A backward dual representation for the quantile hedging of Bermudan options

Bruno Bouchard*
CEREMADE, Université Paris Dauphine
and CREST-ENSAE
bouchard@ceremade.dauphine.fr

Géraldine Bouveret†
Department of Mathematics
Imperial College London
g.bouveret11@imperial.ac.uk

Jean-François Chassagneux
Department of Mathematics
Imperial College London
j.chassagneux@imperial.ac.uk

October 31, 2014

Abstract

Within a Markovian complete financial market, we consider the problem of hedging a Bermudan option with a given probability. Using stochastic target and duality arguments, we derive a backward algorithm for the Fenchel transform of the pricing function. This algorithm is similar to the usual American backward induction, except that it requires two additional Fenchel transformations at each exercise date. We provide numerical illustrations.

Keywords: stochastic target problems, quantile hedging, Bermudan options.

AMS 2010 Subject Classification Primary, 91G20, 91G60 ; Secondary, 60J60, 49L20.
JEL Classification G11, G13, C61.

1 Introduction

We study the problem of hedging a claim of Bermudan style with a given probability p. More precisely, we want to characterize the minimal initial value $v(\cdot, p)$ of an hedging

*Research supported by ANR Liquirisk and Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).
†Research supported by the Natixis Foundation for Quantitative Finance.

portfolio for which we can find a financial strategy such that, with a probability p, it remains above the exercise value of the Bermudan option at any possible exercise date.

This problem is referred to as *quantile hedging*, and it was popularized by Föllmer and Leukert [12, 13] . For claims of European type, they explained how the so-called *quantile hedging price* can be computed explicitly when the market is complete, by using duality arguments or the Neyman-Pearson lemma. A similar question was studied in Bouchard et al. [6] but in a Markovian setting. They showed that, even in incomplete markets and for general *loss functions*, one can characterize the pricing function as the solution of a non-linear parabolic second order differential equation, by using tools developed in the context of stochastic target problems by Soner and Touzi [16]-[17]. When the market is complete, they also observed that taking a Legendre-Fenchel transform in the equation reduces the computation of the price to the resolution of a linear parabolic second order differential equation, which can be solved *explicitly* by using the Feynman-Kac formula.

As far as super-hedging is concerned, the pricing of a Bermudan option reduces to a backward sequence of pricing problems for European claims. It is therefore natural to ask whether a similar result holds for the quantile hedging price, and whether one can extend the closed-form solutions of [12] and [6] to Bermudan options.

This paper answers to the positive. Namely, we provide a backward induction algorithm for the Fenchel transform w of the quantile hedging price $v(\cdot, p)$, with respect to the parameter p which prescribes the probability of hedging, see (2.21) and Theorem 2.1. The algorithm (2.21) is in a sense very similar to the one used for the pricing of Bermudan options. However it is written on the Fenchel transform w, rather than v, and it involves two additional Fenchel transformations at each exercise date.

To derive this, we first build on the original idea of [6] which consists in increasing the state space in order to reduce to a stochastic target problem of American type, as studied in Bouchard and Vu [8]. We then follow a very different route. Instead of appealing to stochastic target technics, we derive from this formulation a first dynamic programming algorithm for v, see Proposition 2.3, which relates to a series of optimal control of martingale problems. This is in the spirit of Bouchard et al. [5]. This dynamic programming principle suggests a backward algorithm for the computation of the Fenchel transform. It is defined in (2.21). We analyze it in details in Section 3.2: the main difficulty consists in controlling the propagation of the differentiability and the growth properties of the corresponding value function, backward in time. Then, as in [6, 12], a martingale representation argument allows us to show, by backward induction, that the algorithms in (2.21) and Proposition 2.3 provides the Fenchel transform of one another.

Before concluding this introduction, we would like to point out that a similar problem has been studied recently by Jiao et al. [14] in the form of general *lookback-style contraints*. They provide an alternative formulation in terms of an optimal control of martingales problem. This has to be compared with [5] and our Proposition 2.3. No Markovian structure is required, but they do not provide an explicit algorithm as we do. Moreover, the smoothness conditions they impose on their loss functions are not satisfied in the quantile hedging case. They also study the case of several constraints in expectation set (independently) at the different exercise times, which is close to the P&L matching problems of Bouchard and Vu [7].

Finally, in this paper, we focus on the quantile hedging problem for sake of simplicity. It is an archetype of an irregular loss function, and it should be clear that a similar analysis can be carried out for a wide class of (more regular) loss functions. Also note that we only use probabilistic arguments, as opposed to PDE technics as in [6], which opens the door to the study of more general non Markovian settings. We leave this for future research.

Notations: Let d be a positive integer. Any vector x of $\mathbb{R}^d$ is seen as a column vector. Its norm and transpose are denoted by $|x|$ and $x^\top$. We set $\mathbb{M}^d := \mathbb{R}^{d\times d}$ and denote by $M^\top$ the transpose of $M \in \mathbb{M}^d$, while $\mathrm{Tr}\,[M]$ is its trace. For ease of notations, we set $\mathcal{O}^d_+ := (0,\infty)^d$.

We fix a finite time horizon $T > 0$. Let $\psi : (t,x,p) \in [0,T] \times \mathcal{O}^d_+ \times \mathbb{R} \mapsto \psi(t,x,p)$. If it is smooth enough, we denote by $\partial_t\psi$ and $\partial_p\psi$ its derivative with respect t and p, and by $\partial_x\psi$ its Jacobian matrix with respect to x, as a row vector. The Hessian with respect to x is $\partial^2_{xx}\psi$, $\partial^2_{pp}\psi$ is the second order derivative with respect to p, and $\partial^2_{xp}\psi$ is the vector of cross second order derivatives. We denote by $\psi^\sharp$ its Fenchel transform with respect to the last argument,

$$\psi^\sharp(t,x,q) := \sup_{p\in\mathbb{R}} \left(pq - \psi(t,x,p)\right) , \tag{1.1}$$

and define

$\mathrm{co}[\psi]$, the closed convex envelope of ψ with respect to its last argument.

If ψ is convex with respect to its last variable, we denote by $D^+_p\psi$ and $D^-_p\psi$ its corresponding right- and left-derivatives. We refer to [15] for the various notions related to convex analysis.

We fix a complete probability space $(\Omega,\mathbb{F},\mathbb{P})$ supporting a d-dimensional Brownian motion W. We denote by $\mathbb{F} = (\mathcal{F}_t)_{0\leqslant t\leqslant T}$ the usual augmented Brownian filtration. All over the paper, inequalities between random variables have to be understood in the $\mathbb{P}$-a.s. sense.

2 Problem formulation and main results

2.1 Financial market and hedging problem

Our financial market consists in a non-risky asset, whose price process is normalized to unity, and d risky assets $X = (X^1, ..., X^d)$ whose dynamics are given by

$$X_s^{t,x} = x + \int_t^s \mu(r, X_r^{t,x})\mathrm{d}r + \int_t^s \sigma(r, X_r^{t,x})\mathrm{d}W_r\,, \tag{2.1}$$

given the initial data $(t,x) \in [0,T] \times \mathcal{O}_+^d$. To ensure that the above is well-defined, we assume that

$$\mu : [0,T] \times \mathcal{O}_+^d \to \mathbb{R}^d \text{ and } \sigma : [0,T] \times \mathcal{O}_+^d \to \mathbb{M}^d \text{ are Lipschitz continuous}\,, \tag{2.2}$$

and that the unique strong solution to (2.1) takes its values in $\mathcal{O}_+^d$ when the initial data lies in $\mathcal{O}_+^d$.

In order to enforce the absence of arbitrage and the completeness of the financial market, we also impose that

$$\begin{aligned} &\sigma \text{ is invertible}\,, \quad \lambda := \sigma^{-1}\mu \text{ is bounded} \\ &\text{and Lipschitz continuous in space}\,. \end{aligned} \tag{2.3}$$

The Lipschitz continuity condition is not required to define the risk neutral measure[1]

$$\mathbb{Q}_{t,x} := \frac{1}{Q_T^{t,x,1}} \cdot \mathbb{P} \;\text{ with }\; \frac{1}{Q^{t,x,q}} := \frac{1}{q}\mathcal{E}\left(-\int_t^\cdot \lambda(s, X_s^{t,x})^\top \mathrm{d}W_s\right)\,,\; q > 0\,, \tag{2.4}$$

but will be used in some of our forthcoming arguments.

In this model, an admissible financial strategy is a d-dimensional predictable process ν such that

$$\mathbb{E}^{\mathbb{Q}_{t,x}}\left[\int_t^T |\nu_s^\top \sigma(s, X_s^{t,x})|^2 ds\right] < \infty\,, \tag{2.5}$$

and the corresponding wealth process remains non-negative

$$Y^{t,x,y,\nu} := y + \int_t^\cdot \nu_r^\top \mathrm{d}X_r^{t,x} \geqslant 0\,, \quad \text{on } [t,T]\,,$$

[1] $\mathcal{E}$ denotes here the Doléans-Dade exponential.

given the initial data (t, x) of the market and the initial dotation $y \geqslant 0$. We denote by $\mathcal{U}_{t,x,y}$ the collection of admissible financial strategies. As usual, each ν_t^i should be interpreted as the number of units of asset i in the portfolio at time t.
We now fix a finite collection of times

$$\mathbb{T}_t := \{t_0 = 0 \leqslant \cdots \leqslant t_i \leqslant \cdots \leqslant t_n = T\} \cap (t, T] ,$$

together with non-negative payoff functions

$$x \in \mathcal{O}_+^d \mapsto g(t_i, x), \text{ Lipschitz continuous for all } i \leqslant n . \tag{2.6}$$

The quantile hedging problem consists in finding the minimal initial wealth $v(t, x, p)$ which ensures that the stream of Bermudan payoffs $\{g(s, X_s^{t,x}),\ s \in \mathbb{T}_t\}$ can be hedged with a given probability p

$$v(t, x, p) := \inf \Gamma(t, x, p) , \tag{2.7}$$

where

$$\Gamma(t, x, p) := \left\{ y \geqslant 0 \ : \ \exists\, \nu \in \mathcal{U}_{t,x,y} \text{ s.t. } \mathbb{P}\left[\bigcap_{s \in \mathbb{T}_t} \mathrm{S}_s^{t,x,y,\nu}\right] \geqslant p \right\} ,$$
$$\text{with } \mathrm{S}_s^{t,x,y,\nu} := \Omega \mathbf{1}_{\{t \geqslant s\}} + \mathbf{1}_{\{t<s\}} \{Y_s^{t,x,y,\nu} \geqslant g(s, X_s^{t,x})\} .$$

Observe that $v(t, \cdot)$ must be interpreted as a continuation value, i.e. the price at time t knowing that the option has not been exercised on $[0, t]$. In particular, $v(T, \cdot) = 0$. For $p = 1$, $v(t, \cdot, 1)$ coincides with the continuation value of the super-hedging price of the Bermudan option. In this complete market, it satisfies the usual dynamic programming principle

$$v(t, x, 1) \quad = \quad \mathbb{E}^{\mathbb{Q}_{t,x}}[(v \vee g)(t_{i+1}, X_{t_{i+1}}^{t,x}, 1)] , \quad \text{for } t \in [t_i, t_{i+1}) ,\ i < n . \tag{2.8}$$

Above and in the following, we use the notation

$$g(t, x, p) := g(t, x) \mathbf{1}_{\{0<p\leqslant 1\}} + \infty \mathbf{1}_{\{p>1\}} , \quad \text{for } p \in \mathbb{R} .$$

Note that Γ can also be formulated in terms of stopping times, see the Appendix for the proof.

