Electronic Journal of Probability

Electron. J. Probab. **18** (2013), no. 15, 1–24.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-1967

A splitting method for fully nonlinear degenerate parabolic PDEs

Xiaolu Tan*

Abstract

Motivated by applications in Asian option pricing, optimal commodity trading etc., we propose a splitting scheme for fully nonlinear degenerate parabolic PDEs. The splitting scheme generalizes the probabilistic scheme of Fahim, Touzi and Warin [13] to the degenerate case. General convergence as well as rate of convergence are obtained under reasonable conditions. In particular, it can be used for a class of Hamilton-Jacobi-Bellman equations, which characterize the value functions of stochastic control problems or stochastic differential games. We also provide a simulation-regression method to make the splitting scheme implementable. Finally, we give some numerical tests in an Asian option pricing problem and an optimal hydropower management problem.

Keywords: Numerical scheme ; nonlinear degenerate PDE ; splitting method ; viscosity solution.
AMS MSC 2010: Primary 65C05 ; secondary 49L25.
Submitted to EJP on April 23, 2012, final version accepted on January 26, 2013.

1 Introduction

Numerical methods for parabolic partial differential equations (PDEs) are largely developed in the literature, on finite difference scheme, finites elements scheme, semi-Lagrangian scheme, Monte-Carlo method, etc. For nonlinear PDEs, and especially in high dimensional cases, the numerical resolution becomes a big challenge.

A typical kind of nonlinear parabolic PDEs is the Hamilton-Jacobi-Bellman (HJB) equation, which characterizes the solution of the optimal control problems. In this context, for finite difference method, one can only use the explicit scheme, since the implicit scheme needs to invert too many matrices. In the one dimensional case, the explicit finite difference scheme can be easily constructed and the monotonicity is guaranteed by the CFL condition. In high dimensional cases, Bonnans and Zidani [4] propose a numerical algorithm to construct a monotone scheme. Another numerical method for general HJB equations is the semi-Lagrangian scheme proposed in Debrabant and Jakobsen [12]. It can be easily constructed to be monotone, but they need next to use a finite

*CMAP, École Polytechnique, Paris. E-mail: xiaolu.tan@polytechnique.edu

difference grid as well as an interpolation method to make it implementable. It hence can be viewed as a finite difference scheme.

Generally speaking, finite difference and semi-Lagrangian schemes are easily implemented and perform quite well in low dimensional cases; and in high dimensional cases, the Monte-Carlo method is preferred. Recently, Fahim, Touzi and Warin [13] proposed a probabilistic method for nonlinear parabolic PDEs, which is closely related to the second order backward stochastic differential equation (2BSDE) developed in Cheridito et al. [9] and Soner et al. [18]. With simulations of a diffusion process, they propose the estimations of the value function and its derivatives by conditional expectations, by which they can approximate the nonlinear part of the PDE and then get a convergent scheme. However, their scheme can only be applied in the non-degenerate cases.

We want to generalize the probabilistic scheme of Fahim, Touzi and Warin [13] to the degenerate case, motivated by its applications in finance. For example, in Asian option pricing problems, we must consider the cumulative average stock prices A_t; for lookback options, we consider also the historical maximum and/or minimum stock prices M_t, m_t. They are all degenerate variables without a diffusion generator, and hence the pricing equation turns to be a degenerate parabolic equation. In some optimal commodity trading models(see e.g. [1], [7] and [8]), the storage amount of commodities is an important state variable, and the optimization problem induces a PDE which degenerates on storage amount variable. In life insurance, Dai et al. [11] proposed a financial pricing model for a Variable Annuities product Guaranteed Minimum Withdrawal Benefit (GMWB). In their model, the price of GMWB depends on two variables: the reference account and the guaranteed account, where the latter degenerates and the pricing equation is a degenerate parabolic PDE.

For these degenerate PDEs, the degenerate part is separable. Therefore, a natural solution is the splitting scheme. Our idea is to use the probabilistic scheme to treat the non-degenerate part, and use the semi-Lagrangian scheme to solve the degenerate part, and by combining the two methods, we get a splitting scheme. In particular, it generalizes the probabilistic scheme of Fahim, Touzi and Warin [13] to the degenerate case.

Another contribution of the paper is to propose a simulation-regression technique to make the semi-Lagrangian scheme implementable, in place of the classical finite difference method together with interpolation technique as used in Debrabant and Jakobsen [12], or Chen and Forsyth [8]. In the simulation-regression method, we can use global polynomials, or local hypercubes or local polynomials etc. as regression function basis. The global polynomial method means to approximate a function with some polynomials on the whole space, while the local basis method means to discretize first the space into local rectangles, and then to approximate the corresponding function with some polynomials on every local rectangle. As illustrated in Gobet, Lemor and Warin [14] and also in Bouchard and Warin [6], the local hypercubes and local polynomials basis method are very efficient in concrete cases. Moreover, they show that in practice, it is enough to choose a small number (about five or six) of discretization points in every dimension for the local basis method, while for finite difference method, one needs many more discretization points (more than 50 points in [8] for example) in every dimension. In particular, it permits to treat problems in high dimensions (up to 5 dimensions in [13] and up to 6 dimensions in [6]). In our context, we shall provide a four dimensional numerical example.

The rest of the paper is organized as follows. In Section 2, we introduce a degenerate PDE and a splitting scheme which combines the probabilistic scheme in [13] and semi-Lagrangian scheme. Then we provide a local uniform convergence result as well as a rate of convergence, where the main idea is to adapt the viscosity solution technique proposed in Barles and Souganidis [3] and Barles and Jakobsen [2]. In Section 3, we propose a simulation-regression technique to approximate the conditional expectations used in the splitting scheme, making the scheme implementable. We shall also discuss the choices of function basis used in the regression and then provide some convergence results for this implementable scheme. Finally, Section 4 provides some experimental examples.

Notation: Let $|\eta| := \eta^1 + \cdots + \eta^d$ for $\eta \in \mathbb{N}^d$. Given $T \in \mathbb{R}^+$ and $d,\ d' \in \mathbb{N}$, we denote $Q_T := [0,T) \times \mathbb{R}^d \times \mathbb{R}^{d'}$, $\overline{Q}_T := [0,T] \times \mathbb{R}^d \times \mathbb{R}^{d'}$ and

$$C^{0,1}(Q_T) \;\; := \;\; \big\{ \varphi \,:\, Q_T \to \mathbb{R} \text{ such that } |\varphi|_1 < \infty \big\},$$

where

$$|\varphi|_0 \;:=\; \sup_{Q_T} |\varphi(t,x,y)| \quad \text{and} \quad |\varphi|_1 \;:=\; |\varphi|_0 \;+ \sup_{Q_T \times Q_T} \frac{|\varphi(t,x,y) - \varphi(t',x',y')|}{|x-x'| + |y-y'| + |t-t'|^{\frac{1}{2}}}.$$

In this paper, the constant C is used in many inequalities, its value may vary from line to line.

2 The degenerate PDE and splitting scheme

In this section, we first introduce a nonlinear parabolic PDE which has a separable degenerate part. We next propose a splitting scheme, and for which we provide a local uniform convergence result of the splitting scheme when the PDE satisfies a comparison result for bounded viscosity solutions, as well as a rate of convergence when the nonlinear part of the PDE is a concave Hamiltonian.

2.1 A degenerate nonlinear PDE

Let $T \in \mathbb{R}^+$, $\mu : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to S_d$ be continuous, denote $a(t,x) := \sigma(t,x)\sigma(t,x)^T$, we define a linear operator $\mathcal{L}^X$ on the smooth functions $\varphi : Q_T \to \mathbb{R}$ by

$$\mathcal{L}^X \varphi(t,x,y) \;\; := \;\; \partial_t \varphi(t,x,y) \;+\; \mu(t,x) \cdot D_x \varphi(t,x,y) \;+\; \frac{1}{2} a(t,x) \cdot D^2_{xx} \varphi(t,x,y).$$

We say that $\mathcal{L}^X$ is a linear operator associated to the diffusion process $X = (X_t)_{0 \le t \le T}$ defined by the stochastic differential equation:

$$dX_t \;\; = \;\; \mu(t,X_t)\, dt \;+\; \sigma(t,X_t)\, dW_t, \tag{2.1}$$

where $W = (W_t)_{0 \le t \le T}$ is a d-dimensional standard Brownian motion.

Given a nonlinear function

$$F :\; (t,x,y,r,p,\Gamma) \in \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^{d'} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d \;\;\mapsto\;\; F(t,x,y,r,p,\Gamma) \in \mathbb{R},$$

we then get a nonlinear operator $F(t,x,y,\varphi,D_x\varphi,D^2_{xx}\varphi)$ on φ. We denote by F_p and F_Γ the derivative of function F w.r.t. p and Γ.

Next, we give the degenerate part which involves with the partial gradient with respect to y. Given functions

$$\big(l^{\alpha,\beta},\ c^{\alpha,\beta},\ f_i^{\alpha,\beta},\ g_j^{\alpha,\beta}\ \big)_{\alpha \in \mathcal{A},\ \beta \in \mathcal{B},\ 1 \le i \le d,\ 1 \le j \le d'}$$

defined on Q_T with index space $\mathcal{A}$ and $\mathcal{B}$, we denote $f^{\alpha,\beta} := (f_i^{\alpha,\beta})_{1 \le i \le d}$ and $g^{\alpha,\beta} := (g_j^{\alpha,\beta})_{1 \le j \le d'}$, and define the Lagrangian $\mathcal{L}^{\alpha,\beta}$ by

$$\begin{aligned}
\mathcal{L}^{\alpha,\beta}\varphi(t,x,y) \;\; := \;\; & l^{\alpha,\beta}(t,x,y) \;+\; c^{\alpha,\beta}(t,x,y)\varphi(t,x,y) \\
& +\; f^{\alpha,\beta}(t,x,y) \cdot D_x\varphi(t,x,y) \;+\; g^{\alpha,\beta}(t,x,y) \cdot D_y\varphi(t,x,y),
\end{aligned}$$

and the Hamiltonian by

$$H(t,x,y,\varphi(t,x,y),D_x\varphi(t,x,y),D_y\varphi(t,x,y)) \quad := \quad \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \mathcal{L}^{\alpha,\beta}\varphi(t,x,y).$$

Finally, let us introduce the degenerate fully nonlinear parabolic PDE which will be considered throughout the paper:

$$\big[-\mathcal{L}^X v \;-\; F(\cdot,v,D_x v,D^2_{xx}v) \;-\; H(\cdot,v,D_x v,D_y v) \,\big](t,x,y) \;=\; 0, \quad \text{on } Q_T, \tag{2.2}$$

with terminal condition

$$v(T,x,y) \quad = \quad \Phi(x,y). \tag{2.3}$$

The PDE (2.2) is composed by three separable parts: the linear part $\mathcal{L}^X$, the nonlinear part F, and the first order degenerate part H.

2.2 A splitting scheme

As observed above, the three parts in PDE (2.2) are separable, we can then propose a splitting numerical scheme to solve it. The idea is to split (2.2) into the following two equations:

$$-\mathcal{L}^X v(t,x,y) \;-\; F(\cdot,v,D_x v,D^2_{xx}v)(t,x,y) \quad = \quad 0 \tag{2.4}$$

and

$$-\partial_t v(t,x,y) \;-\; H(\cdot,v,D_x v,D_y v)(t,x,y) \quad = \quad 0, \tag{2.5}$$

then to solve them separately. Equation (2.4) is nonlinear and non-degenerate for every fixed y, then it can be treated by the probabilistic scheme proposed in Fahim et al.[13]. Equation (2.5) is a first order Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation, we shall solve it by semi-Lagrangian scheme. Then, combining the two schemes sequentially, we get the splitting scheme.

Let us first give a time discrete grid $(t_n)_{n=0,\cdots,N}$ with $t_n := nh$, where $h := T/N$ for $N \in \mathbb{N}$. As in [13], we define $\hat{X}^{t,x}_h$ by the Euler scheme of the diffusion process X in (2.1):

$$\hat{X}^{t,x}_h \quad := \quad x \;+\; \mu(t,x)\,h \;+\; \sigma(t,x)\cdot(W_{t+h}-W_t), \quad \forall (t,x)\in[0,T]\times\mathbb{R}^d.$$

Let v^h denote the numerical solution, then the probabilistic scheme of [13] for equation (2.4) is given by

$$v^h(t_n,x,y) = \mathbf{T}_h[v^h](t_n,x,y) := \mathbb{E}\big[v^h(t_{n+1},\hat{X}^{t_n,x}_h,y)\big] + hF(t_n,x,y,\mathbb{E}\mathcal{D}_h v^h(t_n,x,y)), \tag{2.6}$$

where

$$\mathbb{E}\mathcal{D}_h\varphi(t_n,x,y) \quad := \quad \Big(\, \mathbb{E}\big[\varphi(t_{n+1},\hat{X}^{t_n,x}_h,y)H^{t_n,x,h}_i(\Delta W_{n+1})\big] : i=0,1,2 \,\Big),$$

with $\Delta W_{n+1} := W_{t_{n+1}} - W_{t_n}$ and the Hermite polynomials are defined by $H^{t,x,h}_0(w) := 1$, $H^{t,x,h}_1(w) := \sigma^T(t,x)^{-1}\frac{w}{h}$ and $H^{t,x,h}_2(w) := \sigma^T(t,x)^{-1}\frac{ww^T-hI_d}{h^2}\sigma(t,x)^{-1}$.

