A Probabilistic Numerical Method for Fully Non-linear Parabolic Partial Differential Equations

Arash Fahim

ÉCOLE POLYTECHNIQUE, PARIS, FRANCE & SHARIF UNIVERSITY OF TECHNOLOGY, TEHRAN, IRAN

ÉCOLE DOCTORAL
CENTRE DE MATHÉMATIQUE APPLIQUÉE

P H D T H E S I S

to obtain the title of

PhD of Science

Specialty : Applied Mathematics

Defended by: Arash FAHIM

A PROBABILISTIC NUMERICAL METHOD FOR FULLY NON–LINEAR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

Thesis Adviser: Nizar TOUZI

Jury :

Reviewers :	Denis TALAY	-	Institut National de Recherche en Informatique et en Automatique (Sophia Antipolis, Valbonne, FRANCE)
	Gilles PAGES	-	Université de Paris VI (Paris)
	Jianfeng ZHANG	-	USC (Los Angeles, CA, USA)
Advisor :	Nizar TOUZI	-	École Polytechnique (Paris)
President :	Nicole EL KAROUI	-	Université de Paris VI (Paris)
Examinators :	Nicole EL KAROUI	-	Université de Paris VI (Paris)
	Tom HURD	-	Mc Master University (Hamilton, ON ,CANADA)
	Tom SULISBURY	-	York University (Toronto, ON, CANADA)
	Bijan Z. ZANGENEH	-	Sharif University of Technology (Tehran, IRAN)

Acknowledgments

I am grateful to thank Nizar Touzi for his Comprehensive support during my PhD, the staff of CMAP for their administrative supports during my PhD, all the members of board of examiners for the time they spend on the manuscript of my thesis and on the defense section and also for their useful suggestion about this works, Tom Hurd for his fruitful suggestions on the structure of the these,*Fields Institute for Research in Mathematics* for the hosting my defense,
and
my wife, Nasrin, to be extremely patient and helpful
and I would like to devote this Thesis to her.

A Probabilistic Numerical Method for Fully Non–linear Parabolic Partial Differential Equations

Abstract en Français: Cet these a deux patrtie. La partie premiere introduit une methode probabiliste numerique pour les EDPs parabolique et completement non-linaire et puis on consider sa properietes asymptotiques (convergence et taux de convergence) et aussi l'analyse de l'erreur due à l'approximation de l'espérance conditionnelle par une méthode de type Monte Carlo. Les EDPs complètement non-linaires apparaissent dans plusieurs applications en ingénierie, economie et finance. Citons par example le probleme de propagation de front par courbure moyenne, ou le problème de selection de portefeuille. Une classe importante d'EDP completement non-linéaire est constituée par les équations de HJB découlant du contrôle optimal stochastique. Dans la plupart des cas, il n'existe pas de solution dans le sense classique. Par conséquent, la notion de solution de viscosité est utilisé pour les EDP compeletement non-linéaires. En raison de manque de de solution explicite dans de nombreuses applications, les schémas d'approximation sont devenus très importants. Pour montrer la convergence, la méthode utilisée dans cette thèse a été introduite par Barles et Souganidis. Leurs travaux fournissent le résultat de convergence vers des solutions de viscosité pour une solution approchée obtenue à partir cohérente, monotone et stable régime. Afin d'obtenir le taux de convergence, nous avons supposé que le EDP a non-linéarité concave de type HJB. En d'autres termes, la non-linéarité est une borne inférieure des opérateurs linéaires. La thèse a utilisé la méthode de Krylov des coefficients secoué et d'approximation par un système d'équations HJB couplées pour obtenir des bornes sur les taux de convergence. La mise en œuvre du schémas requiert d'introduire une approximation des espérances conditionnelles. Pour une classe d'estimateurs, nous avons obtenu une borne inférieure sur le nombre de chemins échantillon qui préserve la vitesse de convergence obtenue avant. La généralisation de la méthode à des équations intégro-différentielles est simple et on peut utiliser les mêmes agruments que dans le cas local pour obtenir la convergence et le taux de convergence. Notons cependant que le cas non local introduit la difficulté supplémentaire d'approximation des termes non locaux. La première partie sera terminée est illustrée par quelques expériences numériques. La méthode est utilisée pour résoudre le problème géometrique des flux de courbure moyenne, le problème de la sélection sur un portefeuille d'actifs avec volatilité stochastique dans le modèle de Heston, et le problème de sélection de portefeuille de deux actifs à la fois avec une volatilité stochastique, on satisfait modèle de Heston et l'autre CEV modèle.

La deuxième partie de la thèse traite de la politique de production optimale dans le marché des allocations des permis d'émission de carbone. Le marché des permis d'émissions de carbone est une approche de marché pour mettre en œuvre le protocole de Kyoto. Nous avons calculé la production optimale dans 4 cas: quand il n'y a pas un tel marché, quand il y a un tel marché, mais sans grand producteur de carbone, quand il y a un gros producteur qui n'est pas teneur de marché, et quand il existe un marché avec un grande producteur. Nous avons montré que

dans les premiers, la production optimale est toujours diminuée. Cependant, dans le dernier cas, nous avons montré que le gros producteur peut bénéficier du marché en changeant la prime de risque de l'allocation de carbone en raison de sa production d'appoint. Cette partie est illustrée par quelques expériences numériques qui montre des cas que le grand producteur peut bénéficier d'une production d'appoint.

Abstract in English: This thesis is divided into two parts. First part introduces a probabilistic numerical method for fully non–linear parabolic PDEsand consider its asymptotic properties (convergence and rate of convergence) and the error analysis due to approximation of conditional expectation. Fully non–linear PDEs appear in many applications in engineering, economics and finance (see e.g. problem of portfolio selection and mean curvature flow). An important class of Fully non–linear PDEs is the HJB equations arising in stochastic optimal control. In most cases, there exists no solution in classical sense. Therefore, the notion of viscosity solution is used for the Fully non-linear PDEs. Due to the lack of closed form solution in many applications, the approximation schemes have become appealing. Then one needs to guarantee the convergence of the approximate solution to the viscosity solution of Fully non-linear PDEs. The method hired in this thesis to obtain the convergence result, is introduced by Barles and Souganidis in citebarlessouganidis. Their work provides the convergence result to viscosity solutions for any approximate solution obtained from consistent, monotone and stable scheme. In order to achieve rate of convergence, we supposed that the PDE has a concave non–linearity of HJB type. In other words, the non–linearity is an infimum of linear operators. The thesis used the Krylov method of shaking coefficients and switching system approximation of HJB equations to obtain convergence rates from above and below. The implementation of the scheme needs the conditional expectations inside the method to be replaced by an appropriate estimator. For a class of estimators, we obtained a lower bound on the number of sample paths which preserves the rate of convergence obtained before. The generalization of the method to non–local PDEs is straight forward and one can use the same agruments as the local case to achieve the convergence and the rate of convergence. There is one exception in non-local case which differs from local case i.e. the Monte Carlo approximation of intergal (non–local) term. This is done by using suitable jump–diffusion process. The first part will be ended by some numerical experiments. The method is used to solve the geometric problem of mean curvature flow, the problem of portfolio selection on one asset with stochastic volatility in Heston model, and the problem of portfolio selection on two assets both with stochastic volatility, one satisfies Heston model and the other CEV model.
The second part of the thesis deals with the optimal production policy under the carbon emission allowance market. The carbon emission allowance market is a market approach to implement Kyoto protocol. We calculated the optimal production in 3 cases: when there is such a market but without any large carbon producer, when there is a large producer who is not market maker, and when there is a large producer market maker. We showed that in second cases, the optimal production is always less than the first case and in the third case it is even less than the second case. On the other hand, we showed that the market maker (if there exist any) can benefit from the market by changing the risk premium of the carbon allowance due to her extra production. The model we used here for the price of carbon allowance is a BSDE. Then we introduce a stochastic optimization problem. The carbon producer wants to maximaze her utility from her wealth. Her wealth consists

of two parts: a self–financing portfolio over the carbon emission allowance papers and the benefit from her production. As expected, the optimal production does not depend on the utility. One could pass to a new optimization problem which gives the optimal production. We choose to solve the stochastic optimization problem by the means of HJB equations. We obtained the verification and uniqueness result for the HJB equation. This part is closed by some numerical experiments which shows cases which the large producer can benefit from extra production.

Contents

Notations

For scalars $a, b \in \mathbb{R}$, $a \wedge b := \min\{a, b\}$, $a \vee b := \max\{a, b\}$, and $a^+ := \max\{a, 0\}$.

$\mathbb{R}^d_* := \mathbb{R}^d \setminus \{0\}$.

$\mathcal{C}_d$ is the collection of all bounded real functions on $\mathbb{R}^d$.

$\mathbb{M}(n, d)$ is the collection of all $n \times d$ matrices with real entries.

The collection of all symmetric matrices of size d is denoted by $\mathcal{S}_d$, and its subset of nonnegative symmetric matrices is denoted by $\mathcal{S}_d^+$. By "$\leq$" we denote the partial order induced by the positive cone $\mathcal{S}_d^+$.

For a matrix $A \in \mathbb{M}(n, d)$, A^{T} is the transpose of A. For $A, B \in \mathbb{M}(n, d)$, $A \cdot B := \mathrm{Tr}[A^{\mathrm{T}} B]$. In particular, for $d = 1$, A and B are vectors of $\mathbb{R}^n$ and $A \cdot B$ reduces to the Euclidean scalar product.

A^- is the pseudo–inverse of the matrix A.

For a suitably smooth function φ on $Q_T := (0, T] \times \mathbb{R}^d$, define

$$|\varphi|_\infty := \sup_{(t,x) \in Q_T} |\varphi(t, x)| \quad \text{and} \quad |\varphi|_1 := |\varphi|_\infty + \sup_{Q_T \times Q_T} \frac{|\varphi(t, x) - \varphi(t', x')|}{|x - x'| + |t - t'|^{\frac{1}{2}}}.$$

Finally, the $\mathbb{L}^p-$norm of a r.v. R is denoted by $\|R\|_p := (\mathbb{E}[|R|^p])^{1/p}$.

Chapter 1
Introduction

In the areas of engineering and mathematics; including finance, the Monte Carlo methods are always referred to as the computational methods based on the random sampling. In the approximation of the solutions of PDEs, the Monte Carlo methods play an important role especially when the dimension of problem is large. The finite difference and finite element methods usually are not implementable in large dimensions. However, the Monte Carlo methods are generally less sensitive with respect to dimension and could provide implementable schemes.

The Monte Carlo methods for PDEs starts by the famous Feynman–Kac formula for linear PDEs. The extension of Feynman–Kac to the non–linear PDEs can not easily be done by a simple conditional expectation. However, it could be extended for the semi–linear parabolic equations through Backward Stochastic Differential equations (BSDEs). For more details see [51], [52] and [53]. Semi–linear parabolic equations have the general form

$$\begin{aligned} -\mathcal{L}^X v(t,x) - F(t,x,v(t,x),\sigma^{\mathrm{T}} Dv(t,x)) &= 0 \ \text{ on } \ [0,T)\times\mathbb{R}^d \\ v(T,\cdot) &= g(\cdot) \ \text{ on } \ \mathbb{R}^d, \end{aligned}$$

where $\mathcal{L}^X\varphi := \frac{\partial\varphi}{\partial t} + \mu\cdot D\varphi + \frac{1}{2}a\cdot D^2\varphi$ is the infinitesimal generator of a diffusion process X and $a := \sigma\sigma^{\mathrm{T}}$. The Monte Carlo approximation of the solution of the semi–linear equation is given by a decoupled system which consists of the stochastic differential equation (SDE) and a backward stochastic differential equation (BSDE):

$$\begin{aligned} dY_t &= F(t,X_t,Y_t,Z_t)dt + Z_t dW_t \\ Y_T &= g(X_T). \end{aligned}$$

More precisely, assuming sufficient regularity for the solution of PDE, one has the correspondence $v(t,X_t) = Y_t$ and $Dv(t,X_t) = Z_t$. The numerical methods for the BSDEs are initially developed by the use of the classical solutions of semi–linear parabolic PDEs in [49]. In that work, the authors imposed a restrictive regularity condition over coefficients which implies the existence of classical solutions for the semi–linear PDEs . Moreover, this method depends on the approximation of the solution of PDEs which appears to be difficult in high dimensions.

The theory of BSDEs provides an extension of Feynmen–Kac to the semi–linear case. The purely Monte Carlo method for BSDEs relies on the discretization of the forward diffusion process X and then to find a solution for discretized BSDE backward in time. The advantage of this approach is that it could also be used to approximate the solution of semi–linear parabolic PDEs. See for instance Chevance [26], El Karoui, Peng and Quenez [32], Bally and Pagès [2], Bouchard and Touzi [18] and Zhang [59]. In particular, the latter papers provide the convergence of the "natural" discrete-time approximation of the value function and its partial space gradient with the same L^2 error of order $\sqrt{h}$, where h is the length of time step. The discretization involves the computation of conditional expectations, which need to be further approximated in order to result into an implementable scheme. We refer to [2], [18] and [35] for an complete asymptotic analysis of the approximation, including the regression error.

Therefore, instead of using PDE to approximate the solution of BSDE, we use BSDE to approximate the solution of PDE. More precisely, for a time discretization $\{t_i\}_{i=0}^N$ of $[0,T]$, the approximation for Y and Z could be done by:

$$\begin{aligned}
\hat{Y}^N_{t_N} &= g(\hat{X}^N_T) \\
\hat{Z}^N_{t_i} &= \frac{1}{\Delta t_{i+1}}\mathbb{E}_i[\hat{Y}^N_{t_{i+1}}\Delta W_{i+1}] \\
\hat{Y}^N_{t_i} &= \mathbb{E}_i[\hat{Y}^N_{t_{i+1}}] - \Delta t_{i+1}F(t_i,\hat{X}^N_{t_i},\hat{Y}^N_{t_i},\hat{Z}^N_{t_i}),
\end{aligned}$$

where $\mathbb{E}_i = \mathbb{E}[\cdot|F_{t_i}]$, $\Delta t_{i+1} = t_{i+1} - t_i$ and $\Delta W_{i+1} = W_{t_{i+1}} - W_{t_i}$. For more details on error analysis of discretization of BSDEs, we refer to [18], [35], [59], [60], and [17]. The optimal error of this discretization is the same as for forward SDEs e.g. $|\pi|^{1/2}$ where $|\pi| := \sup\{\Delta t_i | i = 1,\cdots,N\}$.

For fully non–linear parabolic equations, the starting point is [25] where they proposed a system called second order BSDE (2BSDE) corresponding to the following final value problem.

$$-\mathcal{L}^X v(t,x) - F\big(t,x,v(t,x),\sigma^{\mathrm{T}}Dv(t,x),D^2v(t,x)\big) = 0, \quad \text{on } [0,T)\times\mathbb{R}^d, \tag{1.0.1}$$

$$v(T,\cdot) = g, \quad \text{on } \mathbb{R}^d, \tag{1.0.2}$$

where

$$\mathcal{L}^X\varphi \;:=\; \frac{\partial\varphi}{\partial t} + \mu\cdot D\varphi + \frac{1}{2}a\cdot D^2\varphi.$$

and μ and σ are two maps from $\mathbb{R}_+\times\mathbb{R}^d$ to $\mathbb{M}(d,d)$ and $\mathbb{R}^d$, $a := \sigma\sigma^{\mathrm{T}}$ is a map from $\mathbb{R}_+\times\mathbb{R}^d$ to $\mathcal{S}_d^+$, and

$$F : (t,x,r,p,\gamma)\in\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times\mathcal{S}_d \;\longmapsto\; F(x,r,p,\gamma)\in\mathbb{R}.$$

A 2BSDE is a system of SDEs given by

$$\begin{aligned}
dY_t &= F(t,X_t,Y_t,Z_t,\Gamma_t)dt - Z_t\circ dW_t \\
dZ_t &= A_t dt + \Gamma_t dW_t \\
Y_T &= g(X_T),
\end{aligned}$$

where $\circ$ stands for Stratonovich integral. The solution of the 2BSDE is an adapted quadruple (Y_t,Z_t,A_t,Γ_t) which satisfies the above equations. Under the regularity of the solution of the the final value problem (1.0.1)–(1.0.2), the correspondence between the fully non–linear PDE and the system of 2BSDE is given by

$$\begin{aligned} Y_t &= v(t, X_t) \\ Z_t &= \sigma^{\mathrm{T}} Dv(t, X_t) \\ \Gamma_t &= D^2 v(t, X_t) \\ A_t &= \mathcal{L}^X Dv(t, X_t). \end{aligned}$$

By discretizing the 2BSDE, one can propose the following scheme:

$$\begin{aligned} \hat{\Gamma}^N_{t_i} &= \frac{1}{\Delta t_i} \mathbb{E}_i[\hat{Z}^N_{t_{i+1}} \Delta W_{i+1}] \\ \hat{Z}^N_{t_i} &= \frac{1}{\Delta t_{i+1}} \mathbb{E}_i[\hat{Y}^N_{t_{i+1}} \Delta W_{i+1}] \\ \hat{Y}^N_{t_i} &= \mathbb{E}_i[\hat{Y}^N_{t_{i+1}}] - \Delta t_{i+1} F(t_i, \hat{X}^N_{t_i}, \hat{Y}^N_{t_i}, \hat{Z}^N_{t_i}, \hat{\Gamma}^N_{t_i}). \end{aligned} \tag{1.0.3}$$

The main subject of this thesis is to introduce a probabilistic numerical method for the fully non–linear parabolic PDE (1.0.1)–(1.0.2) based on the (1.0.3). Fully non–linear PDEs arise in many problems in applied mathematics and engineering including finance. For example the problem of motion by curvature, portfolio optimization under different type of constraints, option pricing under illiquidity cost, etc. Non–local fully non–linear PDEs arise from stochastic optimization problems for controlled jump–diffusion processes e.g. problem of portfolio optimization in Lévy markets. There are only few examples with explicit and quasi-explicit solution; for example see [8] or [9]. We consider local PDEs and non–local PDEs case separately in two chapters.

Now, we briefly discuss the contents of each chapter together with a review on the relevant literatures.

Chapter 2

In this Chapter, we observe that the backward probabilistic scheme of [25] can be introduced naturally without appealing to the notion of backward stochastic differential equation. This is shown is Section 2.1 where the scheme is decomposed into three steps:
(i) The Monte Carlo step consists in isolating the linear generator of some underlying diffusion process, so as to split the PDE into this linear part and a remaining non–linear one.
(ii) Evaluating the PDE along the underlying diffusion process, we obtain a natural discrete-time approximation by using a kind of finite differences approximation of derivatives in the remaining non–linear part of the equation.
(iii) Finally, the backward discrete–time approximation obtained by the above steps (i)-(ii) involves the conditional expectation operator which is not computable in explicit form. An implementable probabilistic numerical scheme therefore requires

to replace such conditional expectations by a convenient approximation, and induces a further Monte Carlo type of error.

In the present chapter, we do not require the fully non–linear PDE to have a smooth solution, and we only assume that it satisfies a comparison result in the sense of viscosity solutions. Our main objective is to establish the convergence of this approximation towards the unique viscosity solution of the fully-non–linear PDE, and to provide an asymptotic analysis of the approximation error.

Our main results are the following. We first prove the convergence of the discrete-time approximation for general non–linear PDEs, and we provide bounds on the corresponding approximation error for a class of Hamilton-Jacobi-Bellman PDEs. Then, we consider the implementable scheme involving the Monte Carlo error, and we similarly prove a convergence result for general non–linear PDEs, and we provide bounds on the error of approximation for Hamilton-Jacobi-Bellman PDEs. We observe that our convergence results place some restrictions on the choice of the diffusion of the underlying diffusion process. First, an ellipticity condition is needed; we believe that this technical condition can be relaxed in some future work. More importantly, the diffusion coefficient is needed to dominate the partial gradient of the remaining non–linearity with respect to its Hessian component. Although we have no theoretical result that this condition is necessary, our numerical experiments show that the violation of this condition leads to a serious mis-performance of the method, see Figure 2.5.

Our proofs rely on the monotonic scheme method developed by Barles and Souganidis [6] in the theory of viscosity solutions, and the recent method of shaking coefficients of Krylov [43], [44] and [45] and Barles and Jakobsen [5], [4] and [3]. The use of the latter type of methods in the context of a stochastic scheme seems to be new. Notice however, that our results are of a different nature than the classical error analysis results in the theory of backward stochastic differential equations, as we only study the convergence of the approximation of the value function, and no information is available for its gradient or Hessian with respect to the space variable.

The followings are two related numerical methods based on finite differences in the context of Hamilton-Jacobi-Bellman non–linear PDEs:

- Bonnans and Zidani [14] introduced a finite difference scheme which satisfies the crucial monotonicity condition of Barles and Souganidis [6] so as to ensure its convergence. Their main idea is to discretize both time and space, approximate the underlying controlled forward diffusion for each fixed control by a controlled local Markov chain on the grid, approximate the derivatives in certain directions which are found by solving some further optimization problem, and optimize over the control. Beyond the curse of dimensionality problem which is encountered by finite differences schemes, we believe that our method is much simpler as the monotonicity is satisfied without any need to treat separately the linear structures for each fixed control, and without any further investigation of some direction of discretization for the finite differences.

- An alternative finite-differences scheme is the semi-Lagrangian method which solves the monotonicity requirement by absorbing the dynamics of the underlying state in the finite difference approximation, see e.g. Debrabant and Jakobsen [29]. Loosely speaking, this methods is close in spirit to ours, and corresponds to freezing the Brownian motion W_h, over each time step h, to its average order $\sqrt{h}$. However it does not involve any simulation technique, and requires the interpolation of the value function at each time step. Thus it is also subject to the curse of dimensionality problems.

We finally observe a connection with the recent work of Kohn and Serfaty [42] who provide a deterministic game theoretic interpretation for fully non–linear parabolic problems. The game is time limited and consists of two players. At each time step, one tries to maximize her gain and the other to minimize it by imposing a penalty term to her gain. The non–linearity of the fully non–linear PDE appears in the penalty. Also, although the non–linear penalty does not need to be elliptic, a parabolic non–linearity appears in the limiting PDE. This approach is very similar to the representation of [25] where such a parabolic envelope appears in the PDE, and where the Brownian motion plays the role of Nature playing against the player.

Chapter 3

The present Chapter generalizes the probabilistic numerical method in [34] for approximation of the solution of fully non–linear parabolic PDEs to non–local PDEs. Here by non–local PDEs, we mean the integro–partial differential equations which sometimes are referred to as integro–partial differential equations (IPDE). As mentioned in the previous Chapter, the method is originated from [25] where a similar probabilistic numerical method is suggested based on 2BSDEs[1].

As in Chapter 2, the main idea is to separate the equation into a purely linear part and a fully non–linear part. Then, we use the time discretization of a suitable jump–diffusion process to approximate the derivatives and integral term in the non–linear part. The separation into linear and non–linear part is arbitrary up to the satisfaction of some assumptions. The assumptions needed for this result are degenerate ellipticity condition for the remaining non–linearity and that the diffusion coefficient is needed to dominate the partial gradient of the remaining non–linearity with respect to its Hessian component.

The other contribution of this Chapter is the Monte Carlo method for approximation of the integral with respect to Lévy measure which appears in the non–local PDEs. The method is referred to in this Chapter as *Monte Carlo Quadrature (MCQ)*. We treat the jumps as in [16] for finite activity jump–diffusion processes. For infinite activity jump–diffusion processes, we truncate the Lévy measure near zero and then treat them as in the finite measure case. We introduce bounds for the truncation error with respect to the derivatives of integrand and truncation level.

Although MCQ is independent of the numerical scheme, we choose to approximate the Lévy integral inside the non–linearity by MCQ. In this case, we also need

[1] Second order backward stochastic differential equations

to choose appropriate truncation bound with respect to time step which retains the convergence and rate of convergence as in the local case in Chapter 2.

The idea of the proof is captured from [6] for the convergence result and from [12] for the rate of convergence. However, in the non–local PDEs, we need to conquer the new difficulties due to lack of Lipschitz continuity of non–linearities appearing in many interesting PDEs e.g. HJB equations. More precisely, if the non–local non–linearity is of HJB type, then it is Lipschitz if and only if Lévy measure inside the non–local integral is finite. This difficulty makes it impossible to use directly the methods in [6] and [12]. We showed that if the truncation threshold; κ; is properly dependent on time step; h; then one can produce the approximate solution which converges to the solution of the non–local problem.

The first result concerns the convergence of the approximate solution obtained from the scheme; (3.3.3); to the viscosity solution of the final value problem. The difficulty which makes the direct use of the method in Chapter 2 impossible, is that when we have a Lévy integral with respect to infinite Lévy measure in the non–linearity, the non–linearity is no more Lipschitz. If we truncate the Lévy measure, the non–linearity is Lipschitz but as truncation threshold tends to 0, the Lipschitz constant blows up. We solved this problem through manipulatiing the original final value problem to an other whose corresponding scheme is monotone. Turning the manipulation back, we obtain a bounded approximate solution. This approximation is near the approximate function created by the scheme (3.3.3), if the truncation threshold depends appropriately on h.

The second result provides rate of convergence in the case of concave non–linearity. The proof of the rate of convergence uses the results in [12] and [13] which generalizes the result of [6] to non–local case. The method is based on the approximation of the solution of the equation with regular sub and super–solutions. Plugging the regular sub or super–solution into scheme and then usage of the consistency provides the upper and lower bounds. Here, we also need to impose the condition that the truncation threshold depend appropriately on the time step in order to preserve the rate of convergence after truncation. For the rate of convergence, we also need to manipulate the equation to obtain a strictly monotonicity for the scheme which is a crucial requirement in using the method in [12].

Finally, as mentioned in Chapter 2 for non–local case, it is worthy of noticing the relation with the generalization of [42] to non–local case introduced in [38] which provides a deterministic game theoretic interpretation for fully non–linear parabolic problems. The game consists of two players. At each time step in a predetermined time horizon, one tries to maximize her gain and the other to minimize it by imposing a penalty term to her gain. More precisely, she starts in an initial position and chooses a vector p, a matrix Γ, and a function φ. Then, he will plug an arbitrary vector w together with p, Γ and φ in a non–linear penalty term which should be paid by her and change her position by taking one step with appropriate length in the direction of vector w. At the final stage, she will earn as much as a function of her final position. As time step goes to zero, her value function at any time and any position will converge to the solution of a fully non–linear parabolic PDE whose non–linearity consists of the (elliptic envelop of the) penalty term. Vector p, a matrix Γ and a function φ represent the first and second derivatives and the solution function, respectively.

Chapter 4

The long term costs of global warming is believed to be significantly more than the cost of controlling it by reducing the pollution due to greenhouse gases (see [50]). One direct way to reduce the emission is to impose the taxation on the installations whose production increases the pollution. One can propose the standard taxation system which imposes a limitation level on the production of each installation over a time period and any amount of production above this level will be penalized. This taxation method has some significant disadvantages. First, there is no change in the production of the installations whose current optimal production policy does not reach the level. Second, there is no benefit for those who are below their level to keep their position. This effect also creates incentive to merge with other installation who needs to produce above their level.

The Kyoto protocol in 1997 concerns with the reduction of the greenhouse gases including CO_2 and is accepted by several countries e.g. European Union members. In 2000, the European Commission launch European Climate Change Program (ECCP) to implement Kyoto protocol in Europe. As an alternative to standard taxation, ECCP proposed European Union Emission Trading Scheme (EU ETS) which provides a way to control the emission of CO_2 within carbon polluters through trading the papers which allows them extra emission. More precisely, ETS imposes a cap over the total carbon emission. Within ETS, certain industrial installations with intensive carbon pollution are given free allowances. If any installation wants to produce more than her initial allowance, she should buy allowance through EU ETS. However, the allowances will be needed if the total carbon emission per member state violates imposed cap. On the other hand, if such installations, are far away from their production limit, they could sell their allowance through the market.

First phase of the program was run from January 2005 to the end of 2007. All the included installations who violate their limits, were supposed to provide enough allowances, if the cap on total emission is reached. The cap for the second phase (2008–2012) has been revised after the collapse in the first phase in April 2006 due to the release of the information about the unreachability to total carbon emission cap. Moreover, in the second phase, ECCP proposed to relevant installations to put off execution of the first phase emission allowance to the second phase by paying 40 euros per tone. The same mechanism is determined between the second phase and the third phase by the cost of 100 euros per tone. This mechanism, which is referred to as *banking*, proposes an option for the allowance holder to execute the allowance to offset the excess production or to keep it for the next phase. For more details see [19], [20], [21], [22], [23] and [50].

Nowadays, there are other regional markets implementing similar schemes as EU

ETS, e.g. the US REgenial CLean Air Incentive Market (RECLAIM) or Regional Greenhouse Gas Incentive (RGGI). Throughout this chapter, by emission market we mean the emission trading scheme EU ETS.

In this chapter, we analyze the effect of emission market in reducing the carbon emission through the change on production policy of the relevant firms. The firm's objective is to maximize her utility on her wealth which is made of both the profit of her production and the value of her carbon allowance portfolio over her production and her portfolio strategy. We solve the utility maximization problem on portfolio strategy by the duality argument and then on the production by the use of Hamilton–Jacobi–Bellman (HJB) equations.

We observe that the market always reduces the optimal production policy of the small producers and large producers who can not affect the risk premia. However, under certain cases, the large producer can have a larger optimal production in the market. The comparison is based on the fact that negative of the derivative of the value function with respect to the total emission imposed by the firm is equal to the price of the carbon allowance, under some assumptions.

More precisely, we define the rate of profit of the firm for the production rate q by $\pi(q)$ where π a strictly concave function π on its production with $\pi(0) = 0$, $\pi(\infty) = -\infty$ and $\pi'(0+) > 0$, and the the rate of emission of the firm caused by the production rate q by $e(q)$ where e is an increasing concave function.

In the as-usual-business case , optimal production $\bar{q}^{(0)}$ is such that $\pi'(\bar{q}^{(0)}) = 0$. When the standard taxation is applied the optimal production, $q^{(0)}$ should satisfies

$$\pi'(q^{(0)}) - \mathbb{E}_t^{\bar{\mathbb{Q}}^0}[\alpha \mathbb{1}_{\{E_T^{q^{(0)}} \geq E_{\max}\}}]e'(\bar{q}^{(0)}) = 0,$$

where $E_T^{\bar{q}^{(0)}}$ is the accumulated emission of the firm , $E_{\max}$ is the cap on the emission the firm and is the so-called risk-neutral measure or the stochastic discount factor of the firm. By the concavity assumption on e. It is clear that $\bar{q}^{(0)} > q^{(0)}$.

On the existence of the market, on has the relation $\pi'(q^{(1)}) + V_e^{(2)} e'(q^{(1)}) = 0$ for the small producers. In order to have the comparison with respect to previous cases, we need to pass through the crucial step of verifying $V_e^{(2)} = -S_t$ which indicates that $q^{(0)} > q^{(1)}$. Despite $\bar{q}^{(1)}$, $q^{(1)}$ does not depend on the utility of the firm and so the market approach provides an externality for the carbon price which allows to manage the production without knowing the utility of the firm.

For large producers with no impact on the risk premium of the market, comparison is provided by

$$\pi'(q^{(2)}) - e'(q^{(2)}) \left(S_t - V_y^{(2)}(t, E_t^{q^{(2)}}, Y_t^{q^{(2)}})\right)$$

where $V^{(2)}$ is the value function of the firm which corresponds to the optimization problem and V_y is the sensitivity of the value function with respect to the total emission of CO_2 and Y_t^q is the total emission process according to the production activity q of the large producer. We show that V_y is non–positive and therefore, $q^{(2)} \leq q^{(1)}$ which means that the large producer should even reduce his production policy more than the case of small producer.

For the large producer which has impact on the risk premium of the market, we have

$$\pi'(q^{(3)}) + \frac{1}{\eta}(\lambda\lambda')(q^{(3)}) + e'(q^{(3)})(V_e^{(3)} + \beta V_y^{(3)}) - \gamma\lambda'(q^{(3)})V_y^{(3)} = 0$$

where $\lambda(q)$ is the the risk premium according to the production activity q of large producer, $V^{(3)}$ is the value function of the firm, $V_y^{(3)}$ is the sensitivity of the value function with respect to the total emission of CO_2, $V_e^{(3)}$ is the sensitivity of the value function with respect to the production policy, and γ, η and β are positive constants in the model. In order to have the comparison with respect to previous cases, one need to verify $V_e^{(3)} = -S_t$. Then, the comparison of $q^{(3)}$ by $q^{(1)}$ and $q^{(2)}$ depends on the sign of the following term:

$$-e'(q^{(3)})\beta V_y^{(3)} + \lambda'(q^{(3)})\left(\gamma V_y^{(3)} - \frac{1}{\eta}\lambda(q^{(3)})\right)$$

We provided numerical examples to show that this is possible to have $q^{(3)}$ greater than $q^{(2)}$.

Chapter 2

A Probabilistic Numerical Method for Fully Nonlinear Parabolic PDEs

This Chapter[1] is organized as follows. In Section 2.1, we provide a natural presentation of the scheme without appealing to the theory of backward stochastic differential equations. Section 2.2 is dedicated to the asymptotic analysis of the discrete-time approximation, and contains our first main convergence result and the corresponding error estimate. In Section 2.3, we introduce the implementable backward scheme, and we further investigate the induced Monte Carlo error. We again prove convergence and we provide bounds on the approximation error. Finally, Section 2.4 contains some numerical results for the mean curvature flow equation on the plane and space, and for a five-dimensional Hamilton-Jacobi-Bellman equation arising in the problem of portfolio optimization in financial mathematics.

2.1 Discretization

Let μ and σ be two maps from $\mathbb{R}_+ \times \mathbb{R}^d$ to $\mathbb{R}^d$ and $\mathbb{M}(d,d)$, respectively. With $a := \sigma\sigma^{\mathrm{T}}$. We define the linear operator:

$$\mathcal{L}^X\varphi \quad := \quad \frac{\partial\varphi}{\partial t} + \mu \cdot D\varphi + \frac{1}{2} a \cdot D^2\varphi.$$

Given a map

$$F : (t,x,r,p,\gamma) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \quad \longmapsto \quad F(x,r,p,\gamma) \in \mathbb{R}$$

we consider the Cauchy problem:

$$-\mathcal{L}^X v - F\left(\cdot, v, Dv, D^2 v\right) = 0, \quad \text{on } [0,T) \times \mathbb{R}^d, \tag{2.1.1}$$

$$v(T,\cdot) = g, \quad \text{on } \in \mathbb{R}^d. \tag{2.1.2}$$

Under some conditions, a stochastic representation for the solution of this problem was provided in [25] by means of the newly introduced notion of second order backward stochastic differential equations. As an important implication, such a stochastic representation suggests a probabilistic numerical scheme for the above Cauchy problem.

[1] This work is reported on a paper co–authored with Nizar Touzi and Xavier Warin.

The chief goal of this section is to obtain the probabilistic numerical scheme suggested in [25] by a direct manipulation of (2.1.1)-(2.1.2) without appealing to the notion of backward stochastic differential equations.

To do this, we consider an $\mathbb{R}^d$-valued Brownian motion W on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F} = \{\mathcal{F}_t, t \in [0,T]\}$ satisfies the usual completeness conditions, and $\mathcal{F}_0$ is trivial.

For a positive integer n, let $h := T/n$, $t_i = ih$, $i = 0, \ldots, n$, and consider the one step ahead Euler discretization

$$\hat{X}_h^{t,x} := x + \mu(t,x)h + \sigma(t,x)(W_{t+h} - W_t), \tag{2.1.3}$$

of the diffusion X corresponding to the linear operator $\mathcal{L}^X$. Our analysis does not require any existence and uniqueness result for the underlying diffusion X. However, the subsequent formal discussion assumes it in order to provides a natural justification of our numerical scheme.

Assuming that the PDE (2.1.1) has a classical solution, it follows from Itô's formula that

$$\mathbb{E}_{t_i,x}\left[v\left(t_{i+1}, X_{t_{i+1}}\right)\right] \quad = \quad v\left(t_i, x\right) + \mathbb{E}_{t_i,x}\left[\int_{t_i}^{t_{i+1}} \mathcal{L}^X v(t, X_t)dt\right]$$

where we ignored the difficulties related to local martingale part, and $\mathbb{E}_{t_i,x} := \mathbb{E}[\cdot | X_{t_i} = x]$ denotes the expectation operator conditional on $\{X_{t_i} = x\}$. Since v solves the PDE (2.1.1), this provides

$$v(t_i, x) \quad = \quad \mathbb{E}_{t_i,x}\left[v\left(t_{i+1}, X_{t_{i+1}}\right)\right] + \mathbb{E}_{t_i,x}\left[\int_{t_i}^{t_{i+1}} F(\cdot, v, Dv, D^2 v)(t, X_t)dt\right].$$

By approximating the Riemann integral, and replacing the process X by its Euler discretization, this suggest the following approximation of the value function v

$$v^h(T,.) := g \quad \text{and} \quad v^h(t_i, x) := \mathbf{T}_h[v^h](t_i, x), \tag{2.1.4}$$

where we denoted for a function $\psi : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}$ with exponential growth:

$$\mathbf{T}_h[\psi](t,x) := \mathbb{E}\left[\psi(t+h, \hat{X}_h^{t,x})\right] + hF\left(\cdot, \mathcal{D}_h\psi\right)(t,x), \tag{2.1.5}$$

$$\mathcal{D}_h^k\psi(t,x) := \mathbb{E}[D^k\psi(t+h, \hat{X}_h^{t,x})], \quad k = 0,1,2, \quad \mathcal{D}_h\psi := \left(\mathcal{D}_h^0\psi, \mathcal{D}_h^1\psi, \mathcal{D}_h^2\psi\right)^{\mathrm{T}}, \tag{2.1.6}$$

and D^k is the $k-$th order partial differential operator with respect to the space variable x. The differentiations in the above scheme are to be understood in the sense of distributions. This algorithm is well-defined whenever g has exponential growth and F is a Lipschitz map. To see this, observe that any function with exponential growth has weak gradient and Hessian because the Gaussian kernel is a Schwartz function, and the exponential growth is inherited at each time step from the Lipschitz property of F.

At this stage, the above backward algorithm presents the serious drawback of involving the gradient $Dv^h(t_{i+1}, .)$ and the Hessian $D^2v^h(t_{i+1}, .)$ in order to compute $v^h(t_i, .)$. The following result avoids this difficulty by an easy integration by parts argument.

Lemma 2.1.1 *For every function $\varphi : Q_T \rightarrow \mathbb{R}$ with exponential growth, we have:*

$$\mathcal{D}_h\varphi(t_i,x) \quad = \quad \mathbb{E}\left[\varphi(t_{i+1},\hat{X}_h^{t_i,x})H_h(t_i,x)\right],$$

where $H_h = (H_0^h, H_1^h, H_2^h)^T$ and

$$H_0^h = 1, \quad H_1^h = \left(\sigma^{\mathrm{T}}\right)^{-1}\frac{W_h}{h}, \quad H_2^h = \left(\sigma^{\mathrm{T}}\right)^{-1}\frac{W_hW_h^{\mathrm{T}} - h\mathbf{I}_d}{h^2}\,\sigma^{-1}. \qquad (2.1.7)$$

Proof. The main ingredient is the following easy observation. Let G be a one dimensional Gaussian random variable with unit variance. Then, for any function $f : \mathbb{R} \longrightarrow \mathbb{R}$ with exponential growth, we have:

$$\mathbb{E}[f(G)H^k(G)] \quad = \quad \mathbb{E}[f^{(k)}(G)], \qquad (2.1.8)$$

where $f^{(k)}$ is the $k-$th order derivative of f in the sense of distributions, and H^k is the one-dimensional Hermite polynomial of degree k.

1 Now, let $\varphi : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a function with exponential growth. Then, by direct conditioning, it follows from (2.1.8) that

$$\mathbb{E}\left[\varphi(\hat{X}_h^{t,x})W_h^i\right] \quad = \quad h\sum_{j=1}^{d}\mathbb{E}\left[\frac{\partial\varphi}{\partial x_j}(\hat{X}_h^{t,x})\sigma_{ji}(t,x)\right],$$

and therefore:

$$\mathbb{E}\left[\varphi(\hat{X}_h^{t,x})H_1^h(t,x)\right] \quad = \quad \sigma(t,x)^{\mathrm{T}}\mathbb{E}\left[\nabla\varphi(\hat{X}_h^{t,x})\right].$$

2 For $i \neq j$, it follows from (2.1.8) that

$$\begin{aligned}\mathbb{E}\left[\varphi(\hat{X}_h^{t,x})W_h^iW_h^j\right] \quad &= \quad h\sum_{k=1}^{d}\mathbb{E}\left[\frac{\partial\varphi}{\partial x_k}(\hat{X}_h^{t,x})W_h^j\sigma_{ki}(t,x)\right] \\ &= \quad h^2\sum_{k,l=1}^{d}\mathbb{E}\left[\frac{\partial^2\varphi}{\partial x_k\partial x_l}(\hat{X}_h^{t,x})\sigma_{lj}(t,x)\sigma_{ki}(t,x)\right],\end{aligned}$$

and for $j = i$:

$$\mathbb{E}\left[\varphi(\hat{X}_h^{t,x})((W_h^i)^2 - h)\right] \quad = \quad h^2\sum_{k,l=1}^{d}\mathbb{E}\left[\frac{\partial^2\varphi}{\partial x_k\partial x_l}(\hat{X}_h^{t,x})\sigma_{li}(t,x)\sigma_{ki}(t,x)\right].$$

This provides:

$$\mathbb{E}\left[\varphi(\hat{X}_h^{t,x})H_2^h(t,x)\right] \quad = \quad \sigma(t,x)^{\mathrm{T}}\mathbb{E}\left[\nabla^2\varphi(\hat{X}_h^{t,x})\sigma(t,x)\right].$$

□

In view of Lemma 2.1.1, the iteration which computes $v^h(t_i,.)$ out of $v^h(t_{i+1},.)$ in (2.1.4)-(3.1.8) does not involve the gradient and the Hessian of the latter function.

