Politecnico di Milano

Scuola di Ingegneria Industriale e dell'Informazione
Corso di Laurea Magistrale in Ingegneria Matematica

Tesi di Laurea Magistrale

Representation of solutions to Hamilton-Jacobi-Bellman equations using constrained Backward Stochastic Differential Equations

Relatore:
Prof. Marco Fuhrman

Autore:
Nahuel Foresta
Matr.798775

Anno Accademico 2013-2014

Abstract

The aim of this work is to provide, following the paper [1], a representation formula for the Hamilton-Jacobi-Bellman equation using a Forward Backward Stochastic Differential Equations (FBSDE) system. In order to do this, we introduce a class of BSDE where the generator and the final condition are adapted to a bigger filtration than the one generated by the Brownian motion. This equation can be reformulated as a BSDE with constraints on the gain process. We discuss the existence and uniqueness of a minimal solution to this BSDE under reasonable assumptions. We show how this class of equations, when the generator and terminal data are given by a forward diffusion process, provides a representation formula for the solution (in a viscosity sense) to HJB fully non-linear PDE arising in stochastic control problems. In addition, as it is done in [1], we introduce an auxiliary *dual* control problem to which the solution to the BSDE provides the optimal value. This implies a dual representation for stochastic control problems as both are represented by the solution to the BSDE.

[1] Idris Kharroubi and Huyên Pham. 'Feynman-Kac representation for Hamilton-Jacobi-Bellman IPDE'. In: *arXiv preprint arXiv:1212.2000, to appear on Annals Of Probability* (2012)

Keywords: BSDE with constraints, stochastic optimal control, Hamilton-Jacobi-Bellman equation, viscosity solutions.

Sommario

Il presente lavoro ha l'obiettivo di trovare una rappresentazione probabilistica per la soluzione alle equazioni di Hamilton-Jacobi-Bellman del tipo

$$-\frac{\partial u}{\partial t} - \sup_{a \in A} \left[\mathcal{L}^a u + f(\cdot, a, u, \sigma^{\mathrm{T}}(\cdot, a) D_x u) \right] = 0 \text{ su } [0, T) \times \mathbb{R}^n \tag{1}$$

$$u(T, x) = \sup_{a \in \mathbb{R}^l} g(x, a) \qquad x \in \mathbb{R}^n, \tag{2}$$

con $\mathcal{L}^a f(t,x) = {}^1\!/_2 \operatorname{tr}(\sigma\sigma^{\mathrm{T}}(x,a) D_x^2 f(t,x)) + (D_x f(t,x))^{\mathrm{T}} b(x,a)$, utilizzando equazioni differenziali stocastiche backward (BSDE). Con rappresentazione probabilistica si intende che la soluzione è data sotto forma di valore atteso di una variabile aleatoria o come soluzione di un'equazione differenziale stocastica. In questo lavoro si utilizza una speciale classe di equazioni stocastiche backward con vincoli su una parte della soluzione.

Recentemente Kharroubi e Pham hanno pubblicato l'articolo [1] nel quale, utilizzando una BSDE con vincoli sulla parte di salto, trovano una rappresentazione probabilistica per le equazioni integro-differenziali di tipo HJB. L'idea in [1], utilizzata inizialmente in [2] per i problemi di switching e in [3] per le disuguaglianze variazionali associate a problemi di controllo impulsivo, è la seguente: il processo di controllo viene sostituito da un processo aleatorio di salto e al problema viene associata una BSDE con vincoli sulla parte di salto. Il presente lavoro parte da tale risultato e ne ricerca uno analogo per l'equazione (1) utilizzando una BSDE con vincoli sulla parte Browniana. Sebbene il presente elaborato metodologicamente segua passo per passo il lavoro di Kharroubi e Pham, l'uso di una diversa classe di BSDE introduce delle difficoltà tecniche diverse. Per questo si configura come un'estensione a un altro caso. Una prima differenza è che se in [1] il sup nell'equazione (1) è preso su un insieme compatto, è qui invece calcolato su tutto $\mathbb{R}^l$. Torneremo in seguito sulle ulteriori differenze fra il presente elaborato e il precedente.

L'equazione (1), fortemente non lineare, è di notevole importanza nell'ambito dei problemi di Controllo Ottimo Stocastico, in quanto il valore ottimo di un problema di controllo ottimo stocastico è soluzione di un'equazione di questo tipo. Per problema di Controllo Ottimo Stocastico si intende il problema di massimizzare (o minimizzare) un dato funzionale che dipende da una variabile di stato X_t, che evolve secondo un modello stocastico e sulla quale si può agire tramite un parametro di controllo. Esiste un'ampia letteratura a riguardo, si veda per esempio [4], [5], [6], [7]. Di solito il problema si configura come la ricerca della funzione valore ottimo:

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^T f(s, X_s, \alpha_s) ds + g(X_T)\right],$$

dove le funzioni f e g rappresentano dei guadagni (o dei costi nel caso di minimizzazione) e l'insieme $\mathcal{A}(t,x)$ è costituito da tutti i controlli ammissibili specifici del problema.
È possibile dimostrare che la funzione v è caratterizzabile come soluzione di un equazione di tipo HJB (vedere per esempio [6],[7]). Questo è uno degli approcci standard al controllo ottimo stocastico.

Data la complessità delle equazioni di tipo HJB, spesso si cerca una soluzione in un senso particolare detto di *viscosità*. Il concetto di soluzione viscosa è una generalizzazione di quello di soluzione classica, introdotto all'inizio degli anni '80 da P. L. Lions e M. G. Crandall in [8]. Per un'introduzione a questo tipo di soluzioni si rimanda a [9] e [10]. In breve, secondo questo approccio, una funzione non regolare può essere soluzione di un'equazione alle derivate parziali se un'opportuna classe di funzioni test soddisfa l'equazione. Questa teoria è applicabile in molti campi, in particolare alle equazioni di tipo HJB. Diamo qui la definizione di soluzione di viscosità per tale equazione

***Definizione* 1**

1. Una funzione w inferiormente (*risp. superiormente*) semi-continua su $[0,T] \times \mathbb{R}^n$ è soprasoluzione (*risp. sottosoluzione*) di viscosità di (1)-(2) se per ogni $(t,x) \in [0,T) \times \mathbb{R}^n$ e per ogni funzione test φ tale che $w - \varphi$ ha un minimo (*risp. massimo*) locale in (t,x) si abbia

$$-\frac{\partial \varphi}{\partial t} - \sup_{a \in \mathbb{R}^l} \left[\mathcal{L}^a \varphi + f(., a, w, \sigma^{\mathrm{T}}(., a) D_x \varphi)\right] \underset{(\leq)}{\geq} 0,$$

   e se per ogni $x \in \mathbb{R}^n$ vale

$$w(T,x) \underset{(\leq)}{\geq} \sup_{a \in \mathbb{R}^l} g(x,a).$$

2. Una funzione localmente limitata è soluzione viscosità di (1)-(2) se il suo *inviluppo semi-continuo inferiore* è soprasoluzione di viscosità per (1)-(2) e il suo *inviluppo semi-continuo superiore* è sottosoluzione di viscosità per (1)-(2).

Questo tipo di soluzione, nonostante copra un'ampia classe di equazioni, non è semplice da trattare e tanto meno da trovare. Inoltre, i metodi numerici per le soluzioni viscose sono spesso complessi e costosi, e l'uso di metodi probabilistici è competitivo, sopratutto se in dimensione elevata. Per questo metodo la ricerca di rappresentazioni probabilistiche ha conosciuto un forte impulso negli ultimi anni. Idealmente, si possono utilizzare metodi di tipo Monte Carlo per ottenere un'approssimazione numerica della soluzione.

Uno strumento molto utilizzato per questi scopi è quello delle BSDE. Si tratta di equazioni stocastiche delle quali è data la condizione finale. Nel caso deterministico dare una condizione iniziale o finale è equivalente a meno di un inversione dell'asse dei tempi. Nel caso stocastico, il fatto che la condizione sia data al tempo finale, e sia misurabile per la σ-algebra relativa al tempo finale, introduce problemi di misurabilità nella soluzione. Peng e Pardoux hanno introdotto all'inizio degli anni '90 in [11] una teoria specifica per questo tipo di equazioni, che, da allora, ha conosciuto un forte sviluppo. In pratica la soluzione a un'equazione di questo tipo è una coppia di processi, costituita dalla soluzione e da un processo che permette di rispettare i vincoli di misurabilità. Per una discussione completa si rimanda il lettore a [12], [13], [14].

Descriviamo ora brevemente il lavoro svolto, evidenziando i punti in cui si discosta dal lavoro [1] di Kharroubi e Pham.

BSDE con vincoli sulla parte martingala Nella prima parte del lavoro, si studia l'esistenza e unicità della seguente BSDE:

$$Y_t = \xi + \int_t^T \bar{F}(s, Y_s, Z_s)ds + K_T - K_t - \int_t^T Z_s dW_s, \tag{3}$$

dove la soluzione è però adattata a una filtrazione più grande rispetto a quella generata dal processo W. Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità su cui sono definiti due moti Browniani indipendenti W (d-dimensionale) e B (l-dimensionale). Sia inoltre $\widetilde{W} = (W, B)^{\mathrm{T}}$. Si può riformulare la (3) come una BSDE con vincolo:

$$\begin{cases} Y_t = \xi + \int_t^T F(s, Y_s, Z_s, Cs)ds + K_T - K_t - \int_t^T Z_s dW_s - \int_t^T C_s dB_s & \text{(4a)} \\ |C_t| = 0 \qquad \forall 0 \le t \le T. & \text{(4b)} \end{cases}$$

Il termine K_t è un processo crescente che permette di rispettare il vincolo di misurabilità. Si cerca una soluzione minimale secondo la definizione seguente:

***Definizione* 2** Una soluzione minimale di (4a) e (4b) è una quadrupla di processi (Y, Z, C, K) che soddisfa (4a) e (4b) tale che per ogni altra quadrupla $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right)$ che soddisfa (4a) e (4b), si ha che

$$Y_t \le \bar{Y}_t, \qquad 0 \le t \le T \text{ q.c.}$$

Nella precedente definizione è stata omessa per brevità l'elenco degli spazi in cui deve risiedere la soluzione. Si tratta infatti dei consueti spazi utilizzati nella teoria delle BSDE. In questa sede l'accento è posto sul concetto di soluzione *minima*. Per la definizione precisa si rimanda al capitolo 3.
Più interessanti da riportare sono le ipotesi richieste, standard per la teoria delle BSDE, riportate di seguito. Denotiamo con $\mathcal{P}$ la σ-algebra progressiva, ossia quella generata da tutti i processi progressivi.

[H0]

i) $\mathbb{E}\left[|\xi|^2\right] < \infty$

ii) ξ è $\mathcal{F}_T^{\widetilde{\mathbf{W}}}$-misurabile

iii) F è $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^l)$ measurable.

iv) F soddisfa una condizione di Lipschitzianità uniforme: $\exists C_F > 0$ t.c.

$$|F(t, y, z, c) - F(t, y', z', c')| \le C_F \left(|y - y'| + |z - z'| + |c - c'|\right)$$

v) $\mathbb{E}\left[\int_0^T |F(t, 0, 0, 0)|^2 \, dt\right] < \infty,$

e la seguente

[H1] Esiste una quadrupla $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right)$ che soddisfa (4a)-(4b).

Sotto queste ipotesi, è possibile utilizzare un risultato formulato da Peng in [15] per dimostrare che esiste una soluzione minimale a (4a) tramite una tecnica di penalizzazione. Si introduce a tal fine una BSDE penalizzata della forma:

$$Y_t^n = \xi + \int_t^T F\left(s, Y_s^n, Z_s^n, C_s^n\right) ds + K_T^n - K_t^n - \int_t^T Z_s^n dW_s - \int_t^T C_s^n dB_s, \quad (5)$$

dove il termine K_t^n è definito come

$$K_t^n = n \int_0^t |C_s^n| \, ds,$$

e si ottiene il seguente risultato:

Proposizione 1 *Sotto le ipotesi sopraindicate, esiste un'unica soluzione minimale a* (4a)*, limite crescente della successione* $(Y_t^n)_{n \geq 1}$*, dove le* Y^n *sono le soluzioni a* (5)*.*

Per l'esistenza si sfrutta il risultato già fornito in [15], mentre l'unicità viene mostrata con argomenti standard.

Stabilito questo risultato, nel caso in cui la funzione f sia un processo dato (ossia non dipende da y e z), si introduce un problema di controllo stocastico *duale*. Per tale problema, la soluzione alla BSDE appena introdotta fornisce il valore ottimo. Si dimostra infatti che sotto le stesse ipotesi, la soluzione è rappresentabile come

$$Y_t = \operatorname*{ess\,sup}_{u \in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right], \quad (6)$$

dove l'insieme U_A è definito come

$$U_A = \left\{ u_s \text{ processi in } \mathbb{R}^l, \mathcal{F}_t^{\widetilde{W}}\text{-progressivi, limitati } \right\}$$

e il valore atteso è calcolato sotto una probabilità $\mathbb{Q}^u$ che dipende dal processo u. La (6) si può interpretare come il problema di massimizzare un funzionale su un insieme di probabilità equivalenti, dove ogni probabilità è definita rispetto a $\mathbb{P}$ come:

$$\frac{\mathrm{d}\mathbb{Q}^u|_{\mathcal{F}_t^{\widetilde{W}}}}{\mathrm{d}\mathbb{P}|_{\mathcal{F}_t^{\widetilde{W}}}} = L_t^u,$$

dove L_t^u è una martingala di passaggio *alla Girsanov*, definita da:

$$L_t^u = \exp\left\{\int_0^t u_s dB_s - \frac{1}{2}\int_0^t u_s^2 ds\right\}.$$

La tecnica utilizzata per dimostrare questo risultato è anche qui simile a quella in [1]: si dimostra un risultato simile per la soluzione dell'equazione penalizzata Y^n e poi si estende il risultato a Y. Questa interpretazione è stata introdotta sempre in [1] nel caso di vincoli sulla parte di salto e in questo lavoro viene estesa al caso di una BSDE con vincoli sulla parte Browniana.

BSDE con vincoli e HJB Si passa quindi alla parte principale del lavoro. Sempre seguendo le linee guida di [1], si dimostra che la classe di BSDE sopra definita può essere utilizzata per rappresentare la soluzione delle equazioni di Hamilton-Jacobi-Bellman tipiche del controllo ottimo stocastico, quando il dato finale e il generatore sono dati da un processo stocastico diffusivo in avanti.
L'idea sottostante il procedimento che segue è di sostituire il processo di controllo nella diffusione controllata con un processo Browniano e lavorare su questa nuova equazione in congiunzione con la BSDE. Si ottiene così un sistema Forward Backward della forma:

$$\begin{cases} dX_s = b(X_s, X'_s)ds + \sigma(X_s, X'_s)dW_s \\ dX'_s = dB_s \\ Y_t = g(X_T, X'_T) + \int_t^T f(X_s, X'_s, Y_s, Z_s)ds + K_T - K_t - \int_t^T Z_s dW_s - \int_t^T C_s dB_s \\ |C_t| = 0. \end{cases} \tag{7}$$

Sia data ora una condizione iniziale (t, x, a) tale che $(X_t, X'_t) = (x, a)$. Il sistema va risolto in avanti per la prima parte in quanto disaccoppiato, e poi all'indietro per la BSDE (che come è stato mostrato sopra, ammette un'unica soluzione minimale). Denotiamo con $(X^{t,x,a}, X'^{t,x,a}, Y^{t,x,a})$ i processi quando il dato iniziale è (t, x, a). A questo punto, definendo la funzione v come

$$v(t, x, a) := Y_t^{t,x,a}. \tag{8}$$

Tale funzione è soluzione, in senso viscoso, della seguente PDE (1)-(2). Oggetto della restante parte del lavoro è la verifica di questa affermazione. In particolare, la parte più complessa è la dimostrazione che la funzione v è, fissati (t, x), constante in a. In questa fase, le specificità del problema qui considerato, hanno portato a sostanziali modifiche rispetto al lavoro svolto in [1]. Infatti, nell'articolo di Pham e Kharroubi, il processo di controllo viene sostituito con un processo di puro salto in un compatto A, dando luogo a una BSDE con salti che introducono termini integrali. In questo caso, invece, si ha a che fare con termini Browniani che portano ad introdurre termini differenziali. Pertanto non si ha più la compattezza dello spazio A (che nel nostro caso è $\mathbb{R}^l$), tuttavia non avendo termini integrali è possibile utilizzare minimi e massimi locali nelle definizioni di soluzione viscosa. Sì tornerà in seguito su altre differenze introdotte da questo cambiamento.