Proposition 2.1. *For $(t, x, p) \in [0, T] \times \mathcal{O}_+^d \times [0, 1]$,*

$$\begin{aligned} \Gamma(t, x, p) \quad &= \quad \{y \geqslant 0 \ : \ \exists\, \nu \in \mathcal{U}_{t,x,y} \ \text{ s.t. } \mathbb{P}[\mathrm{S}_\tau^{t,x,y,\nu}] \geqslant p,\ \forall\ \tau \in \mathcal{T}_t\} \\ &= \quad \{y \geqslant 0 \ : \ \exists\, \nu \in \mathcal{U}_{t,x,y} \ \text{ s.t. } \mathbb{P}[\mathrm{S}_{\hat{\tau}_\nu}^{t,x,y,\nu}] \geqslant p\} \mathbf{1}_{\{t<T\}} + \mathbb{R}_+ \mathbf{1}_{\{t=T\}} , \end{aligned} \tag{2.9}$$

in which $\mathcal{T}_t$ is the set of stopping times with values in $\mathbb{T}_t$, and $\hat{\tau}_\nu := \min\{s \in \mathbb{T}_t \ : \ Y_s^{t,x,y,\nu} < g(s, X_s^{t,x})\} \wedge T$.

Remark 2.1. The function $p \mapsto v(\cdot, p)$ is non-decreasing. It takes the value 0 if $p \leqslant p_{\min}(t,x)$ where

$$p_{\min}(t,x) := \mathbb{P}[g(s, X_s^{t,x})\mathbf{1}_{\{s<T\}} = 0 \text{ for all } s \in \mathbb{T}_t] \,. \tag{2.10}$$

To avoid trivial statements, we assume that $p_{\min}(t,\cdot) < 1$, for $t < T$, which implies

$$v(t,x,1) > 0\,, \text{ for } t < T\,. \tag{2.11}$$

Moreover, it follows from (2.6) that we can find $C > 0$ such that $g(s,x) \leqslant C(1+\sum_{i=1}^d x^i)$, for $x \in \mathcal{O}_+^d$, $s \in \mathbb{T}_0$. This implies that we can restrict to strategies ν such that

$$0 \leqslant Y^{t,x,y,\nu} \leqslant C(1 + |X^{t,x}|)\,, \tag{2.12}$$

by possibly adopting a buy-and-hold strategy after the first time when the wealth process hits the right-hand side term, recall that $X^{t,x}$ has positive components. In particular,

$$0 \leqslant v(t,x,p) \leqslant C(1+|x|)\,. \tag{2.13}$$

2.2 Equivalent formulation as a stochastic target problem

The first step in our analysis consists in reducing the problem to a stochastic target problem of American type as studied in [8]. As in [6], we first increase the dimension of the controlled process by introducing the family of martingales

$$P^{t,p,\alpha} := p + \int_t^{\cdot} \alpha_s^\top \mathrm{d}W_s\,,$$

where α is a square integrable predictable process. The process $P^{t,p,\alpha}$ will be later on interpreted as the conditional probability of success. It is therefore natural to restrict to the class of controls such that

$$P^{t,p,\alpha} \in [0,1]\,, \quad \text{on} \quad [t,T]\,.$$

We denote by $\mathcal{A}_{t,p}$ the set of predictable square integrable processes such that the above holds, and set $\hat{\mathcal{U}}_{t,x,y,p} := \mathcal{U}_{t,x,y} \times \mathcal{A}_{t,p}$.

Proposition 2.2. *Fix* $(t,x,p) \in [0,T] \times \mathcal{O}_+^d \times [0,1]$, *then*

$$\Gamma(t,x,p) = \Big\{ y \geqslant 0 : \exists\ (\nu,\alpha) \in \hat{\mathcal{U}}_{t,x,y,p} \ \ s.t.\ \ Y^{t,x,y,\nu} \geqslant g(\cdot, X^{t,x}, P^{t,p,\alpha}) \ \ on\ \mathbb{T}_t \Big\}\,. \tag{2.14}$$

Proof. At time T both sets are $\mathbb{R}_+$ by definition of $\mathbb{T}_T$. We now fix $t < T$. Let $\bar{\Gamma}(t,x,p)$ denote the right-hand side in (2.14) and let y be one of his elements. Fix $(\nu,\alpha) \in \hat{\mathcal{U}}_{t,x,y,p}$ such that $Y^{t,x,y,\nu} \geqslant g(\cdot, X^{t,x}, P^{t,p,\alpha})$ on $\mathbb{T}_t$. Then, $\mathrm{S}^{t,x,y,\nu} \supset \{P^{t,p,\alpha} > 0\}$ on $\mathbb{T}_t$. Since $P^{t,p,\alpha} \in [0,1]$ and therefore $\mathbf{1}_{\{P^{t,p,\alpha}>0\}} \geqslant P^{t,p,\alpha}$, this implies

$$\mathbb{P}\left[\bigcap_{s\in\mathbb{T}_t} \mathrm{S}_s^{t,x,y,\nu}\right] \;\geqslant\; \mathbb{P}\left[\bigcap_{s\in\mathbb{T}_t} \{P_s^{t,p,\alpha} > 0\}\right] \geqslant \mathbb{E}\left[P_T^{t,p,\alpha} \prod_{s\in\mathbb{T}_t\setminus\{T\}} \mathbf{1}_{\{P_s^{t,p,\alpha}>0\}}\right].$$

The process $P^{t,p,\alpha}$ being a martingale, for $s \in \mathbb{T}_t$, $\{P_s^{t,p,\alpha} = 0\} \subset \{P_T^{t,p,\alpha} = 0\}$. Hence

$$\mathbb{P}\left[\bigcap_{s\in\mathbb{T}_t} \mathrm{S}_s^{t,x,y,\nu}\right] \;\geqslant\; \mathbb{E}\left[P_T^{t,p,\alpha}\right] = p\,.$$

Therefore, $y \in \Gamma(t,x,p)$ and this argument proves that $\bar{\Gamma}(t,x,p) \subset \Gamma(t,x,p)$.
We now fix $y \in \Gamma(t,x,p)$ and choose $\nu \in \mathcal{U}_{t,x,p}$ such that $p' := \mathbb{P}\left[\bigcap_{s\in\mathbb{T}_t} \mathrm{S}_s^{t,x,y,\nu}\right] \geqslant p$. By the martingale representation theorem, we can find $\alpha \in \mathcal{A}_{t,p'}$ such that

$$\mathbf{1}_{\bigcap_{s\in\mathbb{T}_t} \mathrm{S}_s^{t,x,y,\nu}} = P_T^{t,p',\alpha} \geqslant P_T^{t,p,\alpha}\,.$$

By possibly replacing α by the constant process 0 after the first time after t at which $P^{t,p,\alpha}$ reaches the level 0, we can assume that $\alpha \in \mathcal{A}_{t,p}$. Moreover, the above implies

$$\mathbf{1}_{\mathrm{S}_s^{t,x,y,\nu}} \geqslant P_T^{t,p,\alpha}\,,\; s \in \mathbb{T}_t\,,$$

which by taking conditional expectation and using the fact that $P^{t,p,\alpha}$ is a martingale leads to $\mathbf{1}_{\mathrm{S}^{t,x,y,\nu}} \geqslant P^{t,p,\alpha}$ on $\mathbb{T}_t$. The latter is equivalent to $Y^{t,x,y,\nu} \geqslant g(\cdot, X^{t,x}, P^{t,p,\alpha})$ on $\mathbb{T}_t$. Hence, $y \in \bar{\Gamma}(t,x,p)$. □

2.3 Dynamic programming and dual backward algorithm

With the formulation obtained in Proposition 2.2 at hand, one can now derive a first dynamic programming algorithm. Its proof is postponed to the Appendix.

Proposition 2.3. *Fix* $0 \leqslant i \leqslant n-1$ *and* $(t,x,p) \in [t_i, t_{i+1}) \times \mathcal{O}_+^d \times [0,1]$,

$$v(t,x,p) = \inf_{\alpha\in\mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right)\right]. \tag{2.15}$$

As a consequence, there exists $C > 0$ *such that*

$$|v(t,x,p) - v(t,x',p)| \leqslant C(1 + |x| + |x'|)|x - x'|\,, \tag{2.16}$$

for all $(t,p) \in [0,T] \times [0,1]$ *and* $x, x' \in \mathcal{O}_+^d$.

Remark 2.2. We shall see in Section 3 that $(v \vee g)$ can be replaced by its convex envelope with respect to p in (2.15). This phenomenon was already observed in [5] and [6].

Note that this provides a first way to compute the value function v. Indeed, standard arguments should lead to a characterization of v as a viscosity solution on each interval $[t_i, t_{i+1})$, $i < n$ of[2]

$$\sup_{a \in \mathbb{R}^d} \left\{ -\partial_t \varphi + a^\top \lambda \partial_p \varphi - \frac{1}{2} \left(\mathrm{Tr}\left[\sigma\sigma^\top \partial^2_{xx}\varphi\right] + 2\mathrm{Tr}\left[a^\top \sigma^\top \partial^2_{xp}\varphi\right] + |a|^2 \partial^2_{pp}\varphi \right) \right\} = 0\,, \tag{2.17}$$

with the boundary condition

$$v(t_{i+1}-, \cdot) = (v \vee g)(t_{i+1}, \cdot)\,. \tag{2.18}$$

However, the fact that the control $a \in \mathbb{R}^d$ in the above is not bounded renders the use of numerical schemes delicate in practice.
This can actually be simplified by considering the Fenchel transform $v^\sharp$ of v, see (1.1) in the notations section.
Indeed, as already observed in [6] in the case $n = 1$, a formal change of variable argument in (2.17) suggests that the dual function $v^\sharp$ should be a viscosity solution of the linear partial differential equation

$$-\partial_t \varphi - \frac{1}{2} \left(\mathrm{Tr}[\sigma\sigma^\top \partial^2_{xx}\varphi] + 2q\mathrm{Tr}\left[\lambda^\top \sigma^\top \partial^2_{xq}\varphi\right] + |\lambda|^2 q^2 \partial^2_{qq}\varphi \right) = 0\,, \tag{2.19}$$

on the different time steps, with the boundary conditions obtained by taking the Fenchel transform in (2.18):

$$v^\sharp(t_{i+1}-, \cdot) = (v \vee g)^\sharp(t_{i+1}, \cdot)\,. \tag{2.20}$$

By the Feynman-Kac representation this corresponds to the following backward algorithm

$$\left\{ \begin{array}{rcl} w(T,x,q) & := & q + \infty \mathbf{1}_{\{q<0\}}\,, \\ w(t,x,q) & := & \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(w^\sharp \vee g)^\sharp (t_{i+1}, X^{t,x}_{t_{i+1}}, Q^{t,x,q}_{t_{i+1}}) \right]\,, \quad \text{for} \quad t \in [t_i, t_{i+1})\,,\ i < n\,, \end{array} \right. \tag{2.21}$$

in which $Q^{t,x,q}$ is defined in (2.4).
The main result of this paper shows that this algorithm actually allows to compute the value function v.