Remark 2.1.1 Clearly, one can proceed to different choices for the integration by parts in Lemma 2.1.1. One such possibility leads to the representation of $\mathcal{D}_2^h\varphi$ as:

$$\mathcal{D}_2^h\varphi(t,x) \quad = \quad \mathbb{E}\left[\varphi(\hat{X}_h^{t,x})(\sigma^{\mathrm{T}})^{-1}\frac{W_{h/2}}{(h/2)}\frac{W_{h/2}^{\mathrm{T}}}{(h/2)}\sigma^{-1}\right].$$

This representation shows that the backward scheme (2.1.4) is very similar to the probabilistic numerical algorithm suggested in [25].

Observe that the choice of the drift and the diffusion coefficients μ and σ in the nonlinear PDE (2.1.1) is arbitrary. So far, it has been only used in order to define the underlying diffusion X. Our convergence result will however place some restrictions on the choice of the diffusion coefficient, see Remark 2.2.3.

Once the linear operator $\mathcal{L}^X$ is chosen in the nonlinear PDE, the above algorithm handles the remaining nonlinearity by the classical finite differences approximation. This connection with finite differences is motivated by the following formal interpretation of Lemma 2.1.1, where for ease of presentation, we set $d = 1$, $\mu \equiv 0$, and $\sigma(x) \equiv 1$:

- Consider the binomial random walk approximation of the Brownian motion $\hat{W}_{t_k} := \sum_{j=1}^k w_j$, $t_k := kh$, $k \geq 1$, where $\{w_j, j \geq 1\}$ are independent random variables distributed as $\frac{1}{2}\left(\delta_{\sqrt{h}} + \delta_{-\sqrt{h}}\right)$. Then, this induces the following approximation:

$$\mathcal{D}_h^1\psi(t,x) := \mathbb{E}\left[\psi(t+h, X_h^{t,x})H_1^h\right] \quad \approx \quad \frac{\psi(t, x+\sqrt{h}) - \psi(t, x-\sqrt{h})}{2\sqrt{h}},$$

  which is the centered finite differences approximation of the gradient.

- Similarly, consider the trinomial random walk approximation $\hat{W}_{t_k} := \sum_{j=1}^k w_j$, $t_k := kh$, $k \geq 1$, where $\{w_j, j \geq 1\}$ are independent random variables distributed as $\frac{1}{6}\left(\delta_{\{\sqrt{3h}\}} + 4\delta_{\{0\}} + \delta_{\{-\sqrt{3h}\}}\right)$, so that $\mathbb{E}[w_j^n] = \mathbb{E}[W_h^n]$ for all integers $n \leq 4$. Then, this induces the following approximation:

$$\mathcal{D}_h^2\psi(t,x) := \mathbb{E}\left[\psi(t+h, X_h^{t,x})H_2^h\right] \quad \approx \quad \frac{\psi(t, x+\sqrt{3h}) - 2\psi(t,x) + \psi(t, x-\sqrt{3h})}{3h},$$

  which is the centered finite differences approximation of the Hessian.

In view of the above interpretation, the numerical scheme studied in this paper can be viewed as a mixed Monte Carlo–Finite Differences algorithm. The Monte Carlo component of the scheme consists in the choice of an underlying diffusion process X. The finite differences component of the scheme consists in approximating the remaining nonlinearity by means of the integration-by-parts formula of Lemma 2.1.1.

2.2 Asymptotics of the discrete-time approximation

2.2.1 The main results

Our first main convergence results follow the general methodology of Barles and Souganidis [6], and requires that the nonlinear PDE (2.1.1) satisfies a comparison result in the sense of viscosity solutions.

We recall that an upper-semicontinuous (resp. lower semicontinuous) function $\underline{v}$ (resp. $\overline{v}$) on $[0,T]\times\mathbb{R}^d$, is called a viscosity subsolution (resp. supersolution) of (2.1.1) if for any $(t,x)\in[0,T)\times\mathbb{R}^d$ and any smooth function φ satisfying

$$0=(\underline{v}-\varphi)(t,x)=\max_{[0,T]\times\mathbb{R}^d}(\underline{v}-\varphi)\left(\text{resp. } 0=(\overline{v}-\varphi)(t,x)=\min_{[0,T]\times\mathbb{R}^d}(\overline{v}-\psi)\right),$$

we have:

$$-\mathcal{L}^X\varphi-F(t,x,\mathcal{D}\varphi(t,x))\quad\le\quad(\text{resp. }\ge)\quad 0.$$

Definition 2.2.1 *We say that (2.1.1) has comparison for bounded functions if for any bounded upper semicontinuous subsolution $\underline{v}$ and any bounded lower semicontinuous supersolution $\overline{v}$ on $[0,T)\times\mathbb{R}^d$, satisfying*

$$\underline{v}(T,\cdot)\le\overline{v}(T,\cdot),$$

we have $\underline{v}\le\overline{v}$.

Remark 2.2.1 Barles and Souganidis [6] use a stronger notion of comparison by accounting for the final condition, thus allowing for a possible boundary layer. In their context, a supersolution $\overline{v}$ and a subsolution $\underline{v}$ satisfy:

$$\min\left\{-\mathcal{L}^X\overline{v}(T,x)-F(T,x,\mathcal{D}\overline{v}(T,x)),\overline{v}(T,x)-g(x)\right\}\quad\le\quad 0 \tag{2.2.1}$$

$$\max\left\{-\mathcal{L}^X\underline{v}(T,x)-F(T,x,\mathcal{D}\underline{v}(T,x)),\underline{v}(T,x)-g(x)\right\}\quad\ge\quad 0. \tag{2.2.2}$$

We observe that, by the nature of our equation, (2.2.1) and (2.2.2) imply that the subsolution $\underline{v}\le g$ and the supersolution $\overline{v}\ge g$, i.e. the final condition holds in the usual sense, and no boundary layer can occur. To see this, without loss of generality we suppose that $F(t,x,r,p,\gamma)$ is decreasing with respect to r (see Remark 2.2.7). Let φ be a function satisfying

$$0=(\underline{v}-\varphi)(T,x)=\max_{[0,T]\times\mathbb{R}^d}(\underline{v}-\varphi).$$

Then define $\varphi_K(t,\cdot)=\varphi(t,\cdot)+K(T-t)$ for $K>0$. Then $\underline{v}-\varphi_K$ also has a maximum at (T,x), and the subsolution property (2.2.1) implies that

$$\min\left\{-\mathcal{L}^X\varphi(T,x)-F(T,x,\mathcal{D}\varphi(T,x))+K,\underline{v}(T,x)-g(x)\right\}\quad\le\quad 0.$$

For a sufficiently large K, this provides the required inequality $\underline{v}(T,x)-g(x)\le 0$. A similar argument shows that (2.2.1) implies that $\overline{v}-g\ge 0$.

In the sequel, we denote by F_r, F_p and F_γ the partial gradients of F with respect to r, p and γ, respectively. We also denote by F_γ^- the pseudo-inverse of the non-negative symmetric matrix F_γ. We recall that any Lipschitz function is differentiable a.e.

Assumption F (i) *The nonlinearity F is Lipschitz-continuous with respect to (x, r, p, γ) uniformly in t, and $|F(\cdot,\cdot,0,0,0)|_\infty < \infty$;*
(ii) F is elliptic and dominated by the diffusion of the linear operator $\mathcal{L}^X$, i.e.

$$\nabla_\gamma F \le a \quad \text{on} \quad \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d; \tag{2.2.3}$$

(iii) $F_p \in \text{Image}(F_\gamma)$ and $\left|F_p^{\mathrm{T}} F_\gamma^- F_p\right|_\infty < +\infty$.

Remark 2.2.2 Assumption **F** (iii) is equivalent to

$$|m_F^-|_\infty < \infty \quad \text{where} \quad m_F := \min_{w \in \mathbb{R}^d} \left\{ F_p \cdot w + w^{\mathrm{T}} F_\gamma w \right\}. \tag{2.2.4}$$

This is immediately seen by recalling that, by the symmetric feature of F_γ, any $w \in \mathbb{R}^d$ has an orthogonal decomposition $w = w_1 + w_2 \in \text{Ker}(F_\gamma) \oplus \text{Image}(F_\gamma)$, and by the nonnegativity of F_γ:

$$\begin{aligned} F_p \cdot w + w^{\mathrm{T}} F_\gamma w &= F_p \cdot w_1 + F_p \cdot w_2 + w_2^{\mathrm{T}} F_\gamma w_2 \\ &= -\frac{1}{4} F_p^{\mathrm{T}} F_\gamma^- F_p + F_p \cdot w_1 + \left| \frac{1}{2} (F_\gamma^-)^{1/2} \cdot F_p - F_\gamma^{1/2} w_2 \right|^2. \end{aligned}$$

Remark 2.2.3 The above Condition (3.3.2) places some restrictions on the choice of the linear operator $\mathcal{L}^X$ in the nonlinear PDE (2.1.1). First, F is required to be uniformly elliptic, implying an upper bound on the choice of the diffusion matrix σ. Since $\sigma\sigma^{\mathrm{T}} \in \mathcal{S}_d^+$, this implies in particular that our main results do not apply to general degenerate nonlinear parabolic PDEs. Second, the diffusion of the linear operator σ is required to dominate the nonlinearity F which places implicitly a lower bound on the choice of the diffusion σ.

Example 2.2.1 Let us consider the nonlinear PDE in the one-dimensional case $-\frac{\partial v}{\partial t} - \frac{1}{2}\left(a^2 v_{xx}^+ - b^2 v_{xx}^-\right)$ where $0 < b < a$ are given constants. Then if we restrict the choice of the diffusion to be constant, it follows from Condition **F** that $\frac{1}{3}a^2 \le \sigma^2 \le b^2$, which implies that $a^2 \le 3b^2$. If the parameters a and b do not satisfy the latter condition, then the diffusion σ has to be chosen to be state and time dependent.

Theorem 2.2.1 (Convergence) *Let Assumption **F** hold true, and $|\mu|_1$, $|\sigma|_1 < \infty$ and σ is invertible. Also assume that the fully nonlinear PDE (2.1.1) has comparison for bounded functions. Then for every bounded Lipschitz function g, there exists a bounded function v so that*

$$v^h \longrightarrow v \qquad \textit{locally uniformly.}$$

In addition, v is the unique bounded viscosity solution of problem (2.1.1)-(2.1.2).

Remark 2.2.4 Under the boundedness condition on the coefficients μ and σ, the restriction to a bounded terminal data g in the above Theorem 2.2.1 can be relaxed by an immediate change of variable. Let g be a function with $\alpha-$exponential growth for some $\alpha > 0$. Fix some $M > 0$, and let ρ be an arbitrary smooth positive function with:

$$\rho(x) = e^{\alpha|x|} \quad \text{for} \quad |x| \geq M,$$

so that both $\rho(x)^{-1}\nabla\rho(x)$ and $\rho(x)^{-1}\nabla^2\rho(x)$ are bounded. Let

$$u(t,x) \;:=\; \rho(x)^{-1}v(t,x) \quad \text{for} \quad (t,x) \in [0,T] \times \mathbb{R}^d.$$

Then, the nonlinear PDE problem (2.1.1)-(2.1.2) satisfied by v converts into the following nonlinear PDE for u:

$$\begin{aligned} -\mathcal{L}^X u - \tilde{F}\left(\cdot, u, Du, D^2u\right) = 0 \qquad & \text{on } [0,T) \times \mathbb{R}^d \\ v(T,\cdot) = \tilde{g} := \rho^{-1}g \qquad & \text{on } \mathbb{R}^d, \end{aligned} \tag{2.2.5}$$

where

$$\begin{aligned} \tilde{F}(t,x,r,p,\gamma) \;&:=\; r\mu(x)\cdot\rho^{-1}\nabla\rho + \frac{1}{2}\mathrm{Tr}\left[a(x)\left(r\rho^{-1}\nabla^2\rho + 2p\rho^{-1}\nabla\rho^{\mathrm{T}}\right)\right] \\ &\quad +\rho^{-1}F\left(t,x,r\rho,r\nabla\rho + p\rho, r\nabla^2\rho + 2p\nabla\rho^{\mathrm{T}} + \rho\gamma\right). \end{aligned}$$

Recall that the coefficients μ and σ are assumed to be bounded. Then, it is easy to see that $\tilde{F}$ satisfies the same conditions as F. Since $\tilde{g}$ is bounded, the convergence Theorem 2.2.1 applies to the nonlinear PDE (2.2.5). □

Remark 2.2.5 Theorem 2.2.1 states that the inequality (3.3.2) (i.e. diffusion must dominate the nonlinearity in γ) is sufficient for the convergence of the Monte Carlo–Finite Differences scheme. We do not know whether this condition is necessary:
• Subsection 2.2.4 suggests that this condition is not sharp in the simple linear case,
• however, our numerical experiments of Section 2.4 reveal that the method may have a poor performance in the absence of this condition, see Figure 2.5.

We next provide bounds on the rate of convergence of the Monte Carlo–Finite Differences scheme in the context of nonlinear PDEs of the Hamilton-Jacobi-Bellman type in the same context as [5]. The following assumptions are stronger than Assumption **F** and imply that the nonlinear PDE (2.1.1) satisfies a comparison result for bounded functions.

Assumption HJB *The nonlinearity F satisfies Assumption* **F***(ii)-(iii), and is of the Hamilton-Jacobi-Bellman type:*

$$\begin{aligned} \frac{1}{2}a\cdot\gamma + b\cdot p + F(t,x,r,p,\gamma) \;&=\; \inf_{\alpha\in\mathcal{A}}\{\mathcal{L}^\alpha(t,x,r,p,\gamma)\} \\ \mathcal{L}^\alpha(t,x,r,p,\gamma) \;&:=\; \frac{1}{2}Tr[\sigma^\alpha\sigma^{\alpha\mathrm{T}}(t,x)\gamma] + b^\alpha(t,x)p + c^\alpha(t,x)r + f^\alpha(t,x) \end{aligned}$$

where the functions μ, σ, σ^α, b^α, c^α and f^α satisfy:

$$|\mu|_\infty + |\sigma|_\infty + \sup_{\alpha\in\mathcal{A}} \left(|\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1\right) \quad < \quad \infty.$$

Assumption HJB+ *The nonlinearity F satisfies* **HJB**, *and for any $\delta > 0$, there exists a finite set $\{\alpha_i\}_{i=1}^{M_\delta}$ such that for any $\alpha \in \mathcal{A}$:*

$$\inf_{1\le i\le M_\delta} |\sigma^\alpha - \sigma^{\alpha_i}|_\infty + |b^\alpha - b^{\alpha_i}|_\infty + |c^\alpha - c^{\alpha_i}|_\infty + |f^\alpha - f^{\alpha_i}|_\infty \quad \le \quad \delta.$$

Remark 2.2.6 The assumption HJB+ is satisfied if $\mathcal{A}$ is a separable topological space and $\sigma^\alpha(\cdot)$, $b^\alpha(\cdot)$, $c^\alpha(\cdot)$ and $f^\alpha(\cdot)$ are continuous maps from $\mathcal{A}$ to $C_b^{\frac{1}{2},1}$; the space of bounded maps which are Lipschitz in x and $\frac{1}{2}$–Hölder in t.

Theorem 2.2.2 (Rate of Convergence) *Assume that the final condition g is bounded Lipschitz-continuous. Then, there is a constant $C > 0$ such that:*

(i) *under Assumption* HJB, *we have $v - v^h \le Ch^{1/4}$,*

(ii) *under the stronger condition* HJB+, *we have $-Ch^{1/10} \le v - v^h \le Ch^{1/4}$.*

The above bounds can be improved in some specific examples. See Subsection 2.2.4 for the linear case where the rate of convergence is improved to $\sqrt{h}$.

We also observe that, in the PDE Finite Differences literature, the rate of convergence is usually stated in terms of the discretization in the space variable $|\Delta x|$. In our context of stochastic differential equation, notice that $|\Delta x|$ is or the order of $h^{1/2}$. Therefore, the above upper and lower bounds on the rate of convergence corresponds to the classical rate $|\Delta x|^{1/2}$ and $|\Delta x|^{1/5}$, respectively.

2.2.2 Proof of the convergence result

We now provide the proof Theorem 2.2.1 by building on Theorem 2.1 and Remark 2.1 of Barles and Souganidis [6] which requires the scheme to be consistent, monotone and stable. Moreover, since we are assuming the (weak) comparison for the equation, we also need to prove that our scheme produces a limit which satisfies the terminal condition in the usual sense, see Remark 2.2.1.

Throughout this section, all the conditions of Theorem 2.2.1 are in force.

Lemma 2.2.1 *Let φ be a smooth function with bounded derivatives. Then for all $(t,x) \in [0,T]\times\mathbb{R}^d$:*

$$\lim_{\substack{(t',x')\to(t,x)\\(h,c)\to(0,0)\\t'+h\le T}} \frac{[c+\varphi](t',x') - \mathbf{T}_h[c+\varphi](t',x')}{h} \quad = \quad -\left(\mathcal{L}^X\varphi + F(\cdot,\varphi,D\varphi,D^2\varphi)\right)(t,x).$$

The proof is a straightforward application of Itô's formula, and is omitted.

Lemma 2.2.2 *Let $\varphi, \psi : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ be two Lipschitz functions. Then:*

$$\varphi \le \psi \implies \mathbf{T}_h[\varphi](t,x) \le \mathbf{T}_h[\psi](t,x) + Ch\,\mathbb{E}[(\psi-\varphi)(t+h, \hat{X}_h^{t,x})] \text{ for some } C > 0$$

where C depends only on constant K in (3.3.1).

Proof. By Lemma 2.1.1 the operator $\mathbf{T}_h$ can be written as:

$$\mathbf{T}_h[\psi](t,x) = \mathbb{E}\left[\psi(\hat{X}_h^{t,x})\right] + hF\left(t, x, \mathbb{E}[\psi(\hat{X}_h^{t,x})H_h(t,x)]\right).$$

Let $f := \psi - \varphi \ge 0$ where φ and ψ are as in the statement of the lemma. Let F_τ denote the partial gradient with respect to $\tau = (r,p,\gamma)$. By the mean value Theorem:

$$\begin{aligned}\mathbf{T}_h[\psi](t,x) - \mathbf{T}_h[\varphi](t,x) &= \mathbb{E}\left[f(\hat{X}_h^{t,x})\right] + hF_\tau(\theta)\cdot\mathcal{D}_h f(\hat{X}_h^{t,x}) \\ &= \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(1 + hF_\tau(\theta)\cdot H_h(t,x)\right)\right],\end{aligned}$$

for some $\theta = (t, x, \bar{r}, \bar{p}, \bar{\gamma})$. By the definition of $H_h(t,x)$:

$$\mathbf{T}_h[\psi] - \mathbf{T}_h[\varphi] = \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(1 + hF_r + F_p.(\sigma^{\mathrm{T}})^{-1}W_h + h^{-1}F_\gamma\cdot(\sigma^{\mathrm{T}})^{-1}(W_hW_h^{\mathrm{T}} - hI)\sigma^{-1}\right)\right],$$

where the dependence on θ and x has been omitted for notational simplicity. Since $F_\gamma \le a$ by (3.3.1) of Assumption **F**, we have $1 - a^{-1}\cdot F_\gamma \ge 0$ and therefore:

$$\begin{aligned}\mathbf{T}_h[\psi] - \mathbf{T}_h[\varphi] &\ge \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(hF_r + F_p.\sigma^{\mathrm{T}^{-1}}W_h + h^{-1}F_\gamma\cdot\sigma^{\mathrm{T}^{-1}}W_hW_h^{\mathrm{T}}\sigma^{-1}\right)\right] \\ &= \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(hF_r + hF_p.\sigma^{\mathrm{T}-1}\frac{W_h}{h} + hF_\gamma\cdot\sigma^{\mathrm{T}-1}\frac{W_hW_h^{\mathrm{T}}}{h^2}\sigma^{-1}\right)\right].\end{aligned}$$

Let $m_F^- := \max\{-m_F, 0\}$, where the function m_F is defined in (3.3.1). Under Assumption **F**, we have $K := |m_F^-|_\infty < \infty$, then

$$F_p.\sigma^{\mathrm{T}-1}\frac{W_h}{h} + hF_\gamma\cdot\sigma^{\mathrm{T}-1}\frac{W_hW_h^{\mathrm{T}}}{h^2}\sigma^{-1} \ge -K$$

one can write,

$$\mathbf{T}_h[\psi] - \mathbf{T}_h[\varphi] \ge \mathbb{E}\left[f(\hat{X}_h^{t,x})(hF_r - hK)\right] \ge -C'h\mathbb{E}\left[f(\hat{X}_h^{t,x})\right]$$

for some constant $C > 0$, where the last inequality follows from (3.3.1). □

The following observation will be used in the proof of Theorem 2.2.2 below.

Remark 2.2.7 The monotonicity result of the previous Lemma 2.2.2 is slightly different from that required in [6]. However, as it is observed in Remark 2.1 in [6], their convergence theorem holds under this approximate monotonicity. From the previous proof, we observe that if the function F satisfies the condition:

$$F_r - \frac{1}{4}F_p^{\mathrm{T}}F_\gamma^- F_p \quad \geq \quad 0, \tag{2.2.6}$$

then, the standard monotonicity condition

$$\varphi \ \leq \ \psi \quad \Longrightarrow \quad \mathbf{T}_h[\varphi](t,x) \ \leq \ \mathbf{T}_h[\psi](t,x) \tag{2.2.7}$$

holds. Using the parabolic feature of the equation, we may introduce a new function $u(t,x) := e^{\theta(T-t)}v(t,x)$ which solves a nonlinear PDE satisfying (2.2.6). Indeed, direct calculation shows that the PDE inherited by u is:

$$-\mathcal{L}^X u - \overline{F}\left(\cdot, u, Du, D^2u\right) = 0, \quad \text{on} \quad [0,T)\times\mathbb{R}^d \tag{2.2.8}$$

$$u(T,x) = g(x), \quad \text{on} \quad \mathbb{R}^d, \tag{2.2.9}$$

where $\overline{F}(t,x,r,p,\gamma) = e^{\theta(T-t)}F(t,x,e^{-\theta(T-t)}r, e^{-\theta(T-t)}p, e^{-\theta(T-t)}\gamma) + \theta r$. Then, it is easily seen that $\overline{F}$ satisfies the same conditions as F together with (2.2.6) for sufficiently large θ.

Lemma 2.2.3 *Let $\varphi, \psi : [0,T]\times\mathbb{R}^d \longrightarrow \mathbb{R}$ be two $\mathbb{L}^\infty$−bounded functions. Then there exists a constant $C > 0$ such that*

$$|\mathbf{T}_h[\varphi] - \mathbf{T}_h[\psi]|_\infty \leq |\varphi - \psi|_\infty (1 + Ch)$$

In particular, if g is $\mathbb{L}^\infty$−bounded, the family $(v^h)_h$ defined in (2.1.4) is $\mathbb{L}^\infty$−bounded, uniformly in h.

Proof. Let $f := \varphi - \psi$. Then, arguing as in the previous proof,

$$\mathbf{T}_h[\varphi] - \mathbf{T}_h[\psi] \quad = \quad \mathbb{E}\left[f(\hat{X}_h)\left(1 - a^{-1}\cdot F_\gamma + h|A_h|^2 + hF_r - \frac{h}{4}F_p^{\mathrm{T}}F_\gamma^- F_p\right)\right].$$

where

$$A_h = \frac{1}{2}(F_\gamma^-)^{1/2}F_p - F_\gamma^{1/2}\sigma^{\mathrm{T}\,-1}\frac{W_h}{h}.$$

Since $1 - \mathrm{Tr}[a^{-1}F_\gamma] \geq 0$, $|F_r|_\infty < \infty$, and $|F_p^{\mathrm{T}}F_\gamma^- F_p|_\infty < \infty$ by Assumption **F**, it follows that

$$|\mathbf{T}_h[\varphi] - \mathbf{T}_h[\psi]|_\infty \quad \leq \quad |f|_\infty \left(1 - a^{-1}\cdot F_\gamma + h\mathbb{E}[|A_h|^2] + Ch\right)$$

But, $\mathbb{E}[|A_h|^2] = \frac{h}{4}F_p^{\mathrm{T}}F_\gamma^- F_p + a^{-1}\cdot F_\gamma$. Therefore, by Assumption **F**

$$|\mathbf{T}_h[\varphi] - \mathbf{T}_h[\psi]|_\infty \quad \leq \quad |f|_\infty\left(1 + \frac{h}{4}F_p^{\mathrm{T}}F_\gamma^- F_p + Ch\right) \ \leq \ |f|_\infty(1 + \bar{C}h).$$

To prove that the family $(v^h)_h$ is bounded, we proceed by backward induction. By the assumption of the lemma $v^h(T,.) = g$ is $\mathbb{L}^\infty-$bounded. We next fix some $i < n$ and we assume that $|v^h(t_j,.)|_\infty \le C_j$ for every $i+1 \le j \le n-1$. Proceeding as in the proof of Lemma 2.2.2 with $\varphi \equiv v^h(t_{i+1},.)$ and $\psi \equiv 0$, we see that

$$\left|v^h(t_i,.)\right|_\infty \quad \le \quad h\,|F(t,x,0,0,0)| + C_{i+1}(1+Ch).$$

Since $F(t,x,0,0,0)$ is bounded by Assumption **F**, it follows from the discrete Gronwall inequality that $|v^h(t_i,.)|_\infty \le Ce^{CT}$ for some constant C independent of h. □

Remark 2.2.8 The approximate function v^h defined by (2.1.4) is only defined on $\{ih | i = 0, \cdots, N\} \times \mathbb{R}^d$. Our methodology requires to extend it to any $t \in [0,T]$. This can be achieved by any interpolation, as long as the regularity property of v^h mentioned in Lemma 2.2.4 below is preserved. For instance, on may simply use linear interpolation.

Lemma 2.2.4 *The function v^h is Lipschitz in x, uniformly in h.*

Proof. We report the following calculation in the one-dimensional case $d = 1$ in order to simplify the presentation.
1. For fixed $t \in [0, T-h]$, we argue as in the proof of Lemma 2.2.2 to see that for $x, x' \in \mathbb{R}^d$ with $x > x'$:

$$v^h(t,x) - v^h(t,x') \quad = \quad A + hB, \tag{2.2.10}$$

where, denoting $\delta^{(k)} := D^k v^h(t+h, \hat{X}_h^{t,x}) - D^k v^h(t+h, \hat{X}_h^{t,x'})$ for $k = 0,1,2$:

$$\begin{aligned} A \quad &:= \quad \mathbb{E}\big[\delta^{(0)}\big] + h\Big(F\Big(t,x',\mathcal{D}v^h(t+h,\hat{X}_h^{t,x})\Big) - F\Big(t,x',\mathcal{D}v^h(t+h,\hat{X}_h^{t,x'})\Big)\Big) \\ &= \quad \mathbb{E}\big[(1+hF_r)\delta^{(0)} + hF_p\delta^{(1)} + hF_\gamma\delta^{(2)}\big], \\ |B| \quad &:= \quad \Big|F\Big(t,x,\mathcal{D}v^h(t+h,\hat{X}_h^{t,x})\Big) - F\Big(t,x',\mathcal{D}v^h(t+h,\hat{X}_h^{t,x})\Big)\Big| \le |F_x|_\infty |x-x'|, \end{aligned}$$

by Assumption **F** (i). By Lemma 2.1.1 we write for $k = 1,2$:

$$\begin{aligned} \mathbb{E}\big[\delta^{(k)}\big] \quad &= \quad \mathbb{E}\big[\delta^{(0)}H_k^h(t,x) + v^h(t+h,\hat{X}_h^{t,x'})\Big(H_k^h(t,x) - H_k^h(t,x')\Big)\big] \\ &= \quad \mathbb{E}\Big[\delta^{(0)}H_k^h(t,x) + Dv^h(t+h,\hat{X}_h^{t,x'})\left(\frac{W_h}{h}\right)^{k-1}\Big(\sigma(t,x)^{-k} - \sigma(t,x')^{-k}\Big)\sigma(t,x')\Big]. \end{aligned}$$

Then, dividing both sides of (2.2.10) by $x - x'$ and taking limsup, if follows from the above equalities that

$$\begin{aligned} &\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x) - v^h(t,x')|}{(x-x')} \\ &\le \quad \mathbb{E}\Bigg[\Bigg|\limsup_{|x-x'|\searrow 0} \frac{v^h(t+h,\hat{X}_h^{t,x}) - v^h(t+h,\hat{X}_h^{t,x'})}{(x-x')}\Big(1 + hF_r + F_p\frac{W_h}{\sigma(t,x)} + F_\gamma\frac{W_h^2 - h}{\sigma(t,x)^2 h}\Big) \\ &\qquad + Dv^h(t+h,\hat{X}_h^{t,x})\left(W_hF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2} + hF_p\frac{\sigma_x(t,x)}{\sigma(t,x)}\right)\Bigg|\Bigg] + Ch. \end{aligned}$$

2. Assume $v^h(t+h,.)$ is Lipschitz with constant L_{t+h}. Then

$$\begin{aligned}
&\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{(x-x')} \\
&\leq \;\; L_{t+h}\mathbb{E}\Bigg[\Bigg|(1+\mu_x(t,x)h+\sigma_x(t,x)\sqrt{h}N)\Big(1+hF_r+F_p\frac{\sqrt{h}N}{\sigma(t,x)}+F_\gamma\frac{N^2}{\sigma(t,x)^2}-\frac{F_\gamma}{\sigma(t,x)^2}\Big) \\
&\qquad\qquad +\sqrt{h}NF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2}+hF_p\frac{\sigma_x(t,x)}{\sigma(t,x)}\Bigg|\Bigg]+Ch.
\end{aligned}$$

Observe that

$$F_p\frac{\sigma_x}{\sigma}=\sigma_x\frac{F_p}{\sqrt{F_\gamma}}\frac{\sqrt{F_\gamma}}{\sigma}\;\mathbb{1}_{F_\gamma\neq 0}.$$

Since all terms on the right hand-side are bounded, under our assumptions, it follows that $|F_p\frac{\sigma_x}{\sigma}|_\infty<\infty$ (we emphasize that the geometric structure imposed in Assumption **F** (iii) provides this result in any dimension). Then:

$$\begin{aligned}
&\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{(x-x')} \\
&\leq \;\; L_{t+h}\Bigg(\mathbb{E}\Bigg[\Bigg|(1+\mu_x(t,x)h+\sigma_x(t,x)\sqrt{h}N)\Big(1+F_p\frac{\sqrt{h}N}{\sigma(t,x)}+F_\gamma\frac{N^2}{\sigma(t,x)^2}-\frac{F_\gamma}{\sigma(t,x)^2}\Big) \\
&\qquad\qquad +\sqrt{h}NF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2}\Bigg|\Bigg]+Ch\Bigg)+Ch.
\end{aligned}$$

3. Let $\tilde{\mathbb{P}}$ be the probability measure equivalent to $\mathbb{P}$ defined by the density

$$Z:=1-\alpha+\alpha N^2 \quad \text{where} \quad \alpha=\frac{F_\gamma}{\sigma(t,x)^2}.$$

Then,

$$\begin{aligned}
\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{(x-x')} \leq & L_{t+h}\Bigg(\mathbb{E}^{\tilde{\mathbb{P}}}\Bigg[\Bigg|\Big(1+\mu_x(t,x)h+\sigma_x(t,x)\sqrt{h}N\Big)\Big(1+Z^{-1}F_p\frac{\sqrt{h}N}{\sigma(t,x)}\Big) \\
&+Z^{-1}\sqrt{h}NF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2}\Bigg|\Bigg]+Ch\Bigg)+Ch.
\end{aligned}$$

By Cauchy–Schwartz inequality, we have

$$\begin{aligned}
\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{x-x'} \leq & L_{t+h}\Bigg(\mathbb{E}^{\tilde{\mathbb{P}}}\Bigg[\Bigg|\Big(1+\mu_x(t,x)h+\sigma_x(t,x)\sqrt{h}N\Big)\Big(1+Z^{-1}F_p\frac{\sqrt{h}N}{\sigma(t,x)}\Big) \\
&+Z^{-1}\sqrt{h}NF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2}\Bigg|^2\Bigg]^{\frac{1}{2}}+Ch\Bigg)+Ch
\end{aligned}$$

By writing back the expectation in terms of probability $\mathbb{P}$,

$$\begin{aligned}
\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{x-x'} \leq & L_{t+h}\Bigg(\mathbb{E}\Bigg[Z\Bigg|\Big(1+\mu_x(t,x)h+\sigma_x(t,x)\sqrt{h}N\Big)\Big(1+Z^{-1}F_p\frac{\sqrt{h}N}{\sigma(t,x)}\Big) \\
&+Z^{-1}\sqrt{h}NF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2}\Bigg|^2\Bigg]^{\frac{1}{2}}+Ch\Bigg)+Ch.
\end{aligned}$$

By expanding the quadratic term inside the expectation, we observe that expectation of all the terms having $\sqrt{h}$, is zero. Therefore,

$$\begin{aligned}\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{(x-x')} \le & L_{t+h}\Bigg(\mathbb{E}^{\tilde{\mathbb{P}}}\Bigg[\Bigg|\Big(1+\mu_x(t,x)h+\sigma_x(t,x)\sqrt{h}N\Big)\Big(1+Z^{-1}F_p\frac{\sqrt{h}N}{\sigma(t,x)}\Big) \\ & +Z^{-1}\sqrt{h}NF_\gamma\frac{-2\sigma_x(t,x)}{\sigma(t,x)^2}\Bigg|^2\Bigg]^{\frac{1}{2}}+Ch\Bigg)+Ch \\ \le & L_{t+h}\Big((1+C'h)^{\frac{1}{2}}+Ch\Big)+Ch,\end{aligned}$$

which leads to

$$\limsup_{|x-x'|\searrow 0} \frac{|v^h(t,x)-v^h(t,x')|}{(x-x')} \quad \le \quad Ce^{C'T/2},$$

for some constants $C, C' > 0$. □

Finally, we prove that the terminal condition is preserved by our scheme as the time step shrinks to zero.

Lemma 2.2.5 *For each $x \in \mathbb{R}^d$ and $t_k = kh$ with $k = 1, \cdots, n$, we have;*

$$|v^h(t_k,x)-g(x)| \quad \le \quad C(T-t_k)^{\frac{1}{2}}.$$

Proof. **1.** By the same argument as in the proof of Lemma 2.2.3, we have: and for $j \ge i$:

$$\begin{aligned}v^h(t_j,\hat{X}^{t_i,x}_{t_j}) \quad = \quad & \mathbb{E}_{t_j}\Big[v^h(t_{j+1},\hat{X}^{t_i,x}_{t_{j+1}})\left(1-\alpha_j+\alpha_j N_j^2\right)\Big] \\ & +h\Big(F_0^j+F_r^j\mathbb{E}_{t_j}[v^h(t_{j+1},\hat{X}^{t_i,x}_{t_{j+1}})]+F_p^j\cdot\mathbb{E}_{t_j}[Dv^h(t_{j+1},\hat{X}^{t_i,x}_{t_{j+1}})]\Big),\end{aligned}$$

where $F_0^j := F(t_j,\hat{X}^{t_i,x}_{t_j},0,0,0)$, α_j, F_r^j, F_p^j are $\mathcal{F}_{t_j}-$adapted random variables defined as in the proof of Lemma 2.2.3 at t_j, and $N_j = \frac{W_{t_{j+1}}-W_{t_j}}{\sqrt{h}}$ has a standard Gaussian distribution. Combine the above formula for j from i to $n-1$, we see that

$$v^h(t_i,x)=\mathbb{E}\left[g(\hat{X}^{t_i,x}_T)P_{i,n}\right]+h\mathbb{E}\sum_{j=i}^{n-1}F_0^j+F_r^j\mathbb{E}_{t_j}[v^h(t_{j+1},\hat{X}^{t_i,x}_{t_{j+1}})]+F_p^j\cdot\mathbb{E}_{t_j}[Dv^h(t_{j+1},\hat{X}^{t_i,x}_{t_{j+1}})],$$

where $P_{i,k} := \prod_{j=i}^{k-1}\left(1-\alpha_j+\alpha_j N_j^2\right) > 0$ a.s. for all $1 \le i < k \le n$ and $P_{i,i} = 1$. Obviously $\{P_{i,k}, i \le k \le n\}$ is a martingale for all $i \le n$, a property which will be used later. Since $|F(\cdot,\cdot,0,0,0)|_\infty < +\infty$, and using Assumption **F** and Lemmas 2.2.4 and 2.2.3:

$$|v^h(t_i,x)-g(x)| \quad \le \quad \left|\mathbb{E}\left[\left(g(\hat{X}^{t_i,x}_T)-g(x)\right)P_{i,n}\right]\right|+C(T-t_i). \quad (2.2.11)$$

2. Let $\{g_\varepsilon\}_\varepsilon$ be the family of smooth functions obtained from g by convolution with a family of mollifiers $\{\rho_\varepsilon\}$, i.e. $g_\varepsilon = g * \rho_\varepsilon$. Note that we have

$$|g_\varepsilon - g|_\infty \le C\varepsilon, \quad |Dg_\varepsilon|_\infty \le |Dg|_\infty \quad \text{and} \quad |D^2 g_\varepsilon|_\infty \le \varepsilon^{-1}|Dg|_\infty. \tag{2.2.12}$$

Then:

$$\begin{aligned}
\left|\mathbb{E}\left[\left(g(\hat X_T^{t_i,x}) - g(x)\right) P_{i,n}\right]\right| &\le \mathbb{E}\left[\left|g(\hat X_T^{t_i,x}) - g_\varepsilon(\hat X_T^{t_i,x}) P_{i,n}\right|\right] \\
&\quad + \left|\mathbb{E}\left[\left(g_\varepsilon(\hat X_T^{t_i,x}) - g_\varepsilon(x)\right) P_{i,n}\right]\right| + |g_\varepsilon - g|_\infty \\
&\le C\varepsilon + \left|\mathbb{E}\left[\left(g_\varepsilon(\hat X_T^{t_i,x}) - g_\varepsilon(x)\right) P_{i,n}\right]\right| \\
&\le C\varepsilon + \left|\mathbb{E}\big[P_{i,n}\int_{t_i}^T \left(Dg_\varepsilon \hat b + \frac{1}{2}\mathrm{Tr}\left[D^2 g_\varepsilon)\hat a\right]\right)(s, \hat X_s^{t_i,x}) ds\big]\right| \\
&\quad + \left|\mathbb{E}\big[P_{i,n}\int_{t_i}^T Dg_\varepsilon(\hat X_s^{t_i,x})\hat\sigma(s) dW_s\big]\right|,
\end{aligned} \tag{2.2.13}$$

where we denoted $\hat b(s) = b(t_j, \hat X_{t_j}^{t_i,x})$ and $\hat\sigma(s) = \sigma(t_j, \hat X_{t_j}^{t_i,x})$ for $t_j \le s < t_{j+1}$ and $\hat a = \hat\sigma^T\hat\sigma$. We next estimate each term separately.

2.a. First, since $\{P_{i,k}, i \le k \le n\}$ is a martingale:

$$\begin{aligned}
\left|\mathbb{E}\big[P_{i,n}\int_{t_i}^T Dg_\varepsilon(\hat X_s^{t_i,x})\hat\sigma(s) dW_s\big]\right| &= \left|\sum_{j=i}^{n-1}\mathbb{E}\big[P_{i,n}\int_{t_j}^{t_{j+1}} Dg_\varepsilon(\hat X_s^{t_i,x})\hat\sigma(s) dW_s\big]\right| \\
&\le \sum_{j=i}^{n-1}\left|\mathbb{E}\big[P_{i,j+1}\int_{t_j}^{t_{j+1}} Dg_\varepsilon(\hat X_s^{t_i,x})\hat\sigma(s) dW_s\big]\right| \\
&= \sum_{j=i}^{n-1}\left|\mathbb{E}\big[P_{i,j}\hat\sigma(t_j)\mathbb{E}_{t_j}\big[P_{j,j+1}\int_{t_j}^{t_{j+1}} Dg_\varepsilon(\hat X_s^{t_i,x}) dW_s\big]\big]\right|.
\end{aligned}$$

Notice that

$$\begin{aligned}
\mathbb{E}_{t_j}\left[P_{j,j+1}\int_{t_j}^{t_{j+1}} Dg_\varepsilon(\hat X_s^{t_i,x}) dW_s\right] &= \mathbb{E}_{t_j}\left[(W_{t_{j+1}} - W_{t_j})^2\int_{t_j}^{t_{j+1}} Dg_\varepsilon(\hat X_s^{t_i,x}) dW_s\right] \\
&= \mathbb{E}_{t_j}\left[\int_{t_j}^{t_{j+1}} 2W_s Dg_\varepsilon(\hat X_s^{t_i,x}) ds\right].
\end{aligned}$$

Using Lemma 2.1.1 and (2.2.12), this provides:

$$\left|\mathbb{E}\big[P_{i,n}\int_{t_i}^T Dg_\varepsilon(\hat X_s^{t_i,x})\hat\sigma(s) dW_s\big]\right| \tag{2.2.14}$$

$$\le 2\sum_{j=i}^{n-1}\left|\mathbb{E}\big[P_{i,j+1}\hat\sigma(t_j)^2\frac{\alpha_j}{h}\mathbb{E}_{t_j}\big[\int_{t_j}^{t_{j+1}} sD^2 g_\varepsilon(\hat X_s^{t_i,x}) ds\big]\big]\right|,$$

$$\le C\varepsilon^{-1}\sum_{j=i}^{n-1} h \le C'(T - t_i)\varepsilon^{-1}. \tag{2.2.15}$$

2.c. By (2.2.12) and the boundedness of b and σ, we also estimate that:

$$\left| Dg_\varepsilon(\hat{X}_s^{t_i,x})\hat{b}(s,\hat{X}_s^{t_i,x}) + \frac{1}{2}\mathrm{Tr}\left[D^2 g_\varepsilon(\hat{X}_s^{t_i,x})\hat{a}(s,\hat{X}_s^{t_i,x})\right]\right| \leq C + C\varepsilon^{-1} \quad (2.2.16)$$

2.b. Plugging (2.2.15) and (2.2.16) into (2.2.13), we obtain:

$$\left|\mathbb{E}\left[\left(g_\varepsilon(\hat{X}_T^{t_i,x}) - g_\varepsilon(x)\right) P_{i,n}\right]\right| \leq C(T-t_i) + C(T-t_i)\varepsilon^{-1},$$

which by (2.2.11) provides:

$$|v^h(t_i,x) - g(x)| \leq C\varepsilon + C(T-t_i)\varepsilon^{-1} + C(T-t_i).$$

The required result follows from the choice $\varepsilon = \sqrt{T-t_i}$. □

Corollary 2.2.1 *The function v^h is 1/2-Hölder continuous on t uniformly on h.*

Proof. The proof of $\frac{1}{2}$-Hölder continuity with respect to t could be easily provided by replacing g and $v^h(t_k,\cdot)$ in the assertion of Lemma respectively by $v^h(t,\cdot)$ and $v^h(t',\cdot)$ and consider the scheme from 0 to time t' with time step equal to h. Therefore, we can write;

$$|v^h(t,x) - v^h(t',x)| \leq C(t'-t)^{\frac{1}{2}},$$

where C could be chosen independent of t' for $t' \leq T$. □

2.2.3 Derivation of the rate of convergence

The proof of Theorem 2.2.2 is based on Barles and Jakobsen [5], which uses switching systems approximation and the Krylov method of shaking coefficients [43].