Come al solito, W e B sono moti Browniani in $\mathbb{R}^d$ e $\mathbb{R}^l$ rispettivamente. Siano date le funzioni $b : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^{n\times d}$, $f : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ e $g : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$. Si pongono le seguenti ipotesi su b e σ:

[HFD]

(i) b e σ soddisfano una condizione di Lipschitzianità, i.e.

$$\exists C > 0 \text{ t.c. } |b(x,a) - b(x',a')| + |\sigma(x,a) - \sigma(x',a')| \leq C(|x-x'| + |a-a'|)$$

$$\forall x, x' \in \mathbb{R}^n, \qquad \forall a, a' \in \mathbb{R}^l$$

(ii) $\exists M < \infty$ t.c. $\sup_{a\in\mathbb{R}^l} |b(0,a)| + \sup_{a\in\mathbb{R}^l} |\sigma(0,a)| \leq M$

e su f e g:

[HFC]

(i) f soddisfa una condizione di Lipschitzianità in y e z per qualche costante C, uniformemente in (x,a)

$$\big|f(x,a,y,z) - f(x,a,y',z')\big| \leq C(\big|y-y'\big| + \big|z-z'\big|)$$

(ii) f e g soddisfano una condizione crescita polinomiale in x, i.e. $\exists m \geq 0$ e $\exists C_g > 0$ t.c.

$$\sup_{a\in\mathbb{R}^l} |g(x,a)| \leq C_g(1+|x|^m)$$
$$\sup_{a\in\mathbb{R}^l} |f(x,a,0,0)| \leq C_g(1+|x|^m).$$

Sotto queste ipotesi, esiste la soluzione unica al processo in avanti per ogni dato iniziale (t, x, a). Inoltre sono soddisfatti le ipotesi (H0) sulla BSDE con vincoli della prima parte del lavoro. Queste nuove ipotesi permettono di dimostrare (in modo simile a [1]) che si può costruire una quadrupla di processi che soddisfa in modo non minimale la BSDE in (7):

Lemma 1 *Sotto le ipotesi sopracitate,* $\forall(t, x, a)$ *punto iniziale,* $(t, x, a) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^l$*, esiste una soluzione* $\left(\bar{Y}_s^{t,x,a}, \bar{Z}_s^{t,x,a}, \bar{C}_s^{t,x,a}, \bar{K}_s^{t,x,a}\right)_{t \leq s \leq T}$ *alla BSDE in* (7) $(X, X') = (X_s^{t,x,a}, X_s'^{t,x,a})_{t \leq s \leq T}$*. In questo caso* $\bar{Y}_s = \bar{v}(s, X_s^{t,x,a})$ *per una funzione deterministica* $\bar{v} : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ *che soddisfa una condizione di crescita polinomiale per qualche* $p \geq 2$ *e qualche* $C_l > 0$*:*

$$\bar{v}(t, x) \leq C_l(1 + |x|^p).$$

Tale risultato fa sì che l'ipotesi (H1) sia automaticamente verificata.
Si introduce anche qui una BSDE penalizzata (che è una BSDE classica) e, legata ad essa, una PDE semilineare *penalizzata* della forma

$$\begin{aligned} -\frac{\partial v_n}{\partial t}(t, x, a) - \mathcal{L}^a v_n(t, x, a) - \Delta_a v_n(t, x, a) \\ - f(x, a, v_n(t, x, a), \sigma^{\mathrm{T}}(x, a) D_x v_n(t, x, a)) - n\, |D_a v_n(t, x, a)| = 0 \quad (9) \end{aligned}$$

su $[0, T) \times \mathbb{R}^n \times \mathbb{R}^l$, con condizione finale

$$v_n(T, x, a) = g(x, a) \text{ su } \mathbb{R}^n \times \mathbb{R}^l,$$

dove il collegamento consiste nel fatto che la funzione

$$v_n(t, x, a) := Y_t^{n,t,x,a}$$

fornisce una rappresentazione della soluzione di (9), come dimostrato per esempio in [6]. Notiamo che la funzione v definita in (8) è il limite puntuale non-decrescente delle v_n qua definite.
Come si può vedere, il termine di penalizzazione appare davanti al gradiente in a della funzione. Idealmente si vorrebbe passare al limite nell'equazione penalizzata e concludere che la funzione v, limite di v_n, non dipende da a in quanto a gradiente nullo. Questo elemento costituisce una delle principali differenze con il lavoro [1]: invece di avere un termine penalizzante di natura integrale, ci si ritrova con un

termine penalizzante differenziale e un termine con il laplaciano in a. Questo introduce parecchie differenze nel passaggio al limite dell'equazione.
L'uso di soluzioni viscose complica leggermente il lavoro, ma è possibile a ottenere i seguenti risultati in modo analogo a [1]:

Lemma 2 *Sotto le ipotesi sopracitate, la funzione v definita in (8) è una soprasoluzione viscosa a :*

$$- |D_a v(t,x,a)| = 0 \ , \qquad (t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^q, \tag{10}$$

ossia $\forall (t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^q$ e per ogni funzione $\varphi \in C^{1,2}([0,T] \times (\mathbb{R}^d \times \mathbb{R}^q))$ tali che (t,x,a) è un punto di minimo locale per $(v - \varphi)$, si ha

$$- |D_a \varphi(t,x,a)| \geq 0 \ , \qquad i.e. \qquad D_a \varphi(t,x,a) = 0.$$

La tecnica dimostrativa è (quasi) sempre la seguente: data una funzione test φ per v in (t,x,a), si costruisce una successione di funzioni test φ_n in una successione di punti (t_n, x_n, a_n) che convergono a (t,x,a) e si trasportano le proprietà di v_n a v.
Questo lemma permette di dimostrare il seguente risultato centrale:

Teorema 1 *Sotto le ipotesi sopracitate, per ogni $(t,x) \in [0,T) \times \mathbb{R}^n$ la funzione v definita in (8) non dipende dalla variabile a su $\mathbb{R}^l$, i.e.*

$$v(t,x,a) = v(t,x,a') \quad \forall a, a' \in \mathbb{R}^l.$$

La tecnica usata è molto simile al lavoro [1], ma in questo caso, mancando la compattezza dell'insieme A, si ricorre a procedimenti diversi. Come anticipato, l'uso di minimi locali per la definizione di viscosità permette di risolvere tali problemi. In breve, si approssima la funzione v con la sua inf-convoluzione u_n e si dimostra che tale inf-convoluzione eredita la proprietà di soprasoluzione a (10). Viste la maggiore regolarità della u_n, si riesce a dimostrare che è constante in a. Si estende poi questo risultato a v.

La proprietà di soluzione Una volta affermato che v non dipende da a, si dimostra che costituisce effettivamente una soluzione viscosa della PDE. Per farlo si utilizza la stessa tecnica sopracitata, ossia la costruzione di una successione di funzioni test φ_n, al fine di trasportare la proprietà di sottosoluzione e soprasoluzione viscosa di v_n a v. La scelta delle φ_n è tuttavia in alcuni punti delicata e tecnica, in quanto esse differiscono sostanzialmente da quelle usate in [1] proprio per la presenza di termini differenziali invece che integrali. Alla fine si ottiene il seguente risultato:

Teorema 2 *La funzione v definita in* (8) *è soluzione viscosa della PDE di HJB* (1)-(2).

Nel presente lavoro non si tratta l'unicità della soluzione. Nel caso in cui l'equazione è associata a un problema di controllo con ipotesi di Lipschitzianità sui coefficienti e le funzioni costo, è provata in [7] l'esistenza e unicità della soluzione viscosa in un opportuna classe di funzioni (sublineari, Lipschitz in x e Hölderiane in t). Nel caso generale non è stata trovata in letteratura alcuna informazione sull'unicità dell'equazione di HJB con dominio $\mathbb{R}^n$ e spazio dei controlli illimitato. Resta dunque una questione aperta, che potrebbe essere affrontata con le tecniche usuali per le soluzioni viscose. Rimandiamo il lettore al testo di riferimento [9] per approfondimenti.

Conclusione È stata ottenuta una rappresentazione probabilistica per una soluzione della PDE (1)-(2). Nel caso in cui questa è associata a un problema di controllo stocastico, la soluzione della PDE è unica e costituisce il valore ottimo del problema di controllo. In tal caso, la soluzione della BSDE fornisce una rappresentazione della funzione valore. Inoltre, usando la rappresentazione duale del problema di controllo artificiale, si ha l'equivalenza fra i due problemi di controllo in quanto la funzione di valore di ciascuno di essi coincide.
Il contributo innovativo consiste nell'estensione di [1] al caso del moto Browniano, rilassando leggermente alcune ipotesi ma introducendone altre. Infatti, nel presente lavoro non è richiesta la Lipschitzianità della funzione f negli argomenti (x, a), ipotesi non molto comune, mentre le ipotesi sullo spazio dei controlli sono più delicate: da una parte si impone che lo spazio dei controlli sia tutto $\mathbb{R}^l$, dall'altro non si hanno più le limitazioni sullo spazio A introdotte dal fatto che si usa un processo di puro salto. Inoltre la proprietà di sottosoluzione è ottenuta senza ipotesi aggiuntive.
Il passo successivo potrebbe essere proprio quello di generalizzare lo spazio dei controlli. Una prima pista di ricerca è l'utilizzo di mappa biunivoca fra $\mathbb{R}^l$ e A che permetta di replicare il risultato ottenuto. E' inoltre da valutare la possibilità di costruire uno schema di simulazione di tipo Monte Carlo per simulare la BSDE con vincoli, ottenendo quindi la soluzione.

Contents

1

Introduction

The aim of this work is to provide a representation formula, using probabilistic tools, for the solution to the following equation

$$-\frac{\partial u}{\partial t} - \sup_{a\in A}\left[\mathcal{L}^a u + f(\cdot, a, u, \sigma^{\mathrm{T}}(\cdot, a)D_x u)\right] = 0 \text{ on } [0,T)\times\mathbb{R}^n, \tag{1.1}$$

$$u(T,x) = \sup_{a\in A} g(x,a) \qquad x\in\mathbb{R}^n, \tag{1.2}$$

with $\mathcal{L}^a f(t,x) = {}^1\!/\!_2\,\mathrm{tr}(\sigma\sigma^{\mathrm{T}}(x,a)D_x^2 f(t,x)) + (D_x f(t,x))^{\mathrm{T}} b(x,a)$, when the set A is the whole space $\mathbb{R}^l$. The equation above is commonly known as Hamilton-Jacobi-Bellman equation (HJB for short). By probabilistic tools we mean that the solution is given in terms of an expected value or as a solution to a Stochastic Differential Equation. In this particular case, we will use Backward Stochastic Differential Equations (BSDE).

This kind of equation arises for example in stochastic optimal control problems. Given its strongly non-linear nature, it is common to use the concept of viscosity solution for this equation. This type of solution, useful for a large class of equations, is not always easy to find. Numerical methods for them are often complex and resource intensive. Probabilistic methods, like Monte Carlo simulation, are thus competitive, especially in high dimension. The recently developed theory of BSDE is of great help in treating this kind of problem, as we will see later. In chapter 2 a quick and formal introduction to these tools and concepts is given.

Recently Kharroubi and Pham published the work [1] where, using a BSDE with constraint on the jump component, a representation formula for integro-differential

HJB equations is given. This work extends it to a different case, and establishes a similar result for the equation (1.1)-(1.2) with, as stated before, $A = \mathbb{R}^l$. Although we follow step by step the work [1], the use of a different class of constrained BSDE introduces different technical difficulties. One easily noticeable difference is that the set A in [1] is required to be compact, while here it is required to be the whole set $\mathbb{R}^l$. Through the work we will point out other key differences as we encounter them.

The idea behind the work, developed in [2] for switching type problems and in [3] for quasi variational inequalities, is to substitute the control process with a random process. Then a BSDE with constraints is associated to it and it is shown that a minimal solution to it represents a solution to the control problem.

Other approaches to the problem of representing the solution of a nonlinear PDEs through probabilistic means are present in literature. One example is [16], where a second order BSDE is introduced to represent fully non linear PDEs. The second order BSDE is closely related to the theory of nonlinear G-expectation introduced recently in [17]. As for representation of semilinear PDEs, we refer the reader to [13].

The paper is organized as follows: in chapter 2 we give a brief and basic introduction to stochastic control problems, Hamilton-Jacobi-Bellman equations, viscosity solutions and Backward Stochastic Differential Equations. If the reader is comfortable with these subjects the chapter can be skipped altogether.
In chapter 3 we introduce the class of BSDE with constraints on the Brownian part. Detailed hypotheses are introduced, and existence and uniqueness of the solution are discussed. In chapter 4 we give a different interpretation to the solution of the BSDE, in terms of solution to an auxiliary control problem. Chapter 5 contains the main results of this work: after introducing the framework, we show that the solution to the BSDE gives a representation formula for viscosity solutions to HJB equation. We first state some properties of the candidate solution, then prove it is actually a viscosity solution to the equation. We conclude in chapter 6 with some remarks and observations.

2

Mathematical setting

In this chapter we introduce the basic framework and the basic tools we will use though this work. First, we briefly introduce the concept of optimal stochastic control problem, and how their optimal value can be characterized as solution to a PDE. We then introduce the concept of viscosity solution to partial differential equations. And last we introduce the Backward Stochastic Differential Equations (BSDE for short), a class of SDE that needs a special theory.
Since we are only interested in briefly describing these concepts, the discussion that follows is purely introductory, and in some parts we will omit a detailed discussion of the needed hypotheses. A precise formulation of the mathematical tools used in this work can be found from chapter 3 onwards.

2.1 Optimal stochastic control problems

In many applications it is interesting to control the performance of a system where the state variable evolves randomly, and most of these problems can be reformulated as the problem of maximizing a gain or minimizing a cost that depends on the state of the system as well as the control action.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined a filtration $\mathbb{F} = ((\mathcal{F}_t)_{t\in[0,T]})$ for some fixed time instant T. Consider a state variable $X \in \mathbb{R}^n$ whose dynamic is described by the following stochastic differential equation

$$dX_t = b(X_t, \alpha_t)dt + \sigma(X_t, \alpha_t)dW_t, \tag{2.1}$$

where W is a Brownian motion in $\mathbb{R}^d$, $b : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^{n\times d}$ are two measurable functions. In the equation (2.1) appears the control process

$(\alpha_t)_{t\in[0,T]}$ valued in $A \subset \mathbb{R}^l$, that models the control action an agent can take on the system. The process $(\alpha_t)_{t\in[0,T]}$ is called *admissible control* if it satisfies some conditions we will see later. For now let $\mathcal{A}$ be the space of admissible controls. We call (2.1) a controlled diffusion.
Let $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be two measurable functions representing a gain. Our goal is to maximize an objective functional $J(\alpha; t, x)$ of α, of the form

$$J(\alpha; t, x) \coloneqq \mathbb{E}\left[\int_t^T f(s, X_s, \alpha_s) ds + g(X_T)\right], \tag{2.2}$$

where (t, x) are the starting time and point of the controlled diffusion. Our optimal stochastic control problem is thus to find a value function defined as

$$v(t, x) \coloneqq \sup_{\alpha\in\mathcal{A}} J(\alpha; t, x). \tag{2.3}$$

There is a large literature on this subject, we refer the reader to [5], [6], [7].

2.1.1 Hamilton-Jacobi-Bellman equation

Under the following hypothesis, the function defined in (2.3) can be characterized as solution to a PDE, the Hamilton-Jacobi-Bellman (HJB for short) equation. We require the functions $b(x, a)$ and $\sigma(x, a)$ to satisfy a Lipschitz condition on x for all $a \in A$. The set of admissible controls $\mathcal{A}$ is defined as those progressive processes α such that

$$\mathbb{E}\left[\int_0^T \left(|b(0, \alpha_s)| + |\sigma(0, \alpha_s)|\right) ds\right] < \infty.$$

Under these conditions, (2.1) admits a unique solution for every starting point (t, x). We ask g either to satisfy a quadratic growth condition or to be bounded from below. As for the conditions on f, we define $\mathcal{A}(t, x)$ a subset of controls $\alpha \in \mathcal{A}$ such that

$$\mathbb{E}\left[\int_t^T f(s, X_s^{t,x,\alpha}, \alpha_s)\right] < \infty,$$

when $X_s^{t,x,\alpha}$ is the solution to (2.1) when the starting point is (t, x) and the control α.