Remark 2.1. *The scheme $\mathbf{T}_h$ is well defined as soon as Det$(\sigma(t,x))\neq 0$ for each $(t,x)\in[0,T)\times\mathbb{R}^d$. When φ is smooth, by integration by parts, one can verify that*

$$\mathbb{E}\left[\varphi\big(t_{n+1},\hat{X}^{t_n,x}_h,y\big)\; H^{t_n,x,h}_i(\Delta W_{n+1})\right] \;=\; \mathbb{E}\; D^i_{x^i}\varphi\big(t_{n+1},\hat{X}^{t_n,x}_h,y\big), \quad i=0,1,2.$$

For more details on this fact and of the probabilistic scheme $\mathbf{T}_h$ of (2.6), we refer to Fahim et al. [13].

The second PDE (2.5) is a first order HJBI equation, its semi-Lagrangian scheme is given by

$$v^h(t_n,x,y) \;=\; \mathbf{S}_h[v^h](t_n,x,y) := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \Big\{ hl^{\alpha,\beta}(t_n,x,y) + hc^{\alpha,\beta}(t_n,x,y)v^h(t_{n+1},x,y) + v^h\big(t_{n+1},\ x+hf^{\alpha,\beta}(t_n,x,y),\ y+hg^{\alpha,\beta}(t_n,x,y)\big)\Big\}. \tag{2.7}$$

Remark 2.2. *The semi-Lagrangian scheme $\mathbf{S}_h$ is deduced intuitively from the discrete version of equation* (2.5)*:*

$$\frac{v^h(t_{n+1},x,y)-v^h(t_n,x,y)}{h} \;+\; \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \Big\{\ l^{\alpha,\beta}(t_n,x,y) \;+\; c^{\alpha,\beta}(t_n,x,y)v^h(t_{n+1},x,y) \;+\; \frac{v^h(t_{n+1},\ x+hf^{\alpha,\beta}(t_n,x,y),\ y+hg^{\alpha,\beta}(t_n,x,y)) \;-\; v^h(t_{n+1},x,y)}{h}\ \Big\} = 0.$$

Finally, we are ready to introduce the splitting scheme $\mathbf{S}_h\circ\mathbf{T}_h$ for the original PDE (2.2), (2.3). Concretely, with terminal condition

$$v^h(t_N,x,y) \;:=\; \Phi(x,y), \tag{2.8}$$

we compute $v^h(t_n,\cdot)$ in a backward iteration. Given $v^h(t_{n+1},\cdot)$, we introduce the fictitious time $t_{n+\frac{1}{2}}$ and compute $v^h(t_n,\cdot)$ by

$$v^h(t_{n+\frac{1}{2}},x,y) \;:=\; \mathbf{T}_h[v^h](t_n,x,y) \quad \text{with } \mathbf{T}_h \text{ defined in (2.6)}, \tag{2.9}$$

and

$$\begin{aligned} v^h(t_n,x,y) \;&=\; \mathbf{S}_h\circ\mathbf{T}_h[v](t_n,x,y) \\ :=&\; \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \Big\{\ h\ l^{\alpha,\beta}(t_n,x,y) \;+\; h\ c^{\alpha,\beta}(t_n,x,y)\ v^h(t_{n+\frac{1}{2}},x,y) \\ &\qquad +\; v^h\big(t_{n+\frac{1}{2}},\ x+f^{\alpha,\beta}(t_n,x,y)h,\ y+g^{\alpha,\beta}(t_n,x,y)h\big)\ \Big\}. \end{aligned} \tag{2.10}$$

Clearly, when $\mathrm{Det}(\sigma(t,x))\neq 0$ for every $(t,x)\in[0,T)\times\mathbb{R}^d$, the scheme $\mathbf{S}_h\circ\mathbf{T}_h$ is well defined and it gives a unique numerical solution v^h.

2.3 The convergence results

We shall provide two convergence results for the splitting scheme $\mathbf{S}_h\circ\mathbf{T}_h$ in (2.10), similar to Fahim et al.[13]. The first one is the local uniform convergence in the context of Barles and Souganidis [3], and the second is a rate of convergence.

We first recall that an upper semicontinuous (resp., lower semicontinuous) function $\underline{v}$ (resp. $\overline{v}$) on Q_T is called a viscosity subsolution (resp., supersolution) of (2.2) if, for any $(t,x,y)\in Q_T$ and any smooth function φ satisfying

$$0=(\underline{v}-\varphi)(t,x,y)=\max_{Q_T}(\underline{v}-\varphi) \qquad \Big(\text{resp., } 0=(\overline{v}-\varphi)(t,x,y)=\min_{Q_T}(\overline{v}-\varphi)\Big),$$

we have

$$-\,\mathcal{L}^X\varphi \;-\; F(t,x,y,\varphi,D_x\varphi,D^2_{xx}\varphi) \;-\; H(t,x,y,D_x\varphi,D_y\varphi) \quad\leq (\text{resp.,}\ \ \geq) \quad 0.$$

Definition 2.3. *We say that the PDE* (2.2) *satisfies a comparison result for bounded functions if, for any bounded upper semicontinuous subsolution $\underline{v}$ and any bounded lower semicontinuous supersolution $\overline{v}$ on $\overline{Q}_T$ satisfying*

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

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

Let us now give some assumptions on the equation (2.2), and then provide a first convergence result.

Assumption F : (i) The diffusion coefficients μ and σ are Lipschitz in x and continuous in t, $\sigma\sigma^T(t,x) > 0$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$ and $\int_0^T \big|\sigma\sigma^T(t,0) + \mu(t,0)\big| dt < \infty$.
(ii) The nonlinear operator F is uniformly Lipschitz in (x,y,r,p,Γ), continuous in t and $\sup_{(t,x,y)\in Q_T} |F(t,x,y,0,0,0)| < \infty$.
(iii) F is elliptic and satisfies

$$a^{-1} \cdot F_\Gamma \quad \leq \quad 1 \quad \text{on } \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d'} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d. \tag{2.11}$$

(iv) $F_p \in \text{Image}(F_\Gamma)$ and $\big|F_p^T F_\Gamma^{-1} F_p\big|_\infty < +\infty$.

Remark 2.4. *Assumption* **F** *is almost the same as the Assumption* **F** *in [13], here we just add a variable y in the nonlinear operator F.*

Assumption H : The coefficients in Hamiltonian H are all uniformly bounded, i.e.

$$\sup_{(\alpha,\beta)\in\mathcal{A}\times\mathcal{B},\ 1\leq i\leq d,\ 1\leq j\leq d'} \Big\{ \, |l^{\alpha,\beta}|_0 \ + \ |c^{\alpha,\beta}|_0 \ + \ |f_i^{\alpha,\beta}|_0 \ + \ |g_j^{\alpha,\beta}|_0 \, \Big\} \quad < \quad \infty.$$

Assumption M : $F_r - \ \frac{1}{4}\ F_p^T F_\Gamma^{-1} F_p \ \geq \ 0$ and $c^{\alpha,\beta} \geq 0$ for every $\alpha \in \mathcal{A}, \beta \in \mathcal{B}$.

Remark 2.5. *Assumption* **M** *is imposed to guarantee the monotonicity of the splitting scheme $\mathbf{S}_h \circ \mathbf{T}_h$. However, it is not crucial as soon as Assumptions* **F** *and* **H** *hold true. In fact, as discussed in Remark 3.13 of [13], since the equation is parabolic, we can introduce a new function $u(t,x,y) := e^{\theta(T-t)} v(t,x,y)$ for some positive constant θ large enough, then the new PDE for $u(t,x,y)$ satisfies Assumption* **M** *under Assumptions* **F** *and* **H**. *Here, we impose this assumption only to simplify the presentation and the arguments.*

Theorem 2.6. *Let Assumptions* **F**, **H** *and* **M** *hold true, and assume that the degenerate fully nonlinear parabolic PDE* (2.2) *satisfies a comparison result for bounded viscosity solutions. Then for every bounded Lipschitz terminal condition function Φ, there exists a bounded function v such that*

$$v^h \ \longrightarrow \ v \quad \textit{locally uniformly as} \ \ h \to 0,$$

where v^h is the numerical solution of scheme $\mathbf{S}_h \circ \mathbf{T}_h$ defined by (2.8), (2.9) *and* (2.10). *Moreover, v is the unique bounded viscosity solution of the equation* (2.2) *with terminal condition* (2.3).

It is clear that Assumptions **F** and **H** hold true for a class of HJB equations as well as a class of HJBI equations which characterize the value functions of the stochastic differential game problems. We next provide a rate of convergence in case that F and H are both concave Hamiltonians, i.e. when the nonlinear equation (2.2) is a HJB equation. We shall use the arguments developed by Barles and Jakobsen [2]. The following stronger assumptions implies that the nonlinear PDE (2.2) satisfies a comparison result for bounded functions, and has a unique bounded viscosity solution given a bounded and Lipschitz continuous function Φ, see e.g. Proposition 2.1 of [2].

Assumption HJB : Assumptions **F** and **M** hold and F is a concave Hamiltonian, i.e.

$$\mu \cdot p \ + \ \frac{1}{2} a \cdot \Gamma \ + \ F(t,x,y,r,p,\Gamma) \quad = \quad \inf_{\gamma\in\mathcal{C}} \ \mathcal{L}^\gamma(t,x,y,r,p,\Gamma),$$

with

$$\mathcal{L}^\gamma(t,x,y,r,p,\Gamma) \ := \ l^\gamma(t,x,y) \ + \ c^\gamma(t,x,y) r \ + \ f^\gamma(t,x,y) \cdot p \ + \ \frac{1}{2} a^\gamma(t,x,y) \cdot \Gamma.$$

And $\mathcal{B} = \{\beta\}$ is a singleton, hence H is also a concave Hamiltonian, so that it can be written as

$$H(t,x,y,r,p,q) \;=\; \inf_{\alpha\in\mathcal{A}} \left\{ l^\alpha(t,x,y) \;+\; c^\alpha(t,x,y)r \;+\; f^\alpha(t,x,y)\cdot p \;+\; g^\alpha(t,x,y)\cdot q \right\}$$

Moreover, the functions l, c, f, g and σ satisfy that

$$\sup_{\alpha\in\mathcal{A},\gamma\in\mathcal{C}} \Big(|l^\alpha + l^\gamma|_1 + |c^\alpha + c^\gamma|_1 + |f^\alpha + f^\gamma|_1 + |g^\alpha|_1 + |\sigma^\gamma|_1 \Big) \;<\; \infty$$

Assumption HJB+ : Assumption **HJB** holds true, and for any $\delta > 0$, there exists a finite set $\{\alpha_i, \gamma_i\}_{i=1}^{I_\delta}$ such that for any $(\alpha,\gamma) \in \mathcal{A}\times\mathcal{C}$:

$$\inf_{1\le i\le I_\delta} \Big(|l^\alpha - l^{\alpha_i}|_0 \;+\; |c^\alpha - c^{\alpha_i}|_0 \;+\; |f^\alpha - f^{\alpha_i}|_0 \;+\; |\sigma^\alpha - \sigma^{\alpha_i}|_0 \Big) \;\le\; \delta,$$

and

$$\inf_{1\le i\le I_\delta} \Big(|l^\gamma - l^{\gamma_i}|_0 \;+\; |c^\gamma - c^{\gamma_i}|_0 \;+\; |f^\gamma - f^{\gamma_i}|_0 \;+\; |g^\gamma - g^{\gamma_i}|_0 \Big) \;\le\; \delta.$$

Theorem 2.7. *Suppose that the terminal condition function Φ is bounded and Lipschitz-continuous. Then there is a constant C such that* (i) *under Assumption* **HJB***, we have* $v - v^h \;\le\; Ch^{\frac{1}{4}}$, (ii) *under Assumption* **HJB+***, we have* $-Ch^{\frac{1}{10}} \;\le\; v - v^h \;\le\; Ch^{\frac{1}{4}}$*, where v is the unique bounded viscosity solution of* (2.2) *introduced in Theorem 2.6.*

Remark 2.8. *The above convergence rate is the same as that obtained in Fahim et al.[13]. It may not be the best rate in general. However, to the best of our knowledge, it is the optimal rate that we can prove in this stochastic control problem context so far.*

2.4 Proof of local uniform convergence

To prove the local uniform convergence in Theorem 2.6, we shall verify the criteria proposed in Theorem 2.1 of Barles and Souganidis [3]: the monotonicity, the consistency of the scheme and the stability of the numerical solutions. Moreover, as discussed in Remark 3.2 of [13], we need also to show that

$$\liminf_{(t',x',y',h)\to(T,x,y,0)} v^h(t',x',y') \ge \Phi(x,y) \text{ and } \limsup_{(t',x',y',h)\to(T,x,y,0)} v^h(t',x',y') \le \Phi(x,y). \tag{2.12}$$

Remark 2.9. *By the definition of the numerical scheme $\mathbf{S}_h \circ \mathbf{T}_h$ in* (2.10)*, the numerical solution v^h is only defined on the time grid $(t_n)_{0\le n\le n}$ product $\mathbb{R}^d \times \mathbb{R}^{d'}$. However, we can use linear interpolation method to extend v^h on the whole space Q_T.*

Proposition 2.10. *Let Assumptions* **F**, **H** *and* **M** *hold true, then for two functions φ and ψ defined on Q_T with exponential growth, we have*

$$\varphi \le \psi \quad\Longrightarrow\quad \mathbf{S}_h \circ \mathbf{T}_h[\varphi]\;(t,x,y) \;\le\; \mathbf{S}_h \circ \mathbf{T}_h[\psi]\;(t,x,y).$$

Proof. By Lemma 3.12 and Remark 3.13 of [13], $\varphi \le \psi$ implies that $\mathbf{T}_h[\varphi](t,x,y) \le \mathbf{T}_h[\psi](t,x,y)$. Then since $c^{\alpha,\beta} \ge 0$ according to Assumption **M**, it follows immediately by (2.10) that $\mathbf{S}_h \circ \mathbf{T}_h[\varphi](t,x,y) \;\le\; \mathbf{S}_h \circ \mathbf{T}_h[\psi](t,x,y)$. □

We first define a consistency error function, then prove that our splitting scheme $\mathbf{S}_h \circ \mathbf{T}_h$ is consistent.

