AN APPROXIMATION SCHEME FOR SEMILINEAR PARABOLIC PDES WITH CONVEX AND COERCIVE HAMILTONIANS *

SHUO HUANG†, GECHUN LIANG‡, AND THALEIA ZARIPHOPOULOU§

Abstract. We propose an approximation scheme for a class of semilinear parabolic equations that are convex and coercive in their gradients. Such equations arise often in pricing and portfolio management in incomplete markets and, more broadly, are directly connected to the representation of solutions to backward stochastic differential equations. The proposed scheme is based on splitting the equation in two parts, the first corresponding to a linear parabolic equation and the second to a Hamilton-Jacobi equation. The solutions of these two equations are approximated using, respectively, the Feynman-Kac and the Hopf-Lax formulae. We establish the convergence of the scheme and determine the convergence rate, combining Krylov's shaking coefficients technique and Barles-Jakobsen's optimal switching approximation.

Key words. Splitting, Feynman-Kac formula, Hopf-Lax formula, viscosity solutions, shaking coefficients technique, optimal switching approximation.

AMS subject classifications. 35K65, 65M12, 93E20

1. Introduction.

We consider semilinear parabolic equations of the form

$$
\begin{cases} -\partial_t u - \dfrac{1}{2}\text{Trace}\left(\sigma\sigma^T(t,x)\partial_{xx}u\right) - b(t,x)\cdot\partial_x u + H(t,x,\partial_x u) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^n, \end{cases} \tag{1.1}
$$

where $Q_T = [0,T)\times\mathbb{R}^n$. A key feature is that the Hamiltonian $H(t,x,p)$ is convex and coercive in p. In particular, the coercivity covers the case that H has quadratic growth in p, a case that corresponds to a rich class of equations in mathematical finance arising in optimal investment with homothetic risk preferences ([20]), exponential indifference valuation ([18, 19]), entropic risk measures ([11]) and others.

More broadly, these equations are inherently connected to (quadratic) backward stochastic differential equations (BSDE), a central area of stochastic analysis ([12] [13] and [23]). Specifically, the Hamiltonian $H(t,x,p)$ is directly related to the BSDE's driver and, moreover, the solution of (1.1) yields a functional-form representation of the processes solving the BSDE.

General existence and uniqueness results can be found, among others in [23] as well as in [20], where BSDE techniques have been mainly applied. Closed-form solutions can be constructed only in one-dimensional cases ([38]). Furthermore, approximation schemes have been developed; see [8] and [10] for more references.

Herein, we contribute to further studying problem (1.1) by proposing a new approximation scheme. The key idea is to use in an essential way the *convexity* of

*The authors thank the Editor, the Associate Editor, and the two referees for their valuable comments and suggestions, and T. Lyons and C. Reisinger for their helpful discussions. This work was presented at seminars in Oxford University, Fudan University, Shanghai Jiao Tong University and Jilin University. The authors thank the participants for helpful comments and suggestions.

†Department of Statistics, The University of Warwick, Coventry CV4 7AL, U.K. s.huang.13@warwick.ac.uk

‡Department of Statistics, The University of Warwick, Coventry CV4 7AL, U.K. Partially supported by Royal Society International Exchanges (Grant No. 170137), NSFC (Grant No. 11771158) and a Senior Fellowship at Freiburg Institute of Advanced Studies (FRIAS), University of Freiburg. g.liang@warwick.ac.uk

§Departments of Mathematics and IROM, The University of Texas at Austin, U.S.A. and the Oxford-Man Institute, University of Oxford, U.K. zariphop@math.utexas.edu

the Hamiltonian with respect to the gradient. This property is natural in all above applications but it has not been adequately exploited in the existing approximation studies.

To highlight the main ideas and build intuition, we start with some preliminary informal arguments, considering for simplicity slightly simpler equations. To this end, consider the Hamilton-Jacobi (HJ) equation

$$\begin{cases} -\partial_t u + H(\partial_x u) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^n, \end{cases} \tag{1.2}$$

where the Hamiltonian H is convex and coercive, and the terminal datum U is bounded and Lipschitz continuous. Let L be the Legendre (convex dual) transform of H, $L(q) = \sup_{p\in\mathbb{R}^n}\{p\cdot q - H(p)\}$. The Fenchel-Moreau theorem then yields that $H(p) = \sup_{q\in\mathbb{R}^n}\{p\cdot q - L(q)\}$ and, thus, the HJ equation in (1.2) can be alternatively written as

$$-\partial_t u + \sup_{q\in\mathbb{R}^n}\{\partial_x u\cdot q - L(\partial_x u)\} = 0.$$

Classical arguments from control theory then imply the deterministic optimal control representation

$$u(t,x) = \inf_{q\in\mathbb{L}^2[t,T]}\left[\int_t^T L(q_s)ds + U(X_T^{t,x;q})\right],$$

with the controlled state equation $X_s^{t,x;q} = x - \int_t^s q_u du$, for $s\in[t,T]$.

Hopf and Lax observed that, instead of considering the controls in $\mathbb{L}^2[t,T]$, it suffices to optimize over the controls generating geodesic paths of $X^{t,x;q}$, i.e. the controls $\hat{q}$ such that $X_T^{t,x;\hat{q}} = y$, for any $y\in\mathbb{R}^n$. Such controls are given by $\hat{q}_s = \frac{x-y}{T-t}$, for $s\in[t,T]$. The above "infinite dimensional" optimal control problem is thus reduced to the "finite dimensional" minimization problem

$$u(t,x) = \inf_{y\in\mathbb{R}^n}\left\{(T-t)L(\frac{x-y}{T-t}) + U(y)\right\}. \quad \text{(Hopf-Lax formula)} \tag{1.3}$$

There exist several well established algorithms to study this type of minimization problems (see, for example, [32] for the Nelder-Mead simplex algorithm). We also refer to section 3.3.2.b in [14] for the introduction of the Hopf-Lax formula from a classical calculus of variations perspective.

Adding a diffusion term to equation (1.2) yields the semilinear parabolic equation

$$\begin{cases} -\partial_t u - \frac{1}{2}\mathrm{Trace}\left(\sigma\sigma^T(t,x)\partial_{xx}u\right) + H(\partial_x u) = 0 & \text{in } Q_T; \\ u(T,x) = U(x) & \text{in } \mathbb{R}^n. \end{cases} \tag{1.4}$$

In analogy to the deterministic case, classical arguments from control theory imply the stochastic optimal control representation

$$u(t,x) = \inf_{q\in\mathbb{H}^2[t,T]}\mathbf{E}\left[\int_t^T L(q_s)ds + U(X_T^{t,x;q})|\mathcal{F}_t\right],$$

with the controlled state equation $X_s^{t,x;q} = x - \int_t^s q_u du + \int_t^s \sigma(u, X_u^{t,x;q})dW_u$, for $s\in[t,T]$, and $\mathbb{H}^2[t,T]$ being the space of square-integrable progressively measurable processes q.

Naturally, due to the stochasticity of the state $X^{t,x;q}$, the Hopf-Lax formula (1.3) does not hold for the solution of problem (1.4). On the other hand, we observe that if we still choose, as in the deterministic case, controls of the form $\hat{q}_s = \frac{x-y}{T-t}$, for $y \in \mathbb{R}^n$ and $s \in [t,T]$, then

$$X_s^{t,x;\hat{q}} = \frac{T-s}{T-t}x + \frac{s-t}{T-t}y + \int_t^s \sigma(u, X_u^{t,x;\hat{q}})dW_u,$$

for $s \in [t,T]$. Therefore, for $T-t = o(1)$, we have $X_T^{t,x;\hat{q}} \approx Y_T^{t,y}$, where $Y^{t,y}$ solves the *uncontrolled* stochastic differential equation

$$Y_s^{t,y} = y + \int_t^s \sigma(u, Y_u^{t,y})dW_u,$$

for $s \in [t,T]$. In turn, since y is arbitrary, we readily obtain an *upper* bound of the solution $u(t,x)$ of (1.4), namely,

$$u(t,x) \leq \inf_{y\in\mathbb{R}^n} \left\{ (T-t)L(\frac{x-y}{T-t}) + \mathbf{E}[U(Y_T^{t,y})|\mathcal{F}_t] \right\}. \tag{1.5}$$

Furthermore, the convexity of H yields that L is also convex and, therefore, for any control process $q \in \mathbb{H}^2[t,T]$, we deduce that

$$\begin{aligned}
&\mathbf{E}\left[\int_t^T L(q_s)ds + U(X_T^{t,x;q})|\mathcal{F}_t\right] \\
&\geq (T-t)L\left(\mathbf{E}\left[\frac{1}{T-t}\int_t^T q_u du|\mathcal{F}_t\right]\right) + \mathbf{E}[U(X_T^{t,x;q})|\mathcal{F}_t] \\
&= (T-t)L\left(\mathbf{E}\left[\frac{x - X_T^{t,x;q} + \int_t^T \sigma(u, X_u^{t,x;q})dW_u}{T-t}|\mathcal{F}_t\right]\right) + \mathbf{E}[U(X_T^{t,x;q})|\mathcal{F}_t] \\
&= (T-t)L\left(\frac{x - \mathbf{E}[X_T^{t,x;q}|\mathcal{F}_t]}{T-t}\right) + \mathbf{E}[U(X_T^{t,x;q})|\mathcal{F}_t].
\end{aligned}$$

Therefore, for $T-t = o(1)$, we have $X_T^{t,x;q} \approx Y_T^{t,\hat{y}}$, with $\hat{y} := \mathbf{E}[X_T^{t,x;q}|\mathcal{F}_t]$. Thus, we also obtain a *lower* bound of the solution $u(t,x)$ of (1.4), namely,

$$u(t,x) \geq \inf_{\hat{y}\in\mathbb{R}^n} \left\{ (T-t)L(\frac{x-\hat{y}}{T-t}) + \mathbf{E}[U(Y_T^{t,\hat{y}})|\mathcal{F}_t] \right\}. \tag{1.6}$$

Note that when σ degenerates to 0, inequalities (1.5) and (1.6) give us an equality, which is precisely the Hopf-Lax formula (1.3).

We now see how the above ideas can be combined to develop an approximation scheme for the original problem (1.1). Equation (1.1) can be “split” into a first-order nonlinear equation of Hamilton-Jacobi type and a linear parabolic equation. The solution of the former is represented via the Hopf-Lax formula and corresponds to the value function of a deterministic control problem. The solution of the latter corresponds to a conditional expectation of an uncontrolled diffusion and is given by the Feynman-Kac formula. The scheme is then naturally based on a backwards in time recursive combination of the Hopf-Lax and the Feynman-Kac formula; see (2.2) and (2.12) for further details.

We establish the convergence of the scheme to the unique (viscosity) solution of (1.1) and determine the rate of convergence. We do this by deriving upper and lower bounds on the approximation error (Theorems 3.3 and 3.6, respectively). The main tools come from the *shaking coefficients technique* introduced by Krylov [24] [25] and the *optimal switching approximation* introduced by Barles and Jakobsen [1] [2].

While various arguments follow from adaptations of these techniques, the main difficulty is to derive a consistency error estimate. This is one of the key steps herein and it is precisely where the convexity of the Hamiltonian with respect to the gradient is used in an essential way. Specifically, we obtain this estimate by applying convex duality and using the properties of the optimizers in the related minimization problems (Proposition 2.5 (vi)). Using this estimate and the comparison result for the approximation scheme (Proposition 2.9), we in turn derive an upper bound for the approximation error by perturbing the coefficients of the equation. The lower bound for the approximation error is obtained by another layer of approximation of the equation by using an auxiliary optimal switching system.

Approximation schemes for viscosity solutions were first studied by Barles and Souganidis [4], who showed that any monotone, stable and consistent approximation scheme converges to the correct solution, provided that there exists a comparison principle for the limiting equation. The corresponding convergence rate had been an open problem for a long time until late 1990s when Krylov introduced the shaking coefficients technique to construct a sequence of smooth subsolutions/supersolutions. This technique was further developed by Barles and Jakobsen in a sequence of papers (see [3] and [22] and more references therein), and has recently been applied to solve various problems (see, among others, [5] [7] [16] and [19]).

Krylov's technique depends crucially on the convexity/concavity of the underlying equation with respect to its terms. As a result, unless the approximate solution has enough regularity (so one can interchange the roles of the approximation scheme and the original equation), the shaking coefficients technique only gives either an upper or a lower bound for the approximation error, but not both. A further breakthrough was made by Barles and Jakobsen in [1] and [2], who combined the ideas of optimal switching approximation of Hamilton-Jacobi-Bellman (HJB) equations (initially proposed by Evans and Friedman [15]) with the shaking coefficients technique. They obtained both upper and lower bounds of the error estimate, but with a lower convergence rate due to the introduction of another approximation layer.

The splitting approach (fractional step, prediction and correction, etc.) is dated back to Marchuk [30] in the late 1960s. Its application to nonlinear PDEs was firstly proposed by Lions and Mercier [27] and has been subsequently used by many others. For semilinear parabolic equations related to problems in mathematical finance, splitting methods have been applied by Tourin [36] (see also more references therein). More recently, Nadtochiy and Zariphopoulou [31] proposed a splitting algorithm to the marginal HJB equation arising in optimal investment problems in a stochastic factor model and general utility functions. Henderson and Liang [19] proposed a splitting approach for utility indifference pricing in a multi-dimensional non-traded assets model with intertemporal default risk, and established its convergence rate. Tan [35] proposed a splitting method for a class of fully nonlinear degenerate parabolic PDEs and applied it to Asian options and commodity trading.