Theorem 2.1. *$v = w^\sharp$ on $[0,T] \times \mathcal{O}^d_+ \times [0,1]$.*

[2] A precise statement would require a relaxation of the operator, see [6].

The proof of this result is the object of the subsequent sections. Although it is in the spirit of [6], our proof is different and more involved. The main difficulty comes from the induction. At each time step, we have to verify that $(w^\sharp \vee g)$ behaves in a sufficiently nice way. In the one step case, [6] has only to consider the terminal payoff g. Moreover, we only use probabilistic arguments as opposed to PDE arguments.

Clearly, the algorithm (2.21) provides a way to compute the value function easily. One can for instance use the fact that $w = v^\sharp$ is the unique viscosity solution (2.19) with the boundary conditions (2.20). Let us make this statement more precise.

Definition 2.1. We say that a lower-semicontinuous function u is a viscosity super-solution of the system $(\mathcal{S})$ if, on each $[t_i, t_{i+1}) \times \mathcal{O}_+^d \times (0,\infty)$, $i < n$, it is a viscosity super-solution of (2.19) with the boundary conditions

$$\begin{aligned}
\liminf_{t'\uparrow t_i, (x',q')\to(x,q)} u(t',x',q') &\geqslant (u^\sharp \vee g)^\sharp(t_i,x,q) \;\text{ for }\; (x,q) \in \mathcal{O}_+^d \times (0,\infty)\,,\; i<n\,,\\
\liminf_{t'\uparrow T, (x',q')\to(x,q)} u(t',x',q') &\geqslant g^\sharp(T,x,q) \;\text{ for }\; (x,q) \in \mathcal{O}_+^d \times (0,\infty)\,.
\end{aligned}$$

We define accordingly the notion of sub-solution for upper-semicontinuous functions. A function is a viscosity solution if its lower- (resp. upper-) semicontinuous envelope is a viscosity super- (resp. sub-) solution.

Note that in the above definition we have to understand u as being $+\infty$ on $[0,T] \times \mathcal{O}_+^d \times (-\infty, 0)$ to compute the Fenchel transforms involved in the time boundary conditions.

We now provide a version of the comparison principle for $(\mathcal{S})$ which pertains for the usual extensions of the Black and Scholes model. The assumptions used below are here to avoid the boundary of $\mathcal{O}_+^d$ - when this is not the case, one has to specify additional boundary conditions.

Proposition 2.4. *The function w is continuous on $[0,T) \times \mathcal{O}_+^d \times (0,\infty)$, non-negative, has a linear growth in its x and q variable and is a viscosity solution of $(\mathcal{S})$. Moreover, if there exists two functions $\bar\sigma$ and $\bar\mu$ such that $\sigma(\cdot,x) = \mathrm{diag}[x]\bar\sigma(\cdot,x)$ and $\mu(\cdot,x) = \mathrm{diag}[x]\bar\mu(\cdot,x)$, then $u_1 \geqslant u_2$ on $[0,T) \times \mathcal{O}_+^d \times (0,\infty)$ whenever u_1 and u_2 are respectively a super- and a subsolution of $(\mathcal{S})$, which are non-negative and have a polynomial growth in their x and q variables on $[0,T) \times \mathcal{O}_+^d \times (0,\infty)$.*

The proof is postponed to the Appendix. Given the latter, it is not difficult to follow the arguments of [3] to construct a convergent finite difference scheme for the resolution of $(\mathcal{S})$. Alternatively, one could also use quantization methods to tackle the approximation of w, see [1, 2], or a regression based Monte-Carlo method, see the survey paper [9] and the references therein.

2.4 Examples of application

In this section, we present two examples of application. The numerical results are obtained using the following procedure which is based on the above algorithm to compute $w = v^\sharp$: for $i \leqslant n-1$,

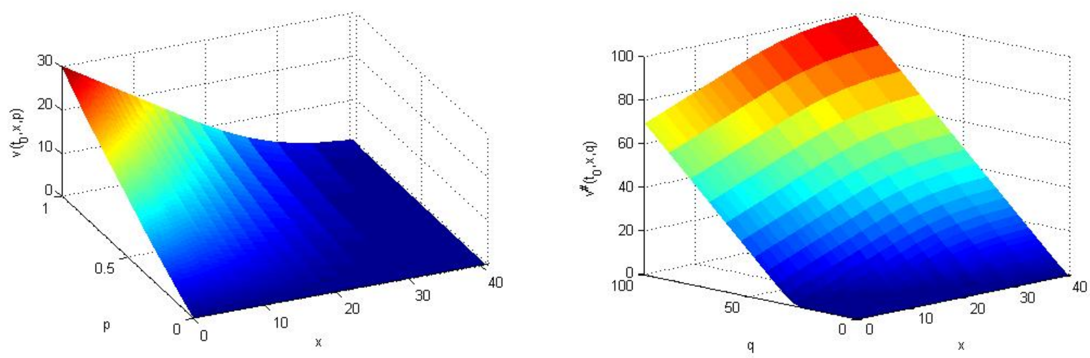
1) Compute the value of $(w^\sharp \vee g)^\sharp(t_{i+1}, \cdot)$ by approximating the Fenchel-Legendre transform numerically.

2) Solve the PDE (2.19)-(2.20) for w, using e.g. finite difference methods, on $[t_i, t_{i+1}] \times \mathcal{O}_+^d \times \mathbb{R}_+$.

We now fix $T = 1$ and $\mathbb{T}_t := \{t_0 = 0,\, t_1 = \frac{1}{3},\, t_2 = \frac{2}{3},\, t_3 = 1\} \cap (t, t_3]$, $t \in [0, T]$. We work in a Black-Scholes setting with market parameters: $d = 1$, $\sigma(t,x) = 0.25x$, $\lambda(t,x) = 0.2$.

For our first numerical application, we consider a put option, i.e. $g(t,x) = [K - x]^+$, with strike $K = 30$.

In figure 1, we plot the functions v and $v^\sharp$ at $t = t_0$. In figure 2(a-b-c), we plot for different values of x the function v and $\text{co}[v \vee g]$. This shows the rather complicated behaviour of the transformation $v \mapsto \text{co}[v \vee g]$, as predicted by Proposition 3.3 (b) below. With the notation of this proposition, figure 2(a) corresponds to the case A_1, figure 2(b) corresponds to the case A_3 and figure 2(c) corresponds to the case A_2. Because of the interest rate being set to 0 and the payoff being convex, we always have $v(t,x,1) \geqslant g(t,x)$. Figure 2(d) shows the decrease of value for v, when p decreases.



(a) (b)

Figure 1: Surface of $v(t,x,p)$ and $v^\sharp(t,x,q)$ at $t = t_0$.

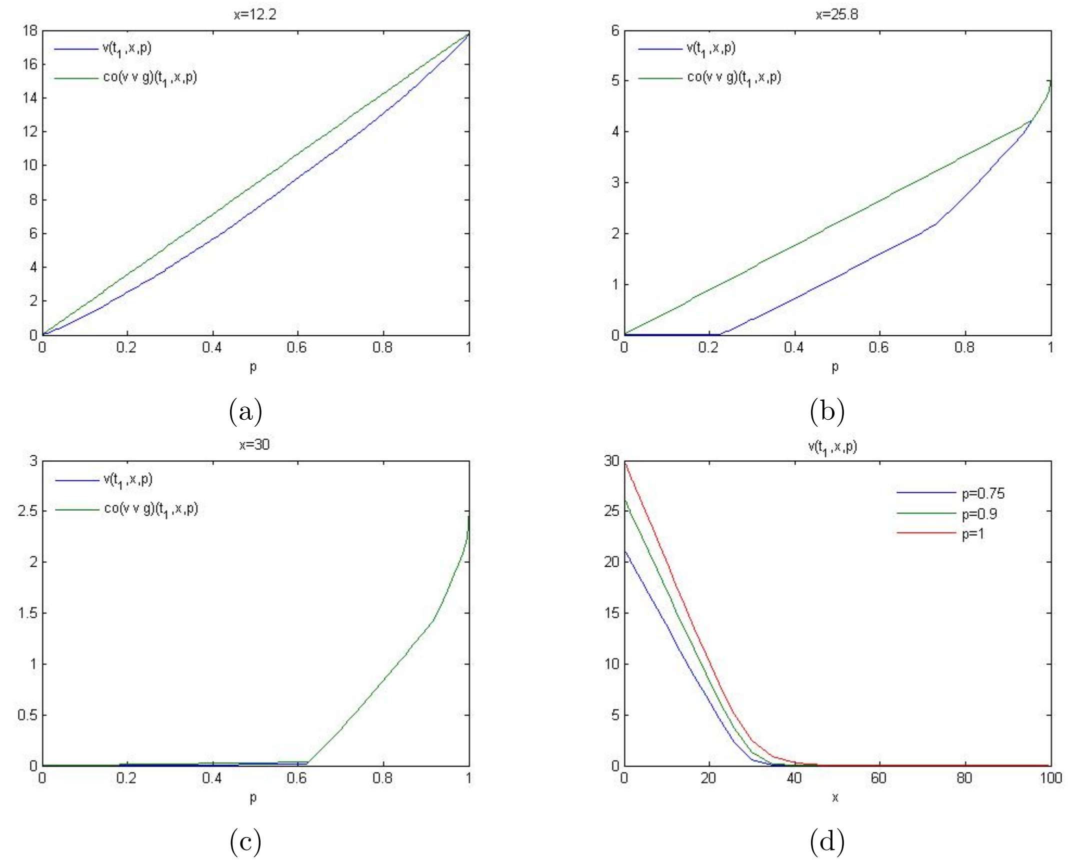
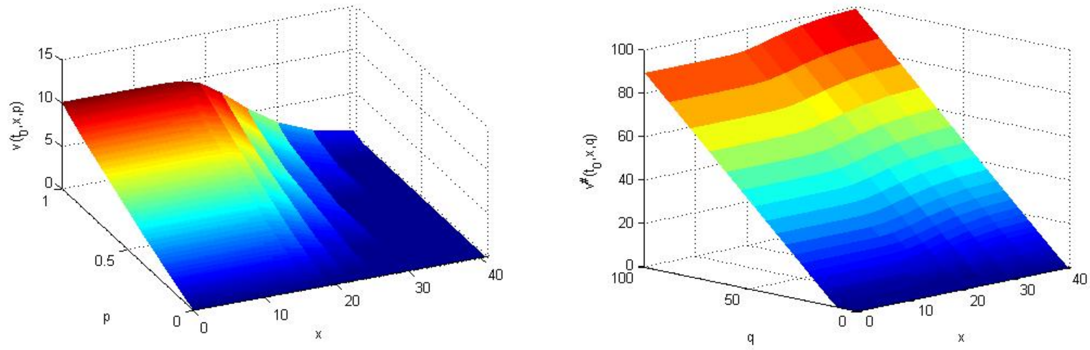