Definition 2.11. *Given a smooth function φ defined on Q_T, the consistency error function of scheme $\mathbf{S}_h \circ \mathbf{T}_h$ is given by*

$$\Lambda_h^\varphi(\cdot) \;:=\; \frac{\varphi(\cdot) - \mathbf{S}_h \circ \mathbf{T}_h[\varphi](\cdot)}{h} + \mathcal{L}^X \varphi(\cdot) + F(\cdot, \varphi, D_x\varphi, D^2_{xx}\varphi) + H(\cdot, \varphi, D_x\varphi, D_y\varphi). \tag{2.13}$$

The scheme $\mathbf{S}_h \circ \mathbf{T}_h$ is said consistent if

$$\Lambda_h^{\varphi+c}(t', x', y') \to 0 \;\text{ as } (c, h, t', x', y') \to (0, 0, t, x, y), \tag{2.14}$$

for every $(t,x,y) \in Q_T$ and every smooth function φ with bounded derivatives.

Proposition 2.12. *Let Assumptions* **F**, **H** *and* **M** *hold true, then the scheme $\mathbf{S}_h \circ \mathbf{T}_h$ is consistent. In addition, if μ and σ are uniformly bounded, then the consistency error function Λ_h^φ is uniformly bounded by $h\, E(\varphi)$, where*

$$E(\varphi) \;:=\; C \Big(1 \;+\; |\partial_{tt}\varphi|_0 \;+\; \sum_{i=0}^{2} |\partial_t D^i_{z^i}\varphi|_0 \;+\; \sum_{i=0}^{4} |D^i_{z^i}\varphi|_0 \Big) \text{ with } z := (x,y) \in \mathbb{R}^{d+d'},$$

for a constant C independent of φ and h.

Proof. For every $(t,x,y) \in Q_T$, the value $\Lambda_h^\varphi(t,x,y)$ is independent of the value of $(\mu(\bar t, \bar x), \sigma(\bar t, \bar x))$ when $(\bar t, \bar x) \neq (t,x)$. Hence we can always change the value of μ and σ outside the neighborhood of (t,x) without influence on the definition of consistency in (2.14). Therefore, without loss of generality, we can just suppose that μ and σ are uniformly bounded and show that for every smooth function φ with bounded derivatives of any order, the consistency error function Λ_h^φ defined in (2.13) satisfies

$$\big|\Lambda_h^\varphi(\cdot)\big|_0 \;\;\le\;\; h\, E(\varphi). \tag{2.15}$$

First, let us denote

$$\mathcal{L}^{\hat X^{t,x}} \varphi(t', x', y) \;\;:=\;\; \partial_t \varphi(t', x', y) \;+\; \mu(t,x) \cdot D_x \varphi(t', x', y) \;+\; \frac{1}{2} a(t,x) \cdot D^2_{xx} \varphi(t', x', y),$$

then by Itô's formula,

$$\begin{aligned} E^h(t,x,y,\varphi) \;\;&:=\;\; \mathbf{T}_h[\varphi](t,x,y) \;-\; \varphi(t,x,y) \\ &=\;\; h\,\big(\,\mathcal{L}^X\varphi(\cdot) \;+\; F(\cdot, \varphi, D_x\varphi, D^2_{xx}\varphi)\,\big)\,(t,x,y) \\ &\quad + h^2\,\Big(\,\frac{1}{h^2}\,\mathbb{E}\int_t^{t+h}\int_t^u \mathcal{L}^{\hat X^{t,x}}\mathcal{L}^{\hat X^{t,x}}\varphi(s, \hat X^{t,x}_s, y)\,ds\,du\,\Big) \\ &\quad + h^2\,\Big[\,\frac{1}{h}\big(F(\cdot, \mathbb{E}\mathcal{D}_h\varphi)(t,x,y) \;-\; F(\cdot, \varphi, D\varphi, D^2_{xx}\varphi)(t,x,y)\big)\,\Big]. \end{aligned} \tag{2.16}$$

Denote $E_1(t,x,y,\varphi) := \mathcal{L}^X\varphi(t,x,y) + F(\cdot, \varphi, D_x\varphi, D^2_{xx}\varphi)(t,x,y)$ and by $E_2(t,x,y,\varphi)$ the last two terms of the above equality (2.16) divided by h^2, then $E^h(t,x,y,\varphi)$ can rewritten as

$$E^h(t,x,y,\varphi) \;\;=\;\; h\, E_1(t,x,y,\varphi) \;+\; h^2\, E_2(t,x,y,\varphi).$$

Clearly, by the boundedness of μ and σ, together with Assumption **F**, there is a constant C independent of h such that

$$\big|E_2(\cdot,\varphi)\big|_0 \;\;\le\;\; C\,\Big(1 + |\partial_{tt}\varphi|_0 + \sum_{i=0}^{2} |\partial_t D^i_{x^i}\varphi|_0 + \sum_{i=0}^{4} |D^i_{x^i}\varphi|_0\Big),$$

and moreover, E_1 is Lipschitz in $z := (x,y)$ with coefficient

$$L_{E_1} \;\leq\; C\big(\, 1 \;+\; |\partial_t D_z \varphi|_0 \;+\; |D_z\varphi|_0 \;+\; |D^2_{zz}\varphi|_0 \;+\; |D^3_{zzz}\varphi|_0 \,\big).$$

By simplifying $\big(c^{\alpha,\beta}(t,x,y), l^{\alpha,\beta}(t,x,y), f^{\alpha,\beta}(t,x,y), g^{\alpha,\beta}(t,x,y)\big)$ into $(c^{\alpha,\beta}, l^{\alpha,\beta}, f^{\alpha,\beta}, g^{\alpha,\beta})$, we deduce that

$$\begin{aligned}
& \frac{1}{h}\,\big(\, \mathbf{S}_h[(\varphi + E^h(\cdot,\varphi))](t,x,y) \;-\; \varphi(t,x,y) \;-\; E^h(t,x,y,\varphi)\,\big) \\
= \;\; & \frac{1}{h}\, \inf_{\alpha\in\mathcal{A}}\, \sup_{\beta\in\mathcal{B}} \Big[\, h l^{\alpha,\beta} \;+\; h c^{\alpha,\beta}\varphi(t,x,y) \;+\; \varphi(t, x + f^{\alpha,\beta}h,\; y + g^{\alpha,\beta}h) \;-\; \varphi(t,x,y) \\
& \qquad\qquad + \; h c^{\alpha,\beta} E^h(t,x,y,\varphi) \;+ E^h(t, x + f^{\alpha,\beta}h, y + g^{\alpha,\beta}h) - E^h(t,x,y,\varphi)\Big] \\
= \;\; & \inf_{\alpha\in\mathcal{A}}\, \sup_{\beta\in\mathcal{B}} \Big[\; l^{\alpha,\beta} \;+\; c^{\alpha,\beta}\varphi(t,x,y) \;+\; (f^{\alpha,\beta}\cdot D_x\varphi \;+\; g^{\alpha,\beta}\cdot D_y\varphi)(t,x,y) \\
& \qquad + \; \frac{1}{h}\big[\varphi(t, x + f^{\alpha,\beta}h, y + g^{\alpha,\beta}h) - \varphi(t,x,y)\big] - (f^{\alpha,\beta}D_x\varphi + g^{\alpha,\beta}D_y\varphi)(t,x,y) \\
& \qquad\quad + \; c^{\alpha,\beta}\; E^h(t,x,y) \;+\; \frac{1}{h}\big[E^h(t, x + f^{\alpha,\beta}h, y + g^{\alpha,\beta}h, \varphi) \;-\; E^h(t,x,y,\varphi)\big]\Big] \\
=: \;\; & H(\cdot, \varphi, D_x\varphi, D_y\varphi)(t,x,y) + hE_3(t,x,y,\varphi),
\end{aligned} \tag{2.17}$$

where $E_3(t,x,y,\varphi)$ is defined by the last equality of (2.17), and it satisfies

$$|E_3(t,x,y,\varphi)| \;\leq\; C\,\Big(\, |D^2_{zz}\varphi|_0 \;+\; \frac{1}{h}E^h(t,x,y,\varphi) \;+\; 2|E_2(t,x,y,\varphi)|\,\Big) \;+\; L_{E_1} \;\leq\; E(\varphi).$$

Combining the estimations (2.16) and (2.17), and by (2.13) as well as the equality

$$\begin{aligned}
& \frac{\varphi(t,x,y) - \mathbf{S}_h\circ\mathbf{T}_h[\varphi](t,x,y)}{h} \\
= \;\; & \frac{\varphi(t,x,y) - \mathbf{T}_h[\varphi](t,x,y)}{h} \;+\; \frac{\varphi(t,x,y) + E^h(t,x,y,\varphi) - \mathbf{S}_h[\varphi + E^h(\cdot,\varphi)](t,x,y)}{h},
\end{aligned}$$

it follows that (2.15) holds true. □

Proposition 2.13. *Let Assumptions* **F**, **H** *and* **M** *hold true, and the terminal condition function* Φ *be* L^∞*-bounded, then* $(v^h)_h$ *is* L^∞*-bounded, uniformly in* h *for* h *small enough.*

Proof. Suppose that $|v^h(t_{n+1},\cdot)|_0 \leq C_{n+1}$, then from Lemma 3.14 of [13], there exists a constant C independent of h such that

$$\big|v^h\big(t_{n+\frac{1}{2}},\cdot\big)\big|_0 \;\;\leq\;\; C_{n+1}(1+hC) \;+\; hC.$$

It follows from (2.10) that when $h < C^{-1}$,

$$|v^h(t_n,\cdot)|_0 \;\;\leq\;\; (1+hC)(C_{n+1}(1+hC)+hC) \;+\; hC \;\;\leq\;\; (1+3hC)C_{n+1} \;+\; 3hC.$$

Therefore, $|v^h(t_n,\cdot)|_0 \leq C' e^{C'T}$ for some constant C' (independent of h) from the discrete Gronwall inequality. □

We have shown in the above the monotonicity, consistency and stability of scheme $\mathbf{S}_h\circ\mathbf{T}_h$, the rest is to confirm (2.12). In fact, we will provide a little stronger property of $(v^h)_{h>0}$ which implies that

$$\lim_{(t',x',y',h)\to(T,x,y,0)} \; v^h(t',x',y') \;\;=\;\; \Phi(x,y).$$

Proposition 2.14. *Let Assumptions* **F**, **H** *and* **M** *hold true, and* Φ *be Lipschitz and uniformly bounded. Then* $(v^h)_h$ *is Lipschitz in* (x,y)*, uniformly in* h.

Proof. To prove the that v^h is Lipschitz in (x,y), we shall use the discrete Gronwall inequality as in the proof of Lemma 3.16 of [13].

Suppose that $v^h(t_{n+1},\cdot)$ is Lipschitz with coefficient L_{n+1}, then by the proof of Lemma 3.16 of [13], the function $v^h(t_{n+\frac{1}{2}},\cdot) = \mathbf{T}_h[v^h](t_n,\cdot)$ is Lipschitz in x with coefficient $L_{n+1}((1+Ch)^{1/2}+Ch)+Ch$; moreover, $v^h(t_{n+\frac{1}{2}},\cdot)$ is Lipschitz in y with coefficient $L_{n+1}(1+Ch)$ by Lemma 3.14 of [13]. It follows that $v^h(t_{n+\frac{1}{2}},\cdot)$ is Lipschitz in (x,y) with coefficient $L_{n+\frac{1}{2}} \le L_{n+1}((1+Ch)^{1/2}+Ch)+Ch$.

Next, we can easily verify by (2.10) that $v^h(t_n,\cdot)$ is Lipschitz in (x,y) with coefficient $L_n \le L_{n+\frac{1}{2}}(1+Ch)+Ch$. Therefore, the proof is concluded by the discrete Gronwall inequality. □

We can also prove that v^h is $1/2-$Hölder in t as was done in Lemma 3.17 of [13] for their numerical solution. However, to avoid the heavy calculation in their proof, we shall give a weaker result which is enough to guarantee the condition (2.12).