Finally, we mention that most of the existing algorithms (see, among others, Howard's finite difference scheme [6]) provide approximations only at certain time grids. In contrast, the splitting approximation can be used to approximate the solution at any time point. Furthermore, since the existing algorithms are often based on finite difference approximation, the "curse of dimensionality" issue arises. We remark that the splitting approximation itself does not involve finite difference formulation, as long as one can find an efficient way to compute conditional expectations, e.g. the multi-level Monte Carlo approach [17], the least squares Monte Carlo approach [28], the cubature approach [29], and etc. This advantage is also shared by existing BSDE time discretization algorithms (see, for example, [8] and [10]). However, the commonly used BSDE time discretization algorithms for (1.1) require that the Hamiltonian has the form $H(t,x,\sigma^{tr}(t,x)\partial_x u)$ (see [10]), which is not the case herein. Indeed, we do not require the last variable in the Hamiltonian H to depend on the diffusion coefficient σ.

The paper is organized as follows. In section 2 we introduce the approximation scheme. In section 3, we prove its convergence rate using the shaking coefficients technique and optimal switching approximation. We provide a numerical test in section 4 and conclude in section 5. Some technical proofs are provided in the appendix.

2. The approximation scheme using the Hopf-Lax formula and splitting. For $T>0$, let $Q_T=[0,T)\times\mathbb{R}^n$. Let also d be a positive integer and $\delta>0$. For a function $f:Q_T\to\mathbb{R}^d$, we introduce its (semi)norms

$$|f|_0:=\sup_{(t,x)\in Q_T}|f(t,x)|,$$

$$[f]_{1,\delta}:=\sup_{\substack{(t,x),(t',x)\in Q_T\\ t\neq t'}}\frac{|f(t,x)-f(t',x)|}{|t-t'|^{\delta}},\quad [f]_{2,\delta}:=\sup_{\substack{(t,x),(t,x')\in Q_T\\ x\neq x'}}\frac{|f(t,x)-f(t,x')|}{|x-x'|^{\delta}}.$$

Furthermore, $[f]_\delta:=[f]_{1,\delta/2}+[f]_{2,\delta}$ and $|f|_\delta:=|f|_0+[f]_\delta$. Similarly, the (semi)norms of a function $g:\mathbb{R}^n\to\mathbb{R}^d$ are defined as

$$|g|_0:=\sup_{x\in\mathbb{R}^n}|g(x)|,\quad [g]_\delta:=\sup_{\substack{x,x'\in\mathbb{R}^n\\ x\neq x'}}\frac{|g(x)-g(x')|}{|x-x'|^{\delta}},\quad |g|_\delta:=|g|_0+[g]_\delta.$$

For $S=Q_T$, $\mathbb{R}^n$ or $Q_T\times\mathbb{R}^n$, we denote by $\mathcal{C}(S)$ the space of continuous functions on S, and by $\mathcal{C}_b^\delta(S)$ the space of bounded and continuous functions on S with finite norm $|f|_\delta$. We also set $\mathcal{C}_b^0(S)\equiv\mathcal{C}_b(S)$ and denote by $\mathcal{C}_b^\infty(S)$ the space of smooth functions on S with bounded derivatives of any order.

We throughout assume the following conditions for equation (1.1).

Assumption 2.1.

(i) The diffusion coefficient $\sigma\in\mathcal{C}_b^1(Q_T)$, the drift coefficient $b\in\mathcal{C}_b^1(Q_T)$, and the terminal datum $U\in\mathcal{C}_b^1(\mathbb{R}^n)$ have norms $|\sigma|_1,|b|_1,|U|_1\le M$, for some $M>0$.

(ii) The Hamiltonian $H(t,x,p)\in\mathcal{C}(Q_T\times\mathbb{R}^n)$ is convex in p, and satisfies the coercivity condition

$$\lim_{|p|\to\infty}\frac{H(t,x,p)}{|p|}=\infty,$$

uniformly in $(t,x)\in Q_T$. Moreover, for every p, $[H(\cdot,\cdot,p)]_1\le M$, and there exist two locally bounded functions H^* and $H_*:\mathbb{R}^n\to\mathbb{R}$ such that

$$H_*(p)=\inf_{(t,x)\in Q_T}H(t,x,p)\quad\textit{and}\quad H^*(p)=\sup_{(t,x)\in Q_T}H(t,x,p).$$

Under the above assumptions, we have the following existence, uniqueness and regularity results for equation (1.1). Their proofs are provided in Appendix A.

PROPOSITION 2.2. *Suppose that Assumption 2.1 is satisfied. Then, there exists a unique viscosity solution* $u \in \mathcal{C}_b^1(\bar{Q}_T)$ *of equation (1.1), with* $|u|_1 \leq C$*, for some constant* C *depending only on* M *and* T*.*

2.1. The backward operator $\mathbf{S}_t(\Delta)$. Using that $H(t,x,p)$ is convex in p, we define its Legendre (convex dual) transform $L : Q_T \times \mathbb{R}^n \to \mathbb{R}$, given by

$$L(t,x,q) := \sup_{p \in \mathbb{R}^n} \{p \cdot q - H(t,x,p)\}. \tag{2.1}$$

For any t and Δ with $0 \leq t < t + \Delta \leq T$, and any $\phi \in \mathcal{C}_b(\mathbb{R}^n)$, we introduce the *backward operator* $\mathbf{S}_t(\Delta) : \mathcal{C}_b(\mathbb{R}^n) \to \mathcal{C}_b(\mathbb{R}^n)$,

$$\begin{cases} \mathbf{S}_t(\Delta)\phi(x) = \min\limits_{y \in \mathbb{R}^n} \left\{ \Delta L\left(t, x, \dfrac{x-y}{\Delta}\right) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y})|\mathcal{F}_t] \right\}, & x \in \mathbb{R}^n, \\ Y_s^{t,y} = y + b(t,y)(s-t) + \sigma(t,y)(W_s - W_t), & s \in [t, t+\Delta], \end{cases} \tag{2.2}$$

on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbf{P})$, where W is an n-dimensional Brownian motion with its augmented filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

We start with some auxiliary properties of H and L.

PROPOSITION 2.3. *Suppose that Assumption 2.1 (ii) is satisfied. Then, the following assertions hold:*

(i) H *is the Legendre transform of* L*, i.e.* $H(t,x,p) = \sup_{q \in \mathbb{R}^n}\{p \cdot q - L(t,x,q)\}$*, for* $(t,x) \in Q_T$*.*

(ii) The functions

$$L_*(q) := \sup_{p \in \mathbb{R}^n} \{p \cdot q - H^*(p)\} \quad \textit{and} \quad L^*(q) := \sup_{p \in \mathbb{R}^n} \{p \cdot q - H_*(p)\}$$

are locally bounded and satisfy, for $(t,x) \in Q_T$*,* $L_*(q) \leq L(t,x,q) \leq L^*(q)$*.*

(iii) For $(t,x) \in Q_T$*,* $L(t,x,q)$ *is convex in* q *with* $[L(\cdot,\cdot,q)]_1 \leq 2M$*. Furthermore, it satisfies the coercivity condition*

$$\lim_{|q| \to \infty} \frac{L(t,x,q)}{|q|} = \infty,$$

uniformly in $(t,x) \in Q_T$*.*

(iv) For each $(t,x) \in Q_T$ *and* $p, q \in \mathbb{R}^n$*, there exist* $p^*, q^* \in \mathbb{R}^n$ *such that*

$$L(t,x,q) = q \cdot p^* - H(t,x,p^*) \quad \textit{and} \quad H(t,x,p) = p \cdot q^* - L(t,x,q^*).$$

Furthermore, $|p^*| \leq \xi(|q|)$ *and* $|q^*| \leq \xi(|p|)$*, for some real-valued increasing function* $\xi(\cdot)$ *independent of* (t,x)*.*

Proof. Parts (i) and (ii) are immediate and, thus, we only prove (iii) and (iv).

(iii) For fixed $(t,x) \in Q_T$, $q_1, q_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$, we have

$$\begin{aligned} L(t,x,\lambda q_1 + (1-\lambda)q_2) &= \sup_{p \in \mathbb{R}^n} \{(\lambda q_1 + (1-\lambda)q_2) \cdot p - H(t,x,p)\} \\ &\leq \lambda \sup_{p \in \mathbb{R}^n} \{q_1 \cdot p - H(t,x,p)\} + (1-\lambda) \sup_{p \in \mathbb{R}^n} \{q_2 \cdot p - H(t,x,p)\} \\ &= \lambda L(t,x,q_1) + (1-\lambda) L(t,x,q_2). \end{aligned}$$

From the definition of L, we further have, for any $q \in \mathbb{R}^n$,

$$\begin{aligned}[L(\cdot,\cdot,q)]_1 &= [L(\cdot,\cdot,q)]_{1,1/2} + [L(\cdot,\cdot,q)]_{2,1} \\ &\le \sup_p\{[H(\cdot,\cdot,p)]_{1,1/2}\} + \sup_p\{[H(\cdot,\cdot,p)]_{2,1}\} \le 2M.\end{aligned}$$

Next, for any $K > 0$, we deduce, by setting $p = K\frac{q}{|q|}$, that

$$L(t,x,q) \ge q \cdot K\frac{q}{|q|} - H(t,x,K\frac{q}{|q|}) \ge K|q| - \sup_{r\in B(0,K)} H^*(r).$$

Dividing both sides by $|q|$ and sending $|q| \to \infty$, the coercivity condition for L follows.

(iv) From (i) and (ii), we deduce that L and H are symmetric to each other and, thus, we only establish the assertions for L. To this end, for each $(t,x) \in Q_T$, we obtain, by setting $p = 0$ in (2.1), that $L(t,x,q) \ge -H(t,x,0)$. Therefore, it suffices to find a real-valued increasing function, say $\xi(\cdot)$, such that, if $|p| > \xi(|q|)$, then

$$p \cdot q - H(t,x,p) < -H(t,x,0).$$

Indeed, it follows from Assumption 2.1 (ii) that there exists a real-valued increasing function, say $K_H(y)$, such that, for any $(t,x) \in Q_T$ and $|p| \ge K_H(y)$, we have $\frac{H(t,x,p)}{|p|} \ge y$. Setting $\xi(x) := \max\{K_H(|H^*(0)| + x), 1\}$, we deduce that, for $|p| > \xi(|q|)$,

$$p \cdot q - H(t,x,p) \le |p|(|q| - \frac{H(t,x,p)}{|p|}) < |q| - (|H^*(0)| + |q|) \le -H(t,x,0),$$

and we easily conclude. $\square$

Next, we show that the minimum in (2.2) is actually achieved, i.e. for any $\phi \in \mathcal{C}_b(\mathbb{R}^n)$, there always exists an associated minimizer y^*.

Proposition 2.4. *Suppose that Assumption 2.1 is satisfied. Then, for each t and Δ with $0 \le t < t + \Delta \le T$, $x \in \mathbb{R}^n$ and $\phi \in \mathcal{C}_b(\mathbb{R}^n)$, there exists a minimizer $y^* \in \mathbb{R}^n$ such that*

$$\mathbf{S}_t(\Delta)\phi(x) = \Delta L\left(t, x, \frac{x - y^*}{\Delta}\right) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y^*})|\mathcal{F}_t].$$

Moreover, there exists a constant $C > 0$, depending only on M and T, such that

$$\left|\frac{x - y^*}{\Delta}\right| \le \xi(C[\phi]_1), \tag{2.3}$$

for some real-valued increasing function $\xi(\cdot)$ independent of (t,x).

Proof. Let $q = \frac{x-y}{\Delta}$. Then $|q| \to \infty$ as $|y| \to \infty$. In turn, from Proposition 2.3 (iii), we deduce that, as $|y| \to \infty$,

$$\Delta L\left(t, x, \frac{x-y}{\Delta}\right) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y})|\mathcal{F}_t] = |x - y|\frac{L(t,x,q)}{|q|} + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y})|\mathcal{F}_t] \to \infty.$$

Furthermore, using that the mapping $y \mapsto \Delta L(t,x,\frac{x-y}{\Delta}) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y})|\mathcal{F}_t]$ is continuous, we deduce that it must admit a minimizer $y^* \in \mathbb{R}^n$.