2.2.3.1 Comparison result for the scheme

Because F does not satisfy the standard monotonicity condition (2.2.7) of Barles and Souganidis [6], we need to introduce the nonlinearity $\overline{F}$ of Remark 2.2.7 so that $\overline{F}$ satisfies (2.2.6). Let u^h be the familiy of functions defined by

$$u^h(T,.) = g \text{ and } u^h(t_i,x) = \overline{\mathbf{T}}_h[u^h](t_i,x), \quad (2.2.17)$$

where for a function ψ from $[0,T]\times\mathbb{R}^d$ to $\mathbb{R}$ with exponential growth:

$$\overline{\mathbf{T}}_h[\psi](t,x) := \mathbb{E}\left[\psi(t+h,\hat{X}_h^{t,x})\right] + h\overline{F}\left(\cdot,\mathcal{D}_h\psi\right)(t,x),$$

and set

$$\overline{v}^h(t_i,x) := e^{-\theta(T-t_i)}u^h(t_i,x), \qquad i=0,\ldots,n. \quad (2.2.18)$$

The following result shows that the difference $v^h - \overline{v}^h$ is of higher order, and thus reduces the error estimate problem to the analysis of the difference $\overline{v}^h - v$.

Lemma 2.2.6 *Under Assumption* **F**, *we have*

$$\limsup_{h\searrow 0} h^{-1}|(v^h - \overline{v}^h)(t,.)|_\infty \quad < \quad \infty.$$

Proof. By definition of $\overline{F}$, we directy calculate that:

$$\overline{v}^h(t,x) \quad = \quad e^{-\theta h}(1+h\theta)\mathbb{E}[\overline{v}^h(t+h,\hat{X}_h^{t,x})] + hF\left(t+h,x,\mathcal{D}_h\overline{v}^h(t,x)\right).$$

Since $1+h\theta = e^{\theta h}+O(h^2)$, this shows that $\overline{v}^h(t,x) = \mathbf{T}_h[\overline{v}^h](t,x)+O(h^2)$. By lemma 2.2.3, we conclude that:

$$|(\overline{v}^h - v^h)(t,\cdot)|_\infty \quad \le \quad (1+Ch)|(\overline{v}^h - v^h)(t+h,\cdot)|_\infty + O(h^2),$$

which shows by the Gronwall inequality that $|(\overline{v}^h - v^h)(t,\cdot)|_\infty \le O(h)$ for all $t \le T-h$. □

By Remark 2.2.7, the operator $\overline{\mathbf{T}}_h$ satisfies the standard monotonicity condition (2.2.7):

$$\varphi \;\le\; \psi \quad\Longrightarrow\quad \overline{\mathbf{T}}_h[\varphi] \;\le\; \overline{\mathbf{T}}_h[\psi]. \tag{2.2.19}$$

The key-ingredient for the derivation of the error estimate is the following comparison result for the scheme.

Proposition 2.2.1 *Let Assumption* **F** *holds true, and set* $\beta := |F_r|_\infty$. *Consider two arbitrary bounded functions* φ *and* ψ *satisfying:*

$$h^{-1}\left(\varphi - \overline{\mathbf{T}}_h[\varphi]\right) \;\le\; g_1 \quad \textit{and} \quad h^{-1}\left(\psi - \overline{\mathbf{T}}_h[\psi]\right) \;\ge\; g_2 \tag{2.2.20}$$

for some bounded functions g_1 *and* g_2. *Then, for every* $i=0,\cdots,n$:

$$(\varphi-\psi)(t_i,x) \quad \le \quad e^{\beta(T-t_i)}|(\varphi-\psi)^+(T,\cdot)|_\infty + (T-h)e^{\beta(T-t_i)}|(g_1-g_2)^+|_\infty \tag{2.2.21}$$

To prove this comparison result, we need the following strengthening of the monotonicity condition:

Lemma 2.2.7 *Let Assumption* **F** *hold true and let* $\beta := |F_r|_\infty$. *Then, for every* $a,b\in\mathbb{R}_+$, *and every bounded functions* $\varphi\le\psi$, *the function* $\delta(t) := e^{\beta(T-t)}(a+b(T-t))$ *satisfies:*

$$\overline{\mathbf{T}}_h[\varphi+\delta](t,x) \quad \le \quad \overline{\mathbf{T}}_h[\psi](t,x)+\delta(t)-hb, \quad t\le T-h,\ x\in\mathbb{R}^d.$$

Proof. Because δ does not depend on x, we have $\mathcal{D}_h[\varphi+\delta] = \mathcal{D}_h\varphi+\delta(t+h)e_1$, where $e_1 := (1,0,0)$. Then, it follows from the regularity of $\overline{F}$ that there exist some ξ such that:

$$\overline{F}\big(t+h,x,\mathcal{D}_h[\varphi+\delta](t,x)\big) = \;\; \overline{F}\big(t+h,x,\mathcal{D}_h\varphi(t,x)\big)+\delta(t+h)\overline{F}_r\big(t+h,x,\xi e_1+\mathcal{D}_h\varphi(t,x)\big),$$

and

$$\begin{aligned}\overline{\mathbf{T}}_h[\varphi+\delta](t,x) &= \delta(t+h)+\mathbb{E}[\varphi(t+h,\hat{X}_h^{t,x})]+h\overline{F}\big(t+h,x,\mathcal{D}_h\varphi(t,x)\big)\\ &\quad +h\delta(t+h)\overline{F}_r\big(t+h,x,\xi e_1+\mathcal{D}_h\varphi(t,x)\big)\\ &= \overline{\mathbf{T}}_h[\varphi](t,x)+\delta(t+h)\left\{1+h\overline{F}_r\big(t+h,x,\xi e_1+\mathcal{D}_h\varphi(t,x)\big)\right\}\\ &\leq \overline{\mathbf{T}}_h[\varphi](t,x)+(1+\beta h)\,\delta(t+h).\end{aligned}$$

Since $\overline{\mathbf{T}}_h$ satisfies the standard monotonicity condition (2.2.19), this provides:

$$\overline{\mathbf{T}}_h[\varphi+\delta](t,x)\leq\overline{\mathbf{T}}_h[\psi](t,x)+\delta(t)+\zeta(t),\quad\text{where}\quad\zeta(t):=(1+\beta h)\,\delta(t+h)-\delta(t).$$

It remains to prove that $\zeta(t)\leq -hb$. From the smoothness of δ, we have $\delta(t+h)-\delta(t)=h\delta'(\bar{t})$ for some $\bar{t}\in[t,t+h)$. Then, since δ is decreasing in t, we see that

$$h^{-1}\zeta(t) \;=\; \delta'(\bar{t})+\beta\delta(t+h)\;\leq\;\delta'(\bar{t})+\beta\delta(\bar{t})\;\leq\;-be^{\beta(T-\bar{t})},$$

and the required estimate follows from the restriction $b\geq 0$. □

Proof of Proposition 2.2.1. We may refer directly to the similar result of [5]. However in our context, we give the following simpler proof. Observe that we may assume without loss of generality that

$$\varphi(T,\cdot)\leq\psi(T,\cdot)\quad\text{and}\quad g_1\leq g_2. \tag{2.2.22}$$

Indeed, one can otherwise consider the function

$$\bar{\psi}:=\psi+e^{\beta(T-t)}\,(a+b(T-t))\quad\text{where}\quad a=|(\varphi-\psi)^+(T,\cdot)|_\infty,\ b=|(g_1-g_2)^+|_\infty,$$

and β is the parameter defined in the previous Lemma 2.2.7, so that $\bar{\psi}(T,\cdot)\geq\varphi(T,\cdot)$ and, by Lemma (2.2.7), $\bar{\psi}(t,x)-\overline{\mathbf{T}}_h[\bar{\psi}](t,x)\geq h(g_1\vee g_2)$. Hence (2.2.22) holds true for φ and $\bar{\psi}$.

We now prove the required result by induction. First $\varphi(T,\cdot)\leq\psi(T,\cdot)$ by (2.2.22). We next assume that $\varphi(t+h,\cdot)\leq\psi(t+h,\cdot)$ for some $t+h\leq T$. Since $\overline{\mathbf{T}}_h$ satisfies the standard monotonicity condition (2.2.19), it follows from (2.2.22) that

$$\overline{\mathbf{T}}_h[\varphi](t,x)\quad\leq\quad\overline{\mathbf{T}}_h[\psi](t,x).$$

On the other hand, under (2.2.22), the hypothesis of the lemma implies:

$$\varphi(t,x)-\overline{\mathbf{T}}_h[\varphi](t,x)\quad\leq\quad\psi(t,x)-\overline{\mathbf{T}}_h[\psi](t,x).$$

Then $\varphi(t,\cdot)\leq\psi(t,\cdot)$. □

2.2.3.2 Proof of Theorem 2.2.2 (i)

Under the conditions of Assumption HJB on the coefficients, we may build a bounded subsolution $\underline{v}^\varepsilon$ of the nonlinear PDE, by the method of shaking the coefficients, which is Lipschitz in x, $1/2-$Hölder continuous in t, and approximates uniformly the solution v:

$$v-\varepsilon\leq\underline{v}^\varepsilon\leq v.$$

Let $\rho(t,x)$ be a C^∞ positive function supported in $\{(t,x) : t \in [0,1], |x| \le 1\}$ with unit mass, and define

$$\underline{w}^\varepsilon(t,x) := v^\varepsilon * \rho^\varepsilon \quad \text{where} \quad \rho^\varepsilon(t,x) := \frac{1}{\varepsilon^{d+2}}\rho\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) \tag{2.2.23}$$

so that, from the convexity of the operator F,

$$\underline{w}^\varepsilon \quad \text{is a subsolution of (2.1.1)}, \quad |\underline{w}^\varepsilon - v| \le 2\varepsilon. \tag{2.2.24}$$

Moreover, since v^ε is Lipschitz in x, and $1/2-$Hölder continuous in t,

$$\underline{w}^\varepsilon \text{ is } C^\infty, \text{ and } \left|\partial_t^{\beta_0} D^\beta \underline{w}^\varepsilon\right| \le C\varepsilon^{1-2\beta_0-|\beta|_1} \quad \text{for any} \quad (\beta_0, \beta) \in \mathbb{N} \times \mathbb{N}^d \setminus \{0\}, \tag{2.2.25}$$

where $|\beta|_1 := \sum_{i=1}^d \beta_i$, and $C > 0$ is some constant. As a consequence of the consistency result of Lemma 2.2.1 above, we know that

$$\mathcal{R}_h[\underline{w}^\varepsilon](t,x) \quad := \quad \frac{\underline{w}^\varepsilon(t,x) - \mathbf{T}_h[\underline{w}^\varepsilon](t,x)}{h} \; + \; \mathcal{L}^X \underline{w}^\varepsilon(t,x) + F(\cdot, \underline{w}^\varepsilon, D\underline{w}^\varepsilon, D^2\underline{w}^\varepsilon)(t,x)$$

converges to 0 as $h \to 0$. The next key-ingredient is to estimate the rate of convergence of $\mathcal{R}_h[\underline{w}^\varepsilon]$ to zero:

Lemma 2.2.8 *For a family $\{\varphi_\varepsilon\}_{0<\varepsilon<1}$ of smooth functions satisfying (3.3.21), we have:*

$$|\mathcal{R}_h[\varphi_\varepsilon]|_\infty \; \le \; R(h,\varepsilon) := C\, h\varepsilon^{-3} \quad \textit{for some constant} \quad C > 0.$$

The proof of this result is reported at the end of this section. From the previous estimate together with the subsolution property of w^ε, we see that $\underline{w}^\varepsilon \le \overline{\mathbf{T}}_h[\underline{w}^\varepsilon] + Ch^2\varepsilon^{-3}$. Then, it follows from Proposition 2.2.1 that

$$\underline{w}^\varepsilon - \overline{v}^h \le C|(\underline{w}^\varepsilon - \overline{v}^h)(T,.)|_\infty + Ch\varepsilon^{-3} \le C(\varepsilon + h\varepsilon^{-3}). \tag{2.2.26}$$

We now use (2.2.24) and (2.2.26) to conclude that

$$v - \overline{v}^h \le v - \underline{w}^\varepsilon + \underline{w}^\varepsilon - \overline{v}^h \le C(\varepsilon + h\varepsilon^{-3}).$$

Minimizing the right hand-side estimate over the choice of $\varepsilon > 0$, this implies the upper bound on the error $v - v^h$:

$$v - \overline{v}^h \quad \le \quad Ch^{1/4}. \tag{2.2.27}$$

2.2.3.3 Proof of Theorem 2.2.2 (ii)

The results of the previous section, together with the reinforced assumption HJB+, allow to apply the switching system method of Barles and Jakobsen [5] which provides the lower bound on the error:

$$v - \overline{v}^h \quad \ge \quad -\inf_{\varepsilon>0}\{C\varepsilon^{1/3} + R(h,\varepsilon)\} \;\; = \;\; -C'h^{1/10},$$

for some constants $C, C' > 0$. The required rate of convergence follows again from Lemma 2.2.6 which states that the difference $v^h - \overline{v}^h$ is dominated by the above rate of convergence.

Proof of Lemma 2.2.8 Notice that the evolution of the Euler approximation $\hat{X}_h^{t,x}$ between t and $t+h$ is driven by a constant drift $\mu(t,x)$ and a constant diffusion $\sigma(t,x)$. Since $D\varphi_\varepsilon$ is bounded, it follows from Itô's formula that:

$$\frac{1}{h}\left[\mathbb{E}\varphi_\varepsilon(t+h,\hat{X}_h^x)-\varphi_\varepsilon(t,x)\right]-\mathcal{L}^X\varphi_\varepsilon(t,x)=\frac{1}{h}\mathbb{E}\int_t^{t+h}\left(\mathcal{L}^{\hat{X}^{t,x}}\varphi_\varepsilon(u,\hat{X}_u^x)-\mathcal{L}^X\varphi_\varepsilon(t,x)\right)du,$$

where $\mathcal{L}^{\hat{X}^{t,x}}$ is the Dynkin operator associated to the Euler scheme:

$$\mathcal{L}^{\hat{X}^{t,x}}\varphi(t',x') \quad = \quad \partial_t\varphi(t',x')+\mu(t,x)D\varphi(t',x')+\frac{1}{2}\mathrm{Tr}\left[a(t,x)D^2\varphi(t',x')\right].$$

Applying again Itô's formula, and using the fact that $\mathcal{L}^{\hat{X}^{t,x}}D\varphi_\varepsilon$ is bounded, leads to

$$\frac{1}{h}\left[\mathbb{E}\varphi_\varepsilon(t+h,\hat{X}_h^x)-\varphi_\varepsilon(t,x)\right]-\mathcal{L}^X\varphi_\varepsilon(t,x) \quad = \quad \frac{1}{h}\mathbb{E}\int_t^{t+h}\int_t^u\mathcal{L}^{\hat{X}^{t,x}}\mathcal{L}^{\hat{X}^{t,x}}\varphi_\varepsilon(s,\hat{X}_s^x)dsdu.$$

Using the boundedness of the coefficients μ and σ, it follows from (3.3.21) that for $\varepsilon\in(0,1)$:

$$\left|\frac{\mathbb{E}\varphi_\varepsilon(t+h,\hat{X}_h^x)-\varphi_\varepsilon(t,x)}{h}-\mathcal{L}^X\varphi_\varepsilon(t,x)\right| \quad \le \quad R_0(h,\varepsilon):=C\,h\varepsilon^{-3}.$$

Step 2 This implies that

$$\begin{aligned}|\mathcal{R}_h[\varphi_\varepsilon](t,x)| \quad &\le \quad \left|\frac{\mathbb{E}\varphi_\varepsilon(t+h,\hat{X}_h^{t,x})-\varphi_\varepsilon(t,x)}{h}-\mathcal{L}^X\varphi_\varepsilon(t,x)\right| \\ &\quad +\left|F(x,\varphi_\varepsilon(t,x),D\varphi_\varepsilon(t,x),D^2\varphi_\varepsilon(t,x))-F(\cdot,\mathcal{D}_h[\varphi_\varepsilon](t,x))\right| \\ &\le \quad R_0(h,\varepsilon)+C\sum_{k=0}^{2}\left|\mathbb{E}D^k\varphi_\varepsilon(t+h,\hat{X}_h^{t,x})-D^k\varphi_\varepsilon(t,x)\right| \qquad (2.2.28)\end{aligned}$$

by the Lipschitz continuity of the nonlinearity F.

By a similar calculation as in Step 1, we see that:

$$|\mathbb{E}D^i\varphi_\varepsilon(t+h,\hat{X}_h^{t,x})-D\varphi_\varepsilon(t,x)| \quad \le \quad Ch\varepsilon^{-1-i},\ i=0,1,2,$$

which, together with (2.2.28), provides the required result. □

2.2.4 The rate of convergence in the linear case

In this subsection, we specialize the discussion to the linear one-dimensional case

$$F(\gamma) \;=\; c\gamma, \tag{2.2.29}$$

for some $c > 0$. The multi-dimensional case $d > 1$ can be handled similarly. Assuming that g is bounded, the linear PDE (2.1.1)-(2.1.2) has a unique bounded solution

$$v(t,x) \;=\; \mathbb{E}\left[g\left(x+\sqrt{1+2c}\,W_{T-t}\right)\right] \quad \text{for} \quad (t,x)\in[0,T]\times\mathbb{R}^d. \tag{2.2.30}$$

We also observe that this solution v is $C^\infty([0,T)\times\mathbb{R})$ with

$$D^k v(t,x) \;=\; \mathbb{E}\left[g^{(k)}\left(x+\sqrt{1+2c}\,W_{T-t}\right)\right], \quad t<T, \quad x\in\mathbb{R}. \tag{2.2.31}$$

This shows in particular that v has bounded derivatives of any order, whenever the terminal data g is C^∞ and has bounded derivatives of any order.

Of course, one can use the classical Monte Carlo estimate to produce an approximation of the function v of (2.2.30). The objective of this section is to analyze the error of the numerical scheme outlined in the previous sections. Namely:

$$v^h(T,\cdot)=g,\; v^h(t_{i-1},x)=\mathbb{E}\left[v^h(t_i,x+W_h)\right]+ch\mathbb{E}\left[v^h(t_i,x+W_h)H_2^h\right],\; i\le n. \tag{2.2.32}$$

Here, $\sigma=1$ and $\mu=0$ are used to write the above scheme.

Proposition 2.2.2 *Consider the linear F of (2.2.29), and assume that $D^{(2k+1)}v$ is bounded for every $k\ge 0$. Then*

$$\limsup_{h\to 0} h^{-1/2}|v^h-v|_\infty \;<\; \infty.$$

Proof. Since v has bounded first derivative with respect to x, it follows from Itô's formula that:

$$v(t,x) \;=\; \mathbb{E}\left[v(t+h,x+W_h)\right]+c\mathbb{E}\left[\int_0^h \triangle v(t+s,v+W_s)ds\right],$$

Then, in view of Lemma 2.1.1, the error $u:=v-v^h$ satisfies $u(t_n,X_{t_n})=0$ and for $i\le n-1$:

$$\begin{aligned} u\left(t_i,X_{t_i}\right) \;=\;& \mathbb{E}_i\left[u\left(t_{i+1},X_{t_{i+1}}\right)\right]+ch\,\mathbb{E}_i\left[\triangle u\left(t_{i+1},X_{t_{i+1}}\right)\right] \\ &+c\mathbb{E}_i\int_0^h\left[\triangle v\left(ih+s,X_{ih+s}\right)-\triangle v\left((i+1)h,X_{(i+1)h}\right)\right]ds, \end{aligned} \tag{2.2.33}$$

where $\mathbb{E}_i:=\mathbb{E}[\cdot|\mathcal{F}_{t_i}]$ is the expectation operator conditional on $\mathcal{F}_{t_i}$.

Step 1 Set

$$a_i^k:=\mathbb{E}\left[\triangle^k u\left(t_i,X_{t_i}\right)\right], \qquad b_i^k:=\mathbb{E}\int_0^h\left[\triangle^k v\left(t_{i-1}+s,X_{t_{i-1}+s}\right)-\triangle^k v\left(t_i,X_{t_i}\right)\right]ds,$$

and we introduce the matrices

$$A := \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \qquad B := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix},$$

and we observe that (2.2.33) implies that the vectors $a^k := (a_1^k, \dots, a_n^k)^{\mathrm{T}}$ and $b^k := (b_1^k, \dots, b_n^k)^{\mathrm{T}}$ satisfy $Aa^k = chBa^{k+1} + cBb^k$ for all $k \geq 0$, and therefore:

$$a^k \;=\; chA^{-1}Ba^{k+1} + cA^{-1}Bb^k \quad \text{where} \quad A^{-1} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \tag{2.2.34}$$

By direct calculation, we see that the powers $(A^{-1}B)^k$ are given by:

$$(A^{-1}B)^k_{i,j} = \mathbb{1}_{\{j \geq i+k\}} \binom{j-i-1}{k-1} \quad \text{for all} \quad k \geq 1 \text{ and } i, j = 1, \dots, n.$$

In particular, because $a_n^k = 0$, $(A^{-1}B)^{n-1}a^k = 0$. Iterating (2.2.34), this provides:

$$a^0 \;=\; ch(A^{-1}B)a^1 + c(A^{-1}B)b^0 \;=\; \dots \;=\; \sum_{k=0}^{n-2} c^{k+1}h^k(A^{-1}B)^{k+1}b^k,$$

and therefore:

$$u(0,x) \;=\; a_1^0 \;=\; c\sum_{k=0}^{n-2}(ch)^k(A^{-1}B)^{k+1}_{1,j}b^k. \tag{2.2.35}$$

Because of

$$(A^{-1}B)^k_{1,j} = \mathbb{1}_{\{j \geq 1+k\}} \binom{j-2}{k-1} \quad \text{for all} \quad k \geq 1 \text{ and } j = 1, \dots, n\,,$$

we can write (2.2.35):

$$u(0,x) \;=\; c\sum_{k=0}^{n-2}(ch)^k \sum_{j=k+2}^{n} \binom{j-2}{k} b_j^{k-1}.$$

By changing the order of the summations in the above we conclude that:

$$u(0,x) \;=\; c\sum_{j=2}^{n}\sum_{k=0}^{j-2}(ch)^k \binom{j-2}{k} b_j^{k-1}. \tag{2.2.36}$$

Step 2 From our assumption that $D^{2k+1}v$ is $\mathbb{L}^\infty-$bounded for every $k \geq 0$, it follows that

$$|b_j^k| \;\leq\; \mathbb{E}\left[\int_{t_{i-1}}^{t_i} \left|\triangle^k v(s,X_s) - \triangle^k v(t_j,X_{t_j})\right| ds\right] \;\leq\; Ch^{3/2}$$

for some constant C. We then deduce from (2.2.36) that:

$$|u(0,x)| \;\leq\; cCh^{3/2}\sum_{j=2}^{n}\sum_{k=0}^{j-2}(ch)^k\binom{j-2}{k}.$$

So,

$$|u(0,x)| \;\leq\; cCh^{3/2}\sum_{j=2}^{n}(1+ch)^{j-2} \;=\; cCh^{3/2}\,\frac{(1+ch)^{n-1}-1}{ch} \;\leq\; C\sqrt{h}.$$

2.3 Probabilistic Numerical Scheme

In order to implement the backward scheme (2.1.4), we still need to discuss the numerical computation of the conditional expectations involved in the definition of the operators $\mathbf{T}_h$ in (3.1.8). In view of the Markov feature of the process X, these conditional expectations reduce to simple regressions. Motivated by the problem of American options in financial mathematics, various methods have been introduced in the literature for the numerical approximation of these regressions. We refer to [18] and [35] for a detailed discussion.

The chief object of this section is to investigate the asymptotic properties of our suggested numerical method when the expectation operator $\mathbb{E}$ in (2.1.4) is replaced by some estimator $\hat{\mathbb{E}}^N$ corresponding to a sample size N:

$$\tilde{\mathbf{T}}_h^N[\psi](t,x) \;:=\; \hat{\mathbb{E}}^N\left[\psi(t+h,\hat{X}_h^x)\right] + hF\left(\cdot,\hat{\mathcal{D}}_h\psi\right)(t,x), \tag{2.3.1}$$

$$\hat{\mathbf{T}}_h^N[\psi](t,x) \;:=\; -K_h[\psi] \vee \tilde{\mathbf{T}}_h^N[\psi](t,x) \wedge K_h[\psi] \tag{2.3.2}$$

where

$$\hat{\mathcal{D}}_h\psi(t,x) := \hat{\mathbb{E}}^N\left[\psi(t+h,\hat{X}_h^{t,x})H_h(t,x)\right],\;\; K_h[\psi] := \|\psi\|_\infty(1+C_1h)+C_2h,$$

where

$$C_1 = \frac{1}{4}|F_p^{\mathrm{T}}F_\gamma^- F_p|_\infty + |F_r|_\infty \quad\text{and}\quad C_2 = |F(t,x,0,0,0)|_\infty.$$

The above bounds are needed for technical reasons which were already observed in [18].

With these notations, the implementable numerical scheme is:

$$\hat{v}_N^h(t,x,\omega) = \hat{\mathbf{T}}_h^N[\hat{v}_N^h](t,x,\omega), \tag{2.3.3}$$

where $\hat{\mathbf{T}}_h^N$ is defined in (2.3.1)-(2.3.2), and the presence of ω throughout this section emphasizes the dependence of our estimator on the underlying sample.

Let $\mathcal{R}_b$ be the family of random variables R of the form $\psi(W_h)H_i(W_h)$ where ψ is a function with $|\psi|_\infty \le b$ and H_i's are the Hermite polynomials:

$$H_0(x) = 1, \; H_1(x) = x \text{ and } H_2(x) = x^T x - h \;\; \forall x \in R^d.$$

Assumption E *There exist constants* $C_b, \lambda, \nu > 0$ *such that* $\left\|\hat{\mathbb{E}}^N[R] - \mathbb{E}[R]\right\|_p \le C_b h^{-\lambda} N^{-\nu}$ *for every* $R \in \mathcal{R}_b$*, for some* $p \ge 1$.

Example 2.3.1 Consider the regression approximation based on the Malliavin integration by parts as introduced in Lions and Reigner [46], Bouchard, Ekeland and Touzi [15], and analyzed in the context of the simulation of backward stochastic differential equations by [18] and [28]. Then Assumption **E** is satisfied for every $p > 1$ with the constants $\lambda = \frac{d}{4p}$ and $\nu = \frac{1}{2p}$, see [18].

Our next main result establishes conditions on the sample size N and the time step h which guarantee the convergence of $\hat{v}_N^h$ towards v.

Theorem 2.3.1 *Let Assumptions* **E** *and* **F** *hold true, and assume that the fully nonlinear PDE* (2.1.1) *has comparison with growth* q*. Suppose in addition that*

$$\lim_{h\to 0} h^{\lambda+2} N_h^{\nu} \;\; = \;\; \infty. \tag{2.3.4}$$

Assume that the final condition g *is bounded Lipschitz, and the coefficients* μ *and* σ *are bounded. Then, for almost every* ω*:*

$$\hat{v}_{N_h}^h(\cdot,\omega) \longrightarrow v \qquad \textit{locally uniformly,}$$

where v *is the unique viscosity solution of* (2.1.1)*.*

Proof. We adapt the argument of [6] to the present stochastic context. By Remark 2.2.7 and Lemma 2.2.6, we may assume without loss of generality that the strict monotonicity (2.2.6) holds.

By (2.3.2), we see that $\hat{v}^h$ is uniformly bounded. So, we can define:

$$\hat{v}_*(t,x) := \liminf_{\substack{(t',x')\to(t,x)\\ h\to 0}} \hat{v}^h(t',x') \;\text{ and }\; \hat{v}^*(t,x) := \limsup_{\substack{(t',x')\to(t,x)\\ h\to 0}} \hat{v}^h(t',x') \tag{2.3.5}$$

Our objective is to prove that $\hat{v}_*$ and $\hat{v}^*$ are respectively viscosity superpersolution and subsolution of (2.1.1). By the comparison assumption, we shall then conclude that they are both equal to the unique viscosity solution of the problem whose existence is given by Theorem 2.2.1. In particular, they are both deterministic functions.

We shall only report the proof of the supersolution property, the subsolution property follows from the same type of argument.

In order to prove that $\hat{v}_*$ is a supersolution of (2.1.1), we consider $(t_0, x_0) \in [0,T) \times \mathbb{R}^n$ together with a test function $\varphi \in C^2([0,T) \times \mathbb{R}^n)$, so that

$$0 = \min\{\hat{v}_* - \varphi\} = (\hat{v}_* - \varphi)(t_0, x_0).$$

By classical manipulations, we can find a sequence $(t_n, x_n, h_n) \to (t_0, x_0, 0)$ so that $\hat{v}^{h_n}(t_n, x_n) \to \hat{v}_*(t_0, x_0)$ and

$$(\hat{v}^{h_n} - \varphi)(t_n, x_n) = \min\{\hat{v}^{h_n} - \varphi\} =: C_n \to 0.$$

Then, $\hat{v}^{h_n} \geq \varphi + C_n$, and it follows from the monotonicity of the operator $\mathbf{T}_h$ that:

$$\mathbf{T}_{h_n}[\hat{v}^{h_n}] \geq \mathbf{T}_{h_n}[\varphi + C_n].$$

By the definition of $\hat{v}^{h_n}$ in (2.3.3), this provides:

$$\hat{v}^{h_n}(t,x) \geq \mathbf{T}_{h_n}[\varphi + C_n](t,x) - (\mathbf{T}_{h_n} - \hat{\mathbf{T}}_{h_n})[\hat{v}_n^h](t,x),$$

where, for ease of notations, the dependence on N_h has been dropped. Because $\hat{v}^{h_n}(t_n, x_n) = \varphi(t_n, x_n) + C_n$, the last inequality gives:

$$\varphi(t_n, x_n) + C_n - \mathbf{T}_{h_n}[\varphi + C_n](t_n, x_n) + h_n R_n \geq 0, \qquad R_n := h_n^{-1}(\mathbf{T}_{h_n} - \hat{\mathbf{T}}_{h_n})[\hat{v}^{h_n}](t_n, x_n).$$

We claim that

$$R_n \longrightarrow 0 \qquad \mathbb{P} - \text{a.s. along some subsequence.} \tag{2.3.6}$$

Then, after passing to the subsequence, dividing both sides by h_n, and sending $n \to \infty$, it follows from Lemma 2.2.1 that:

$$-\mathcal{L}^X \varphi - F\left(\cdot, \varphi, D\varphi, D^2\varphi\right) \geq 0,$$

which is the required supersolution property.

It remains to show (2.3.6). We start by bounding R_n with respect to the error of estimation of conditional expectation. By Lemma 2.2.3, $|\mathbf{T}_{h_n}[\hat{v}^{h_n}]|_\infty \leq K_{h_n}$ and so by (2.3.2), we can write:

$$\left|\left(\mathbf{T}_{h_n} - \hat{\mathbf{T}}_{h_n}\right)[\hat{v}^{h_n}](t_n, x_n)\right| \;\; \leq \;\; \left|\left(\mathbf{T}_{h_n} - \tilde{\mathbf{T}}_{h_n}\right)[\hat{v}^{h_n}](t_n, x_n)\right|. \tag{2.3.7}$$

By the Lipschitz-continuity of F, we have:

$$\left|\left(\mathbf{T}_{h_n} - \hat{\mathbf{T}}_{h_n}\right)[\hat{v}^{h_n}](t_n, x_n)\right| \;\; \leq \;\; C\left(\mathcal{E}_0 + h_n\mathcal{E}_1 + h_n\mathcal{E}_2\right).$$

where:

$$\mathcal{E}_i \;\; = \;\; |(\mathbb{E} - \hat{\mathbb{E}})[\hat{v}^{h_n}(t_n + h_n, X_{h_n}^{x_n})H_i^{h_n}(t_n, x_n)]|$$

$$\left|\left(\mathbf{T}_{h_n} - \hat{\mathbf{T}}_{h_n}\right)[\hat{v}^{h_n}](t_n, x_n)\right| \;\; \leq \;\; C\left(\left|(\mathbb{E} - \hat{\mathbb{E}})[R_n^0]\right| + \left|(\mathbb{E} - \hat{\mathbb{E}})[R_n^1]\right| + h_n^{-1}\left|(\mathbb{E} - \hat{\mathbb{E}})[R_n^2]\right|\right).$$

where $R_n^i = \hat{v}^{h_n}\big(t_n + h_n, x_n + \sigma(x)W_h\big)H_i(W_h)$, $i = 1, 2, 3$ and H_i is Hermite polynomial of degree i. This leads the following estimate for the error R_n:

$$|R_n| \;\le\; \frac{C}{h_n}\left(\left|(\mathbb{E}-\hat{\mathbb{E}})[R_n^0]\right| + \left|(\mathbb{E}-\hat{\mathbb{E}})[R_n^1]\right| + h_n^{-1}\left|(\mathbb{E}-\hat{\mathbb{E}})[R_n^2]\right|\right). \quad (2.3.8)$$

Because $R_n^i \in \mathcal{R}_b$ with bound obtained in Lemma 2.2.3 by Assumption **E** we have,:

$$\|R_n\|_p \;\le\; Ch_n^{-\lambda-2}N_{h_n}^{-\nu},$$

so by (2.3.4) we have $\|R_n\|_p \longrightarrow 0$ which implies (2.3.6). □

We finally discuss the choice of the sample size so as to keep the same rate for the error bound.

Theorem 2.3.2 *Let the nonlinearity F be as in Assumption* **HJB***, and consider a regression operator satisfying Assumption* **E***. Let the sample size N_h be such that*

$$\lim_{h\to 0} h^{\lambda+\frac{21}{10}}N_h^{\nu} \;>\; 0. \quad (2.3.9)$$

Then, for any bounded Lipschitz final condition g, we have the following $\mathbb{L}^p-$bounds on the rate of convergence:

$$\|v-\hat{v}^h\|_p \;\le\; Ch^{1/10}.$$

Proof. By Remark 2.2.7 and Lemma 2.2.6, we may assume without loss of generality that the strict monotonicity (2.2.6) holds true.

We proceed as in the proof of Theorem 2.2.2 to see that

$$v-\hat{v}^h \;\le\; v-v^h+v_h-\hat{v}^h \;=\; \varepsilon+R(h,\varepsilon)+v_h-\hat{v}^h.$$

Since $\hat{v}^h$ satisfies (2.3.3),

$$h^{-1}\left(\hat{v}^h-\mathbf{T}_h[\hat{v}^h]\right) \;\ge\; -R_h[\hat{v}^h] \quad \text{where} \quad R_h[\varphi] \;:=\; \frac{1}{h}\left|\left(\mathbf{T}_h-\hat{\mathbf{T}}_h\right)[\varphi]\right|,$$

where, in the present context, $R_h[\hat{v}^h]$ is a non-zero stochastic term. By Proposition 2.2.1, it follows from the last inequality that:

$$v-\hat{v}^h \le C\left(\varepsilon+R(h,\varepsilon)+R_h[\hat{v}^h]\right),$$

where the constant $C>0$ depends only on the Lipschitz coefficient of F, β in Lemma 2.2.7 and the constant in Lemma 2.2.8.

Similarly, we follow the line of argument of the proof of Theorem 2.2.2 to show that a lower bound holds true, and therefore:

$$|v-\hat{v}^h| \le C\left(\varepsilon^{1/3}+R(h,\varepsilon)+R_h[\hat{v}^h]\right),$$

We now use (2.3.9) and proceed as in the last part of the proof of Theorem 2.3.1 to deduce from (2.3.8) and Assumption **F** that

$$\|R_h[\hat{v}^h]\|_p \quad \leq \quad C h^{1/10}.$$

With this choice of the sample size N, the above error estimate reduces to

$$\|\hat{v}^h - v\|_p \quad \leq \quad C\left(\varepsilon^{1/3} + R(h,\varepsilon) + h^{1/10}\right),$$

and the additional term $h^{1/10}$ does not affect the minimization with respect to ε.
□

Example 2.3.2 Let us illustrate the convergence results of this section in the context of the Malliavin integration by parts regression method of [46] and [18] where $\lambda = \frac{d}{4p}$ and $\nu = \frac{1}{2p}$ for every $p > 1$. So, for the convergence result we need to choose N_h of the order of $h^{-\alpha_0}$ with $\alpha_0 > \frac{d}{2} + 4p$. For the L^p-rate of convergence result, we need to choose N_h of the order of $h^{-\alpha_1}$ with $\alpha_1 \geq \frac{d}{2} + \frac{21p}{5}$.

2.4 Numerical Results

In this section, we provide an application of the Monte Carlo-finite differences scheme suggested in this paper in the context of two different types of problems. We first consider the classical mean curvature flow equation as the simplest front propagation example. We test our backward probabilistic scheme on the example where the initial data is given by a sphere, for which an easy explicit solution is available. A more interesting geometric example in space dimensions 2 is also considered. We next consider the Hamilton-Jacobi-Bellman equation characterizing the classical optimal investment problem in financial mathematics. Here, we again test our scheme in dimension two where an explicit solution is available, and we consider more involved examples in space dimension 5, in addition to the time variable.

In all examples considered in this section the operator $F(t, x, r, p, \gamma)$ does not depend on the $r-$variable. We shall then drop this variable from our notations, and we simply write the scheme as:

$$\begin{aligned} &v^h(T,.) := g \quad \text{and} \\ &v^h(t_i,x) := \mathbb{E}[v^h(t_{i+1},\hat{X}_h^x)] + hF\left(t_i,x,\mathcal{D}_h v^h(t_i,x)\right) \end{aligned} \tag{2.4.1}$$

where

$$\mathcal{D}_h\psi := \left(\mathcal{D}_h^1\psi, \mathcal{D}_h^2\psi\right),$$

and $\mathcal{D}_h^1$ and $\mathcal{D}_h^2$ are defined in Lemma 2.1.1. We recall from Remark 2.1.1 that:

$$\begin{aligned} \mathcal{D}_{2h}^2\varphi(t_i,x) &= \mathbb{E}\left[\varphi(t_i+2h,\hat{X}_{2h}^{t_i,x})\left(\sigma^{\mathrm{T}}\right)^{-1}\frac{(W_{t_i+h}-W_{t_i})(W_{t_i+h}-W_{t_i})^{\mathrm{T}}-h\mathbf{I}_d}{h^2}\,\sigma^{-1}\right] \\ &= \mathbb{E}\left[\mathcal{D}_h^1\varphi(t_i+h,\hat{X}_h^{t_i,x})\left(\sigma^{\mathrm{T}}\right)^{-1}\frac{W_{t_i+h}-W_{t_i}}{h}\right] \end{aligned} \tag{2.4.2}$$

The second representation is the one reported in [25] where the present backward probabilistic scheme was first introduced. These two representations induce two different numerical schemes because once the expectation operator $\mathbb{E}$ is replaced by an approximation $\hat{E}^N$, equality does not hold anymore in the latter equation for finite N. In our numerical examples below, we provide results for both methods. The numerical schemes based on the first (resp. second) representation will be referred to as *scheme* 1 (resp. 2). An important outcome of our numerical experiments is that scheme 2 turns out to have a significantly better performance than scheme 1.

Remark 2.4.1 The second scheme needs some final condition for $\mathcal{D}_h^1\varphi(T, X_h^{T-h,x})$. Since g is smooth in all our examples, we set this final condition to ∇g. Since the second scheme turns out to have a better performnace, we may also use the final condition for Z suggested by the first scheme.

We finally discuss the choice of the regression estimator in our implemented examples. Two methods have been used:

- The first method is the basis projection *a la* Longstaff and Schwartz [47], as developed in [35]. We use regression functions with localized support : on each support the regression functions are chosen linear and the size of the support is adaptative according to the Monte Carlo distribution of the underlying process.

- The second method is based on the Malliavin integration by parts formula as suggested in [46] and further developed in [15]. In particular, the optimal exponential localization function $\phi^k(y) = exp(-\eta^k y)$ in each direction k is chosen as follows. The optimal parameter η_k is provided in [15] and should be chosen for each conditional expectation depending on k. Our numerical experiments however revealed that such optimal parameters do not provide sufficiently good performance, and more accurate results are obtained by choosing $\eta_k = 5/\sqrt{\Delta t}$ for all values of k.

2.4.1 Mean curvature flow problem

The mean curvature flow equation describes the motion of a surface where each point moves along the inward normal direction with speed proportional to the mean curvature at that point. This geometric problem can be characterized as the zero-level set $S(t) := \{x \in \mathbb{R}^d : v(t,x) = 0\}$ of a function $v(t,x)$ depending on time and space satisfying the geometric partial differential equation:

$$v_t - \Delta v + \frac{Dv \cdot D^2 v Dv}{|Dv|^2} = 0 \quad \text{and} \quad v(0,x) = g(x) \tag{2.4.3}$$

and $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ is a bounded Lipschitz-continuous function. We refer to [55] for more details on the mean curvature problem and the corresponding stochastic representation.

To model the motion of a sphere in $\mathbb{R}^d$ with radius $2R > 0$, we take $g(x) := 4R^2 - |x|^2$ so that g is positive inside the sphere and negative outside. We first solve the sphere problem in dimension 3. In this case, it is well-known that the surface $S(t)$ is a sphere with a radius $R(t) = 2\sqrt{R^2 - t}$ for $t \in (0, R^2)$. Reversing time, we rewrite (2.4.3) for $t \in (0, T)$ with $T = R^2$:

$$- v_t - \frac{1}{2}\sigma^2 \Delta v + F(x, Dv, D^2 v) = 0 \quad \text{and} \quad v(T, x) = g(x), \qquad (2.4.4)$$

where

$$F(x, z, \gamma) \quad := \quad \gamma \left(\frac{1}{2}\sigma^2 - 1 \right) + \frac{z \cdot \gamma z}{|z|^2}.$$

We implement our Monte Carlo-finite differences scheme to provide an approximation $\hat{v}^h$ of the function v. As mentioned before, we implement four methods: Malliavin integration by parts-based or basis projection-based regression, and scheme 1 or 2 for the representation of the Hessian.

