term ν. We assume now that the market impact λ is a function of the intensity. If the initial move from S to $S + dS$ is still transformed into $S + \mu dS$ by the market impact of the re-hedging trade, the intensity of the hedging-induced trading is $\Gamma S^2\sigma\mu$, and we now have $\lambda = \lambda(\Gamma S^2\mu\sigma)$. From (3), σ being fixed, μ is implicitly deduced from Γ_c through the relation

$$\mu - \Gamma_c\mu\lambda(\Gamma_c\mu\sigma) = 1. \tag{12}$$

We now introduce $F(\Gamma_c) := \Gamma_c\mu$, then F solves

$$\frac{1}{F} + \lambda(\sigma F) = \frac{1}{\Gamma_c}, \tag{13}$$

which implies the ode on F

$$F' = \frac{F^2}{\Gamma_c^2(1-\sigma F^2\lambda'(\sigma F))}. \tag{14}$$

The natural assumption $F(0) = 0$ implies the existence locally around 0 of a non-decreasing solution . Conversely, given F the condition for λ to be well behaved around 0 is that $\mu(\Gamma_c) = \frac{F}{\Gamma_c}$ be differentiable at 0 with $\mu(0) = 1$.

Then the same P&L calculations as above lead to the Black and Scholes pde

$$\partial_t u + \frac{1}{2}\Gamma_c\mu(\Gamma_c)\sigma^2 = 0, \tag{15}$$

which reads also

$$\partial_t u + \frac{1}{2}F(\Gamma_c)\sigma^2 = 0. \tag{16}$$

Note that the conditions stated above guarantee the ellipticity of the operator $u \to F(\gamma_c)$, since F is increasing. Equation (16) can then be rewritten as

$$\theta + \frac{\sigma^2}{2} F(\Gamma_c) = 0,$$

hence in view of (13), the non-linear B&S equation takes the simple form

$$\frac{\sigma^2}{2}\frac{1}{\theta} + \frac{1}{\Gamma_c} = \lambda(\frac{2}{\sigma}\theta), \tag{17}$$

or equivalently

$$\frac{\sigma^2}{2}\frac{1}{\partial_t u} + \frac{1}{s^2\partial_{ss}u} = \lambda(\frac{2}{\sigma}\partial_t u). \tag{18}$$

Note that a simple linear change of variable in time simplifies the above equation in

$$\frac{1}{\partial_t u} + \frac{1}{s^2\partial_{ss}u} = \frac{1}{\Lambda(\partial_t u)}, \tag{19}$$

where $\tilde{\Lambda} = \lambda^{-1}(\sigma \cdot)$.

We summarize the results concerning the compatibility conditions onf F, λ, μ in the following Proposition.

Proposition 1. *For any $\lambda : \mathbb{R} \to [0, +\infty]$ continuous, there exits $\mu(\Gamma_c)$ such that $F : \Gamma_c \to \Gamma_c\mu(\Gamma_c)$ is defined locally around 0 and non decreasing, and $\mu(0) = 1$. Then for all $\Gamma_c \in \mathbb{R}$, either (12) holds or $\mu = +\infty$.*

Conversely, if $\mu(0) = 1$, μ is differentiable at 0 and positive, then there exists λ satisfying (12) whenever μ is defined.

2.3. **The Verification Theorem.** The following result shows that, granted there is a smooth solution of (15), one is able to replicate a contingent claim equal to the terminal value of u at time T, taking into account the market impact effect on the volatility of the underlying and on the P&L of the strategy.

Theorem 2.1. *Let u be a smooth solution of (15) with terminal condition $u(T, \cdot) = \Phi$. Let B_t be the standard Brownian motion with $\mathcal{F}_t$ the associated filtration. Let $\lambda(\Gamma_c\mu\sigma), f(\Gamma_c\mu), \mu(\Gamma_c)$ be defined as above. Let S_t satisfy*

$$\frac{dS_t}{S_t} = \sigma dB_t.$$

Let $\tilde{S}_t$ satisfy

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = \sigma\mu(\Gamma_c)dB_t,$$

with $\Gamma_c = \tilde{S}_t^2\partial_{ss}u(t, \tilde{S}_t)$. Then, almost surely

$$u(0, S_0) + \int_0^T \partial_s u d\tilde{S} + \frac{1}{2}\int_0^T \Gamma_c\sigma^2(\mu^2(\Gamma_c) - \mu(\Gamma_c))dt = \Phi(\tilde{S}_T). \tag{20}$$

Proof. The proof is a simple application of Ito's formula, since we restrict ourselves to the case where the solution of (15) is smooth, and thus solution in the classical sense. □

Remark 1. Note that the P&L generated by the hedging strategy is no more the usual expression $\int_0^T \partial_s u d\tilde{S}$, but includes an additional term due to the market impact.