Proposition 2.15. *Let Assumptions* **F**, **H** *and* **M** *hold true, and* Φ *be Lipschitz and uniformly bounded. Then* $|v^h(t_n,x,y)-\Phi(x,y)| \le C\sqrt{T-t_n}$.

Proof. We first introduce $\bar{v}^h$ as the numerical solution of (2.4) computed by scheme $\mathbf{T}_h$, i.e. $\bar{v}^h(T,\cdot) := \Phi(\cdot)$ and $\bar{v}^h(t_n,\cdot) := \mathbf{T}_h[\bar{v}^h](t_n,\cdot)$. Clearly, by Lemmas 3.14 and 3.17 of [13], $(\bar{v}^h)_{h>0}$ is uniformly bounded and satisfies

$$|\bar{v}^h(t_n,\cdot)-\Phi(\cdot)| \quad \le \quad C(T-t_n)^{1/2}, \text{ uniformly in } h. \tag{2.18}$$

We claim that

$$|\bar{v}^h(t_n,x,y)-v^h(t_n,x,y)| \quad \le \quad C(T-t_n). \tag{2.19}$$

Then by (2.18), we conclude the proof. Thus it is enough to prove the claim (2.19).

We first recall that by Assumption **F** and (2.6), for a constant $c\in\mathbb{R}$, we have $\mathbf{T}_h[v^h+c](t,x,y) \le \mathbf{T}_h[v^h](t,x,y)+c+hF_r|c|$. Suppose that for L large enough,

$$|\bar{v}^h(t_{n+1},x,y)-v^h(t_{n+1},x,y)| \quad \le \quad L(T-t_{n+1}).$$

It follows by the monotonicity of $\mathbf{T}_h$ and the uniform boundedness of v^h and $\bar{v}^h$ that

$$|\bar{v}^h(t_n,x,y)-v^h(t_{n+\frac{1}{2}},x,y)| \quad \le \quad L(T-t_{n+1})+Ch.$$

And hence by (2.10),

$$|\bar{v}^h(t_n,x,y)-v^h(t_n,x,y)| \quad \le \quad L(T-t_{n+1})+2Ch \quad \le \quad L(T-t_n),$$

which confirms (2.19). □

We remark finally that with Propositions 2.10, 2.12, 2.13, 2.14 and 2.15 together with Theorem 2.1 of Barles and Souganidis [3], Theorem 2.6 holds true.

2.5 Proof for rate of convergence

As in [13], our arguments to prove the rate of convergence in Theorem 2.7 are based on Krylov's shaking coefficient method, and our analysis stays in the context of Barles and Jakobsen [2]. We first derive some technical Lemmas similar to that in [13].

Lemma 2.16. *Let Assumptions* **F**, **H** *and* **M** *hold true and* $h\le 1$, *define* $\lambda_1 := |F_r|_\infty$, $\lambda_2 := \sup_{\alpha,\beta}|c^{\alpha,\beta}|_0$, $\lambda := \lambda_1+\lambda_2+\lambda_1\lambda_2$. *Then, for every* $(a,b,c)\in\mathbb{R}^3_+$, *and every bounded function* $\varphi\le\psi$ *defined on* Q_T, *with function* $\delta(t) := e^{\lambda(T-t)}(a+b(T-t))+c$, *we have*

$$\mathbf{S}_h\circ\mathbf{T}_h[\varphi+\delta](t,x,y) \ \le \ \mathbf{S}_h\circ\mathbf{T}_h[\psi](t,x,y) \ + \ \delta(t) \ - \ h(b-\lambda c), \ \forall t\le T-h \text{ and } x\in\mathbb{R}^d.$$

Proof. First, from the proof of Lemma 3.21 in [13], we have

$$\mathbf{T}_h[\varphi+\delta](t,x,y) \;\; \le \;\; \mathbf{T}_h[\varphi](t,x,y) \;+\; (1+h\lambda_1)\,\delta(t+h).$$

It follows by the definition of the splitting scheme $\mathbf{S}_h \circ \mathbf{T}_h$ in (2.10) that

$$\mathbf{S}_h \circ \mathbf{T}_h[\varphi+\delta](t,x,y) \;\; \le \;\; \mathbf{S}_h \circ \mathbf{T}_h[\varphi](t,x,y) \;+\; (1+h\lambda_1)(1+h\lambda_2)\,\delta(t+h).$$

By the monotonicity of the splitting scheme $\mathbf{S}_h \circ \mathbf{T}_h$, we get

$$\mathbf{S}_h \circ \mathbf{T}_h[\varphi+\delta](t,x,y) \le \mathbf{S}_h \circ \mathbf{T}_h[\psi](t,x,y) + \delta(t) + \zeta(t), \text{ where } \zeta(t) := (1+h\lambda)\delta(t+h) - \delta(t).$$

Finally, using exactly the same arguments as in the proof of Lemma 3.5 of [13], it follows that

$$\zeta(t) \;\; \le \;\; -\,h(b-\lambda c),$$

which concludes the proof. □

Proposition 2.17. *Let Assumptions* **F**, **H** *and* **M** *hold true,* $h \le 1$ *and* φ, ψ *be two bounded functions defined on* $\mathbb{Q}_T$ *satisfying*

$$\frac{1}{h}\big(\varphi - \mathbf{S}_h \circ \mathbf{T}_h[\varphi]\big) \;\; \le \;\; g_1 \quad \textit{and} \quad \frac{1}{h}(\psi - \mathbf{S}_h \circ \mathbf{T}_h[\psi]) \;\; \ge \;\; g_2, \quad \textit{on } Q_T$$

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

$$(\varphi-\psi)(t_n,x,y) \;\; \le \;\; e^{\lambda(T-t_n)}|(\varphi-\psi)^+(T,\cdot)|_0 \;+\; (T-h)e^{\lambda(T-t_n)}|(g_1-g_2)^+|_0,$$

with some constant $\lambda \ge |F_r|_\infty + \sup_{\alpha,\beta}|c^{\alpha,\beta}|_0 + |F_r|_\infty\ \sup_{\alpha,\beta}|c^{\alpha,\beta}|_0$.

Proof. With Lemma 2.16, the proof is exactly the same as in Proposition 3.20 of [13]. Note that we replace β by λ in our proposition. □

Now, we are ready to give the

Proof of Theorem 2.7 (i). First, under Assumption **HJB**, we can rewrite the original PDE (2.2) as a standard HJB

$$-\,\partial_t v \;\; - \;\; \inf_{\alpha\in\mathcal{A},\gamma\in\mathcal{C}} \Big\{ \,(l^\alpha + l^\gamma) \;+\; (c^\alpha + c^\gamma)v \;+\; (f^\alpha + f^\gamma)\cdot D_x v$$
$$+\; g^\alpha \cdot D_y v \;+\; \frac{1}{2}(\sigma^\gamma \sigma^{\gamma T})\cdot D^2_{xx} v \,\Big\} \;=\; 0.$$

With Assumption **HJB** and the Lipschitz terminal condition, it satisfies a comparison result and admits a unique viscosity solution in $C^{0,1}(Q_T)$ (see e.g. Proposition 2.1 of [2]). Then by the shaking coefficients method, we can construct a bounded subsolution $\underline{v}^\varepsilon \in C^{0,1}(Q_T)$ such that

$$v \;-\; \varepsilon \;\; \le \;\; \underline{v}^\varepsilon \;\; \le \;\; v.$$

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

$$\underline{w}^\varepsilon(t,x,y) \;\; := \;\; \underline{v}^\varepsilon * \rho^\varepsilon, \quad \text{ where } \rho^\varepsilon(t,x,y) \;\; := \;\; \frac{1}{\varepsilon^{d+d'+2}}\,\rho\Big(\frac{t}{\varepsilon^2},\frac{x}{\varepsilon},\frac{y}{\varepsilon}\Big).$$

Then $\underline{w}^\varepsilon$ is a smooth subsolution of (2.2) and satisfies $|\underline{w}^\varepsilon - v| \le 2\varepsilon$. Moreover, since $\underline{v}^\varepsilon \in C^{0,1}(Q_T)$ is uniformly Lipschitz in (x,y) and $1/2-$Hölder in t, it follows that

$$\underline{w}^\varepsilon \in C^\infty, \text{ and } \big|\partial_t^{\eta_0} D^{\eta_1+\eta_2}_{x^{\eta_1}y^{\eta_2}}\underline{w}^\varepsilon\big| \le C\varepsilon^{1-2\eta_0-|\eta_1|-|\eta_2|},\ \forall(\eta_0,\eta_1,\eta_2)\in N^{1+d+d'}\setminus\{0\}. \quad (2.20)$$

Now, let us consider the consistency error function $\Lambda_h^{\underline{w}^\varepsilon}(t,x,y)$ defined in (2.13). By Proposition 2.12 and (2.20), it follows that there exists a constant C independent of ε and h for $0 \le h \le 1$ such that

$$|\Lambda_h^{\underline{w}^\varepsilon}|_0 \;\le\; R(h,\varepsilon) \;:=\; Ch\varepsilon^{-3}. \tag{2.21}$$

Moreover, since $\underline{w}^\varepsilon$ is a subsolution of equation (2.2), it follows by the definition of $\Lambda_h^{\underline{w}^\varepsilon}$ in (2.13) that

$$\underline{w}^\varepsilon \;\le\; \mathbf{S}_h \circ \mathbf{T}_h[\underline{w}^\varepsilon] \;+\; Ch^2\varepsilon^{-3}.$$

Finally, by Proposition 2.17, we get

$$\underline{w}^\varepsilon - v^h \;\le\; C\,(\varepsilon + h\varepsilon^{-3}), \quad \text{and} \quad v \,-\, v^h \;=\; v \,-\, \underline{w}^\varepsilon \,+\, \underline{w}^\varepsilon \,-\, v^h \;\le\; C(\varepsilon \,+\, h\varepsilon^{-3})$$

and it follows by a minimization technique on ε that

$$v \,-\, v^h \;\le\; C \inf_{\varepsilon>0} \big(\, \varepsilon \,+\, h\varepsilon^{-3} \,\big) \;\le\; C' h^{\frac{1}{4}}. \tag{2.22}$$

$\square$

Proof of Theorem 2.7 (ii) : Under Assumption **HJB+**, we can apply the switching system method of Barles and Jakobsen [2] which constructs a smooth supersolution closed to viscosity solution to PDE (2.2) and provides the lower bound:

$$v \,-\, v^h \;\ge\; -\inf_{\varepsilon>0} \big(\, C\varepsilon^{\frac{1}{3}} \,+\, R(h,\varepsilon) \,\big) \;=\; -\,C' h^{\frac{1}{10}}, \tag{2.23}$$

where $R(h,\varepsilon)$ is defined in (2.21). $\square$

3 Basis projection and simulation-regression method

To get an implementable scheme, we need to specify how to compute the expectations $\mathbb{E}\Big[\varphi(t_{n+1}, \hat{X}_h^{t_n,x}, y) H_i^{t_n,x,h}(\Delta W_{n+1})\Big]_{i=0,1,2}$ in the splitting scheme $\mathbf{S}_h \circ \mathbf{T}_h$. When analytic closed formulas are not available in the concrete examples, we usually use Monte-Carlo simulation-regression method to estimate them. Some estimations were discussed in recent works, e.g. Malliavin estimations [5], function basis regression [14] and cubature method [10], etc.

All of these methods need the simulations of X. Given a discrete time grid $(t_n)_{0\le n\le N}$, where $t_n := n\,h$ and $h := T/N$, we define a Euler approximation $\hat{X}$ of X

$$\hat{X}_{t_{n+1}} \;:=\; \hat{X}_{t_n} \,+\, \mu(t_n, \hat{X}_{t_n})h \,+\, \sigma(t_n, \hat{X}_{t_n})\Delta W_{n+1}, \tag{3.1}$$

where $\Delta W_{n+1} := W_{t_{n+1}} - W_{t_n}$. Then with simulations of process $\hat{X}$ as well as W, one can estimate the conditional expectations

$$\mathbb{E}\;\Big[\; \varphi(t_{n+1}, \hat{X}_{t_{n+1}}, y)\; H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1}) \;\Big|\; \hat{X}_{t_n} \;\Big]_{i=0,1,2}.$$

However, these methods are usually discussed in a non-degenerate context, in other words, they can be used for a given fixed y, which is not appropriate for the implementation of our splitting scheme $\mathbf{S}_h \circ \mathbf{T}_h$.

One solution is to discretize the space of Y into a discrete grid $(y_i)_{i\in I}$, and then for each fixed y_i, we simulate the diffusion process X and get estimations of the conditional expectations for all x with every fixed y_i, then use the interpolation method to get the estimation of theses expectations for all x and y. This is a combination of finite difference method and Monte-Carlo method, which may lose the advantages of Monte-Carlo method in high dimensional cases.

Therefore, we propose to simulate the diffusion process X with Euler scheme and to simulate Y with a continuous probability distribution (e.g. normal distribution, uniform distribution, etc.) independent of X. And then we use a regression method like in Longstaff and Schwartz [15] in American option pricing context or Gobet, Lemor and Warin [14] in BSDE context to estimate the conditional expectations

$$\mathbb{E}\left[\varphi(t_{n+1},\hat{X}_{t_{n+1}},Y)H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1})\;\Big|\;\hat{X}_{t_n},\;Y\right]_{i=0,1,2}, \tag{3.2}$$

with which we shall make the splitting scheme $\mathbf{S}_h\circ\mathbf{T}_h$ implementable.