Given the approximation $\hat{v}^h$, we deduce an approximation of the surface $\hat{S}^h(t) := \{x \in \mathbb{R}^3 : \hat{v}^h(t, x) = 0)\}$ by using a dichotomic gradient descent method using the estimation of the gradient $\mathcal{D}^1 v$ estimated along the resolution. The dichotomy is stopped when the solution is localized within 0.01 accuracy.

Remark 2.4.2 Of course the use of the gradient is not necessary in the present context where we know that $S(t)$ is a sphere at any time $t \in [0, T)$. The algorithm described above is designed to handle any type of geometry.

Remark 2.4.3 In our numerical experiments, the nonlinearity F is truncated so that it is bounded by an arbitrary value taken equal to 200.

Our numerical results show that Malliavin and basis projection methods give similar results. However, for a given number of sample paths, the basis projection method of [35] are slightly more accurate. Therefore, all results reported for this example correspond to the basis projection method.

Figure 2.1 provides results obtained with one million particles and $10 \times 10 \times 10$ mesh with a time step equal to 0.0125. The diffusion coefficient σ is taken to be either 1 or 1.8. We observe that results are better with $\sigma = 1$. We also observe that the error increases near time 0.25 corresponding to an acceleration of the dynamics of the phenomenon, and suggesting that a thinner time step should be used at the end of simulation.

Figure 2.2 plots the difference between our calculation and the reference for scheme 1 and volatility 1 and 1.8 for varying time step. The corresponding results with scheme 2 are reported in figure 2.3. We notice that some points at time $T = 0.25$ are missing due to a non convergence of the gradient method for a diffusion $\sigma = 1.8$. We observe that results for scheme 2 are slightly better than results for scheme 1. With $\sigma = 1$, it takes 150 seconds on a Nehalem intel processor 2.9 GHz to obtain the result at time $t = 0.15$ with the regression method, while it takes 1500 seconds with

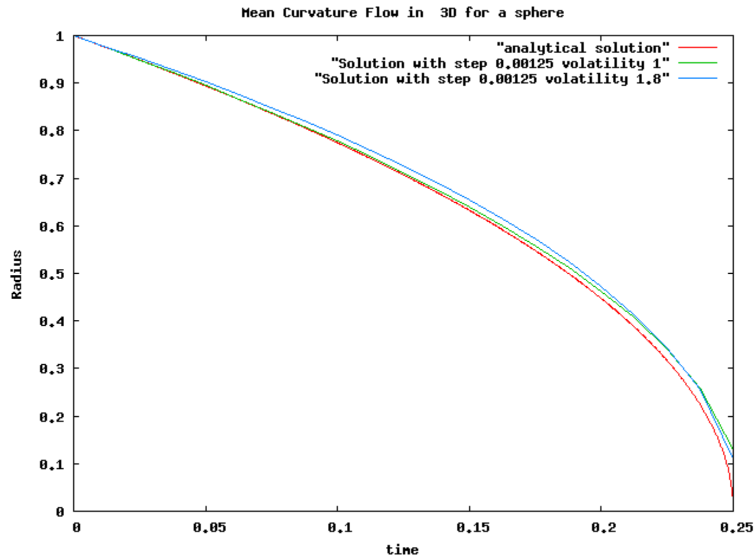

Figure 2.1: Solution of the mean curvature flow for the sphere problem

the Malliavin method (notice that the dichotomy used with the gradient method is a very inefficient method).

We finally report in Figure 2.4 some numerical results for the mean curvature flow problem in dimension 2 with a more interesting geometry: the initial surface (i.e. the zero-level set for v) consists of two disks with unit radius, with centers positioned at -1.5 and 1.5 and connected by a stripe of unit width. We give the resulting deformation with scheme 2 for a diffusion $\sigma = 1$, a time step $h = 0.0125$, and one million particles. Once again, the Malliavin integration by parts based regression method and the basis projection method with 10×10 meshes produce similar results. We used 1024 points to describe the surface.

One advantage of this method is the total parallelization that can be performed to solve the problem for different points on the surface : for the results given parallelization by Message Passing (MPI) was achieved.

2.4.2 Continuous-time portfolio optimization

We next report an application to the continuous-time portfolio optimization problem in financial mathematics. Let $\{S_t, t \in [0,T]\}$ be an Itô process modeling the price evolution of n financial securities. The investor chooses an adapted process $\{\theta_t, t \in [0,T]\}$ with values in $\mathbb{R}^n$, where θ_t^i is the amount invested in the $i-$th security held at time t. In addition, the investor has access to a non-risky security (bank account) where the remaining part of his wealth is invested. The non-risky asset S^0 is defined by an adapted interest rates process $\{r_t, t \in [0,T]\}$, i.e. $dS_t^0 = S_t^0 r_t dt$, $t \in [0,1]$. Then, the dynamics of the wealth process is described by:

$$dX_t^\theta = \theta_t \cdot \frac{dS_t}{S_t} + (X_t^\theta - \theta_t \cdot \mathbb{1})\frac{dS_t^0}{S_t^0} = \theta_t \cdot \frac{dS_t}{S_t} + (X_t^\theta - \theta_t \cdot \mathbb{1}) r_t dt,$$

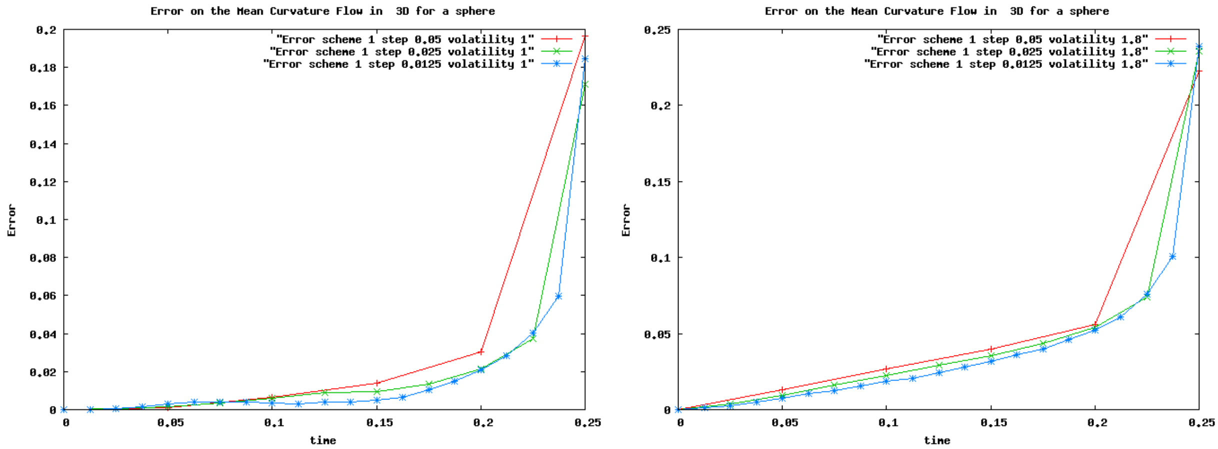

Figure 2.2: Mean curvature flow problem for different time step and diffusion: scheme 1

where $\mathbb{1} = (1, \cdots, 1) \in \mathbb{R}^d$. Let $\mathcal{A}$ be the collection of all adapted processes θ with values in $\mathbb{R}^d$, which are integrable with respect to S and such that the process X^θ is uniformly bounded from below. Given an absolute risk aversion coefficient $\eta > 0$, the portfolio optimization problem is defined by:

$$v_0 \quad := \quad \sup_{\theta \in \mathcal{A}} \mathbb{E}\left[-\exp\left(-\eta X_T^\theta\right)\right]. \tag{2.4.5}$$

Under fairly general conditions, this linear stochastic control problem can be characterized as the unique viscosity solution of the corresponding HJB equation. The main purpose of this subsection is to implement our Monte Carlo-finite differences scheme to derive an approximation of the solution of the fully nonlinear HJB equation in non-trivial situations where the state has a few dimensions. We shall first start by a two-dimensional example where an explicit solution of the problem is available. Then, we will present some results in a five dimensional situation.

2.4.2.1 A two dimensional problem

Let $d = 1$, $r_t = 0$ for all $t \in [0, 1]$, and assume that the security price process is defined by the Heston model [36]:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{Y_t} S_t dW_t^{(1)} \\ dY_t &= k(m - Y_t)dt + c\sqrt{Y_t}\left(\rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)}\right), \end{aligned}$$

where $W = (W^{(1)}, W^{(2)})$ is a Brownian motion in $\mathbb{R}^2$. In this context, it is easily seen that the portfolio optimization problem (2.4.5) does not depend on the state variable s. Given an initial state at the time origin t given by $(X_t, Y_t) = (x, y)$, the

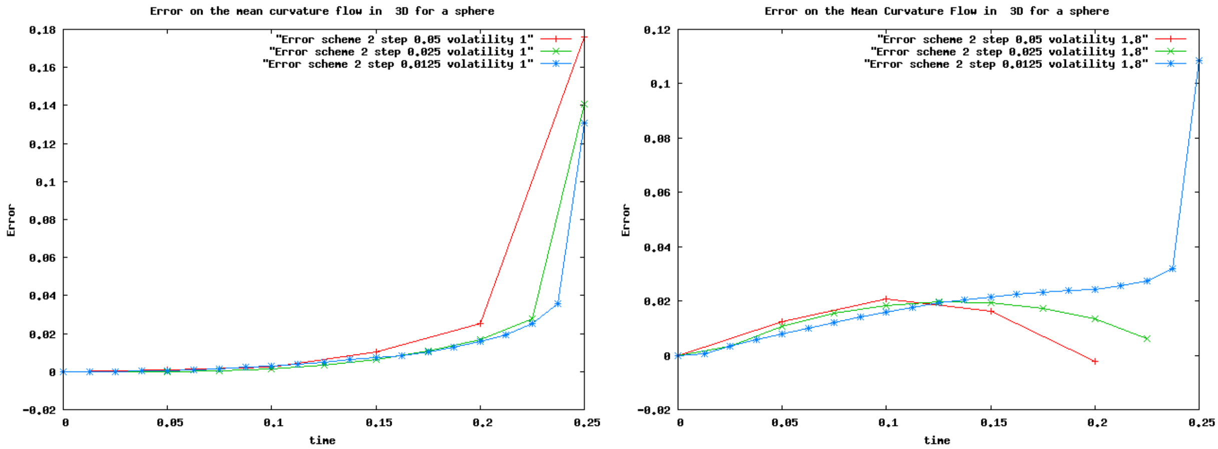

Figure 2.3: Mean curvature flow problem for different time step and diffusions: scheme 2

value function $v(t,x,y)$ solves the HJB equation:

$$v(T,x,y) = -e^{-\eta x} \text{ and } 0 = -v_t - k(m-y)v_y - \tfrac{1}{2}c^2 y v_{yy} - \sup_{\theta\in\mathbb{R}}\left(\frac{1}{2}\theta^2 y v_{xx} + \theta(\mu v_x + \rho c y v_{xy})\right)$$
$$= -v_t - k(m-y)v_y - \tfrac{1}{2}c^2 y v_{yy} + \frac{(\mu v_x + \rho c y v_{xy})^2}{2 y v_{xx}}. \tag{2.4.6}$$

A quasi explicit solution of this problem was provided by Zariphopoulou [61]:

$$v(t,x,y) = -e^{-\eta x}\left\|\exp\left(-\frac{1}{2}\int_t^T \frac{\mu^2}{\tilde{Y}_s}ds\right)\right\|_{\mathbb{L}^{1-\rho^2}} \tag{2.4.7}$$

where the process $\tilde{Y}$ is defined by

$$\tilde{Y}_t = y \quad \text{and} \quad d\tilde{Y}_t = (k(m-\tilde{Y}_t) - \mu c\rho)dt + c\sqrt{\tilde{Y}_t}dW_t.$$

In order to implement our Monte Carlo-finite differences scheme, we re-write (2.4.6) as:

$$-v_t - k(m-y)v_y - \frac{1}{2}c^2 y v_{yy} - \frac{1}{2}\sigma^2 v_{xx} + F\left(y, Dv, D^2 v\right) = 0, \; v(T,x,y) = -e^{-\eta x}, \tag{2.4.8}$$

where $\sigma > 0$ and the nonlinearity $F : \mathbb{R}\times\mathbb{R}^2\times\mathbb{S}_2$ is given by:

$$F(y,z,\gamma) \;=\; \frac{1}{2}\sigma^2\gamma_{11} + \frac{(\mu z_1 + \rho c y \gamma_{12})^2}{2y\gamma_{11}}.$$

Notice that the nonlinearity F does not to satisfy Assumption **F**, we consider the truncated nonlinearity:

$$F_{\varepsilon,M}(y,z,\gamma) \;:=\; \frac{1}{2}\sigma^2\gamma_{11} - \sup_{\varepsilon\le\theta\le M}\left(\frac{1}{2}\theta^2(y\vee\varepsilon)\gamma_{11} + \theta(\mu z_1 + \rho c(y\vee\varepsilon)\gamma_{12}\right),$$

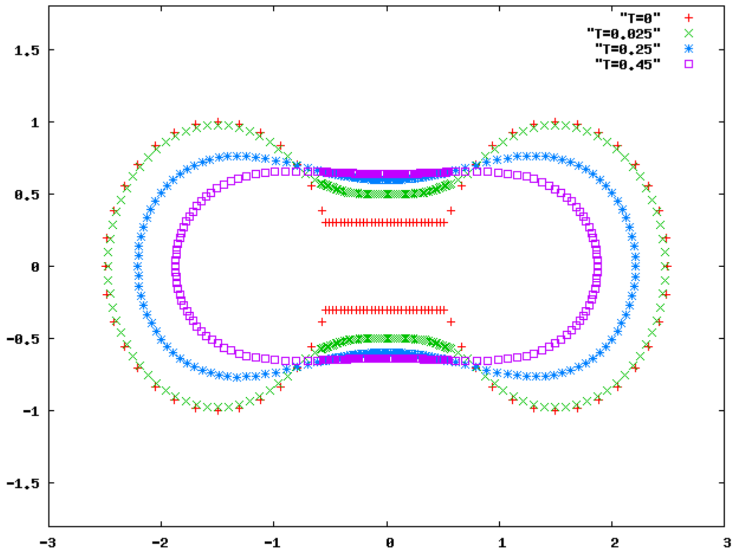

Figure 2.4: Mean curvature flow problem in 2D

for some $\varepsilon, n > 0$ jointly chosen with σ so that Assumption **F** holds true. Under this form, the forward two-dimensional diffusion is defined by:

$$dX_t^{(1)} = \sigma dW_t^{(1)}, \quad \text{and} \quad dX_t^{(2)} = k(m - X_t^{(2)})dt + c\sqrt{X_t^{(2)}}dW_t^{(2)}. \qquad (2.4.9)$$

In order to guarantee the non-negativity of the discrete-time approximation of the process $X^{(2)}$, we use the implicit Milstein scheme [41]:

$$X_n^{(2)} \quad = \quad \frac{X_{n-1}^{(2)} + km\Delta t + c\sqrt{X_{n-1}^{(2)}}\xi_n\sqrt{\Delta t} + \frac{1}{4}c^2\Delta(\xi_n^2 - 1)}{1 + k\Delta t} \qquad (2.4.10)$$

where $(\xi_n)_{n\geq 1}$ is a sequence of independent random variable with distribution $\mathbf{N}(0,1)$.

Our numerical results correspond to the following values of the parameter: $\mu = 0.15$, $c = 0.2$, $k = 0.1$, $m = 0.3$, $Y_0 = m$, $\rho = 0$. The initial value of the portfolio is $x_0 = 1$, the maturity T is taken equal to one year. With this parameters, the value function is computed from the quasi-explicit formula (2.4.7) to be $v_0 = -0.3534$.

We also choose $M = 40$ for the truncation of the nonlinearity. This choice turned out to be critical as an initial choice of $M = 10$ produced an important bias in the results.

The two schemes have been tested with the Malliavin and basis projection methods. The latter was applied with 40×10 basis functions. We provide numerical results corresponding to 2 millions particles. Our numerical results show that the Malliavin and the basis projection methods produce very similar results, and achieve a good accuracy: with 2 millions particles, we calculate the variance of our estimates by performing 100 independent calculations:

- the results of the Malliavin method exhibit a standard deviation smaller than 0.005 for scheme one (except for a step equal to 0.025 and a volatility equal to 1.2 where standard deviation jumped to 0.038), 0.002 for scheme two with a computing time of 378 seconds for 40 time steps,

- the results of the basis projection method exhibit a standard deviation smaller than 0.002 for scheme 1 and 0.0009 for scheme two with a computing time of 114 seconds for 40 time steps.

Figure 2.5 provides the plots of the errors obtained by the integration by parts-based regression with Schemes one and two. All solutions have been calculated as the average of 100 calculations. We first observe that for a small diffusion coefficient $\sigma = 0.2$, the numerical performance of the algorithm is very poor: surprisingly, the error increases as the time step shrinks to zero and the method seems to be biased. This numerical result hints that the requirement that the diffusion should dominate the nonlinearity in Theorem 2.2.1, might be a sharp condition. We also observe that

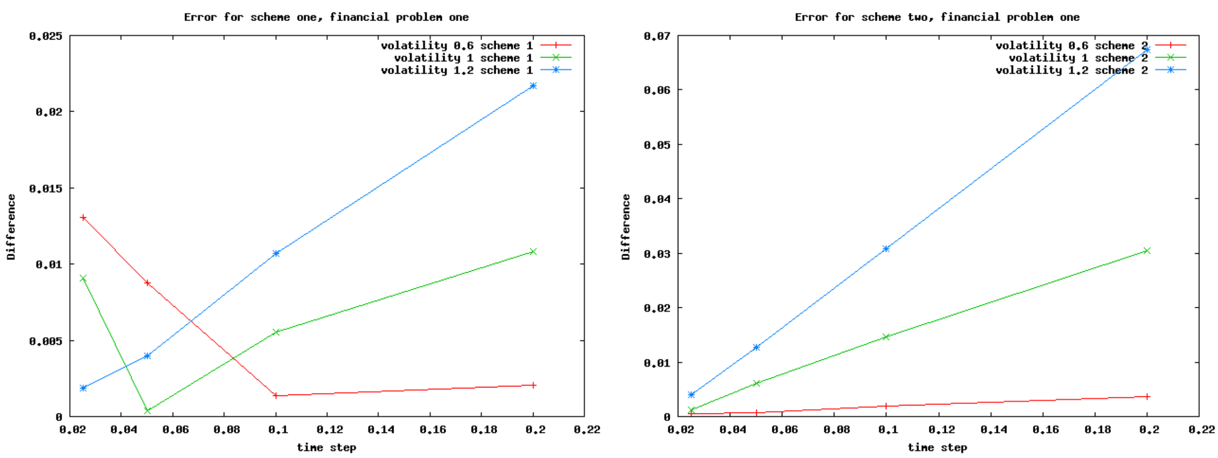

Figure 2.5: Difference between calculation and reference for scheme one and two

scheme one has a persistent bias even for a very small time step, while scheme two exhibits a better convergence towards the solution.

2.4.2.2 A five dimensional example

We now let $n = 2$, and we assume that the interest rate process is defined by the Ornstein-Uhlenbeck process:

$$dr_t = \kappa(b - r_t)dt + \zeta dW_t^{(0)}.$$

While the price process of the second security is defined by a Heston model, the first security's price process is defined by a CEV-SV models, see e.g. [48] for a presentation of these models and their simulation:

$$\begin{aligned} dS_t^{(i)} &= \mu_i S_t^{(i)} dt + \sigma_i \sqrt{Y_t^{(i)}} S_t^{(i)^{\beta_i}} dW_t^{(i,1)}, \quad \beta_2 = 1, \\ dY_t^{(i)} &= k_i \left(m_i - Y_t^{(i)} \right) dt + c_i \sqrt{Y_t^{(i)}} dW_t^{(i,2)} \end{aligned}$$

where $\left(W^{(0)}, W^{(1,1)}, W^{(1,2)}, W^{(2,1)}, W^{(2,2)}\right)$ is a Brownian motion in $\mathbb{R}^5$, and for simplicity we considered a zero-correlation between the security price process and its volatility process.

Since $\beta_2 = 1$, the value function of the portfolio optimization problem (2.4.5) does not depend on the $s^{(2)}-$variable. Given an initial state $(X_t, r_t, S_t^{(1)}, Y_t^{(1)}, Y_t^{(2)}) = (x, r, s_1, y_1, y_2)$ at the time origin t, the value function $v(t, x, r, s_1, y_1, y_2)$ satisfies the HJB equation:

$$\begin{aligned} 0 &= -v_t - (\mathbf{L}^r + \mathbf{L}^Y + \mathbf{L}^{S^1})v - rxv_x \\ &\quad - \sup_{\theta_1, \theta_2} \left\{ \theta_1 \cdot (\mu - r\mathbb{1})v_x + \theta_1 \sigma_1^2 y_1 s_1^{2\beta_1 - 1} v_{xs_1} + \frac{1}{2}(\theta_1^2 \sigma_1^2 y_1 s_1^{2\beta_1 - 2} + \theta_2^2 \sigma_2^2 y_2) v_{xx} \right\} \\ &= -v_t - (\mathbf{L}^r + \mathbf{L}^Y + \mathbf{L}^{S^1})v - rxv_x \\ &\quad + \frac{((\mu_1 - r)v_x + \sigma_1^2 y_1 s_1^{2\beta_1 - 1} v_{xs_1})^2}{2\sigma_1^2 y_1 s_1^{2\beta_1 - 2} v_{xx}} + \frac{((\mu_2 - r)v_x)^2}{2\sigma_2^2 y_2 v_{xx}} \end{aligned} \tag{2.4.11}$$

where

$$\mathbf{L}^r v = \kappa(b - r)v_r + \frac{1}{2}\zeta^2 v_{rr}, \qquad \mathbf{L}^Y v = \sum_{i=1}^{2} k_i (m_i - y_i) v_{y_i} + \frac{1}{2} c_i^2 y_i v_{y_i y_i},$$

$$\text{and} \quad \mathbf{L}^{S^1} v \;=\; \mu_1 s_1 v_{s_1} - \frac{1}{2}\sigma_1^2 s_1 y_1 v_{s_1 s_1}.$$

In order to implement our Monte Carlo-finite differences scheme, we re-write (2.4.11) as:

$$\begin{aligned} &-v_t - (\mathbf{L}^r + \mathbf{L}^Y + \mathbf{L}^{S^1})v - \tfrac{1}{2}\sigma^2 v_{xx} + F\left((x, r, s_1, y_1, y_2), Dv, D^2 v\right) = 0, \\ &v(T, x, r, s_1, y_1, y_2) = -e^{-\eta x}, \end{aligned} \tag{2.4.12}$$

where $\sigma > 0$, and the nonlinearity $F : \mathbb{R}^5 \times \mathbb{R}^5 \times \mathbb{S}_2$ is given by:

$$F(u, z, \gamma) = \frac{1}{2}\sigma^2 \gamma_{11} - x_1 x_2 z_1 + \frac{((\mu_1 - x_2)z_1 + \sigma_1^2 x_4 x_3^{2\beta_1 - 1}\gamma_{1,3})^2}{2\sigma_1^2 x_4 x_3^{2\beta_1 - 2}\gamma_{11}} + \frac{((\mu_2 - x_2)z_1)^2}{2\sigma_2^2 x_5 \gamma_{11}},$$

where $u = (x_1, \cdots, x_5)$. We next consider the truncated nonlinearity:

$$\begin{aligned} F_{\varepsilon,M}(u, z, \gamma) &:= \frac{1}{2}\sigma^2 \gamma_{11} - x_1 x_2 z_1 + \sup_{\varepsilon \le |\theta| \le M} \Big\{ (\theta \cdot (\mu - r\mathbb{1})z_1 + \theta_1 \sigma_1^2 (x_4 \vee \varepsilon)(x_3 \vee \varepsilon)^{2\beta_1 - 1}\gamma_{13} \\ &\quad + \frac{1}{2}(\theta_1^2 \sigma_1^2 (x_3 \vee \varepsilon)(x_4 \vee \varepsilon)^{2\beta_1 - 2} + \theta_2^2 \sigma_2^2 (x_5 \vee \varepsilon))\gamma_{11} \Big\}, \end{aligned}$$

where $\varepsilon, M > 0$ are jointly chosen with σ so that Assumption **F** holds true. Under this form, the forward two-dimensional diffusion is defined by:

$$\begin{aligned}
dX_t^{(1)} &= \sigma dW_t^{(0)}, & dX_t^{(2)} &= \kappa(b - X_t^{(2)})dt + \zeta dW_t^{(1)}, \\
dX_t^{(3)} &= \mu_1 X_t^{(3)} dt + \sigma_1 \sqrt{X_t^{(4)}} X_t^{(3)^{\beta_1}} dW_t^{(1,1)}, & dX_t^{(4)} &= k_1(m_1 - X_t^{(4)})dt + c_1 \sqrt{X_t^{(4)}} dW_t^{(1,2)}, \\
dX_t^{(5)} &= k_2(m_2 - X_t^{(5)})dt + c_2 \sqrt{X_t^{(5)}} dW_t^{(2,2)}.
\end{aligned} \tag{2.4.13}$$

The component $X_t^{(2)}$ is simulated according to the exact discretization:

$$X_{t_n}^{(2)} = b + e^{-k\Delta t}\left(X_{t_{n-1}}^{(2)} - b\right) + \zeta \sqrt{\frac{1 - \exp(-2\kappa\Delta t)}{2\kappa}} \xi_n,$$

where $(\xi_n)_{n \geq 1}$ is a sequence of independent random variable with distribution $\mathbf{N}(0,1)$. The following scheme for the price of the asset guarantees non-negativity (see [1]) :

$$\ln X_n^{(3)} = \ln X_{n-1}^{(3)} + \left(\mu_1 - \frac{1}{2}\sigma_1^2 \left(X_{n-1}^{(3)}\right)^{2(\beta_1 - 1)} X_{n-1}^{(4)}\right) \Delta t + \sigma_1 \left(X_{n-1}^{(3)}\right)^{\beta_i - 1} \sqrt{X_{n-1}^{(4)}} \Delta W_n^{(1,2)}$$

where $\Delta W_n^{(1,2)} := W_n^{(1,2)} - W_{n-1}^{(1,2)}$. We take the following parameters $\mu_1 = 0.10$, $\sigma_1 = 0.3$, $\beta_1 = 0.5$ for the first asset, $k_1 = 0.1$, $m_1 = 1.$, $c_1 = 0.1$ for the diffusion process of the first asset. The second asset is defined by the same parameters as in the two dimensional example: $\mu_2 = 0.15$, $c_2 = 0.2$, $m = 0.3$ and $Y_0^{(2)} = m$. As for the interest rate model we take $b = 0.07$, $X_0^{(2)} = b$, $\zeta = 0.3$.

The initial values of the portfolio the assets prices are all set to 1. For this test case we first use the basis projection regression method with $4 \times 4 \times 4 \times 4 \times 10$ meshes and three millions particles which, for example, takes 520 seconds for 20 time steps. Figure 2.6 contains the plot of the solution obtained by scheme 2, with different time steps. We only provide results for the implementation of scheme 1 with a coarse time step, because the method was diverging with a thinner time step. We observe that there is still a difference for very thin time step with the three considered values of the diffusion. This seems to indicate that more particles and more meshes are needed. While doing many calculation we observed that for the thinner time step mesh, the solution sometimes diverges. We therefore report the results corresponding to thirty millions particles with $4 \times 4 \times 4 \times 4 \times 40$ meshes. First we notice that with this discretization all results are converging as time step goes to zero: the exact solution seems to be very closed to -0.258. During our experiments with thirty millions particles, the scheme was always converging with a very low variance on the results. A single calculation takes 5100 seconds with 20 time steps.

Remark 2.4.4 With thirty millions particles, the memory needed forced us to use 64-bit processors with more than four gigabytes of memory.

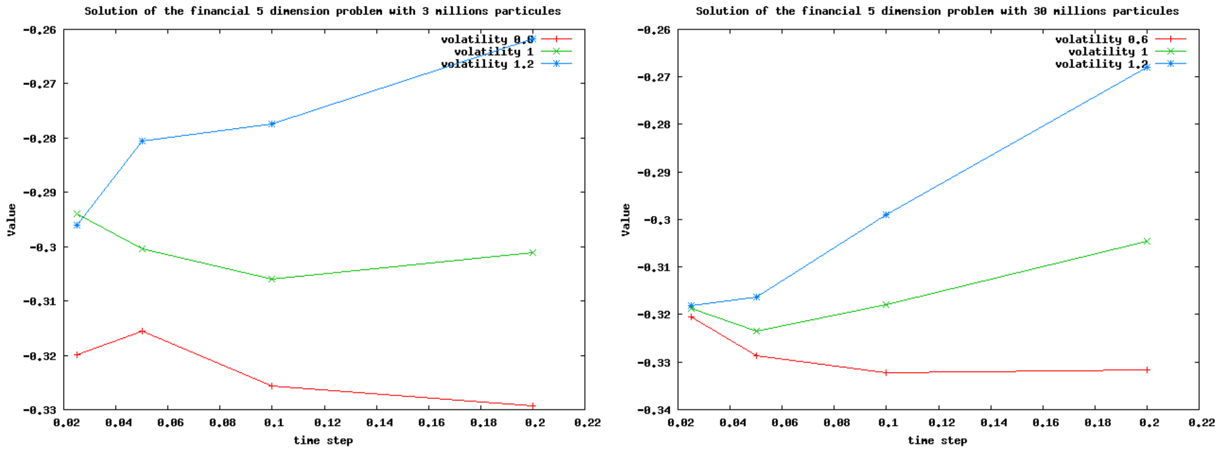

Figure 2.6: Five dimensional financial problem and its results for different volatilities with 3 millions and 30 millions particles

2.4.2.3 Conclusion on numerical results

The Monte Carlo-Finite Differences algorithm has been implemented with both schemes suggested by (2.4.2), using the basis projection and Malliavin regression methods. Our numerical experiments reveal that the second scheme performs better both in term of results and time of calculation for a given number of particles, independently of the regeression method.

We also provided numerical results for different choices of the diffusion parameter in the Monte Carlo step. We observed that small diffusion coefficients lead to poor results, which hints that the condition that the diffusion must dominate the nonlinearity in Assumption **F** (iii) may be sharp. On the other hand, we also observed that large diffusions require a high refinement of the meshes meshes, and large number of particles, leading to a high computational time.

Finally, let us notice that a reasonable choice of the diffusion could be time and state dependent, as in the classical importance sampling method. We have not tried any experiment in this direction, and we hope to have some theoretical results on how to choose optimally the drift and the diffusion coefficient of the Monte Carlo step.

CHAPTER 3

Probabilistic Numerical Methods for Fully non–linear non–local Parabolic PDEs

This Chapter is organized as follows: In Section 3.1, the problematic features of non–local fully non–linear PDEs is discussed on a naïve generalization of the Monte Carlo method from local case in Chapter 2 to non–local case. In Section 3.2 the Monte Carlo quadrature (MCQ) is presented as a purely Monte Carlo approximation of Lévy integral together with the error analysis. Section 3.3 contains the results of convergence and asymptotic properties of the scheme.

3.1 Preliminaries and features for non–local PDEs

Let μ and σ be functions from $[0,T]\times\mathbb{R}^d$ to $\mathbb{R}^d$ and $\mathbb{M}(d,d)$ respectively, η be a function from $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d_*$ to $\mathbb{R}^d$, and $a=\sigma^{\mathrm{T}}\sigma$. Suppose the following non–local Cauchy problem:

$$-\mathcal{L}^X v(t,x) - F\big(t,x,v(t,x),Dv(t,x),D^2v(t,x),v(t,\cdot)\big) = 0, \quad \text{on } [0,T)\times\mathbb{R}^d, \tag{3.1.1}$$

$$v(T,\cdot) = g, \quad \text{on } \in\mathbb{R}^d. \tag{3.1.2}$$

where $F:\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times\mathcal{S}_d\times\mathcal{C}_d\to\mathbb{R}$ and $\mathcal{L}^X$ given by:

$$\begin{aligned}\mathcal{L}^X\varphi(t,x) \;:=\; & \left(\frac{\partial\varphi}{\partial t}+\mu\cdot D\varphi+\frac{1}{2}a\cdot D^2\varphi\right)(t,x)\\ & +\int_{\mathbb{R}^d_*}\Big(\varphi(t,x+\eta(t,x,z))-\varphi(t,x)-\mathbb{1}_{\{|z|\le 1\}}D\varphi(t,x)\cdot\eta(t,x,z)\Big)d\nu(z).\end{aligned}$$

$\mathcal{L}^X$ is the infinitesimal generator of a jump–diffusion, X_t, satisfying SDE:

$$dX_t=\mu(t,X_t)dt+\sigma(t,X_t)dW_t+\int_{\{|z|>1\}}\eta(t,X_{t-},z)J(dt,dz)+\int_{\{|z|\le 1\}}\eta(t,X_{t-},z)\tilde{J}(dt,dz),$$

where J and $\tilde{J}$ are respectively a Poisson jump measure and its compensation who associate to Lévy measure ν by:

$$\begin{aligned}\nu(A) &= \mathbb{E}\left[\int_A J([0,1],dz)\right]\\ \tilde{J}(dt,dz) &= J(dt,dz)-dt\times\nu(dz).\end{aligned}$$

For more details on jump–diffusion processes, see [7] and the references therein or the classic work of [57].

Now, we provide a discretization for the process X. Suppose that $h = \frac{T}{n}$, $t_i = ih$, and $\kappa \geq 0$. We define the Euler discretization of jump–diffusion process X_t with truncated Lévy measure by:

$$\hat{X}_h^{t,x,\kappa} = x + \tilde{\mu}(t,x)h + \sigma(t,x)W_h + \int_{\{|z|>\kappa\}} \eta(t,x,z)\tilde{J}([0,h],dz), \quad (3.1.3)$$

$$\hat{X}_{t_{i+1}}^{x,\kappa} = \hat{X}_h^{t_i,\hat{X}_{t_i}^{x,\kappa},\kappa} \quad \text{and} \quad \hat{X}_0^{x,\kappa} = x. \quad (3.1.4)$$

where $\tilde{\mu}(t,x) = \mu(t,x) + \int_{\{|z|>1\}} \eta(t,x,z)\nu(dz)$ and we make the choice of $\kappa = 0$ when ν is a finite measure. Let $\tilde{N}_t^\kappa$ and N_t^κ be respectively the Poisson process derived from jump measure J by counting all jumps of size greater than κ which happen in time interval $[0,t]$ and its compensation, i.e.

$$N_t^\kappa = \int_{\{|z|>\kappa\}} J([0,t],dz) \quad \text{and} \quad \tilde{N}_t^\kappa = \int_{\{|z|>\kappa\}} \tilde{J}([0,t],dz). \quad (3.1.5)$$

One can write the jump part of $\hat{X}_h^{t,x,\kappa}$ as a compound Poisson process (see for example [27])

$$\hat{X}_h^{t,x,\kappa} \;=\; x + \mu_\kappa(t,x)h + \sigma(t,x)W_h + \sum_{i=1}^{N_h^\kappa} \eta(t,x,Z_i), \quad (3.1.6)$$

where $\mu_\kappa(t,x) = \mu(t,x) - \int_{\{\kappa<|z|\leq 1\}} \eta(t,x,z)\nu(dz)$, Z_is are i.i.d. $\mathbb{R}_*^d$–valued random variables, independent of W and N^κ, and distributed as $\mathbb{1}_{\{|z|>\kappa\}}\frac{1}{\lambda_\kappa}\nu(dz)$.

The classical solution for the problem (3.1.1)–(3.1.2) does not exist in general and therefore, we appeal to the notion of viscosity solutions for non–local parabolic PDEs. We remind that:

Definition 3.1.1 • *The viscosity sub(super)–solution of* (3.1.1)-(3.1.2) *is a upper semi–continuous (lower semi–continuous) function* $\underline{v}(\overline{v})$: $[0,T] \times \mathbb{R}^d \to \mathbb{R}$ *such that:*

1. *for any* $(t_0,x_0) \in [0,T) \times \mathbb{R}^d$ *and any smooth function* φ *with:*

$$0 = \max(\min)\{\underline{v} - \varphi\} = (\overline{v} - \varphi)(t_0,x_0)$$

We have:

$$0 \;\geq (\leq)\; -\mathcal{L}^X\varphi(t_0,x_0) - F\left(\cdot,\varphi,D\varphi,D^2\varphi,\varphi(\cdot)\right)(t_0,x_0).$$

2. $g(\cdot) \geq \underline{v}(T,\cdot)(\leq \overline{v}(T,\cdot))$.

The function v *which is both viscosity sub and super solution, is called viscosity solution of* (3.1.1)-(3.1.2).

• *We say that* (3.1.1) *has comparison for bounded functions if for any bounded lower semi–continuous viscosity super–solution* $\overline{v}$ *and any bounded upper semi–continuous sub–solution* $\underline{v}$, *satisfying*

$$\overline{v}(T,\cdot) \geq \underline{v}(T,\cdot),$$

we have $\overline{v} \geq \underline{v}$ *on* $[0,T] \times \mathbb{R}^d$.

3.1.1 The scheme for non–local fully non–linear parabolic PDEs

In this section, we introduce a probabilistic scheme by following directly the same idea as the scheme for the local PDEs. Then, we consider some problems which prevent us to utilize the scheme in many interesting applications. Therefore, we introduce a modified version of the scheme which works for the class of non–linearities of HJB type (Hamilton–Jacobi–Bellman).

Following the same idea as in Chapter 2, one can obtain the following immature scheme.

$$v^h(T,.) = g \quad \text{and} \quad v^h(t_i,x) = \mathbf{T}_h[v^h](t_i,x), \tag{3.1.7}$$

where for every function $\psi : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}$ with exponential growth:

$$\begin{aligned}\mathbf{T}_h[\psi](t,x) &:= \mathbb{E}\left[\psi\left(t+h, \hat{X}_h^{t,x}\right)\right] + hF\left(t,x,\mathcal{D}_h\psi, \psi(t+h,\cdot)\right), \\ \mathcal{D}_h\psi &:= \left(\mathcal{D}_h^0\psi, \mathcal{D}_h^1\psi, \mathcal{D}_h^2\psi\right),\end{aligned} \tag{3.1.8}$$

where

$$\mathcal{D}_h^k\psi(t,x) \quad := \quad \mathbb{E}\left[\psi(t+h, \hat{X}_h^{t,x,\kappa})H_k^h(t,x)\right], k = 0,1,2, \tag{3.1.9}$$

where

$$H_0^h = 1, \quad H_1^h = \left(\sigma^{\mathrm{T}}\right)^{-1} \frac{W_h}{h}, \quad H_2^h = \left(\sigma^{\mathrm{T}}\right)^{-1} \frac{W_h W_h^{\mathrm{T}} - h\mathbf{I}_d}{h^2}\, \sigma^{-1}.$$

The details of approximation of derivatives with Hermit polynomials can be found in Lemma 2.1 in Chapter 2.

For the above scheme, there is an obvious extension which could be done immediately by the following assumptions analogous to Assumption **F** in Chapter 2, i.e.

Assumption F (i) *The nonlinearity F is Lipschitz-continuous with respect to (x,r,p,γ,ψ) uniformly in t, and $|F(\cdot,\cdot,0,0,0,0)|_\infty < \infty$;*
(ii) *F is elliptic and dominated by the diffusion of the linear operator $\mathcal{L}^X$, i.e.*

$$\nabla_\gamma F \le a \quad \text{on} \quad \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \times \mathbb{C}_d; \tag{3.1.10}$$

(iii) *$F_p \in \mathrm{Image}(F_\gamma)$ and $\left|F_p^T F_\gamma^- F_p\right|_\infty < +\infty$.*
We remind that the non–local non–linearity F is called elliptic if

1. F is non–decreasing on the second derivative component, i.e.

$$F(t,x,r,p,\gamma_1,\psi) \le F(t,x,r,p,\gamma_2,\psi) \quad \text{for} \quad \gamma_1 \le \gamma_2.$$

2. F is non–decreasing on the non–local component, i.e.

$$F(t,x,r,p,\gamma,\psi_1) \le F(t,x,r,p,\gamma,\psi_2) \quad \text{for} \quad \psi_1 \le \psi_2.$$

Then, we have the following Theorem.

Theorem 3.1.1 *Let Assumption* **F** *hold true, and* $|\mu|_1$*,* $|\sigma|_1 < \infty$ *and* σ *is invertible. Also assume that the fully nonlinear PDE (3.1.1) has comparison for bounded functions. Then, for every bounded Lipschitz function* g*, there exists a bounded function* v *such that*

$$v^h \longrightarrow v \qquad \text{locally uniformly.}$$

In addition, v *is the unique bounded viscosity solution of problem (3.1.1)-(3.1.2).*

The proof is an straight forward implementation of the Subsection 3.2 of Chapter 2.

Example 3.1.1 *Let* ν *be a finite positive measure and* $F(t,x,r,p,\gamma,\psi) = G(t,x,r,p,\gamma,\int_{\mathbb{R}^d_*}\psi(x+\eta(t,x,z))\zeta(t,x,z)\nu(dz))$ *for some function* G *such that Assumption* **F** *is valid for* F*. Then, the above Theorem is applicable.* □

However, in the rest if this Section, we show that there are many interesting applications for which Theorem 3.1.1 fails to provide the convergence result. One of the major class of fully non–linear PDEs is the class of HJB equations which come from stochastic control problems arising in many applications including finance. The non–linearity of HJB equations do not satisfies Assumption **F** in general. Even for local PDEs of HJB type, Assumption **F** is not valid, because F is not uniformly Lipschitz with respect to x. In addition, when the Lévy measure ν is an infinite Lévy measure, there is no chance for F to be uniformly Lipschitz with respect to ψ.