Next, we prove inequality (2.3). For $\phi \in \mathcal{C}_b^1(\mathbb{R}^n)$, following the same reasoning as in the proof of Proposition 2.3 (iv), it suffices to find a real-valued increasing function $\xi(\cdot)$ such that

$$\Delta L\,(t,x,q) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,x-\Delta q})|\mathcal{F}_t] > \Delta L\,(t,x,0) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,x})|\mathcal{F}_t], \tag{2.4}$$

if $|q| > \xi(C[\phi]_1)$, for some constant $C > 0$ depending only on M and T. To prove this, note that Assumption 2.1 (i) on the coefficients σ and b implies that

$$\begin{aligned}\mathbf{E}[\phi(Y_{t+\Delta}^{t,x})|\mathcal{F}_t] - \mathbf{E}[\phi(Y_{t+\Delta}^{t,x-\Delta q})|\mathcal{F}_t] \ &\leq [\phi]_1 \mathbf{E}\left[\left|Y_{t+\Delta}^{t,x} - Y_{t+\Delta}^{t,x-\Delta q}\right| |\mathcal{F}_t\right] \\ &\leq C[\phi]_1 \Delta |q|.\end{aligned} \tag{2.5}$$

On the other hand, from Proposition 2.3 (iv), there exists a real-valued increasing function, say $K_L(y)$, such that, for any $(t,x) \in Q_T$ and $|q| \geq K_L(y)$, we have $\frac{L(t,x,q)}{|q|} \geq y$. Setting $\xi(x) := \max\{K_L(|L^*(0)| + x), 1\}$, we deduce that, for $|q| > \xi(C[\phi]_1)$,

$$\frac{L(t,x,q)}{|q|} > |L^*(0)| + C[\phi]_1 \geq \frac{L^*(0)}{|q|} + C[\phi]_1 \geq \frac{L(t,x,0)}{|q|} + C[\phi]_1.$$

Using the above inequality, together with (2.5), we obtain (2.4). Finally, the case $[\phi]_1 = \infty$ follows trivially. □

Next, we derive some key properties of the backward operator $\mathbf{S}_t(\Delta)$.

Proposition 2.5. *Suppose that Assumption 2.1 is satisfied. Then, for each t and Δ with $0 \leq t < t + \Delta \leq T$, the operator $\mathbf{S}_t(\Delta)$ has the following properties:*

(i) (Constant preserving) For any $\phi \in \mathcal{C}_b(\mathbb{R}^n)$ and $c \in \mathbb{R}$,

$$\mathbf{S}_t(\Delta)(\phi + c) = \mathbf{S}_t(\Delta)\phi + c.$$

(ii) (Monotonicity) For any $\phi, \psi \in \mathcal{C}_b(\mathbb{R}^n)$ with $\phi \geq \psi$,

$$\mathbf{S}_t(\Delta)\phi \geq \mathbf{S}_t(\Delta)\psi.$$

(iii) (Concavity) For any $\phi \in \mathcal{C}_b(\mathbb{R}^n)$, $\mathbf{S}_t(\Delta)\phi$ is concave in ϕ.

(iv) (Stability) For any $\phi \in \mathcal{C}_b(\mathbb{R}^n)$,

$$|\mathbf{S}_t(\Delta)\phi|_0 \leq C\Delta + |\phi|_0,$$

where $C = \max\{|L^*(0)|, |H^*(0)|\}$, with L^* and H^* as in Proposition 2.3 (ii) and Assumption 2.1 (ii). Therefore, the operator $\mathbf{S}_t(\Delta)$ is indeed a mapping from $\mathcal{C}_b(\mathbb{R}^n)$ to $\mathcal{C}_b(\mathbb{R}^n)$.

(v) For any $\phi \in \mathcal{C}_b^1(\mathbb{R}^n)$, there exists a constant C depending only on $[\phi]_1$, M and T, such that

$$|\mathbf{S}_t(\Delta)\phi - \phi|_0 \leq C\sqrt{\Delta}.$$

(vi) For any $\phi \in \mathcal{C}_b^\infty(\mathbb{R}^n)$, define

$$\mathcal{E}(t,\Delta,\phi) := \left|\frac{\phi - \mathbf{S}_t(\Delta)\phi}{\Delta} - \mathbf{L}_t\phi\right|_0, \tag{2.6}$$

where the operator $\mathbf{L}_t$ is given by

$$\mathbf{L}_t\phi(x) = -\frac{1}{2}\,\mathit{Trace}\left(\sigma\sigma^T(t,x)\partial_{xx}\phi(x)\right) - b(t,x)\cdot\partial_x\phi(x) + H(t,x,\partial_x\phi(x)).$$

Then,

$$\mathcal{E}(t,\Delta,\phi) \leq \ C\Delta\left(|\partial_{xxxx}\phi|_0 + \mathcal{R}(\phi)\right),$$

where the constant C depends only on $[\phi]_1$, M and T, and $\mathcal{R}(\phi)$ represents the "insignificant" terms containing the derivatives of ϕ up to third order.

Proof. Parts (i)-(iii) are immediate. We only prove (iv)-(vi) and, in particular, for the case $n=1$, since the general case follows along similar albeit more complicated arguments.

(iv) Choosing $y=x$ in (2.2) gives

$$\mathbf{S}_t(\Delta)\phi(x) \leq \Delta L^*(0) + |\phi|_0. \tag{2.7}$$

It follows from the definition of L_* in Proposition 2.3 (ii) that $L_*(q) \geq -H^*(0) \geq -|H^*(0)|$, for $q\in\mathbb{R}^n$. In turn, Proposition 2.4 further yields

$$\begin{aligned}\mathbf{S}_t(\Delta)\phi(x) &= \Delta L(t,x,\frac{x-y^*}{\Delta}) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y^*})|\mathcal{F}_t]\\ &\geq \Delta L_*(\frac{x-y^*}{\Delta}) - |\phi|_0\\ &\geq \ -\Delta|H^*(0)| - |\phi|_0. \end{aligned}\tag{2.8}$$

The assertion then follows by combining (2.7) and (2.8).

(v) From Proposition 2.3 (ii) and Proposition 2.4, we deduce that

$$\begin{aligned}|\mathbf{S}_t(\Delta)\phi(x) - \phi(x)| = &\left|\Delta L(t,x,\frac{x-y^*}{\Delta}) + \mathbf{E}[\phi(Y_{t+\Delta}^{t,y^*}) - \phi(x)|\mathcal{F}_t]\right|\\ &\leq \Delta\max\left\{|L^*(\frac{x-y^*}{\Delta})|, |L_*(\frac{x-y^*}{\Delta})|\right\} + [\phi]_1\mathbf{E}\left[\left|Y_{t+\Delta}^{t,y^*} - x\right||\mathcal{F}_t\right]\\ &\leq C\Delta + (C\Delta + M\Delta + CM\sqrt{\Delta})[\phi]_1 \leq C\sqrt{\Delta},\end{aligned}$$

where the constant C depends only on $[\phi]_1$, M and T.

(vi) For $(t,x)\in[0,T-\Delta]\times\mathbb{R}$, let $q^*\in\mathbb{R}$ be such that

$$H(t,x,\partial_x\phi(x)) = \max_{q\in\mathbb{R}}\{q\partial_x\phi(x) - L(t,x,q)\} = q^*\partial_x\phi(x) - L(t,x,q^*).$$

From Proposition 2.3 (iv), we have $|q^*| \leq \xi(|\partial_x\phi(x)|) \leq C$, where the constant C depends only on $[\phi]_1$, M and T.

Choosing $y = x-\Delta q^*$ in (2.2) and applying Itô's formula to $\phi(Y_{t+\Delta}^{t,x-\Delta q^*})$ yield

$$\begin{aligned}&\phi(x) - \mathbf{S}_t(\Delta)\phi(x) - \Delta\mathbf{L}_t\phi(x)\\ &\geq \phi(x) - \Delta L(t,x,q^*) - \phi(x-\Delta q^*) - \mathbf{E}[\phi(Y_{t+\Delta}^{t,x-\Delta q^*}) - \phi(x-\Delta q^*)|\mathcal{F}_t] - \Delta\mathbf{L}_t\phi(x)\\ &= \ (\phi(x) - \phi(x-\Delta q^*) - \Delta q^*\partial_x\phi(x))\\ &\quad - \left(\mathbf{E}\left[\int_t^{t+\Delta}\left(b(t,y)\partial_x\phi(Y_s^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2\partial_{xx}\phi(Y_s^{t,x-\Delta q^*})\right)ds|\mathcal{F}_t\right]\right.\\ &\quad \left. - \ \Delta b(t,x)\partial_x\phi(x) - \frac{1}{2}\Delta|\sigma(t,x)|^2\partial_{xx}\phi(x)\right) := (I) - (II).\end{aligned}$$

Next, we obtain a lower and an upper bound for terms (I) and (II), respectively. To this end, Taylor's expansion yields

$$
\begin{aligned}
&\phi(x) - \phi(x - \Delta q^*) - \Delta q^* \partial_x \phi(x) \\
&= \int_{x-\Delta q^*}^{x} \left(\partial_x \phi(x) - \int_s^x \partial_{xx}\phi(u)du \right) ds - \Delta q^* \partial_x \phi(x) \\
&\geq -C\Delta^2 |\partial_{xx}\phi|_0. \qquad (2.9)
\end{aligned}
$$

For term (II), applying Itô's formula to $\partial_x \phi(Y_s^{t,x-\Delta q^*})$ and $\partial_{xx}\phi(Y_s^{t,x-\Delta q^*})$ gives

$$
\begin{aligned}
&\mathbf{E}\left[\partial_x \phi(Y_s^{t,x-\Delta q^*})|\mathcal{F}_t\right] \\
&= \partial_x\phi(y) + \int_t^s \mathbf{E}\left[b(t,y)\partial_{xx}\phi(Y_u^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2 \partial_{xxx}\phi(Y_u^{t,x-\Delta q^*})|\mathcal{F}_t \right] du,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbf{E}\left[\partial_{xx} \phi(Y_s^{t,x-\Delta q^*})|\mathcal{F}_t\right] \\
&= \partial_{xx}\phi(y) + \int_t^s \mathbf{E}\left[b(t,y)\partial_{xxx}\phi(Y_u^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2 \partial_{xxxx}\phi(Y_u^{t,x-\Delta q^*})|\mathcal{F}_t \right] du.
\end{aligned}
$$

Keeping the terms involving the derivatives of ϕ and using Assumption 2.1 on b and σ, we further have

$$
\begin{aligned}
&\mathbf{E}\left[\int_t^{t+\Delta} \left(b(t,y)\partial_x\phi(Y_s^{t,x-\Delta q^*}) + \frac{1}{2}|\sigma(t,y)|^2 \partial_{xx}\phi(Y_s^{t,x-\Delta q^*}) \right) ds|\mathcal{F}_t \right] \\
&- \Delta b(t,x)\partial_x\phi(x) - \frac{1}{2}\Delta|\sigma(t,x)|^2\partial_{xx}\phi(x) \\
&\leq C\Delta^2(|\partial_x\phi|_0 + |\partial_{xx}\phi|_0 + |\partial_{xxx}\phi|_0 + |\partial_{xxxx}\phi|_0). \qquad (2.10)
\end{aligned}
$$

In turn, combining estimates (2.9) and (2.10) above, we deduce that

$$
\frac{\phi(x) - \mathbf{S}_t(\Delta)\phi(x)}{\Delta} - \mathbf{L}_t\phi(x) \geq -C\Delta(|\partial_x\phi|_0 + |\partial_{xx}\phi|_0 + |\partial_{xxx}\phi|_0 + |\partial_{xxxx}\phi|_0),
$$

where the constant C depends only on $[\phi]_1$, M and T.

To prove the reverse inequality, we work as follows. For $(t,x) \in [0, T-\Delta] \times \mathbb{R}$, let $y^* \in \mathbb{R}$ be the minimizer in (2.2) and set $p^* := \frac{x-y^*}{\Delta}$. Then, we deduce from Proposition 2.4 that $|p^*| \leq C$, where the constant C depends only on $[\phi]_1$, M and T. In turn, similar calculations as above yield

$$
\begin{aligned}
&\phi(x) - \mathbf{S}_t(\Delta)\phi(x) - \Delta\mathbf{L}_t\phi(x) \\
&= \phi(x) - \Delta L(t,x,p^*) - \phi(x-\Delta p^*) - \mathbf{E}[\phi(Y_{t+\Delta}^{t,x-\Delta p^*}) - \phi(x-\Delta p^*)|\mathcal{F}_t] - \Delta\mathbf{L}_t\phi(x) \\
&= \Delta\left(p^*\partial_x\phi(x) - L(t,x,p^*)\right) - \Delta H(t,x,\partial_x\phi(x)) - \int_x^{x-\Delta p^*} \left(\int_x^s \partial_{xx}\phi(u)du \right) ds \\
&- \Bigg(\mathbf{E}\left[\int_t^{t+\Delta} \left(b(t,y)\partial_x\phi(Y_s^{t,x-\Delta p^*}) + \frac{1}{2}|\sigma(t,y)|^2\partial_{xx}\phi(Y_s^{t,x-\Delta p^*}) \right) ds|\mathcal{F}_t \right] \\
&- \Delta b(t,x)\partial_x\phi(x) - \frac{1}{2}\Delta|\sigma(t,x)|^2\partial_{xx}\phi(x) \Bigg) \\
&\leq C\Delta^2(|\partial_x\phi|_0 + |\partial_{xx}\phi|_0 + |\partial_{xxx}\phi|_0 + |\partial_{xxxx}\phi|_0),
\end{aligned}
$$

for some constant C depending only on $[\phi]_1$, M and T. We easily conclude. □

2.2. The approximation scheme. We present the approximation scheme for equation (1.1). For $(t,x) \in \bar{Q}_{T-\Delta}$, we introduce the iterative algorithm

$$u^{\Delta}(t,x) = \mathbf{S}_t(\Delta)u^{\Delta}(t+\Delta,\cdot)(x), \tag{2.11}$$

with $u^{\Delta}(T,\cdot) = U(\cdot)$ and $\mathbf{S}_t(\Delta)$ defined in (2.2). The values between $T-\Delta$ and T are obtained by a standard linear interpolation.