Remark 3.1. (i) *The distribution of Y may be chosen arbitrarily according to the concrete context.*
(ii) *In practice, if we choose local hypercubes or local polynomials as functions basis for the regression method, we still need to discretize the space. However, as discussed in the introduction, this discretization can be coarse in practice, which permits to keep the advantage of the simulation-regression method in high-dimensional cases (see also the numerical examples in Section 4).*

In the following, we first give a basis projection scheme as well as a similation-regression method to estimate the regression coefficient. Then we discuss the convergence of Monte-Carlo errors in our context.

3.1 Basis projection scheme and simulation-regression method

3.1.1 The basis projection scheme

To compute the conditional expectations (3.2), we first project them on a functional space spanned by the basis functions $(e_k(x,y))_{1\le k\le K}$, where $K\in\mathbb{N}\cup\{+\infty\}$. We recall that $H_2^{t,x,h}$ is a matrix of dimension $d\times d$, $H_1^{t,x,h}$ is a vector of dimension d and $H_0^{t,x,h}=1$. In order to simplify the presentation, we shall suppose that $d=d'=1$. All of the results can be easily extended to the case $d>1$ and/or $d'>1$. Let

$$\tilde{\lambda}^i\;:=\;\arg\min_{\lambda}\;\mathbb{E}\left(\varphi(t_{n+1},\hat{X}_{t_{n+1}},Y)H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1})\;-\;\sum_{k=1}^{K}\lambda_k e_k(\hat{X}_{t_n},Y)\right)^2, \tag{3.3}$$

then the projected approximation of (3.2) is denoted by

$$\tilde{\mathbb{E}}\left[\,\varphi(t_{n+1},\hat{X}_{t_{n+1}},Y)H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1})\;\big|\;\hat{X}_{t_n},Y\,\right]\;:=\;\sum_{k=1}^{K}\tilde{\lambda}_k^i e_k(\hat{X}_{t_n},Y). \tag{3.4}$$

Remark 3.2. *There are several choices for function basis $(e_k(x,y))_{1\le k\le K}$, for example global polynomials, local hypercubes or local polynomials, we refer to Bouchard and Warin [6] for some interesting discussions.*

We replace the conditional expectations (3.2) in scheme $\mathbf{S}_h\circ\mathbf{T}_h$ by their projected approximations (3.4), and denote the new splitting scheme by $\mathbf{S}_h\circ\tilde{\mathbf{T}}_h$. Concretely, it is defined as follows:

$$\tilde{\mathbf{T}}_h[\tilde{v}^h](t_n,x,y)\;\;:=\;\;\tilde{\mathbb{E}}\left[\tilde{v}^h(t_{n+1},\hat{X}_h^{t_n,x},y)\right]\;+\;hF(\cdot,\tilde{\mathbb{E}}\mathcal{D}\tilde{v}^h(\cdot))(t_n,x,y),$$

where

$$\tilde{\mathbb{E}}\mathcal{D}_h\varphi(t_n,x,y)\;\;=\;\;\left(\;\tilde{\mathbb{E}}\left[\varphi(t_{n+1},\hat{X}_h^{t_n,x},y)H_i^{t_n,x,h}(\Delta W_{n+1})\right]:i=0,1,2\;\right),$$

and

$$\begin{aligned}\tilde{v}^h(t_n,x,y) &= \mathbf{S}_h \circ \tilde{\mathbf{T}}_h[\tilde{v}^h](t_n,x,y) \\ &:= \inf_\alpha \sup_\beta \Big\{ hl^{\alpha,\beta}(t_n,x,y) \ + \ hc^{\alpha,\beta}(t_n,x,y)\tilde{\mathbf{T}}_h[\tilde{v}^h](t_n,x,y) \\ &\qquad + \ \tilde{\mathbf{T}}_h[\tilde{v}^h]\big(t_n,\ x+f^{\alpha,\beta}(t_n,x,y)h,\ y+g^{\alpha,\beta}(t_n,x,y)h\big) \Big\}. \end{aligned} \tag{3.5}$$

3.1.2 Simulation-regression method

Next, we propose to use a simulation-regression method to approximate $\tilde{\lambda}$. We still suppose that $d = d' = 1$ for simplicity.

Let $\big((\hat{X}^m_{t_n})_{0\le n\le N}, (\Delta W^m_n)_{0<n\le N}, Y^m\big)_{1\le m\le M}$ be M independent simulations of $\hat{X}$, ΔW and Y, where $\hat{X}$ is defined in (3.1), the regression method with function basis $(e_k(x,y))_{1\le k\le K}$ is to get the solution of the least square problem:

$$\hat{\lambda}^{i,M} = \arg\min_\lambda \ \sum_{m=1}^M \Big(\varphi(t_{n+1},\hat{X}^m_{t_{n+1}},Y^m)H_i^{t_n,\hat{X}^m_{t_n},h}(\Delta W^m_{n+1}) - \sum_{k=1}^K \lambda_k e_k(\hat{X}^m_{t_n},Y^m)\Big)^2. \tag{3.6}$$

A raw regression estimation of the conditional expectations (3.2) from these M samples is given by

$$\bar{\mathbb{E}}^M\Big[\varphi(t_{n+1},\hat{X}_{t_{n+1}},Y)H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1})\big|\hat{X}_{t_n},Y\Big] \ := \ \sum_{k=1}^K \hat{\lambda}^{i,M}_k e_k(\hat{X}_{t_n},Y),\ i=0,1,2. \tag{3.7}$$

Then with a priori upper bounds $\overline{\Gamma}_i(\hat{X}_{t_n},Y)$ and lower bounds $\underline{\Gamma}_i(\hat{X}_{t_n},Y)$, we define the regression estimation of (3.2):

$$\begin{aligned} &\hat{\mathbb{E}}^M\Big[\ \varphi(t_{n+1},\hat{X}_{t_{n+1}},Y)H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1})\ \big|\ \hat{X}_{t_n},Y\ \Big] \qquad (3.8)\\ :=\ &\underline{\Gamma}_i(\hat{X}_{t_n},Y)\ \vee\ \bar{\mathbb{E}}^M\Big[\varphi(t_{n+1},\hat{X}_{t_{n+1}},Y)H_i^{t_n,\hat{X}_{t_n},h}(\Delta W_{n+1})\big|\hat{X}_{t_n},Y\Big]\ \wedge\ \overline{\Gamma}_i(\hat{X}_{t_n},Y). \end{aligned}$$

Remark 3.3. *As observed in Bouchard and Touzi [5], the truncation method is an important technique to obtain a L^p-convergence. By Lemma* (2.15)*, we can choose $\overline{\Gamma}_0(x,y) = \Gamma_0(x,y)$ and $\underline{\Gamma}_0(x,y) = -\Gamma_0(x,y)$ with a function Γ_0 satisfying*

$$\Gamma_0(x,y) \quad \le \quad \Phi(x,y)\ +\ C\sqrt{T-t_n} \quad \text{for some constant } C. \tag{3.9}$$

Remark 3.4. *In Gobet et al. [14], the authors propose the following minimization problem in place of* (3.6)*:*

$$\min_{\lambda^0,\lambda^1} \ \sum_{m=1}^M \Big(\varphi(t_{n+1},\hat{X}^m_{t_{n+1}},Y^m)\ -\ \sum_{k=1}^K \lambda^0_k e_k(\hat{X}^m_{t_n},Y^m)\ -\ \sum_{k=1}^K \lambda^1_k e_k(\hat{X}^m_{t_n},Y^m)\Delta W^m_{n+1}\Big)^2,$$

which gives also a good estimation for $\tilde{\lambda}^i$ by the fact that ΔW_{n+1} is independent of the $\sigma-$field generated by $Y, W_0, \Delta W_1, \cdots, \Delta W_n$.

We replace the conditional expectations (3.2) in scheme $\mathbf{S}_h \circ \mathbf{T}_h$ by their regression estimations (3.8) and denote the new numerical splitting scheme by $\mathbf{S}_h \circ \hat{\mathbf{T}}^M_h$, which is

$$\hat{\mathbf{T}}^M_h[\hat{v}^h](t_n,x,y) \quad := \quad \hat{\mathbb{E}}^M\big[\hat{v}^h(t_{n+1},\hat{X}^{t_n,x}_h,y)\big]\ +\ h\ F(\cdot,\hat{\mathbb{E}}^M\mathcal{D}\hat{v}^h(\cdot))(t_n,x,y),$$

and

$$\hat{\mathbb{E}}^M \mathcal{D}_h \varphi(t_n, x, y) \;=\; \left(\hat{\mathbb{E}}^M \left[\varphi(t_{n+1}, \hat{X}_h^{t_n,x}, y) H_i^{t_n,x,h}(\Delta W_{n+1}) \right] \;:\; i = 0,1,2 \right),$$

so that $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$ is defined by

$$\begin{aligned} \hat{v}^h(t_n, x, y) &= \mathbf{S}_h \circ \hat{\mathbf{T}}_h^M[\hat{v}^h](t_n, x, y) \\ &:= \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in \mathcal{B}} \Big\{ h l^{\alpha,\beta}(t_n, x, y) \;+\; h c^{\alpha,\beta}(t_n, x, y) \hat{\mathbf{T}}_h^M[\hat{v}^h](t_n, x, y) \\ &\qquad + \; \hat{\mathbf{T}}_h^M[\hat{v}^h]\big(t_n,\; x + f^{\alpha,\beta}(t_n, x, y)h,\; y + g^{\alpha,\beta}(t_n, x, y)h\big) \Big\}. \end{aligned} \tag{3.10}$$

3.2 The convergence results of simulation-regression scheme

To get a convergence result of schemes $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$ and $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$, we can no longer use the same arguments as in Fahim et al. [13], since there is no uniform convergence property in L^p for the Monte-Carlo error $(\hat{\mathbb{E}}^M - \mathbb{E})(R)$ as in the Assumption **E** of [13]. To see this, let us consider the extreme case where the equation is totally degenerate (i.e. $d = 0$ and $d' > 0$), and then we need to approximate an arbitrary bounded function in a functional space with finite number of basis functions, which does not give a uniform convergence. Also, since we are in the viscosity solution analysis context of Barles and Souganidis [3], we can not hope to obtain a probabilistic $L^2(\Omega)-$convergence as in Gobet et al. [14].

However, we can get a convergence result if we choose the local hypercubes as function basis. Let us restrict the numerical resolution on $[0,T] \times D$ instead of Q_T, where $D \subset \mathbb{R}^{d+d'}$ is a bounded domain. Clearly, we need to assume that the boundary conditions on the domain $D^c := \mathbb{R}^{d+d'} \setminus D$ are available for scheme $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$.

Definition 3.5. *Given a domain $D \subseteq \mathbb{R}^{d+d'}$, a class of hypercube sets $(B_k)_{1 \le k \le K}$ is called a partition of D whenever $\cup_{k=1}^K B_k = D$ and $B_i \cap B_j = \emptyset$.*

Remark 3.6. *The simplest examples of partition of D is the uniform partition. With uniform interval $[x_k, x_k')$ and $[y_k, y_k')$, B_k are of the form $[x_k, x_k') \times [y_k, y_k')$. Recently, Bouchard and Warin [6] proposed a partition based on the simulations. They first sort all the simulations and then divide the space in a non-uniform way such that they have the same number of simulation particles in every hypercube B_k.*

Remark 3.7. *If we use hypercubes $(\mathbf{1}_{B_k})_{\mathbf{1} \le k \le K}$ as basis function in the projections* (3.3)*, where $(B_k)_{1 \le k \le K}$ is a partition of $D \subseteq \mathbb{R}^{d+d'}$, then the projection approximation is equivalent to taking another conditional expectation on the σ-field generated by $\big\{(X_{t_n}, Y) \in B_k\big\}_{1 \le k \le K}$, in other words,*

$$\begin{aligned} &\tilde{\mathbb{E}}\left[\, \varphi(t_{n+1}, \hat{X}_{t_{n+1}}, Y) H_i^{t_n, \hat{X}_{t_n}, h}(\Delta W_{n+1}) \;\middle|\; \hat{X}_{t_n}, Y \,\right] \\ = \;& \sum_{k=1}^K \mathbb{E}\Big[\, \varphi(t_{n+1}, \hat{X}_{t_{n+1}}, Y) H_i^{t_n, \hat{X}_{t_n}, h}(\Delta W_{n+1}) \;\Big|\; (\hat{X}_{t_n}, Y) \in 1_{B_k} \,\Big] \, \mathbf{1}_{B_k}(\hat{X}_{t_n}, Y). \end{aligned} \tag{3.11}$$

Let us use $(e_k)_{1 \le k \le K} = (\mathbf{1}_{B_k})_{1 \le k \le K}$ as projection basis in (3.3) and (3.6), where $(B_k)_{1 \le k \le K}$ is a partition of D. Given a bounded function φ on D, a process $\hat{X}$ and a random variable Y, we shall consider the random variables of the form

$$R_i(\varphi) \;:=\; \varphi(t_{n+1}, \hat{X}_{t_{n+1}}, Y)\; H_i^{t_n, \hat{X}_{t_n}, h}(\Delta W_{n+1}), \quad i = 0,1,2, \tag{3.12}$$

and then give an estimation for the regression error $(\hat{\mathbb{E}}^M - \tilde{\mathbb{E}})\,[R_i(\varphi) \mid \hat{X}_{t_n} = x, Y = y]$.