The other problem which occurs in many applications is the lack of explicit form for non–linearity F. We present the following example in order to mention this problem.

Example 3.1.2 *Suppose that we want to implement the scheme for the fully non–linear equation:*

$$\begin{aligned} -v_t - F(x, Dv(t,x), D^2v(t,x), v(t,\cdot)) &= 0 \\ v(T,\cdot) &= g(\cdot), \end{aligned}$$

where

$$F(x,p,\gamma,\psi) := \sup_{\theta\in\mathbb{R}_+}\left\{\mathcal{L}^\theta(p,\gamma) + \int_{\mathbb{R}_*}\psi(x+\theta z)\nu(dz)\right\} \tag{3.1.11}$$

$$\mathcal{L}^\theta(p,\gamma) := \theta bp + \frac{1}{2}\theta^2 a^2\gamma \tag{3.1.12}$$

$$\mathcal{I}(x,\psi)^\theta := \int_{\mathbb{R}_*}\psi(x+\theta z)\nu(dz). \tag{3.1.13}$$

This fully non-linear equation solves the problem of portfolio management for one asset in the Black–Scholes model including jumps in asset price. For the sake of *simplicity, for the moment we forget about infinite activity jumps. Observe that if* $\nu = 0$ *(the asset price never jumps) then* F *becomes of the form:*

$$F(x,p,\gamma,\psi) \quad := \quad \sup_{\theta\in\mathbb{R}_+}\left\{\theta b p + \frac{1}{2}\theta^2 a^2\gamma\right\}.$$

which could be given in explicit form by:

$$F(x,p,\gamma,\psi) \quad := \quad -\frac{(bp)^2}{2a^2\gamma},$$

and the scheme could be easily implemented as in Chapter 2 and even for more complicated examples (see Section 2.4). But, when $\nu \neq 0$ *(jump do exists), the explicit form for* F *is not known and the supremum should be approximated. This problem is in common with other numerical methods for fully non–linear PDEs e.g. finite difference.*

Although his problem is obviously beyond the subject of this thesis, we address it here in order to mention that why we provide *MCQ* in Section 3.2 to approximate the integral inside the supremum. More precisely, when there is no explicite form for the non–linearity, one has to calculate the Lévy integral inside the supremum for each θ and then, apply some numerical methods to approximate the supremum over all possible θs. Therefore, we proposed a Monte Carlo Quadrature method to approximate the integral in a purely probabilistic way. The MCQ could be considered independently in other applications.

Now, suppose that ν is an infinite measure and therefore in Example 3.1.2, (3.1.13) should be written of the form

$$\mathcal{I}(x,\psi) \quad := \quad \int_{\mathbb{R}_*}\left(\psi(x+\theta z) - \psi(x) - \mathbb{1}_{\{|z|\le 1\}}\theta D\psi(x)\cdot z\right)\nu(dz).$$

In this case, there are two ways to treat with singular Lévy measure; one is to truncate Lévy measure near zero (as we did for discretization of X) and the other is to approximate infinite small jumps by a Brownian motion. In both cases, the general form for the approximate F is

$$F_\kappa(x,r,p,\gamma,\psi) \quad := \quad \sup_{\theta\in\mathbb{R}_+}\left\{c_\kappa r + \theta b_\kappa p + \frac{1}{2}\theta^2 a^2\gamma + \int_{\{|z|>\kappa\}}\psi(x+\theta z)\nu(dz)\right\}.$$

where

$$c_\kappa := \int_{\{|z|>\kappa\}}\nu(dz) \quad \text{and} \quad b_\kappa := b\int_{\{1\ge|z|>\kappa\}} z\nu(dz).$$

Examining the assumptions of Theorem 3.1.1 to function F_κ, one can easily check that derivatives of F_κ with respect to r, p and ψ blow up to infinity as κ vanishes which destroys the convergence result. To overcome this problem, we will show that κ could be chosen dependent on h, so that the corresponding scheme satisfies the requirements of [6] for the proof of convergence.

Thus in Section 3.3, we will introduce the modified scheme (3.3.3) based on the approximation of non–linearity F obtained from truncation of infinite Lévy measure and MCQ in Section 3.2 and then provide asymptotic results like in Chapter 2 for non–local case.

3.2 Monte Carlo Quadrature (MCQ)

In this section, we propose a Monte Carlo method the value of the following Lévy generator:

$$\mathcal{I}[\varphi](x) \;\; := \;\; \int_{\mathbb{R}^d_*} \left(\varphi\left(x+\eta(z)\right) - \varphi(x) - \mathbb{1}_{\{|z|\le 1\}}\eta(z)\cdot D\varphi(x)\right)\nu(dz). \quad (3.2.1)$$

The method is a pure Monte Carlo method to approximate (3.2.1) and, therefore could be used in the approximation of Lévy integral inside the scheme (3.3.3). Because, the result of this section is independent of the numerical scheme (3.3.3) introduced in this Chapter, one can read it independently from other Section.
Through out this Section, we drop the dependency with respect to (t,x) or other variables and for the sake of simplicity and just write $\eta(z)$.
Notice that in order for (3.2.1) to be well–defined for regular functions, we impose the following assumption on η:

$$\frac{|\eta(z)|}{|z|\wedge 1} \;\; \le \;\; C, \quad \text{for some constant} \;\; C. \quad (3.2.2)$$

We present MCQ in three cases with respect to the behavior of Lévy measure near zero:

- finite measure; $\int_{\{|z|\le 1\}} \nu(dz) < \infty$,
- infinite measure;
  - case I: $\int_{\{|z|\le 1\}} |\eta(z)|\nu(dz) < \infty$,
  - case II: $\int_{\{|z|\le 1\}} |\eta(z)|^2\nu(dz) < \infty$.

3.2.1 Finite Lévy Measure

When Lévy measure is finite, we choose $\kappa = 0$. In this case, we introduce Lemma 3.2.1 which proposes a way to approximate the Lévy integral of general form:

$$\int_{\mathbb{R}^d_*} \varphi(x+\eta(z))\zeta(z)d\nu(z), \quad (3.2.3)$$

and then we use this Lemma to approximate the Lévy infinitesimal generator (3.2.1).

Let J be a jump Poisson measure with intensity given by Lévy measure ν, and $\{N_t\}_{t\ge 0}$ be the Poisson process given by $N_t = \int_0^t \int_{\mathbb{R}^d_*} J(ds,dz)$ whose intensity is

$\lambda := \int_{\mathbb{R}^d_*} \nu(dz)$. By (3.1.6), we can write $\hat{X}^x$ by

$$\hat{X}^x_t = x + \mu_0 t + \sigma W_t + \sum_{i=1}^{N_t} \eta(Z_i) \tag{3.2.4}$$

where Z_is are i.i.d. random variables with law $\frac{1}{\lambda}\nu(dz)$. We also introduce a Lévy process Y_t by

$$Y_t = \sum_{i=1}^{N_t} \zeta(Z_i). \tag{3.2.5}$$

Next Lemma shows that (3.2.3) could be approximated by a Monte Carlo formula purely free of integration.

Lemma 3.2.1 *Let*

$$\hat{\nu}^{\eta,\zeta}_h(\varphi)(x) := \mathbb{E}\left[\int_{\mathbb{R}^d_*} \varphi(\hat{X}^x_h + \eta(z))\zeta(z)d\nu(z)\right]. \tag{3.2.6}$$

Then, for every bounded function $\varphi : \mathbb{R}^d \to \mathbb{R}$:

$$\hat{\nu}^{\eta,\zeta}_h(\varphi)(x) = \frac{1}{h}\mathbb{E}[\varphi(\hat{X}^x_h)Y_h].$$

Proof. For the sake of simplicity, we just concentrate on the jump part of process $\hat{X}^x$ and without loss of generality, we write $\hat{X}^x_h = x + \sum_{i=1}^{N_h} \eta(Y_i)$. The right hand side can be expressed as:

$$\mathbb{E}\left[\varphi(\hat{X}^x_h)Y_h\right] = e^{-\lambda h}\sum_{n=0}^{\infty} \mathbb{E}\left[\varphi(\hat{X}^x_h)Y_h | N_h = n\right]\frac{(\lambda h)^n}{n!}.$$

Then by (3.2.4)–(3.2.5),

$$\begin{aligned}\mathbb{E}\left[\varphi(\hat{X}^x_h)Y_h\right] &= e^{-\lambda h}\lambda h \sum_{n=1}^{\infty}\mathbb{E}\left[\varphi\left(x + \sum_{i=1}^{n}\eta(Z_i)\right)\left(\sum_{j=1}^{n}\zeta(Z_j)\right)\right]\frac{(\lambda h)^{n-1}}{n!} \\ &= e^{-\lambda h}\lambda h \sum_{n=1}^{\infty}\frac{(\lambda h)^{n-1}}{n!}\sum_{j=1}^{n}\mathbb{E}\left[\varphi\left(x + \sum_{i=1}^{n}\eta(Z_i)\right)\zeta(Z_j)\right].\end{aligned}$$

Notice that in the above expression, the summation starts from $n = 1$ because $Y_h = 0$ when $N_h = 0$. Because Z_is are i.i.d. one can conclude that,

$$\sum_{j=1}^{n}\mathbb{E}\left[\varphi\left(x + \sum_{i=1}^{n}\eta(Z_i)\right)\zeta(Z_j)\right] = n\mathbb{E}\left[\varphi\left(x + \sum_{i=1}^{n}\eta(Z_i)\right)\zeta(Z_1)\right]$$

Then, one can write

$$\mathbb{E}\left[\varphi\left(x + \eta(Z_1) + \sum_{i=2}^{n}\eta(Z_i)\right)\zeta(Z_1)\right] = \mathbb{E}\left[\varphi\left(\eta(Z) + \hat{X}^x_h\right)\zeta(Z) | N_h = n-1\right],$$

where Z is dependent of Z_is but has the same law as Z_is. Therefore, we can conclude that:

$$\mathbb{E}\left[\varphi(\hat{X}_h^x)Y_h\right] = e^{-\lambda h}\lambda h\sum_{n=1}^{\infty}\mathbb{E}\left[\varphi(\eta(Z)+\hat{X}_h^x)\zeta(Z)|N_h=n-1\right]\frac{(\lambda h)^{n-1}}{(n-1)!}.$$

But, we know that

$$e^{-\lambda h}\sum_{n=1}^{\infty}\mathbb{E}\left[\varphi(\eta(Z)+\hat{X}_h^x)\zeta(Z)|N_h=n-1\right]\frac{(\lambda h)^{n-1}}{(n-1)!}=\mathbb{E}\left[\varphi(\eta(Z)+\hat{X}_h^x)\zeta(Z)\right]$$

Therefore,

$$\mathbb{E}\left[\varphi(\hat{X}_h^x)Y_h\right] = \lambda h\mathbb{E}\left[\varphi(\eta(Z)+X_h^x)\zeta(Z)\right].$$

Because the density of Z is $\frac{\nu(dz)}{\lambda}$,

$$\mathbb{E}\left[\varphi(\hat{X}_h^x)Y_h\right] = h\mathbb{E}\left[\int_{\mathbb{R}_*^d}\varphi(\eta(z)+\hat{X}_h^x)\zeta(z)d\nu(z)\right].$$

□

In the light of Lemma (3.2.1), we propose the following approximation for (3.2.1):

$$\mathcal{I}_h[\varphi](x) := \hat{\nu}_h^{\eta,1} - \varphi(x)\int_{\mathbb{R}_*^d}\nu(dz) - D\varphi(x)\cdot\int_{\mathbb{R}_*^d}\eta(z)\nu(dz).$$

Next Lemma provide error bound for this approximation.

Lemma 3.2.2 *For any Lipschitz function φ we have:*

$$|(\mathcal{I}_h-\mathcal{I})[\varphi]|_\infty \leq C\sqrt{h}|D\varphi|_\infty. \tag{3.2.7}$$

Proof. As a direct consequence of Lemma (3.2.1), $\hat{\nu}_h^{\eta,1}=\frac{1}{h}\mathbb{E}[\varphi(\hat{X}_h^x)N_h]$. Therefore, one can conclude that,

$$|(\mathcal{I}-\mathcal{I}_h)[\varphi]|_\infty \leq C|D\varphi|_\infty\mathbb{E}\left[|\hat{X}_h^x-x|\right].$$

So, because

$$\mathbb{E}\left[|\hat{X}_h^x-x|\right] \leq C\left(h\int_{\mathbb{R}_*^d}|\eta(z)|\nu(dz)+\sqrt{h}\right), \tag{3.2.8}$$

which provides the result. □

3.2.2 Infinite Lévy Measure

In the case of singular Lévy measure, we truncate Lévy measure near zero and reduce the problem to a finite measure. In other words, for any $\kappa > 0$ we have the truncation approximation of integral operator (3.2.1).

$$\mathcal{I}_\kappa[\varphi](x) \quad := \quad \int_{\{|z|>\kappa\}} \left(\varphi\left(x+\eta(z)\right) - \varphi(x) - \mathbb{1}_{\{|z|\le 1\}}\eta(z)\cdot D\varphi(x)\right)\nu(dz).$$

Then, we use Lemma (3.2.1) to present the MCQ approximation for (3.2.1).

$$\mathcal{I}_{\kappa,h}[\varphi](x) := \hat{\nu}^{\eta,1}_{\kappa,h} - \varphi(x)\int_{\{|z|>\kappa\}} d\nu(z) - \int_{\{1\ge|z|>\kappa\}} \eta(t,x,z)\cdot D\varphi(x)d\nu(z),$$

where by Lemma (3.2.1)

$$\hat{\nu}^{\eta,1}_{\kappa,h} := \int_{\{|z|>\kappa\}} \varphi(\hat{X}^{x,\kappa}_h + \eta(t,x,z))\nu(dz) = h^{-1}\mathbb{E}\left[\varphi(\hat{X}^{x,\kappa}_h)N^\kappa_h\right]$$

Following Lemma provides the error of MCQ approximation of (3.2.1) in the case of infinite Lévy measure.

Lemma 3.2.3 *Let function φ be Lipschitz.*

1. *If $\int_{\{|z|\le 1\}} |z|\nu(dz) < \infty$, then*

$$|(\mathcal{I}_{\kappa,h} - \mathcal{I})[\varphi]|_\infty \le C|D\varphi|_\infty\left(\sqrt{h} + \int_{\{0<|z|\le\kappa\}} |z|\nu(dz)\right). \tag{3.2.9}$$

2. *If $\int_{\{|z|\le 1\}} |z|^2\nu(dz) < \infty$, then*

$$|(\mathcal{I}_{\kappa,h} - \mathcal{I})[\varphi]|_\infty \le C\left(|D\varphi|_\infty\left(\sqrt{h} + h\int_{\{|z|>\kappa\}} |z|\nu(dz)\right) + |D^2\varphi|_\infty\int_{\{0<|z|\le\kappa\}} |z|^2\nu(dz)\right) \tag{3.2.10}$$

Proof.

1. Notice that,

$$|(\mathcal{I} - \mathcal{I}_{\kappa,h})[\varphi]|_\infty \le |(\mathcal{I} - \mathcal{I}_\kappa)[\varphi]|_\infty + |(\mathcal{I}_\kappa - \mathcal{I}_{\kappa,h})[\varphi]|_\infty.$$

By (3.2.2), the truncation error is given by:

$$|(\mathcal{I} - \mathcal{I}_\kappa)[\varphi]|_\infty \le 2|D\varphi|_\infty \int_{\{0<|z|\le\kappa\}} |\eta(z)|\nu(dz). \tag{3.2.11}$$

On the other hand, by (3.2.8) and (3.2.2), we observe that

$$\begin{aligned} |(\mathcal{I}_\kappa - \mathcal{I}_{\kappa,h})[\varphi]|_\infty &\le C|D\varphi|_\infty\left(h\int_{\{|z|>\kappa\}} |\eta(z)|\nu(dz) + \sqrt{h}\right) \\ &\le C|D\varphi|_\infty\left(h\int_{\{|z|>\kappa\}} |z|\nu(dz) + \sqrt{h}\right) \end{aligned}$$

which together with (3.2.11) provides the result.

2. By (3.2.2), the truncation error is given by:

$$|(\mathcal{I}-\mathcal{I}_\kappa)[\varphi]|_\infty \le C|D^2\varphi|_\infty \int_{\{0<|z|\le\kappa\}} |z|^2\nu(dz), \tag{3.2.12}$$

for any function φ with bounded derivatives up to second order. On the other hand, (3.2.8) allows us to calculate the Monte Carlo error by:

$$|(\mathcal{I}_\kappa-\mathcal{I}_{\kappa,h})[\varphi]|_\infty \quad \le \quad C|D\varphi|_\infty \left(h\int_{\{|z|>\kappa\}} |z|\nu(dz) + \sqrt{h}\right)$$

which completes the proof. □

3.3 Asymptotic results

This section is devoted to the convergence result for the scheme (3.3.3). We first remind the notion of viscosity solution and provide the assumptions required for the main results together with the statement of main results. Then, we provide the proof of the results in two separate subsection.

we need to impose the following assumption on the non–linearity F to obtain the convergence Theorem.

Assumption IHJB1: *Function F satisfies:*

$$\frac{1}{2}a(t,x)\cdot\gamma + \mu(t,x)\cdot p + F(t,x,r,p,\gamma,\psi) := \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}}\Big\{\mathcal{L}^{\alpha,\beta}(t,x,r,p,\gamma) + \mathcal{I}^{\alpha,\beta}(t,x,r,p,\gamma,\psi)\Big\}$$

for given sets $\mathcal{A}$ and $\mathcal{B}$ where

$$\mathcal{L}^{\alpha,\beta}(t,x,r,p,\gamma) := \frac{1}{2}a^{\alpha,\beta}(t,x)\cdot\gamma + b^{\alpha,\beta}(t,x)\cdot p + c^{\alpha,\beta}(t,x)r + k^{\alpha,\beta}(t,x),$$

and

$$\mathcal{I}^{\alpha,\beta}(t,x,r,p,\psi) := \int_{\mathbb{R}^d_*} \left(\psi\left(x+\eta^{\alpha,\beta}(t,x,z)\right) - r - \mathbb{1}_{\{|z|\le 1\}}\eta^{\alpha,\beta}(t,x,z)\cdot p\right)\nu(dz)$$

where for any $(\alpha,\beta)\in\mathcal{A}\times\mathcal{B}$, $a^{\alpha,\beta}$, $b^{\alpha,\beta}$, $c^{\alpha,\beta}$, $k^{\alpha,\beta}$ and $\eta^{\alpha,\beta}$ satisfy

$$\sup_{\alpha\in\mathcal{A},\beta\in\mathcal{B}}\left\{|a^{\alpha,\beta}|_1 + |b^{\alpha,\beta}|_1 + |c^{\alpha,\beta}|_1 + |k^{\alpha,\beta}|_1 + \frac{|\eta^{\alpha,\beta}(\cdot,z)|_1}{|z|\wedge 1}\right\} < \infty.$$

The non–linearity is dominated by the diffusion of the linear operator $\mathcal{L}^X$, i.e. for *any t, x, z, α and β*

$$
\begin{aligned}
|a^{-}\cdot a^{\alpha^*,\beta^*}|_1 < \infty \quad &\textit{and} \quad 0\le a^{\alpha,\beta}\le a, && (3.3.1)\\
\eta^{\alpha,\beta},\ b^{\alpha,\beta}\in \mathrm{Image}(a^{\alpha,\beta}) \quad &\textit{and} \quad \sup_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} |(b^{\alpha,\beta})^{\mathrm{T}}(a^{\alpha,\beta})^{-}b^{\alpha,\beta}|_\infty<\infty, && (3.3.2)\\
&\quad\ \sup_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} \frac{|(\eta^{\alpha,\beta})^{\mathrm{T}}(a^{\alpha,\beta})^{-}b^{\alpha,\beta}|_\infty}{1\wedge|z|}<\infty\\
&\quad\ \sup_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} \frac{|(\eta^{\alpha,\beta})^{\mathrm{T}}(a^{\alpha,\beta})^{-}\eta^{\alpha,\beta}|_\infty}{1\wedge|z|^2}<\infty.
\end{aligned}
$$

Remark 3.3.1 A function F which satisfies Assumption **IHJB1** is not well–defined for arbitrary $(t,x,r,p,\gamma,\psi)\in\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times\mathcal{S}_d\times\mathcal{C}_d$. But, for any second order differentiable function, ψ, with bounded derivatives with respect to x, $F(t,x,\psi(t,x),D\psi(t,x),D^2\psi(t,x),\psi(t,\cdot))$ is well–defined.

Now, we propose a Monte Carlo scheme for (3.1.1)-(3.1.2) based on the same idea as in Chapter 2, and also the approximation of the non–linearity.

$$
v^{\kappa,h}(T,.)=g \quad \text{and} \quad v^{\kappa,h}(t_i,x)=\mathbf{T}_{\kappa,h}[v^{\kappa,h}](t_i,x), \tag{3.3.3}
$$

where for every function $\psi:\mathbb{R}_+\times\mathbb{R}^d\longrightarrow\mathbb{R}$ with exponential growth:

$$
\begin{aligned}
\mathbf{T}_{\kappa,h}[\psi](t,x) &:= \mathbb{E}\left[\psi\left(t+h,\hat{X}_h^{t,x,\kappa}\right)\right]+hF_{\kappa,h}\left(t,x,\mathcal{D}_h\psi,\psi(t+h,\cdot)\right), && (3.3.4)\\
\mathcal{D}_h\psi &:= \left(\mathcal{D}_h^0\psi,\mathcal{D}_h^1\psi,\mathcal{D}_h^2\psi\right),
\end{aligned}
$$

$$
\begin{aligned}
F_{\kappa,h}(t,x,r,p,\gamma,\psi) = \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}}\Bigg\{ &\frac{1}{2}a^{\alpha,\beta}(t,x)\cdot\gamma+b^{\alpha,\beta}(t,x)\cdot p+c^{\alpha,\beta}(t,x)r+k^{\alpha,\beta}(t,x)\\
&+\int_{\{|z|\ge\kappa\}}\Big(\hat{\nu}_h^{\eta^{\alpha,\beta},1}(\psi(t,\cdot))(x)-r-\eta^{\alpha,\beta}(t,x,z)\cdot p\Big)\nu(dz)\Bigg\},
\end{aligned}
$$

and

$$
\mathcal{D}_h^k\psi(t,x) := \mathbb{E}\left[\psi(t+h,\hat{X}_h^{t,x,\kappa})H_k^h(t,x)\right], k=0,1,2, \tag{3.3.5}
$$

where

$$
H_0^h=1,\quad H_1^h=\left(\sigma^{\mathrm{T}}\right)^{-1}\frac{W_h}{h},\quad H_2^h=\left(\sigma^{\mathrm{T}}\right)^{-1}\frac{W_hW_h^{\mathrm{T}}-h\mathbf{I}_d}{h^2}\,\sigma^{-1}.
$$

The details of approximation of derivatives with (3.3.5) can be found in Lemma 2.1 in Chapter 2. In order to have the convergence result, we also need to impose the following assumption over $F_{\kappa,h}$.

Assumption Inf–Sup: *For any $\kappa>0$, $t\in[0,T]$, x and $x'\in\mathbb{R}^d$ and any Lipschitz functions ψ and φ, there exists a $(\alpha^*,\beta^*)\in\mathcal{A}\times\mathcal{B}$ such that*

$$
\Phi_\kappa^{\alpha^*,\beta^*}[\psi,\varphi](t,x,x')=\mathcal{J}_\kappa^{\alpha^*,\beta^*}[\psi](t,x)-\mathcal{J}_\kappa^{\alpha^*,\beta^*}[\varphi](t,x')
$$

where

$$\Phi^{\alpha,\beta}_{\kappa}[\psi,\varphi](t,x) := \inf_{\alpha} \mathcal{J}^{\alpha,\beta}_{\kappa}[\psi](t,x) - \sup_{\beta} \mathcal{J}^{\alpha,\beta}_{\kappa}[\varphi](t,x'), \tag{3.3.6}$$

and

$$\begin{aligned}\mathcal{J}^{\alpha,\beta}_{\kappa}[\phi](t,x) \quad := \quad & \frac{1}{2}a^{\alpha,\beta}\cdot D^2\phi(t,x) + b^{\alpha,\beta}\cdot D\phi(t,x) + c^{\alpha,\beta}\phi(t,x) + k^{\alpha,\beta}(t,x) \\ & + \int_{\{|z|\geq\kappa\}} \big(\hat{\nu}_h^{\eta^{\alpha,\beta},1}(\phi(t,\cdot))(x) - \phi(t,x) - \eta^{\alpha,\beta}(t,x,z)\cdot D\phi(t,x)\big)\nu(dz).\end{aligned}$$

The first result concerns the convergence of the convergence of $v^{\kappa,h}$ for κ appropriately chosen with respect to h.

Theorem 3.3.1 (Convergence) *Let η, μ and σ be bounded and Lipschitz continuous on x uniformly on t and z, σ is invertible and Assumptions* **IHJB1** *and* **Inf–Sup** *hold true, and assume that* (3.1.1) *has comparison for bounded functions. Then, if κ_h is such that:*

$$\lim_{h\to 0} \kappa_h = 0 \qquad \textit{and} \qquad \limsup_{h\to 0} \theta^2_{\kappa_h} h = 0 \tag{3.3.7}$$

where

$$\theta_\kappa \quad := \quad \sup_{\alpha,\beta} |\theta^{\alpha,\beta}_{\kappa}|_\infty, \tag{3.3.8}$$

with

$$\begin{aligned}\theta^{\alpha,\beta}_{\kappa} \quad := \quad & c^{\alpha,\beta} + \int_{\{|z|\geq\kappa\}}\nu(dz) + \frac{1}{4}\left(b^{\alpha,\beta} - \int_{\{1>|z|\geq\kappa\}}\eta^{\alpha,\beta}(z)\nu(dz)\right)^T \\ & \times(a^{\alpha,\beta})^-\left(b^{\alpha,\beta} - \int_{\{1>|z|\geq\kappa\}}\eta^{\alpha,\beta}(z)\nu(dz)\right),\end{aligned}$$

then $v^{\kappa_h,h}$ converges to some function v locally uniform. In addition, v is the unique viscosity solution of (3.1.1)-(3.1.2).
Specially, if Lévy measure is finite, for the choice of $\kappa_h = 0$ the assertion of the Theorem hold true.

Remark 3.3.2 It is always possible to choose κ_h such that (3.3.7) is satisfied. To see this, notice that θ_κ in (3.3.8) is non–increasing on κ

$$\lim_{\kappa\to 0}\theta_\kappa = +\infty \quad \text{and} \quad \limsup_{\kappa\to\infty}\theta_\kappa < \infty.$$

Then, we define $\kappa_h := \inf\{\kappa | \theta_\kappa \leq h^{-\frac{1}{2}}\} + h$. By the definition of κ_h, $\theta_{\kappa_h} \leq h^{-\frac{1}{2}}$. Because Observe that κ_h is non–decreasing with respect to h and $\lim_{h\to 0}\kappa_h = 0$.
If there exists a q such that, $q := \lim_{h\to 0}\kappa_h > 0$, then, for $\kappa < q$, we would have $\theta_\kappa = \infty$ which obviously contradicts the fact that for $\kappa > 0$, $\theta_\kappa < \infty$. Therefore, κ_h satisfies (3.3.8).

Remark 3.3.3 The choice of κ_h in the above Theorem seems to be crucial for the convergence. Otherwise, we only have the following convergence result.

Proposition 3.3.1 *Under the same assumption as Theorem 3.3.1, when Lévy measure ν is infinite, for every Lipschitz bounded function g, we have*

$$\lim_{\kappa\to 0}\lim_{h\to 0} v^{\kappa,h} = v$$

where v is the unique viscosity solution of (3.1.1)–(3.1.2) assuming that it exists.

Proof. . Let v^κ be the solution of the following problem:

$$-\mathcal{L}^X v^\kappa(t,x) - F_\kappa\big(t,x,v^\kappa(t,x),Dv^\kappa(t,x),D^2v^\kappa(t,x),v^\kappa(t,\cdot)\big) = 0, \text{on}[0,T)\times\mathbb{R}^d, \quad (3.3.9)$$

$$v^\kappa(T,\cdot) = g(\cdot), \qquad \text{on } \in \mathbb{R}^d. \quad (3.3.10)$$

where $F_\kappa : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \times \mathcal{C}_d \to \mathbb{R}$ is given by:

$$F_\kappa(t,x,r,p,\gamma,\psi) \quad := \quad \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}}\Big\{\mathcal{L}^{\alpha,\beta}(t,x,r,p,\gamma) + \mathcal{I}_\kappa^{\alpha,\beta}(t,x,r,p,\gamma,\psi)\Big\}$$

where

$$\mathcal{I}_\kappa^{\alpha,\beta}(t,x,r,p,\gamma,\psi) := \int_{\{|z|\geq\kappa\}}\big(\psi\big(x+\eta^{\alpha,\beta}(t,x,z)\big) - r - \mathbb{1}_{\{|z|\leq 1\}}\eta^{\alpha,\beta}(t,x,z)\cdot p\big)\nu(dz) \quad (3.3.11)$$

where $a^{\alpha,\beta}$, $b^{\alpha,\beta}$, $c^{\alpha,\beta}$, $k^{\alpha,\beta}$ and $\eta^{\alpha,\beta}$ are as in Assumption **IHJB1**. Let $v^{\kappa,h}$ be the approximate solution given by the scheme (3.3.3). Let $\kappa > 0$ be fixed. Because the truncated Lévy measure is finite, by Theorem 3.3.1, $v^{\kappa,h}$ converges to v^κ locally uniformly as $h \to 0$. Let v^κ be the solution of (3.3.9)–(3.3.10). By Theorem 5.1 of [13] and Assumption **IHJB1**, we have:

$$|v - v^\kappa|_\infty \quad \leq \quad C\sup_{\alpha,\beta}\left\{\left(\int_{0<|z|<\kappa}|\eta^{\alpha,\beta}(\cdot,z)|_\infty^2\nu(dz)\right)^{\frac{1}{2}}\right\} \quad (3.3.12)$$

$$\leq \quad C\left(\int_{0<|z|<\kappa}|z|_\infty^2\nu(dz)\right)^{\frac{1}{2}}. \quad (3.3.13)$$

Therefore, one can choose $\kappa > 0$ so that $|v^\kappa - v|_\infty$ be small enough. Then, when h goes to 0, $v^{\kappa,h}$ converges to v^κ. □

The above limit proposes to implement the numerical scheme in two steps:

- First by choosing κ so that v^κ is near enough to v, we obtain a uniform approximation of v.
- Second by sending $h \to 0$, we obtain locally uniform convergence of $v^{\kappa,h}$ to v^κ.

Notice that the above convergence is not uniformly on (κ, h). However, the convergence in Theorem 3.3.1, is uniform on h when the choice of κ is made suitably dependent on h.

Remark 3.3.4 By Remark 3.7 in Chapter 2,the boundedness condition on g can be relaxed.

In order to obtain the rate of convergence result, we impose Assumptions **IHJB2** and **IHJB2+** which restrict us to concave non–linearities.

Assumption IHJB2 *The non–linearity F satisfies Assumption* **IHJB1** *with $\mathcal{B}$ be a singleton set*

Remark 3.3.5 Therefore, when the non–linearity F satisfies **IHJB2**, we can drop the super script β and write F by

$$\begin{aligned}\frac{1}{2}a(t,x)\cdot\gamma+\mu(t,x)\cdot p+F(t,x,r,p,\gamma,\psi) \quad &:= \quad \inf_{\alpha\in\mathcal{A}}\Big\{\mathcal{L}^\alpha(t,x,r,p,\gamma)\\ &\qquad +\mathcal{I}^\alpha(t,x,r,p,\gamma,\psi)\Big\}\end{aligned}$$

where

$$\mathcal{L}^\alpha(t,x,r,p,\gamma) \quad := \quad \frac{1}{2}\mathrm{Tr}\left[(a^\alpha)^{\mathrm{T}}\right](t,x)\gamma+b^\alpha(t,x)p+c^\alpha(t,x)r+k^\alpha(t,x),$$

and

$$\mathcal{I}^\alpha(t,x,r,p,\psi) \quad := \quad \int_{\mathbb{R}^d_*}\left(\psi\left(x+\eta^\alpha(t,x,z)\right)-r-\mathbb{1}_{\{|z|\le 1\}}\eta^\alpha(t,x,z)\cdot p\right)\nu(dz).$$

In this case, the non–linearity is a concave function of (r,p,γ,ψ).

Assumption IHJB2+ *The non–linearity F satisfies* **IHJB2** *and for any $\delta>0$, there exists a finite set $\{\alpha_i\}_{i=1}^{M_\delta}$ such that for any $\alpha\in\mathcal{A}$:*

$$\inf_{1\le i\le M_\delta}\left\{\begin{array}{l}|\sigma^\alpha-\sigma^{\alpha_i}|_\infty+|b^\alpha-b^{\alpha_i}|_\infty+|c^\alpha-c^{\alpha_i}|_\infty\\ \qquad +|k^\alpha-k^{\alpha_i}|_\infty+\int\limits_{\mathbb{R}^d_*}|(\eta^\alpha-\eta^{\alpha_i})(\cdot,z)|^2_\infty d\nu(z)\end{array}\right\}\le\delta.$$

Remark 3.3.6 The Assumption IHJB2+ is satisfied if $\mathcal{A}$ is a compact separable topological space and $\sigma^\alpha(\cdot)$, $b^\alpha(\cdot)$, and $c^\alpha(\cdot)$ are continuous maps from $\mathcal{A}$ to $C_b^{\frac{1}{2},1}([0,T]\times\mathbb{R}^d)$; the space of bounded maps which are Lipschitz on x and $\frac{1}{2}$–Hölder on t and $\eta^\alpha(\cdot)$ is continuous maps from $\mathcal{A}$ to $\Big\{\varphi : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d_*\to\mathbb{R}\Big|\int_{R^d_*}|\varphi(\cdot,z)|^2_\infty\nu(dz)<\infty\Big\}$.

Theorem 3.3.2 (Rate of Convergence) *Assume that the final condition g is bounded and Lipschitz-continuous. Then, there is a constant $C > 0$ such that*

- *under Assumption IHJB,*
  $v - v^{\kappa,h} \le C\left(h^{\frac{1}{4}} + h\theta_\kappa^2 + h\varepsilon^{-3} + h^{\frac{3}{4}}\theta_\kappa + h\sqrt{\theta_\kappa} + h^{-\frac{1}{4}}\int_{\{|z|\le\kappa\}}|z|^2\nu(dz)\right).$
- *under Assumption IHJB2+,*
  $-C\left(h^{1/10} + h^{\frac{7}{10}}\theta_\kappa + h\sqrt{\theta_\kappa} + h^{-\frac{3}{10}}\int_{\{|z|\le\kappa\}}|z|^2\nu(dz)\right) \le v - v^{\kappa,h}.$

In addition, if it is possible to find κ_h such that

$$\lim_{h\to 0}\kappa_h = 0, \quad \limsup_{h\to 0} h^{\frac{3}{4}}\theta_{\kappa_h}^2 < \infty \quad \textit{and} \quad \limsup_{h\to 0} h^{-\frac{1}{2}}\int_{0<|z|<\kappa_h}|z|^2\nu(dz) < \infty, \tag{3.3.14}$$

then, there is a constant $C > 0$ such that

- *under Assumption IHJB, $v - v^{\kappa_h,h} \le Ch^{1/4}$.*
- *under Assumption IHJB2+, $-Ch^{1/10} \le v - v^{\kappa_h,h}$.*

Example 3.3.1 *For the Lévy measure*

$$\nu(dz) = \mathbb{1}_{\mathbb{R}^d_*}|z|^{-d-1}dz,$$

one can always find κ_h such that the condition of Theorem 3.3.2 is satisfied. In the other words, it is always enough to choose κ_h such that

$$\limsup_{h\to 0} h^{-\frac{1}{2}}\kappa_h = 0.$$

3.3.1 Convergence

We suppose the all the assumptions of Theorem 3.3.1 holds true throughout this subsection.

We first manipulate the scheme to provide strict monotonicity by the similar idea as in Remark 3.13 and Lemma 3.19 in Chapter 2. Let $u^{\kappa,h}$ be the solution of

$$u^{\kappa,h}(T,\cdot) = g \quad \text{and} \quad u^{\kappa,h}(t_i,x) = \overline{\mathbf{T}}_{\kappa,h}[u^{\kappa,h}](t_i,x), \tag{3.3.15}$$

where

$$\overline{\mathbf{T}}_{\kappa,h}[\psi](t,x) := \mathbb{E}\left[\psi\left(t+h, \hat{X}_h^{t,x,\kappa}\right)\right] + h\overline{F}_{\kappa,h}\left(t,x,\mathcal{D}_h\psi,\psi(t+h,\cdot)\right) \tag{3.3.16}$$

and

$$\overline{F}_{\kappa,h}(t,x,r,p,\gamma,\psi) = \sup_\alpha\inf_\beta\left\{\frac{1}{2}a^{\alpha,\beta}\cdot\gamma + b^{\alpha,\beta}\cdot p + (c^{\alpha,\beta}+\theta_\kappa)r + e^{\theta_\kappa(T-t)}k^{\alpha,\beta}(t,x)\right.$$
$$\left. + \int_{\{|z|\ge\kappa\}}\left(\hat{\nu}_{\kappa,h}^{\eta^{\alpha,\beta},1}(\psi) - r - \mathbb{1}_{\{|z|\le 1\}}\eta^{\alpha,\beta}(z)\cdot p\right)\nu(dz)\right\}.$$

Remark 3.3.7 Assumption **Inf–Sup** is also true if we replace $\mathcal{J}_\kappa^{\alpha,\beta}$ by

$$\bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\psi](t,x) = \frac{1}{2}a^{\alpha,\beta}\cdot D^2\phi(t,x) + b^{\alpha,\beta}\cdot D\phi(t,x) + (c^{\alpha,\beta}+\theta_\kappa)\phi(t,x) + e^{\theta_\kappa(T-t)}k^{\alpha,\beta}(t,x)$$
$$+\int_{\{|z|\geq\kappa\}}\big(\hat{\nu}_h^{\eta^{\alpha,\beta},1}(\phi(t,\cdot))(x) - \phi(t,x) - \eta^{\alpha,\beta}(t,x,z)\cdot D\phi(t,x)\big)\nu(dz).$$

The proof is straight forward.

We have the following Lemma which shows that for proper choice of θ_κ the scheme (3.3.15) is strictly monotone.

Lemma 3.3.1 *Let θ_κ be as in (3.3.8) and φ and ψ : $[0,T]\times\mathbb{R}^d \longrightarrow \mathbb{R}$ be two bounded functions. Then:*

$$\varphi \leq \psi \implies \overline{\mathbf{T}}_{\kappa,h}[\varphi] \leq \overline{\mathbf{T}}_{\kappa,h}[\psi].$$

Proof. Let $f := \psi-\varphi \geq 0$ where φ and ψ are as in the statement of the lemma. For simplicity, we drop the dependence on (t,x) when it is not necessary. By Assumption **IHJB1** and Lemma (3.2.1), we can write:

$$\overline{\mathbf{T}}_{\kappa,h}[\psi] - \overline{\mathbf{T}}_{\kappa,h}[\varphi] = \mathbb{E}[f(t+h,\hat{X}_h)]$$
$$+h\left(\inf_\alpha\sup_\beta \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\widehat{\psi}](t+h,x) - \inf_\alpha\sup_\beta \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\widehat{\varphi}](t+h,x)\right),$$

where $\widehat{\phi}(t,x) := \mathbb{E}[\phi(t,\hat{X}_h^x)]$ for $\phi = \varphi$ or ψ. Therefore,

$$\overline{\mathbf{T}}_{\kappa,h}[\psi] - \overline{\mathbf{T}}_{\kappa,h}[\varphi] \quad\geq\quad \mathbb{E}[f(t+h,\hat{X}_h)] + h\bar{\Phi}_\kappa^{\alpha,\beta}[\widehat{\psi},\widehat{\varphi}](t+h,x,x),$$

where $\bar{\Phi}_\kappa^{\alpha,\beta}$ is defined by

$$\bar{\Phi}_\kappa^{\alpha,\beta}[\psi,\varphi](t,x) := \inf_\alpha \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\psi](t,x) - \sup_\beta \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\varphi](t,x').$$

By Assumption **Inf–Sup**, there exists (α^*,β^*) so that

$$\overline{\mathbf{T}}_{\kappa,h}[\psi] - \overline{\mathbf{T}}_{\kappa,h}[\varphi] \quad\geq\quad \mathbb{E}[f(t+h,\hat{X}_h)] + h\left(\bar{\mathcal{J}}_\kappa^{*}[\widehat{\psi}](t+h,x) - \bar{\mathcal{J}}_\kappa^{*}[\widehat{\varphi}](t+h,x)\right).$$

Observe that by the linearity of $\bar{\mathcal{J}}_\kappa^{\alpha,\beta}$, one can write:

$$\bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\hat{\phi}](t+h,x) \quad=\quad \mathbb{E}\left[\bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\phi](t+h,\hat{X}_h)\right].$$

By the definition of $\bar{\mathcal{J}}_\kappa^{\alpha,\beta}$ and Lemma 2.1 in Chapter 2,

$$\overline{\mathbf{T}}_{\kappa,h}[\psi]-\overline{\mathbf{T}}_{\kappa,h}[\varphi] \geq \mathbb{E}\Bigg[f(\hat{X}_h)\bigg(1+h\Big(c_\kappa^{\alpha^*,\beta^*}+\theta_\kappa + b_\kappa^{\alpha^*,\beta^*}\cdot(\sigma^{\mathrm{T}})^{-1}\frac{W_h}{h}$$
$$+\frac{1}{2}a^{\alpha^*,\beta^*}\cdot(\sigma^{\mathrm{T}})^{-1}\frac{W_hW_h^{\mathrm{T}}-h\mathrm{I}_d}{h^2}\sigma^{-1}\Big)\bigg) + h\hat{\nu}_h^{\eta^{\alpha^*,\beta^*},1}(f)\Bigg],$$

where $b_\kappa^{\alpha,\beta} = b^{\alpha,\beta} - \int_{\{1>|z|\geq\kappa\}}\eta^{\alpha,\beta}(z)\nu(dz)$ and $c_\kappa^{\alpha,\beta} = c^{\alpha,\beta} - \int_{\{|z|\geq\kappa\}}\nu(dz)$.