Remark 2. The dynamic of S will not be directly observed, it is the trajectory that S *would have followed if one had not delta hedged the option.* The observed spot trajectory is $\tilde{S}$. It is still driven by the same noise source as S, and measurable with respect to $\mathcal{F}_t$.

Remark 3. Note that the 3rd term of the left hand side of (20) accounts for extra P&L due to re-hedging. It is always positive (i.e. in favor of the option's seller). As observed above this seems surprising, but note that the change of volatility from σ to $\mu\sigma$ acts always against the option's seller (μ is greater (resp. lower) than one when Γ is positive (resp. negative)) , and the sum of the two impacts is always against the option's seller.

2.4. **Rewriting the market impact in terms of transaction costs on the volatility.** We consider here that there is an active option trading market, so that it is possible to hedge not only the delta, but also the gamma of a position, through short term at the money options. The replication strategy is thus achieved by trading *zero delta* very short term options physically settled (i.e. at maturity the options buyer of a call receives the stock and pays the strike in cash). One can show quite easily that when the hedging frequency goes to infinity, such a strategy allows to perfectly replicate the final payoff. Assume that a given amount of Gamma, say Γ_c can be bought (resp. sold if this amount is negative) at a given implied volatility level $\sigma(\Gamma_c)$. In this case, following a standard replication portfolio argument, the B&S pricing equation becomes (still in the absence of interest rates)

$$\partial_t u + \frac{1}{2}\Gamma_c \sigma^2(\Gamma_c) = 0. \tag{21}$$

One sees that there will be a correspondence with the previous pde

$$\partial_t u + \frac{1}{2}\Gamma_c \mu(\Gamma_c)\sigma^2 = 0,$$

namely through the relation

$$\sigma(\Gamma_c) = \mu^{\frac{1}{2}}(\Gamma_c)\sigma. \tag{22}$$

So to a market impact model (parametrized either by the function λ or directly by μ) one can associate a transaction cost function on the short term volatility, i.e. a *supply curve* (in the spirit of [4]) of volatility $\Gamma_c \to \sigma(\Gamma_c)$.

2.5. **A limit case: The Gamma-max Theta-max model.** An input meaningful for practitioners is the value of the Gamma-max, i.e. the quantity of stocks that one is able to buy/sell on a given market move. Again, this relies on some linearity, i.e. on a twice bigger move, one is able to do a twice bigger transaction.

It is important to understand that this constraint in the "Gamma short" case (one sells the convex payoff) is the most natural one: a trader can not take too much "Gamma-short", moreover it is also the easiest to deal with in terms of smoothing: if suffices to replace the payoff by the smallest function majoring it and satisfying (8). This case has already been treated , see [10] for example. It is shown there that the constraint

$$\partial_{ss} u \leq K(S)$$

is enforced in the B&S pde by solving

$$\max\{\partial_t u + \frac{1}{2}\sigma^2 S^2 \partial_{ss} u, \partial_{ss} u - K(S)\} = 0.$$

Conversely, it is difficult to enforce the constraint $\Gamma \geq -\Gamma_{max}$ by a function majoring the payoff (this would amount to find a semi-convex upper envelope ...). In this case, the constraint realizes as a bound on the theta paid by the option's seller (which in this case has bought a convex payoff / sold a concave payoff). This can be still be understood as a bound on the number of stocks bought/sold on a move of S to $S + dS$: If the option held is has a positive Gamma, following a move of the stock by dS, assuming that one is delta-hedged before the move, the P&L realized is $\Gamma(dS)^2/2 + \Theta dt$. In this case Θ is negative and Γ positive. This P&L is locked only if one can re-hedge after the spot move. The lack of liquidity (i.e. the Gamma-max) prevents from being able to lock this P&L: possibly after the spot move, the market impact of the re-hedge will make the spot go partially backward, cancelling some of the P&L. Hence instead of paying a theta equal to $\frac{1}{2}\Gamma S^2\sigma^2/2$, one agrees to pay only $\frac{1}{2}\Gamma_{max} S^2\sigma^2/2$. The constraint will appear more like a bound on the theta than on the second derivative of the payoff. (The induced hedging strategy will be discussed hereafter).