Remark 2.1 If f satisfies a quadratic growth condition on x, we have that $\mathcal{A}(t, x) = \mathcal{A}$.

Under these conditions the value function v of the control problem

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^T f(s, X_s, \alpha_s)ds + g(X_T)\right] \tag{2.4}$$

can be characterized as solution to the following PDE:

$$-\frac{\partial v}{\partial t} - \sup_{a \in A}\left[\mathcal{L}^a v(t,x) + f(t,x,a)\right] = 0 \text{ on } [0,T] \times \mathbb{R}^n, \tag{2.5}$$

$$v(T,x) = g(x) \text{ on } \mathbb{R}^n, \tag{2.6}$$

where $\mathcal{L}^a f(x) = {}^1\!/_2 \operatorname{tr}\left[\sigma\sigma^{\mathrm{T}}(x,a) D_x^2 f(x)\right] + b(x,a)^{\mathrm{T}} D_x f(x)$.
When there is a $C^{1,2}([0,T] \times \mathbb{R}^n)$ function w that solves (2.5)-(2.6), satisfying a quadratic growth condition, then it is also the value function to the associated control problem (see [6]).

2.2 HJB PDE and viscosity solutions

(2.5)-(2.6) does not always admit a solution in the classical sense. It is a strongly nonlinear equation, especially if the diffusion coefficient σ depends on the control action. When dealing with this kind of equations, it is not unusual to work with viscosity type solutions. The concept of viscosity solution was first introduced in the early 1980s by Pierre-Louis Lions and Michael G. Crandall in [8]. A detailed introduction to viscosity solutions can be found in [9], we outline here only the main concepts in this particular "parabolic" case. The equation must be of the form

$$F\left(t, x, v(t,x), \frac{\partial v}{\partial t}(t,x), D_x v(t,x), D_x^2 v(t,x)\right) = 0 \qquad (t,x) \in [0,T) \times \mathcal{O}, \tag{2.7}$$

where $\mathcal{O}$ is an open subset of $\mathbb{R}^n$ and the function F on $[0,T) \times \mathcal{O} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n \times n}$ satisfy an ellipticity and parabolicity condition:

$$M \leq \bar{M} \Rightarrow F(t,x,r,q,P,M) \geq F(t,x,r,q,P,\bar{M}) \text{ (ellipticity)},$$
$$q \leq \bar{q} \Rightarrow F(t,x,r,q,P,M) \geq F(t,x,r,\bar{q},P,M) \text{ (parabolicity)},$$

where the ordering $M \leq \bar{M}$ in the space of symmetric $n \times n$ matrices means that $\bar{M} - M$ is a positive semidefinite matrix.

It is clear that the equation (2.5) falls in this category with F defined as

$$F(t,x,r,q,P,M) = -q - \sup_{a\in A}\left[\frac{1}{2}\operatorname{tr}(\sigma\sigma^{\mathrm{T}}(x,a)M) + P^{\mathrm{T}}b(x,a) + f(x,a)\right].$$

The idea is that a discontinuous function w can be the solution to the equation above if a class of regular test functions satisfies the equation in a classical sense. First we recall the definition of semi-continuous envelope:

Definition **2.1** Let u be a locally bounded function on $[0,T)\times\mathbb{R}^n$. The greatest lower semi-continuous function smaller than u is called its lower semi-continuous envelope and its denoted by u_*

$$u_*(t,x) = \liminf_{\substack{(t',x')\to(t,x)\\ t'<T}} u(t',x') \qquad \forall (t,x)\in[0,T]\times\mathbb{R}^n.$$

Definition **2.2** Let u be a locally bounded function on $[0,T)\times\mathbb{R}^n$. The smallest upper semi-continuous function greater than u is called its upper semi-continuous envelope and its denoted by u^*

$$u^*(t,x) = \limsup_{\substack{(t',x')\to(t,x)\\ t'<T}} u(t',x') \qquad \forall (t,x)\in[0,T]\times\mathbb{R}^n.$$

We can now give the definition of viscosity supersolution and subsolution to (2.7):

Definition **2.3** A locally bounded, lower (*resp upper*) semi-continuous function w is said to be a viscosity supersolution (*resp subsolution*) to (2.7) if for every point $(\bar{t},\bar{x})\in[0,T)\times\mathcal{O}$ and for every $\varphi\in C^{1,2}([0,T]\times\mathbb{R}^n)$ such that $(w-\varphi)(\bar{t},\bar{x})$ is a local minimum (*resp maximum*), it holds that

$$F\left(\bar{t},\bar{x},w(\bar{t},\bar{x}),\frac{\partial\varphi}{\partial t}(\bar{t},\bar{x}),D_x\varphi(\bar{t},\bar{x}),D_x^2\varphi(\bar{t},\bar{x})\right) \underset{(\text{resp } \leq)}{\geq} 0.$$

Definition **2.4** A function v is a viscosity solution to (2.7) if v_* is a supersolution and v^* is a subsolution.

Remark 2.2 The notion of viscosity solution is consistent with the notion of classical solution, as shown in [9].

If the functions $b(x,a), \sigma(x,a), f(x,a), g(x)$ are Lipschitz on x and bounded on a, the value function (2.4) is a continuous viscosity solution to the HJB equation defined in (2.5)-(2.6) as shown in [6],[7]. Under the same hypotheses, in [7], it is also shown that the solution is unique in the class of functions u satisfying

$$|u(s,y)| \leq K(1+|y|) \qquad \forall (s,y) \in [0,T] \times \mathbb{R}^n,$$
$$|u(s,y) - u(\hat{s},\hat{y})| \leq K \left\{ |y-\hat{y}| + (1+|y| \vee |\hat{y}|) \, |s-\hat{s}|^{1/2} \right\}, \forall s,\hat{s} \in [0,T],\ y,\hat{y} \in \mathbb{R}^n.$$

Even if the existence and uniqueness of a solution has been determined, it is not simple to find the solution as typical simulation methods do not work with viscosity solutions and specific methods are needed, which are often complex and resource intensive. Thus it would be interesting to find a probabilistic representation of the solution, in terms of an expected value or as a solution to a SDE, especially for high dimension problems.

2.3 Backward Stochastic Differential Equations

Of particular interest in the representation of solutions to "parabolic" equations are the Backward Stochastic Differential Equations, a class of SDE first introduced by Peng and Pardoux in [11], with final condition, of the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dWs. \tag{2.8}$$

The fact a final condition is given introduces measurability issues, since we want the solution to be adapted to an increasing filtration and non-anticipating. That is why the solution of a BSDE is a pair of processes (Y,Z), where the process Z in $\mathbb{R}^d$ allows Y to be adapted.
In (2.8), W is a Brownianm motion, the function f is called the generator and ξ is the final condition, which can be a random variable. We fix as filtration $\mathcal{F}^W$, the one generated by W augmented with the $\mathbb{P}$-null sets.

Let $\mathbf{S^2(0,T)}$ denote the space of progressive processes Y such that

$$\mathbb{E}\left[\sup_{0\leq t \leq T} |Y_t|^2 \right] < \infty$$

and $\mathbf{L}^p([0,T])$ the set of progressive processes such that

$$\mathbb{E}\left[\int_0^T Z_t^p dt\right] < \infty.$$

***Definition* 2.5** Given a pair (ξ, f), a solution to the BSDE (2.8) is a pair of processes $(Y, Z) \in \mathbf{S^2(0,T)} \times \mathbf{L}^2([0,T])$ that satisfy it.

Proposition 2.1 *Under the following hypotheses*

- $\xi \in \mathbf{L}^2(\Omega)$ *and it is* $\mathcal{F}_T$ *measurable*
- $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ *is*
- $\forall (y,z)(\omega,t) \to f(t,y,z)$ *is a progressive process*
- $\mathbb{E}\left[\int_0^T |f(t,0,0)|^2 dt\right] < \infty$
- f *is Lipschitz continuous on* (y,z) $dt \times d\mathbb{P}$ *a.e.*

there exists a unique solution (Y,Z) *to the BSDE* (2.8).

For a proof see [6]. In the last few years, BSDE have known a great research impulse and their theory has been greatly developed thanks to their applications to several problems in mathematics.

3

Constrained Backward Stochastic Differential Equation

In this chapter we want to find a minimal solution to the following problem

$$Y_t = \xi + \int_t^T \bar{F}(s, Y_s, Z_s)ds + K_T - K_t - \int_t^T Z_s dW_s, \tag{3.1}$$

where the solution is a triple of processes (Y, Z, K) in a space we will introduce in the following section. The difference with a standard BSDE is that we require initial data ξ and the generator function $\bar{F}$ to be adapted to a filtration generated by a Brownian motion with more components than W. The term K_t is a non-decreasing process that allows the measurability constraint to be satisfied. It is possible to write the equation above in a different form, introducing a fourth process with the constraint that it must be equal to zero at all time. We obtain thus the following form for the equation above:

$$\begin{cases} Y_t = \xi + \int_t^T F(s, Y_s, Z_s, Cs)ds + K_T - K_t - \int_t^T Z_s dW_s - \int_t^T C_s dB_s & \text{(3.2a)} \\ |C_t| = 0 \qquad \forall 0 \leq t \leq T, & \text{(3.2b)} \end{cases}$$

where B is another Brownian motion independent from W. Most of the times it will be simpler to work with the latter formulation, especially when we want to emphasize the role of the constraint. Different ways to reformulate this problem can be found in literature. In [13], the authors study a similar equation using a projection operation, whereas in [18] a similar problem where the solution is

constrained to a convex set is analysed.
Since (3.1) is not a classical BSDE, it must be show that it admits a unique solution (in a sense we will precise later). This is obtained with a method used in [15], where a sequence of BSDE with a penalizing term converging to (3.1) is constructed.

3.1 Setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Fix $T < \infty$ a finite time instant. Let W_t and B_t be two independent Brownian motions, d-dimensional and l-dimensional respectively. Let $\widetilde{W}_t$ be the $(d+l)$-dimensional Brownian motion $(W_t, B_t)^{\mathrm{T}}$. Let $\mathcal{F}_t^{\mathbf{W}}$, $\mathcal{F}_t^{\mathbf{B}}$ and $\mathcal{F}_t^{\widetilde{\mathbf{W}}}$ the filtrations generated by W_t, B_t and $\widetilde{W}_t$ respectively.
We introduce the following spaces

$$\mathbf{S^2(0,T)} = \left\{ (Y_t)_{0\le t\le T} \in \mathbb{R}\ \mathcal{F}_t^{\widetilde{\mathbf{W}}}\text{-adapted and càdlàg such that } \|Y\|_{\mathbf{S^2}} < \infty \right\},$$
$$\mathbf{L}^p(\mathbf{0},\mathbf{T};\mathbb{R}^n) = \left\{ (A_t)_{0\le t\le T} \in \mathbb{R}^n\ \mathcal{F}_t^{\widetilde{\mathbf{W}}}\text{-adapted and càdlàg such that} \|A\|_{\mathbf{L}^p} < \infty \right\},$$
$$\mathbf{K}^2 = \left\{ K \in \mathbf{S^2} \text{ such that } K_0 = 0 \text{ and } K_t \text{ is nondecreasing } \right\},$$

where the norms are defined as

$$\|Y\|_{\mathbf{S^2}} = \left(\mathbb{E}\left[\sup_{0\le t\le T} |Y_t|^2 \right] \right)^{1/2},$$
$$\|X\|_{\mathbf{L}^p} = \left(\mathbb{E}\left[\int_0^T |X_t|^p \, dt \right] \right)^{1/p}.$$

We also introduce the progressive σ-algebra $\mathcal{P}$, i.e. the σ-algebra generated by the progressive processes on $\Omega \times [0,T]$.
We fix ξ random variable in $\mathbb{R}$ and a function $F : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^l$ on which we put the following hypothesis:

[H0]

i) $\mathbb{E}\left[|\xi|^2 \right] < \infty$

ii) ξ is $\mathcal{F}_T^{\widetilde{\mathbf{W}}}$-measurable

iii) F is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^l)$ measurable.

iv) F satisfies a uniform Lipschitz condition: $\exists C_F > 0$ such that

$$|F(t,y,z,c) - F(t,y',z',c')| \le C_F \left(|y-y'| + |z-z'| + |c-c'|\right)$$

$\forall y, y' \in \mathbb{R}$, $\forall z, z' \in \mathbb{R}^d$ and $\forall c, c' \in \mathbb{R}^l$

v) $\mathbb{E}\left[\int_0^T |F(t,0,0,0)|^2 \, dt\right] < \infty$.

Now that we have a setting for the BSDE in (3.2a), we need a concept of solution for it. We consider the minimal solution in the sense specified by the following definition.

***Definition* 3.1** A minimal solution to (3.2a) and (3.2b) with generator F and terminal condition ξ is a quadruple of processes $(Y, Z, C, K) \in \mathbf{S^2} \times \mathbf{L}^2(\mathbf{0}, \mathbf{T}; \mathbb{R}^n) \times \mathbf{L}^2(\mathbf{0}, \mathbf{T}; \mathbb{R}^l) \times \mathbf{K^2}$ satisfying (3.2a) and (3.2b) such that for every other quadruple $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right) \in \mathbf{S^2} \times \mathbf{L}^2(\mathbf{0}, \mathbf{T}; \mathbb{R}^n) \times \mathbf{L}^2(\mathbf{0}, \mathbf{T}; \mathbb{R}^l) \times \mathbf{K^2}$ satisfying (3.2a) and (3.2b), we have that

$$Y_t \le \bar{Y}_t, \qquad 0 \le t \le T \text{ a.s.}$$

This kind of problem, a BSDE with a positive constraint on the solution, is analysed in [15]. In addition to the hypothesis (H0), the following hypothesis is needed

[H1] There exists a quadruple $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right)$ that satisfies (3.2a) and (3.2b)

Under (H0) and (H1), in the last part of [15] it is shown that there exists a minimal solution (Y, Z, A), in the sense specified above, to the following problem

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s) ds + A_T - A_t - \int_t^T Z_s dW_s,$$

subject to the constraint

$$\phi(s, Y_s, Z_s) = 0 \text{ a.e., a.s.}$$

Here ϕ is a non negative function, Lipschitz on (y, z) and such that, for any fixed (y, z), $\phi(\cdot, y, z) \in \mathbf{L}^2(\mathbf{0}, \mathbf{T}; \mathbb{R})$. The function g is the generator function with the same properties as F in (H0), ξ a square integrable random variable and A a nondecreasing process.
Y is the limit of a nondecreasing sequence of processes Y^n in $\mathbf{S^2}(\mathbf{0}, \mathbf{T})$, solution to a BSDE of the form

$$Y_t^n = \xi + \int_t^T g(s, Y_s^n, Z_s^n)ds + A_T^n - A_t^n - \int_t^T Z_s^n dW_s,$$

where A_t^n is the penalizing term defined as

$$A_t^n := n \int_0^t \phi(s, Y_s^n, Z_s^n)ds.$$

3.2 The penalized BSDE

As anticipated in the last section, to prove the existence of a minimal solution to (3.2a) we need to introduce a sequence of backward SDEs with a penalizing term that reflects the constraint. In our case this yields

$$Y_t^n = \xi + \int_t^T F\left(s, Y_s^n, Z_s^n, C_s^n\right) ds + K_T^n - K_t^n - \int_t^T Z_s^n dW_s - \int_t^T C_s^n dB_s, \quad (3.3)$$

where K_t^n is the penalizing term, defined as

$$K_t^n = n \int_0^t |C_s^n| \, ds.$$

The reasoning used in [15] is the following: (3.3) is a classical BSDE that admits a solution (Y^n, Z^n, C^n) with Y^n and K^n continuous. Through standard comparison theorems for BSDEs, it is possible to show that $Y_t^n \leq Y_t^{n+1}$ for all n and that $Y_t^n \leq \bar{Y}_t$ from (H1) for all n. Thus the Y^n part of the solutions forms a sequence converging monotonically to some Y. Thanks to the hypothesis on ξ and F, and to the hypothesis (H1) it is possible to show that the sequence Y^n is uniformly bounded on $\mathbf{S^2}$ and so is its increasing pointwise limit Y. It is then possible to apply the main theorem in [15] and state that Y solves (3.1) and the penalized terms Y^n, Z^n, C^n and K^n converge to Y, Z, C and K. So we have the following result:

Proposition 3.1 *Under (H0) and (H1) the following results hold:*

1. *The constrained BSDE* (3.3) *admits a unique solution* (Y^n, Z^n, C^n, K^n) *in* $\mathbf{S^2(0,T)} \times \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^n) \times \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l) \times \mathbf{K}^2$ *for all* n

2. *The components of the penalized solution* $(Y^n, Z^n, C^n, K^n)_{n\geq 1}$ *are bounded, uniformly on* n*, by* ξ*,* F *and* $\bar{Y}$ *in the following way*

$$\|Y^n\|_{\mathbf{S}^2} + \|Z^n\|_{\mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^n)} + \|C^n\|_{\mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l)} + \|K^n\|_{\mathbf{K}^2} \leq C\left(\mathbb{E}\left[|\xi|^2\right] + \mathbb{E}\left[\int_0^T |F(s,0,0,0)|^2\, ds\right] + \mathbb{E}\left[\sup_{0\leq t\leq T}\left|\bar{Y}_t\right|^2\right]\right), \quad (3.4)$$

where C *depends only on the Lipschitzianity condition on* F *and on* T*.*

3. *The sequence* $(Y^n, Z^n, C^n, K^n)_{n\geq 1}$ *converges to a* (Y, Z, C, K) *minimal solution to* (3.2a)-(3.2b) *in* $\mathbf{S^2(0,T)} \times \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^n) \times \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l) \times \mathbf{K^2}$*. More precisely,* Y_t *is the monotonic limit of* Y^n_t*,* Z *and* C *are the weak (resp. strong) limits in* $\mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^n)$ *and* $\mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l)$ *(resp. in* $\mathbf{L}^p(\mathbf{0},\mathbf{T};\mathbb{R}^n)$ *and* $\mathbf{L}^p(\mathbf{0},\mathbf{T};\mathbb{R}^l)$ *for* $p < 2$*) of* Z^n *and* C^n*, and* K *is the weak limit of* K^n *in* $\mathbf{L}^2$*.*

3.3 Uniqueness of the solution

The work [15] does not address the uniqueness of the solution to the problem with constraints. In this case, it is pretty easy to show that the solution is unique.