Therefore, by the same argument as in Lemma 3.12 in Chapter 2, one can write:

$$\begin{aligned}\overline{\mathbf{T}}_{\kappa,h}[\psi] - \overline{\mathbf{T}}_{\kappa,h}[\varphi] \geq \mathbb{E}\Big[f(\hat{X}_h)\Big(1 - \frac{1}{2}a^{\alpha^*,\beta^*}\cdot a^{-1} + h\Big(|A_h^{\alpha^*,\beta^*}|^2 + c_\kappa^* + \theta_\kappa \\ - \frac{1}{4}(b_\kappa^{\alpha^*,\beta^*})^{\mathrm{T}}(a^{\alpha^*,\beta^*})^{-}b_\kappa^{\alpha^*,\beta^*}\Big)\Big) + h\hat{\nu}_h^{\eta^{\alpha^*,\beta^*},1}(f)\Big],\end{aligned}$$

where

$$A_h^{\alpha^*,\beta^*} \;:=\; \frac{1}{h}(\sigma^{\alpha^*,\beta^*})^{1/2}(\sigma^{\mathrm{T}})^{-1}W_h + \frac{1}{2}((\sigma^{\alpha^*,\beta^*})^{-})^{1/2}b_\kappa^{\alpha^*,\beta^*}. \qquad (3.3.17)$$

Therefore, by positivity of f and Assumption **IHJB1**, one can deduce:

$$\overline{\mathbf{T}}_{\kappa,h}[\psi] - \overline{\mathbf{T}}_{\kappa,h}[\varphi] \;\geq\; h\mathbb{E}\Big[f(\hat{X}_h)\Big(c_\kappa^* + \theta_\kappa - \frac{1}{4}(b_\kappa^{\alpha^*,\beta^*})^{\mathrm{T}}(a^{\alpha^*,\beta^*})^{-}b_\kappa^{\alpha^*,\beta^*}\Big)\Big]$$

By the choice of θ_κ in (3.3.8), we have

$$\overline{\mathbf{T}}_{\kappa,h}[\psi] - \overline{\mathbf{T}}_{\kappa,h}[\varphi] \;\geq\; 0.$$

Then, sending ε to zero provides the result. □

The following Corollary shows the monotonicity of scheme 3.3.3.

Corollary 3.3.1 *Let $\varphi,\psi:\ [0,T]\times\mathbb{R}^d \longrightarrow \mathbb{R}$ be two bounded functions. Then:*

$$\varphi\leq\psi \implies \mathbf{T}_{\kappa,h}[\varphi] \leq \mathbf{T}_{\kappa,h}[\psi] - \frac{\theta_\kappa^2 h^2}{2}e^{-\theta_\kappa h}\mathbb{E}[(\psi-\varphi)(t+h,\hat{X}_h^{t,x,\kappa})].$$

In particular, if κ_h satisfies (3.3.7), then

$$\varphi\leq\psi \implies \mathbf{T}_{\kappa_h,h}[\varphi] \leq \mathbf{T}_{\kappa_h,h}[\psi] + Ch\mathbb{E}[(\psi-\varphi)(t+h,\hat{X}_h^{t,x,\kappa_h})]$$

for some constant C.

Proof. Let θ_κ be as in Lemma 3.3.1 and define $\varphi_\kappa(t,x) := e^{\theta_\kappa(T-t)}\varphi(t,x)$ and $\psi_\kappa(t,x) := e^{\theta_\kappa(T-t)}\psi(t,x)$. By Lemma 3.3.1,

$$\overline{\mathbf{T}}_{\kappa,h}[\varphi_\kappa] \;\leq\; \overline{\mathbf{T}}_{\kappa,h}[\psi_\kappa].$$

By multiplying both sides by $e^{-\theta_\kappa(T-t)}$, we have

$$\begin{aligned}&\Big(e^{-\theta_\kappa h}(1+\theta_\kappa h) - 1\Big)\mathbb{E}[\varphi(t+h,\hat{X}_h^{t,x,\kappa})] + \mathbf{T}_{\kappa,h}[\varphi] \\ &\qquad\leq \Big(e^{-\theta_\kappa h}(1+\theta_\kappa h) - 1\Big)\mathbb{E}[\psi(t+h,\hat{X}_h^{t,x,\kappa})] + \mathbf{T}_{\kappa,h}[\psi].\end{aligned}$$

So,

$$\mathbf{T}_{\kappa,h}[\varphi] \;\leq\; \Big(e^{-\theta_\kappa h}(1+\theta_\kappa h) - 1\Big)\mathbb{E}[(\psi-\varphi)(t+h,\hat{X}_h^{t,x,\kappa})] + \mathbf{T}_{\kappa,h}[\psi].$$

But, $e^{-\theta_\kappa h}(1+\theta_\kappa h)-1 \le -\frac{\theta_\kappa^2 h^2}{2}e^{-\theta_\kappa h}$. So,

$$\mathbf{T}_{\kappa,h}[\varphi] \quad \le \quad -\frac{\theta_\kappa^2 h^2}{2}e^{-\theta_\kappa h}\mathbb{E}[(\psi-\varphi)(t+h,\hat{X}_h^{t,x,\kappa})]+\mathbf{T}_{\kappa,h}[\psi].$$

which (3.3.7) provides the result. □

In order to provide a uniform bound on $v^{\kappa,h}$, we bound $u^{\kappa,h}$ with respect to θ_κ as in the following Lemma.

Lemma 3.3.2 *Let φ and $\psi : [0,T]\times\mathbb{R}^d \longrightarrow \mathbb{R}$ be two $L^\infty-$bounded functions. Then*

$$|\overline{\mathbf{T}}_{\kappa,h}[\varphi]-\overline{\mathbf{T}}_{\kappa,h}[\psi]|_\infty \le |\varphi-\psi|_\infty(1+(C+\theta_\kappa)h)$$

where $C=\sup_{\alpha,\beta}|c^{\alpha,\beta}|_\infty$. In particular, if g is $L^\infty-$bounded, for a fixed κ the family $(u^{\kappa,h}(t,\cdot))_h$ defined in (3.3.3) is $L^\infty-$bounded, uniformly in h by

$$(\overline{C}+|g|_\infty)e^{(C+\theta_\kappa)(T-t_i)}.$$

Proof. Let $f:=\varphi-\psi$. Then, by Assumption **Inf–Sup** and the same argument as in the proof of Lemma 3.3.1,

$$\begin{aligned}\overline{\mathbf{T}}_{\kappa,h}[\varphi]-\overline{\mathbf{T}}_{\kappa,h}[\psi] \le \mathbb{E}\Big[f(\hat{X}_h)\Big(1-a^{-1}\cdot a^{\alpha^*,\beta^*}+h\Big(|A_h^{\alpha^*,\beta^*}|^2+c^{\alpha^*,\beta^*}+\theta_\kappa \\ -\int_{\{|z|\ge\kappa\}}\nu(dz)-\frac{1}{4}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\ge\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)^{\mathrm{T}}(a^{\alpha^*,\beta^*})^- \\ \times\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\ge\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)\Big)\Big)+h\hat{\nu}_h^{\eta^{\alpha^*,\beta^*},1}(f)\Big],\end{aligned}$$

where $A_h^{\alpha^*,\beta^*}$ is given by (3.3.17). On the other hand,

$$\Big|\hat{\nu}_h^{\eta^{\alpha^*,\beta^*},1}(f)\Big| \le |f|_\infty\int_{\{|z|\ge\kappa\}}\nu(dz)$$

Therefore ,

$$\begin{aligned}\overline{\mathbf{T}}_{\kappa,h}[\varphi]-\overline{\mathbf{T}}_{\kappa,h}[\psi] \le |f|_\infty\mathbb{E}\Big[\Big|1-a^{-1}\cdot a^{\alpha^*,\beta^*}+h\Big(|A_h^{\alpha^*,\beta^*}|^2+c^{\alpha^*,\beta^*}+\theta_\kappa \\ -\frac{1}{4}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\ge\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)^{\mathrm{T}}(a^{\alpha^*,\beta^*})^-\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\ge\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)\Big)\Big|\Big].\end{aligned}$$

By Assumption **IHJB1** and (3.3.8), $1-a^{-1}\cdot a^{\alpha^*,\beta^*}$ and

$$c^{\alpha^*,\beta^*}+\theta_\kappa-\frac{1}{4}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\ge\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)^{\mathrm{T}}(a^{\alpha^*,\beta^*})^-\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\ge\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)$$

are positive. Therefore, one can write

$$\overline{\mathbf{T}}_{\kappa,h}[\varphi]-\overline{\mathbf{T}}_{\kappa,h}[\psi]\leq|f|_\infty\Big(1-a^{-1}\cdot a^{\alpha^*,\beta^*}+h\Big(\mathbb{E}[|A_h^{\alpha^*,\beta^*}|^2]+c^{\alpha^*,\beta^*}+\theta_\kappa \tag{3.3.18}$$
$$-\frac{1}{4}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\geq\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)^{\mathrm{T}}(a^{\alpha^*,\beta^*})^{-}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\geq\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)\Big)\Big).$$

But, Notice that

$$\mathbb{E}[|A_h^{\alpha^*,\beta^*}|^2]=h^{-1}a^{-1}\cdot a^{\alpha^*,\beta^*}$$
$$+\frac{1}{4}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\geq\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big)^{\mathrm{T}}a^{\alpha^*,\beta^*-1}\Big(b^{\alpha^*,\beta^*}-\int_{\{1>|z|\geq\kappa\}}\eta^{\alpha^*,\beta^*}(z)\nu(dz)\Big).$$

By replacing $\mathbb{E}[|A_h^{\alpha^*,\beta^*}|^2]$ into (3.3.18), one obtains

$$\begin{aligned}\overline{\mathbf{T}}_{\kappa,h}[\varphi]-\overline{\mathbf{T}}_{\kappa,h}[\psi] &\leq |f|_\infty(1+h(c^{\alpha^*,\beta^*}+\theta_\kappa))\\ &\leq |f|_\infty(1+(C+\theta_\kappa)h),\end{aligned}$$

with $C=\sup_{\alpha,\beta}|c^{\alpha,\beta}|_\infty$. By changing the role of φ and ψ and implementing the same argument, one obtains

$$\big|\overline{\mathbf{T}}_{\kappa,h}[\varphi]-\overline{\mathbf{T}}_{\kappa,h}[\psi]\big|_\infty \leq |f|_\infty(1+(C+\theta_\kappa)h).$$

To prove that the family $(u^{\kappa,h})_h$ is bounded, we proceed by backward induction as in Lemma 3.14 in Chapter 2. By choosing in the first part of the proof $\varphi\equiv\bar{u}^{\kappa,h}(t_{i+1},.)$ and $\psi\equiv 0$, we see that

$$|u^{\kappa,h}(t_i,\cdot)|_\infty \leq h\overline{C}e^{\theta_\kappa(T-t_i)}+|u^{\kappa,h}(t_{i+1},\cdot)|_\infty(1+(C+\theta_\kappa)h),$$

where $\overline{C}:=\sup_{\alpha,\beta}|k^{\alpha,\beta}|_\infty$. It follows from the discrete Gronwall inequality that

$$|u^{\kappa,h}(t_i,\cdot)|_\infty\leq(\overline{C}(T-t_i)+|g|_\infty)e^{(C+\theta_\kappa)(T-t_i)}.$$

□

Define

$$\bar{v}^{\kappa,h}:=e^{-\theta_\kappa(T-t)}u^{\kappa,h}. \tag{3.3.19}$$

Next Corollary provides a bound for $v^{\kappa,h}$ uniformly on κ and h.

Corollary 3.3.2 *$\bar{v}^{\kappa,h}$ is bounded uniformly on h and κ, and*

$$|v^{\kappa,h}-\bar{v}^{\kappa,h}|_\infty \leq K\theta_\kappa^2 h \quad \textit{for some constant} \quad K.$$

If also, κ_h satisfies (3.3.7), then

$$\lim_{h\to0}|v^{\kappa_h,h}-\bar{v}^{\kappa_h,h}|_\infty = 0.$$

Proof. By Lemma 3.3.2 for fixed κ, we have:

$$|u^{\kappa,h}(t,.)|_\infty \leq (\overline{C}+|g|_\infty)e^{(C+\theta_\kappa)(T-t)}.$$

Therefore,

$$|\bar{v}^{\kappa,h}(t,.)|_\infty \leq (\overline{C}+|g|_\infty)e^{C(T-t)}.$$

For the next part, define $\bar{u}^{\kappa,h}(t,x)=e^{\theta_\kappa(T-t)}v^{\kappa,h}(t,x)$. Direct calculations shows that

$$\bar{u}^{\kappa,h}=e^{\theta_\kappa h}(1-\theta_\kappa h)\mathbb{E}\left[\bar{u}^{\kappa,h}\left(t+h,\hat{X}_h^{t,x,\kappa}\right)\right]+h\overline{F}_{\kappa,h}\left(t,x,\mathcal{D}_h\bar{u}^{\kappa,h},\bar{u}^{\kappa,h}(t+h,\cdot)\right).$$

By an argument similar to Lemma 3.19 in Chapter 2, we have

$$\begin{aligned}|(u^{\kappa,h}-\bar{u}^{\kappa,h})(t,\cdot)|_\infty &\leq \frac{1}{2}\theta_\kappa^2h^2|\bar{u}^{\kappa,h}(t+h,\cdot)|_\infty \\ &\quad +(1+(C+\theta_\kappa)h)|(u^{\kappa,h}-\bar{u}^{\kappa,h})(t+h,\cdot)|_\infty,\end{aligned} \tag{3.3.20}$$

where C is as in Lemma 3.3.2. By repeating the proof of Lemma 3.3.2 for $\bar{u}^{\kappa,h}$, one can conclude,

$$|\bar{u}^{\kappa,h}(t,\cdot)|_\infty \leq (\overline{C}+|g|_\infty)e^{(C+\theta_\kappa)(T-t)}(1+\frac{\theta_\kappa h}{2}).$$

So, by multiplying 3.3.20 by $e^{\theta_\kappa(T-t)}$, we have

$$\begin{aligned}|(\bar{v}^{\kappa,h}-v^{\kappa,h})(t,\cdot)|_\infty &\leq \frac{1}{2}\tilde{C}\theta_\kappa^2h^2e^{C(T-t)}(1+\frac{\theta_\kappa h}{2})e^{-\theta_\kappa h} \\ &\quad +e^{-\theta_\kappa h}(1+(C+\theta_\kappa)h)|(\bar{v}^{\kappa,h}-v^{\kappa,h})(t+h,\cdot)|_\infty,\end{aligned}$$

for some constant $\tilde{C}$. Because $e^{-\theta_\kappa h}(1+(C+\theta_\kappa)h)\leq e^{Ch}$, one can deduce from discrete Gronwall inequality that

$$|(\bar{v}^{\kappa,h}-v^{\kappa,h})(t,\cdot)|_\infty \leq K\theta_\kappa^2h,$$

for some constant K independent of κ which provides the second part of the theorem. □

We continue with the following consistency Lemma.

Lemma 3.3.3 *Let φ be a smooth function with the bounded derivatives. Then for all $(t,x)\in[0,T]\times\mathbb{R}^d$:*

$$\lim_{\substack{(t',x')\to(t,x)\\(h,c)\to(0,0)\\t'+h\leq T}}\frac{\varphi(t',x')-\mathbf{T}_{\kappa,h}[c+\varphi](t',x')}{h} = -\left(\mathcal{L}^X\varphi+F(\cdot,\varphi,D\varphi,D^2\varphi,\varphi(t,\cdot))\right)(t,x).$$

Proof. The proof is straightforward by Lebesgue dominated convergence Theorem. □

To complete the convergence argument, we need to proof the the approximate solution $v^{\kappa_h,h}$ converge to the final condition as

Lemma 3.3.4 *Let κ_h satisfy (3.3.7), then $\bar{v}^{\kappa_h,h}$ is uniformly Lipschitz with respect to x.*

Proof. We report the following calculation in the one-dimensional case $d = 1$ in order to simplify the presentation.
For fixed $t \in [0, T-h]$, we argue as in the proof of Lemma 3.3.2 to see that for $x, x' \in \mathbb{R}$ with $x > x'$:

$$\begin{aligned}
u^{\kappa,h}(t,x) - u^{\kappa,h}(t,x') &= \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}^{t,x}) - u^{\kappa,h}(t+h,\hat{X}^{t,x'})\Big)\Big] \\
&\quad + h\left(\inf_\alpha \sup_\beta \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\widehat{u^{\kappa,h}}](t+h,x) - \inf_\alpha \sup_\beta \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\widehat{u^{\kappa,h}}](t+h,x')\right) \\
&\le \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}^{t,x}) - u^{\kappa,h}(t+h,\hat{X}^{t,x'})\Big)\Big] \\
&\quad + h\left(\sup_\beta \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\widehat{u^{\kappa,h}}](t+h,x) - \inf_\alpha \bar{\mathcal{J}}_\kappa^{\alpha,\beta}[\widehat{u^{\kappa,h}}](t+h,x')\right).
\end{aligned}$$

Observe that by (3.3.6), one can write

$$\begin{aligned}
u^{\kappa,h}(t,x) - u^{\kappa,h}(t,x') \quad &\le \quad \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}^{t,x}) - u^{\kappa,h}(t+h,\hat{X}^{t,x'})\Big)\Big] \\
&\qquad + h\Big(\bar{\Phi}^{\alpha,\beta}[\widehat{u^{\kappa,h}},\widehat{u^{\kappa,h}}](t+h,x,x')\Big),
\end{aligned}$$

where $\bar{\Phi}$ is defined in the proof of Lemma 3.3.2. By Assumption **Inf–Sup**, there exists (α^*, β^*) such that

$$\bar{\Phi}^{\alpha^*,\beta^*}[\widehat{u^{\kappa,h}},\widehat{u^{\kappa,h}}](t+h,x,x') = \bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\widehat{u^{\kappa,h}}](t+h,x) - \bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\widehat{u^{\kappa,h}}](t+h,x').$$

Therefore,

$$\begin{aligned}
u^{\kappa,h}(t,x) - u^{\kappa,h}(t,x') &\le \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}^{t,x}) - u^{\kappa,h}(t+h,\hat{X}^{t,x'})\Big)\Big] \\
&\quad + h\Big(\bar{\mathcal{J}}_\kappa^*[\widehat{u^{\kappa,h}}](t+h,x) - \bar{\mathcal{J}}_\kappa^*[\widehat{u^{\kappa,h}}](t+h,x')\Big).
\end{aligned}$$

For the other in equality we do the same except that when we

$$u^{\kappa,h}(t,x) - u^{\kappa,h}(t,x') \quad \le \quad A + hB + hC,$$

where

$$\begin{aligned}
A \quad := \quad & \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}^{t,x}) - u^{\kappa,h}(t+h,\hat{X}^{t,x'})\Big)\Big] \\
& + h\left(\bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\widehat{u^{\kappa,h}}](t+h,x) - \bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\widehat{\tilde{u}^{\kappa,h}}](t+h,x)\right),
\end{aligned}$$

with $\tilde{u}^{\kappa,h}(y) = u^{\kappa,h}(y + x' - x)$,

$$B \;:=\; \bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\widehat{\tilde{u}^{\kappa,h}}](t+h,x) - \bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\widehat{u^{\kappa,h}}](t+h,x'),$$

and

$$C := \hat{\nu}_h^{\alpha^*,\beta^*,1}(u^{\kappa,h}(t+h,\cdot))(x) - \hat{\nu}_h^{\alpha^*,\beta^*,1}(u^{\kappa,h}(t+h,\cdot))(x').$$

We continue the proof in the following steps.

Step 1.

$$C = h^{-1}\mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}^{*,x}) - u^{\kappa,h}(t+h,\hat{X}^{*,x'})\Big)N_h^\kappa\Big],$$

where $\hat{X}^{*,x} := x + \sum_{i=1}^{N_h^\kappa} \eta^{\alpha^*,\beta^*}(x,Z_i)$ with Z_is are i.i.d. random variables distributed as $\frac{\nu(dz)}{\lambda_\kappa}$.

Step 2. By the definition of $\bar{\mathcal{J}}_\kappa^{\alpha,\beta}$,

$$\begin{aligned} B \;&=\; \frac{1}{2}(a^{\alpha^*,\beta^*}(x) - a^{\alpha^*,\beta^*}(x'))\mathcal{D}_h^2 u^{\kappa,h}(t+h,x') + (b_\kappa^{\alpha^*,\beta^*}(x) - b_\kappa^{\alpha^*,\beta^*}(x')) \\ &\quad \times \mathcal{D}_h^1 u^{\kappa,h}(t+h,x') + (c^{\alpha^*,\beta^*}(x) - c^{\alpha^*,\beta^*}(x'))\mathcal{D}_h^0 u^{\kappa,h}(t+h,x') \\ &\quad + k^{\alpha^*,\beta^*}(x) - k^{\alpha^*,\beta^*}(x'), \end{aligned}$$

where $b_\kappa^{\alpha,\beta}(x) := b^{\alpha,\beta}(x) - \int_{\{1>|z|\geq\kappa\}} \eta^{\alpha,\beta}(x,z)\nu(dz)$. On the other hand,

$$\mathcal{D}_h^k = \mathbb{E}\left[Du^{\kappa,h}(t+h,\hat{X}_h^{x'})\left(\frac{W_h}{h}\sigma^{-1}(x')\right)^{k-1}\right], \quad \text{for} \quad k = 1,2.$$

So,

$$\begin{aligned} B \leq \mathbb{E}\Big[&\frac{1}{2}(a^{\alpha^*,\beta^*}(x) - a^{\alpha^*,\beta^*}(x'))Du^{\kappa,h}(t+h,\hat{X}_h^{x'})\frac{W_h}{h}\sigma^{-1}(x') + (b_\kappa^{\alpha^*,\beta^*}(x) - b_\kappa^{\alpha^*,\beta^*}(x')) \\ &\times Du^{\kappa,h}(t+h,\hat{X}_h^{x'}) + (c^{\alpha^*,\beta^*}(x) - c^{\alpha^*,\beta^*}(x'))u^{\kappa,h}(t+h,\hat{X}_h^{x'})\Big] + f^{\alpha^*,\beta^*}(x) - f^{\alpha^*,\beta^*}(x'). \end{aligned}$$

Step 3. By the definition of $\bar{\mathcal{J}}_\kappa^{\alpha,\beta}$, one can observe that

$$\begin{aligned} &\bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[u^{\kappa,h}](t+h,x) - \bar{\mathcal{J}}_\kappa^{\alpha^*,\beta^*}[\tilde{u}^{\kappa,h}](t+h,x) \\ &\qquad = \frac{1}{2}a^{\alpha^*,\beta^*}(x)\delta^{(2)} + b_\kappa^*(x)\delta(1) + c_\kappa^*(x)\delta(0) \end{aligned}$$

where c_κ^* and b_κ^* are defined in the proof of Lemma 3.3.1, and

$$\delta^{(k)} \;=\; \mathbb{E}\left[D^k u^{\kappa,h}(t+h,\hat{X}_h^x) - D^k u^{\kappa,h}(t+h,\hat{X}_h^{x'})\right] \quad \text{for} \quad k = 0,1,2.$$

By Lemma 2.1 in Chapter 2, for $k = 1$ and 2

$$\begin{aligned}
\delta^{(k)} &= \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}_h^x) - u^{\kappa,h}(t+h,\hat{X}_h^{x'})\Big) H_h^k(t,x) \\
&\quad + u^{\kappa,h}(t+h,\hat{X}_h^{x'}) H_h^k(t,x)\left(1 - \frac{\sigma^k(x)}{\sigma^k(x')}\right)\Big] \\
&= \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}_h^x) - u^{\kappa,h}(t+h,\hat{X}_h^{x'})\Big) H_h^k(t,x) \\
&\quad + Du^{\kappa,h}(t+h,\hat{X}_h^{x'})\left(\frac{W_h}{h}\right)^{k-1}\sigma(x')\left(\sigma^{-k}(x) - \sigma^{-k}(x')\right)\Big].
\end{aligned}$$

Therefore, one can write

$$\begin{aligned}
A \;\le\; & \mathbb{E}\Big[\Big(u^{\kappa,h}(t+h,\hat{X}_h^x) - u^{\kappa,h}(t+h,\hat{X}_h^{x'})\Big) \\
&\quad \times\Big(1 - \bar{a}^* + \bar{a}^* N^2 + hc_\kappa^* + b_\kappa^* N\sqrt{h}\Big)(x) \\
&\quad + hb_\kappa^*(x') Du^{\kappa,h}(t+h,\hat{X}_h^{x'})\sigma(x')\left(\sigma^{-1}(x) - \sigma^{-1}(x')\right) \\
&\quad + a^*(x') Du^{\kappa,h}(t+h,\hat{X}_h^{x'})\sqrt{h} N\sigma(x')\left(\sigma^{-2}(x) - \sigma^{-2}(x')\right)\Big],
\end{aligned}$$

where $a^* := \frac{1}{2}a^{\alpha^*,\beta^*}$, $\bar{a}^* := \frac{1}{2}a^{-1}a^{\alpha^*,\beta^*}$, $c^* := c^{\alpha^*,\beta^*}$, $c_\kappa^* := c^* + \theta_\kappa$, and $b_\kappa^* := b_\kappa^{\alpha^*,\beta^*}$.

Step 4. By dividing both sides by $x - x'$ and taking the limit we have:

$$\begin{aligned}
Du^{\kappa,h}(t,x) \;\le\; & \mathbb{E}\Big[Du^{\kappa,h}(t+h,\hat{X}_h^x)\bigg(\Big(1 + h\tilde{\mu}_\kappa' + \sqrt{h}\sigma' N + \tilde{J}_{\kappa,h}\Big) \\
&\quad \times\Big(1 - \bar{a}^* + \bar{a}^* N^2 + hc_\kappa^* + b_\kappa^* N\sqrt{h}\Big) \\
&\quad + h\Big((b_\kappa^*)' - b_\kappa^*\frac{\sigma'}{\sigma}\Big) + \Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1} - a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}N\bigg) \\
&\quad + Du^{\kappa,h}(t+h,\hat{X}_h^{*,x})\left(1 + \mu^* h + \tilde{J}_{\kappa,h}'^*\right) N_h^\kappa\Big] + Ce^{\theta\kappa(T-t)}h,
\end{aligned}$$

where $\tilde{J}_{\kappa,h} := \int_{\{|z|>\kappa\}} \eta(z)\tilde{J}([0,h],dz)$, $\tilde{J}_{\kappa,h}'^* := \int_{\{|z|>\kappa\}} \eta'(z)\tilde{J}([0,h],dz)$, and N_h^κ is a Poisson process with intensity $\lambda_\kappa := \int_{\{|z|>\kappa\}} \nu(dz)$.

Let $L_t := |Du^{\kappa,h}(t,\cdot)|_\infty$. Then

$$\mathbb{E}\Big[Du^{\kappa,h}(t+h,\hat{X}_h^{*,x})\left(1 + \mu^* h + \tilde{J}_{\kappa,h}'^*\right) N_h^\kappa\Big] \;\le\; L_{t+h} Ch\left(\lambda_\kappa + \lambda_\kappa'^*\right),$$

where $\lambda_\kappa'^* := \int_{\{|z|>\kappa\}} \eta'^*(z)\nu(dz)$. Let $G := N + \frac{b_\kappa^*\sigma}{2}\sqrt{h}$. By the change of measure

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \exp\left(-\frac{(b_\kappa^*\sigma)^2}{4}h + \frac{b_\kappa^*\sigma}{2}\sqrt{h}N\right),$$

we have $G \sim \mathcal{N}(0,1)$ under $\mathbb{Q}$ and one can write

$$\begin{aligned} Du^{\kappa,h}(t,x) \quad \leq \quad & \mathbb{E}^{\mathbb{Q}}\Big[\frac{d\mathbb{P}}{d\mathbb{Q}} Du^{\kappa,h}(t+h,\hat{X}^x_h)\Big(\Big(1+h(\tilde{\mu}'_\kappa - \frac{b^*_\kappa\sigma}{2}) + \sqrt{h}\sigma' G + \tilde{J}_{\kappa,h}\Big) \\ & \times\Big(1-\bar{a}^* + \bar{a}^* G^2 + h(c^*_\kappa - \frac{(b^*_\kappa\sigma)^2}{2})\Big) \\ & + h\Big((b^*_\kappa)' - b^*_\kappa\frac{\sigma'}{\sigma} - \frac{b^*_\kappa\sigma}{2}\Big) + \Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1} - a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}G\Big)\Big] \\ & + L_{t+h}Ch\left(\lambda_\kappa + \lambda'^*_\kappa\right) + Ce^{\theta\kappa(T-t)}h, \end{aligned}$$

Step 5. Notice that $1-\bar{a}^* + a^*G^2 + h(c^*_\kappa - \frac{(b^*_\kappa\sigma)^2}{2})$ is positive and therefore, one can take $\frac{Z}{\mathbb{E}^{\mathbb{Q}}[Z]}$ as a density for the new measure $\mathbb{Q}^Z$. So,

$$\begin{aligned} Du^{\kappa,h}(t,x) \leq \mathbb{E}^{\mathbb{Q}^Z}\Big[& \frac{d\mathbb{P}}{d\mathbb{Q}} Du^{\kappa,h}(t+h,\hat{X}^x_h)\Big(\Big(1+h(\tilde{\mu}'_\kappa - \frac{b^*_\kappa\sigma}{2}) + \sqrt{h}\sigma' G + \tilde{J}_{\kappa,h}\Big) \\ & + Z^{-1}\Big(h\Big((b^*_\kappa)' - b^*_\kappa\frac{\sigma'}{\sigma} - \frac{b^*_\kappa\sigma}{2}\Big) + \Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1} - a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}G\Big)\Big)\Big] \\ & + L_{t+h}Ch\left(\lambda_\kappa + \lambda'^*_\kappa\right) + Ce^{\theta\kappa(T-t)}h. \end{aligned}$$

So,

$$\begin{aligned} Du^{\kappa,h}(t,x) \leq \mathbb{E}^{\mathbb{Q}^Z}\Big[&\Big(\frac{d\mathbb{P}}{d\mathbb{Q}}\Big)^2 (Du^{\kappa,h}(t+h,\hat{X}^x_h))^2\Big]^{\frac{1}{2}} \\ & \times\mathbb{E}^{\mathbb{Q}^Z}\Big[\Big(\Big(1+h(\tilde{\mu}'_\kappa - \frac{b^*_\kappa\sigma}{2}) + \sqrt{h}\sigma' G + \tilde{J}_{\kappa,h}\Big) \\ & + Z^{-1}\Big(h\Big((b^*_\kappa)' - b^*_\kappa\frac{\sigma'}{\sigma} - \frac{b^*_\kappa\sigma}{2}\Big) + \Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1} - a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}G\Big)\Big)^2\Big]^{\frac{1}{2}} \\ & + L_{t+h}Ch\left(\lambda_\kappa + \lambda'^*_\kappa\right) + Ce^{\theta\kappa(T-t)}h. \end{aligned}$$

Notice that

$$\mathbb{E}^{\mathbb{Q}^Z}\Big[\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big)^2 (Du^{\kappa,h}(t+h,\hat{X}^x_h))^2\Big] \quad \leq \quad L^2_{t+h}\exp(\frac{1}{4}(b^*_\kappa\sigma)^2 h).$$

On the other hand,

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}^Z}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}}\Big(\Big(1+h(\tilde{\mu}'_\kappa - \frac{b^*_\kappa\sigma}{2}) + \sqrt{h}\sigma' G + \tilde{J}_{\kappa,h}\Big) + Z^{-1}\Big(h\Big((b^*_\kappa)' - b^*_\kappa\frac{\sigma'}{\sigma} - \frac{b^*_\kappa\sigma}{2}\Big) \\ & \qquad + \Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1} - a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}G\Big)\Big)^2\Big] \\ & = \mathbb{E}\Big[Z\Big(\Big(1+h(\tilde{\mu}'_\kappa - \frac{b^*_\kappa\sigma}{2}) + \sqrt{h}\sigma' G + \tilde{J}_{\kappa,h}\Big) + Z^{-1}\Big(h\Big((b^*_\kappa)' - b^*_\kappa\frac{\sigma'}{\sigma} - \frac{b^*_\kappa\sigma}{2}\Big) \\ & \qquad + \Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1} - a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}G\Big)\Big)^2\Big]. \end{aligned}$$

By calculation of the right hand side of the above equality, one can observe that all the terms of order $\sqrt{h}$ vanish and we have:

$$\mathbb{E}^{\mathbb{Q}^Z}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}}\Big(\Big(1+h(\tilde{\mu}'_\kappa-\frac{b^*_\kappa\sigma}{2})+\sqrt{h}\sigma' G+\tilde{J}_{\kappa,h}\Big)+Z^{-1}\Big(h\Big((b^*_\kappa)'-b^*_\kappa\frac{\sigma'}{\sigma}-\frac{b^*_\kappa\sigma}{2}\Big)$$
$$+\Big(\frac{1}{2}(a^{\alpha^*,\beta^*})'\sigma^{-1}-a^{\alpha^*,\beta^*}\frac{\sigma'}{\sigma^2}\Big)\sqrt{h}G\Big)\Big)^2\Big]^{\frac{1}{2}}$$
$$\leq\Big(1+h\Big(c^*+\theta_\kappa-\frac{(b^*_\kappa)^2}{4a^*}-b^*_\kappa\sigma\sigma'+(b^*_\kappa)'-\frac{b^*_\kappa\sigma'}{\sigma}-\frac{b^*_\kappa\sigma}{2}+O(h\theta^2_\kappa)\Big)\Big)^{\frac{1}{2}}.$$

Therefore, by the choice of κ_h, for h small enough we have

$$L_t\leq L_{t+h}\exp\Big(\frac{1}{2}h(C+\theta\kappa_h-b\kappa_h{}^*\sigma\sigma'+(b^*\kappa_h)'-\frac{b\kappa_h{}^*\sigma'}{\sigma}-\frac{b^*_{\kappa_h}\sigma}{2}+2\lambda_{\kappa_h}+2\lambda'^*\kappa_h)\Big)$$
$$+Ce^{\theta_{\kappa_h}(T-t)}h$$
$$\leq L_{t+h}\exp\Big(h(C+\theta\kappa_h)\Big)+Ce^{\theta_{\kappa_h}(T-t)}h.$$

By discrete Gronwall inequality,

$$L_t \;\leq\; (|Dg|_\infty+C(T-t))e^{(\theta_{\kappa_h}+C)(T-t)}.$$

Therefore, by definition of $\bar{v}^{\kappa,h}$, we have

$$|D\bar{v}^{\kappa_h,h}|_1 \;\leq\; e^{C(T-t)}(|Dg|_\infty+C(T-t)).$$

□

Lemma 3.3.5 *Let κ_h satisfies (3.3.7), then*

$$\lim_{t\to T}\bar{v}^{\kappa,h}(t,x) \;=\; g(x).$$

Proof. We follow the same notations as in the proof of the previous Lemma and write

$$u^{\kappa,h}(t,x)\;=\;\mathbb{E}\Big[u^{\kappa,h}(t+h,\hat{X}^{t,x})\Big]+h\inf_\alpha\sup_\beta\bar{\mathcal{J}}^{\alpha,\beta}_\kappa[\widehat{u^{\kappa,h}}](t+h,x)$$
$$\leq\;\mathbb{E}\Big[u^{\kappa,h}(t+h,\hat{X}^{t,x})\Big]+h\sup_\beta\bar{\mathcal{J}}^{\alpha,\beta}_\kappa[\widehat{u^{\kappa,h}}](t+h,x).$$

Observe that by (3.3.6), one can write

$$u^{\kappa,h}(t,x)\;\leq\;\mathbb{E}\Big[u^{\kappa,h}(t+h,\hat{X}^{t,x})\Big]+h\Big(\bar{\Phi}^{\alpha,\beta}[\widehat{u^{\kappa,h}},0](t+h,x,x')\Big)+h\sup_{\alpha,\beta}|f^{\alpha,\beta}|_\infty,$$

By Assumption **Inf–Sup**, there exists (α^*,β^*) so that

$$u^{\kappa,h}(t,x)\;\leq\;\mathbb{E}\Big[u^{\kappa,h}(t+h,\hat{X}^{t,x})\Big]+h\bar{\mathcal{J}}^{\alpha^*,\beta^*}_\kappa[\widehat{u^{\kappa,h}}](t+h,x)+h\bar{C},$$

where $\bar{C}:=\sup_{\alpha,\beta}|f^{\alpha,\beta}|_\infty$. Therefore, for any $j=i,\cdots,n-1$ one can write

$$u^{\kappa,h}(t_j,\hat{X}^{t_i,x}_{t_j})\;\leq\;\mathbb{E}^{\mathbb{Q}}_{t_j}\Big[u^{\kappa,h}(t_{j+1},\hat{X}^{t_i,x}_{t_{j+1}})\Big(1-a^*_j+a^*_jG^2_j+hC^*_j\Big)\Big]+h\bar{C}.$$

where $a_j^* := \bar{a}^*(t_j, \hat{X}_{t_j}^{t_i,x})$, $C_j^* := (c_\kappa^* - \frac{(b_\kappa^*\sigma)^2}{2})(t_j, \hat{X}_{t_j}^{t_i,x})$ and G_js are independent standard Gaussian random variables under the new equivalent measure $\mathbb{Q}$. By the consecutive use of the above inequality and the fact that $1 - a_j^* + a_j^* G^2 + hC_j^*$ is positive, one can write

$$u^{\kappa,h}(t_i,x) \le \mathbb{E}^{\mathbb{Q}}\Big[g(\hat{X}_T^{t_i,x})\prod_{j=i}^{n-1}\Big(1 - a_j^* + a_j^* G^2 + hC_j^*\Big)\Big] + \bar{C}h\sum_{j=i}^{n-1} e^{\theta_\kappa t_j}.$$

Notice that in the above inequality we used the fact that

$$\mathbb{E}_{t_j}^{\mathbb{Q}}\Big[1 - a_j^* + a_j^* G_j^2 + hC_j^*\Big] \;=\; 1 + h\mathbb{E}_{t_j}^{\mathbb{Q}}[C_j^*] \le 1 + \theta_\kappa h.$$

On the other hand, $Z := \prod_{j=i}^{n-1}\Big(1 - a_j^* + a_j^* G_j^2 + hC_j^*\Big)$ is positive there for $\frac{Z}{E^{\mathbb{Q}}[Z]}$ could be considered as a density of a new measure $\mathbb{Q}^Z$ with respect to $\mathbb{P}$. Therefore,

$$u^{\kappa,h}(t_i,x) \le \mathbb{E}^{\mathbb{Q}}[Z]\mathbb{E}^{\mathbb{Q}^Z}\Big[g(\hat{X}_T^{t_i,x})\Big] + \bar{C}h\sum_{j=i}^{n-1} e^{\theta_\kappa t_j}.$$

By the definition of $\bar{v}^{\kappa,h}$, one can write

$$\bar{v}^{\kappa,h}(t_i,x) \le e^{-\theta_\kappa(T-t_i)}\mathbb{E}^{\mathbb{Q}}[Z]\mathbb{E}^{\mathbb{Q}^Z}\Big[g(\hat{X}_T^{t_i,x})\Big] + e^{-\theta_\kappa(T-t_i)}\bar{C}h\sum_{j=i}^{n-1} e^{\theta_\kappa t_j}.$$

Therefore,

$$\bar{v}^{\kappa,h}(t_i,x) - g(x) \le e^{-\theta_\kappa(T-t_i)}\mathbb{E}^{\mathbb{Q}}[Z]\mathbb{E}^{\mathbb{Q}^Z}\Big[|g(\hat{X}_T^{t_i,x}) - g(x)|\Big] + C|g(x)|(T-t_i) + e^{-\theta_\kappa(T-t_i)}\bar{C}(T-t_i).$$

Notice that $g(\hat{X}_T^{t_i,x}) - g(x)$ converges to zero $\mathbb{P}$–a.s. and therefore $\mathbb{Q}^Z$ a.s. as $(t_i,h)\to(T,0)$. So, by Lebesgue dominated convergence Theorem,

$$\limsup_{(t_i,h)\to(T,0)} \bar{v}^{\kappa,h}(t_i,x) - g(x) \le 0.$$

By the similar argument one can prove that:

$$\liminf_{(t_i,h)\to(T,0)} \bar{v}^{\kappa,h}(t_i,x) - g(x) \ge 0,$$

which completes the proof. □

Remark 3.3.8 By extending the above proof as in the Lemma 3.17 and Corollary 3.18 of Chapter 2, one can proof that

$$|\bar{v}^{\kappa,h}(t,x) - g(x)| \;\le\; C(T-t)^{\frac{1}{2}}.$$

Also, observe that by the similar argument as in Chapter 2, $v^{\kappa_h,h}$ is $\frac{1}{2}$–Hölder on t uniformly on h and x.

So, the approximate solution $\bar{v}^{\kappa_h,h}$ both satisfies the requirement of the convergence established in [6] and converges to a function v locally uniformly. Moreover, v is the unique viscosity solution of (3.1.1)–(3.1.2). So, by Corollary 3.3.2, the same assertion is true for $v^{\kappa,h}$.

3.3.2 Rate of Convergence

The proof of the rate of convergence for the non–local scheme is the same as the local case; Subsection 2.2.3. More precisely, The generalization of the method we used in Subsection 2.2.3 for the rate of convergence to non–local case, is developed in [12] and [13] where the scheme needs to be consistent and satisfies comparison principle. Therefore in this Subsection, we only present the results which enable us to apply the generalization in [12] and [13] to the scheme 3.3.3.

Before, providing consistency and comparison principle result for the scheme (3.3.15), we show that truncation error could be handled by the Theorem of continuous dependence for (3.1.1)–(3.1.2). More precisely, if v and v^κ are solutions of (3.1.1)–(3.1.2) and (3.3.9)–(3.3.10), respectively; then by Theorem 5.1 in [13]

$$|v - v^\kappa|_\infty \quad \leq \quad C\left(\int_{0<|z|<\kappa} |z|^2 \nu(dz)\right)^{\frac{1}{2}}.$$

Therefore, By choosing κ_h so that $\int_{0<|z|<\kappa_h} |z|^2\nu(dz) \leq Ch^{\frac{1}{2}}$, one can just concentrate on the rate of convergence of $v^{\kappa,h}$ to v^κ.