Specifically, the approximation scheme is given by

$$\begin{cases} S(\Delta,t,x,u^{\Delta}(t,x),u^{\Delta}(t+\Delta,\cdot)) = 0 & \text{in } \bar{Q}_{T-\Delta}; \\ u^{\Delta}(t,x) = g^{\Delta}(t,x) & \text{in } \bar{Q}_T \backslash \bar{Q}_{T-\Delta}, \end{cases} \tag{2.12}$$

where $S : \mathbb{R}^+ \times \bar{Q}_{T-\Delta} \times \mathbb{R} \times \mathcal{C}_b(\mathbb{R}^n) \to \mathbb{R}$ and $g^{\Delta} : \bar{Q}_T \backslash \bar{Q}_{T-\Delta} \to \mathbb{R}$ are defined, respectively, by

$$S(\Delta,t,x,p,v) = \frac{p - \mathbf{S}_t(\Delta)v(x)}{\Delta} \tag{2.13}$$

and

$$g^{\Delta}(t,x) = \omega_1(t)U(x) + \omega_2(t)\mathbf{S}_{T-\Delta}(\Delta)U(x), \tag{2.14}$$

with $\omega_1(t) = (t+\Delta-T)/\Delta$ and $\omega_2(t) = (T-t)/\Delta$ being the linear interpolation weights.

Note that when $T-\Delta < t \leq T$, the approximation term g^{Δ} corresponds to the usual linear interpolation between $T-\Delta$ and T. When $t = T-\Delta$, we have $\omega_1(t) = 0$ and $\omega_2(t) = 1$ and, thus, $g^{\Delta}(T-\Delta,x) = u^{\Delta}(T-\Delta,x)$.

We first prove the well-posedness of the approximation scheme (2.12).

LEMMA 2.6. *Suppose that Assumption 2.1 is satisfied. Then, the approximation scheme (2.12) admits a unique solution $u^{\Delta} \in \mathcal{C}_b(\bar{Q}_T)$, with $|u^{\Delta}|_0 \leq C$, where the constant C depends only on M and T.*

Proof. By the stability property (iv) in Proposition 2.5, we have that $\mathbf{S}_t(\Delta)\phi$ is uniformly bounded if so is ϕ. Therefore, equation (2.12) is always well defined in $\bar{Q}_{T-\Delta}$, which yields the existence and uniqueness of the solution u^{Δ}. Furthermore, for $0 \leq t \leq T-\Delta$, $|u^{\Delta}(t,\cdot)|_0 \leq C\Delta + |u^{\Delta}(t+\Delta,\cdot)|_0$. By backward induction and the definition of g^{Δ} in (2.12), we conclude that

$$|u^{\Delta}|_0 \leq CT + \sup_{t\in(T-\Delta,T]} |g^{\Delta}(t,\cdot)|_0 \leq C,$$

where the constant C depends only on M and T. ∎

LEMMA 2.7. *Suppose that Assumption 2.1 holds. Let $u^{\Delta} \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (2.12) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (1.1). Then, there exists a constant C, depending only on M and T, such that*

$$|u - u^{\Delta}| \leq C\sqrt{\Delta} \quad \text{in } \bar{Q}_T \backslash \bar{Q}_{T-\Delta}. \tag{2.15}$$

Proof. From (2.12), we have, for $(t,x) \in \bar{Q}_T \backslash \bar{Q}_{T-\Delta}$,

$$\begin{aligned} |u(t,x) - u^{\Delta}(t,x)| &= |u(t,x) - g^{\Delta}(t,x)| \\ &= |u(t,x) - u(T,x) + \omega_2(t)(U(x) - \mathbf{S}_{T-\Delta}(\Delta)U(x))| \\ &\leq |u(t,x) - u(T,x)| + |U(x) - \mathbf{S}_{T-\Delta}(\Delta)U(x)| \\ &\leq C(\sqrt{|T-t|} + \sqrt{\Delta}) \leq C\sqrt{\Delta}, \end{aligned}$$

where the second to last inequality follows from the regularity property of the solution u (cf. Proposition 2.2) and property (v) of the operator $\mathbf{S}_t(\Delta)$ (cf. Proposition 2.5). □

Using the properties of $\mathbf{S}_t(\Delta)$ established in Proposition 2.5, we next obtain the following key properties of the approximation scheme (2.12).

PROPOSITION 2.8. *Suppose that Assumption 2.1 is satisfied. Then, for each t and Δ with $0 \le t < t+\Delta \le T$, $x \in \mathbb{R}^n$, $p \in \mathbb{R}$ and $v \in \mathcal{C}_b(\mathbb{R}^n)$, the approximation scheme $S(\Delta, t, x, p, v)$ has the following properties:*

(i) (Constant preserving) For any $c \in \mathbb{R}$,

$$S(\Delta, t, x, p+c, v+c) = S(\Delta, t, x, p, v).$$

(ii) (Monotonicity) For any $u \in \mathcal{C}_b(\mathbb{R}^n)$ with $u \le v$,

$$S(\Delta, t, x, p, u) \ge S(\Delta, t, x, p, v).$$

(iii) (Convexity) $S(\Delta, t, x, p, v)$ is convex in p and v.

(iv) (Consistency) For any $\phi \in \mathcal{C}_b^\infty(\bar{Q}_T)$, there exists a constant C, depending only on $[\phi]_{2,1}$, M and T, such that

$$\begin{aligned} &|-\partial_t\phi(t,x) + \mathbf{L}_t\phi(t,x) - S(\Delta, t, x, \phi(t,x), \phi(t+\Delta, \cdot))| \\ &\le C\Delta\left(|\partial_{tt}\phi|_0 + |\partial_{xxxx}\phi|_0 + |\partial_{xxt}\phi|_0 + \mathcal{R}(\phi)\right). \end{aligned} \tag{2.16}$$

Proof. Parts (i)-(iii) follow easily from Proposition 2.5, so we only prove (iv). To this end, we split the consistency error into three parts. Specifically,

$$\begin{aligned} &|-\partial_t\phi(t,x) + \mathbf{L}_t\phi(t,x) - S(\Delta, t, x, \phi(t,x), \phi(t+\Delta, \cdot))| \\ &\le \mathcal{E}(t, \Delta, \phi(t+\Delta, \cdot)) + |\phi(t+\Delta, x) - \phi(t,x) - \Delta\partial_t\phi(t,x)|\Delta^{-1} \\ &\quad + |\mathbf{L}_t\phi(t,x) - \mathbf{L}_t\phi(t+\Delta, x)| := (I) + (II) + (III), \end{aligned}$$

where $\mathcal{E}$ was defined in (2.6). For term (I), Proposition 2.5 (vi) yields

$$\begin{aligned} \mathcal{E}(t, \Delta, \phi(t+\Delta, \cdot)) &\le C\Delta\left(|\partial_{xxxx}\phi(t+\Delta, \cdot)|_0 + \mathcal{R}(\phi(t+\Delta, \cdot))\right) \\ &\le C\Delta\left(|\partial_{xxxx}\phi|_0 + \mathcal{R}(\phi)\right), \end{aligned} \tag{2.17}$$

for some constant C depending only on $[\phi]_{2,1}$, M and T. For term (II), Taylor's expansion gives

$$\begin{aligned} &|\phi(t+\Delta, x) - \phi(t,x) - \Delta\partial_t\phi(t,x)|\Delta^{-1} \\ &\le |\int_t^{t+\Delta}\left(\partial_t\phi(t,x) - \int_v^t \partial_{tt}\phi(u,x)du\right)dv - \Delta\partial_t\phi(t,x)|\Delta^{-1} \\ &\le \Delta|\partial_{tt}\phi|_0. \end{aligned} \tag{2.18}$$

Finally, for term (III), we have from Assumption 2.1 that

$$\begin{aligned} &|\mathbf{L}_t\phi(t,x) - \mathbf{L}_t\phi(t+\Delta, x)| \\ &\le C(|\partial_{xx}\phi(t,x) - \partial_{xx}\phi(t+\Delta, x)| + |\partial_x\phi(t,x) - \partial_x\phi(t+\Delta, x)|) \\ &\quad + |H(t, x, \partial_x\phi(t,x)) - H(t, x, \partial_x\phi(t+\Delta, x))| \\ &\le C\Delta(|\partial_{xxt}\phi|_0 + |\partial_{xt}\phi|_0), \end{aligned} \tag{2.19}$$

for some constant C depending only on $[\phi]_{2,1}$ and M. Combining estimates (2.17)-(2.19), we easily conclude. □

The following "comparison-type" result for the approximation scheme (2.12) will be used frequently in the next section. Most of the arguments follow from Lemma 3.2 of [2], but we highlight some key steps for the reader's convenience.

PROPOSITION 2.9. *Suppose that Assumption 2.1 is satisfied, and that* $u,\ v \in \mathcal{C}_b(\bar{Q}_T)$ *are such that*

$$S(\Delta, t, x, u, u(t+\Delta, \cdot)) \le h_1 \quad \text{in } \bar{Q}_{T-\Delta};$$
$$S(\Delta, t, x, v, v(t+\Delta, \cdot)) \ge h_2 \quad \text{in } \bar{Q}_{T-\Delta},$$

for some $h_1,\ h_2 \in \mathcal{C}_b(\bar{Q}_{T-\Delta})$. *Then,*

$$u - v \le \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u-v)^+ + (T-t) \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \quad \text{in } \bar{Q}_T. \tag{2.20}$$

Proof. We first note that without loss of generality, we may assume that

$$u \le v \ \text{ in } \bar{Q}_T \setminus \bar{Q}_{T-\Delta} \ \text{ and } \ h_1 \le h_2 \ \text{ in } \bar{Q}_{T-\Delta}, \tag{2.21}$$

since, otherwise, the function $w := v + \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u-v)^+ + (T-t) \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+$ satisfies $u \le w$ in $\bar{Q}_T \setminus \bar{Q}_{T-\Delta}$ and, by the monotonicity property (ii) in Proposition 2.8,

$$\begin{aligned} S(\Delta, t, x, w, w(t+\Delta, \cdot)) &\ge S(\Delta, t, x, v, v(t+\Delta, \cdot)) + \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \\ &\ge h_2 + \sup_{\bar{Q}_{T-\Delta}} (h_1 - h_2)^+ \ge h_1 \ \text{ in } \bar{Q}_{T-\Delta}. \end{aligned}$$

Thus, it suffices to prove that $u \le v$ in $\bar{Q}_T$ when (2.21) holds.

To this end, for $b \ge 0$, let $\psi_b(t) := b(T-t)$ and $M(b) := \sup_{\bar{Q}_T} \{u - v - \psi_b\}$. We need to show that $M(0) \le 0$. We argue by contradiction. If $M(0) > 0$, then by the continuity of M, we must have $M(b) > 0$, for some $b > 0$. For such b, consider a sequence, say $\{(t_n, x_n)\}$ in $\bar{Q}_T$, such that for $\delta(t,x) := M(b) - (u - v - \psi_b)(t,x)$, we have $\lim_{n\to\infty} \delta(t_n, x_n) = 0$. Since $M(b) > 0$ but $u - v - \psi_b \le 0$ in $\bar{Q}_T \setminus \bar{Q}_{T-\Delta}$, we must have $t_n \le T - \Delta$ for sufficiently large n. Then, for such n, we have

$$\begin{aligned} h_1(t_n, x_n) &\ge S(\Delta, t_n, x_n, u(t_n, x_n), u(t_n + \Delta, \cdot)) \\ &\ge S(\Delta, t_n, x_n, (v + \psi_b + M(b) - \delta)(t_n, x_n), (v + \psi_b + M(b))(t_n + \Delta, \cdot)) \\ &\ge S(\Delta, t_n, x_n, v(t_n, x_n), v(t_n + \Delta, \cdot)) + (\psi_b(t_n) - \psi_b(t_n + \Delta) - \delta(t_n, x_n))\Delta^{-1} \\ &\ge h_2(t_n, x_n) + b - \delta(t_n, x_n)\Delta^{-1}. \end{aligned}$$

On the other hand, since $h_1 \le h_2$ in $\bar{Q}_{T-\Delta}$, we must have $b - \delta(t_n, x_n)\Delta^{-1} \le 0$. Then, letting $n \to \infty$, we deduce that $b \le 0$, which is a contradiction. □

Following along similar argument, we also obtain the comparison inequality (2.20) on the partition grid $\bar{\mathcal{G}}_T^\Delta : \{0 < \Delta < 2\Delta < \cdots < T - \Delta < T\}$.

COROLLARY 2.10. *Let* $\mathcal{G}_T^\Delta := \bar{\mathcal{G}}_T^\Delta \setminus \{T\}$ *be the partition grid before terminal time* T. *Suppose that* $u,\ v \in \mathcal{C}_b(\bar{Q}_T)$ *are such that*

$$S(\Delta, t, x, u, u(t+\Delta, \cdot)) \le h_1 \quad \text{in } \mathcal{G}_T^\Delta;$$
$$S(\Delta, t, x, v, v(t+\Delta, \cdot)) \ge h_2 \quad \text{in } \mathcal{G}_T^\Delta,$$

for some h_1, $h_2 \in \mathcal{C}_b(\mathcal{G}_T^\Delta)$. *Then,*

$$u - v \le |(u(T,\cdot) - v(T,\cdot))^+|_0 + (T-t)|(h_1 - h_2)^+|_0 \quad \text{in } \bar{\mathcal{G}}_T^\Delta. \tag{2.22}$$

3. Convergence rate of the approximation scheme. The classical convergence theory of Barles-Souganidis (see [4]) will only imply the convergence of the approximate solution u^Δ to the viscosity solution u of equation (1.1). To further determine the convergence rate of u^Δ to u, we establish upper and lower bounds on the approximation error.