Proposition 3.2 *If a minimal solution to* (3.2a)-(3.2b) *exists, it is also unique.*

Proof. Let (Y, Z, C, K) and (Y^1, Z^1, C^1, K^1) be two solutions to (3.2a)-(3.2b), and let $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right)$ be their difference.
Let us address the components separately. It is clear that, by definition of minimal solution, $\bar{Y} = 0$. And, by the constraint on the problem, $\bar{C} = 0$. $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right)$ satisfies the following equation

$$0 = \bar{Y}_t = \int_t^T \left(F(s, Y_s, Z_s, C_s) - F(s, Y^1_s, Z^1_s, C^1_s)\right) ds - \int_t^T \bar{Z}_s dW_s + \bar{K}_T - \bar{K}_t \quad (3.5)$$

for $0 \leq t \leq T$. By taking the joint quadratic variation with the Brownian motion W we get

$$0 = \langle \int_t^T \left(F(s, Y_s, Z_s, C_s) - F(s, Y^1_s, Z^1_s, C^1_s)\right) ds + \bar{K}_T - \bar{K}_t, W\rangle - \langle \int_t^T \bar{Z}_s dW_s, W\rangle \quad 0 \leq t \leq T,$$

and since the F and K term has finite variation and thus its variation with W equals zero, we obtain

$$0 = -\langle \int_t^T \bar{Z}_s dW_s, W \rangle = -\int_t^T \bar{Z}_s ds \quad 0 \leq t \leq T.$$

Differentiating w.r.t t we obtain

$$0 = \bar{Z}_s \quad 0 \leq t \leq T \qquad \Rightarrow \qquad Z = Z^1 \quad dt \otimes d\mathbb{P} \text{ - a.s.}$$

By considering (3.5) and remembering that $(Y, Z, C) = (Y^1, Z^1, C^1)$ we have

$$\bar{K}_T - \bar{K}_t = 0 \quad \forall t \in [0, T].$$

Since $\bar{K}_0 = 0$ by definition, we have that also $\bar{K}_T = 0$ and thus

$$\bar{K}_t = 0 \qquad \forall t \in [0, T].$$

and the a.s. uniqueness of the solution is proven. □

4

Dual Control Problem

Consider again the equation (3.1) from chapter 3. We analyse here the case where F does not depend on y, z and c, which means $F(s) = f_s$ a known process. In this case, it is possible to find an explicit representation for Y_t from chapter 3 in terms of essential supremum over a family of change of probabilities.
Consider the family of processes

$$U_A = \left\{ u_s \text{ processes in } \mathbb{R}^l, \mathcal{F}_t^{\widetilde{W}}\text{-progressive, essentially bounded } \right\} \tag{4.1}$$

where, as in chapter 3, the Brownian motion $\widetilde{W}$ is defined as $\widetilde{W} = (W, B)^{\mathrm{T}}$ and consider for each element $u \in U_A$ the local martingale

$$L_t^u = \exp\left\{ \int_0^t u_s dB_s - \frac{1}{2} \int_0^t u_s^2 ds \right\}.$$

It holds that for every process in U_A, L_t^u is a martingale.

Proposition 4.1 *For every $u \in U_A$, L_t^u is actually a martingale*

Proof. Since u is bounded, we have that

$$\int_0^T u_s^2 ds \leq MT \leq K,$$

with K generic constant. Thus by Novikov's condition L_t^u is a martingale. □

Consider now the probability measure defined by

$$\left.\frac{\mathrm{d}\mathbb{Q}^u}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t^{\widetilde{W}}} = L_t^u$$

and consider also the gain functional

$$J(u;t) = \mathbb{E}^u\left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right],$$

where $\mathbb{E}^u[\cdot]$ denotes the expectation w.r.t $\mathbb{Q}^u$. We introduce the following *auxiliary* control problem:

$$\text{Maximize } J(u;t) \text{ over the set of processes in } U_A \tag{4.2}$$

We will show now that Y_t solution to the BSDE in chapter 3 is the optimal value to (4.2), i.e.

$$Y_t = \operatorname*{ess\,sup}_{u\in U_A} \mathbb{E}^u\left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right].$$

Let us see how.
First we analyse what happens to W_t and B_t after a change of probability. Fix u in U_A and consider the process in $\mathbb{R}^{d+l}$

$$\tilde{u}_s = \begin{pmatrix} 0_d \\ u_s \end{pmatrix}.$$

We have that

$$L_t^{\tilde{u}} = \exp\left\{\int_0^t \tilde{u}_s d\widetilde{W}_s - \frac{1}{2}\int_0^t |\tilde{u}_s|^2 ds\right\} = \exp\left\{\int_0^t u_s dB_s - \frac{1}{2}\int_0^t |u_s|^2 ds\right\} = L_t^u,$$

which is a martingale. Now by Girsanov's theorem

$$\widetilde{W}_t^{\tilde{u}} = \widetilde{W}_t - \int_0^t \tilde{u}_s ds$$

is a $\mathbb{Q}^u$-Brownian motion, where the probability $\mathbb{Q}^u$ is defined by

$$\left.\frac{\mathrm{d}\mathbb{Q}^u}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t^{\widetilde{W}}} = L_t^{\tilde{u}} = L_t^u.$$

It is clear that

$$\widetilde{W}_t^{\tilde{u}} = \begin{pmatrix} W_t \\ B_t^u \end{pmatrix}, \qquad \text{where } B_t^u = B_t - \int_0^t u_s ds$$

and W_t is the same as before. In conclusion W_t is still a d-dimensional Brownian motion on $[0, T]$ under $\mathbb{Q}^u$, and B_t^u is a Brownian motion in $\mathbb{R}^l$ under $\mathbb{Q}^u$ (B_t is not), and, more importantly, they are still independent.
We introduce now a subset of U_A

$$U_A^n = \{u \in U_A \text{ such that } |u_t| \leq n \qquad \text{a.s., a.e. } t \in [0, T]\}. \tag{4.3}$$

We start by stating a result on the solution to the penalized BSDE (3.3)

Proposition 4.2 *For every $n \in \mathbb{N}$, the solution Y_t^n to the penalized* (3.3) *is represented as*

$$Y_t^n = \operatorname*{ess\,sup}_{u \in U_A^n} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]. \tag{4.4}$$

Proof. We have that Y_t^n satisfies

$$Y_t^n = \xi + \int_t^T (f_s + n\,|C_s^n|)\, ds - \int_t^T Z_s^n dW_s - \int_t^T C_s^n dB_s. \tag{4.5}$$

Fix $u \in U_A^n$ and consider the probability measure $\mathbb{Q}^u$. Under $\mathbb{Q}^u$, substituting the b.m. B, (4.5) becomes

$$Y_t^n = \xi + \int_t^T f_s + \int_t^T (n\,|C_s^n| - u_s C_s^n)\, ds - \int_t^T Z_s^n dW_s - \int_t^T C_s^n dB_s^u.$$

By tanking expectation value w.r.t. $\mathbb{Q}^u$ conditional on $\mathcal{F}_t^{\widetilde{W}}$ we obtain

$$\begin{aligned} Y_t^n = \mathbb{E}^u &\left[\xi + \int_t^T f_s + \int_t^T (n\,|C_s^n| - u_s C_s^n)\, ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right] \\ &- \mathbb{E}^u \left[\int_t^T Z_s^n dW_s \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right] - \mathbb{E}^u \left[\int_t^T C_s^n dB_s^u \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]. \end{aligned} \tag{4.6}$$

Now both

$$M_t^C = \int_0^t C_s^n dB_s^u \qquad \text{and} \qquad M_t^Z = \int_0^t Z_s^n dW_s$$

are $\mathbb{Q}^u$-martingales. In order to show this, consider $\varphi \in \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^n)$ and $\psi \in \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l)$. Since $\int_0^t \psi_s^2 ds < \infty$,

$$\int_0^t \psi_s dB_s^u$$

is a local martingale starting from zero. By Burkholder-Davis-Gundy inequality

$$\mathbb{E}^u\left[\sup_{t\in[0,T]}\left|\int_0^t \psi_s dB_s^u\right|\right] \leq C\mathbb{E}^u\left[\left(\int_0^T \psi_s^2 ds\right)^{\frac{1}{2}}\right] = C\mathbb{E}\left[L_T^u\left(\int_0^T \psi_s^2 ds\right)^{\frac{1}{2}}\right]$$
$$\leq \mathbb{E}\left[(L_T^u)^2\right]^{\frac{1}{2}}\mathbb{E}\left[\int_0^T \psi_s^2 ds\right]^{\frac{1}{2}}.$$

qich are both finite since L_t^u is a $\mathbb{P}$-martingale bounded in L^2 and $\psi \in \mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l)$, and thus $\int_0^t \psi_s dB_s$ is a martingale. The same reasoning works for ϕ. It is then clear that

$$\mathbb{E}^u\left[\int_t^T C_s^n dB_s^u \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right] = \mathbb{E}^u\left[\int_0^T C_s^n dB_s^u - \int_0^t C_s^n dB_s^u \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right] = 0$$

by the martingale property.
Coming back to (4.6) we have that

$$Y_t^n = \mathbb{E}^u\left[\xi + \int_t^T f_s ds + \int_t^T \left(n\left|C_s^n\right| - u_s C_s^n\right) ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right]. \tag{4.7}$$

Since $u \in U_A^n$, we have that

$$n\left|C_s^n\right| - u_s C_s^n \geq 0 \qquad \forall \text{ a.s., a.e.} t \in [0,T],$$

and so

$$Y_t^n \geq \mathbb{E}^u\left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right].$$

By the arbitrariness of u in U_A^n we have that

$$Y_t^n \geq \operatorname*{ess\,sup}_{u \in U_A^n} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]. \tag{4.8}$$

To prove the opposite inequality, we choose the following process in U_A^n

$$\tilde{u} = n \cdot \operatorname{sgn}(C_s^n). \tag{4.9}$$

Inserting it in (4.7) we have that

$$Y_t^n = \mathbb{E}^{\tilde{u}} \left[\xi \int_t^T f_s ds + \int_t^T (n\,|C_s^n| - \tilde{u}_s C_s^n)\, ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right] = \mathbb{E}^{\tilde{u}} \left[\xi \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]$$

by (4.9). Combining it with (4.8) we obtain that

$$Y_t^n = \operatorname*{ess\,sup}_{u \in U_A^n} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]$$

and the ess sup is attained on $\tilde{u}$. □

Now, by taking n to infinity, we can state a similar result for Y solution to (3.2a) and (3.2b)

Theorem 4.1 *Under (H0) and (H1), the minimal solution to* (3.2a)-(3.2b) *is represented as*

$$Y_t = \operatorname*{ess\,sup}_{u \in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right].$$

Proof. Let (Y, Z, C, K) be the minimal solution to (3.2a)(3.2b).
Let $\tilde{Y}$ be defined as follows

$$\tilde{Y}_t = \operatorname*{ess\,sup}_{u \in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]. \tag{4.10}$$

Since $U_A^n \subset U_A$ we have that

$$Y_t^n = \operatorname*{ess\,sup}_{u \in U_A^n} J(u;t) \leq \operatorname*{ess\,sup}_{u \in U_A} J(u;t) = \tilde{Y}_t,$$

and since $Y_t = \lim_{n\to\infty} Y_t^n$ we have

$$Y_t \le \tilde{Y}_t = \operatorname*{ess\,sup}_{u\in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right]. \tag{4.11}$$

To prove the opposite inequality, we write the equation satisfied by Y

$$Y_t = \xi + \int_t^T f_s ds + K_T - K_t - \int_t^T Z_s dW_s,$$

where the integral on B does not appear due to the constraint.
By taking the $\mathbb{Q}^u$-expectation, and remembering that the stochastic integral is a $\mathbb{Q}^u$-martingale, we obtain that

$$Y_t = \mathbb{E}^u \left[\xi + \int_t^T f_s ds + K_t - K_t \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right].$$

Since K is a non-decreasing process, it holds that

$$Y_t \ge \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right].$$

Since this holds for any u in U_A, this yields

$$Y_t \ge \operatorname*{ess\,sup}_{u\in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right],$$

and combining it with (4.11) we obtain the claim:

$$Y_t = \operatorname*{ess\,sup}_{u\in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}} \right].$$

□

Remark 4.1 Given additional conditions on the random variable ξ and the process f, it is possible to forget the need of the hypothesis (H1). If the process defined in (4.10) is in $\mathbf{S^2(0,T)}$, then we got an upper bound in $\mathbf{S^2(0,T)}$ for the sequence $(Y_n)_{n\ge1}$, and that is enough to prove the existence of the minimal solution Y. A simple condition could be requiring ξ and f_s to be bounded.

5

A representation formula for Hamilton-Jacobi-Bellman equation

In this chapter we show that the minimal solution to a BSDE with constraints, like the one in chapter 3, gives a representation formula to a Hamilton-Jacobi-Bellman nonlinear equation associated to a control problem of a forward diffusion process. As already stated, representation formulas for semilinear PDEs using BSDE are well known, like in section 5.3 of [6], but the representation of fully nonlinear equations like the HJB one is a different matter.

The idea used here, introduced by Kharroubi and Pham in [1], is to replace the control process in a controlled diffusion by a random process, in our case a standard Brownian motion. The solution of this combined diffusion acts then as terminal data and generator for a constrained BSDE as the one in chapter 3, which solved backwards provides a representation for the solution to a HJB type PDE.

5.1 Setting

As before, we introduce here two independent Brownian motions W_t and B_t, in $\mathbb{R}^d$ and $\mathbb{R}^l$ respectively, with their natural filtrations. We introduce as well as the filtration $\mathcal{F}^{\bar{W}}$ generated by the combined $\mathbb{R}^{d+l}$ Brownian motion.
Given two measurable functions

$$
\begin{aligned}
b &: \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^n \\
\sigma &: \mathbb{R}^n \times \mathbb{R}^l \longrightarrow \mathbb{R}^{n\times d},
\end{aligned}
$$

we introduce the following forward diffusion

$$\begin{cases} dX_s = b(X_s, X'_s)ds + \sigma(X_s, X'_s)dW_s \\ dX'_s = dB_s. \end{cases} \tag{5.1}$$

We put the following hypothesis on b and σ

[HFD]

(i) b and σ satisfy a Lipschitz condition, i.e.

$$\exists C > 0 \text{ s.t. } |b(x,a) - b(x',a')| + |\sigma(x,a) - \sigma(x',a')| \leq C(|x - x'| + |a - a'|)$$

$$\forall x, x' \in \mathbb{R}^n, \qquad \forall a, a' \in \mathbb{R}^l$$

(ii) $\exists M < \infty$ s.t. $\sup_{a\in\mathbb{R}^l} |b(0,a)| + \sup_{a\in\mathbb{R}^l} |\sigma(0,a)| \leq M$.