We introduce the parameter Λ, equal to the Gamma max in currency. In the Gamma max Theta max model, the B&S pde becomes

$$\partial_t u + \frac{1}{2} F(\Gamma_c)\sigma^2 = 0,$$

with

$$\begin{aligned} F(\Gamma_c) &= -\Lambda \text{ for } \Gamma_c \leq -\Lambda, \\ &= \Gamma_c \text{ for } -\Lambda \leq \Gamma_c \leq \Lambda, \\ &= +\infty \text{ for } \Gamma_c \geq \Lambda. \end{aligned}$$

The last line must be understood as "*At all times one enforce the constraint* $\Gamma_c \leq \Lambda$". This last model can be seen as the "cheapest" way smoothing of the option: when the constraint is not saturated, no smoothing is applied. Somehow, the Gamma-max model behaves like "all or nothing", compared to the market impact model. It is easy to see that the function F of the Gamma max model can be approximated by smooth functions of the form $\Gamma_c \to \Gamma_c\mu(\Gamma_c)$, and, indeed, as we will show in next section, this case is a limiting case of the general intensity dependent model. Note also that the Gamma max parameter Λ is closely related to the market impact parameter λ. More precisely, in the linear case, the theta max paid is equal to $\frac{1}{2}\lambda^{-1}$, and the Gamma in currency can not be grater than λ^{-1}. Therefore we hereafter denote $\lambda = \Lambda^{-1}$ to unify this case with the linear case. The functions $\Gamma \to F(\Gamma)$ are represented in Fig. 1.

3. Passing to the limit

Here we give a rigorous result implying that for a well chosen sequence of market impact functions λ (or μ) the solution converges to the solution of the Gamma max Theta max problem. Hence the problem with constraint is recovered as the limit case of unconstrained penalized problems. This result also clarifies the intuition that the non-linearity F has the effect of taking a semi-concave envelope in the regions where the second derivative becomes large. We show the following result:

Theorem 3.1. *Let $\mu_n, n \geq 0$ be a sequence of functions satisfying the conditions of Proposition 1, such that for all n, μ_n is smooth on all $\mathbb{R}$. Assume that μ_n converges to μ such that μ is continuous on $(-\infty, \Gamma_{max}]$ and $\mu \equiv= +\infty$ for $x > \Gamma_{max}$, for some Γ_{max}. Let Φ be a given terminal payoff and for all n let u_n solve (15) on $[0,T] \times (0,+\infty)$ with terminal condition $u_T = \Phi$. Then for all n, u_n is smooth on $[0,T) \times \mathbb{R}$, and the sequence u_n converges locally uniformly in $[0,T) \times (0,+\infty)$ to u viscosity solution of*

$$\max\{\partial_t u + \frac{1}{2}\sigma^2\Gamma_c\mu(\Gamma_c), S^2\partial_{ss}u - \Gamma_{max}\} = 0. \tag{23}$$

Proof. The fact that u_n is smooth comes from the smoothness of μ_n and then follows from classical results on one dimensional diffusion equations. Then the convergence is a consequence of the following intermediary result:

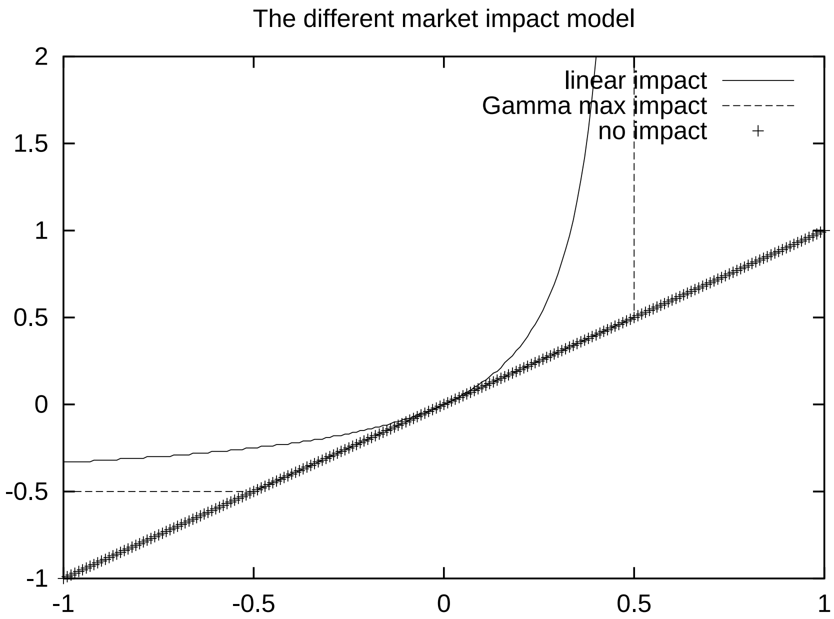

FIGURE 1. The function $\Gamma_c \to F(\Gamma_c)$ used for the different models

Theorem 3.2. *Let $F : \mathbb{R} \to \mathbb{R}$ be smooth and increasing with*

$$\liminf_{\gamma \to +\infty} F'(\gamma) = A > 0. \tag{24}$$

Let u be solution on $[0, T] \times (0, +\infty)$ of

$$\partial_t u + F(x^2 \partial_{xx} u) = 0, \tag{25}$$

with $u_T \in C^\infty$. Then for all $M > 0, \epsilon > 0$, there exists $\tau > 0$ depending only on $A, \|u_T\|_{L^1}$ such that $\forall t \leq T - \tau$, $\|F(x^2 \partial_{xx} u(t, \cdot))^+\|_{L^\infty[\epsilon, \epsilon^{-1}]} \leq M$.

We denote as before $F_n(\gamma) = \mu_n(\gamma)\gamma$. Given the growth condition assumed on μ_n, Theorem 3.2 implies that on any compact set of $[0, T) \times (0, +\infty)$, for any $\epsilon > 0$, there will exists n_ϵ such that for all $n \geq n_\epsilon, s^2 \partial_{ss} u_n \leq \Gamma_{max} + \epsilon$. Then u_n will solve on every compact set of $[0, T) \times (0, +\infty)$

$$\max\{\partial_t u_n + \frac{1}{2}\sigma^2 F(S^2 \partial_{ss} u_n), S^2 \partial_{ss} u_n - \Gamma_{max}\} = \eta_n, \tag{26}$$

for some sequence η_n converging uniformly to 0 in L^∞. This in turn implies Theorem 3.1.

□

Proof of Theorem 3.2. This is a classical result on non-linear semi-groups that can be retrieved from Veron [12] or Evans [7], we will therefore only sketch the proof. First by differentiation, we get that $w = \partial_{xx} u$ solves

$$\partial_t w + \partial_{xx}(F(x^2 w)) = 0. \tag{27}$$

Then, by applying a classical sequence of change of variables, we transform (27) into

$$\partial_t w + \partial_{xx}(F(w)) - x\partial_x(F(w)) = 0, \tag{28}$$

for $x \in \mathbb{R}$. For simplicity we will now drop the first order term that can be treated without difficulty, and then obtain

$$\partial_t(F(w)) + F'(w)\partial_{xx}(F(w)) = 0.$$

Consider now G smooth, convex on $\mathbb{R}$ and $k(t,x) = G(F(w(t,x)))$. We have

$$\partial_{xx}k = G'\partial_{xx}(F(w)) + G''(\partial_x(F(w)))^2,$$

and using the convexity of G we get

$$\partial_t k + F'(w)\partial_{xx}k \geq 0.$$

Now we choose G such that $G = 0$ on $\{F \leq M-1\}$ and $G \geq (F-M)^+$. We then have $F'(w) \geq F'(F^{-1}(M-1))$ on $\{k > 0\}$. Using then (24) we get that $F' \geq A/2$ on $\{k > 0\}$ if M is large enough, and then by standard comparison arguments, (following for example Veron [12] or Evans [7]) we get the existence of τ such that $G(t) \leq M$ for $t \leq T - \tau$. This in turn leads to $F \leq 2M$ for $t \leq T - \tau$.

□

Remark. We do not claim any bound on the negative part of $\partial_s su$.