We start with the special case when (1.1) has a unique smooth solution u with bounded derivatives of any order.

THEOREM 3.1. *Suppose that Assumption 2.1 is satisfied and that equation (1.1) admits a unique smooth solution* $u \in \mathcal{C}_b^\infty(\bar{Q}_T)$. *Then, there exists a constant* C*, depending only on* M *and* T*, such that*

$$|u - u^\Delta| \le C\Delta \quad \text{in } \bar{Q}_T.$$

Proof. Using that $u \in \mathcal{C}_b^\infty(\bar{Q}_T)$, the consistency error estimate (2.16) yields

$$\begin{aligned} &|S(\Delta, t, x, u(t,x), u(t+\Delta, \cdot))| \\ &\le C\Delta\left(|\partial_{tt}u|_0 + |\partial_{xxxx}u|_0 + |\partial_{xxt}u|_0 + \mathcal{R}(u)\right) \le C\Delta, \end{aligned}$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. On the other hand, from the definition of the approximation scheme (2.12), we have

$$S(\Delta, t, x, u^\Delta(t,x), u^\Delta(t+\Delta, \cdot)) = 0,$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. In turn, the comparison result in Proposition 2.9 yields

$$u(t,x) - u^\Delta(t,x) \le \sup_{(t,x)\in\bar{Q}_T\backslash\bar{Q}_{T-\Delta}} (u(t,x) - u^\Delta(t,x))^+ + (T-t)C\Delta,$$

and

$$u^\Delta(t,x) - u(t,x) \le \sup_{(t,x)\in\bar{Q}_T\backslash\bar{Q}_{T-\Delta}} (u^\Delta(t,x) - u(t,x))^+ + (T-t)C\Delta,$$

for $(t,x) \in \bar{Q}_T$. It is left to prove that $|u - u^\Delta| < C\Delta$ in $\bar{Q}_T\backslash\bar{Q}_{T-\Delta}$. Indeed, the comparison result in Corollary 2.10 yields

$$|u(T-\Delta, x) - u^\Delta(T-\Delta, x)| \le (T - (T-\Delta))C\Delta = C\Delta^2,$$

and thus, in $\bar{Q}_T\backslash\bar{Q}_{T-\Delta}$,

$$\begin{aligned} &|u(t,x) - u^\Delta(t,x)| = |u(t,x) - \omega_1(t)U(x) - \omega_2(t)u^\Delta(T-\Delta, x)| \\ &= |u(t,x) - U(x) + \omega_2(t)(U(x) - u^\Delta(T-\Delta, x))| \\ &\le |u(t,x) - U(x)| + \omega_2(t)\left(|U(x) - u(T-\Delta, x)| + |u(T-\Delta, x) - u^\Delta(T-\Delta, x)|\right) \\ &\le [u]_{1,1}\Delta + \omega_2(t)[u]_{1,1}\Delta + \omega_2(t)C\Delta^2 \le C\Delta. \end{aligned}$$

We easily conclude. □

In general, the above result might not hold as (1.1) only admits a viscosity solution $u \in \mathcal{C}_b^1(\bar{Q}_T)$ due to possible degeneracies. A natural idea is then to approximate the viscosity solution u by a sequence of smooth sub- and supersolutions u_ε and, in turn, compare them with u^Δ using the comparison result for the approximation scheme developed in Proposition 2.9. We carry out this procedure next.

3.1. Upper bound for the approximation error. We derive an upper bound for the approximation error $u - u^\Delta$. We do so by first constructing a sequence of smooth subsolutions to equation (1.1) by perturbing its coefficients. As we mentioned in the introduction, this approach, known as the *shaking coefficients technique*, was initially proposed by Krylov [24] [25], and further developed by Barles and Jakobsen [3] [22].

To this end, for $\varepsilon \in [0,1]$, we extend the functions $f := \sigma, b$ and H to $Q_{T+\varepsilon^2}^{-\varepsilon^2} := [-\varepsilon^2, T+\varepsilon^2) \times \mathbb{R}^n$ and $Q_{T+\varepsilon^2}^{-\varepsilon^2} \times \mathbb{R}^n$, respectively, so that Assumption 2.1 still holds. We then define $f^\theta(t,x) := f(t+\tau, x+e)$ and $H^\theta(t,x,p) := H(t+\tau, x+e, p)$, where $\theta = (\tau, e)$ with $\theta \in \Theta^\varepsilon := [-\varepsilon^2, 0] \times \varepsilon B(0,1)$, and consider the perturbed version of equation (1.1), namely,

(3.1)
$$\begin{cases} -\partial_t u^\varepsilon + \sup\limits_{\theta \in \Theta^\varepsilon} \left\{ -\dfrac{1}{2} \mathrm{Trace}\left(\sigma^\theta {\sigma^\theta}^T(t,x) \partial_{xx} u^\varepsilon \right) - b^\theta(t,x) \cdot \partial_x u^\varepsilon + H^\theta(t,x,\partial_x u^\varepsilon) \right\} = 0 \\ \hfill \text{in } Q_{T+\varepsilon^2}; \\ \hfill u^\varepsilon(T+\varepsilon^2, x) = U(x) \text{ in } \mathbb{R}^n. \end{cases}$$

Note that when the perturbation parameter $\varepsilon = 0$, equations (3.1) and (1.1) coincide.

We establish existence, uniqueness and regularity results for the HJB equation (3.1), and a comparison between u and u^ε. Their proofs are provided in Appendix A.

Proposition 3.2. *Suppose that Assumption 2.1 is satisfied. Then, there exists a unique viscosity solution $u^\varepsilon \in \mathcal{C}_b^1(\bar{Q}_{T+\varepsilon^2})$ of the HJB equation (3.1), with $|u^\varepsilon|_1 \le C$, for some constant C depending only on M and T. Moreover,*

$$|u - u^\varepsilon| \le C\varepsilon \quad \text{in } \bar{Q}_T. \tag{3.2}$$

Next, we regularize u^ε by a standard mollification procedure. For this, let $\rho(t,x)$ be a $\mathbb{R}_+$-valued smooth function with compact support $\{-1 < t < 0\} \times \{|x| < 1\}$ and mass 1, and introduce the sequence of mollifiers ρ_ε,

$$\rho_\varepsilon(t,x) := \frac{1}{\varepsilon^{n+2}} \rho\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right). \tag{3.3}$$

For $(t,x) \in \bar{Q}_T$, we then define

$$u_\varepsilon(t,x) = u^\varepsilon * \rho_\varepsilon(t,x) = \int_{-\varepsilon^2 < \tau < 0} \int_{|e| < \varepsilon} u^\varepsilon(t-\tau, x-e) \rho_\varepsilon(\tau, e)\, de\, d\tau.$$

Standard properties of mollifiers imply that $u_\varepsilon \in \mathcal{C}_b^\infty(\bar{Q}_T)$,

$$|u^\varepsilon - u_\varepsilon|_0 \le C\varepsilon, \tag{3.4}$$

and, moreover, for positive integers i and j,

$$|\partial_t^i \partial_x^j u_\varepsilon|_0 \leq C\varepsilon^{1-2i-|j|}, \tag{3.5}$$

where the constant C is independent of ε.

We observe that the function $u^\varepsilon(t-\tau, x-e)$, $(t,x) \in Q_T$, is a viscosity subsolution of equation (1.1) in Q_T, for any $(\tau, e) \in \Theta^\varepsilon$. On the other hand, a Riemann sum approximation shows that $u_\varepsilon(t,x)$ can be viewed as the limit of convex combinations of $u^\varepsilon(t-\tau, x-e)$, for $(\tau, e) \in \Theta^\varepsilon$. Since the equation in (1.1) is convex in $\partial_x u$, and linear in $\partial_t u$ and $\partial_{xx} u$, the convex combinations of $u^\varepsilon(t-\tau, x-e)$ are also subsolutions of (1.1) in Q_T. Using the stability of viscosity solutions, we deduce that $u_\varepsilon(t,x)$ is also a subsolution of (1.1) in Q_T.

We are now ready to establish an upper bound for the approximation error.

THEOREM 3.3. *Suppose that Assumption 2.1 holds. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (2.12) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (1.1). Then, there exists a constant C, depending only on M and T, such that*

$$u - u^\Delta \leq C\Delta^{\frac{1}{4}} \quad \text{in } \bar{Q}_T.$$

Proof. Substituting u_ε into the consistency error estimate (2.16) and using (3.5) give

$$\begin{aligned}
&|-\partial_t u_\varepsilon(t,x) + \mathbf{L}_t u_\varepsilon(t,x) - S(\Delta, t, x, u_\varepsilon(t,x), u_\varepsilon(t+\Delta, \cdot))| \\
&\leq C\Delta\left(|\partial_{tt} u_\varepsilon|_0 + |\partial_{xxxx} u_\varepsilon|_0 + |\partial_{xxt} u_\varepsilon|_0 + \mathcal{R}(u_\varepsilon)\right) \leq C\Delta\varepsilon^{-3},
\end{aligned}$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. Since u_ε is a subsolution of (1.1) in Q_T, we have

$$S(\Delta, t, x, u_\varepsilon(t,x), u_\varepsilon(t+\Delta, \cdot)) \leq C\Delta\varepsilon^{-3},$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. Furthermore, by the definition of the approximation scheme (2.12), we also have

$$S(\Delta, t, x, u^\Delta(t,x), u^\Delta(t+\Delta, \cdot)) = 0,$$

for $(t,x) \in \bar{Q}_{T-\Delta}$. In turn, Proposition 2.9 implies

$$u_\varepsilon - u^\Delta \leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u_\varepsilon - u^\Delta)^+ + C(T-t)\Delta\varepsilon^{-3} \;\text{ in } \bar{Q}_T.$$

Next, using estimates (3.2) and (3.4), we obtain that $|u - u_\varepsilon| \leq C\varepsilon$ and, thus,

$$\begin{aligned}
u - u^\Delta &= (u - u_\varepsilon) + (u_\varepsilon - u^\Delta) \\
&\leq C\varepsilon + \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u_\varepsilon - u^\Delta)^+ + C(T-t)\Delta\varepsilon^{-3} \\
&\leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u - u^\Delta)^+ + C(\varepsilon + \Delta\varepsilon^{-3}) \;\text{ in } \bar{Q}_T.
\end{aligned}$$

By choosing $\varepsilon = \Delta^{\frac{1}{4}}$, we further deduce that

$$u - u^\Delta \leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u - u^\Delta)^+ + C\Delta^{\frac{1}{4}} \;\text{ in } \bar{Q}_T.$$

We conclude using estimate (2.15) in Lemma 2.7. ∎

3.2. Lower bound for the approximation error. To obtain a lower bound of $u-u^{\Delta}$, we cannot follow the above perturbation procedure to construct approximate smooth supersolutions to equation (1.1). This is because if we perturb its coefficients to obtain a viscosity supersolution, its convolution with the mollifier may no longer be a supersolution due to the convexity of equation (1.1) with respect to its terms. Furthermore, interchanging the roles (as in [19]) of equation (1.1) and its approximation scheme (2.12) does not work either, because the solution u^{Δ} of the approximation scheme (and its perturbation solution) may in general lose the Hölder and Lipschitz continuity in (t,x). This is due to the lack of the continuous dependence result for the approximation scheme, compared with the continuous dependence result for equation (1.1) and its perturbation equation (3.1) (see Lemma A.1).