Under HFD-(i), we have the existence of a unique time continuous solution

$$(X_s^{t,x,a}, X'^{t,x,a}_s)_{t\leq s\leq T}$$

to 5.1 for any initial condition $(t,x,a) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^l$ (see [19], [20]). Moreover $\forall p \geq 2$, we have the standard estimate that there exists a C_s such that $\forall (t,x,a) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^l$:

$$\mathbb{E}\left[\sup_{t\leq s\leq T} (\left|X_s^{t,x,a}\right|^p + \left|X'^{t,x,a}_s\right|^p)\right] \leq C_s(1 + |x|^p + |a|^p). \tag{5.2}$$

We introduce now other two measurable functions f and g which can be seen as the gains in the control problem, as well as the generator and terminal value of the BSDE when composed with the processes $(X_s^{t,x,a}, X'^{t,x,a}_s)_{t\leq s\leq T}$.

$$g : \mathbb{R}^n \times \mathbb{R}^l, \longrightarrow \mathbb{R} \qquad\qquad f : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R} \times \mathbb{R}^d \longrightarrow \mathbb{R}.$$

Remark 5.1 If we are only concerned with control problems we could choose f to be a real function defied only on $\mathbb{R}^n \times \mathbb{R}^l$.

We put the following hypotheses on f and g:

[HFC]

(i) f satisfies a Lipschitz condition on y and z for some constant C, uniformly on (x, a)

$$\big|f(x, a, y, z) - f(x, a, y', z')\big| \leq C(\big|y - y'\big| + \big|z - z'\big|)$$

(ii) f and g satisfy a polynomial growth condition on x, i.e. $\exists m \geq 0$ and $\exists C_g > 0$ such that

$$\sup_{a \in \mathbb{R}^l} |g(x, a)| \leq C_g(1 + |x|^m)$$
$$\sup_{a \in \mathbb{R}^l} |f(x, a, 0, 0)| \leq C_g(1 + |x|^m).$$

5.2 The nonlinear PDE

We have all the ingredients to introduce the nonlinear PDE for which we want to find a representation formula:

$$-\frac{\partial u}{\partial t} - \sup_{a \in \mathbb{R}^l} \big[\mathcal{L}^a u + f(., a, u, \sigma^{\mathrm{T}}(., a) D_x u)\big] = 0 \qquad \text{on } [0, T) \times \mathbb{R}^n, \quad (5.3)$$

$$u(T, x) = \sup_{a \in \mathbb{R}^l} g(x, a) \qquad x \in \mathbb{R}^n, \quad (5.4)$$

where the operator $\mathcal{L}^a$ is

$$\mathcal{L}^a u(t, x) = D_x^{\mathrm{T}} u(t, x) b(x, a) + \frac{1}{2} \operatorname{tr}(\sigma\sigma^{\mathrm{T}}(x, a) D_x^2 u(t, x)).$$

We search a probabilistic representation for a viscosity solution to this equation. We recall here the definition of viscosity solution to the equation (5.3)-(5.4)

***Definition* 5.1** A lower semi-continuous function $u : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ is said to be a viscosity supersolution to (5.3)-(5.4) if

- $u(T, x) \geq \sup_{a \in \mathbb{R}^l} g(x, a) \qquad \forall x \in \mathbb{R}^n.$

- For all $(t, x) \in [0, T) \times \mathbb{R}^n$ and for every $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$ such that $u - \varphi$ has a local minimum in (t, x) then

$$-\frac{\partial \varphi}{\partial t}(t, x) - \sup_{a \in \mathbb{R}^l} \big[\mathcal{L}^a \varphi(t, x) + f(x, a, u(t, x), \sigma^{\mathrm{T}}(., a) D_x \varphi(t, x))\big] \geq 0 \quad (5.5)$$

on $[0, T) \times \mathbb{R}^n$.

***Definition* 5.2** An upper semi-continuous function $u : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ is said to be a viscosity subsolution to (5.3)-(5.4) if

- $u(T, x) \leq \sup_{a \in \mathbb{R}^l} g(x, a) \qquad \forall x \in \mathbb{R}^n.$
- For all $(t, x) \in [0, T) \times \mathbb{R}^n$ and for every $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$ such that $u - \varphi$ has a local maximum in (t, x) then

$$-\frac{\partial \varphi}{\partial t}(t, x) - \sup_{a \in \mathbb{R}^l} \left[\mathcal{L}^a \varphi(t, x) + f(x, a, u(t, x), \sigma^{\mathrm{T}}(., a) D_x \varphi(t, x)) \right] \leq 0 \quad (5.6)$$

on $[0, T) \times \mathbb{R}^n$.

***Definition* 5.3** A locally bounded function $u : [0, T) \times \mathbb{R}^n \to \mathbb{R}$ is said to be a viscosity solution to (5.3)-(5.4) its lower semi-continuous envelope u_* is a viscosity supersolution to (5.3)-(5.4) and its upper semi-continuous envelope u^* is a viscosity subsolution to (5.3)-(5.4).

Remark 5.2 Since we are dealing only with differential terms, we can use local minima (maxima) for the viscosity solutions definitions.

5.3 The constrained BSDE

To the PDE (5.3)-(5.4) we will associate a constrained BSDE similar to that in chapter 3. In this contest, the constrained BSDE is

$$Y_t = g(X_T, X'_T) + \int_t^T f(X_s, X'_s, Y_s, Z_s) ds + K_T - K_t - \int_t^T Z_s dW_s - \int_t^T C_s dB_s, \quad (5.7)$$

with the constraint

$$|C_s| = 0 \quad t \leq s \leq T. \quad (5.8)$$

As in chapter 3, we want to find a minimal solution in the sense of definition 3.1. For any initial condition (t, x, a) for the diffusion, the final data $g(X_T, X'_T)$ and the generator $f(X_t, X'_t, \cdot, \cdot)$ satisfy the hypotheses (H0). Indeed the final data $g(X_T, X'_T)$ is in L^2 since

$$\begin{aligned}\mathbb{E}\left[|g(X_T, X'_T)|^2\right] &\leq \mathbb{E}\left[\sup_{a\in\mathbb{R}^l}|g(X_T, a)|^2\right]\\ &\leq \mathbb{E}\left[2C_g(1+|X_T|^{2m})\right]\\ &\leq 2C_g(1+|x|^{2mp})\end{aligned}$$

thanks to HFC-(ii) and the estimate (5.2).
The reasoning to show that, under (HFD) and (HFC), f satisfies the hypothesis (H0) is analogous.
The following proposition shows that the hypothesis (H1) is no longer needed since we have that under (HFD) and (HFC) there exists a quadruple of processes $\left(\bar{Y}, \bar{Z}, \bar{C}, \bar{K}\right)$ that satisfy (5.7)-(5.8).

Lemma 5.1 *Under the assumptions (HFD) and (HFC), $\forall(t, x, a)$ starting point, $(t, x, a) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^l$, there exists a solution $\left(\bar{Y}_s^{t,x,a}, \bar{Z}_s^{t,x,a}, \bar{C}_s^{t,x,a}, \bar{K}_s^{t,x,a}\right)_{t\leq s\leq T}$ to the BSDE (5.7)-(5.8) with $(X, X') = (X_s^{t,x,a}, X_s'^{t,x,a})_{t\leq s\leq T}$. In this case $\bar{Y}_s = \bar{v}(s, X_s^{t,x,a})$ for a deterministic function $\bar{v} : [0, T] \times \mathbb{R}^n$ that satisfies a polynomial growth condition for some $p \geq 2$ and some $C_l > 0$*

$$\bar{v}(t, x) \leq C_l(1 + |x|^p).$$

Proof. In the following proof, C indicates a generic constant. Consider the function

$$\bar{v}(t, x) = \bar{C}e^{\rho(T-t)}(1 + |x|^p), \text{ with } p = \max(2, m), \tag{5.9}$$

m being the exponent in HFC-(ii). $\bar{C}$ and ρ are two constant to be specified later. We show first that (5.9) is a classical supersolution to (5.3). Consider now $\mathcal{L}^a\bar{v}$

- First we consider $D_x^{\mathrm{T}}\bar{v}(t, x)b(x, a)$. We have that

$$(D_x\bar{v})_i = \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\,x_i,$$

thus

$$\begin{aligned}
\left|D_x^{\mathrm{T}}\bar{v}(t,x)b(x,a)\right| &= \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\left|x^{\mathrm{T}}b(x,a)\right| \\
&\leq \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\,|x|\,|b(x,a)| \\
&\leq \bar{C}e^{\rho(T-t)}p\,|x|^{p-1}\,|b(0,a)+C\,|x|| \qquad (5.10)\\
&= \bar{C}e^{\rho(T-t)}p\,|x|^{p-1}\,|b(0,a)| + \bar{C}Ce^{\rho(T-t)}p\,|x|^{p} \\
&\leq C\bar{v}(t,x)\,|b(0,a)| + Cv(t,x) \leq C(1+|b(0,a)|)\bar{v}(t,x), \qquad (5.11)
\end{aligned}$$

where in (5.10) we used the Lipschitz condition in HFD-(i).

- Next we consider the f part

$$\left|f(x,a,\bar{v},\sigma^{\mathrm{T}}(x,a)D_x\bar{v}\right| \leq |f(x,a,0,0)| + C(|\bar{v}| + \left|\sigma^{\mathrm{T}}(x,a)D_x\bar{v}\right|).$$

Since $\bar{v} \geq \bar{C}$ we note that, using HFC-(ii),

$$\begin{aligned}
|f(x,a,0,0)| &\leq C_g(1+|x|^m) \leq C\bar{v} \\
|x| &\leq \bar{v}.
\end{aligned}$$

For the last term we have

$$\left(\sigma^{\mathrm{T}}D_x\bar{v}\right)_k = \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\sum_i \sigma_{i,k}x_i\ ,$$

and thus

$$\begin{aligned}
\left|\sigma^{\mathrm{T}}D_x\bar{v}\right| &= \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\left(\sum_k\left(\sum_i \sigma_{i,k}x_i\right)^2\right)^{1/2} \\
&\leq \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\,C\left(\sum_k\sum_i \sigma_{i,k}^2x_i^2\right)^{1/2} \\
&\leq \bar{C}e^{\rho(T-t)}p\,|x|^{p-2}\,C\left(\max_{i,k}\sigma_{ki}^2\right)^{1/2}|x|\,.
\end{aligned}$$

Since

$$\left(\max_{i,k} \sigma_{ki}^2\right)^{1/2} \leq |\sigma(x,a)| \leq (|\sigma(0,a)| + |x|)\,,$$

we finally have that

$$\left|\sigma^{\mathrm{T}} D_x \bar{v}\right| \leq C\left(|\sigma(0,a)| + 1\right)\bar{v},$$

$$\left|f(x,a,\bar{v},\sigma^{\mathrm{T}}(x,a)D_x\bar{v}\right| \leq C\left(1 + C(|\sigma(0,a)|)\right)\bar{v}. \tag{5.12}$$

- As for the hessian part, we have the following

$$\left(D_x^2\,|x|^p\right)_{i,j} = p(p-2)\,|x|^{p-4}\,x_i x_j + \delta_{ij} p\,|x|^{p-2}\,.$$

Let $S(x,a) = \sigma\sigma^{\mathrm{T}}(x,a)$ and $D = D_x^2\bar{v}$, we have then

$$\left|\mathrm{tr}\left(\sigma\sigma^{\mathrm{T}}(x,a)D_x^2\bar{v}\right)\right| = |\mathrm{tr}(SD)| \leq \sum_{i=1}^{n} |S_{ii}D_{ii}| \leq \max_i |S_{ii}| \sum_{i=1}^{n} |D_{ii}|\,.$$

Since

$$S_{ii} = \sum_k \sigma_{ik}^2 \leq C\,|\sigma(x,a)|^2 \leq C(|\sigma(0,a)|^2 + |x|^2)$$

and

$$\sum_i^n D_{ii} = \bar{C}e^{\rho(T-t)}\left(p(p-2)\,|x|^{p-4}\sum_i^n x_i^2 + np\,|x|^{p-2}\right),$$

we can finally state that

$$\frac{1}{2}\left|\mathrm{tr}\left(\sigma\sigma^{\mathrm{T}}(x,a)D_x^2\bar{v}\right)\right| \leq C_\sigma\left(|\sigma(0,a)|^2 + 1\right)\bar{v}. \tag{5.13}$$

By combining (5.11), (5.12) and (5.13) and using HFD-(ii) and HFC-(ii) we have that there exists a constant $\tilde{C}$ such that

$$\begin{aligned}\left|\mathcal{L}^a\bar{v} + f(x,a,\bar{v},\sigma^{\mathrm{T}}D_x\bar{v})\right| &\leq \bar{v}\sup_{a\in\mathbb{R}^l}\left[C\left(|b(0,a)| + |\sigma(0,a)| + |\sigma(0,a)|^2 + 2\right)\right]\\ &\leq \tilde{C}\bar{v},\end{aligned}$$

and thus

$$\sup_{a\in\mathbb{R}^l}\left|\mathcal{L}^a\bar{v} + f(x,a,\bar{v},\sigma^{\mathrm{T}}D_x\bar{v})\right| \le \tilde{C}\bar{v}. \tag{5.14}$$

Since for the time derivative we have

$$\frac{\partial\bar{v}}{\partial t} = -\rho\bar{C}e^{\rho(T-t)}(1+|x|^p) = -\rho\bar{v},$$

it yelds

$$-\frac{\partial\bar{v}}{\partial t} - \sup_{a\in\mathbb{R}^l}\left|\mathcal{L}^a\bar{v} + f(x,a,\bar{v},\sigma^{\mathrm{T}}D_x\bar{v})\right| \ge (\rho-\tilde{C})\bar{v},$$

and by choosing $\rho > \tilde{C}$, we get that $\bar{v}$ is a supersolution to (5.3). The choice of $\bar{C}$ is not important at this stage.
We can thus define $\left(\bar{Y},\bar{Z},\bar{C},\bar{K}\right)$ as follows:

$$\bar{Y}_t = \bar{v}(t,X_t) \qquad \bar{Y}_T = g(X_T,X'_T) \tag{5.15}$$
$$\bar{Z}_t = \sigma^{\mathrm{T}}(X_t,X'_t)D_x\bar{v}(t,X_t) \tag{5.16}$$
$$\bar{C}_t = 0 \tag{5.17}$$
$$\bar{K}_t = \int_0^t \left[-\frac{\partial\bar{v}}{\partial t}(s,X_s) - \mathcal{L}^{X'_s}\bar{v}(s,X_s) + f(X_s,X'_s,\bar{Y}_s,\bar{Z}_s)\right]ds. \tag{5.18}$$

We first note that $\bar{K}_t$ is non-decreasing since $\bar{v}$ is a supersolution. From the polynomial growth in (5.9) and (5.2) it is clear that $\bar{Y}\in\mathbf{S^2(0,T)}$, and same goes for $\bar{Z}\in\mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^n)$. $\bar{C}\in\mathbf{L}^2(\mathbf{0},\mathbf{T};\mathbb{R}^l)$ since it satisfies the constraint. As for $\bar{K}$ we have, since it starts from zero and is nondecreasing

$$\begin{aligned}\sup_{0\le t\le T}\left|\bar{K}_t\right|^2 &\le \left|\bar{K}_T\right|^2\\ &\le C\int_0^T\left[\left(\frac{\partial\bar{v}}{\partial t}(s,X_s)\right)^2 + \left(\mathcal{L}^{X'_s}\bar{v}(s,X_s)\right)^2 + \left(f(X_s,X'_s,\bar{Y}_s,\bar{Z}_s)\right)^2\right]ds.\end{aligned}$$

By the Lipschitz conditions HFD-(ii) on b and σ, the polynomial growth HFC-(ii) on (x,a) and Lipschitz conditions HFC-(i) on (y,z) conditions on f and the polynomial growth of $\bar{v}$, we get that the last term is bounded by

$$C \int_0^T (1 + |X_s|^h),$$

for some constants C and h. By using (5.2) we get that $\bar{K} \in \mathbf{K}^2$.
Now, by applying Itô's Lemma to $\bar{Y}_t = \bar{v}(t, X_t)$ we get

$$dY_t = \left(\frac{\partial \bar{v}}{\partial t}(t, X_t) + b(X_t, X'_t) D_x \bar{v}(t, X_t) + \frac{1}{2} \operatorname{tr}\left(\sigma\sigma^{\mathrm{T}}(X_t, X'_t) D_x^2 \bar{v}(t, X_t) \right) \right) dt \\ + \sigma^{\mathrm{T}}(X_t, X'_t) D_x \bar{v}(t, X_t) dW_t,$$

and by adding and subtracting $\pm f(X_t, X'_t, \bar{v}(t, X_t), \sigma^{\mathrm{T}}(X_t, X'_t) D_x \bar{v}(X_t, X'_t))$, integrating between t and T and remembering the definitions of $\bar{Y}$,$\bar{Z}$ and $\bar{K}$ we have

$$Y_T - Y_t = \underbrace{- \int_t^T - \left[\frac{\partial \bar{v}}{\partial t}(s, X_s) + \mathcal{L}^{X'_s} \bar{v}(s, X_s) + f(X_s, X'_s, \bar{Y}_s, \bar{Z}_s) \right] ds}_{-\bar{K}_T + \bar{K}_t} \\ - \int_t^T f(X_s, X'_s, \bar{Y}_s, \bar{Z}_s) ds + \int_t^T \bar{Z}_s dW_s,$$

which yields

$$\bar{Y}_t = g(X_T, X'_T) + K_T - K_t + \int_t^T f(X_s, X'_s, \bar{Y}_s, \bar{Z}_s) ds - \int_t^T \bar{Z}_s dW_s - \int_t^T \bar{C}_s dBs$$

with the constraint $\bar{C}_s = 0$ satisfied.
We thus have that the quadruple defined in (5.15)-(5.18) is a solution (not minimal) to the BSDE (5.7)-(5.8). □

From the previous lemma, for every initial condition (t, x, a), we have an upper bound for the solution and thus the hypothesis (H1) is automatically satisfied. From chapter 3, we have the existence and uniqueness of a minimal solution

$$\left(Y_s^{t,x,a}, Z_s^{t,x,a}, C_s^{t,x,a}, K_s^{t,x,a} \right)_{t \le s \le T}$$

to the BSDE (5.7)-(5.8) for any initial condition $(t, x, a) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^l$. We can thus define on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^l$.

$$v(t, x, a) = Y_t^{t,x,a}. \tag{5.19}$$

The function just defined is our candidate viscosity solution to (5.3)-(5.4).