We shift to $\bar{v}^{\kappa_h,h}$ which is is derived from the strictly monotone scheme (3.3.15) and find the rate of convergence for $\bar{v}^{\kappa_h,h}$. The following Corollary shows that this shift do not effect the rate of convergence.

Corollary 3.3.3 *Let F which satisfies* **IHJB1**, *and* $F(t, x, 0, 0, 0, 0) = 0$. *Then,*

$$|\bar{v}^{\kappa_h,h} - v^{\kappa_h,h}| \leq Ch\theta^2_{\kappa_h}.$$

In addition, if κ_h is such that

$$\limsup_{h\to 0} h^{\frac{3}{4}} \theta^2_{\kappa_h} < \infty,$$

then

$$|\bar{v}^{\kappa_h,h} - v^{\kappa_h,h}| \leq Ch^{\frac{1}{4}}$$

Proof. The proof is straightforward by the proof of Lemma 3.3.2. □

Form now on, we concentrate on the approximate solution $\bar{v}^{\kappa,h}$ which is obtained from strictly monotone scheme 3.3.15 through (3.3.19). In order to provide the result, we need to use the consistency of the scheme for the regular approximate solutions. Then, the comparison principle for the scheme provides bounds over the difference between $u^{\kappa,h}$ and regular approximate solutions. Let

$$\begin{aligned}\mathcal{R}_{\kappa,h}[\psi](t,x) \quad := \quad & \frac{\psi(t,x) - \overline{\mathbf{T}}_{\kappa,h}[\psi](t,x)}{h} \; + \; \mathcal{L}^X\psi(t,x) \\ & + \overline{F}_\kappa(\cdot, \psi, D\psi, D^2\psi, \psi(t,\cdot))(t,x).\end{aligned}$$

Lemma 3.3.6 *For a family $\{\varphi_\varepsilon\}_{0<\varepsilon<1}$ of smooth functions satisfying*

$$\left|\partial_t^{\beta_0} D^\beta \varphi^\varepsilon\right| \leq C\varepsilon^{1-2\beta_0-|\beta|_1} \quad \text{for any} \quad (\beta_0,\beta)\in\mathbb{N}\times\mathbb{N}^d\setminus\{0\}, \tag{3.3.21}$$

where $|\beta|_1 := \sum_{i=1}^d \beta_i$, and $C>0$ is some constant, we have:

$$|\mathcal{R}_{\kappa,h}[\varphi_\varepsilon]|_\infty \;\leq\; R(h,\varepsilon) \;:=\; C\left(h\varepsilon^{-3} + h\theta_\kappa\varepsilon^{-1} + h\sqrt{\theta_\kappa} + \varepsilon^{-1}\int_{\{|z|\leq\kappa\}} |z|^2\nu(dz)\right),$$

for some constant $C>0$ independent of κ. If in addition

$$\limsup_{h\to 0} h\theta_{\kappa_h}^2 < \infty \quad \text{and} \quad \limsup_{h\to 0}\sqrt{h}\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)<\infty,$$

we have:

$$|\mathcal{R}_{\kappa_h,h}[\varphi_\varepsilon]|_\infty \;\leq\; R(h,\varepsilon) \;:=\; C\,(h\varepsilon^{-3}+\sqrt{h}\varepsilon^{-1}).$$

Proof. $\mathcal{R}_{\kappa,h}[\varphi_\varepsilon]$ is bounded by

$$\begin{aligned}\sup_\alpha\Big\{\Big|\mathbb{E}\Big[\frac{1}{h}\big(\varphi_\varepsilon(t+h,X_h^{t,x,\kappa})-\varphi_\varepsilon(t,x)\big)+\frac{1}{2}\mathrm{Tr}\left[a^\alpha(D^2\varphi_\varepsilon(t+h,X_h^{t,x,\kappa})-D^2\varphi_\varepsilon(t,x))\right]\\ +b^\alpha(D\varphi_\varepsilon(t+h,X_h^{t,x,\kappa})-D\varphi_\varepsilon(t,x))+(\theta_\kappa+c^\alpha)(\varphi_\varepsilon(t+h,X_h^{t,x,\kappa})-\varphi_\varepsilon(t,x))\\ +\mathcal{I}^\alpha[\varphi_\varepsilon](t,x)-\mathcal{I}^\alpha_{\kappa,h}[\varphi_\varepsilon](t+h,x)\Big]\Big|\Big\}\end{aligned}$$

For the Lévy integral term by Lemma 3.2.3, we have:

$$\begin{aligned}|\mathcal{I}^\alpha[\varphi_\varepsilon](t,x)-\mathcal{I}^\alpha_{\kappa,h}[\varphi_\varepsilon](t+h,x)| &\leq C\Big(|D\varphi_\varepsilon|_\infty\big(\sqrt{h}+h\int_{\{|z|>\kappa\}}|z|\nu(dz)\big)\\ &\quad +h|\partial_t D^2\varphi_\varepsilon|_\infty+|D^2\varphi_\varepsilon|_\infty\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)\Big)\\ &\leq C\left(h\varepsilon^{-3}+h\sqrt{\theta_\kappa}+\varepsilon^{-1}\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)\right).\end{aligned}$$

By the same argument as Lemma 3.22 in Chapter 2 all the other terms are bounded by $h\varepsilon^{-3}$ except

$$\theta_h\left(\varphi_\varepsilon(t+h,X_h^{t,x,\kappa})-\varphi_\varepsilon(t,x)\right)$$

which is bounded by $\theta_h h\varepsilon^{-1}$. The second assertion of the Lemma is straightforward. □

Next we need to have maximum principle for scheme 3.3.15. Note that Lemma 3.21 in Chapter 2 holds true for scheme 3.3.15 with $\beta=\theta_\kappa+C$ where $C=\sup_\alpha|c^\alpha|$. Therefore, Proposition 3.20 in Chapter 2 holds true for non–local case. More precisely, we have the following Proposition.

Proposition 3.3.2 *Let Assumption* **IHJB1** *holds true, and consider two arbitrary bounded functions φ and ψ satisfying:*

$$h^{-1}\left(\varphi - \overline{\mathbf{T}}_h[\varphi]\right) \ \le\ g_1 \quad \textit{and} \quad h^{-1}\left(\psi - \overline{\mathbf{T}}_h[\psi]\right) \ \ge\ g_2$$

for some bounded functions g_1 and g_2. Then, for every $i = 0, \cdots, n$:

$$(\varphi-\psi)(t_i,x) \quad \le \quad e^{(\theta_\kappa+C)}|(\varphi-\psi)^+(T,\cdot)|_\infty + (T-h)e^{(\theta_\kappa+C)(T-t_i)}|(g_1-g_2)^+|_\infty$$

where $C = \sup_\alpha |c^\alpha|$.

The approximation of the solution of non–local PDE by the Krylov method of shaking coefficients and switching system is developed in [12] who provides the result of rate of convergence of general monotone schemes for the non–local PDEs satisfying Assumption IHJB. In the regularity result for (3.3.22), provided by [13], it is proved that (v^i) is Lipschitz with respect to x and locally 1/2-Hölder continuous with respect to t. However, in the case of the scheme (3.3.3), we need the solution of (3.1.1)–(3.1.2) be uniformly $\frac{1}{2}$–Hölder continuous on t. It is because we need the regular approximate solutions obtained from Krylov method and switching solution to satisfy (3.3.21). Therefore, in the present work we need to rebuild Lemma 5.3 in [13] under the Assumption IHJB to obtain global $\frac{1}{2}$–Hölder continuous on t for the solution of the switching system.

Therefore, we continue this subsection by introducing the switching system of non–local PDEs with the regularity result needed for the solution of this system. Let k be a non-negative constant. Suppose the following system of PDEs:

$$\max\left\{-\mathcal{L}^X v_i(t,x) - F_i\left(t,x,v_i(t,x),Dv_i(t,x),D^2v_i(t,x),v_i(t,\cdot)\right), v_i - \mathcal{M}^i v\right\} = 0$$
$$v_i(T,\cdot) = g_i(\cdot), \qquad (3.3.22)$$

where $i = 1,\cdots,M$ and

$$\begin{aligned}
F_i(t,x,r,p,\gamma,\psi) &:= \inf_{\alpha\in\mathcal{A}_i}\left\{\mathcal{L}^\alpha(t,x,r,p,\gamma,\gamma) + \mathcal{I}^\alpha(t,x,r,p,\gamma,\psi)\right\}\\
\mathcal{L}^\alpha(t,x,r,p,\gamma,\gamma) &:= \frac{1}{2}\mathrm{Tr}\left[a^\alpha(t,x)\gamma\right] + b^\alpha(t,x)\cdot p + c^\alpha(t,x)r + k^\alpha(t,x)\\
\mathcal{I}^\alpha(t,x,r,p,\gamma,\psi) &:= \int_{\mathbb{R}^d_*}\left(\psi\left(t,x+\eta^\alpha(t,x,z)\right) - r - \mathbb{1}_{\{|z|\le 1\}}\eta^\alpha(t,x,z)\cdot p\right)d\nu(z)\\
\mathcal{M}^i r &:= \min_{j\neq i} r_j + k.
\end{aligned}$$

We would like to emphasize that g_is need to satisfy $g_i - \mathcal{M}^i\bar{g} \le 0$ where $\bar{g} = (g_1,\cdots,g_M)$. If for all i, $g_i = g$ then we obviously have $g_i - \mathcal{M}^i\bar{g} \le 0$.

Existence and comparison principle result for the above switching system is provided in Proposition 6.1 [12]. Also, it is known from Theorem 6.3 in [12], that if $(v^1,\cdots,v^M)$ and v be respectively the solutions of (3.3.22) and (3.1.1)-(3.1.2) with $\mathcal{A} = \cup_{i=1}^M \mathcal{A}_i$ and $\mathcal{A}_i$s are disjoint sets, then

$$0 \le v^i - v \le Ck^{\frac{1}{3}} \quad \text{for} \quad i = 1,\cdots,M. \qquad (3.3.23)$$

The following Lemma provide the uniform 1/2-Hölder continuity for (v^i).

Lemma 3.3.7 *Assume* **IHJB2** *holds for each i and let (v^i) be the viscosity solution of (3.3.22). Then there exist a constant C such that for any $i = 1, \cdots, M$:*

$$\left|v^i\right|_1 \quad \leq \quad C.$$

Proof. Lipschitz continuity with respect to x is due to Lemma 5.2 in [13]. To obtain uniform $1/2-$Hölder continuity with respect to t, we modify the the proof of Lemma 5.3 in [13] by using assumption **IHJB2**.
Fix $y \in \mathbb{R}^d$ and $t' > 0$. Let $t \in \mathbb{R}_+$ be such that $t \leq t'$. For each $i = 1, \cdots, M$, define:

$$\psi_i(t,x) \quad := \quad \lambda \frac{L}{2}\left[e^{A(t'-t)}|x-y|^2 + B(t'-t)\right] + K(t'-t) + \lambda^{-1}\frac{L}{2} + v^i(t',y)$$

Where $L = \frac{1}{2}|v|_1$ and λ, a and γ will be defined later. Then:

$$\begin{aligned}
\partial_t \psi_i(t,x) &= -\lambda \frac{L}{2}\left(Ae^{A(t'-t)}|x-y|^2 + B\right) - K \\
D\psi_i(t,x) &= 2\lambda L e^{A(t'-t)}(x-y) \\
D^2\psi(t,x) &= \lambda L e^{A(t'-t)} I_{d\times d}.
\end{aligned}$$

So,

$$\begin{aligned}
&-\partial_t\psi_i - \inf_{\alpha\in\mathcal{A}}\left\{\mathcal{L}^\alpha(t,x,\psi_i,D\psi_i,D^2\psi_i) + \mathcal{I}^\alpha(t,x,\psi_i,D\psi_i)\right\} \\
&= \lambda L\left(Ae^{A(t'-t)}|x-y|^2 + B\right) + K - \inf_{\alpha\in\mathcal{A}}\left\{\frac{1}{2}\lambda L e^{A(t'-t)}\mathrm{Tr}\left[a^\alpha(t,x)\right]\right. \\
&\quad + 2\lambda L e^{A(t'-t)} b^\alpha(t,x)\cdot(x-y) + c^\alpha(t,x)\psi_i + k^\alpha(t,x) + \lambda\frac{L}{2}e^{A(t'-t)} \\
&\quad \left.\times \int_{\mathbb{R}^d_*}\left(|x+\eta^\alpha(t,x,z)-y|^2 - |x-y|^2 - 2\mathbb{1}_{\{|z|\leq 1\}}\eta^\alpha(t,x,z)\cdot(x-y)\right)d\nu(z)\right\}.
\end{aligned}$$

By **IHJB2**, we can choose K and λ so that,

$$\begin{aligned}
|a^\alpha|_\infty \leq K, |b^\alpha|_\infty \leq K, |c^\alpha|_\infty \leq K, |k^\alpha|_\infty \leq K, K^{-1} \leq \lambda \leq K \\
|v|_\infty \leq K, |\eta^\alpha(t,x,z)| \leq K(1\wedge|z|).
\end{aligned}$$

Without loss of generality and with the similar argument as in Remark 2.2.7, we can suppose that for any α, $c^\alpha \leq 0$. So by choosing positive large A, there exists non–negative constants C_1, C_2, C_3 and C_4 such that:

$$\begin{aligned}
&-\partial_t\psi_i - \inf_{\alpha\in\mathcal{A}}\left\{\mathcal{L}^\alpha(t,x,\psi_i,D\psi_i,D^2\psi_i) + \mathcal{I}^\alpha(t,x,\psi_i,D\psi_i)\right\} \\
&\geq \quad \lambda L e^{A(t'-t)} K\left(\left(\frac{A}{K} - \frac{1}{2}\right)|x-y|^2 - C_1|x-y| + C_2 B - C_3\right) - C_4.
\end{aligned}$$

Therefore, choice of large B and A makes the right hand side non–negative.

$$-\partial_t\psi_i - \inf_{\alpha\in\mathcal{A}}\left\{\mathcal{L}^\alpha(t,x,\psi_i,D\psi_i,D^2\psi_i) + \mathcal{I}^\alpha(t,x,\psi_i,D\psi_i)\right\} \geq 0.$$

On the other hand,

$$\psi(t',x) \;\; = \;\; \frac{L}{2}\left(\lambda|x-y|^2+\lambda^{-1}\right)+v^i(t',y).$$

Minimizing with respect to λ,

$$\psi(t',x) \geq L|x-y|+v^i(t',y) \geq v^i(t',x).$$

We can conclude that ψ_i is a super solution of (3.3.22). So, by comparison Theorem in [12],

$$\psi_i(t,y) \geq v^i(t,y).$$

So,

$$\frac{L}{2}\left(\lambda B(t'-t)+\lambda^{-1}\right)+v^i(t',y) \geq v^i(t,y).$$

Therefore, for $\lambda=(t'-t)^{-\frac{1}{2}}$ we have

$$v^i(t,y)-v^i(t',y) \leq C\sqrt{t'-t}.$$

The other inequality can be done similarly by choosing:

$$\psi_i(t,x) \;\; := \;\; -\lambda\frac{L}{2}\left[e^{A(t'-t)}|x-y|^2-B(t'-t)\right]-K(t'-t)-\lambda^{-1}\frac{L}{2}+v^i(t',y).$$

$\square$

Remark 3.3.9 Notice that all the result of switching system is correct for (3.1.1)-(3.1.2) satisfying **IHJB2** by simply setting $M=1$ and $k=0$.

Therefore, by [12] there are regular functions $\underline{w}^\kappa_\varepsilon$ and $\overline{w}^\kappa_\varepsilon$ which are respectively the regular sub– and super–solution of

$$\begin{aligned}
&-\mathcal{L}^X u^\kappa(t,x)-\overline{F}_\kappa\left(t,x,u^\kappa(t,x),Du^\kappa(t,x),D^2u^\kappa(t,x),u^\kappa(t,\cdot)\right)=0, && \text{on } [0,T)\times\mathbb{R}^d,\\
&u^\kappa(T,\cdot)=g, && \text{on } \in\mathbb{R}^d.
\end{aligned}$$

where

$$\overline{F}_\kappa(t,x,r,p,\gamma,\psi) \;\; := \;\; \inf_{\alpha\in\mathcal{A}}\left\{\mathcal{L}^\alpha(t,x,r,p,\gamma)+\mathcal{I}^\alpha_\kappa(t,x,r,p,\gamma,\psi)\right\}$$

(one can replace sup inf by inf sup) where

$$\mathcal{L}^\alpha(t,x,r,p,\gamma) \;\; := \;\; \frac{1}{2}\mathrm{Tr}\left[\sigma^\alpha\sigma^{\alpha\mathrm{T}}(t,x)\gamma\right]+b^\alpha(t,x)p+(c^\alpha(t,x)+\theta_\kappa)r,$$

and

$$\mathcal{I}^\alpha_\kappa(t,x,r,p,\gamma,\psi):=\int_{\{|z|>\kappa\}}\left(\psi\big(x+\eta^\alpha(t,x,z)\big)-r-\mathbb{1}_{\{|z|\leq 1\}}\eta^\alpha(t,x,z)\cdot p\right)\nu(dz).$$

Then, by Proposition 6.2 and Theorem 6.3 of [12], Lemma 3.3.6 and Proposition 3.3.2,

$$\begin{aligned}(u^\kappa - u^{\kappa,h})(t,x) &\leq (u^\kappa - \underline{w}^\kappa_\varepsilon + \underline{w}^\kappa_\varepsilon - u^{\kappa,h})(t,x)\\ &\leq Ce^{(\theta_\kappa+C_1)(T-t)}\left(\varepsilon + h\varepsilon^{-3} + h\theta_\kappa\varepsilon^{-1} + h\sqrt{\theta_\kappa} + \varepsilon^{-1}\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)\right)\end{aligned}$$

and

$$\begin{aligned}(u^{\kappa,h} - u^\kappa)(t,x) &\leq (u^{\kappa,h} - \overline{w}^\kappa_\varepsilon + \overline{w}^\kappa_\varepsilon - u^\kappa)(t,x)\\ &\leq Ce^{(\theta_\kappa+C_1)(T-t)}\left(\varepsilon^{\frac{1}{3}} + h\varepsilon^{-3} + h\theta_\kappa\varepsilon^{-1} + h\sqrt{\theta_\kappa} + \varepsilon^{-1}\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)\right).\end{aligned}$$

Notice that $v^\kappa(t,x) = e^{-\theta_\kappa(T-t)}u^\kappa(t,x)$. So,

$$v^\kappa - \bar{v}^{\kappa,h} \leq C\left(\varepsilon + h\varepsilon^{-3} + h\theta_\kappa\varepsilon^{-1} + h\sqrt{\theta_\kappa} + \varepsilon^{-1}\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)\right)$$

and

$$\bar{v}^{\kappa,h} - v^\kappa \leq C\left(\varepsilon^{\frac{1}{3}} + h\varepsilon^{-3} + h\theta_\kappa\varepsilon^{-1} + h\sqrt{\theta_\kappa} + \varepsilon^{-1}\int_{\{|z|\leq\kappa\}}|z|^2\nu(dz)\right).$$

On the other hand, because of (3.3.14) and by Lemma (3.3.2), the second part of Theorem 3.3.2 is provided after choice of optimal ε.

3.4 Conclusion

The scheme presented in this Chapter is the first probabilistic numerical method for fully non–linear non–local problems. As in local case (Chapter 2), it converges to the viscosity solution of the problem and a rate of convergence is known for the convex (concave) non–linearities. Moreover, with the same argument as in Section 4 in Chapter 2, Monte Carlo approximations of expectations inside the scheme do not affect the asymptotic results if enough number of samples would be used. The error analysis for MCQ shows that the appropriate approximation of jump–diffusion process with compound Poisson process could be applied in discretization procedure. On the other hand there are some features where the scheme is not implementable in non–local case, e.g. when the non–linearity is of HJB type. This could be the challenge of future works.

Chapter 4

Optimal Production Policy under the Carbon Emission Market

In this chapter, we analyze the effect of emission market in reducing the carbon emission through the change on production policy of the relevant firms. The firm's objective is to maximize her utility on her wealth which is made of both the profit of her production and the value of her carbon allowance portfolio over her production and her portfolio strategy. We solve the utility maximization problem on portfolio strategy by the duality argument and then on the production by the use of Hamilton–Jacobi–Bellman (HJB) equations[1].

4.1 Small producer with one-period carbon emission market

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space endowed with a one-dimensional Brownian motion W. We denote by $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ the completed canonical filtration of the Brownian motion W, and by $\mathbb{E}_t := \mathbb{E}[\cdot|\mathcal{F}_t]$ the conditional expectation operator given $\mathcal{F}_t$.

We consider a production firm with preferences described by the utility function $U : \mathbb{R} \longrightarrow \mathbb{R} \cup \{\infty\}$ assumed to be strictly increasing, strictly concave and C^1 over $\{U < \infty\}$. We denote by $\pi_t(\omega, q)$ the (random) time t rate of profit of the firm for a production rate q. Here $\pi : \mathbb{R}_+ \times \Omega \times \mathbb{R}_+ \to \mathbb{R}$ is an $\mathbb{F}-$progressively measurable map. As usual we shall omit ω from the notations. For fixed (t, ω), we assume that the function $\pi_t(\cdot) := \pi(t, \cdot)$ is strictly concave, C^1 in q and satisfies

$$\pi_t'(0+) > 0 \quad \text{and} \quad \pi_t'(\infty) < 0.$$

Let us denote by $e_t(q_t)$ the rate of carbon emissions generated by a production rate q. Here, $e_{\cdot}(.) : \Omega \times [0, T] \times \mathbb{R}_+$ is an $\mathbb{F}-$progressively measurable map and C^1 in $q \in \mathbb{R}_+$. Then the total quantity of carbon emissions induced by a production policy $\{q_t, t \in [0, T]\}$ is given by

$$E_T^q := \int_0^T e_t(q_t) dt. \tag{4.1.1}$$

The aim of the carbon emission market is to incur this cost to the producer so as to obtain an overall reduction of the carbon emissions.

[1] This work is reported on a paper co–authored with Redouane Belaouar and Nizar Touzi.

From now on, we analyze the effect of the presence of the carbon emission market within the *cap-and-trade* scheme.

In order to model the carbon emission market, we introduce an (unobservable) state variable Y defined by the dynamics:

$$dY_t \quad = \quad \mu_t dt + \gamma_t dW_t, \tag{4.1.2}$$

where μ and γ are two bounded $\mathbb{F}-$adapted processes and $\gamma > 0$.

We assume that there is one single period $[0, T]$ during which the carbon emission market is in place. At each time $t \geq 0$, the random variable Y_t indicates the market view of the cumulated carbon emissions. At time T, $Y_T \geq \kappa$ (resp. $Y_T < \kappa$) means that the cumulated total emission have (resp. not) exceeded the quotas κ, fixed by the trading scheme. Let α be the penalty per unit of carbon emission. Then, the value of the carbon emission contract at time T is:

$$S_T \quad := \quad \alpha \mathbb{1}_{\{Y_T \geq \kappa\}}.$$

The carbon emission allowance could be viewed as a derivative security defined by the above payoff. The carbon emission market allows for trading this contract in continuous-time throughout the time period $[0, T]$. Assuming that the market is frictionless, it follows from the classical no-arbitrage valuation theory that the price of the carbon emission contract at each time t is given by

$$S_t \;:=\; \mathbb{E}_t^{\mathbb{Q}}\left[S_T\right] \;=\; \alpha \mathbb{Q}_t\left[Y_T \geq \kappa\right], \tag{4.1.3}$$

where $\mathbb{Q}$ is a probability measure equivalent to $\mathbb{P}$, the so-called equivalent martingale measure, $\mathbb{E}_t^{\mathbb{Q}}$ and $\mathbb{Q}_t$ denote the conditional expectation and probability given $\mathcal{F}_t$. Given market prices of the carbon allowances, the risk-neutral measure may be inferred from the market prices. Since the market is frictionless, the value of the initial holdings in (free) allowances, $E^{\max}$, can be expressed equivalently in terms of their value in cash $S_0 E^{\max}$.

In the present context, and in contrast with the standard taxation (Remark 4.1.1), production firms have a clear incentive to reduce emissions as they have the possibility to sell their allowances on the emission market. Hence, the financial market induces a mutualization of carbon emissions, and there is no incentive to merge for the single objective of avoiding the carbon taxes. We will see however that large producers can have a negative impact.

We now formulate the objective function of the firm in the presence of the emission market. The primary activity of the firm is the production modeled by the rate q_t at time t. This generates a gain $\pi_t(q_t)$. The resulting carbon emissions are given by $e_t(q_t)$. Given that the price of the externality is available on the market, the profit on the time interval $[0, T]$ is given by:

$$\int_0^T \pi_t(q_t) dt - S_T \int_0^T e_t(q_t) dt. \tag{4.1.4}$$

In addition to the production activity, the company trades continuously on the carbon emissions market. Let $\{\theta_t, t \geq 0\}$ be an $\mathbb{F}-$adapted process which is $S-$integrable. For every $t \geq 0$, θ_t indicates the number of contracts of carbon emissions held by the company at time t. Under the self-financing condition, the wealth accumulated by trading on the emission market is:

$$x + \int_0^T \theta_t dS_t, \tag{4.1.5}$$

where x is the sum of the initial capital of the company and the market value of its free emission allowances contracts. By (4.1.4) and (4.1.5), together with an integration by parts, the total wealth of the firm at time T is

$$X_T^\theta + B_T^q \tag{4.1.6}$$

where

$$X_T^\theta := x + \int_0^T \theta_t dS_t, \qquad B_T^q := \int_0^T (\pi_t(q_t) - S_t e_t(q_t))\, dt - \int_0^T E_t^q dS_t,$$

and

$$E_t^q := \int_0^t e_u(q_u) du, \quad \text{for all} \quad t \in [0, T].$$

We assume that the firm is allowed to trade without any constraint. Then, the objective of the manager is:

$$V^{(1)} \quad := \quad \sup\left\{\mathbb{E}\left[U\left(X_T^\theta + B_T^q\right)\right] : \ \theta \in \mathcal{A}, \ q \in \mathcal{Q}\right\}, \tag{4.1.7}$$

where $\mathcal{A}$ is the collection of all $\mathbb{F}-$adapted processes such that the process X is bounded from below by a martingale, and $\mathcal{Q}$ is the collection of all non-negative $\mathbb{F}-$adapted processes.

Notice that the stochastic integrals with respect to S can be collected together in the expression of $X_T^\theta + B_T^q$. since $\mathcal{A}$ is a linear subspace, it follows that the maximization with respect to q and θ are completely decoupled, this problem is easily solved by optimizing successively with respect to q and θ. The partial maximization with respect to q provides an optimal production level $q^{(1)}$ defined by the first order condition:

$$\frac{\partial \pi_t}{\partial q}(q_t^{(1)}) \quad = \quad S_t \frac{\partial e_t}{\partial q}(q_t^{(1)}). \tag{4.1.8}$$

Because of the assumptions on $\pi_t(.)$ and $e_t(.)$, we deduce immediately that $q_t^{(1)}$ is less than the optimal production of the firm in the absence of any restriction on the emission, meaning that the emission market leads to a reduction of the production, and therefore a reduction of the carbon emissions.

We next turn to the optimal trading strategy by solving:

$$\sup_{\theta} \mathbb{E}\left[U\left(X_T^{x,\theta-E^{q^{(1)}}} + B^{q^{(1)}}\right)\right] \quad \text{where} \quad B^q := \int_0^T \left(\pi_t(q_t) - S_t e_t(q_t)\right) dt.$$

In the present context of a complete market, the solution is given by:

$$x + \int_0^T \left(\theta_t^{(1)} - E_t^{q^{(1)}}\right) dS_t + B^{q^{(1)}} \quad = \quad (U')^{-1}\left(y^{(1)} \frac{d\mathbb{Q}}{d\mathbb{P}}\right)$$

where the Lagrange multiplier $y^{(1)}$ is defined by:

$$\mathbb{E}^{\mathbb{Q}}\left[(U')^{-1}\left(y^{(1)} \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] \quad = \quad x + \mathbb{E}^{\mathbb{Q}}\left[B^{q^{(1)}}\right].$$

Let us sum up the present context of a small firm:

- the trading activity of the company has no impact on its optimal production policy,
- the firm's optimal production $q^{(1)}$ is smaller than that of the business-as-usual situation, so that the emission market is indeed a good tool for the reduction of carbon emissions,
- the emission market assigns a price to the externality that the firm manager can use in order to optimize his production scheme.

Remark 4.1.1 Let us examine the case where there is no possibility to trade the carbon emission allowances. This is the standard taxation system where α is the amount of tax to be paid at the end of period per unit of carbon emission. Assuming again that the firm's horizon coincides with this end of period, its objective is:

$$V_0 \quad := \quad \sup_{q.\in\mathcal{Q}} \mathbb{E}\left[U\left(\int_0^T \pi_t(q_t)dt - \alpha\left(E_T^q - E^{\max}\right)^+\right)\right]$$

where $E^{\max}$ is the free allowances of the market. Direct calculation leads to the following characterization of the optimal production level:

$$\frac{\partial \pi_t}{\partial q}\left(q_t^{(0)}\right) \quad = \quad \alpha \frac{\partial e_t}{\partial q}\left(q_t^{(0)}\right) \mathbb{E}_t^{\mathbb{Q}^{(0)}}\left[\mathbb{1}_{\mathbb{R}_+}\left(E_T^{q_t^{(0)}} - E^{\max}\right)\right] \tag{4.1.9}$$

where

$$\frac{d\mathbb{Q}^{(0)}}{d\mathbb{P}} \quad = \quad \frac{U'\left(\int_0^T \pi_t(q_t^{(0)})dt - \alpha\left(E_T^{q_t^{(0)}} - E^{\max}\right)^+\right)}{\mathbb{E}\left[U'\left(\int_0^T \pi_t(q_t^{(0)})dt - \alpha\left(E_T^{q_t^{(0)}} - E^{\max}\right)^+\right)\right]}. \tag{4.1.10}$$

The natural interpretation of (4.1.9) and(4.1.10) is that the production firm assigns an individual price to its emissions:

$$S_t \quad := \quad \alpha \mathbb{E}_t^{\mathbb{Q}^{(0)}} \left[\mathbb{1}_{\mathbb{R}_+} \left(E_T^{q_t^{(0)}} - E^{\max} \right) \right], \tag{4.1.11}$$

i.e. the expected value of the amount of tax to be paid under the measure $\mathbb{Q}^{(0)}$ defined by her marginal utility as a density. The probability measure $\mathbb{Q}^{(0)}$ is the so-called risk-neutral measure in financial mathematics, or the stochastic discount factor of the firm. Given this evaluation, the firm optimizes her adjusted profit function, $\pi_t(q) - e_t(q)S_t$;

$$\frac{\partial \pi_t}{\partial q}(q^{(0)}) \quad = \quad \frac{\partial e_t}{\partial q}(q^{(0)})S_t.$$

We continue by commenting on the optimal production policy defined by (4.1.9)-(4.1.10):

- assuming that the firms know the nature of their utility functions, the system of equations (4.1.9)-(4.1.10) is still a nontrivial nonlinear fixed point problem.
- This problem would be considerably simplified if the manager were to know the market price for carbon emissions (4.1.11). But of course, in the present context, this is an individual subjective price which is not quoted on any financial market.
- The present situation, based on a classical taxation policy, offers no incentive to reduce emissions beyond $E^{\max}$. Indeed, if the optimal production in the absence of taxes produces carbon emissions below the level $E^{\max}$, then it is indeed the same as the business-as-usual situation. So, the taxation does not contribute to reduce the carbon emissions. As a consequence, the only way to benefit from having carbon emissions below the level $E^{\max}$ is to merge with another firm whose emissions are above its given free emissions allowances. Hence, such a policy puts a clear incentive to mergers. □

The emission market provides an evaluation of the externality of carbon emissions by firms. Given this information there is no more need to know precisely the utility function of the firm in order to solve the nonlinear system (4.1.9)-(4.1.10). The quoted price of the externality is then very valuable for the managers as it allows them to better optimize their production scheme.

4.2 Large producer with one-period carbon emission market

In this section, we consider the case of a large carbon emitting production firm. We shall see that this leads to different considerations as the trading activity will have an impact on the production policy of the company.

We model this situation by assuming that the state variable Y is affected by the production policy of the firm:

$$dY_t^q = (\mu_t + \beta e_t(q_t))\, dt + \gamma_t dW_t \tag{4.2.1}$$

where $\beta > 0$ is a given impact coefficient. The price process S of the carbon emission allowances is, as in the previous section, given by the no-arbitrage valuation principle:

$$S_t^q = \alpha \mathbb{Q}_t^q \left[Y_T^q \geq \kappa \right], \tag{4.2.2}$$

and is also affected by the production policy q. The equivalent martingale measure $\mathbb{Q}^q$ is defined by

$$\left.\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \exp\left(-\int_0^T \lambda_t(q_t) dW_t - \frac{1}{2}\int_0^T \lambda_t(q_t)^2 dt \right) \tag{4.2.3}$$

where $\lambda : \mathbb{R}_+ \times \Omega \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ is an $\mathbb{F}-$progressively measurable map. The dynamics of the price process S are given by

$$\frac{dS_t^q}{S_t^q} = \sigma_t^q \left(dW_t + \lambda_t(q_t) dt \right), \; t < T, \tag{4.2.4}$$

where the volatility function σ_t^q is progressively measurable and depends on the control process $\{q_s, 0 \leq s \leq T\}$. As in the previous section, the wealth process of the company is given by:

$$X_T^{x,\theta} := x + \int_0^T \theta_t dS_t^q \quad \text{and} \quad B_T^q := \int_0^T \pi_t(q_t) dt - S_T^q \int_0^T e_t(q_t) dt$$

4.2.1 Large Carbon emission with no impact on risk premia

In this subsection, we restrict our attention to the case of large emitting firm with no impact on the risk premia, i.e.

$$\lambda_t(q) \quad \text{is independent of } q \text{ for any } t \geq 0. \tag{4.2.5}$$

The objective of the large emitting firm is:

$$V_0^{(2)} := \sup_{q.\in\mathcal{Q},\; \theta\in\mathcal{A}} \mathbb{E}\left[U\left(X_T^{x,\theta} + B_T^q \right) \right].$$

Proposition 4.2.1 *Assume (4.2.5), and that the market is complete with unique risk-neutral measure $\mathbb{Q}$. Then, the optimal production policy is independent of the utility function of the producer U, and obtained by solving:*

$$\sup_{q.\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\left[B_T^q \right]. \tag{4.2.6}$$

Moreover, if $q^{(2)}$ is an optimal production scheme, then the optimal investment strategy $\theta^{(2)}$ is characterized by

$$X_T^{x,\theta^{(2)}} + B_T^{q^{(2)}} = (U')^{-1}\left(y^{(2)} \frac{d\mathbb{Q}}{d\mathbb{P}} \right), \quad x + \mathbb{E}^{\mathbb{Q}}\left[B_T^{q^{(2)}} \right] = \mathbb{E}^{\mathbb{Q}}\left[(U')^{-1}\left(y^{(2)} \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]. \tag{4.2.7}$$

Proof. We first fix some production strategy q. Since the market is complete, the partial maximization with respect to θ can be performed by the classical duality method:

$$X_T^{x,\theta^q} + B_T^q = (U')^{-1}\left(y^q \frac{d\mathbb{Q}}{d\mathbb{P}}\right), \tag{4.2.8}$$

where the Lagrange multiplier y^q is defined by

$$\mathbb{E}^{\mathbb{Q}}\left[(U')^{-1}\left(y^q \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] \;=\; x + \mathbb{E}^{\mathbb{Q}}\left[B_T^q\right]. \tag{4.2.9}$$

This reduces the problem to:

$$\sup_{q.\geq 0} \mathbb{E}\left[U \circ (U')^{-1}\left(y^q \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]. \tag{4.2.10}$$

Notice that $U \circ (U')^{-1}$ is decreasing and the density $\frac{d\mathbb{Q}}{d\mathbb{P}} > 0$. Then (4.2.10) reduces to

$$\inf\left\{y^q : \; q. \geq 0\right\}.$$

Since $(U')^{-1}$ is also decreasing, (4.2.9) converts the problem into

$$\sup\left\{\mathbb{E}^{\mathbb{Q}}\left[B_T^q\right] : \; q. \in \mathcal{Q}\right\}.$$

Finally, given the optimal strategy $q^{(2)}$, the optimal investment policy is characterized by (4.2.8). □

In order to push further the characterization of the optimal production policy $q^{(2)}$, we specialize the discussion to the Markov case by assuming that $\pi_t(q) = \pi(t, q_t)$, $e_t(q) = e(t, q_t)$, and $\lambda_t(q) = \lambda(t)$ for some deterministic functions $\pi, e : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ in $C^{0,1}(\mathbb{R}_+ \times \mathbb{R}_+)$, $\lambda : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ in $C^0(\mathbb{R}_+)$, and

$$dY_t^q = \left(\mu(t, Y_t^q) + \beta e(t, q_t)\right) dt + \gamma(t, Y_t^q) dW_t,$$

for some continuous deterministic functions $\mu, \gamma : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}$.

The state variable E is now defined by the dynamics

$$dE_t^q \;=\; e(t, q_t)dt \tag{4.2.11}$$

which records the cumulated carbon emissions of the company. The dynamic version of the producer planning problem (4.2.6) is given by:

$$V^{(2)}(t, e, y) \;:=\; \sup_{q. \in \mathcal{Q}} \mathbb{E}_{t,e,y}^{\mathbb{Q}}\left[\int_t^T \pi(t, q_t)dt - \alpha E_T^q \mathbb{1}_{\{Y_T^q > 0\}}\right]. \tag{4.2.12}$$

Then, $V^{(2)}$ solves the dynamic programming equation:

$$\begin{aligned} 0 \;=\; & \frac{\partial V^{(2)}}{\partial t} + (\mu - \lambda\gamma)V_y^{(2)} + \frac{1}{2}\gamma^2 V_{yy}^{(2)} \\ & + \max_{q \geq 0}\left\{\pi(t, q) + e(t, q)V_e^{(2)} + \beta e(t, q)V_y^{(2)}\right\}, \end{aligned} \tag{4.2.13}$$

together with the terminal condition

$$V^{(2)}(T,e,y) \quad = \quad -\alpha e \mathbb{1}_{\{y>0\}}. \tag{4.2.14}$$

For the moment assume that the value function $V^{(2)}$ is smooth. Then, the optimal strategy is given by

$$\frac{\partial \pi}{\partial q}\left(t,q^{(2)}\right) \quad = \quad -\frac{\partial e}{\partial q}\left(t,q^{(2)}\right)\left(V_e^{(2)}+\beta V_y^{(2)}\right)(t,e,y).$$

By the definition of the value function $V^{(2)}$ in (4.2.12), we expect that

$$-V_e^{(2)}(t,E_t,Y_t)=S_t. \tag{4.2.15}$$

Then

$$\frac{\partial \pi}{\partial q}\left(t,q_t^{(2)}\right) \quad = \quad \frac{\partial e}{\partial q}\left(t,q_t^{(2)}\right)\left(S_t-V_y^{(2)}(t,E_t^{q^{(2)}},Y_t^{q^{(2)}})\right) \tag{4.2.16}$$

Also, it is clear that $V^{(2)}$ is non-increasing in y. Then, comparing the previous expression with (4.1.8), it follows from the assumption on π and e that:

$$q^{(2)} \quad < \quad q^{(1)}.$$

In other words, the impact of the production firm on the prices of carbon emission allowances increases the cost of the externatlity for the firm. This immediately affects the profit function of the firm and leads to a decrease of the level of optimal production. Hence, the presence of the emission market is playing a positive role in terms of reducing the carbon emissions.

The following result shows that under certain assumptions, the above formal calculation is valid in our model.

Theorem 4.2.1 *Suppose that μ_t is continuous and deterministic, γ is constant, $\lambda(q)=\lambda_0$, and $e(q)=e_1q+e_0$ where λ_0, e_1 and e_0 are non–negative constants. Assume that π is $C^{0,1}([0,T]\times\mathbb{R}_+)$, strictly concave in q and*

$$\frac{\partial \pi}{\partial q}(t,0+)>0 \quad \text{and} \quad \frac{\partial \pi}{\partial q}(t,\infty)<0.$$

Then $V_e^{(2)}$ exists and (4.2.15) holds true. In addition, if problem (4.2.13)–(4.2.14) has a bounded solution in $C^{1,1,2}([0,T)\times\mathbb{R}_+\times\mathbb{R})$, then there exists an optimal production strategy satisfying (4.2.16).

Proof. The existence of V_e is due to the fact that V is concave on e and Proposition 4.5.1 verifies (4.2.15).
For the last assertion of the Theorem, notice that by Lemma (4.4.1), V is the unique bounded viscosity solution of (4.2.13)–(4.2.14). Therefore, by the assumption of the Theorem, $V\in C^{1,1,2}([0,T)\times\mathbb{R}_+\times\mathbb{R})$ and one can use the dynamic programming principle to deduce $q^{(2)}$ obtained from (4.2.16) is an optimal strategy. $\square$

4.2.2 Large Carbon emission Impacting the Risk-Neutral Measure

We now consider the general case where the risk premium process is impacted by the emissions of the production firm:

$$\frac{d\mathbb{Q}^q}{d\mathbb{P}}\bigg|_{\mathcal{F}_T} = \exp\left(-\int_0^T \lambda(q_t)dW_t - \frac{1}{2}\int_0^T \lambda(q_t)^2 dt\right).$$

The partial maximization with respect to θ, as in the proof of Proposition 4.2.1, is still valid in this context, and reduces the production firm's problem to

$$\sup_{q.\in\mathcal{Q}} \mathbb{E}\left[U\circ(U')^{-1}\left(y^q\,\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right)\right] \tag{4.2.17}$$

where y^q is defined by

$$\mathbb{E}^{\mathbb{Q}^q}\left[(U')^{-1}\left(y^q\,\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right)\right] = x + \mathbb{E}^{\mathbb{Q}^q}\left[B_T^q\right]. \tag{4.2.18}$$

We also assume that the preferences of the production firm are defined by an exponential utility function

$$U(x) := -e^{-\eta x}, \quad x\in\mathbb{R}.$$

Then $U\circ(U')^{-1}(y) = -y/\eta$, and (4.2.17) reduces to

$$\inf_{q.\geq 0} \mathbb{E}\left[y^q\,\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right] = \inf_{q.\geq 0} y^q. \tag{4.2.19}$$

Finally, the budget constraint (4.2.18) is in the present case:

$$\begin{aligned} x + \mathbb{E}^{\mathbb{Q}^q}\left[B_T^q\right] &= \frac{-1}{\eta}\mathbb{E}^{\mathbb{Q}^q}\left[\ln\left(\frac{y^q}{\eta}\,\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right)\right] \\ &= \frac{-1}{\eta}\left\{\ln\left(\frac{y^q}{\eta}\right) + \mathbb{E}^{\mathbb{Q}^q}\left[\ln\left(\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right)\right]\right\}, \end{aligned}$$

so that the optimization problem (4.2.19) is equivalent to:

$$\begin{aligned} &\sup_{q.\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}^q}\left[B_T^q + \frac{1}{\eta}\ln\left(\frac{d\mathbb{Q}^q}{d\mathbb{P}}\right)\right] \\ &= \sup_{q.\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}^q}\left[\int_0^T\left(\pi + \frac{\lambda^2}{2\eta}\right)(t,q_t)dt - S_T^q\int_0^T e_t(q_t)dt\right]. \end{aligned} \tag{4.2.20}$$

Notice the difference between the above optimization problem, which determines the optimal production policy of the production firm, and the problem (4.2.6). In the present situation where the risk premium process is impacted by the carbon emissions of the firm, the firm's optimization criterion is penalized by the entropy of the risk-neutral measure with respect to the statistical measure.