To overcome these difficulties, we follow the idea of Barles and Jakobsen [2] to build approximate supersolutions which are smooth at the "right points" by introducing an appropriate optimal switching stochastic control system. To apply this method to the problem herein, we first observe that, using the convex dual function L introduced in (2.1), we can write equation (1.1) as a HJB equation, namely,

$$\left\{\begin{array}{ll} -\partial_t u+\sup\limits_{q\in\mathbb{R}^n}\mathcal{L}^q\left(t,x,\partial_x u,\partial_{xx}u\right)=0 & \text{in } Q_T;\\ u(T,x)=U(x) & \text{in } \mathbb{R}^n, \end{array}\right. \tag{3.6}$$

with

$$\mathcal{L}^q(t,x,p,X):=-\frac{1}{2}\mathrm{Trace}\left(\sigma\sigma^T(t,x)X\right)-(b(t,x)-q)\cdot p-L(t,x,q).$$

It then follows from Proposition 2.3 (iv) that the supremum can be achieved at some point, say q^*, with $|q^*|\leq\xi(|\partial_x u|)$. Furthermore, Proposition 2.2 implies that $|q^*|\leq C$, for some constant C depending only on M and T. Thus, we rewrite the equation in (3.6) as

$$-\partial_t u+\sup_{q\in K}\mathcal{L}^q\left(t,x,\partial_x u,\partial_{xx}u\right)=0,$$

where $K\subset\mathbb{R}^n$ is a compact set. Since K is separable, it has a countable dense subset, say $K_\infty=\{q_1,q_2,q_3,...\}$ and, in turn, the continuity of $\mathcal{L}^q$ in q implies that

$$\sup_{q\in K}\mathcal{L}^q(t,x,p,X)=\sup_{q\in K_\infty}\mathcal{L}^q(t,x,p,X).$$

Therefore, the equation in (3.6) further reduces to

$$-\partial_t u+\sup_{q\in K_\infty}\mathcal{L}^q\left(t,x,\partial_x u,\partial_{xx}u\right)=0.$$

For $m\geq 1$, we now consider the approximations of (3.6),

$$\left\{\begin{array}{ll} -\partial_t u^m+\sup\limits_{q\in K_m}\mathcal{L}^q\left(t,x,\partial_x u^m,\partial_{xx}u^m\right)=0 & \text{in } Q_T;\\ u^m(T,x)=U(x) & \text{in } \mathbb{R}^n, \end{array}\right. \tag{3.7}$$

where $K_m:=\{q_1,...,q_m\}\subset K_\infty$, i.e. K_m consists of the first m points in K_∞ and satisfies $\cup_{m\geq 1}K_m=K_\infty$. It then follows from Proposition 2.1 of [2] that (3.7) admits a unique viscosity solution $u^m\in\mathcal{C}_b^1(\bar{Q}_T)$, with $|u^m|_1\leq C$, for some constant C

depending only on M and T. Furthermore, Arzela-Ascoli's theorem yields that there exists a subsequence of $\{u^m\}$, still denoted as $\{u^m\}$, such that, as $m \to \infty$,

$$u^m(t,x) \to u(t,x) \ \text{ uniformly in } (t,x) \in \bar{Q}_T. \tag{3.8}$$

Next, we construct a sequence of (local) smooth supersolutions to approximate u^m. For this, we consider the optimal switching system

$$\begin{cases} \max\left\{-\partial_t v_i + \mathcal{L}^{q_i}(t,x,\partial_x v_i, \partial_{xx} v_i), v_i - \mathcal{M}_i^k v\right\} = 0 & \text{in } Q_T; \\ v_i(T,x) = U(x) & \text{in } \mathbb{R}^n, \end{cases} \tag{3.9}$$

where $i \in \mathcal{I} := \{1, ..., m\}$ and $\mathcal{M}_i^k v := \min_{j \neq i,\ j \in \mathcal{I}}\{v_j + k\}$, for some constant $k > 0$ representing the switching cost.

Proposition 3.4. *Suppose that Assumption 2.1 is satisfied. Then, there exists a unique viscosity solution $v = (v_1, \ldots, v_m)$ of the optimal switching system (3.9) such that $|v|_1 \leq C$, for some constant C depending only on M and T. Moreover, for $i \in \mathcal{I}$,*

$$0 \leq v_i - u^m \leq C(k^{\frac{1}{3}} + k^{\frac{2}{3}}) \quad \text{in } \bar{Q}_T. \tag{3.10}$$

The proof essentially follows from Proposition 2.1 and Theorem 2.3 of [2] and it is thus omitted. We only remark that since we do not require the switching cost to satisfy $k \leq 1$, we keep the term $k^{\frac{2}{3}}$ in the above estimate. This will not affect the convergence rate of the approximation scheme.

Next, still following the approach of [2], we construct smooth approximations of v_i. Since in the continuation region of (3.9), the solution v_i satisfies the *linear* equation, namely,

$$-\partial_t v_i + \mathcal{L}^{q_i}(t,x,\partial_x v_i, \partial_{xx} v_i) = 0 \ \text{ in } \{(t,x) \in Q_T : v_i(t,x) < \mathcal{M}_i^k v(t,x)\},$$

we may perturb its coefficients to obtain a sequence of smooth supersolutions. This will in turn give a lower bound of the error $u^m - u^\Delta$. A subtle point herein is how to identify the continuation region by appropriately choosing the switching cost k. For this, we follow the idea used in Lemma 3.4 of [2].

Proposition 3.5. *Suppose that Assumption 2.1 holds. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (2.12) and $u^m \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of the HJB equation (3.7). Then, there exists a constant C, depending only on M and T, such that*

$$u^\Delta - u^m \leq \sup_{\bar{Q}_T \setminus \bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C\Delta^{\frac{1}{10}} \quad \text{in } \bar{Q}_T.$$

Proof. Let $\varepsilon \in [0,1]$. In analogy to (3.1), we perturb the coefficients of the optimal switching system (3.9) and consider

$$\begin{cases} \max\left\{-\partial_t v_i^\varepsilon + \inf_{(\tau,e)\in\Theta^\varepsilon} \mathcal{L}^{q_i}(t+\tau, x+e, \partial_x v_i^\varepsilon, \partial_{xx} v_i^\varepsilon), v_i^\varepsilon - \mathcal{M}_i^k v^\varepsilon\right\} = 0 & \text{in } Q_{T+\varepsilon^2}; \\ v_i^\varepsilon(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^n. \end{cases} \tag{3.11}$$

It then follows from Proposition 2.2 of [2] that (3.11) admits a unique viscosity solution, say $v^\varepsilon = (v_1^\varepsilon, \ldots, v_m^\varepsilon)$, with $|v^\varepsilon|_1 \leq C$ and, moreover, for each $i \in \mathcal{I}$,

$$|v_i^\varepsilon - v_i| \leq C\varepsilon \ \text{ in } \bar{Q}_T, \tag{3.12}$$

where the constant C depends only on M and T. In turn, inequalities (3.10) and (3.12) imply that, for each $i \in \mathcal{I}$,

$$|v_i^\varepsilon - u^m| \le |v_i^\varepsilon - v_i| + |v_i - u^m| \le C(\varepsilon + k^{\frac{1}{3}} + k^{\frac{2}{3}}) \text{ in } \bar{Q}_T. \tag{3.13}$$

Next, we regularize v_i^ε by introducing $v_{i,\varepsilon}(t,x) := v_i^\varepsilon * \rho_\varepsilon(t,x)$, for $(t,x) \in \bar{Q}_T$, where ρ_ε is the mollifer defined in (3.3). Then, $v_{i,\varepsilon} \in \mathcal{C}_b^\infty(\bar{Q}_T)$,

$$|v_{i,\varepsilon} - v_i^\varepsilon|_0 \le C\varepsilon, \tag{3.14}$$

and, moreover, for positive integers m and n,

$$|\partial_t^m \partial_x^n v_{i,\varepsilon}|_0 \le C\varepsilon^{1-2m-|n|}. \tag{3.15}$$

We introduce the function $w_\varepsilon := \min_{i\in\mathcal{I}} v_{i,\varepsilon}$, which is smooth in $\bar{Q}_T$ except for finitely many points. Then, (3.13) and (3.14) yield

$$|u^m - w_\varepsilon| \le C(\varepsilon + k^{\frac{1}{3}} + k^{\frac{2}{3}}) \text{ in } \bar{Q}_T. \tag{3.16}$$

For each $(t,x) \in \bar{Q}_T$, let $j := \arg\min_{i\in\mathcal{I}} v_{i,\varepsilon}(t,x)$. Then, $w_\varepsilon(t,x) = v_{j,\varepsilon}(t,x)$ and, for such j, we obtain that

$$v_{j,\varepsilon}(t,x) - \mathcal{M}_j^k v_\varepsilon(t,x) = \max_{i\neq j, i\in\mathcal{I}} \{v_{j,\varepsilon}(t,x) - v_{i,\varepsilon}(t,x) - k\} \le -k.$$

In turn, inequality (3.14) implies that

$$v_j^\varepsilon(t,x) - \mathcal{M}_j^k v^\varepsilon(t,x) \le v_{j,\varepsilon}(t,x) - \mathcal{M}_j^k v_\varepsilon(t,x) + C\varepsilon \le -k + C\varepsilon.$$

Furthermore, since $|v^\varepsilon|_1 \le C$ for $v^\varepsilon = (v_1^\varepsilon, \ldots, v_m^\varepsilon)$, we also have

$$\begin{aligned} v_j^\varepsilon(t-\tau, x-e) - \mathcal{M}_j^k v^\varepsilon(t-\tau, x-e) &\le v_j^\varepsilon(t,x) - \mathcal{M}_j^k v^\varepsilon(t,x) + C(|\tau|^{\frac{1}{2}} + |e|) \\ &\le -k + C\varepsilon + 2C\varepsilon, \end{aligned}$$

for any $(\tau, e) \in \Theta^\varepsilon$. If we then choose $k = 4C\varepsilon$, we obtain that, for any $(\tau, e) \in \Theta^\varepsilon$,

$$v_j^\varepsilon(t-\tau, x-e) - \mathcal{M}_j^k v^\varepsilon(t-\tau, x-e) < 0.$$

Therefore, the point $(t-\tau, x-e)$, for $(\tau, e) \in \Theta^\varepsilon$, is in the continuation region of (3.11). Thus,

$$-\partial_t v_j^\varepsilon(t-\tau, x-e) + \inf_{(\tau,e)\in\Theta^\varepsilon} \mathcal{L}^{q_j}\left(t, x, \partial_x v_j^\varepsilon(t-\tau, x-e), \partial_{xx} v_j^\varepsilon(t-\tau, x-e)\right) = 0,$$

and, in turn,

$$-\partial_t v_j^\varepsilon(t-\tau, x-e) + \mathcal{L}^{q_j}\left(t, x, \partial_x v_j^\varepsilon(t-\tau, x-e), \partial_{xx} v_j^\varepsilon(t-\tau, x-e)\right) \ge 0.$$

Using the definition of $v_{j,\varepsilon}$ and that $\mathcal{L}^{q_j}$ is linear in $\partial_x v_j^\varepsilon$ and $\partial_{xx} v_j^\varepsilon$, we further have

$$\begin{aligned} &- \partial_t v_{j,\varepsilon}(t,x) + \mathcal{L}^{q_j}\left(t, x, \partial_x v_{j,\varepsilon}(t,x), \partial_{xx} v_{j,\varepsilon}(t,x)\right) \\ &= \int_{-\varepsilon^2<\tau<0} \int_{|e|<\varepsilon} \left(-\partial_t v_j^\varepsilon(t-\tau, x-e) + \mathcal{L}^{q_j}\left(t, x, \partial_x v_j^\varepsilon(t-\tau, x-e), \partial_{xx} v_j^\varepsilon(t-\tau, x-e)\right)\right) \\ &\quad \times \rho_\varepsilon(\tau, e) de d\tau \ge 0. \end{aligned} \tag{3.17}$$

Next, we observe that, for $(t,x) \in \bar{Q}_{T-\Delta}$, the definition of j implies that $w_\varepsilon(t,x) = v_{j,\varepsilon}(t,x)$ and $w_\varepsilon(t+\Delta,\cdot) \le v_{j,\varepsilon}(t+\Delta,\cdot)$. Then, applying Proposition 2.8 (ii) (iv) and estimate (3.15), we obtain that, for any $(t,x) \in \bar{Q}_{T-\Delta}$,

$$\begin{aligned}
& S(\Delta,t,x,w_\varepsilon(t,x),w_\varepsilon(t+\Delta,\cdot)) \\
\ge\ & S(\Delta,t,x,v_{j,\varepsilon}(t,x),v_{j,\varepsilon}(t+\Delta,\cdot)) \\
\ge\ & -\partial_t v_{j,\varepsilon}(t,x) + \sup_{q\in\mathbb{R}^n} \mathcal{L}^q(t,x,\partial_x v_{j,\varepsilon}(t,x),\partial_{xx} v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \\
\ge\ & -\partial_t v_{j,\varepsilon}(t,x) + \mathcal{L}^{q_j}(t,x,\partial_x v_{j,\varepsilon}(t,x),\partial_{xx} v_{j,\varepsilon}(t,x)) - C\Delta\varepsilon^{-3} \ge -C\Delta\varepsilon^{-3},
\end{aligned}$$

for some constant C depending only on M and T, where we used (3.17) in the last inequality. In turn, the comparison result in Proposition 2.9 implies that

$$u^\Delta - w_\varepsilon \le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - w_\varepsilon)^+ + C(T-t)\Delta\varepsilon^{-3} \ \text{ in } \bar{Q}_T.$$

Combining the above inequality with (3.16), we further get

$$\begin{aligned}
u^\Delta - u^m &= (u^\Delta - w_\varepsilon) + (w_\varepsilon - u^m) \\
&\le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - w_\varepsilon)^+ + C(T-t)\Delta\varepsilon^{-3} + C(\varepsilon + k^{\frac{1}{3}} + k^{\frac{2}{3}}) \\
&\le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C(\varepsilon + \varepsilon^{\frac{1}{3}} + \varepsilon^{\frac{2}{3}} + \Delta\varepsilon^{-3}) \\
&\le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C\Delta^{\frac{1}{10}} \ \text{ in } \bar{Q}_T,
\end{aligned}$$

where we used $k = 4C\varepsilon$ in the second to last inequality, and chose $\varepsilon = \Delta^{\frac{3}{10}}$ in the last inequality. ∎

We are now ready to establish a lower bound for the approximation error.