3.1. **The induced hedging strategy.** The previous result shows that *any* market impact model enforces the Gamma max Theta max constraint when the market impact becomes large enough outside of the admissible range for Γ_c. For the upper constraint, this is clear that the diffusion becoming infinite has the effect of taking a semi-concave envelope of the payoff. On the gamma long side (we buy a call) assume that the function λ rises very quickly when Γ_c goes under the limit. Imagine that you are long a convex option, and that the spot has gone up by δS. The "linear" pde tells you to sell $\Gamma\delta S$ stocks. The "non-linear" approach developed above tells you to that any attempt to sell more than $\Gamma_{max}\delta S$ stocks would make the spot immediately go back under its original position. The pde tells you that you pay a theta equal to the worst case scenario: you sell $\Gamma_{max}\delta S$, and this makes the spot go back to the point where you are delta hedged. If the spot ends up anywhere else, your P& L is positive. Notice however that following this strategy, the number of shares that you hold in your delta hedge will not be equal to the delta of the option. Your strategy is to *follow as much as you can this theoretical delta*, and the P& L that you can not lock is considered as lost as you pay a lower theta. Hence, the replication is not exact: assume that you are long a call, with zero delta (just under the strike) and that the spot goes up quickly and keeps rising for several days until the expiry: then every day you would sell stocks as much as you can, but your theoretical δ would be far above this value, and at the expiry your net position will be very positive. If the spot went down rapidly before expiry, there would exist a time when your delta would be the "good" delta. If the maturity falls at this precise moment, your P& L will be 0, otherwise it will be positive. This strategy is therefore a super-replication strategy.

3.2. **Consistency with earlier results.** Our approach is consistent with the one proposed in [11], [9]. In this work, the authors introduce the operator

$$G(p,A) = \min\{-p - \frac{1}{2}\sigma^2 A, \Gamma_{max} - A, -\Gamma_{min} + A\}.$$

Intuitively, the solution of the super replication problem satisfying the gamma constraint $\Gamma_{min} \leq S^2\partial_{ss}u \leq \Gamma_{max}$ would satisfy

$$G(\partial_t u, S^2\partial_{ss}u) = 0.$$

However, it appears clearly that F is not monotone in A, which means that the problem is ill-posed. The authors therefore introduce the modified operator (adapted to our notations)

$$\hat{G}(p,A) = \sup\{G(p, A+\beta), \beta \geq 0\}.$$

This operator is by construction non increasing in the variable A. The above equation can be rewritten as

$$\partial_t u = \min\{-\frac{1}{2}\sigma^2 S^2 \Gamma_{min}, -\frac{1}{2}\sigma^2 S^2 \partial_{ss} u\}$$

if $S^2 \partial_{ss} u \leq \Gamma$,

$$u := \inf\{v > u, \partial_{ss} v \leq \Gamma\}$$

otherwise. Here again the ill-posed constraint $S^2 \partial_{ss} u \geq \Gamma_{min}$ is changed into the constraint $\partial_t u \leq -\frac{1}{2}\sigma^2 \Gamma_{min}$, which shows the consistency between our approach and the one of [11], [9].

We mention also the works [4] and [5] by Cetin, Jarrow, Touzi and Protter which set the problem in a more general formalism: indeed they introduce the notion of *supply curve*, that is a price process $S(t, x, \omega)$, which describes the price when trading a quantity x of the stock. This approach is somehow equivalent to our intensity based model. However note that they consider a model with liquidity costs but without market impact, therefore the volatility of the underlying is not affected.

4. Numerical simulations

We present our simulation in the linear impact model. We propose the following numerical scheme: We set ϵ to a small constant ($\epsilon = 10^{-3}$ in our applications), and divide the time interval $[0, T]$ into $[0, t_1, \ldots, t_N = T]$.

- Define

$$\begin{aligned} F(\Gamma_c, \bar{\Gamma}_c) &= \frac{\Gamma_c}{1 - \lambda \bar{\Gamma}_c} \text{ if } \bar{\Gamma}_c \leq \frac{1}{\lambda}, \\ &= \epsilon^{-1} \text{ otherwise.} \end{aligned}$$

- Initialize $i = N$.
- **Terminal condition** Initialize $u(t_N) = \Phi(T)$
- **time loop** For $i = N$ down to $i = 1$
- initialize $v_1 = u(t_i)$.
- **Non linear iterations** for j in $[1..N_{itnl}]$ solve on $[t_{i-1}, t_i]$

$$\begin{aligned} \partial_t w &= -\sigma^2 F(s^2 \partial_{ss} w, s^2 \partial_{ss} v_j), \\ w(t_i) &= u(t_i). \end{aligned}$$

- Set $v_{j+1} = w(t_{i-1})$ and loop on j until $j = N_{itnl}$.
- Set $u(t_{i-1} = w_(t_{i-1})$ and iterate on the time step i.