Lemma 3.8. *Suppose that the a priori estimations used in* (3.8) *satisfy*

$$\underline{\Gamma}_i(x,y)^2 \ + \ \overline{\Gamma}_i(x,y)^2 \quad \leq \quad C\,\Gamma(x,y)^2\,h^{-i}, \quad \textit{for some function}\ \Gamma(x,y).$$

Then for every $(x,y) \in B_k$*,*

$$\mathbb{E}\Big[\ (\hat{\mathbb{E}}^M - \tilde{\mathbb{E}})^2\ [R_i(\varphi) \mid \hat{X}_{t_n} = x, Y = y]\ \Big] \quad \leq \quad C\ \frac{1}{M}\ h^{-i}\,\frac{|\varphi|_0^2\ +\ \Gamma^2(x,y)}{\mathbb{P}((\hat{X}_{t_n},Y) \in B_k)}. \qquad (3.13)$$

The proof is almost the same as that of Theorem 5.1 of Bouchard and Touzi [5], we report it in Appendix for completeness.

Let φ be bounded by constant b, δ denote the longest edge of the hypercubes $(B_k)_{1\leq k\leq K}$, then the volume of B_k is of order $\delta^{d+d'}$, and $\mathbb{P}((\hat{X}_{t_n},Y) \in B_k) \approx C\delta^{d+d'}$, where C depends on the density of $(\hat{X}_{t_n},Y)$. As the total volume of D is fixed and finite, let

$$\hat{C}(\delta) \quad := \quad \sup_{N,n,k,x,y}\ C\ \frac{1}{M}\ h^{-i}\,\frac{b^2\ +\ \Gamma^2(x,y)}{\mathbb{P}((\hat{X}_{t_n},Y) \in B_k)}, \qquad (3.14)$$

it follows that $\hat{C}(\delta) \approx C\delta^{-(d+d')}$.

Now, let us give a local uniform convergence as well as a rate of convergence for the simulation-regression scheme $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$.

Theorem 3.9. *Let Assumptions* **F***,* **H** *and* **M** *hold true,* F *be uniformly bounded,* Φ *be bounded and Lipschitz continuous, and the PDE* (2.2) *satisfy a comparison result for bounded viscosity solutions. In addition, given a time step* h*, there is a* D*-partition hypercubes* $(B_k^h)_{1\leq k\leq K_h}$ *with edge* δ_h *such that* $\delta_h\ h^{-1} \to 0$ *as* $h \to 0$*. Let the truncation function* Γ_0 *satisfies* (3.9)*, and we use hypercubes* $(1_{B_k^h})_{1\leq k\leq K_h}$ *as projection basis functions and with sample number* $M = M_h$ *such that* $\hat{C}(\delta_h)\ h^{-2}M_h^{-1} \to 0$*, where* $\hat{C}(\delta_h)$ *is defined in* (3.14)*. Then there exists a function* v*, such that*

$$\hat{v}^h \quad \to \quad v \quad \textit{locally uniformly, a.s.}$$

where $\hat{v}^h$ *is the numerical solution of scheme* $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$ *defined in* (3.10) *with terminal condition* Φ*. Moreover,* v *is the unique bounded viscosity solution of* (2.2) *and* (2.3)*.*

Theorem 3.10. *Let Assumption* **HJB+** *hold,* Φ *be bounded and Lipschitz continuous, and assume that we use hypercubes* $(1_{B_k^h})_{1\leq k\leq K_h}$ *as projection basis functions whose longest edge satisfies* $\delta_h \leq Ch^{\frac{11}{10}}$*, and we choose simulation number* $M = M_h$ *such that*

$$\limsup_{h\to 0}\ h^{-\frac{1}{20}-2}\ \hat{C}(\delta)\ M_h^{-1} \quad < \quad \infty.$$

Then there is a constant $C > 0$*, s.t.*

$$\|v - \hat{v}^h\|_{L^2(\Omega)} \quad \leq \quad Ch^{\frac{1}{10}},$$

where $\hat{v}^h$ *is the numerical solution of scheme* $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$ *in* (3.10) *with terminal condition* Φ *and* v *is the unique bounded viscosity solution of* (2.2) *and* (2.3)*.*

3.3 Some analysis on the basis projection scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$

In preparation of the proof for Theorems 3.9 and 3.10, we give some analysis on the scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$. In general, we shall show that if we use the local hypercubes as projection function basis, then $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$ is still monotone, consistent and stable.

Proposition 3.11. *Let* $(B_k)_{1\leq k\leq K}$ *be a partition of domain* D*, and the three projections (*$i = 0,1,2$*) of* (3.3) *use the same hypercubes* $(1_{B_k})_{1\leq k\leq K}$ *as projection function basis. Then under Assumptions* **F***,* **H** *and* **M***,*

- *i) The basis projection scheme* $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$ *is monotone.*
- *ii) If the terminal condition* Φ *is uniformly bounded, then the numerical solution* $\tilde{v}^h$ *of scheme* $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$ *in* (3.5) *is uniformly bounded for* h *small enough.*

Proof. In view of Remark 3.7, we replace the conditional expectations in $\mathbf{S}_h \circ \mathbf{T}_h$ by the new conditional expectations (3.11), and then get the projection scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$. Therefore, all the arguments still hold in the proof of Lemma 3.2 and 3.3 of [13] for $\tilde{\mathbf{T}}_h$, so do Propositions 2.10 and 2.13. Therefore, Proposition 3.11 holds true. □

Similar to the consistency error function Λ_h^φ for scheme $\mathbf{S}_h \circ \mathbf{T}_h$ defined in (2.13), we define the consistency error function $\tilde{\Lambda}_h^\varphi$ for scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$ by

$$\tilde{\Lambda}_h^\varphi(\cdot) := \frac{\varphi(\cdot) - \mathbf{S}_h \circ \tilde{\mathbf{T}}_h[\varphi](\cdot)}{h} + \mathcal{L}^X \varphi(\cdot) + F(\cdot, \varphi, D_x\varphi, D_{xx}^2\varphi) + H(\cdot, \varphi, D_x\varphi, D_y\varphi). \tag{3.15}$$

Proposition 3.12. *Let* δ *denote the longest edge of hypercubes* $(B_k)_{1\le k\le K}$*, then the projection error for a Lipschitz continuous function is proportional to* δ*. Moreover, if we use hypercubes* $(1_{B_k})_{1\le k\le K}$ *as projection function basis, then under Assumptions* **F**, **H** *and* **M**, *the consistency error function* $\tilde{\Lambda}_h^\varphi$ *is uniformly bounded by* $h\tilde{E}(\varphi)$, *where*

$$\tilde{E}(\varphi) \quad := \quad E(\varphi) + Ch^{-1}\,\delta\big(\,|D_z\varphi|_0 \;+\; h|D_{zz}^2\varphi|_0 \;+\; h|D_{zzz}^3\varphi|_0\,\big), \quad \textit{for } z := (x,y),$$

with $E(\varphi)$ *defined in Proposition 2.12.*

Proof. In view of Remark 3.7, the error caused by conditional expectation on hypercube is bounded by $C\delta|D_{z^{i+1}}^{i+1}\varphi|_0$ for $D_{z^i}^i\varphi$. Thus we get immediately the new consistency error $\tilde{E}(\varphi)$ with Proposition 2.12. □

Proposition 3.13. *Suppose that the three projections in* (3.3) *use the same* D*-partition hypercubes as projection function basis, then Lemma 2.16 and Proposition 2.17 hold true if we replace the scheme* $\mathbf{S}_h \circ \mathbf{T}_h$ *by* $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$.

Proof. With the Proposition 3.11 and under Assumptions **F**, **H** and **M**, we see that all the arguments are still true in the proofs of Lemma 2.16 and Proposition 2.17 for scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$, in view of Remark 3.7. So we get the same results for the basis projection scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$. □

3.4 The proof for convergence results of scheme $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$

To prove Theorem 3.9, we shall mimic the proof of Theorem 4.1 in [13], which uses the arguments of [3] in a stochastic context.

Proof of Theorem 3.9. Given $\hat{v}^h$ the numerical solution of scheme $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$, we denote

$$\hat{v}_*(t,x,y) := \liminf_{(t',x',y',h)\to(t,x,y,0)} \hat{v}^h(t',x',y'), \quad \hat{v}^*(t,x,y) := \limsup_{(t',x',y',h)\to(t,x,y,0)} \hat{v}^h(t',x',y').$$

First, it is clear by the truncation function (3.9) as well as the boundedness of F that $|v(t_n,x,y) - \Phi(x,y)| \le C(T-t_n)$ for some constant C, which implies that $\hat{v}_*(T,x,y) = \hat{v}^*(T,x,y) = \Phi(x,y)$. Then it is enough to prove that $\hat{v}_*$ and $\hat{v}^*$ are respectively viscosity supersolution and subsolution of (2.2) to conclude the proof with the comparison assumption. Here we shall only prove the supersolution property, since the subsolution property holds true with the same kind of argument.

Given $(t_0,x_0,y_0) \in Q_T$ and a test function $\varphi \in C_c^\infty(Q_T)$ such that

$$0 \;=\; \min(\hat{v}_* - \varphi) \;=\; (\hat{v}_* - \varphi)(t_0,x_0,y_0),$$

by uniform boundedness of $\hat{v}^h$ and manipulation on φ, there is a sequence $(t_k, x_k, y_k, h_k) \to (t_0, x_0, y_0, 0)$ such that $\hat{v}^{h_k}(t_k, x_k, y_k) \to \hat{v}_*(t_0, x_0, y_0)$ and

$$C_k \;:=\; \min(\hat{v}^{h_k} \,-\, \varphi) \;=\; (\hat{v}^{h_k} \,-\, \varphi)(t_k, x_k, y_k) \;\to\; 0.$$

From the monotonicity of scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$, it follows that

$$\mathbf{S}_h \circ \tilde{\mathbf{T}}_h[\,\hat{v}^{h_k}\,] \;\; \geq \;\; \mathbf{S}_h \circ \tilde{\mathbf{T}}_h[\,\varphi \,+\, C_k\,],$$

and hence

$$\begin{aligned} 0 \;&=\; \hat{v}^{h_k}(t_k, x_k, y_k) \,-\, \mathbf{S}_h \circ \hat{\mathbf{T}}_h^M[\hat{v}^{h_k}](t_k, x_k, y_k) \\ &=\; \hat{v}^{h_k}(t_k, x_k, y_k) \,-\, \mathbf{S}_h \circ \tilde{\mathbf{T}}_h[\hat{v}^{h_k}](t_k, x_k, y_k) \,+\, h_k R_k \\ &\leq\; \varphi(t_k, x_k, y_k) \,+\, C_k \,-\, \mathbf{S}_h \circ \tilde{\mathbf{T}}_h[\varphi + C_k](t_k, x_k, y_k) \,+\, h_k R_k, \end{aligned}$$

where $R_k := h_k^{-1}(\mathbf{S}_{h_k} \circ \hat{\mathbf{T}}_{h_k}^M - \mathbf{S}_{h_k} \circ \tilde{\mathbf{T}}_{h_k})[\hat{v}^{h_k}](t_k, x_k, y_k)$. We claim that

$$R_k \;\to\; 0 \quad \mathbb{P}\text{-a.s. along some subsequence.} \tag{3.16}$$

Then, from the consistence of scheme $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$ in Proposition 3.12,

$$\left[-\mathcal{L}^X\varphi - F(\cdot, \varphi, D_x\varphi, D_{xx}^2\varphi) - H(\cdot, \varphi, D_x\varphi, D_y\varphi)\right](t_0, x_0, y_0) \;\;\geq\;\; 0,$$

which is the required supersolution property.