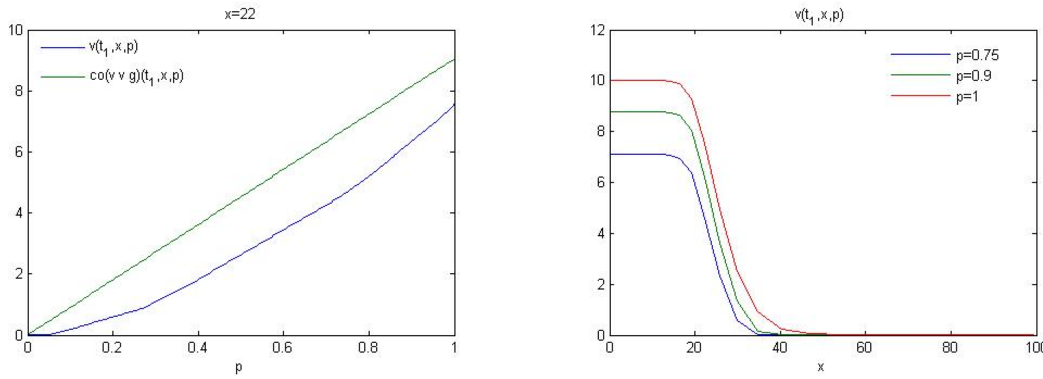
Figure 2: (a)-(c): plots of $v(t,x,\cdot)$ and $\mathrm{co}[v \vee g](t,x,\cdot)$ at $t = t_1$ and for different values of x. (d): plot of $v(t,\cdot,p)$ at $t = t_1$ and for different values of p.

In our second example, we consider a put spread option with strikes 20 and 30, i.e. $g(t,x) = [30-x]^+ - [20-x]^+$. The numerical results are displayed in Figure 3 and 4. It may happen here that $v(t,x,1) < g(t,x)$, see figure 4(a).



(a) (b)

Figure 3: Surface of $v(t,x,p)$ and $v^\sharp(t,x,q)$ at $t=t_0$.



(a) (b)

Figure 4: (a): plot of $v(t,x,\cdot)$ and $\text{co}[v \vee g](t,x,\cdot)$ at $t=t_1$ and for $x=22$. (b): plot of $v(t,.,p)$ at $t=t_1$ and for different values of p.

We conclude this section with the following remark on the behaviour of v near $p=1$.

Remark 2.3. (a) We know from the identification $v=w^\sharp$ and Proposition 3.2 (b) that $p \mapsto v(t,x,p)$ is convex and continuous on $[0,1]$.
(b) Nothing prevents $D_p^- v(\cdot,1)$ to be equal to $+\infty$. This can be checked by direct calculation in the European case and the Black-Scholes setting using the explicit formula [12, Equation (3.15)].

3 Proof of the backward dual representation

From now on, we extend v to $[0,T] \times \mathcal{O}_+^d \times \mathbb{R}$ by setting

$$v(\cdot, p) = 0 \text{ if } p < 0 \text{ and } v(\cdot, p) = +\infty \text{ if } p > 1\,. \tag{3.1}$$

Using the convention $\inf \varnothing = +\infty$, this extension is consistent with (2.7).

3.1 The backward algorithm as a lower bound

We first show that the backward algorithm (2.21) actually provides a lower bound for the value function v.

Proposition 3.1. *$v \geqslant w^\sharp$ on $[0,T] \times \mathcal{O}_+^d \times [0,1]$.*

Proof. First note that $v(T,\cdot) = 0 = w^\sharp(T,\cdot)$, by definition. Thus, $(v \vee g)(T,\cdot) = (w^\sharp \vee g)(T,\cdot)$. We now assume that $v \geqslant w^\sharp$ on $[t_{i+1}, T] \times \mathcal{O}_+^d \times [0,1]$ for some $i \leqslant n-1$. Then, $(v \vee g)^\sharp(t_{i+1},\cdot) \leqslant (w^\sharp \vee g)^\sharp(t_{i+1},\cdot)$ and therefore

$$\begin{aligned}(v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right) &\geqslant P_{t_{i+1}}^{t,p,\alpha} q Q_{t_{i+1}}^{t,x,1} - (v \vee g)^\sharp\left(t_{i+1}, X_{t_{i+1}}^{t,x}, q Q_{t_{i+1}}^{t,x,1}\right)\\ &\geqslant P_{t_{i+1}}^{t,p,\alpha} q Q_{t_{i+1}}^{t,x,1} - (w^\sharp \vee g)^\sharp\left(t_{i+1}, X_{t_{i+1}}^{t,x}, q Q_{t_{i+1}}^{t,x,1}\right)\,.\end{aligned}$$

Fix $t \in [t_i, t_{i+1})$. Taking the expectation on both sides and recalling (2.21), we obtain

$$\mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right)\right] \geqslant pq - w(t,x,q)\,.$$

Taking first the supremum over $q \in \mathbb{R}$ in the right-hand side and then the infimum over $\alpha \in \mathcal{A}_{t,p}$ in the left-hand side, we get from Proposition 2.3 that $v(t,x,p) \geqslant w^\sharp(t,x,p)$. □

3.2 Representation and differentiability of the backward dual algorithm

This section is devoted to the study of the function $(w^\sharp \vee g)^\sharp$ which appears in the dual algorithm (2.21) and of its Fenchel transform $(w^\sharp \vee g)^{\sharp\sharp}$. We first provide a decomposition in simple terms in Proposition 3.3. They only contain w, g and auxiliary functions that are easy to handle, see (3.3)-(3.4) below. In view of (2.21), this will then allow us to study the subdifferential of $w(t_i,\cdot)$ in terms of the subdifferential of $w(t_{i+1},\cdot)$. This analysis is reported in Lemma 3.2. These results will be of important use in the final proof of Theorem 2.1 as it will require to find a particular value p in the subdifferential of $w(t_i,\cdot)$ and then to apply a martingale representation argument between elements of the subdifferential of $(w^\sharp \vee g)^\sharp$ at t_{i+1} and p at t_i, see the proof of Theorem 3.1.

We start with properties that stem directly from the definition of w and standard results in convex analysis. The proof is postponed to the Appendix.

Proposition 3.2. *For all $(t,x) \in [0,T] \times \mathcal{O}_+^d$:*
(a) *The functions $q \in \mathbb{R} \mapsto w(t,x,q)$ is a proper convex non-decreasing and non-negative function. Moreover, $w(\cdot,0) = 0$ and $w(\cdot,q) = \infty$ for $q < 0$.*
(b) *The function $p \in \mathbb{R} \mapsto w^\sharp(t,x,p)$ and $q \in \mathbb{R} \mapsto (w^\sharp \vee g)^\sharp(t,x,q)$ are convex, non-negative, non-decreasing and continuous on their domains. Finally, $w^\sharp(\cdot,0) = 0$, $(w^\sharp \vee g)^\sharp(\cdot,0) = 0$ and $(w^\sharp \vee g)^\sharp(\cdot,q) = +\infty$ for $q < 0$.*

The next result is key to get the representation of $(w^\sharp \vee g)^\sharp$ and $(w^\sharp \vee g)^{\sharp\sharp}$. Recall that $g(t,x,p) = g(t,x)\mathbf{1}_{\{0<p\leqslant 1\}} + \infty\mathbf{1}_{\{p>1\}}$.

Lemma 3.1. *Let $p_1 \geqslant 0$ and f be a non-decreasing convex function such that $f(0) = 0$, $f \geqslant g(t,x,\cdot)$ on $[p_1,\infty)$, $f \leqslant g(t,x,\cdot)$ on $(-\infty,p_1]$.*
(a) *The convex envelope of $f \vee g$ is given by*

$$(f \vee g)^{\sharp\sharp}(p) = \mathrm{co}[f \vee g](p) = pq_1\mathbf{1}_{\{0\leqslant p<p_1\}} + f(p)\mathbf{1}_{\{p_1\leqslant p\leqslant 1\}} + \infty\mathbf{1}_{(1,\infty)}\,,$$

with $q_1 = g(t,x)/p_1\mathbf{1}_{\{p_1>0\}}$.
(b) *Moreover, we have*

$$(f \vee g)^\sharp(\cdot,q) = p_1[q-q_1]^+\mathbf{1}_{\{q\leqslant D_p^+ f(p_1)\}} + f^\sharp(q)\mathbf{1}_{\{q>D_p^+ f(p_1)\}}\,,\ q \geqslant 0\,,$$

which is a closed proper convex function. In particular, it is continuous at $D_p^+ f(p_1)$ when $0 < D_p^+ f(p_1) < +\infty$.

Proof.
1. The left-hand side identity in (a) follows from see [15, Theorem 12.2]. We set $\varphi : p \mapsto pq_1\mathbf{1}_{\{p>0\}} \vee f(p)$, which is convex. By assumption, we already know that $f(p) \leqslant g(t,x,p) = 0$ for $p \leqslant 0$. Since $f(0) = 0$ and $f(p_1) = g(t,x)$, we have by convexity that $f(p) \leqslant pq_1$, $p \in [0,p_1]$, which implies $\varphi(p)\mathbf{1}_{\{p\leqslant p_1\}} = pq_1\mathbf{1}_{\{0\leqslant p<p_1\}}$, for $p \leqslant p_1$. Since $f(p) \leqslant pq_1$ for $p \in [0,p_1]$ and $f(p_1) = p_1q_1$, we compute that $D_p^- f(p_1) \geqslant q_1$. By convexity, we also have $f(p) \geqslant f(p_1) + D_p^- f(p_1)(p-p_1) \geqslant pq_1$ for $p \geqslant p_1$ and then $\varphi\mathbf{1}_{[p_1,\infty)} = f\mathbf{1}_{[p_1,\infty)}$. In particular, we observe that $\varphi \leqslant f \vee g$. It is straightforward to check that any candidate for the convex envelope of $f \vee g$ is below φ. The above shows also that $D_p^+ f(p_1) > 0$ whenever $q_1 > 0$.
2. Let us now observe that $f^\sharp(q) < \infty$, for $q \geqslant 0$, since $f(\cdot,p) = g(\cdot,p) = \infty$ for