5.4 The penalized BSDE and the associated penalized equation

From chapter 3 we have also the existence of a sequence $(Y^{n,t,x,a})_{n\geq 1}$, solution to a sequence of penalized BSDE, such that $Y_s^{n,t,x,a} \nearrow Y_s^{t,x,a}$ for all $s \in [0, T]$, when the processes X and X' are given by $\left(X_s^{t,x,a}, {X'}_s^{t,x,a}\right)$.

In this setting, the penalized BSDE associated to (5.7) is:

$$Y_t^n = g(X_T, X_T') + \int_t^T f(X_s, X_s', Y_s^n, Z_s^n)ds + n \int_t^T |C_s^n|\, ds \\ - \int_t^T Z_s^n dW_s - \int_t^T C_s^n dB_s. \tag{5.20}$$

Let $(Y_s^{n,t,x,a}, Z_s^{n,t,x,a}, C_s^{n,t,x,a})_{t\leq s\leq T}$ be the solution when

$$(X, X') = \left(X_s^{t,x,a}, {X'}_s^{t,x,a}\right)_{t\leq s\leq T}.$$

Define now

$$v_n(t, x, a) := Y_t^{n,t,x,a}. \tag{5.21}$$

From chapter 3 we know that

$$v_n(t, x, a) = Y_t^{n,t,x,a} \nearrow Y_t^{t,x,a} = v(t, x, a)$$

on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^l$.
From (3.4) we know that the solution Y^n is bounded in $\mathbf{S^2(0, T)}$ and, using HFC-

(ii) and (5.2), we obtain

$$
\begin{aligned}
|v_n(t,x,a)|^2 &\leq C\left(\mathbb{E}\left[g(X_T^{t,x,a}, X_t'^{t,x,a})^2\right] + \mathbb{E}\left[\int_t^T f\left(X_s^{t,x,a}, X_s'^{t,x,a}, 0, 0\right)^2 ds\right] + \right.\\
&\left. + \mathbb{E}\left[\sup_{t\leq s\leq T}\left|\bar{v}(s, X_s^{t,x,a})\right|^2\right]\right) \qquad \forall (t,x,a)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^l\\
&\leq \tilde{C}\left(\mathbb{E}\left[1+|X_T|^{2m}\right] + \mathbb{E}\left[\int_t^T(1+|X_T|^{2m})\right] + \mathbb{E}\left[\sup_{t\leq s\leq T}(1+|X_T|^{2m})\right]\right)\\
&\leq C(1+|x|^{\bar{p}}+|a|^{\bar{p}}),
\end{aligned}
$$

for some $\bar{p}\geq 2$. By taking limits we obtain

$$
|v_n(t,x,a)| + |v(t,x,a)| \leq C(1+|x|^p+|a|^p), \tag{5.22}
$$

which tells us that both v_n and v satisfy a polynomial growth condition for some $p\geq 2$.
It is well known (Theorem 5.3.3 in [6]) that (5.21) is a viscosity solution to the following PDE:

$$
\begin{aligned}
-\frac{\partial v_n}{\partial t}(t,x,a) &- \mathcal{L}^a v_n(t,x,a) - \Delta_a v_n(t,x,a)\\
&- f(x,a,v_n(t,x,a),\sigma^{\mathrm{T}}(x,a)D_x v_n(t,x,a)) - n\,|D_a v_n(t,x,a)| = 0
\end{aligned} \tag{5.23}
$$

on $[0,T)\times\mathbb{R}^n\times\mathbb{R}^l$ with final condition

$$
v_n(T,x,a) = g(x,a) \text{ on } \mathbb{R}^n\times\mathbb{R}^l, \tag{5.24}
$$

which we can call the penalized PDE. More explicitly we have the following result:

Proposition 5.1 *$v_n(t,x,a)$ is a continuous viscosity solution to the penalized PDE* (5.23)-(5.24)

The whole point is now to work on v by using what we know on v_n. Ideally, the gradient term on (5.23) tells us that the first derivative with respect to a of the sequence v_n converges to zero, and thus that $|D_a v| = 0$, at least in the viscosity sense. This will allow us to show that the function v does not depend on the variable a and it can be a solution to (5.3)-(5.4).

5.5 v does not depend on a

We show in this section that the function v defined in (5.19) is constant on the variable a and thus we can consider it as not dependent on it. Since we are not dealing with classical derivatives, but with partial differential equations in the viscosity sense, it will be slightly more complicated. Notice that we do not know if the function v is continuous, but we do know that is l.s.c. as monotonic increasing pointwise limit of a sequence of continuous functions.

We shall now prove the non dependence of v on a. To this end we introduce the equation

$$-|D_a v(t,x,a)| = 0\ , \qquad (t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^q. \tag{5.25}$$

We show that v defined as in (5.19)

$$v(t,x,a) = Y_t^{t,x,a}$$

is a viscosity supersolution to it.
From now on, we'll use the notation $B_r(x)$ and $\overline{B_r(x)}$ to indicate respectively the open and closed ball with centre x and radius r.

Lemma 5.2 *Under the assumptions (HFD) and (HFC), the function v defined in* (5.19) *is a viscosity supersolution to* (5.25)*: $\forall (t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^q$ and for any function $\varphi \in C^{1,2}([0,T] \times (\mathbb{R}^d \times \mathbb{R}^q))$ such that $(v-\varphi)(t,x,a)$ is a local minimum, we have*

$$-|D_a\varphi(t,x,a)| \geq 0\ , \qquad i.e. \qquad D_a\varphi(t,x,a) = 0.$$

Proof. v is the pointwise limit of the nondecreasing sequence of continuous functions v_n. This means v is lsc and we have (see [10]):

$$v = v_* = \liminf_{n\to\infty} {}_* v_n, \tag{5.26}$$

where

$$\liminf_{n\to\infty} {}_* v_n(t,x,a) := \liminf_{\substack{n\to\infty \\ (t',x',a')\to(t,x,a) \\ t'<T}} v_n(t',x',a') \qquad (t,x,a) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^l.$$

Let $(t, x, a) \in [0, T) \times \mathbb{R}^d \times \mathbb{R}^q$, and $\varphi \in C^{1,2}([0, T] \times (\mathbb{R}^d \times \mathbb{R}^q))$, such that $(v - \varphi)(t, x, a)$ is a local minimum:

$$(v - \varphi)(t, x, a) = \min_{(s,y,b) \in \overline{B_r(t,x,a)}} (v - \varphi)(s, y, b)$$

on the closed neighbourhood $\overline{B_r(t, x, a)}$ for some $r > 0$. We can assume that the minimum is strict. If we consider the continuous function $v_n - \varphi$ on the compact $\overline{B_r(t, x, a)}$, there exists $\forall n$ a triplet $(t_n, x_n, a_n) \in \overline{B_r(t, x, a)}$ such that

$$(v_n - \varphi)(t_n, x_n, a_n) = \min_{(s,y,b) \in \overline{B_r(t,x,a)}} (v_n - \varphi)(s, y, b).$$

The (t_n, x_n, a_n) form a bounded sequence in $\overline{B_r(t, x, a)}$ and thus, up to a subsequence,

$$(t_n, x_n, a_n) \longrightarrow (\bar{t}, \bar{x}, \bar{a}).$$

We have that $(\bar{t}, \bar{x}, \bar{a}) = (t, x, a)$. Indeed, since $(\bar{t}, \bar{x}, \bar{a})$ is a minimum point for v_n it holds that

$$\begin{aligned}(v_n - \varphi)(t_n, x_n, a_n) &\leq (v_n - \varphi)(t, x, a) \\ &\leq (v - \varphi)(t, x, a).\end{aligned}$$

Suppose that $(\bar{t}, \bar{x}, \bar{a}) \neq (t, x, a)$, then

$$\begin{aligned}(v - \varphi)(\bar{t}, \bar{x}, \bar{a}) = \liminf_{n \to \infty} {}_*(v_n - \varphi)(\bar{t}, \bar{x}, \bar{a}) &\leq \liminf_{n \to \infty} (v_n - \varphi)(t_n, x_n, a_n) \\ &\leq (v - \varphi)(t, x, a),\end{aligned}$$

which is a contradiction since the minimum of $v - \varphi$ on (t, x, a) is strict, and thus $(\bar{t}, \bar{x}, \bar{a}) = (t, x, a)$. We also obtain from the last inequalities that

$$(v - \varphi)(t, x, a) \leq \liminf_{n \to \infty} (v_n - \varphi)(t_n, x_n, a_n) \leq (v - \varphi)(t, x, a),$$

that means that $\liminf_{n \to \infty} (v_n - \varphi)(t_n, x_n, a_n) = (v - \varphi)(t, x, a)$ and up to another subsequence

$$v_n(t_n, x_n, a_n) \to v(t, x, a).$$

We have established that

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \longrightarrow (t, x, a, v(t, x, a)) \text{ as } n \to \infty. \qquad (5.27)$$

From the viscosity supersolution property[1] of v_n at (t_n, x_n, a_n) with the test function φ, we have

$$-\frac{\partial \varphi}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^{a_n}\varphi(t_n, x_n, a_n) - \Delta_a \varphi(t_n, x_n, a_n) - f(x_n, a_n) \\ -n\,|D_a\varphi(t_n, x_n, a_n)| \geq 0,$$

which implies

$$|D_a\varphi(t_n, x_n, a_n)| \leq \frac{1}{n}\Bigg[-\frac{\partial \varphi}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^{a_n}\varphi(t_n, x_n, a_n) - \Delta_a \varphi(t_n, x_n, a_n) \\ - f(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^{\mathrm{T}}(x_n, a_n) D_x\varphi(t_n, x_n, a_n))\Bigg].$$

Thus, by sending $n \to \infty$, since $\varphi \in C^{1,2}([0, T] \times (\mathbb{R}^d \times \mathbb{R}^q))$ we obtain

$$|D_a\varphi(t, x, a)| \leq 0 \Rightarrow -|D_a\varphi(t, x, a)| \geq 0$$

and thus v is a supersolution to (5.25). □

We notice that the equation (5.25) does not have the derivatives in t or x, so we can consider the function v as dependant on the variable a only, for any fixed (t, x). We then have the following lemma:

Lemma 5.3 *Under (HFD) and (HFC), $\forall (t, x) \in [0, T) \times \mathbb{R}^n$, the function of the variable a $v(t, x, .)$ on $\mathbb{R}^l$ is a viscosity super solution to*

$$-|D_a v(t, x, a)| = 0, \qquad a \in \mathbb{R}^l, \qquad (5.28)$$

i.e. $\forall a \in \mathbb{R}^l, \quad \forall \varphi \in C^2(\mathbb{R}^l)$ such that $(v(t, x, .) - \varphi)(a)$ is a local minimum, then

$$-|D_a\varphi(a)| = 0.$$

[1] note that since (t_n, x_n, a_n) converges to (t, x, a) the sequence is inside $B_r(t, x, a)$ for n large enough

Proof. Fix $(t,x) \in [0,T) \times \mathbb{R}^n$. Fix $a \in \mathbb{R}^l$ and $\varphi \in C^2(\mathbb{R}^l)$ such that

$$(v(t,x,.) - \varphi)(a) = \min_{a' \in B_{r'}(a)} (v(t,x,.) - \varphi)(a')$$

for some $r' > 0$. Define

$$\varphi^n(t',x',a') = \varphi(a') - n(|t-t'|^2 + |x-x'|^{2p}) - |a-a'|^{2p}$$

on a closed neighbourhood $[t,t+r] \times \overline{B_r(x,a)}$ of radius $r > 0$. Since v is l.s.c. and φ is continuous, there exists a minimum on the closed neighbourhood:

$$\forall n, \quad \exists (t_n,x_n,a_n) \in [t,t+r] \times \overline{B_r(x,a)} : \\ (v-\varphi^n)(t_n,x_n,a_n) = \min_{[t,t+r]\times\overline{B_r(x,a)}} (v-\varphi^n)(t',x',a').$$

Now (t_n,x_n,a_n) form a sequence on the compact $[t,t+r] \times \overline{B_r(x,a)}$, thus, up to a subsequence, we have:

$$(t_n,x_n,a_n) \to (\bar{t},\bar{x},\bar{a}).$$

Now

$$\begin{aligned} v(t,x,a) - \varphi(a) &= (v-\varphi^n)(t,x,a) \\ &\geq (v-\varphi^n)(t_n,x_n,a_n) \\ &= v(t_n,x_n,a_n) - \varphi(a_n) + n(|t_n-t|^2 + |x_n-x|^{2p}) + |a_n-a|^{2p} \\ &= v(t_n,x_n,a_n) + v(t,x,a_n) - v(t,x,a_n) - \varphi(a_n) \\ &+ n(|t_n-t|^2 + |x_n-x|^{2p}) + |a_n-a|^{2p} \\ &\geq v(t_n,x_n,a_n) + v(t,x,a) - \varphi(a) - v(t,x,a_n) \\ &+ n(|t_n-t|^2 + |x_n-x|^{2p}) + |a_n-a|^{2p}. \end{aligned} \tag{5.29}$$

Thus,

$$\begin{aligned} n(|t_n-t|^2 + |x_n-x|^{2p}) + |a_n-a|^{2p} &\leq |v(t,x,a_n) - v(t_n,x_n,a_n)| \\ &\leq C(1+|x|^p + |x_n|^p + |a_n|^p) < K, \end{aligned}$$

since the sequences (x_n),(a_n) are convergent thus bounded. This implies that

$$t_n \to t, \qquad x_n \to x.$$

To prove that also $a_n \to a$, we have by (5.29)

$$v(t_n, x_n, a_n) - \varphi(a_n) \le v(t, x, a) - \varphi(a) \Rightarrow \\ \underline{\lim}(v(t_n, x_n, a_n) - \varphi(a_n)) \le v(t, x, a) - \varphi(a).$$

And since $v - \varphi$ is *lsc* we have the following

$$v(t, x, \bar{a}) - \varphi(\bar{a}) \le v(t, x, a) - \varphi(a),$$

which is a contradiction since $v(t, x, .) - \varphi$ has a strict local minimum on a. Now, from lemma 5.2 we know that v is a viscosity solution to (5.25) at the point (t_n, x_n, a_n). φ^n is a test function at the point (t_n, x_n, a_n) and thus we have

$$0 = D_a\varphi^n(t_n, x_n, a_n) = D_a\varphi(a_n) - 2p(a_n - a)\left|a_n - a\right|^{2p-1}.$$

By sending $n \to \infty$, since φ is continuous and $a_n \to a$ we finally obtain that

$$D_a\varphi(a) = 0,$$

which proves the claim.