Therefore, it is enough to justify the claim (3.16) to conclude the proof. Indeed, by the definition of splitting scheme $\mathbf{S}_h \circ \hat{\mathbf{T}}_h^M$ and $\mathbf{S}_h \circ \tilde{\mathbf{T}}_h$, and the boundedness of $c^{\alpha,\beta}$,

$$\begin{aligned} \mathbb{E}|R_k|^2 \;\;&\leq\;\; (1 + Ch_k)^2\, \frac{1}{h_k^2}\, \mathbb{E}\,\Big[\,\tilde{\mathbf{T}}_{h_k}[\hat{v}^{h_k}] - \hat{\mathbf{T}}_{h_k}^M[\hat{v}^{h_k}]\,\Big]^2(t_k, x_k, y_k) \\ &\leq\;\; C\,(1 + Ch_k)^2\, \frac{1}{h_k^2}\, (\tilde{\mathbb{E}} - \hat{\mathbb{E}}^M)^2\, \Big[R_0(\hat{v}^{h_k}) \,+\, h_k R_1(\hat{v}^{h_k}) \,+\, h_k R_2(\hat{v}^{h_k})\Big], \end{aligned}$$

where $R_i(\hat{v}^{h_k})$ is defined in (3.12). And therefore by Lemma 3.8

$$\begin{aligned} \mathbb{E}|R_k|^2 \;\;&\leq\;\; C\,(1 + Ch_k)^2\, \frac{1}{h_k^2}\Big(\hat{C}(\delta) \,+\, h_k^2\,\big(\frac{1}{h_k}\hat{C}(\delta) \,+\, \frac{1}{h_k^2}\hat{C}(\delta)\big)\Big)\, \frac{1}{M_{h_k}} \\ &\leq\;\; C\, h_k^{-2}\, \hat{C}(\delta)\, M_{h_k}^{-1}, \end{aligned}$$

which turns to 0 by assumptions of the theorem. Further, the L^2-convergence implies that $R_k \to 0$ in probability and hence it admits a subsequence which converges to 0 almost surely. We then proved the claim (3.16) and hence conclude the proof of the theorem. □

Proof of Theorem 3.10. With Proposition 3.13, we can proceed as in the proof of Theorem 2.7. Then there is a subsolution $\underline{w}^h$ of (2.2) such that

$$v \;\leq\; \underline{w}^h \,+\, C\epsilon \quad \text{and} \quad \underline{w}^h \,-\, \tilde{v}^h \;\;\leq\;\; C(\,h\epsilon^{-3} \,+\, h^{-1}\delta \,+\, \delta\epsilon^{-2}\,).$$

Moreover, since

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

it follows from Proposition 3.13 that $\tilde{v}^h - \hat{v}^h \leq C|R_h[\hat{v}^h]|$, and then

$$v \,-\, \hat{v}^h \;\;=\;\; v \,-\, \tilde{v}^h \,+\, \tilde{v}^h \,-\, \hat{v}^h \;\;\leq\;\; C\left(\epsilon \,+\, h\epsilon^{-3} \,+\, h^{-1}\delta \,+\, \delta\epsilon^{-2} \,+\, \left|R_h[\hat{v}^h]\right|\right).$$

Similarly, we have the other side of the error boundary and get

$$\left|v-\hat{v}^h\right|^2 \quad \leq \quad C\Big(\,(\epsilon^{\frac{1}{3}}+h\epsilon^{-3}+h^{-1}\delta+\delta\epsilon^{-2})^2 \;+\; \left|R_h[\hat{v}^h]\right|^2\;\Big). \tag{3.17}$$

Finally, it is enough to take expectations on both sides of (3.17) and maximize the right side on ϵ for $\epsilon_h = h^{\frac{3}{10}}$, which implies that

$$\mathbb{E}\left|v-\hat{v}^h\right|^2 \quad \leq \quad C\Big(h^{\frac{1}{20}} \;+\; \frac{1}{M_h}\frac{1}{h^2}\hat{C}(\delta)\Big) \quad \leq \quad C'h^{\frac{1}{20}}.$$

□

4 Numerical examples

We provide here some numerical examples, one is from Asian option pricing problem and the other is from an optimal management problem for a hydropower plant. In every numerical example, we use local polynomial basis for the simulation regression method. The space is discretized to get the local basis, where we use 5 discretization points for every dimension, i.e. the space is divided into 6 parts along every dimension. The polynomials are of second order degree, i.e. they are of the form $a_0+a_1x+a_2x^2$ in the one-dimensional case. Further, we give also the computation time of each numerical example using a computer with 2.4GHz CPU and 4G memory.

4.1 Asian option pricing

Our first example is to price Asian option in an uncertain volatility model (UVM), whose pricing equation is a degenerate and nonlinear PDE. Then we also consider the problem in UVM with Hull-White interest rate.

4.1.1 Asian option pricing in UVM: a two-dimensional case

We consider an uncertain volatility model with risky asset S_t given by

$$dS_t \;\;=\;\; rS_tdt \;\;+\;\; \sigma_tS_tdW_t,$$

where r is the constant interest rate, σ is the volatility process which is bounded between the lower volatility $\underline{\sigma}$ and the upper volatility $\overline{\sigma}$. Denote

$$A_t \;\;:=\;\; \int_0^t S_u\,du,$$

an Asian option is an option with payoff $g(S_T,A_T)$ at maturity T, whose pricing equation is

$$\Big(\;\partial_t v \;+\; rsD_sv \;+\; \frac{1}{2}\max_{\underline{\sigma}\leq\sigma\leq\overline{\sigma}}\sigma^2s^2D^2_{ss}v \;+\; sD_av \;-\; rv\;\Big)(t,s,a) \;\;=\;\; 0, \tag{4.1}$$

with terminal condition $v(T,s,a)=g(s,a)$.

To implement the splitting scheme, we rewrite (4.1) in form of the equation (2.2) with some constant σ_0:

$$\partial_t v \;+\; rsD_sv \;+\; \frac{1}{2}\sigma_0^2s^sD^2_{ss}v \;+\; \frac{1}{2}\max_{\underline{\sigma}\leq\sigma\leq\overline{\sigma}}(\sigma^2-\sigma_0^2)s^2D^2_{ss}v \;+\; sD_av \;-\; rv \;=\; 0. \tag{4.2}$$

Further we consider a call spread type payoff $g(S,A)=(A-K_1)^+-(A-K_2)^+$. With 50 independent computations for every time discretization using the splitting scheme, we get the mean value as well as its standard deviation. Moreover, as comparison, we implemented the Crank-Nicolson finite difference scheme of equation (4.1) with parameters $\Delta S = 1$ and $\Delta A = 0.25$. The results of our splitting scheme and Crank-Nicolson scheme for different Δt are given in Figure 1. We notice that the standard deviation of the splitting method price from 50 independent computations is less than 1% of the mean value and the relative difference between the two schemes are less than 0.3%.

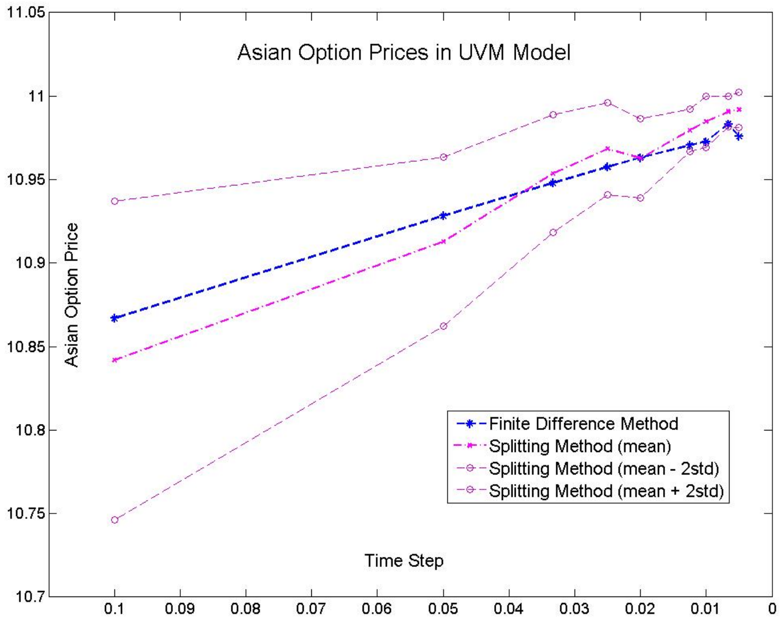

Figure 1: The comparison of some numerical methods for pricing Asian option with payoff $(A - K_1)^+ - (A - K_2)^+$ in UVM, with parameters $S_0 = 100$, $K_1 = 90$, $K_2 = 110$, $T = 1$, $r = 0.05$, $\underline{\sigma} = 0.18$, $\overline{\sigma} = 0.22$ and $\sigma_0 = 0.2$. When $\Delta t = 0.005$, a single computation takes 3.74 seconds for finite difference method, and 131.1 seconds for our splitting method using 5×10^5 simulations.

4.1.2 Asian option pricing in UVM with Hull-White interest rate: a three-dimensional case

We can also consider the uncertain volatility model with a stochastic interest rate, e.g. Hull-White interest rate (HW-IR). In HW-IR model, the interest rate has dynamic

$$dr_t \;=\; b(\theta_t - r_t)dt \;+\; \sigma^r dB_t,$$

where θ_t is determined by the current interst rate curve, b is the drawback force coefficient and $B = (B_t)_{t\geq 0}$ is another Brownian motion with correlation ρ to Brownian motion W which generates the dynamics of risky asset S. Then the value function

$$v(t,s,a,r) \;:=\; \mathbb{E}\left[e^{-\int_t^T r_s ds} g(S_T, A_T) \;\middle|\; S_t = s,\; A_t = a,\; r_t = r \right]$$

solves the pricing equation

$$\Big(\partial_t v \;+\; rsD_s v \;+\; b(\theta_t - r)D_r v \;+\; \frac{1}{2}(\sigma^r)^2 D^2_{rr} v \;-\; rv \\ +\; \max_{\underline{\sigma}\leq\sigma\leq\overline{\sigma}} \big(\rho\sigma\sigma^r s D^2_{rs} v \;+\; \frac{1}{2}\sigma^2 s^2 D^2_{ss} v\big) \;+\; sD_a v \Big)(t,s,a,r) \;=\; 0,$$

with terminal condition $v(T, s, a, r) = g(s, a)$.

Let $S_0 = 100$, $K_1 = 90$, $K_2 = 110$, $T = 1$, $\underline{\sigma} = 0.15$, $\overline{\sigma} = 0.25$, $r_0 = 0.02$, $b = 0.01$, $\sigma^r = 0.01$, $\rho = 0.2$ and interest rate curve is $f_t = 0.02, \quad \forall t > 0$. As in (4.2), we rewrite the pricing equation in form of (2.2) with constant σ_0. For $g(S_T, A_T) = (A_T - K_1)^+ - (A_T - K_2)^+$, we implement our splitting method with different constants σ_0, and take the mean value of 50 independent computations. The results are given in figure 2. We notice that the solution seems to be close to 11.51, and when the time discretization Δt is large, the numerical solution underestimates the value. Another phenomena is that when σ_0 is larger (e.g. $\sigma_0 = 0.25$), the performance of the numerical solutions seems more stable.

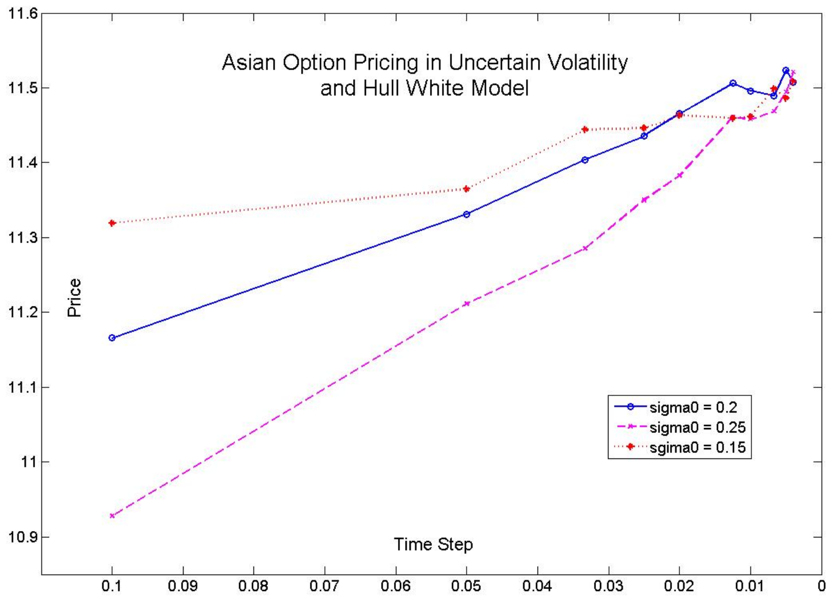

Figure 2: The price of Asian option with payoff $(A - K_1)^+ - (A - K_2)^+$ in UVM with HW IR and in BS model with HW IR. In case that $\Delta t = 0.005$, it takes 309.4 seconds for the splitting method using 5×10^5 simulations.

4.2 Optimal management of hydropower plant: A four-dimensional case

Let us consider an optimal management problem for a hydropower plant, which generalizes a little the model in Chapiter 2 of the thesis of Arnaud Porchet [16].