Typically, the number of non linear iterations N_{itnl} needed for convergence was small: $N_{itnl} = 3$ was enough in our numerical example.

For stability reasons, we used an implicit scheme. This avoids problem that could have arisen with the brutal method that we use to enforce the upper bound on the gamma (i.e. infinite diffusion). Alternatives to this method do exist (it is a semi-concave envelope problem), but we found that our method worked quite well. As noticed in [11], with a constant volatility model, it is enough to enforce the upped bound on Γ on the terminal condition, but this fails to be true with a generic local volatility.

We present some numerical simulations of the linear model in Fig. 2 and Fig. 3. Note that in Figure 2 (resp. 3) the option sold is convex (resp. concave), while in Figure 4 the option is a call spread, therefore the second derivative of the payoff changes sign. One sees that the impact-free solution is below the solution with market impact both in the concave and in the convex area.

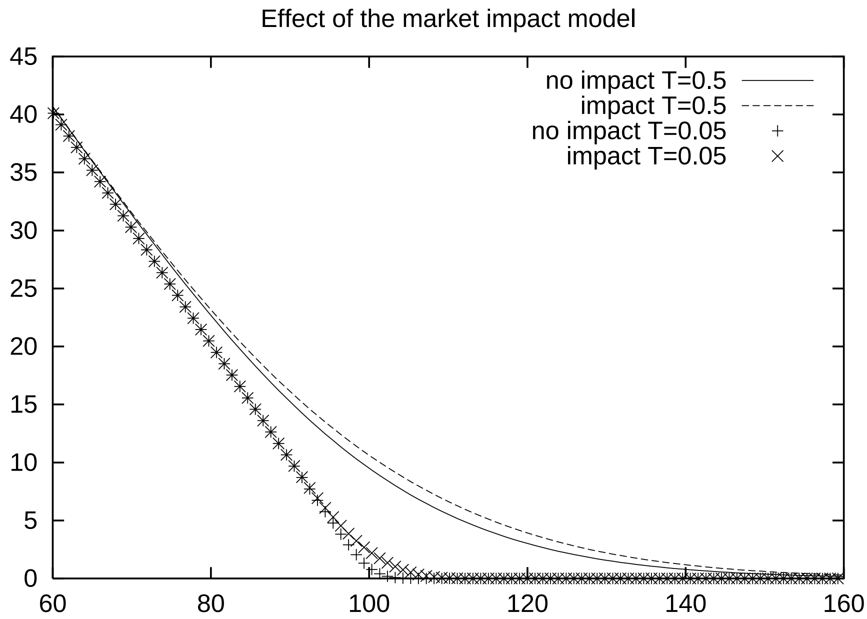

FIGURE 2. Example on a put option of the market impact model (gamma short case)

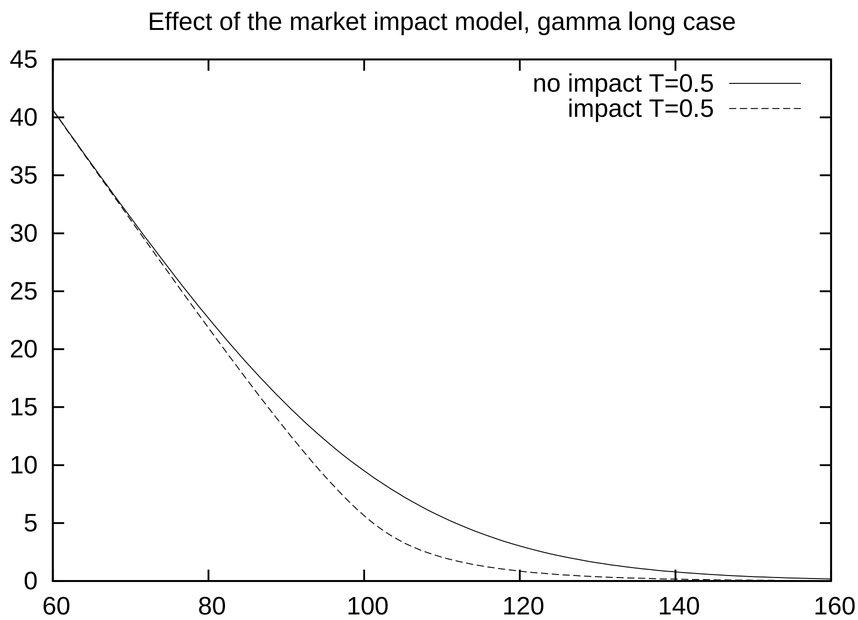

FIGURE 3. Example on a put option of the market impact model (gamma long case)