□

We are now ready to show that v does not depend on the variable a locally. Since we only know that v solves (5.28) only in a viscosity sense, we will need to work around this. This is done by first approximating v with its inf-convolution and showing that such inf-convolution also solves (5.28). Since the inf-convolution of v is more regular, we are able to conclude that it is constant on a, and then show how this property extends also to v.

Theorem 5.1 *Under (HFD) and (HFC), for every* $(t, x) \in [0, T) \times \mathbb{R}^n$ *the function* v *defined in* (5.19) *does not depend on the variable* a *on* $\mathbb{R}^l$, *i.e.*

$$v(t, x, a) = v(t, x, a') \quad \forall a, a' \in \mathbb{R}^l.$$

Proof. We divide the proof in four steps.

Step 1 Fix $R > 0$. $\forall n$ consider $\forall a \in B_R(0)$ the function defined as follows:

$$u_n(t,x,a) = \inf_{a' \in \overline{B_R(0)}} \left[v(t,x,a') + n\left|a'-a\right|^{2p} \right]. \tag{5.30}$$

It is clear that $u_n(t,x,a) \leq v(t,x,a)$ and $u_n(t,x,a) \leq u_{n+1}(t,x,a)$. Moreover, since v is lsc $\forall n \quad \exists a_n \in \overline{B_R(0)}$ such that

$$u_n(t,x,a) = v(t,x,a_n) + n\left|a_n - a\right|^{2p}. \tag{5.31}$$

Now

$$u_n(t,x,a) \nearrow v(t,x,a). \tag{5.32}$$

Indeed, since $u_n(t,x,a)$ are increasing and bounded by $v(t,x,a)$ we have that

$$u_n(t,x,a) \nearrow \bar{u}(t,x,a) \leq v(t,x,a). \tag{5.33}$$

On the other side, we can write from (5.31) and the fact that u_n is bounded by v pointwise

$$\begin{aligned} u_n(t,x,a) = v(t,x,a_n) + n\left|a_n - a\right|^{2p} &\leq v(t,x,a) \\ \Rightarrow n\left|a_n - a\right|^{2p} &\leq |v(t,x,a)| + |v(t,x,a_n)| \\ &\leq C_v(2 + 2\left|x\right|^p + \left|a\right|^p + \left|a_n\right|^p) \\ &\leq C_v(2 + 2\left|x\right|^p + 2R^p) \end{aligned}$$

and we deduce that

$$a_n \to a \qquad \text{for} \qquad n \to \infty.$$

From (5.31) and (5.33) we have

$$v(t,x,a_n) + n\left|a_n - a\right|^{2p} \leq \bar{u}(t,x,a)$$

and thus, since v is lsc,

$$\begin{aligned} \liminf_{n\to\infty} v(t,x,a_n) + \liminf_{n\to\infty} n\left|a_n - a\right|^{2p} &\leq \bar{u}(t,x,a) \\ v(t,x,a) + \liminf_{n\to\infty} n\left|a_n - a\right|^{2p} &\leq \bar{u}(t,x,a) \end{aligned}$$

and, since $\liminf_{n\to\infty} n\,|a_n - a|^{2p}$ is positive,

$$v(t,x,a) \leq \bar{u}(t,x,a). \tag{5.34}$$

By combing (5.33) and (5.34) we obtain finally the (5.32).

Step 2 Fix $(t,x) \in [0,T] \times \mathbb{R}^n$. We now show that

$$\forall \epsilon > 0,\ \forall a \in B_{R-\epsilon}(0),\ \forall n > \frac{C_v(2 + 2\,|x|^p + 2R^p)}{\epsilon^{2p}} \Longrightarrow a_n \in B_R(0)$$

where a_n is the one defined in (5.31) and C_v is the constant in (5.22). Indeed, from the definition of a_n we have

$$u_n(t,x,a) = v(t,x,a_n) + n\,|a_n - a|^{2p} \leq v(t,x,a),$$

thus

$$\begin{aligned} n\,|a_n - a|^{2p} &\leq |v(t,x,a)| + |v(t,x,a_n)| \\ &\leq C_v\left(2 + 2\,|x|^p + |a|^p + |a_n|^p\right) \\ &\leq C_v\left(2 + 2\,|x|^p + 2R^p\right). \end{aligned}$$

Since

$$n > \frac{C_v(2 + 2\,|x|^p + 2R^p)}{\epsilon^{2p}},$$

we have that

$$|a_n - a| < \epsilon \Rightarrow a_n \in B_R(0).$$

Step 3 Fix $(t,x) \in [0,T] \times \mathbb{R}^n$. We show now that $\forall n > {C_v(2 + 2\,|x|^p + 2R^p)}/{\epsilon^{2p}}$, $\forall a \in B_{R-\epsilon}(0)$, $\forall \varphi \in C^2(\mathbb{R}^l)$ that realizes a local minimum, i.e.

$$0 = \left(u_n(t,x,.) - \varphi\right)(a) = \min_{a' \in B_{r'}(a)} \left[(u_n(t,x,.) - \varphi)(a')\right] \tag{5.35}$$

for some $r' > 0$, there exists $a_n \in B_R(0)$ and $\psi \in C^2(\mathbb{R}^l)$ such that

$$D_a\psi(a_n) = D_a\varphi(a) \tag{5.36}$$

$$0 = (v(t,x,.) - \psi)(a_n) = \min_{a' \in B_r(a_n)} [(v(t,x,.) - \psi)(a')] \tag{5.37}$$

for some $r > 0$ to be specified. Consider $a_n \in B_R(0)$ as defined in (5.31). Then, $\exists \eta > 0$ such that

$$B_\eta(a_n) \subset B_R(0).$$

Let $r = r' \wedge \eta$ and define

$$\psi(a') = \varphi(a' - a_n + a) - n\,|a_n - a|^{2p}. \tag{5.38}$$

It is clear that (5.36) is satisfied. Moreover

$$\psi(a_n) = \varphi(a) - n\,|a_n - a|^{2p} = u_n(t,x,a) - n\,|a_n - a|^{2p} = v(t,x,a_n)$$

and, $\forall a' \in B_r(a_n)$,

$$\begin{aligned} \psi(a') &= \varphi(a' - a_n - a) - n\,|a_n - a|^{2p} \\ &\le u_n(t,x,a' - a_n + a) - n\,|a_n - a|^{2p} \\ &\le v(t,x,a') + n\,|a_n - a|^{2p} - n\,|a_n - a|^{2p}, \end{aligned}$$

where we used the definition of $u_n(t,x,a)$ in the last line and the fact that $a' \in B_r(a_n) \subset B_R(0)$. Thus

$$\begin{aligned} \psi(a') &\le v(t,x,a') \qquad \forall a' \in B_r(a_n) \\ \psi(a_n) &= v(t,x,a_n), \end{aligned}$$

and (5.37) is satisfied.

Step 4 For any fixed $(t,x) \in [0,T] \times \mathbb{R}^n$, $\forall \epsilon > 0$, $\forall n > {C_v(2 + 2\,|x|^p + 2R^p)}/{\epsilon^{2p}}$ we have that u_n is a viscosity supersolution to (5.28) on $B_{R-\epsilon}$.
Indeed $\forall a \in B_{R-\epsilon}(0), \forall \varphi \in C^2(\mathbb{R}^l)$ such that $(u_n(t,x,.) - \varphi)(a) = 0$ and has a local minimum in a, there exists $a_n \in B_R(0)$ and $\psi \in C^2(\mathbb{R}^l)$ that work as test function for v, as seen in **Step 3**.
Then

$$0 = D_a\psi(a_n) = D_a\varphi(a),$$

and thus u_n is a supersolution to (5.28). Now, by definition as inf-convolution, u_n is semi-concave on $B_{R-\epsilon}(0)$. By Theorem 2.1.7 on [21] u_n is locally Lipschitz on $B_{R-\epsilon}(0)$. By Rademacher's theorem it is also almost everywhere differentiable on $B_{R-\epsilon}(0)$. Now, using Corollary 2.1 (ii) in [10], which we can apply since u_n is locally Lipschitz and thus continuous, we get that that $D_a u_n = 0$ holds in classical sense. Thus by proposition A in Annex , it is locally constant i.e.

$$u_n(t,x,a) = u_n(t,x,a') \quad \forall a' \in B_\eta(a) \tag{5.39}$$

$\forall \eta$ such that $B_\eta(a) \subset B_{R-\epsilon}(0)$. Now by taking the limit for $n \to \infty$ we obtain that the same holds true for v

$$v(t,x,a) = v(t,x,a') \quad \forall a' \in B_\eta(a). \tag{5.40}$$

Since this happens for all $\epsilon > 0$, we can send $\epsilon \to 0$ and obtain that v does not depend on a locally on $B_R(0)$ $\forall R > 0$. Since $B_R(0)$ is connected, it means it does not depend on a on the whole set $B_R(0)$. And since this is true $\forall R > 0$, it also means that v does not depend on a on all $\mathbb{R}^l$.

□

5.6 The viscosity solution property of v

We are now able to finally show that v is a viscosity solution to (5.3). From now on we will omit the dependence on a since it is constant on a and, by misuse of notation, denote it as $v(t,x)$.
What we obtain is the following result:

Theorem 5.2 *Under (HFD) and (HFC), v defined in* (5.19) *is a viscosity solution to* (5.3)-(5.4).

In what follows we have the detailed proof, divided in proof of supersolution property and subsolution property. It is interesting to note that the proof of subsolution property to (5.4) does not need additional hypotheses, as in [1].

Proposition 5.2 *Under (HFD) and (HFC) v, defined in* (5.19)*, is a viscosity supersolution to* (5.3).

Proof. From (5.26) and the fact that v does not depend on a we have

$$v(t,x) = v_*(t,x) = \liminf_{n\to\infty} {}_*v_n(t,x,a) \tag{5.41}$$

$\forall (t,x,a) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^l$. Let $(t,x) \in [0,T) \times \mathbb{R}^n$ and $\varphi \in C^{1,2}([0,T] \times \mathbb{R}^n)$ a test function such that

$$(v-\varphi)(t,x) = \min_{(t',x')\in\overline{B_r(t,x)}} (v-\varphi)(t',x')$$

for some $r>0$.
Now fix $a\in\mathbb{R}^l$ and define the test function

$$\psi(t',x',a') = \varphi(t',x') - (|t-t'|^2 + |x-x'|^{2p} + |a-a'|^{2p})$$

for all $(t',x',a')\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^l$. We have that $\psi(t,x,a)=\varphi(t,x)$. and $\psi\le\varphi$ with equality only when $(t',x',a')=(t,x,a)$, thus

$$(v-\psi)(t,x,a) = \text{strict}\min_{(t',x',a')\in\overline{B_r(t,x,a)}} (v-\psi)(t',x',a').$$

Now consider the continuous function v_n. $\forall n$, there exists $(t_n,x_n,a_n)\in\overline{B_r(t,x,a)}$ such that

$$(v_n-\psi)(t_n,x_n,a_n) = \min_{(t',x',a')\in\overline{B_r(t,x,a)}} (v-\psi)(t',x',a').$$

As in Lemma 5.2 we have that

$$(t_n,x_n,a_n,v_n(t_n,x_n,a_n)) \to (t,x,a,v(t,x)).$$

For n large enough, ψ is a test function for v_n at the point (t_n,x_n,a_n) and by the fact that v_n is a viscosity supersolution to 5.23 we have

$$\begin{aligned}-\frac{\partial\psi}{\partial t}(t_n,x_n,a_n) - \mathcal{L}^{a_n}\psi(t_n,x_n,a_n) - \Delta_a\psi(t_n,x_n,a_n) - n\,|D_a\psi(t_n,x_n,a_n)|\\ -f(x_n,a_n,v_n(t_n,x_n,a_n),\sigma^{\mathrm{T}}(x_n,a_n)D_x\psi(t_n,x_n,a_n)) \ge 0,\end{aligned}$$

and since $n\,|D_a\psi(t_n,x_n,a_n)|\ge 0$ we have

$$\begin{aligned}&-\frac{\partial\psi}{\partial t}(t_n,x_n,a_n) - \mathcal{L}^{a_n}\psi(t_n,x_n,a_n) - \Delta_a\psi(t_n,x_n,a_n)\\ &-f(x_n,a_n,v_n(t_n,x_n,a_n),\sigma^{\mathrm{T}}(x_n,a_n)D_x\psi(t_n,x_n,a_n)) \ge 0.\end{aligned}$$

By sending n to infinity, since ψ is a $C^{1,2,2}([0,T] \times \mathbb{R}^n \times \mathbb{R}^l)$ function, we get

$$-\frac{\partial \psi}{\partial t}(t,x,a) - \mathcal{L}^a \psi(t,x,a) - f(x,a,v(t,x),\sigma^{\mathrm{T}}(x,a) D_x \psi(t,x)) - \Delta_a \psi(t,x,a) \geq 0.$$

It is clear, by explicit computation, that $\Delta_a \psi(t,x,a) = 0$. Remembering that $\psi(t,x,a) = \varphi(t,x)$ and that the same holds for their respective derivatives on t and x, we have

$$-\frac{\partial \varphi}{\partial t}(t,x) - \mathcal{L}^a \varphi(t,x) - f(x,a,v(t,x),\sigma^{\mathrm{T}}(x,a) D_x \varphi(t,x)) \geq 0.$$

Since a was arbitrarily chosen in $\mathbb{R}^l$, we have that v is a supersolution to (5.3), i.e.

$$-\frac{\partial \varphi}{\partial t}(t,x) - \sup_{a \in \mathbb{R}^l} \left[\mathcal{L}^a \varphi(t,x) - f(x,a,v(t,x),\sigma^{\mathrm{T}}(x,a) D_x \varphi(t,x)) \right] \geq 0.$$

□

Proposition 5.3 *Under (HFD) and (HFC) $v*$, the upper semi-continuous envelope of v defined in* (5.19), *is a viscosity subsolution to* (5.3).

Proof. From [10] p. 91 we know that

$$v^*(t,x) = \limsup_{n \to \infty}{}^* v_n(t,x,a), \tag{5.42}$$

where the $\limsup_{n \to \infty}{}^*$ is defined as

$$\limsup_{n \to \infty}{}^* v_n(t,x,a) = \limsup_{\substack{n \to \infty \\ (t',x',a') \to (t,x,a) \\ t' < T}} v_n(t',x',a').$$

Fix now $(t,x) \in [0,T) \times \mathbb{R}^n$ and $\varphi \in C^{1,2}([0,T] \times \mathbb{R}^n)$ such that $v^* - \varphi$ has a local maximum in (t,x), i.e.

$$(v^* - \varphi)(t,x) = \max_{(t',x') \in B_r(t,x)} (v^* - \varphi)(t',x') \tag{5.43}$$

for some $r > 0$. We may assume that the maximum is strict.

Fix now $a \in \mathbb{R}^l$. By (5.42) we can find a sequence (t_n, x_n, a_n) such that

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \to (t, x, a, v^*(t, x)). \tag{5.44}$$

We define φ_n as follows

$$\varphi_n(t', x', a') = \varphi(t', x') + |t' - t_n|^2 + |x' - x_n|^{2p} + \frac{1}{n^3}\left(\frac{|a' - a_n|}{r/4}\right)^n \tag{5.45}$$

and we notice that $\varphi_n(t', x', a') \geq \varphi(t', x')$ and equality holds only (t_n, x_n, a_n). Consider now the continuous functions v_n. $\forall n$ there exists $(\bar{t}_n, \bar{x}_n, \bar{a}_n) \in \overline{B_r(t, x, a)}$, i.e.

$$(v_n - \varphi_n)(\bar{t}_n, \bar{x}_n, \bar{a}_n) = \max_{(t',x',a') \in \overline{B_r(t,x,a)}} (v_n - \varphi_n)(t', x', a'). \tag{5.46}$$

Since $\overline{B_r(t, x, a)}$ is compact, we have that up to a subsequence

$$(\bar{t}_n, \bar{x}_n, \bar{a}_n) \to (\bar{t}, \bar{x}, \bar{a}),$$

for some $(\bar{t}, \bar{x}, \bar{a}) \in \overline{B_r(t, x, a)}$. Now that we have defined everything, we divide the rest of the proof in steps in order to make it easier to read.