$p > 1 \vee p_1$. It follows that the subdifferential of $f^\sharp$ at non-negative q is non empty. The proof of (b) follows from calculations based on the following results from convex analysis, see e.g. [11, Chapter I Proposition 5.1]. Let ψ be a proper function on $\mathbb{R}$, then q is in the subdifferential of ψ at p if and only if

$$\psi^\sharp(q) + \psi(p) = pq\,. \tag{3.2}$$

(i) At $p = 0$, the subdifferential of $(f \vee g)^{\sharp\sharp} = \mathrm{co}[f \vee g]$ is equal to $[0, q_1]$. This follows directly from the characterisation of the convex envelope of $f \vee g$ given in (a). Using the above equality with $\psi = (f \vee g)^{\sharp\sharp}$, we then have for $q \in [0, q_1]$

$$(f \vee g)^{\sharp\sharp}(0) + (f \vee g)^\sharp(q) = 0 \times q \implies (f \vee g)^\sharp(q) = 0\,,$$

since $(f \vee g)^{\sharp\sharp}(0) = 0$ by our assumption, namely $f(0) = 0 = g(\cdot, 0)$ and $g \geqslant 0$.
(ii) The subdifferential of $(f \vee g)^{\sharp\sharp} = \mathrm{co}[f \vee g]$ at p_1 is equal to $\mathcal{D} := [q_1, D_p^+ f(p_1)]$ if $D_p^+ f(p_1) < +\infty$ or $[q_1, +\infty)$ otherwise. This follows again directly from (a). We recall from the Step 1. that $f(p_1) = q_1 p_1$. Then, using (3.2) with $\psi = (f \vee g)^{\sharp\sharp}$ and (a), we have for $q \in \mathcal{D}$

$$(f \vee g)^{\sharp\sharp}(p_1) + (f \vee g)^\sharp(q) = p_1 q \implies (f \vee g)^\sharp(q) = p_1 q - f(p_1) = p_1(q - q_1) = p_1[q - q_1]^+\,.$$

(iii) If $q > D_p^+ f(p_1)$, an element p of the subdifferential of $f^\sharp$ at q satisfies

$$f(p) + f^\sharp(q) = pq\,.$$

We first note that $p \geqslant p_1$ necessarily. Indeed, by [11, Chapter I Corollary 5.2], $q \in [D_p^- f(p), D_p^+ f(p)]$ while $q > D_p^+ f(p_1)$. Recall that $f = (f \vee g)^{\sharp\sharp}$ on $[p_1, \infty)$. We then deduce from the previous equality that

$$(f \vee g)^{\sharp\sharp}(p) + f^\sharp(q) = pq \implies f^\sharp(q) = pq - (f \vee g)^{\sharp\sharp}(p) \leqslant (f \vee g)^\sharp(q)\,.$$

Observing that the reverse inequality follows from $f \leqslant f \vee g$, we get $f^\sharp(q) = (f \vee g)^\sharp(q)$ for $q \in (D_p^+ f(p_1), +\infty)$. □

We are now in position to provide the decomposition of $(w^\sharp \vee g)^\sharp$ and $(w^\sharp \vee g)^{\sharp\sharp}$. It basically follows from the application of the previous lemma to $f = w^\sharp$.

Proposition 3.3. *For $(t, x, p) \in [0, T] \times \mathcal{O}_+^d \times \mathbb{R}$, we define the following 'facelift' of g*

$$\tilde{g}(t, x, p) = q_g(t, x) p \mathbf{1}_{\{0 \leqslant p \leqslant 1\}} + \infty \mathbf{1}_{\{p > 1\}}\,,$$

with

$$q_g(t, x) := \frac{g(t, x)}{p_g(t, x)} \mathbf{1}_{\{p_g(t,x) > 0\}} \quad \text{and} \quad p_g(t, x) := \sup\left\{p \in \mathbb{R} \,|\, w^\sharp(t, x, p) = g(t, x)\right\} \wedge 1\,.$$

Then,

(a) *The function* $p \mapsto (w^\sharp \vee g)^{\sharp\sharp}(\cdot, p)$ *is continuous on its domain and*

$$(w^\sharp \vee g)^{\sharp\sharp} = \mathrm{co}[w^\sharp \vee g] = w^\sharp \vee \tilde{g}\,. \tag{3.3}$$

(b) *For all* $q \in \mathbb{R}_+$*:*

$$(w^\sharp \vee g)^\sharp(\cdot, q) = [q - g(\cdot)]^+ \mathbf{1}_{A_1}(\cdot) + w(\cdot, q)\mathbf{1}_{A_2}(\cdot) + \kappa(\cdot, q)\mathbf{1}_{A_3}(\cdot)\,, \tag{3.4}$$

where

$$\kappa(\cdot, q) := p_g(\cdot)\,[q - q_g(\cdot)]^+ \mathbf{1}_{\{q<\bar{q}(\cdot)\}} + w(\cdot, q)\mathbf{1}_{\{q\geqslant\bar{q}(\cdot)\}}\,,$$

with $\bar{q}(\cdot) := D_p^+ w^\sharp\,(\cdot, p_g(\cdot))$ *and the subsets of* $[0,T]\times\mathcal{O}_+^d$*:* $A_1 := \{g > 0,\ w^\sharp(\cdot,1) \leqslant g\}$*,* $A_2 := \{g = 0\}$*,* $A_3 := \{g > 0,\ w^\sharp(\cdot,1) > g\}$*.*

Remark 3.1. (a) It follows from Proposition 3.2 that $w^\sharp(\cdot, 0) = 0$. Hence, $g(t,x) > 0$ implies $p_g(t,x) > 0$ and

$$q_g(t,x) = \frac{g(t,x)}{p_g(t,x)}\mathbf{1}_{\{g(t,x)>0\}} \text{ so that } q_g(t,x) = 0 \text{ if and only if } g(t,x) = 0\,.$$

(b) The decomposition on A_1, A_2 and A_3 will prove useful in the sequel, see e.g. proof of Lemma 3.2(c) below.
(c) On A_3, we have $\bar{q} > 0$ since $w^\sharp(\cdot, p_g(\cdot)) \geqslant g > 0$ and $w^\sharp(\cdot, 0) = 0$, see Proposition 3.2.

Proof of Proposition 3.3. The identities in (3.3) are immediate consequences of Lemma 3.1(a), Proposition 3.2(b) and of the definition of p_g. We now prove (3.4). For $(t,x) \in A_1$, we have $w^\sharp(t,x,\cdot) \leqslant g$ and therefore $(w^\sharp \vee g)^\sharp(t,x,\cdot) = g^\sharp(t,x,\cdot) = [\cdot - g(t,x)]^+$ on $\mathbb{R}_+$. For $(t,x) \in A_2$, we have that $w^\sharp \geqslant g$ by Proposition 3.2(b) and the result follows directly. On A_3, the expression is exactly the one given by Lemma 3.1(b). □

We can now turn to the study of the subdifferential of w. Recall the definition of $p_{\min}$ in (2.10).

Lemma 3.2. *Fix* $0 \leqslant i \leqslant n-1$ *and* $(t,x) \in [t_i, t_{i+1}) \times \mathcal{O}_+^d$*. Then:*

(a) $D_q^+ w(t,x,\cdot) \geqslant 0$ *if* $q \geqslant 0$ *and* $D_q^- w(t,x,\cdot) \geqslant 0$ *if* $q > 0$*.*
(b) $\lim_{q\uparrow\infty} D_q^+ w(t,x,q) = 1$*.*
(c) $D_q^+ w(t,x,0) = p_{\min}(t,x)$*.*

Moreover,

$$D_q^- w(t,x,q) = \mathbb{E}\Big[D_q^-(w^\sharp \vee g)^\sharp(t_{i+1}, X_{t_{i+1}}^{t,x}, qQ_{t_{i+1}}^{t,x,1}))\Big] \quad \textit{for } q > 0\,, \quad \textit{and} \tag{3.5}$$

$$D_q^+ w(t,x,q) = \mathbb{E}\Big[D_q^+(w^\sharp \vee g)^\sharp(t_{i+1}, X_{t_{i+1}}^{t,x}, qQ_{t_{i+1}}^{t,x,1}))\Big] \quad \textit{for } q \geqslant 0\,. \tag{3.6}$$

Proof. The proof is based on an induction argument. Our assumptions guarantee that (a)-(b)-(c) are valid at T. Let us assume that it holds true on $[t_{i+1}, T]$ for some $i \leqslant n-1$.
In view of Proposition 3.3, we obtain for $q \geqslant 0$ and $j \leqslant n$ that

$$D_q^+(w^\sharp \vee g)^\sharp(t_j,x,q) = \mathbf{1}_{\{q \geqslant g(t_j,x)\}}\mathbf{1}_{A_1}(t_j,x) + D_q^+ w(t_j,x,q)\mathbf{1}_{A_2}(t_j,x) + D_q^+\kappa(t_j,x,q)\mathbf{1}_{A_3}(t_j,x)\,,$$

with

$$D_q^+\kappa(t_j,x,q) = p_g(t_j,x)\mathbf{1}_{\{q_g(t_j,x)\leqslant q<\bar{q}(t_j,x)\}} + D_q^+ w(t_j,x,q)\mathbf{1}_{\{q\geqslant \bar{q}(t_j,x)\}}\,.$$

For $q > 0$, we have

$$D_q^-(w^\sharp \vee g)^\sharp(t_j,x,q) = \mathbf{1}_{\{q > g(t_j,x)\}}\mathbf{1}_{A_1}(t_j,x) + D_q^- w(t_j,x,q)\mathbf{1}_{A_2}(t_j,x) + D_q^-\kappa(t_j,x,q)\mathbf{1}_{A_3}(t_j,x)\,,$$

with

$$D_q^-\kappa(t_j,x,q) = p_g(t_j,x)\mathbf{1}_{\{q_g(t_j,x)< q<\bar{q}(t_j,x)\}} + D_q^- w(t_j,x,q)\mathbf{1}_{\{q> \bar{q}(t_j,x)\}}\,.$$

Using our induction hypothesis, we have $\lim_{q\uparrow\infty} D_q^+\kappa(t_{i+1},x,q) = 1$, which ensures that $\lim_{q\uparrow\infty} D_q^+(w^\sharp \vee g)^\sharp(t_{i+1},x,q) = 1$. By the convexity of $(w^\sharp \vee g)^\sharp$, this implies that $D_q^+(w^\sharp \vee g)^\sharp(t_{i+1},x,q) \leqslant 1$. In view of (2.21), a dominated convergence argument then leads to (3.5)-(3.6) and $\lim_{q\uparrow+\infty} D_q^+ w(t,x,q) = 1$.
We now use our induction hypothesis again to observe from the decomposition above that

$$D_q^-(w^\sharp \vee g)^\sharp(t_{i+1},x,q) \geqslant 0\,,\ q>0\,,\quad \text{and}\quad D_q^+(w^\sharp \vee g)^\sharp(t_{i+1},x,q) \geqslant 0\,,\ q \geqslant 0\,.$$

Recalling (3.5)-(3.6), this shows that $D_q^- w(t,x,q) \geqslant 0$ for $q > 0$ and $D_q^+ w(t,x,q) \geqslant 0$ for $q \geqslant 0$.
It remains to prove (c). From Remark 3.1 (a) and (c), the above decomposition implies that $D_q^+(w^\sharp \vee g)^\sharp(t_{i+1},x,0) = D_q^+ w(t_{i+1},x,0)\mathbf{1}_{\{g(t_{i+1},x)=0\}}$. By our induction hypothesis, the last term is $D_q^+(w^\sharp \vee g)^\sharp(t_{i+1},x,0) = p_{\min}(t_{i+1},x)\mathbf{1}_{\{g(t_{i+1},x)=0\}}$. This identity combined with (3.6) provides

$$D_q^+ w(t,x,0) = \mathbb{E}\Big[p_{\min}(t_{i+1}, X^{t,x}_{t_{i+1}})\mathbf{1}_{\{g(t_{i+1},X^{t,x}_{t_{i+1}})=0\}}\Big] = p_{\min}(t,x)\,,$$

in which the last identity is a direct consequence of the definition of $p_{\min}$ in (2.10). □

Remark 3.2. Note that the subdifferential of $w(t,x,\cdot)$ at 0 is $(-\infty, p_{\min}(t,x)]$, since $w(t,x,q) = \infty$ for $q < 0$ and $D_q^+ w(t,x,0) = p_{\min}(t,x)$. See (a) of Proposition 3.2 and (c) of Lemma 3.2.