Theorem 3.6. *Suppose that Assumption 2.1 holds. Let $u^\Delta \in \mathcal{C}_b(\bar{Q}_T)$ satisfy the approximation scheme (2.12) and $u \in \mathcal{C}_b^1(\bar{Q}_T)$ be the unique viscosity solution of equation (1.1). Then, there exists a constant C, depending only on M and T, such that*

$$u - u^\Delta \ge -C\Delta^{\frac{1}{10}} \quad \textit{in } \bar{Q}_T.$$

Proof. Proposition 3.5 yields

$$\begin{aligned}
u^\Delta - u &= (u^\Delta - u^m) + (u^m - u) \\
&\le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - u^m)^+ + C\Delta^{\frac{1}{10}} + (u^m - u) \\
&\le \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u^\Delta - u)^+ + C\Delta^{\frac{1}{10}} + \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u - u^m)^+ + (u^m - u) \\
&\le C\Delta^{\frac{1}{10}} + \sup_{\bar{Q}_T\setminus\bar{Q}_{T-\Delta}} (u - u^m)^+ + (u^m - u),
\end{aligned}$$

where we used estimate (2.15) in the last inequality. Sending $m \to \infty$ and using (3.8), we conclude. ∎

4. A numerical example. We present a numerical result, applying the approximation scheme (2.12) for the case

$$\sigma(t,x)=1,\ b(t,x)=0,\ H(t,x,p)=p^2/2,\ T=1.$$

We also choose $U(x)=0\vee x\wedge K$ in the semilinear PDE (1.1). Then the equation in (1.1) becomes the Cole-Hopf equation (see [14]):

$$-\partial_t u(t,x)-\frac{1}{2}\partial_{xx}u(t,x)+\frac{1}{2}(\partial_x u(t,x))^2=0. \tag{4.1}$$

It is well known that, by the Cole-Hopf transformation (see [14] and [38]), the function $v(t,x):=e^{-u(t,x)}$ satisfies the heat equation

$$\partial_t v(t,x)+\frac{1}{2}\partial_{xx}v(t,x)=0,$$

with $v(T,x)=e^{-U(x)}=e^{-0\vee x\wedge K}$. In turn,

$$\begin{aligned}v(t,x)=\Phi(-\frac{x}{\sqrt{T-t}})+e^{-x+(T-t)/2}\left(\Phi(\frac{K-x+T-t}{\sqrt{T-t}})-\Phi(\frac{-x+T-t}{\sqrt{T-t}})\right)\\+e^{-K}\Phi(-\frac{K-x}{\sqrt{T-t}}),\end{aligned}$$

where Φ is the standard normal cumulative distribution function and, thus, we obtain the explicit solution $u(t,x)=-\log v(t,x)$.

We use this exact solution as a benchmark, and compare it with the approximate solution obtained by the approximation scheme (2.12). Moreover, we also compare our results with the ones obtained via the standard Howard's finite difference (FD) algorithm (see, for example, [6] for a detailed discussion of Howard's FD scheme).

Since Howard's scheme is based on the finite difference method, for the comparison purpose, we also numerically compute the conditional expectation appearing in the backward operator $\mathbf{S}_t(\Delta)$ (cf. (2.2)) using the finite difference method. However, we emphasize that, different from Howard's scheme, the splitting approximation itself does not depend on the finite difference method (as long as one can find an efficient way to compute conditional expectations, e.g. the multi-level Monte Carlo approach [17], the least squares Monte Carlo approach [28], the cubature approach [29], and etc). Hence, our approximation scheme can be potentially used to numerically solve high dimensional PDEs without the "curse of dimensionality" issue.

To numerically compute the finite-dimensional minimization problem in the backward operator $\mathbf{S}_t(\Delta)$ (cf. (2.2)), since the finite difference method already provides us with all the points to be compared, we use the simple brute force method to find the minimizers and minimal values[1].

Figures 1 and 2 demonstrate the performance of the approximation scheme (2.12) with the parameter $K=5$. They illustrate how the approximate solutions converge as we increase the number of time steps T/Δ. For our parameter values, $\Delta=0.1$ (so $T/\Delta=10$) is sufficient for the approximate solutions to converge, as the relative error is already negligible (0.056%).

[1]In general, we may implement the standard Nelder-Mead simplex algorithm (see [32]), which is commonly used in the literature when the derivatives of the objective function in the minimization problem are not known.

Figure 3 compares the values numerically computed by the approximation scheme (2.12) and the Howard's FD scheme with different time steps. It shows that the approximation scheme gives a better approximation than the Howard's scheme does. In particular, we observe that when the time step $\Delta = 0.1$ (so $T/\Delta = 10$), the numerical solution computed by our approximation scheme is far more accurate than the one computed by the FD scheme. The relative error is 0.056% for the former and 0.142% for the latter. It also shows that the approximation scheme converges linearly with time step Δ, and this is consistent with our theoretical results in Theorem 3.1. Table 1 further compares the computation errors and costs between the approximation scheme (2.12) and the Howard's FD scheme. Since there involves an additional minimization step in the approximation scheme (2.12), its computation costs are higher than the FD scheme. However, we observe that when the time step is small (e.g. $\Delta = 0.1$), the computation times for both schemes are extremely fast (less than 0.05 second).

5. Conclusions. We proposed an approximation scheme for a class of semilinear parabolic equations whose Hamiltonian is convex and coercive to the gradients. The scheme is based on splitting the equation in two parts, the first corresponding to a linear parabolic equation and the second to a Hamilton-Jacobi equation. The solutions of these equations are approximated using, respectively, the Feynman-Kac and the Hopf-Lax formulae. We established the convergence of the approximation scheme and determined the convergence rate, combining Krylov's shaking coefficients technique and Barles-Jakobsen's optimal switching approximation. One of the key steps is the derivation of a consistency error via convex duality arguments, using the convexity of the Hamiltonian in an essential way.

The approach and the results herein may be extended in various directions. Firstly, one may consider problem (1.1) in a bounded domain, an undoubtedly important case since various applications are cast in such domains (e.g. utilities defined in half-space, constrained risk measures, etc.) However, various non-trivial technical difficulties arise. Some recent works on such problems using other approaches can be found in [9], [26] and [34].

Secondly, one may consider variational versions of problem (1.1). These are naturally related to optimal stopping and to singular stochastic optimization problems, both directly related to various applications with early-exercise, fixed and/or proportional transaction costs, irreversible investment decisions, etc. Recent results in this direction that use some of the ideas developed herein can be found in [21].

Appendix A. Proofs of Propositions 2.2 and 3.2.

We note that equation (1.1) is a special case (choosing $\varepsilon = 0$) of the HJB equation (3.1). Therefore, we omit the proof of Proposition 2.2 and only prove Proposition 3.2.

We first show that there exists a bounded solution to (3.1). To this end, using the convex dual function $L^\theta(t,x,q) := \sup_{p\in\mathbb{R}^n}(p\cdot q - H^\theta(t,x,p))$, we rewrite (3.1) as

$$\text{(A.1)} \qquad \begin{cases} -\partial_t u^\varepsilon + \sup\limits_{\theta\in\Theta^\varepsilon, q\in\mathbb{R}^n} \mathcal{L}^{\theta,q}\left(t,x,\partial_x u^\varepsilon, \partial_{xx} u^\varepsilon\right) = 0 & \text{in } Q_{T+\varepsilon^2}; \\ u^\varepsilon(T+\varepsilon^2, x) = U(x) & \text{in } \mathbb{R}^n, \end{cases}$$

where

$$\mathcal{L}^{\theta,q}(t,x,p,X) = -\frac{1}{2}\mathrm{Trace}\left(\sigma^\theta {\sigma^\theta}^T(t,x) X\right) - (b^\theta(t,x) - q)\cdot p - L^\theta(t,x,q).$$

We also introduce the stochastic control problem

$$u^\varepsilon(t,x)=\inf_{\theta\in\Theta^\varepsilon[t,T+\varepsilon^2],q\in\mathbb{H}^2[t,T+\varepsilon^2]}\mathbf{E}\left[\int_t^{T+\varepsilon^2}L^{\theta_s}\left(s,X_s^{t,x;\theta,q},q_s\right)ds+U\left(X_{T+\varepsilon^2}^{t,x;\theta,q}\right)|\mathcal{F}_t\right],$$

with the controlled state equation

$$dX_s^{t,x;\theta,q}=\left(b^{\theta_s}(s,X_s^{t,x;\theta,q})-q_s\right)ds+\sigma^{\theta_s}\left(s,X_s^{t,x;\theta,q}\right)dW_s,$$

where $\Theta^\varepsilon[t,T+\varepsilon^2]$ is the space of Θ^ε-valued progressively measurable processes (τ_s,e_s) and $\mathbb{H}^2[t,T+\varepsilon^2]$ is the space of square-integrable progressively measurable processes q_s, for $s\in[t,T+\varepsilon^2]$. Next, we identify its value function with a bounded viscosity solution to (A.1). For this, we only need to establish upper and lower bounds for the value function $u^\varepsilon(t,x)$ and, in turn, use standard arguments as in [33] and [37].

To find an upper bound for u^ε, we choose an arbitrary perturbation parameter process $\theta\in\Theta^\varepsilon[t,T+\varepsilon]$ and choose $\hat{q}$ with $\hat{q}_s\equiv 0$. Then, Proposition 2.3 (ii) yields

$$\begin{aligned}u^\varepsilon(t,x)&\le\mathbf{E}\left[\int_t^{T+\varepsilon^2}L^{\theta_s}(s,X_s^{t,x;\theta,\hat{q}},0)ds+U(X_{T+\varepsilon^2}^{t,x;\theta,\hat{q}})|\mathcal{F}_t\right]\\&\le(T+\varepsilon^2-t)|L^*(0)|+M\le(T+1)|L^*(0)|+M.\end{aligned}$$

For the lower bound, we use again Proposition 2.3 (ii) to obtain that $L_*(q)\ge-H^*(0)\ge-|H^*(0)|$, for any $q\in\mathbb{R}^n$. In turn, for any $(\theta,q)\in\Theta^\varepsilon[t,T+\varepsilon^2]\times\mathbb{H}^2[t,T+\varepsilon^2]$,

$$\begin{aligned}&\mathbf{E}\left[\int_t^{T+\varepsilon^2}L^{\theta_s}(s,X_s^{t,x;\theta,q},q_s)ds+U(X_{T+\varepsilon^2}^{t,x;\theta,q})|\mathcal{F}_t\right]\\&\ge\mathbf{E}\left[\int_t^{T+\varepsilon^2}L_*(q_s)ds+U(X_{T+\varepsilon^2}^{t,x;\theta,q})|\mathcal{F}_t\right]\\&\ge-(T+\varepsilon^2-t)|H^*(0)|-M\ge-(T+1)|H^*(0)|-M,\end{aligned}$$

and, thus, $u^\varepsilon(t,x)\ge-(T+1)|H^*(0)|-M$ and $|u^\varepsilon|_0\le C$, for some constant C independent of ε.

The uniqueness of the viscosity solution is a direct consequence of the continuous dependence result, presented next. Its proof follows along similar arguments as in Theorem A.1 of [22] and is thus omitted.

Lemma A.1. *For any $s\in(0,T+\varepsilon^2]$, let $u\in USC(\bar{Q}_s)$ be a bounded from above viscosity subsolution of (3.1) with coefficients σ^θ,b^θ and H^θ, and $\bar{u}\in LSC(\bar{Q}_s)$ be a bounded from below viscosity supersolution of (3.1) with coefficients $\bar{\sigma}^\theta,\bar{b}^\theta$ and $\bar{H}^\theta$. Suppose that Assumption 2.1 holds for both sets of coefficients with respective constants M and $\bar{M}$, uniformly in $\theta\in\Theta^\varepsilon$, and that either $u(s,\cdot)\in\mathcal{C}_b^1(\mathbb{R}^n)$ or $\bar{u}(s,\cdot)\in\mathcal{C}_b^1(\mathbb{R}^n)$. Then, there exists a constant C, depending only on M, $\bar{M}$, $[u(s,\cdot)]_1$ or $[\bar{u}(s,\cdot)]_1$, and s, such that, for $(t,x)\in\bar{Q}_s$,*

(A.2)
$$u-\bar{u}\le C\left(|(u(s,\cdot)-\bar{u}(s,\cdot))^+|_0+\sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta-\bar{\sigma}^\theta|_0+|b^\theta-\bar{b}^\theta|_0\right\}+\sup_{\theta\in\Theta^\varepsilon}|H^\theta-\bar{H}^\theta|_0\right).$$

The x-regularity of u^ε follows easily from (A.2) by choosing $u=u^\varepsilon$, $\bar{u}=u^\varepsilon(\cdot,\cdot+e)$ and $s=T+\varepsilon^2$.

To get the time regularity, we work as follows. Firstly, let $\rho(x)$ be a $\mathbb{R}_+$-valued smooth function with compact support $B(0,1)$ and mass 1, and introduce the sequence of mollifiers $\rho_\varepsilon(x) := \frac{1}{\varepsilon^n}\rho\left(\frac{x}{\varepsilon}\right)$. For $0 \le t < s \le T + \varepsilon^2$, let $u_{\varepsilon'}$ be the unique bounded solution of (3.1) in Q_s with terminal condition $u_{\varepsilon'}(s,x) = u^\varepsilon(s,\cdot) * \rho_{\varepsilon'}(x)$, for some $\varepsilon' > 0$. It then follows from (A.2) that, for $(t,x) \in \bar{Q}_s$,

$$u^\varepsilon - u_{\varepsilon'} \le C|(u^\varepsilon(s,\cdot) - u_{\varepsilon'}(s,\cdot))^+|_0 \le C[u^\varepsilon(s,\cdot)]_1\varepsilon' \le C\varepsilon'.$$

Similarly, we also have $u_{\varepsilon'} - u^\varepsilon \le C\varepsilon'$.