Numerical implementation of the Gamma-max model. For the Gamma max model, we just replace the definition of F by

$$
\begin{aligned}
F(\Gamma_c, \bar{\Gamma}_c) &= \Gamma_c \text{ if } |\bar{\Gamma}_c| \leq \frac{1}{\lambda}, \\
&= 12\epsilon^{-1} \text{ if } \bar{\Gamma}_c \geq \frac{1}{\lambda}, \\
&= \frac{-1}{\lambda} \text{ otherwise.}
\end{aligned}
$$

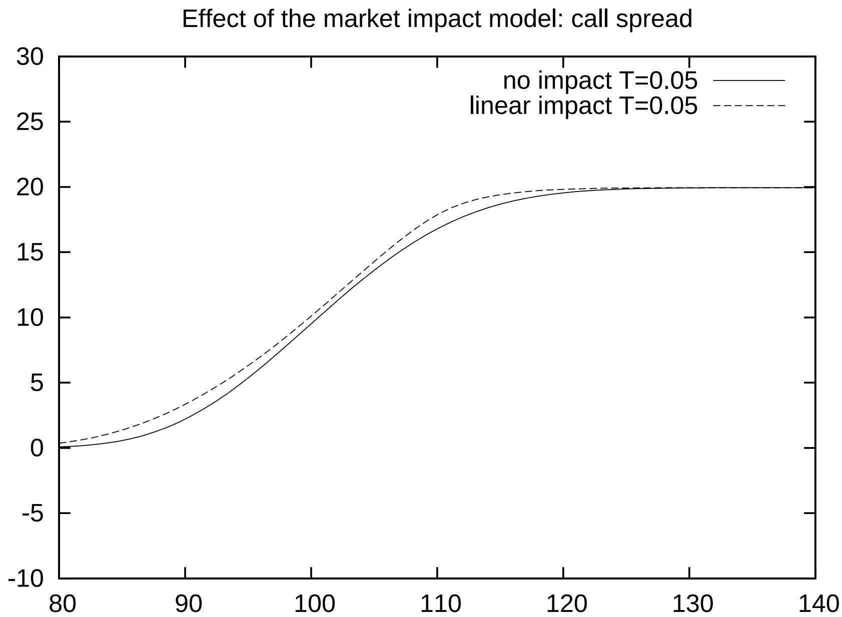

FIGURE 4. The call spread case

REFERENCES

[1] R.A. Almgren, , and Tianhui Michael Li. A fully-dynamic closed-form solution for delta-hedging with market impact. 2011.

[2] R.A. Almgren. Equity market impact. *RISK*, July, 2005.

[3] Jean-Philippe Bouchaud, Doyne J. Farmer, and Fabrizio Lillo. How markets slowly digest changes in supply and demand. Sep 2008.

[4] Umut Çetin, Robert A. Jarrow, and Philip Protter. Liquidity risk and arbitrage pricing theory. *Finance Stoch.*, 8(3):311–341, 2004.

[5] Umut Çetin, H. Mete Soner, and Nizar Touzi. Option hedging for small investors under liquidity costs. *Finance Stoch.*, 14(3):317–341, 2010.

[6] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[7] Lawrence C. Evans. Differentiability of a nonlinear semigroup in L^1. *J. Math. Anal. Appl.*, 60(3):703–715, 1977.

[8] Robert V. Kohn and Sylvia Serfaty. A deterministic-control-based approach to fully nonlinear parabolic and elliptic equations. *Comm. Pure Appl. Math.*, 63(10):1298–1350, 2010.

[9] H. M. Soner P. Cheridito and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Annales de lInstitut Henri Poincar, Srie C: Analyse Non-Linaire*, 22:633–666, 2005.

[10] H. M. Soner and N. Touzi. Superreplication under gamma constraints. *SIAM J. Control Optim.*, 39:73–96, 2000.

[11] H. M. Soner and N. Touzi. Hedging under gamma constraints by optimal stopping and face-lifting. *Mathematical finance*, 17:59–80, 2007.

[12] Laurent Véron. Effets régularisants de semi-groupes non linéaires dans des espaces de Banach. *Ann. Fac. Sci. Toulouse Math. (5)*, 1(2):171–200, 1979.