A hydropower plant manages a dam, which is filled by rain precipitations with non-negative rate A_t, which follows equation

$$dA_t \;\; = \;\; \mu_a A_t dt \; + \; \sigma_a A_t dW_t^1.$$

Denote by B_t the volume of water in the dam, then

$$dB_t \;\; = \;\; (A_t - q_t)\, dt,$$

where q_t represents the water flow sent at time t to generate electricity. It makes a profit $\int_0^T q_t S_t dt$ in period $[0, T]$, where S_t represents the market electricity price, which follows dynamics

$$dS_t \;=\; \mu_s S_t dt \;+\; \sigma_s S_t dW_t^2.$$

At the same time, the power station invests in electricity market with money θ_t, then the total revenue of the power station X_t follows equation

$$dX_t \;=\; \frac{\theta_t}{S_t} dS_t \;+\; q_t S_t dt \;=\; \theta_t \mu_s dt \;+\; \theta_t \sigma_s dW_t^2 \;+\; q_t S_t dt.$$

The power station optimizes its expected utility $\mathbb{E} U(X_T)$ on the strategy $(q_t)_{0\le t\le T}$ and $(\theta_t)_{0\le t\le T}$. Formally, we get a Bellman equation

$$\begin{aligned} u_t \;&+\; \mu_s s D_s u \;+\; \frac{1}{2}\sigma_s^2 s^2 D_{ss}^2 u \;+\; \mu_a a D_a u \;+\; \frac{1}{2}\sigma_a^2 a^2 D_{aa}^2 u \;+\; \rho \sigma_s \sigma_a s a D_{sa}^2 u \\ &+\; \max_{\theta} \Big[\; \theta_s \mu D_x u \;+\; \frac{1}{2}\theta^2 \sigma_s^2 D_{xx}^2 u \;+\; \theta \rho a \sigma_a \sigma_s D_{ax}^2 u \;+\; \theta \sigma_s^2 s D_{sx}^2 u \;\Big] \\ &+\; \max_q \big[\; (a-q) D_b u \;+\; q s D_x u \;\big] \;=\; 0. \end{aligned}$$

As in the examples in Section 5.2 of [13], we truncate the optimization on θ and rewrite the equation in form of (2.2).

$$\begin{aligned} u_t \;&+\; \mu_s s D_s u \;+\; \frac{1}{2}\sigma_s^2 s^2 D_{ss}^2 u \;+\; \mu_a a D_a u \;+\; \frac{1}{2}\sigma_a^2 a^2 D_{aa}^2 u \;+\; \rho \sigma_s \sigma_a s a D_{sa}^2 u \;+\; \frac{1}{2}\sigma_x^2 D_{xx}^2 u \\ &+\; \max_{-n\le\theta\le n} \Big[\; \theta_s \mu D_x u \;+\; \frac{1}{2}\theta^2 \sigma_s^2 D_{xx}^2 u \;+\; \theta \rho a \sigma_a \sigma_s D_{ax}^2 u \;+\; \theta \sigma_s^2 s D_{sx}^2 u \;-\; \frac{1}{2}\sigma_x^2 D_{xx}^2 u \;\Big] \\ &+\; \max_q \big[\; (a-q) D_b u \;+\; q s D_x u \;\big] \;=\; 0. \end{aligned}$$

Let $\mu_a = 0$, $\sigma_a = 0.2$, $\mu_s = 0$, $\sigma_s = 0.2$, $\rho = 0$, $n = 5$ and the utility function is given by $U(x) := -e^{-\rho x}$ with $\rho = 0.2$. Using the different choices of σ_x, we report the numerical result in Figure 3. We notice that the solution seems converge to the value -0.66, and when σ_x is chosen larger (e.g. $\sigma_x = 1.2$), the numerical solution is more stable w.r.t. the time step Δt.

5 Appendix

We give here the proof for Lemma 3.8. Let $(\tilde\lambda_k^i)_{1\le k\le K}$ be the projection coefficients of $R_i(\varphi)$ on basis $(e_k(\hat X_{t_n}, Y))_{1\le k\le K}$ as defined in (3.3), and $\hat\lambda_k^{i,M}$ be simulated regression estimations of $\tilde\lambda_k^i$ with M simulations of X, Y as defined in (3.6). Then for $(x,y) \in B_k$,

$$\tilde{\mathbb{E}}\;\big[\; R_i(\varphi) \mid \hat X_{t_n} = x, Y = y \;\big] \;=\; \tilde\lambda_k^i$$

and

$$\hat{\mathbb{E}}^M\;\big[\; R_i(\varphi) \mid \hat X_{t_n} = x, Y = y \;\big] \;=\; \underline{\Gamma}_i(x,y) \;\vee\; \hat\lambda_k^{i,M} \;\wedge\; \overline{\Gamma}_i(x,y).$$

Moreover,

$$\tilde\lambda_k^i \;=\; \frac{\mathbb{E}[R_i(\varphi) e_k(\hat X_{t_n}, Y)]}{\mathbb{E}[e_k^2(\hat X_{t_n}, Y)]} \quad \text{and} \quad \hat\lambda_k^{i,M} \;=\; \frac{\mathbb{E}^M[R_i(\varphi) e_k(\hat X_{t_n}, Y)]}{\mathbb{E}^M[e_k^2(\hat X_{t_n}, Y)]},$$

where $\mathbb{E}^M$ is the empirical expectation defined as follows: given M simulations $(U^m)_{1\le m\le M}$ of random variable U, $\mathbb{E}^M[U] := \frac{1}{M}\sum_{m=1}^M U^m$.

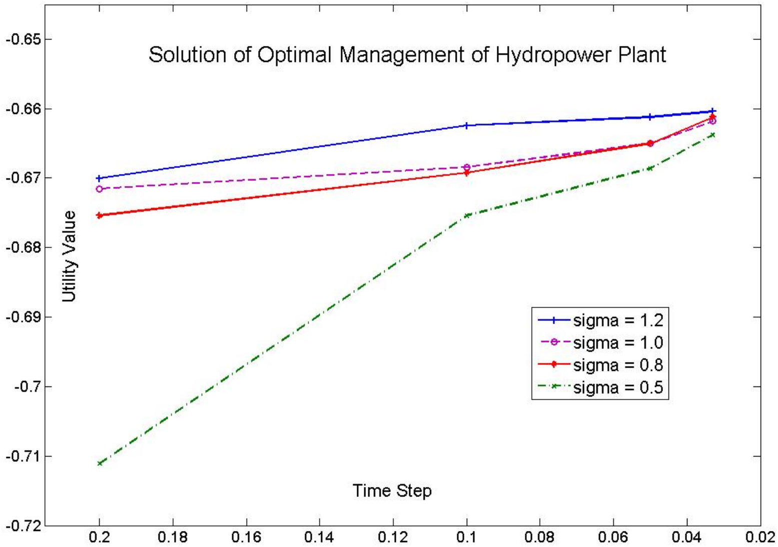

Figure 3: Solution of optimal management for a hydropower plant, with $\sigma_x = 1$ and $\sigma_x = 1.5$. Using 2×10^6 simulations, the splitting scheme takes 639.2 seconds for a single calculation when $\Delta t = 0.0333$.

Proof of Lemma 3.8. We omit the notations $i, k, x, y, \hat{X}_{t_n}, Y$ then simplify the notation as $\tilde{\lambda} = \tilde{\mathbb{E}}[R] = \frac{\mathbb{E}[Re]}{\mathbb{E}[ee]}$ and $\hat{\mathbb{E}}^M[R] = -\Gamma \vee \frac{\hat{\mathbb{E}}^M[Re]}{\hat{\mathbb{E}}^M[ee]} \wedge \Gamma$. Denote $\epsilon^M(Re) := \hat{\mathbb{E}}^M[Re] - \mathbb{E}[Re]$ and $\epsilon^M(ee) := \hat{\mathbb{E}}^M[ee] - \mathbb{E}[ee]$, then

$$\left|\hat{\mathbb{E}}^M[R] - \tilde{\mathbb{E}}[R]\right|^2 \;\le\; \left|\frac{\hat{\mathbb{E}}^M[Re]}{\hat{\mathbb{E}}^M[ee]} - \frac{\mathbb{E}[Re]}{\mathbb{E}[ee]}\right|^2 \wedge 4\Gamma^2 \;=\; \left|\frac{\epsilon^M(Re)}{\hat{\mathbb{E}}^M[ee]} + \tilde{\lambda}\frac{\epsilon^M(ee)}{\hat{\mathbb{E}}^M[ee]}\right|^2 \wedge 4\Gamma^2,$$

and it follows that

$$\begin{aligned}
\mathbb{E}\big[\hat{\mathbb{E}}^M[R] - \tilde{\mathbb{E}}[R]\big]^2 \;&\le\; \mathbb{E}\left|\frac{\epsilon^M(Re)}{\hat{\mathbb{E}}^M[ee]} + \tilde{\lambda}\frac{\epsilon^M(ee)}{\hat{\mathbb{E}}^M[ee]}\right|^2 \;\wedge\; 4\Gamma^2 \\
\le\;& 8\frac{\mathbb{E}[(\epsilon^M(Re))^2]}{(\mathbb{E}[ee])^2} \;+\; 8\tilde{\lambda}^2\frac{\mathbb{E}[(\epsilon^M(ee))^2]}{(\mathbb{E}[ee])^2} \;+\; 4\Gamma^2\mathbb{P}\left(\left|\frac{\hat{\mathbb{E}}^M[ee] \;-\; \mathbb{E}[ee]}{\mathbb{E}[ee]}\right| > \frac{1}{2}\right) \\
\le\;& 8\frac{\mathbb{E}[(\epsilon^M(Re))^2]}{(\mathbb{E}[ee])^2} \;+\; 8\tilde{\lambda}^2\frac{\mathbb{E}[(\epsilon^M(ee))^2]}{(\mathbb{E}[ee])^2} \;+\; 16\Gamma^2\frac{\mathbb{E}[(\epsilon^M(ee))^2]}{(\mathbb{E}[ee])^2} \\
=\;& \frac{1}{M}\;\frac{8}{(\mathbb{E}[ee])^2}\;\Big[\mathrm{Var}(Re) \;+\; \tilde{\lambda}^2\mathrm{Var}(ee) \;+\; 2\Gamma^2\mathrm{Var}(ee)\Big]. \qquad (5.1)
\end{aligned}$$

When $e = 1_{B_k}$, we have $\mathbb{E}[e^2(\hat{X}_{t_n}, Y)] = \mathbb{E}[e(\hat{X}_{t_n}, Y)] = \mathbb{P}((\hat{X}_{t_n}, Y) \in B_k)$, $\mathbb{E}[eR_i] \le C\,|\varphi|_0\,h^{i/2}\,\mathbb{E}[e]$ and $\tilde{\lambda} \le C\,|\varphi|_0\,h^{i/2}$, and then it follows by (5.1) that (3.13) holds true. □

References

[1] Bardou, Olivier; Bouthemy, Sandrine; Pagès, Gilles. When are swing options bang-bang? *Int. J. Theor. Appl. Finance* **13** (2010), no. 6, 867–899. MR-2718983

[2] Barles, Guy; Jakobsen, Espen R. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Math. Comp.*, **76** (2007), no. 260, 1861–1893 (electronic). MR-2336272

[3] Barles, G.; Souganidis, P. E. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.* **4** (1991), no. 3, 271–283. MR-1115933

[4] Bonnans, J. Frédéric; Ottenwaelter, élisabeth; Zidani, Housnaa. A fast algorithm for the two dimensional HJB equation of stochastic control. *M2AN Math. Model. Numer. Anal.* **38** (2004), no. 4, 723–735. MR-2087732

[5] Bouchard, Bruno; Touzi, Nizar. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.* **111** (2004), no. 2, 175–206. MR-2056536

[6] B. Bouchard and X. Warin, Monte-Carlo valorisation of American options: facts and new algorithms to improve existing methods, preprint, 2011.

[7] J.F. Bonnans, Z. Cen and T. Christel, Energy contracts management by stochastic programming techniques, to appear in special issue of "Annals of Operations Research" on "stochastic programming".

[8] Z. Chen and P.A. Forsyth, A Semi-Lagrangian Approach for Natural Gas Storage Valuation and Optimal Operation *SIAM J. Scientific Computing* **30** (2007), 339–368.

[9] Cheridito, Patrick; Soner, H. Mete; Touzi, Nizar; Victoir, Nicolas. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.* **60** (2007), no. 7, 1081–1110. MR-2319056

[10] D. Crisan, and K. Manolarakis, Solving Backward Stochastic Differential Equations using the Cubature Method, preprint, 2011.

[11] Dai, Min; Kwok, Yue Kuen; Zong, Jianping. Guaranteed minimum withdrawal benefit in variable annuities. *Math. Finance* **18** (2008), no. 4, 595–611. MR-2454673

[12] K. Debrabant and E.R. Jakobsen, Semi-Lagrangian schemes for linear and fully nonlinear diffusion equations, preprint, 2009.

[13] Fahim, Arash; Touzi, Nizar; Warin, Xavier. A probabilistic numerical method for fully nonlinear parabolic PDEs. *Ann. Appl. Probab.* **21** (2011), no. 4, 1322–1364. MR-2857450

[14] Gobet, Emmanuel; Lemor, Jean-Philippe; Warin, Xavier. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.* **15** (2005), no. 3, 2172–2202. MR-2152657

[15] F.A. Longstaff and R.S. Schwartz, Valuing American options by simulation: a simple least-square approach, *Review of Finnaical studies*, **14** (2001), 113–147.

[16] A. Porchet, Problems of Valuation and Organization in Energy Markets, PhD Dissertation, Universitlę́ Paris Dauphine, 2008.

[17] Rogers, L. C. G.; Shi, Z. The value of an Asian option. *J. Appl. Probab.* **32** (1995), no. 4, 1077–1088. MR-1363350

[18] H.M. Soner, N. Touzi and J. Zhang, Wellposedness of second order backward SDEs, *Probability Theory and Related Fields*, to appear.

Acknowledgments. The author is grateful to J. Frédéric Bonnans, Damien Lamberton and Nizar Touzi for fruitful discussions and two anonymous referees for helpful comments and suggestions. This research is supported by the Chair *Financial Risks* of the *Risk Foundation* sponsored by Société Générale, the Chair *Derivatives of the Future* sponsored by the Fédération Bancaire Française, and the Chair *Finance and Sustainable Development* sponsored by EDF and CA-CIB.