3.3 The backward algorithm as an upper-bound

Our final proof will proceed by backward induction on the time steps. Part of the induction hypothesis is:

Hypothesis ($\mathrm{H}_{\mathrm{i+1}}$). The following holds

(i) The functions $v(t_{i+1}, \cdot)$ and $\mathrm{co}[v \vee g](t_{i+1}, \cdot)$ are continuous on $\mathcal{O}_+^d \times [0,1]$.

(ii) $\mathrm{co}[v \vee g](t_{i+1}, \cdot, 0) = 0$ and $\mathrm{co}[v \vee g](t_{i+1}, \cdot, 1) = (v \vee g)(t_{i+1}, \cdot, 1)$.

(iii) For all $x \in \mathcal{O}_+^d$, the map $q \in \mathbb{R}_+ \mapsto q - (w^\sharp \vee g)^\sharp(t_{i+1}, x, q)$ is non-decreasing, continuous and converges to $(v \vee g)(t_{i+1}, x, 1)$ as $q \to \infty$.

Before turning to the final argument, we provide three additional results that hold at any time $t \in [t_i, t_{i+1})$ whenever $\mathrm{H}_{\mathrm{i+1}}$ is in force.

3.3.1 Bounds and limits for $w^\sharp$

Our first additional result concerns the behaviour of $w^\sharp$. It shows that $w^\sharp(t_i, x, 1) = v(t_i, x, 1)$. The last assertion will be used in the proof of Lemma 3.4 below to show that (iii) of H_{i} holds if (iii) of $\mathrm{H}_{\mathrm{i+1}}$ does.

Lemma 3.3. *Let (iii) of* $\mathrm{H}_{\mathrm{i+1}}$ *hold. Fix* $(t, x) \in [t_i, t_{i+1}) \times \mathcal{O}_+^d$*. Then,* $w^\sharp(t, x, \cdot)$ *is non-negative, continuous on its domain* $(-\infty, 1]$ *and*

$$0 \leqslant w^\sharp(t, x, \cdot) \leqslant w^\sharp(t, x, 1) = v(t, x, 1) \quad \text{on } (-\infty, 1]\,.$$

Moreover, the map $q \in \mathbb{R} \mapsto q - w(t, x, q)$ *is non-decreasing, continuous on* $\mathbb{R}_+$ *and converges to* $v(t, x, 1)$ *as* $q \to \infty$*.*

Proof. The continuity and non-negativity of $w^\sharp(t, x, \cdot)$ are stated in (b) of Proposition 3.2. We now observe that (2.21) implies that

$$\delta(q) := q - w(t, x, q) = \mathbb{E}^{\mathbb{Q}_{t,x}}\left[q Q_{t_{i+1}}^{t,x,1} - (w^\sharp \vee g)^\sharp(t_{i+1}, X_{t_{i+1}}^{t,x}, q Q_{t_{i+1}}^{t,x,1})\right],$$

which shows that $q \mapsto \delta(q)$ is non-decreasing since (iii) of $\mathrm{H}_{\mathrm{i+1}}$ holds. Applying the monotone convergence Theorem, (iii) of $\mathrm{H}_{\mathrm{i+1}}$ and (2.8), we obtain that $q \in \mathbb{R}_+ \mapsto q - w(t, x, q)$ is continuous and that

$$\lim_{q \to \infty} \delta(q) = \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)(t_{i+1}, X_{t_{i+1}}^{t,x}, 1)\right] = v(t, x, 1)\,.$$

This implies that $w^\sharp(t_i, x, 1) = \sup_{q \geqslant 0} \delta(q) \geqslant \lim_{q \to \infty} \delta(q) = v(t, x, 1)$, while $w^\sharp(t, x, p) \geqslant \lim_{q \to \infty}(q(p-1) + \delta(q)) = \infty$ for $p > 1$. The fact that $w^\sharp(t_i, x, 1) \leqslant v(t, x, 1)$ has been proved in Proposition 3.1. □

3.3.2 Convexification in the dynamic programming algorithms

As already mentioned in Remark 2.2, one can expect that $v \vee g$ can be replaced by its convex envelope, with respect to p, in (2.15). The Hypotheses (i)-(ii) of $\mathrm{H}_{\mathrm{i+1}}$ ensure this, see Proposition 3.4 below. We shall prove a similar result for $w^\sharp$ later on in Theorem 3.1. Note that the two identities (3.7) and (3.9) below already suggest that the equality $v = w^\sharp$ at t_{i+1} should iterate at t_i, since we already know from Proposition 3.1 that $v \geqslant w^\sharp$.

Proposition 3.4. *Let* (i)-(ii) *of* $\mathrm{H}_{\mathrm{i+1}}$ *hold. Then, for all* $t \in [t_i, t_{i+1})$ *and* $(x,p) \in \mathcal{O}^d_+ \times [0,1]$*, we have*

$$v(t,x,p) = \inf_{\alpha \in \mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}} \left[\mathrm{co}[v \vee g] \left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}} \right) \right] . \tag{3.7}$$

Moreover, (ii) *of* H_{i} *holds.*

Proof. We fix $(t,x) \in [t_i, t_{i+1}) \times \mathcal{O}^d_+$. Assuming that (3.7) is true, we deduce that (ii) of H_{i} holds, since $\mathcal{A}_{t,p} = \{0\}$ for $p \in \{0,1\}$ leading to $P^{t,p,\alpha}_{t_{i+1}} = p$ in that case. By (ii) of H_{i}, the same argument combined with Proposition 2.3 implies that (3.7) is valid for $p \in \{0,1\}$.

It remains to prove (3.7) for $0 < p < 1$. In view of Proposition 2.3, this reduces to showing that

$$\inf_{\alpha \in \mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}} \left[\mathrm{co}[v \vee g] \left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}} \right) \right] \geqslant \inf_{\alpha \in \mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}} \left[(v \vee g) \left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}} \right) \right] ,$$

the reverse inequality being straightforward. We argue as in the [5, Proof of Proposition 3.3]. It follows from the Caratheodory theorem that we can find two maps $(\lambda_j, \pi_j) : (x,p) \in \mathcal{O}^d_+ \times [0,1] \mapsto (\lambda_j, \pi_j)(x,p) \in \mathcal{O}^d_+ \times [0,1]$, $j \leqslant 2$, such that

$$\begin{aligned} \textstyle\sum_{j=1}^2 \pi_j(x,p) = 1 \ , \ p = \sum_{j=1}^2 \pi_j(x,p)\lambda_j(x,p) \\ \text{and} \ \ \mathrm{co}[v \vee g](t_{i+1},x,p) = \textstyle\sum_{j=1}^2 \pi_j(x,p)(v \vee g)(t_{i+1},x,\lambda_j(x,p)) \,. \end{aligned} \tag{3.8}$$

We claim that they can be chosen in a measurable way. More precisely, (i) of H_{i} and [4, Proposition 7.49] imply that they can be chosen to be analytically measurable. We can then appeal to [4, Lemma 7.27] to obtain a Borel-measurable version which coincides a.e. for the pull-back measure of $(X^{t,x}_{t_{i+1}-\varepsilon}, P^{t,p,\alpha}_{t_{i+1}-\varepsilon})$, for $\alpha \in \mathcal{A}_{t,p}$ and $0 < \varepsilon < t_{i+1} - t$ fixed. This is this version that we use in the following.

We now let ξ be a $\mathcal{F}_{t_{i+1}}$-measurable random variable such that

$$\mathbb{P}[\xi = \lambda_j(X^{t,x}_{t_{i+1}-\varepsilon}, P^{t,p,\alpha}_{t_{i+1}-\varepsilon}) | \mathcal{F}_{t_{i+1}-\varepsilon}] = \pi_j(X^{t,x}_{t_{i+1}-\varepsilon}, P^{t,p,\alpha}_{t_{i+1}-\varepsilon}) \,.$$

Then, $\mathbb{E}[\xi|\mathcal{F}_{t_{i+1}-\varepsilon}] = P^{t,p,\alpha}_{t_{i+1}-\varepsilon}$ by the above construction, and we can then find $\alpha_\varepsilon \in \mathcal{A}_{t,p}$ such that $P^{t,p,\alpha_\varepsilon}_{t_{i+1}-\varepsilon} = P^{t,p,\alpha}_{t_{i+1}-\varepsilon}$ and $P^{t,p,\alpha_\varepsilon}_{t_{i+1}} = \xi$. Recalling (3.8), we obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}_{t,x}}\left[\mathrm{co}[v \vee g]\left(t_{i+1}, X^{t,x}_{t_{i+1}-\varepsilon}, P^{t,p,\alpha}_{t_{i+1}-\varepsilon}\right)\right] &= \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X^{t,x}_{t_{i+1}-\varepsilon}, P^{t,p,\alpha_\varepsilon}_{t_{i+1}}\right)\right] \\ &\geqslant \inf_{\alpha' \in \mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha'}_{t_{i+1}}\right)\right] \\ &+ \Delta(\epsilon),\end{aligned}$$

with $\Delta(\epsilon) = -C\mathbb{E}^{\mathbb{Q}_{t,x}}\left[(1 + |X^{t,x}_{t_{i+1}-\varepsilon}| + |X^{t,x}_{t_{i+1}}|)|X^{t,x}_{t_{i+1}-\varepsilon} - X^{t,x}_{t_{i+1}}|\right]$, recall (2.6) and (2.16). Moreover, since $0 \leqslant \mathrm{co}[v \vee g](t_{i+1}, x, \cdot) \leqslant v \vee g(t_{i+1}, x, \cdot) \leqslant C(1 + |x|)$, using (i) of H_{i+1}, we can pass to the limit to obtain

$$\mathbb{E}^{\mathbb{Q}_{t,x}}\left[\mathrm{co}[v \vee g]\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}}\right)\right] \geqslant \inf_{\alpha' \in \mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha'}_{t_{i+1}}\right)\right].$$

□

Since our final result is $v = w^\sharp$, the same convexification should appear in the dual algorithm. As already mentioned, it will actually allow us to show that $v = w^\sharp$ at t_i if this true at t_{i+1}.