The firm's optimal production problem (4.2.20) is a standard stochastic control problem. We continue our discussion by considering the Markov case, and introducing the dynamic version of (4.2.20):

$$V^{(3)}(t,e,y) := \sup_{q.\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}^q}_{(t,e,y)}\left[\int_t^T \left(\pi + \frac{\lambda^2}{2\eta}\right)(t,q_t)dt - E^q_T \alpha \mathbb{1}_{\{Y^q_T \geq 0\}}\right], \qquad (4.2.21)$$

where the controlled state dynamics is given by:

$$\begin{aligned} dY^q_t &= \left(\mu(t,Y^q_t) + \beta e(t,q_t) - \gamma(t,Y^q_t)\lambda(t,q_t)\right)dt + \gamma(t,Y^q_t)dW^q_t, \\ dE^q_t &= e(t,q_t)dt, \end{aligned}$$

W^q is a Brownian motion under $\mathbb{Q}^q$, and μ, e, γ, λ are as in **(H1)**–**(H2)**.

By classical arguments, we then see that $V^{(3)}$ solves the dynamic programming equation:

$$\begin{aligned} 0 &= \frac{\partial V^{(3)}}{\partial t} + \mu V^{(3)}_y + \frac{1}{2}\gamma^2 V^{(3)}_{yy} \qquad (4.2.22) \\ &\quad + \max_{q\in\mathbb{R}_+}\left\{\pi(t,q) + \frac{1}{2\eta}\lambda(t,q)^2 + e(t,q)(V^{(3)}_e + \beta V^{(3)}_y) - \gamma\lambda(t,q)V^{(3)}_y\right\} \end{aligned}$$

together with the terminal condition

$$V^{(3)}(T,e,y) = -\alpha e \mathbb{1}_{\{y>0\}}. \qquad (4.2.23)$$

In terms of the value function $V^{(3)}$, the optimal production policy is obtained as the maximizer in the above equation. Under the technical Assumption (4.4.2) below, an interior maximum occurs, and if $V^{(3)}$ is regular enough, then the first order condition is:

$$\frac{\partial \pi}{\partial q}(q^{(3)}) + \frac{1}{\eta}(\lambda\frac{\partial \lambda}{\partial q})(q^{(3)}) + \frac{\partial e}{\partial q}(q^{(3)})(V^{(3)}_e + \beta V^{(3)}_y) - \gamma\frac{\partial \lambda}{\partial q}(q^{(3)})V^{(3)}_y = 0, \quad (4.2.24)$$

where the dependency with respect to (t,e,y) has been omitted for simplicity. As before, we expect that the value function (4.2.21) is regular enough and that the price of the carbon emissions allowance contract, as observed on the emission market, is given by:

$$S_t = -V^{(3)}_e(t,E_t,Y_t). \qquad (4.2.25)$$

Then, it follows that the optimal production policy of the firm is defined by:

$$\begin{aligned} \frac{\partial \pi}{\partial q}(t,q^{(3)}) &= \frac{\partial e}{\partial q}(t,q^{(3)})\left(S_t - \beta V^{(3)}_y(t,Y_t,E_t)\right) \\ &\quad + \frac{\partial \lambda}{\partial q}(t,q^{(3)})\left(\gamma V^{(3)}_y(t,Y_t,E_t) - \frac{1}{\eta}\lambda(t,q^{(3)})\right). \qquad (4.2.26) \end{aligned}$$

The latter expression is the main formula for our financial interpretation and our subsequent numerical experiments. In contrast with the previous case where the risk-premium process was not impacted by the carbon emissions of the large firm, we can not conclude from the above formula that $q^{(3)}$ is smaller than $q^{(1)}$; recall that the optimal production policy in the absence of a financial market defined by

$$\frac{\partial \pi}{\partial q}(t, q^{(1)}) = \frac{\partial e}{\partial q}(t, q^{(1)}) S_t.$$

This is due to the fact that the difference term

$$-\frac{\partial e}{\partial q}(t, q^{(3)}) \beta V_y^{(3)}(t, Y_t, E_t) + \frac{\partial \lambda}{\partial q}(t, q^{(3)}) \left(\gamma V_y^{(3)}(t, Y_t, E_t) - \frac{1}{\eta} \lambda(t, q^{(3)}) \right)$$

has no known sign, and there is no economic argument supporting that it should have some specific sign. The economic intuition hidden in this term is that the large producer may take advantage of his impact on the emission market by manipulating the prices so as to achieve a profit from its trading activity which compensates a higher production activity inducing larger carbon emissions. In the present situation, we see that the emission market has a negative effect on the carbon emissions: the large firm may optimally choose to increase its carbon emissions thus increasing its profit by means of its ability to manipulate the financial market.

Next Theorem shows that for some choice of the coefficients, (4.2.25) holds true and we have the relation (4.2.26).

Theorem 4.2.2 *Suppose that μ_t is continuous and deterministic, γ is constant, $e(q) = e_1 q + e_0$ and $\lambda(q) = \lambda_1 q + \lambda_0$, and $\tilde{\pi}_t(q) := \pi_t(q) + \frac{\lambda(q)^2}{2\eta}$ is deterministic and strictly concave in q with*

$$\tilde{\pi}_t'(0) > 0 \quad \textit{and } \tilde{\pi}_t'(-\infty) < 0.$$

Then $V_e^{(3)}$ exists and (4.2.25) holds true. In addition, if problem (4.2.22)-(4.2.23) has a solution in $C^{1,1,2}([0,T) \times \mathbb{R}_+ \times \mathbb{R})$, then there exists an optimal production strategy satisfying (4.2.26).

Proof. The proof follows the same line of argument as the proof of Theorem 4.2.1. □

4.3 Numerical results

4.3.1 A linear-quadratic example

The main goal of the numerical results is to understand the behavior of the optimal strategy

$$\begin{aligned} \frac{\partial \pi}{\partial q}(t, q^{(3)}) &= \frac{\partial e}{\partial q}(t, q^{(3)}) \left(S_t - \beta V_y^{(3)}(t, Y_t, E_t) \right) \\ &+ \frac{\partial \lambda}{\partial q}(t, q^{(3)}) \left(\gamma V_y^{(3)}(t, Y_t, E_t) - \frac{1}{\eta} \lambda(t, q^{(3)}) \right) \end{aligned} \tag{4.3.1}$$

and more precisely find an example where $q^{(3)} > q^{(1)}$.
We consider the Dynamic Programming Equation

$$V_t + \mu V_y + \frac{1}{2}\gamma^2 V_{yy} + \max_{q\geq 0}\theta(q, V_e, V_y) = 0 \tag{4.3.2}$$

where θ is defined by

$$\theta(q, V_e, V_y) = \pi(t, q) + \frac{1}{2\eta}\lambda(t, q)^2 + e(t, q)(V_e^{(3)} + \beta V_y^{(3)}) - \gamma\lambda(t, q)V_y^{(3)},$$

and with the terminal boundary condition

$$V(T, e, y) = -\alpha e \mathbb{1}_{\{y\geq 0\}}.$$

Here, we consider a simple case where

$$\pi(q) = q(1-q), \;\; e(q) = \lambda(q) = q, \;\; \beta = 1, \text{ and } \alpha = 1.$$

Note that this example satisfies the assumption of Theorem 4.2.2. So, $V_e = -S_t$ and therefore one can compare $q^{(1)}$, $q^{(2)}$ and $q^{(3)}$. It follows that

$$\theta(q, V_e, V_y) = -\left(1 - \frac{1}{2\eta}\right)q^2 + (1 + V_e + (1-\gamma)V_y)\, q.$$

We next assume that $\eta > \frac{1}{2}$ so that the function θ is strictly concave in the q variable. Then, it follows from the first order condition that the optimal production policy is given by:

$$q^{(3)} = \frac{1}{2\rho}(1 + V_e + (1-\gamma)V_y)$$

with $\rho = \left(1 - \frac{1}{2\eta}\right)$, and

$$\max_{q\geq 0}\theta(q, V_e, V_y) = \frac{1}{4\rho}\left(1 + V_e + (1-\gamma)V_y\right)^2.$$

Then, the Dynamic Programming Equation (4.3.2) reduces to:

$$V_t + \mu V_y + \frac{1}{2}\gamma^2 V_{yy} + \frac{1}{4\rho}\left(1 + V_e + (1-\gamma)V_y\right)^2 = 0. \tag{4.3.3}$$

Note that, in order to to compare with $q^{(1)}$, optimal strategy (4.3.1) could be written as:

$$\pi'\left(q^{(3)}\right) = e'\left(q^{(3)}\right)S_t - \tau(e, y),$$

where the correction term $\tau(e, y)$ is defined by

$$\tau(e, y) \;\; = \;\; \frac{2\eta(1-\gamma)}{2\eta - 1}V_y + \frac{1}{2\eta - 1}(1 + V_e).$$

The main objective of our numerical implementation is to exhibit examples of parameters which induce $\tau(e, y) < 0$, or equivalently in terms of the optimal strategy $q^{(3)} > q^{(1)}$.

4.3.2 Numerical scheme

The first step is to set a computational bounded domain $[0, L_e] \times [-L_y, L_y]$ for the (e, y) space domain and discretize the computational domain by the grid $\{(e_i, y_j)\}_{i,j}$. Since we deal with non-linear advection and diffusion phenomena, it is natural to consider Neumann boundary conditions.

Let Δt be the time step and $t^{(k)} = k\Delta t$, for $k = 0, \cdots, n := \lfloor \frac{T}{\Delta t} \rfloor$. We set the discrete terminal data $V_{ij}^{t^{(n)}} = -e_i \mathbb{1}_{\{y_j \geq 0\}}$.

The main difficulty in solving the equation (4.3.3) is the semi-linear terms. In order to overcome this difficulty, we used a time-splitting discretization which divides our scheme into two steps:

- Step 1: we use an implicit finite-differences scheme to solve the diffusion part of the model. This means that on a time step $[t^{(n)}, t^{(n+1)}]$, we solve

$$V_t + \frac{1}{2}\gamma^2 V_{yy} = 0. \tag{4.3.4}$$

- Step 2: we solve the coupling between the advection part with the non-linear effects

$$V_t + \mu V_y + \frac{1}{4\rho}\left(1 + V_e + (1-\gamma)V_y\right)^2 = 0. \tag{4.3.5}$$

  In this important part, we used a relaxation scheme introduced by C. Besse [10]. The scheme is constructed as follow: We rewrite (4.3.5) as the system of two equations:

$$V_t + \mu V_y + \frac{1}{4\rho}\left(1 + V_e + (1-\gamma)V_y\right)\varphi = 0, \tag{4.3.6}$$

$$\varphi = 1 + V_e + (1-\gamma)V_y \tag{4.3.7}$$

  which are solved using a leap-frog scheme in time.

Compared to the Crank-Nicholson scheme, which is also based on a time-centering method, this scheme allows us to avoid a costly numerical treatment of the nonlinearity and to preserve the flexibility of spatial discretization choice.

4.3.3 Results

For parameters $\mu = 0.1$, $\gamma = 0.65$, $\eta = 5$ and the final time is $T = 10$ we produced the following results.

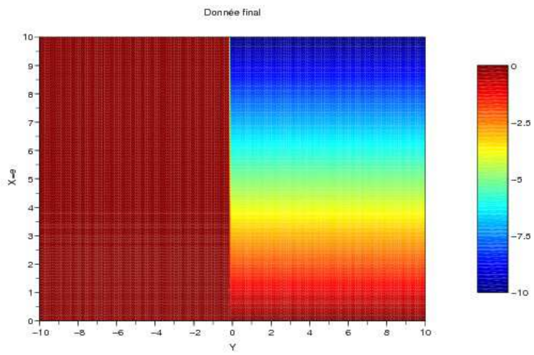

Figure 4.1: Terminal boudary condition $V^{(3)}(T = 10, e, y)$

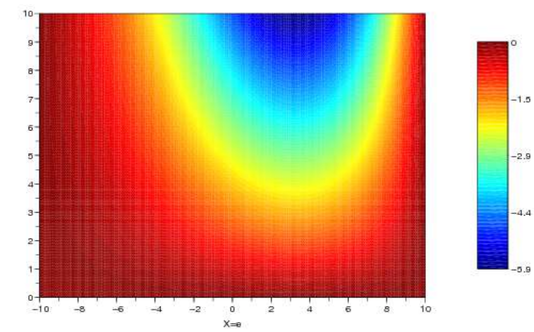

Figure 4.2: The solution of the dynamic programming equation $V^3(e, y)$ at time $t = 0.2$

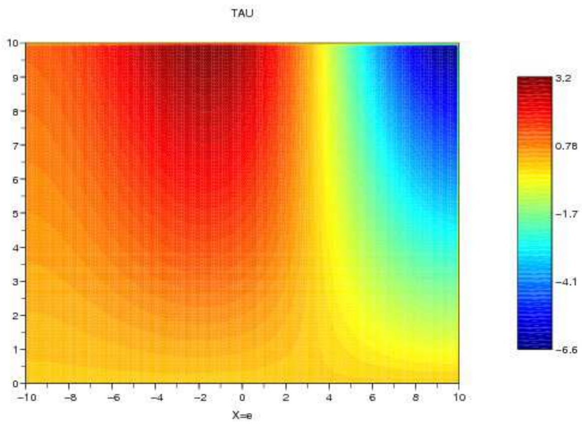

Figure 4.3: The difference term $\tau(e, y)$ at time $t = 0.2$

The blue region shows the couples (e, y) for which we have $q^{(3)} > q^{(2)}$ and therefore within this region the large producer optimally increases her production.

4.4 Uniqueness and verification

Let

$$V(t,e,y) \quad = \quad \sup_{q.\in\mathcal{Q}} \mathbb{E}_{t,e,y}\left[\int_t^T \tilde{\pi}(s,q_s)\mathrm{d}s - \alpha E_T^{q,e}\mathbb{1}_{\{Y_T^{q,y}\geq\kappa\}}\right], \qquad (4.4.1)$$

where

$$\begin{aligned} dY_t^q &= \Big(\mu(t,Y_t^q) + \beta e(t,q_t) + \gamma(t,Y_t)\lambda(t,q_t)\Big)dt + \gamma(t,Y_t^q)dW_t, \\ dE_t^q &= e_t(q)dt \end{aligned}$$

with $\pi, e : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ in $C^{0,1}(\mathbb{R}_+ \times \mathbb{R}_+)$, $\lambda : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ are in $C^0(\mathbb{R}_+)$, $\mu, \gamma : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}$ are continuous in t and Lipschitz in y, and $\gamma \geq 0$.

Notice that $V = V^{(2)}$ or $V^{(3)}$ when $\tilde{\pi} := \pi$ or $\pi + \frac{\lambda^2}{2\eta}$, respectively. Also for simplicity, the dependency of martingale measure with respect to q in the definition of $V^{(2)}$ or $V^{(3)}$ is absorbed in the dynamic of Y_t^q. Therefore in the current Appendix the reference expectation $\mathbb{E}$ is with respect to the measure $\mathbb{P}$ under which the dynamic of Y_t^q is as in the above.

Throughout the Appendix, we suppose

$$\begin{aligned} &(i) \quad \tilde{\pi},\ e,\ \text{and } \lambda \text{ are in } C^{0,1}([0,T]\times\mathbb{R}_+), \\ &(ii) \quad e \text{ is convex and, } \lambda \text{ and } e \text{ are increasing in } q, \qquad (4.4.2) \\ &(iii) \quad \tilde{\pi} \text{ is strictly concave in } q\ ,\ \frac{\partial\tilde{\pi}}{\partial q}(t,0+) > 0 \ \text{ and } \ \frac{\partial\tilde{\pi}}{\partial q}(t,\infty) < 0. \end{aligned}$$

The following Lemma is needed for the proof of Theorems 4.2.2 and 4.2.1.

Lemma 4.4.1 *There exists some $\overline{q}$ such that:*

$$V(t,e,y) = \sup_{q.\in\overline{\mathcal{Q}}} \mathbb{E}_{t,e,y}\left[\int_t^T \tilde{\pi}(t,q_t)dt - E_T^q \alpha \mathbb{1}_{\{Y_T^q \geq 0\}}\right], \tag{4.4.3}$$

where $\overline{\mathcal{Q}}$ is the collection of all $q. \in \mathcal{Q}$ with $0 \leq q \leq \overline{q}$.

Proof. By (4.4.2)(i), we can introduce $\overline{q}$ such that $\tilde{\pi}(\overline{q}) < 0$ and $\tilde{\pi}$ is decreasing in $q \in [\overline{q}, \infty)$. Therefore, if $\tilde{q} := q \wedge \overline{q}$, then $E^{\tilde{q},e} \leq E^{q,e}$ and $\tilde{\pi}(\tilde{q}) \geq \tilde{\pi}(q)$.
On the other hand, by Theorem 1.1 in [37], $Y_T^{\tilde{q},y} \leq Y_T^{q,y}$ a.s.. Therefore,

$$J(\tilde{q}) \geq J(q) \text{ a.s.},$$

where $J(q) := \int_t^T \tilde{\pi}(t,q_t)dt - E_T^q \alpha \mathbb{1}_{\{Y_T^q \geq 0\}}$. □

The next result states that V can be characterized by the PDE Therefore, V solves the dynamic programming equation:

$$\begin{aligned} 0 \;=\; & \frac{\partial V}{\partial t} + \mu V_y + \frac{1}{2}\gamma^2 V_{yy} \\ & + \max_{0\leq q\leq\overline{q}} \{\tilde{\pi}(t,q) + e(t,q)(V_e + \beta V_y) - \gamma\lambda(t,q)V_y\} \end{aligned} \tag{4.4.4}$$

together with the terminal condition

$$V(T,e,y) \;=\; -\alpha e \mathbb{1}_{\{y>0\}}. \tag{4.4.5}$$

Theorem 4.4.1 *Let (4.4.2) hold true. Then V is the unique bounded viscosity solution of (4.4.4)-(4.4.5) on $[0,T] \times \mathbb{R}_+ \times \mathbb{R}$.*

Proof. Notice that one can write (4.4.4) as

$$0 = \frac{\partial V}{\partial t} + H(t,y,V_y,V_e,V_{yy})$$

where

$$\begin{aligned} H(t,y,v_1,v_2,v_{11}) \;:=\; & \mu(t,y)v_1 + \frac{1}{2}\gamma^2(t,y)v_{11} \\ & + \max_{\overline{q}\geq q\geq 0} \{\tilde{\pi}(t,q) + e(t,q)(v_2 + \beta v_1) - \gamma(t,y)\lambda(t,q)v_1\}. \end{aligned}$$

By continuity of H, one can apply Theorem 7.4 in [58] to obtain that V satisfies (4.4.4) in viscosity sense on $[0,T) \times \mathbb{R}_+ \times \mathbb{R}$.
On the other hand, for any $q \in \overline{\mathcal{Q}}$, $\mathbb{1}_{\{Y_T^{t,(q,y)} \geq \kappa\}}$ and $E_T^{t,(q,e)}$ converges to $\mathbb{1}_{\{y\geq\kappa\}}$ and e a.s. as $t \to T$, respectively. Therefore, by Lebesgue dominated convergence Theorem

$$\lim_{t\to T} V(t,e,y) \;=\; -\alpha e \mathbb{1}_{\{y\geq\kappa\}} = V(T,e,y).$$

Consequently, we can deduce that V is the bounded viscosity solution of the boundary value problem (4.4.4)–(4.4.5).
The uniqueness follows from the comparison principle for viscosity solutions in [58]. □

4.5 Existence of optimal production policy

We first show that the existence of an optimal production policyallows to relate the value function V to the market price of carbon allowance S_t.

Lemma 4.5.1 *Let the assumption (4.4.2) hold true. If there exists an optimal control q^* for any (t,e,y) then $\frac{\partial V}{\partial e}(t,e,y) = -\alpha\mathbb{E}[\mathbb{1}_{\{Y_T^{t,y,q^*}\geq\kappa\}}]$.*

Remark 4.5.1 *Lemma 4.5.1 is crucial for the comparison between $q^{(3)}$ and $q^{(2)}$ or $q^{(1)}$. Notice that $S_t = \alpha\mathbb{E}_t[\mathbb{1}_{\{Y_T^{q^*}\geq\kappa\}}]$ is market price which is observable and $\left(\pi + \frac{\lambda^2}{2\eta}\right)$ is concave in q. Therefore, one can replace V_e by $-S_t$ in (4.2.24) and examine the sign of V_y to establish comparison.*

Proof. Notice that by the concavity of V in e, $\frac{\partial V}{\partial e}$ exists almost everywhere. Suppose that $e > e'$. Then, by direct calculations one can write

$$V(t,e,y) - V(t,e',y) + (e-e')\alpha\mathbb{E}\left[\mathbb{1}_{\{Y_T^{t,y,q^*}\geq\kappa\}}\right] \quad\leq\quad 0,$$

where q^* is an optimal strategy for $V(t,e,y)$. This implys that

$$\frac{V(t,e,y) - V(t,e',y)}{e-e'} + \mathbb{E}\left[\mathbb{1}_{\{Y_T^{t,y,q^*}\geq\kappa\}}\right] \quad\leq\quad 0.$$

By passing to the limite as $e' \to e$,

$$V_e(t,e,y) \quad\leq\quad -\mathbb{E}\left[\mathbb{1}_{\{Y_T^{t,y,q^*}\geq\kappa\}}\right].$$

For the other side inequality use $e' > e$. $\square$

We next provide a sufficient condition for the existence of an optimal production policy.

Proposition 4.5.1 *Let μ be deterministic, γ be constant and*

$$e(t,q) := e_1 q + e_0 \quad \textit{and} \quad \lambda(t,q) := \lambda_1 q + \lambda_0, \quad q \geq 0, \tag{4.5.1}$$

where $e_0, \lambda_0, e_1, \lambda_1$ are nonnegative constants. Then the control problem (4.4.1) has an optimal control q^* in $\overline{\mathcal{Q}}$.
In particular, in this setting we have $V_e(t, E_t^{q^*}, Y_t^{q^*}) = -S_t$.

Proof. If $e_1 = \lambda_1 = 0$, the result is trivial. Therefore we suppose that at least one of them is non–zero. Notice that when μ and γ are deterministic, one can write

$$Y_t^q := Y_t^0 + \int_0^t (\beta e(q_s) + \gamma\lambda(q_s))\,dt \quad \text{with} \quad Y_t^0 := y + \int_0^t (\mu_s \mathrm{d}s + \gamma W_s).$$

By Girsanov theorem, we notice that, for every $q \in \overline{\mathcal{Q}}$, the random variable Y_T^q has a Gaussian distribution under the equivalent probability measure $\frac{d\mathbb{Q}}{d\mathbb{P}} := \mathcal{E}\big(-(\beta e(q_t) + \gamma\lambda q_t + \mu_t)\gamma^{-1} dW_t\big)$. Here $\mathcal{E}$ is the Doleans-Dade exponential. Then, the distribution of Y_T^q is absolutely continuous with respect to the Lebesgue measure on $[0, T]$ for all $q \in \overline{\mathcal{Q}}$.

In other words, the distribution of Y_T^q has no atoms, and the cumulative distribution function of the random variable Y_T^q is continuous.
Let $(q^n)_{n\geq 1}$ be a maximizing sequence of V_0, i.e.

$$q^n \in \mathcal{Q} \quad \text{for all } n \geq 1 \text{ and } \quad J(q^n) \longrightarrow V_0.$$

Step 1. Since the processes q^n are uniformly bounded, we deduce from weak convergence and Mazur's lemma that, after possibly passing to a subsequence, there exists a convex combination $\hat{q}^n$ of $(q^j, j \geq n)$ such that:

$$\hat{q}^n := \sum_{j\geq n} \lambda_j^n q^j \quad \longrightarrow \quad q^* \text{ in } \mathbb{L}^1(\Omega \times [0, T]) \text{ and } m \otimes \mathbb{P} - \text{a.s.} \tag{4.5.2}$$

where m is the Lebesgue measure on $[0, T]$. Here $\lambda_j^n \geq 0$ and $\sum_{j\geq n} \lambda_j^n = 1$. Clearly $q^* \in \mathcal{Q}$. Since Y^q is linear in q, this implies that

$$\hat{Y}_T^n := \sum_{j\geq n} \lambda_j^n Y_T^{q^j} \quad \longrightarrow \quad Y_T^{q^*}, \text{ a.s.} \tag{4.5.3}$$

Step 2. By direct estimation and use of Hölder inequality, $Y_T^{q^n}$ is tight under $\mathbb{P}$ and therefore under any equivalent probability measure $\hat{\mathbb{P}}$ with density in $L^2(\mathbb{P})$. Hence after passing to a subsequence, it should converge in distribution to a $\mathcal{F}_T$ random variable Y_T^* which must be equal to $Y_T^{q^*}$;

$$Y_T^{q^n} \longrightarrow Y_T^{q^*} \quad \text{in distribution under} \quad \hat{\mathbb{P}}.$$

Since the convergence in distribution is equivalent to convergence of the corresponding cumulative density functions at all points of continuity, because the probability distribution of Y_T^q is absolutely continuous with respect to Lebesgue measure, it follows that for any positive random variable Z with $\mathbb{E}[Z] = 1$ and $\mathbb{E}[Z^2] < \infty$,

$$\begin{aligned} \mathbb{E}\left[Z \mathbb{1}_{\{Y_T^{q^n} \geq \kappa\}}\right] &= \hat{\mathbb{P}}\left[Y_T^{q^n} \geq \kappa\right] \\ &\longrightarrow \hat{\mathbb{P}}\left[Y_T^{q^*} \geq \kappa\right] = \mathbb{E}\left[Z \mathbb{1}_{\{Y_T^{q^*} \geq \kappa\}}\right]. \end{aligned} \tag{4.5.4}$$

Step 3. Notice that because e and λ are affine, One can write:

$$\int_0^T e(q_s) \mathrm{d}s = \delta\left(Y_T^{q^j} - Y_T^0 - c\right),$$

where $\delta := (\beta e_1 + \gamma\lambda_1)^{-1}$ and $c := \beta e_0 + \gamma\lambda_0$. By the concavity condition (4.4.2), we see that:

$$\begin{aligned}
\sum_{j\geq n} \lambda_j^n J\left(q^j\right) &\leq \mathbb{E}\left[\int_0^T \tilde{\pi}(t, \hat{q}_t^n)dt - \alpha \sum_{j\geq n} \lambda_j^n \mathbb{1}_{\{Y_T^{q^j} \geq \kappa\}} \int_0^T e(q_s^j)ds\right], \\
&= \mathbb{E}\left[\int_0^T \tilde{\pi}(t, \hat{q}_t^n)dt - \alpha \sum_{j\geq n} \lambda_j^n \delta\left(Y_T^{q^j} - Y_T^0 - c\right) \mathbb{1}_{\{Y_T^{q^j} \geq \kappa\}}\right]
\end{aligned}$$

Observe that $Y_T^{q^j} - Y_T^0 - c = \left(Y_T^{q^j} - \kappa\right)^+ + Z^+ - Z^-$ on $\{Y_T^{q^j} \geq \kappa\}$ where $Z^\pm := (Y_T^0 + c - \kappa)^\pm + 1$.

$$\begin{aligned}
\sum_{j\geq n} \lambda_j^n J\left(q^j\right) &\leq \mathbb{E}\Big[\int_0^T \tilde{\pi}(t, \hat{q}_t^n)dt - \alpha \sum_{j\geq n} \lambda_j^n \delta\left(Y_T^{q^j} - \kappa\right)^+\Big] \\
&\quad + \alpha\delta \sum_{j\geq n} \lambda_j^n \mathbb{E}\left[Z^+ \mathbb{1}_{\{Y_T^{q^j} \geq \kappa\}}\right] - \alpha\delta \sum_{j\geq n} \lambda_j^n \mathbb{E}\left[Z^- \mathbb{1}_{\{Y_T^{q^j} \geq \kappa\}}\right].
\end{aligned}$$

By the convexity of the function $y \longmapsto y^+$

$$\begin{aligned}
\sum_{j\geq n} \lambda_j^n J\left(q^j\right) &\leq \mathbb{E}\Big[\int_0^T \tilde{\pi}(t, \hat{q}_t^n)dt - \alpha\delta\left(Y_T^{\hat{q}^n} - \kappa\right)^+\Big] \\
&\quad + \alpha\delta \sum_{j\geq n} \lambda_j^n \mathbb{E}\left[Z^+ \mathbb{1}_{\{Y_T^{q^j} \geq \kappa\}}\right] \\
&\quad - \alpha\delta \sum_{j\geq n} \lambda_j^n \mathbb{E}\left[Z^- \mathbb{1}_{\{Y_T^{q^j} \geq \kappa\}}\right],
\end{aligned}$$

Finally, by applying Step 2 successively to $Z := Z^+$ and Z^-, one can write

$$V(t,e,y) = \lim_{n\to\infty} \sum_{j\geq n} \lambda_j^n J\left(q^j\right) \leq \mathbb{E}\left[\int_0^T \tilde{\pi}(t, q^*)dt - \alpha Y_T^{q^*} \mathbb{1}_{\{Y_T^{q^*} \geq \kappa\}}\right]$$

by dominated convergence. Since $q^* \in \overline{\mathcal{Q}}$, we deduce that $J(q^*) = V_0$. □

Remark 4.5.2 *Proposition (4.5.1) is also valid if we replace Condition (4.5.1) by $\lambda(q) = a + be(q)$ and $\tilde{\pi}(t, e^{-1}(q))$ is convex on q. The modification is straightforward.*

Bibliography

[1] L. B. G. Andersen, R. Brotherton-Ratcliffe, “Extended LIBOR market models with stochastic volatility” Journal of Computational Finance, Vol.9,1, 1-40, 2005. 45

[2] V. Bally, “An Approximation Schemes For BSDEs and Applications to Control and non–linear PDEs”. Prepublication 95-15 du Laboratoire de Statistique et Processus de l’Universite du Maine, (1995). 2

[3] G. Barles, E. R. Jakobsen, “On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations”. Mathematical Modeling and Numerical Analysis, ESAIM, M2AM, Vol. 36 (2002), No. 1, 33-54. 4

[4] G. Barles, E. R. Jakobsen, “Error Bounds For Monotone Approximation Schemes For Hamilton-Jacobi-Bellman Equations”. SIAM J. Numer. Anal., Vol. 43 (2005), No. 2, 540-558. 4

[5] G. Barles, E. R. Jakobsen, “Error Bounds For Monotone Approximation Schemes For Parabolic Hamilton-Jacobi-Bellman Equations”. Math. Comp., 76(240):1861-1893, 2007. 4, 17, 25, 27, 28

[6] G. Barles, P. E. Souganidis, “Convergence of Approximation Schemes for Fully non–linear Second Order Equation”. Asymptotic Anal., 4, pp. 271–283, 1991. 4, 6, 15, 18, 20, 25, 33, 51, 72

[7] R. Bass, “Stochastic differential equations with jumps”. Asymptotic Anal., 4, pp. 271–283, 1991. 48

[8] F.E. Benth, K.H. Karlsen, K. Reikvam “Optimal portfolio management rules in a non-Gaussian market with durability and inter-temporal substitution”. Finance Stochast. 5, 447-467 (2001). 3

[9] F.E. Benth, K.H. Karlsen, K. Reikvam “Optimal portfolio selection with consumption and non–linear integro–differential equations with gradient constraint: A viscosity solution approach”. Finance Stochast. 5, 275-303 (2001). 3

[10] C. Besse. “Schéma de relaxation pour l’équation de Schrödinger non linéaire et les systèmes de Davey et Stewartson.” C.R. Acad. Sci. Paris. Sér. I Math., Vol. 326, pp. 1427-1432, (1998). 91

[11] I. H. Biswas, E. R. Jakobsen, K. H. Karlsen, “Difference-quadrature schemes for non–linear degenerate parabolic integro-PDE”. Preprint, 2009.

[12] I. H. Biswas, E. R. Jakobsen, K. H. Karlsen, "Error estimates for a class of finite difference-quadrature schemes for fully non–linear non–degenerate parabolic integro-PDEs" J. Hyperbolic Differ. Equ. 5(1): 187-219, 2008. 6, 73, 75, 77, 78

[13] I. H. Biswas, E. R. Jakobsen, K. H. Karlsen, "Viscosity solutions for a system of integro-PDEs and connection to optimal switching and control of jump–diffusion processes". Submitted to Asymptotic Anal., 2008. 6, 59, 73, 75, 76

[14] F. Bonnans, H. Zidani, "Consistency of Generalized Finite Difference Schemes for the Stochastic HJB Equation". SIAM J. Numerical Analysis 41-3(2003), 1008-1021 4

[15] B. Bouchard, I. Ekeland, N. Touzi, "On the Malliavin approach to Monte Carlo approximation of conditional expectations". Finance and Stochastics, 8, 45-71 (2004). 33, 37

[16] B. Bouchard, R. Elie, "Discrete time approximation of decoupled Forward-Backward SDE with jumps". Stochastic Processes and their Applications 118, 53–75 (2008). 5

[17] B. Bouchard, R. Elie, N. Touzi, "Discrete–time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs". Radon Series Comp. Appl. Math, 8, 1–34, (2009). 2

[18] B. Bouchard, N. Touzi, "Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations". Stochastic Processes and their Applications 111, 175–206 (2004). 2, 32, 33, 36

[19] R. Carmona, J. Hinz, "Risk-Neutral Modeling of Emission Allowance Prices and Option Valuation." Preprint. 7

[20] R. Carmona, M. Fehr, J. Hinz, "Optimal stochastic control and carbon price formation." SIAM J. Control Optim., 48(4), pp. 2168–2190, 2009. 7

[21] R. Carmona, M. Fehr, J. Hinz, "Properly designed emissions trading schemes do work!" Working paper, 2009. 7

[22] R. Carmona, M. Fehr, J. Hinz, A. Porchet "Market design for emission trading scheme." Preprint. 7

[23] U. Cetin, M. Verschuere. "Pricing and hedging in carbon emissions market." To appear in International Journal of Theoretical and Applied Finance, (2008). 7

[24] F. Camilli, E. R. Jakobsen, "A Finite Element like Scheme for Integro-Partial Differential Hamilton-Jacobi-Bellman Equations". SIAM J. Numer. Anal. 47(4): 2407-2431, 2009.

[25] P. Cheridito, H. M. Soner, N. Touzi, N. Victoir, "Second Order Backward Stochastic Differential Equations and Fully non–linear Parabolic PDEs". Communications on Pure and Applied Mathematics, Volume 60, Issue 7, Date: July 2007, Pages: 1081-1110. 2, 3, 5, 11, 12, 14, 37

[26] D. Chevance, "Ph.D. thesis". Univ. Provence, (1996). 2

[27] R. Cont, P. Tankov, "Lévy process..." 48

[28] D. Crisan, K. Manolarakis, N. Touzi, "On the Monte Carlo simulation Of BSDE's: An improvement on the Malliavin weights". to appear. 33

[29] K. Debrabant and E. R. Jakobsen. "Semi-Lagrangian schemes for linear and fully non-linear diffusion equations". Preprint, 2009. 5

[30] F. Delarue and S. Menozzi, "A Forward-Backward Stochastic Algorithm for Quasi-Linear PDEs". Annals of Applied Probability, Vol. 16, No. 1, 140-184, 2006.

[31] H. Dong, N. V. Krylov, "On The Rate Of Convergence Of Finite-Difference Approximations For Bellman's Equations With Constant Coefficients". St. Petersburg Math. J., Vol. 17 (2005), No. 2, 108-132.

[32] N. El Karoui, S. Peng, M. C. Quenez, "Backward Stochastic Differential Equations in Finance". Mathematical Finance, 7, 1-71 (1997). 2

[33] L. C. Evans and A. Friedman, "Optimal stochastic switching and the Dirichlet problem for the Bellman equation". Trans. Amer. Math. Soc. 253:365-389 (1979).

[34] A. Fahim, N. Touzi and X. Warin, "A Probabilistic Numerical Method for Fully non–linear Parabolic PDEs". Preprint (2009). 5

[35] E. Gobet, J. P. Lemor, X. Warin "A regression-based Monte-Carlo method to solve backward stochastic differential equations". Annals of Applied Probability, Vol.15(3), pp.2172-2002 - 2005. 2, 32, 37, 38

[36] S. L. Heston, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". The Review of Financial Studies, Vol.6(2), 327-343, 1993. 40

[37] N. Ikeda, S. Watanabe "A comparison Theorem for solutions of stochastic differential equations and its application." Osaka J. Math., 14(1977), 619–633. 94

[38] C. Imbert, S. Serfaty, "Repeated games for eikonal equations, integral curvature flows and non–linear parabolic integro–differential equations". Preprints, 2009. 6

[39] E. R. Jakobsen, K. H. Karlsen, “Continuous dependence estimates for viscosity solutions of integro-PDEs”. J. Differential Equations 212(2): 278-318, 2005.

[40] E. R. Jakobsen, K. H. Karlsen, C. La Chioma, “Error estimates for approximate solutions to Bellman equations associated with controlled jump–diffusions”. Numer. Math. 110(2): 221-255, 2008.

[41] C. Kahl, P. Jackel “Fast strong approximation Monte Carlo schemes for stochastic volatility models” Quantitative Finance,Vol.6(6),513-536, 2006. 42

[42] R. V. Kohn, S. Serfaty, “A deterministic-control-based approach to motion by curvature”. Comm. Pure and Appl. Math., 59, (2006), 344-407. 5, 6

[43] N. V. Krylov, “On The Rate Of Convergence Of Finite-Difference Approximations For Bellman’s Equations”. St. Petersburg Math. J., Vol. 9 (1997), No. 3, 245-256. 4, 25

[44] N. V. Krylov, “On The Rate Of Convergence Of Finite-Difference Approximations For Bellman’s Equations With Lipschitz Coefficients”. Applied Math. and Optimization, Vol. 52 (2005), No. 3, 365-399. 4

[45] N. V. Krylov, “On The Rate Of Convergence Of Finite-Difference Approximations For Bellman’s Equations With Variable Coefficients”. Probab. Theory Ralat. fields, 117:1-16, (2000). 4

[46] P. L. Lions, H. Regnier, Calcul du prix et des “Sensibilités d’une option américaine par une méthode de Monte Carlo”, (2001), Preprint. 33, 36, 37

[47] F. A. Longstaff, R. S. Schwartz, “Valuing American options by simulation: a simple least-square approach”, *Review of Financial Studies* 14, 113–147 (2001). 37

[48] R. Lord, R. Koekkoek, D. Van Dijk, “A comparison of biased simulation schemes for stochastic volatility models” forthcoming in Quantitative Finance, 2008. 43

[49] J. Ma, P. Protter, J. Yong, “Solving forward–backward stochastic differential equations explicitly–a four step scheme” Probab. Theory Related Fields, 98, pp. 339–359, 1994. 1

[50] J.J. McCarthy, O.F. Canziani, N. A. Leary, D. J. Dokken, K. S. White, “Climate Change 2001: Impacts , adaption and vulnerability.” Contribution of working group II to the third assessment report of the intergovernmental panel on climate change, Cambridge University Press, Cambridge, http://www.grida.no/climate/ipcc_tar/wg2/index.htm 7

[51] E. Pardoux, “Backward stochastic differencial equations and viscosity solutions of semi–linear PDEs”. Lecture Note, (2003) 1

[52] E. Pardoux and S. Peng, "Adapted solution of a backward stochastic differential equation", Systems and control letters 14 (1990), pp. 55–61. 1

[53] E. Pardoux and S. Peng, "Backward stochastic differential equations and quasi-linear parabolic partial differential equations", Lecture Notes In Control And Inform. Sci. 176 (1992), pp. 200–217. 1

[54] J. Seifert, U. H. Marliese, M. Wagner. "Dynamic Behavior of CO_2 Spot Prices." Journal of Environmental Economics and Management, Vol. 56, No. 2, pp. 180-194, (2008).

[55] H. M. Soner and N. Touzi, "A stochastic representation for mean curvature type geometric flows". The Annals of Probability, Vol 31(3),1145-1165 (2003). 37

[56] H. M. Soner and N. Touzi, "Dynamic programming for stochastic target problems and geometric flows". TJ. Eur. Math. Soc., 4,201-236 (2002).

[57] D. W. Strook, "Diffusion processes associated with Lévy generators". Z. Wahrscheinlichkeitstheprie und Verw. Gebiete 32, No. 3, 209–244, (1975). 48

[58] N. Touzi, "Stochastic control problems, viscosity solutions and application to finance" Scuola Normale Superiore, (2004). 94

[59] J. Zhang, "A numerical scheme for backward stochastic differential equations". Annals of Applied Probability, 14 (1), 459-488, (2004). 2

[60] J. Zhang, "Some fine properties of backward stochastic differential equations". PhD Thesis, Purdue University, (2001). 2

[61] T. Zariphopoulou, "A solution approach to valuation with unhedgeable risks". Finance and Stochastics, 5 (2001) 61-82. 41