On the other hand, standard properties of mollifiers imply that $|\partial_x^j u_{\varepsilon'}(s,\cdot)|_0 \le C\varepsilon'^{1-j}$. Next, define the functions $w_{\varepsilon'}^+(t,x) := u_{\varepsilon'}(s,x) + (s-t)C_{\varepsilon'}$ and $w_{\varepsilon'}^-(t,x) := u_{\varepsilon'}(s,x) - (s-t)C_{\varepsilon'}$, where $C_{\varepsilon'} = (\frac{1}{\varepsilon'} + 1)C$, for some constant C independent of ε. We easily deduce that they are, respectively, bounded supersolution and subsolution of (3.1) in Q_s, with the same terminal condition $w_{\varepsilon'}^+(s,x) = w_{\varepsilon'}^-(s,x) = u_{\varepsilon'}(s,x)$. Thus, by (A.2), we have $w_{\varepsilon'}^-(t,x) \le u_{\varepsilon'}(t,x) \le w_{\varepsilon'}^+(t,x)$, for $(t,x) \in \bar{Q}_s$, which in turn implies that $|u_{\varepsilon'}(t,x) - u_{\varepsilon'}(s,x)| \le C_{\varepsilon'}|s-t|$. Choosing $\varepsilon' = \sqrt{|s-t|}$, we then obtain that

$$\begin{aligned}|u^\varepsilon(t,x) - u^\varepsilon(s,x)| &\le |u^\varepsilon(t,x) - u_{\varepsilon'}(t,x)| + |u_{\varepsilon'}(t,x) - u_{\varepsilon'}(s,x)| + |u_{\varepsilon'}(s,x) - u^\varepsilon(s,x)| \\ &\le 2C\varepsilon' + C_{\varepsilon'}|s-t| \le C(\varepsilon' + \frac{|s-t|}{\varepsilon'} + |s-t|) \le C\sqrt{|s-t|},\end{aligned}$$

which, together with the boundedness and the x-regularity of u^ε, implies that $|u^\varepsilon|_1 \le C$.

Finally, note that $u(t,x)$ is also the bounded viscosity solution of (3.1) when $\sigma^\theta \equiv \sigma$, $b^\theta \equiv b$ and $H^\theta \equiv H$. Applying (A.2) once more and using the regularity of σ, b, H and u^ε, we deduce that

$$\begin{aligned}u^\varepsilon - u &\le C\left(|(u^\varepsilon(T,\cdot) - u(T,\cdot))^+|_0 + \sup_{\theta\in\Theta^\varepsilon}\left\{|\sigma^\theta - \sigma|_0 + |b^\theta - b|_0\right\} + \sup_{\theta\in\Theta^\varepsilon}|H^\theta - H|_0\right) \\ &\le C\left(|u^\varepsilon(T,\cdot) - u^\varepsilon(T+\varepsilon^2,\cdot)|_0 + \varepsilon\right) \le C\varepsilon \quad \text{in } \bar{Q}_T.\end{aligned}$$

Similarly, we also have $u - u^\varepsilon \le C\varepsilon$, and we easily conclude.

REFERENCES

[1] Barles, G. and E. R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM Journal on Numerical Analysis*, 43(2): 540-558, 2005.

[2] Barles, G. and E. R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Mathematics of Computation*, 76: 1861-1893, 2007.

[3] Barles, G. and E. R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.*, 36(1): 33-54, 2002.

[4] Barles, G. and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*, 4(3): 271–283, 1991.

[5] Bayraktar, E. and A. Fahim. A stochastic approximation for fully nonlinear free boundary problems. *Numerical Methods for Partial Differential Equations*, 30(3): 902-929, 2014.

[6] Bokanowski, O., S. Maroso, and H. Zidani. Some convergence results for Howard's algorithm. *SIAM Journal on Numerical Analysis*, 47(4): 3001-3026, 2009.

[7] Bokanowski, O., A. Picarelli, and H. Zidani. Dynamic programming and error estimates for stochastic control problems with maximum cost. *Applied Mathematics & Optimization*, 71(1): 125-163, 2015.

[8] Bouchard, B. and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111(2), 175–206, 2004.

[9] Caffarelli, L.A. and P. E. Souganidis. A rate of convergence for monotone finite difference approximations to fully nonlinear uniformly elliptic PDEs. *Communications on Pure and Applied Mathematics*, 61(1):1-7, 2008.

[10] Chassagneux, J.F. and A. Richou. Numerical simulation of quadratic BSDEs. *The Annals of Applied Probability*, 26(1): 262–304, 2016.

[11] Chong, W.F., Y. Hu, G. Liang and T. Zariphopoulou. An ergodic BSDE approach to forward entropic risk measures: representation and large-maturity behavior. *Finance and Stochastics*, 23(1): 239–273, 2019.

[12] Delbaen, F., Y. Hu, and X. Bao. Backward SDEs with superquadratic growth. *Probability Theory and Related Fields*, 150(1), 145-192, 2011.

[13] El Karoui, N. and R. Rouge. Pricing via utility maximization and entropy. *Mathematical Finance* 10, 259–276. 2000.

[14] Evans, L. C. *Partial Differential Equations*. Graduate Studies in Mathematics, Americal Mathematical Society, 1998.

[15] Evans, L. C. and A. Friedman. Optimal stochastic switching and the Dirichlet problem for the Bellman equation. *Trans. Amer. Math. Soc*, 253:365-389, 1979.

[16] Fahim A., N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic PDEs. *The Annals of Applied Probability*, 21: 1322–1364, 2011.

[17] Giles, M. B. Multilevel monte carlo path simulation. *Operations Research*, 56(3), 607–617, 2008.

[18] Henderson, V. and G. Liang. Pseudo linear pricing rule for utility indifference valuation. *Finance and Stochastics*, 18(3):593–615, 2014.

[19] Henderson, V. and G. Liang. A multidimensional exponential utility indifference pricing model with applications to counterparty risk. *SIAM Journal on Control and Optimization*, 54(3): 690–717, 2016.

[20] Hu, Y., P. Imkeller and M. Müller. Utility maximization in incomplete markets. *The Annals of Applied Probability*, 15: 1691–1712, 2005.

[21] Huang, S. An approximation scheme for variational inequalities with convex and coercive Hamiltonians. *Working paper*, 2018, arXiv:1810.08842.

[22] Jakobsen, E. R. On the rate of convergence of approximation schemes for Bellman equations associated with optimal stopping time problems. *Math. Models Methods Appl. Sci.*, 13(5): 613-644, 2003.

[23] Kobylanski, M. Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*, 28: 558–602, 2000.

[24] Krylov, N. V. On the rate of convergence of finite-difference approximations for Bellman's equation. *St. Petersburg Math. J.*, 9(3):639-650, 1997.

[25] Krylov, N. V. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probability Theory and Related Fields*, 117(1): 1-16, 2000.

[26] Krylov, N. V. On the rate of convergence of finite-difference approximations for elliptic Isaacs equations in smooth domains. *Communications in Partial Differential Equations*, 40(8): 1393-1407, 2015.

[27] Lions, P. L. and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6): 964-979, 1979.

[28] Longstaff, F. A. and E. S. Schwartz. Valuing American options by simulation: a simple least-squares approach. *The Review of Financial Studies*, 14(1), 113-147, 2001.

[29] Lyons, T. and N. Victoir. Cubature on Wiener space. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 460(2041), 169–198, 2004.

[30] Marchuk, G. I. Some application of splitting-up methods to the solution of mathematical physics problems. *Apl. Mat.*, 13: 103-132, 1968.

[31] Nadtochiy, S. and T. Zariphopoulou. An approximation scheme for solution to the optimal investment problem in incomplete markets. *SIAM J. Finan. Math.*, 4(1): 494-538, 2013.

[32] Nelder, J. A. and R. Mead. A simplex method for function minimization. *Computer Journal*, 7: 308-313, 1965.

[33] Pham, H. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer, 2009.

[34] Picarelli, A., C. Reisinger, and J. Rotaetxe Arto. Error bounds for monotone schemes for parabolic Hamilton-Jacobi-Bellman equations in bounded domains. *Working paper*, 2017, arXiv:1710.11284.

[35] Tan X. A splitting method for fully nonlinear degenerate parabolic PDEs. *Electron. J. Probab.*, 18(15): 1-24, 2013.

[36] Tourin, A. Splitting methods for Hamilton-Jacobi equations, *Numerical Methods Partial Differential Equations*, 22: 381-396, 2006.

[37] Touzi, N. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*. Springer, 2012.

[38] Zariphopoulou, T., A solution approach to valuation with unhedgeable risks, *Finance and Stochastics*, 5(1): 61–82, 2001.

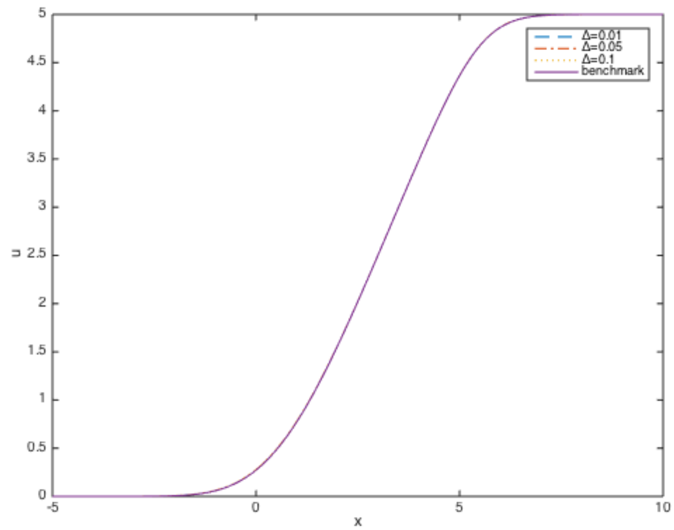

Fig. 1.1: Approximate values of $u(0,x)$ with various time steps $\Delta = 0.01/0.05/0.1$.

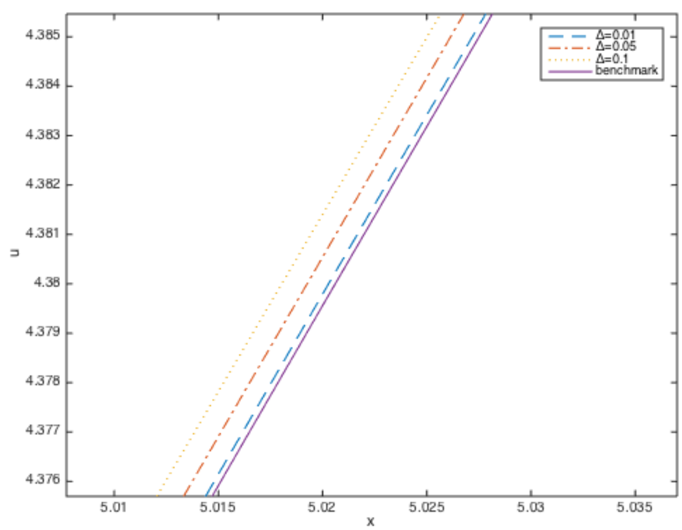

Fig. 1.2: Approximate values of $u(0,x)$ with various time steps $\Delta = 0.01/0.05/0.1$. The figure zooms in Fig. 1.1.

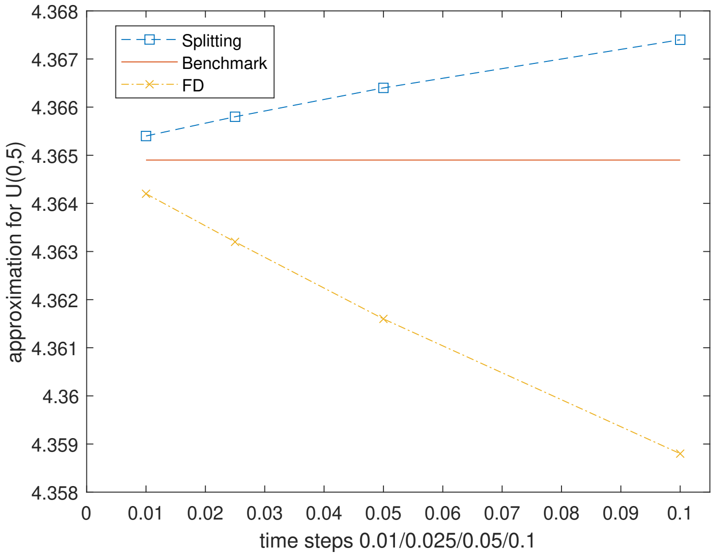

Fig. 1.3: Comparison of exact value and approximate values for $u(0,5)$ via the approximation scheme (2.12) and the Howard's FD scheme with various time steps $\Delta = 0.01/0.025/0.05/0.1$.

time steps / numerical schemes	0.01	0.25	0.05	0.1
splitting approx. value	4.3655	4.3658	4.3664	4.3674
approx. error	0.012%	0.02%	0.032%	0.056%
running time (in seconds)	18.78	1.07	0.16	0.04
FD approx. value	4.3642	4.3632	4.3616	4.3588
approx. error	0.016%	0.039%	0.076%	0.142%
running time (in seconds)	7.01	0.43	0.03	0.01

Table 1.1: Comparison of running errors and costs for approximating $u(0,5)$ via the approximation scheme (2.12) and the Howard's FD scheme with various time steps $\Delta = 0.01/0.025/0.05/0.1$.