Theorem 3.1. *Let (iii) of* H_{i+1} *hold. Fix* $0 \leqslant i \leqslant n-1$, $(t, x, p) \in [t_i, t_{i+1}) \times \mathcal{O}^d_+ \times [0, 1]$. *Then, there exists* $\bar\alpha \in \mathcal{A}_{t,p}$ *such that*

$$w^\sharp(t, x, p) = \mathbb{E}^{\mathbb{Q}_{t,x}}\left[\mathrm{co}[w^\sharp \vee g]\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\bar\alpha}_{t_{i+1}}\right)\right]. \tag{3.9}$$

Proof. Recall the definition of $p_{\min}$ in (2.10).
1. We first assume that $p \in (p_{\min}(t, x), 1)$. We know from Lemma 3.2 (b)-(c) that there exists a $\tilde q \in (0, \infty)$ such that p lies in the subdifferential of $w(t, x, \cdot)$ at $\tilde q$. Then, we can find $\lambda \in [0, 1]$ such that $p = \lambda D^+_q w(t, x, \tilde q) + (1 - \lambda)D^-_q w(t, x, \tilde q)$. In view of (3.5)-(3.6), this implies that

$$p = \mathbb{E}\left[(\lambda D^+_q (w^\sharp \vee g)^\sharp + (1 - \lambda)D^-_q (w^\sharp \vee g)^\sharp)\left(t_{i+1}, X^{t,x}_{t_{i+1}}, \tilde q Q^{t,x,1}_{t_{i+1}}\right)\right]. \tag{3.10}$$

It follows from Lemma 3.2 and its proof that the random variable in the expectation is valued in $[0, 1]$. By the martingale representation Theorem, we can find $\bar\alpha \in \mathcal{A}_{t,p}$ such that

$$(\lambda D^+_q (w^\sharp \vee g)^\sharp + (1 - \lambda)D^-_q (w^\sharp \vee g)^\sharp)\left(t_{i+1}, X^{t,x}_{t_{i+1}}, \tilde q Q^{t,x,1}_{t_{i+1}}\right) = p + \int_t^{t_{i+1}} \bar\alpha_s^\top \mathrm{d}W_s =: P^{t,p,\bar\alpha}_{t_{i+1}}.$$

For later use, note that the above implies

$$P^{t,p,\bar\alpha}_{t_{i+1}} \tilde q Q^{t,x,1}_{t_{i+1}} - (w^\sharp \vee g)^\sharp\left(t_{i+1}, X^{t,x}_{t_{i+1}}, \tilde q Q^{t,x,1}_{t_{i+1}}\right) = (w^\sharp \vee g)^{\sharp\sharp}\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\bar\alpha}_{t_{i+1}}\right), \tag{3.11}$$

where we used (3.2) with $\psi = (w^\sharp \vee g)^\sharp$. On the other hand, we also have, again by (3.2) with $\psi = w$,

$$w(t,x,\tilde{q}) + w^\sharp(t,x,p) = \tilde{q}p\,, \tag{3.12}$$

and, by (2.21),

$$w(t,x,\tilde{q}) = \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(w^\sharp \vee g)^\sharp\left(t_{i+1}, X^{t,x}_{t_{i+1}}, \tilde{q}Q^{t,x,1}_{t_{i+1}}\right)\right]. \tag{3.13}$$

Thus, inserting (3.10) and (3.13) into (3.12), and using (3.11), leads to

$$\begin{aligned} w^\sharp(t,x,p) &= \mathbb{E}^{\mathbb{Q}_{t,x}}\left[P^{t,p,\bar{\alpha}}_{t_{i+1}}\tilde{q}Q^{t,x,1}_{t_{i+1}} - (w^\sharp \vee g)^\sharp\left(t_{i+1}, X^{t,x}_{t_{i+1}}, \tilde{q}Q^{t,x,1}_{t_{i+1}}\right)\right] \\ &= \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(w^\sharp \vee g)^{\sharp\sharp}\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\bar{\alpha}}_{t_{i+1}}\right)\right]. \end{aligned}$$

We conclude by appealing to (3.3).

2. We now assume that $p \in [0, p_{\min}(t,x)]$. Since $[0, p_{\min}(t,x)]$ belongs to the subdifferential of $w(t,x,\cdot)$ at 0, recall Remark 3.2, and $p_{\min}(t,x) = D^+_q w(t,x,0)$, recall Lemma 3.2, we can find $\lambda \in [0,1]$ such that $p = \lambda D^+_q w(t,x,0)$. We then proceed as above up to obvious modifications.

3. We finally assume that $p = 1$. We know from Lemma 3.3 that $w^\sharp(t,x,1) = v(t,x,1)$. Hence, (2.8) implies

$$w^\sharp(t,x,1) = v(t,x,1) = \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X^{t,x}_{t_{i+1}}, 1\right)\right].$$

As in the proof of Lemma 3.3, we deduce from (iii) of H_{i+1}that $\mathrm{co}[w^\sharp \vee g](t_{i+1},\cdot,1) = (w^\sharp \vee g)^{\sharp\sharp}(t_{i+1},\cdot,1) \geqslant (v \vee g)(t_{i+1},\cdot,1)$. In view of Proposition 3.1, this leads to $(v \vee g)(t_{i+1},x,1) = \mathrm{co}[w^\sharp \vee g](t_{i+1},x,1)$. □

3.4 Conclusion of the proof

To conclude the proof of Theorem 2.1, we need to prove the inequality $v \leqslant w^\sharp$.

Proposition 3.5. *$v \leqslant w^\sharp$ on $[0,T] \times \mathcal{O}^d_+ \times [0,1]$.*

Proof. We use a backward induction argument. We assume that H_{i+1} holds and that $v = w^\sharp$ and on $[t_{i+1},T] \times \mathcal{O}^d_+ \times [0,1]$ for some $i \leqslant n-1$. Since it is true for $i = n$ by construction, the proof will be completed if one can show that this implies that H_i holds and that $v = w^\sharp$ on $[t_i,T] \times \mathcal{O}^d_+ \times [0,1]$.

Let us fix $(t,x,p) \in [t_i, t_{i+1}) \times \mathcal{O}^d_+ \times [0,1]$. Then, our induction hypothesis implies that

$$\mathbb{E}^{\mathbb{Q}_{t,x}}\left[\mathrm{co}[v \vee g]\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}}\right)\right] = \mathbb{E}^{\mathbb{Q}_{t,x}}\left[\mathrm{co}[w^\sharp \vee g]\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}}\right)\right],$$

for all $\alpha \in \mathcal{A}_{t,p}$. It then follows from Theorem 3.1 and Proposition 3.4 that $v(t,x,p) \leqslant w^\sharp(t,x,p)$. But, the reverse inequality is proved in Proposition 3.1. This shows that $v = w^\sharp$ on on $[t_i,T] \times \mathcal{O}_+^d \times [0,1]$. Then (i) of $\mathrm{H_i}$ is a consequence of Proposition 3.3 and Proposition 3.2. Proposition 3.4 implies (ii) of $\mathrm{H_i}$. Regarding the validity of (iii) of $\mathrm{H_i}$, it is proved in Lemma 3.4 below. □

Lemma 3.4. *The hypothesis* $\mathrm{H_{i+1}}$ *implies (iii) of* $\mathrm{H_i}$*.*

Proof. It follows from (3.4) that

$$q - (w^\sharp \vee g)^\sharp(t_i,x,q) = \big(q - [q - g(t_i,x)]^+\big)\mathbf{1}_{A_1}(t_i,x) + (q - w(t_i,x,q))\mathbf{1}_{A_2}(t_i,x) + \big(q - \kappa(t_i,x,q)\big)\mathbf{1}_{A_3}(t_i,x)\,, \quad (3.14)$$

in which

$$q - \kappa(t_i,x,q) = (q - p_g(t_i,x)[q - q_g(t_i,x)]^+)\mathbf{1}_{\{q<\bar{q}(t_i,x)\}} + (q - w(t_i,x,q))\mathbf{1}_{\{q\geqslant\bar{q}(t_i,x)\}}\,.$$

By Lemma 3.3, $w^\sharp(t_i,x,1) = v(t_i,x,1)$ so that $A_2 \cup A_3 = \{v(\cdot,1) > g\}$, recall (2.11). In particular, we observe that $\bar{q} < \infty$ on A_3. The fact that the right-hand side in (3.14) converges to $(v \vee g)(t_i,x,1)$ as $q \to \infty$ is then a consequence of Lemma 3.3 and the definition of the $(A_i)_{i\leqslant 3}$.
It remains to show that each term in (3.14) is non-decreasing and continuous. From Lemma 3.3, we know that $q \mapsto (q - w(t_i,x,q)$ is continuous and non-decreasing. The first term in the right-hand side of (3.14) is continuous and non-decreasing as well. Regarding the last term, we know that $q \mapsto \kappa(t_i,x,q)$ is continuous, so that it suffices to check the monotony on each sub-interval $(-\infty,\bar{q}(t_i,x)]$ and $[\bar{q}(t_i,x),\infty)$ distinctly. On the second interval, we have that $q \mapsto q - \kappa(t_i,x,q)$ is non-decreasing by Lemma 3.3. This is also true on first interval since $p_g(t_i,x) \leqslant 1$. □

4 Appendix

We provide here the proofs of some technical results that were used in the proof of Theorem 2.1.

Proof of Proposition 2.1 For $t = T$ the sets in (2.9) are $\mathbb{R}_+$ by definition of $\mathbb{T}_t$ and $\mathcal{T}_t$. For $t < T$, $\mathrm{S}^{t,x,y,\nu}_{\hat{\tau}_\nu} \subset \bigcap_{s\in\mathbb{T}_t} \mathrm{S}^{t,x,y,\nu}_s \subset \mathrm{S}^{t,x,y,\nu}_\tau$, for any $\tau \in \mathcal{T}_t$, which proves (2.9) again.

Proof of Proposition 2.3. 1. We first show that (2.15) holds. Let $\bar{v}(t,x,p)$ denote the right-hand side of (2.15) and set

$$J(t,x,p,\alpha) := \mathbb{E}^{\mathbb{Q}_{t,x}}\left[(v \vee g)\left(t_{i+1}, X^{t,x}_{t_{i+1}}, P^{t,p,\alpha}_{t_{i+1}}\right)\right].$$

Fix y and $\alpha \in \mathcal{A}_{t,p}$ such that $y > J(t,x,p,\alpha)$. Then, it follows from the martingale representation theorem that we can find $\nu \in \mathcal{U}_{t,x,y}$ such that

$$Y_{t_{i+1}}^{t,x,y,\nu} > (v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right) .$$

In particular, $Y_{t_{i+1}}^{t,x,y,\nu} \geqslant g\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right)$. Since, we also have $Y_{t_{i+1}}^{t,x,y,\nu} > v(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha})$, it follows from the same arguments as in the proof of [8, Lemma 2.2] that we can find a predictable process $(\tilde{\nu}, \tilde{\alpha})$ which coincides with (ν, α) on $[t, t_{i+1}]$, in the $\mathrm{d}t \times \mathrm{d}\mathbb{P}$-sense, and such that

$$Y_s^{t,x,y,\tilde{\nu}} \geqslant g\left(s, X_s^{t,x}, P_s^{t,p,\tilde{\alpha}}\right) \text{ , for all } s \in \mathbb{T}_{t_{i+1}} .$$

These processes are elements of $\hat{\mathcal{U}}_{t,x,y,p}$ whenever $\tilde{\nu}$ is square integrable in the sense of (2.5), and $\tilde{\alpha}$ is square integrable in the classical sense. To reduce to this case, we use the fact that $P^{t,p,\tilde{\alpha}}$ is restricted to live in the interval $[0,1]$ while $\tilde{\nu}$ can be modified so that (2.12) holds. By the Itô isometry, this induces the required square integrability property of the financial strategy, recall (2.2). Combining the above with Proposition 2.2 shows that $\bar{v}(t,x,p) \geqslant v(t,x,p)$.
Conversely, let us fix $y > v(t,x,p)$. Then, it follows from the geometric dynamic programming principle of [8, Theorem 2.1] that there exists $(\nu, \alpha) \in \hat{\mathcal{U}}_{t,x,y,p}$ such that

$$Y_{t_{i+1}}^{t,x,y,\nu} \geqslant (v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right) .$$

Since $Y^{t,x,y,\nu}$ is a super-martingale under $\mathbb{Q}_{t,x}$, this implies that $y \geqslant J(t,x,p,\alpha)$. The fact that $v(t,x,p) \geqslant \bar{v}(t,x,p)$ then follows from the arbitrariness of α.
2. We now prove the Lipschitz continuity property. Note that it is true for $t = T$, since $v(T,\cdot) = 0$ by construction. Let us assume that (2.16) holds on $[t_{i+1}, T]$ for some $i < n$ and show that it is then also true on $[t_i, T]$. Let us fix $(t,p) \in [t_i, t_{i+1}) \times [0,1]$ and $x, x' \in \mathcal{O}_+^d$. It follows from (2.15) that $|v(t,x,p) - v(t,x',p)|$ is bounded from above by

$$\sup_{\alpha \in \mathcal{A}_{t,p}} \mathbb{E}^{\mathbb{Q}_{t,x}}\left[\left|(v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x}, P_{t_{i+1}}^{t,p,\alpha}\right) - (Q_{t_{i+1}}^{t,x,1}/Q_{t_{i+1}}^{t,x',1})(v \vee g)\left(t_{i+1}, X_{t_{i+1}}^{t,x'}, P_{t_{i+1}}^{t,p,\alpha}\right)\right|\right] .$$

Since (2.16) holds for $(v \vee g)(t_{i+1}, \cdot, p)$, (2.6) holds, and v has linear growth, see (2.13), we deduce that there exists $C > 0$ such that the above is bounded by

$$C\, \mathbb{E}^{\mathbb{Q}_{t,x}}\left[|X_{t_{i+1}}^{t,x} - X_{t_{i+1}}^{t,x'}|(1 + |X_{t_{i+1}}^{t,x}| + |X_{t_{i+1}}^{t,x'}|) + |Q_{t_{i+1}}^{t,x,1}/Q_{t_{i+1}}^{t,x',1} - 1|\, (1 + |X_{t_{i+1}}^{t,x'}|)\right] .$$

In view of (2.2)-(2.3), this is controlled by $|x - x'|(1 + |x| + |x'|)$ up to a multiplicative constant. □

Proof of Proposition 2.4. The growth property on $[0, T) \times \mathcal{O}_+^d \times (0, \infty)$ follows from Proposition 3.2 (which will be proved just below), Theorem 2.1, (3.1) and (2.13):

$$0 \leqslant w(t, x, q) = \sup_{p \in \mathbb{R}} (pq - v(t, x, p)) = \sup_{p \in [0,1]} (pq - v(t, x, p)) \leqslant q\,.$$

Note that Theorem 2.1 implies that $(w^\sharp \vee g)^\sharp(T, \cdot) = g^\sharp$. The fact that the lower- (resp. upper-) semicontinuous envelope of w is a viscosity super- (resp. sub-) solution of $(\mathcal{S})$ is standard and we omit the proof. Continuity will then follow from the comparison principle. The comparison can be proved by backward induction. It is well-known that (2.19) admits a comparison principle in the class of functions with polynomial growth, see e.g. [10]. Hence, the comparison holds on $[t_{n-1}, T)$. Assume that it holds on $[t_{i+1}, T)$ and that $(u_j^\sharp \mathbf{1}_{[0,T)} \vee g)^\sharp(t_{i+1}, \cdot)$ has polynomial growth, for $j = 1, 2$, then it holds on $[t_i, T)$ too. Indeed $u_1(t_{i+1}, \cdot) \geqslant u_2(t_{i+1}, \cdot)$ implies $(u_1^\sharp \vee g)^\sharp(t_{i+1}, \cdot) \geqslant (u_2^\sharp \vee g)^\sharp(t_{i+1}, \cdot)$. Hence, we just have to prove that $(u_1^\sharp \vee g)^\sharp$ has polynomial growth. By [15, Theorem 16.5], we have $(u_j^\sharp \vee g^{\sharp\sharp})^\sharp = \text{co}[u_j^{\sharp\sharp} \wedge g^\sharp]$. Hence $0 \leqslant (u_j^\sharp \vee g)^\sharp \leqslant (u_j^\sharp \vee g^{\sharp\sharp})^\sharp = \text{co}[u_j^{\sharp\sharp} \wedge g^\sharp] \leqslant \text{co}[u_j \wedge g^\sharp]$. Since the later has polynomial growth, the required property holds. □

Proof of Proposition 3.2. We proceed by backward induction on $\mathbb{T}_0 \cup \{0\}$. Our claims are straightforward from (2.21) at time T. Indeed, direct computations show that $w^\sharp(T, \cdot, p) = 0 + \infty \mathbf{1}_{\{p>1\}}$. Hence, $(w^\sharp \vee g)^\sharp(T, x, q) = g^\sharp(T, x, q) = [q - g(T, x)]^+ + \infty \mathbf{1}_{\{q<0\}}$. The properties (a) and (b) hold.

We now assume that (a) and (b) are satisfied on $[t_{i+1}, T]$ for some $i \leqslant n-1$ and fix $(t, x) \in [t_i, t_{i+1}) \times \mathcal{O}_+^d$. Then, the definition of w in (2.21) implies that $w(t, x, \cdot)$ is non-negative, non-decreasing, convex and that $w(t, x, 0) = 0$ (it is in particular proper). It takes the value $+\infty$ for $q < 0$, by (2.21) and the fact that $(w^\sharp \vee g)^\sharp(t_{i+1}, \cdot, q) = +\infty$ for $q < 0$. Hence (a) holds on $[t_i, T]$. These two last assertions imply that $w^\sharp(\cdot, p) = \sup_{q \geqslant 0} \{pq - w(\cdot, q)\}$ and $w^\sharp(t, \cdot, p) = 0$ for $p \leqslant 0$. We know from [15, Theorem 12.2] that it is closed, convex and continuous on the interior of its domain. Since $w^\sharp$ is non-decreasing, by definition, we get from its closeness that it is continuous on its domain. The fact that $w^\sharp(t, \cdot, \cdot) \geqslant w^\sharp(t, \cdot, 0) = 0$ also implies that $(w^\sharp \vee g)(t, x, \cdot)$ is non-negative; moreover, $(w^\sharp \vee g)(t, \cdot, 0) = 0$. For $q < 0$, we then compute $(w^\sharp \vee g)^\sharp(t, \cdot, q) = \sup_{p \leqslant 1} \{pq - (w^\sharp \vee g)(t, \cdot, p)\} = +\infty$. For $q \geqslant 0$, we get $(w^\sharp \vee g)^\sharp(t, \cdot, q) = \sup_{p \in [0,1]} \{pq - (w^\sharp \vee g)(t, \cdot, p)\} \geqslant 0$. Moreover, $(w^\sharp \vee g)^\sharp(t, x, 0) = 0$ and $(w^\sharp \vee g)^\sharp(t, x, \cdot)$ non-decreasing on its domain $[0, \infty)$. By definition, $(w^\sharp \vee g)^\sharp(t, x, \cdot)$ is convex and then continuous on the interior of its domain. Being non-decreasing and convex, it is also right-continuous at 0. □

References

[1] Vlad Bally and Gilles Pagès. Error analysis of the quantization algorithm for obstacle problems. *Stochastic Processes and their Applications*, 106(1):1–40, 2003.

[2] Vlad Bally and Gilles Pagès. A quantization method for pricing and hedging multidimensional american style options. *Mathematical finance*, 15(1):119–168, 2005.

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

[4] Dimitri P Bertsekas and Steven E Shreve. *Stochastic optimal control: The discrete time case*, volume 139. Academic Press New York, 1978.

[5] Bruno Bouchard, Romuald Elie, and Anthony Réveillac. Bsdes with weak terminal condition. *Annals of Probability, to appear*, 2012.

[6] Bruno Bouchard, Romuald Elie, and Nizar Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2009.

[7] Bruno Bouchard and Thanh Nam Vu. A stochastic target approach for p&l matching problems. *Mathematics of Operations Research*, 37(3):526–558, 2012.

[8] Bruno Bouchard and Thanh Nam Vu. The obstacle version of the geometric dynamic programming principle: Application to the pricing of american options under constraints. *Applied Mathematics and Optimization*, 61(2):235–265, 2010.

[9] Bruno Bouchard and Xavier Warin. Monte-carlo valuation of american options: facts and new algorithms to improve existing methods. In *Numerical methods in finance*, pages 215–255. Springer, 2012.

[10] M.G. Crandall, H. Ishii, and P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Amer. Math. Soc.*, 27:1–67, 1992.

[11] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*, volume 1. North-Holland American Elsevier, 1976.

[12] Hans Föllmer and Peter Leukert. Quantile hedging. *Finance and Stochastics*, 3(3):251–273, 1999.

[13] Hans Föllmer and Peter Leukert. Efficient hedging: cost versus shortfall risk. *Finance and Stochastics*, 4(2):117–146, 2000.

[14] Ying Jiao, Olivier Klopfenstein, and Peter Tankov. Hedging under multiple risk constraints. *arXiv preprint arXiv:1309.5094*, 2013.

[15] R Tyrrell Rockafellar. *Convex analysis*. Number 28. Princeton university press, 1997.

[16] H Mete Soner and Nizar Touzi. Stochastic target problems, dynamic programming, and viscosity solutions. *SIAM Journal on Control and Optimization*, 41(2):404–424, 2002.

[17] H Mete Soner and Nizar Touzi. A stochastic representation for mean curvature type geometric flows. *Annals of probability*, 31(3):1145–1165, 2003.