Step 1 We show now that $(\bar{t}, \bar{x}) = (t, x)$.
For n large enough $(t_n, x_n, a_n) \in B_r(t, x, a)$, so we can write from (5.46) that

$$\begin{aligned} v_n(t_n, x_n, a_n) - \varphi(t_n, x_n) = (v_n - \varphi_n)(t_n, x_n, a_n) &\leq (v_n - \varphi_n)(\bar{t}_n, \bar{x}_n, \bar{a}_n) \\ &\leq (v^* - \varphi)(\bar{t}_n, \bar{x}_n). \end{aligned} \tag{5.47}$$

By taking $\limsup_{n\to\infty}$, using (5.44) and the fact that φ is continuous on the left side of the inequality, the fact that v^* is u.s.c. on the right side we obtain

$$(v^* - \varphi)(t, x) \leq (v^* - \varphi)(\bar{t}, \bar{x}),$$

which contradicts the fact that the maximum in (5.43) is strict. Thus it must be

$$(t, x) = (\bar{t}, \bar{x}).$$

Step 2 We show now that for some N large enough

$$(\bar{t}_n, \bar{x}_n, \bar{a}_n) \in B_r(t, x, a) \qquad \forall n > N.$$

To prove that, we start by noticing from (5.47) that

$$\varphi_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) \leq v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) - v_n(t_n, x_n, a_n) + \varphi(t_n, x_n)$$

and thus, using the polynomial growth conditions,

$$\begin{aligned} \frac{1}{n^3}\left(\frac{|a' - a_n|}{r/4}\right)^n &\leq |v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) - v_n(t_n, x_n, a_n) + \varphi(t_n, x_n) - \varphi(\bar{t}_n, \bar{x}_n)| \\ &\leq C(1 + |x_n|^p + |a_n|^p + |\bar{x}_n|^p + |\bar{a}_n|^p) \\ &\leq K \end{aligned} \tag{5.48}$$

since the sequences (x_n, a_n) and $(\bar{x}_n, \bar{a}_n)$ are convergent and thus bounded.

Consider now $\tilde{N}$ such that $\forall n > \tilde{N}$ $|(\bar{t}_n, \bar{x}_n) - (t, x)| < r/2$ (such $\tilde{N}$ exists since $(\bar{t}_n, \bar{x}_n)$ converges to (t, x)), and let $\hat{N}$ be s.t $|a_n - a| < r/4$ $\forall n > \hat{N}$.
Now, there exists $N > \tilde{N} \vee \hat{N}$ s.t for $n > N$ it holds

$$|\bar{a}_n - a_n| \leq \frac{r}{4}.$$

Suppose this is not true, i.e. for every $M \in \mathbb{N}$ there exists a $\bar{n} > M$ such that $|\bar{a}_{\bar{n}} - a_{\bar{n}}| > r/4$. Then there exists a subsequence $(\bar{a}_{n_k})_{k \geq 0}$ such that

$$|\bar{a}_{n_k} - a_{n_k}| > \frac{r}{4} \qquad \forall k \geq 0,$$

then we would have

$$K \geq \frac{1}{{n_k}^3}\left(\frac{|\bar{a}_{n_k} - a_{n_k}|}{r/4}\right)^{n_k} > \frac{1}{{n_k}^3}\delta^{n_k} \quad \text{for some } \delta > 1 \text{ , } \forall k \geq 0,$$

which is a contradiction for k big enough.
Thus we have that

$$|\bar{a}_n - a_n| \leq \frac{r}{4} \qquad \forall n > N \tag{5.49}$$

and

$$|\bar{a}_n - a| \leq |\bar{a}_n - a_n| + |a_n - a| < \frac{r}{2} \qquad \forall n > N. \tag{5.50}$$

Then if we consider the distance of the point $(\bar{t}_n, \bar{x}_n, \bar{a}_n)$ from (t, x, a) we have

$$\begin{aligned} |(\bar{t}_n, \bar{x}_n, \bar{a}_n) - (t, x, a)|^2 &= |(\bar{t}_n, \bar{x}_n) - (t, x)|^2 + |\bar{a}_n - a|^2 \\ &< \frac{r^2}{4} + \frac{r^2}{4} < \frac{r^2}{2} \\ |(\bar{t}_n, \bar{x}_n, \bar{a}_n) - (t, x, a)| &< \frac{r}{\sqrt{2}} < r \end{aligned}$$

for $n > N$. This proves that

$$(\bar{t}_n, \bar{x}_n, \bar{a}_n) \in B_r(t, x, a) \qquad \forall n > N. \tag{5.51}$$

By its definition in (5.45), through explicit calculation, it is clear that

$$\varphi_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) \to \varphi(t, x) \tag{5.52}$$
$$n\,|D_a \varphi_n(t_n, x_n, a_n)| \to 0 \tag{5.53}$$
$$\Delta_a \varphi_n(t_n, x_n, a_n) \to 0. \tag{5.54}$$

Indeed we have that $\forall n > N$, thanks to (5.49) the term

$$\frac{1}{n^3}\left(\frac{|\bar{a}_n - a_n|}{r/4}\right)^n \leq \frac{1}{n^3} \to 0$$

and the same applies to the norm of its gradient and its laplacian.

Step 3 We also have that $v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) \to v^*(t, x)$.

To show this, we start from (5.47) and take the $\liminf$ on the left side of the inequalities:

$$\begin{aligned} \liminf_{n\to\infty} v_n(t_n, x_n, a_n) - \varphi(t_n, x_n) &\leq \liminf_{n\to\infty} (v_n - \varphi_n)(\bar{t}_n, \bar{x}_n, \bar{a}_n) \\ (v^* - \varphi)(t, x) &\leq \liminf_{n\to\infty} (v_n - \varphi_n)(\bar{t}_n, \bar{x}_n, \bar{a}_n), \end{aligned}$$

and we do the same with the lim sup on the right side inequalities:

$$\begin{aligned}\limsup_{n\to\infty}(v_n-\varphi_n)(\bar{t}_n,\bar{x}_n,\bar{a}_n) &\le \limsup_{n\to\infty}(v^*-\varphi)(\bar{t}_n,\bar{x}_n)\\ &\le (v^*-\varphi)(t,x),\end{aligned}$$

which proves that $v_n(\bar{t}_n,\bar{x}_n,\bar{a}_n) \to v^*(t,x)$ since $\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n) \to \varphi(t,x)$. By explicit calculation we have also that

$$\begin{aligned}D_x\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n) &\to D_x\varphi(t,x)\\ D_x^2\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n) &\to D_x^2\varphi(t,x).\end{aligned}$$

Step 4 We can finally show that v is a subsolution.
For $n > N$ large enough $(\bar{t}_n,\bar{x}_n,\bar{a}_n) \in B_r(t,x,a)$ and so φ_n is a test function for v_n at $(\bar{t}_n,\bar{x}_n,\bar{a}_n)$. Using the subsolution property of v_n to (5.23) we have

$$\begin{aligned}&-\frac{\partial\varphi_n}{\partial t}(\bar{t}_n,\bar{x}_n,\bar{a}_n) - \mathcal{L}^{\bar{a}_n}\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n) - \Delta_a\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n)\\ &- f(\bar{x}_n,\bar{a}_n,v_n(\bar{t}_n,\bar{x}_n,\bar{a}_n),\sigma^{\mathrm{T}}(\bar{x}_n,\bar{a}_n)D_x\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n)) - n\,|D_a\varphi_n(\bar{t}_n,\bar{x}_n,\bar{a}_n)| \le 0,\end{aligned}$$

and by taking its limit and using (5.52), (5.53) and (5.54) we get

$$-\frac{\partial\varphi}{\partial t}(t,x) - \mathcal{L}^{\bar{a}}\varphi(t,x) - f(x,\bar{a},v^*(t,x),\sigma^{\mathrm{T}}(x,\bar{a})D_x\varphi(t,x)) \le 0.$$

Since

$$\begin{aligned}&-\sup_{a'\in\mathbb{R}^l}[\mathcal{L}^{a'}\varphi(t,x) + f(x,a',v^*(t,x),\sigma^{\mathrm{T}}(x,a')D_x\varphi(t,x))] \le\\ &\qquad\qquad - \mathcal{L}^{\bar{a}}\varphi(t,x) - f(x,\bar{a},v^*(t,x),\sigma^{\mathrm{T}}(x,\bar{a})D_x\varphi(t,x))\end{aligned}$$

we obtain the subsolution property, i.e

$$-\frac{\partial\varphi}{\partial t}(t,x) - \sup_{a'\in\mathbb{R}^l}[\mathcal{L}^{a'}\varphi(t,x) + f(x,a',v^*(t,x),\sigma^{\mathrm{T}}(x,a')D_x\varphi(t,x))] \le 0.$$

□

Proposition 5.4 *Under (HFD) and (HFC), v defined in* (5.19) *is a viscosity supersolution to* (5.4).

Proof. Let $(x,a) \in \mathbb{R}^n \times \mathbb{R}^n$. From (5.41), there exists a sequence $(t_n, x_n, a_n)_n$ valued in $[0,T) \times \mathbb{R}^n \times \mathbb{R}^l$ such that

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \to (T, x, a, v_*(T,x)) \text{ as } n \to \infty.$$

The sequence of continuous functions v_n is nondecreasing and $v_n(T, x', a') = g(x', a')$ so we have

$$v_*(T,x) \geq \lim_{n\to\infty} v_1(t_n, x_n, a_n) = g(x,a).$$

Since this was done for an arbitrary a, we obtain that

$$v_*(T,x) \geq \sup_{a\in\mathbb{R}^l} g(x,a),$$

and thus v is a supersolution to (5.4). □

Proposition 5.5 *Under (HFD) and (HFC), $v*$, the upper semi-continuous envelope of v defined in* (5.19), *is a viscosity subsolution to* (5.4).

Proof. Consider a point $\bar{x}$. From (5.42) we can find a sequence such that

$$(t_n, x_n, v_n(t_n, x_n, a_n)) \to (T, \bar{x}, v^*(T, \bar{x})). \tag{5.55}$$

Consider now, just as in lemma 5.1, the smooth function

$$\bar{v}(t,x) = \bar{C} e^{\rho(T-t)}(1 + |x|^p) \tag{5.56}$$

where we can choose the constant $\bar{C}$ as

$$\bar{C} := \frac{\sup_{a\in\mathbb{R}^l} g(\bar{x}, a)}{1 + |\bar{x}|^p}.$$

Like in lemma 5.2, we can show that the function in (5.56) defines a solution $\bar{Y_s^{t,x,a}} = \bar{v}(s, X_s^{t,x,a})$ (and $\bar{Y}_T^{t,x,a} = g(X_T^{t,x,a}, X'^{t,x,a}_T))$ when $\bar{v}$ is composed with the forward diffusion. As in chapter 3, by comparison theorems it is possible to show that $\bar{Y}_s$ is an upper bound for Y_s^n, so we have:

$$v_n(s,y,b) = Y_s^{n,s,y,b} \leq \bar{Y}_s^{s,y,b} = \bar{v}(s, X_s^{s,y,b}) = \bar{v}(s,y) \quad \forall s,y,b \in [0,T) \times \mathbb{R}^n \times \mathbb{R}^l.$$

By the choice on $\bar{C}$, we have that

$$\bar{v}(T, \bar{x}) = \sup_{a \in \mathbb{R}^l} g(\bar{x}, a).$$

By defining, for all $\epsilon > 0$, h_ϵ as:

$$h_\epsilon(t, x) = \sqrt{T - t} + \sup_{a \in \mathbb{R}^l} g(x, a) + \epsilon,$$

it is clear that we have the following inequality on in $(T, \bar{x})$:

$$\bar{v}(T, \bar{x}) < h_\epsilon(T, \bar{x}).$$

Since $\bar{v}$ and h_ϵ are continuous, there must be a neighbourhood of radius r_ϵ centred in (T, x) such that

$$\bar{v}(t, x) \leq h_\epsilon(t, x) \qquad \forall (t, x) \in B_{r_\epsilon}(T, x).$$

For n large enough, say $n > N$, $(t_n, x_n) \in B_{r_\epsilon}(T, x)$ and

$$h_\epsilon(t_n, x_n) \geq \bar{v}(t_n, x_n) \geq v_n(t_n, x_n, a_n) \quad \forall n > N.$$

And taking the limit on n:

$$v^*(T, \bar{x}) \leq h_\epsilon(T, \bar{x}) = \sup_{a \in \mathbb{R}^l} g(\bar{x}, a) + \epsilon. \tag{5.57}$$

Sending ϵ to 0, we obtain the claim in $\bar{x}$:

$$v^*(T, \bar{x}) \leq \sup_{a \in \mathbb{R}^l} g(\bar{x}, a)$$

□

5.7 About the uniqueness of the solution

We have established the existence of a representation formula for a solution, in the class of lower semi-continuous function with polynomial growth, to the equation

(5.3)-(5.4). The next thing one would ask himself is if that solution is unique. In that case we would have an explicit representation formula for the solution of the equation. We do not address this issue here, and we limit ourselves at an analysis of what can be found in the literature.

As already stated in section 2.2, when the equation is associated to a control problem with $f = f(x, a)$ and $g = g(x)$, the existence and uniqueness of the solution to the PDE is assured under the hypotheses that b, σ, f and g are Lipschitz continuous on x and bounded on a (see [7]). More precisely, the solution is unique in the class of functions satisfying:

$$
\begin{aligned}
|u(s,y)| &\leq K(1+|y|) \qquad \forall (s,y) \in [0,T] \times \mathbb{R}^n, \\
|u(s,y) - u(\hat{s},\hat{y})| &\leq K\left\{|y-\hat{y}| + (1+|y| \vee |\hat{y}|)\, |s-\hat{s}|^{1/2}\right\}, \forall s, \hat{s} \in [0,T],\ y, \hat{y} \in \mathbb{R}^n.
\end{aligned}
$$

In the general case, it is harder to find uniqueness results in literature. Usually it is done through comparison theorems for viscosity solutions, but for the second order case with unbounded domain like for our equation, it has been difficult to find the needed tools. We refer the reader to the reference article [9] for classical techniques to prove uniqueness. The uniqueness of the solution is thus an open problem.

6

Conclusions and future work

The aim of the work was to establish a representation formula ("à la Feynman-Kac") for the Hamilton-Jacobi-Bellman equation in the case where the control space is the whole set $\mathbb{R}^l$. Following the discussion on [1], we have shown that the solution to a constrained BSDE (5.7)(5.8) gives a representation formula for the fully nonlinear Hamilton-Jacobi-Bellman PDE (5.3)-(5.4).
When the PDE is associated to an optimal stochastic control problem:

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^T f(X_s^{t,x,\alpha}, \alpha_s)ds + g(X_T^{t,x,\alpha})\right],$$

then it is a representation formula for the *unique* value function of the problem. Thanks to this probabilistic representation, it may be possible to find a numerical scheme to simulate this function through Monte Carlo simulation.
Moreover, when using the fact that the solution to the BSDE is also represented through the dual control problem in 4, we can write the solution to the original control problem as essential supremum over a family of changes of probability, i.e.

$$\sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^T f(X_s^{t,x,\alpha}, \alpha_s)ds + g(X_T^{t,x,\alpha})\right] = \operatorname*{ess\,sup}_{u \in U_A} \mathbb{E}^u \left[\xi + \int_t^T f_s ds \middle| \mathcal{F}_t^{\widetilde{\mathbf{W}}}\right].$$

Since we do not have a uniqueness result in the general case, the first step would be to try classical viscosity solution techniques and comparison principles.

The assumptions made are quite general and nothing unusual, except for the fact that the control space must be the whole $\mathbb{R}^l$ space. This can be fairly limiting as in a lot of applications the control space is much more restricted. One of the

next steps would be thus to investigate if it is possible to find a function ρ that maps the space $\mathbb{R}^l$ into the desired control space A, and what conditions should this function satisfy while keeping the results in this work true. It may even be possible to find a function that maps it to a different A set, for example a finite discrete set.
It is also to note that assumption that f is Lipschitz on the state variable and on the control, used in [1], is not needed here as the sub-polynomial growth condition is enough. Requiring Lipschitz continuity on the control variable is not among the typical hypotheses and could be limiting. Moreover, there is no need for additional assumption to prove the subsolution property to the final condition.

It would also be interesting to study what suitable numerical schemes can be used to simulate the constrained BSDE and obtain the solution to the control problem.

Appendix

The following is a result found in a past version of [1]. The proof was removed from the final paper, we rewrite it here since it is needed in one of the proofs.

Proposition A *Let O be an open nonempty subset of $\mathbb{R}^l$ and $u : A \to \mathbb{R}$ a locally Lipschitz function such that*

$$D_x u = 0$$

almost everywhere on O. Then u is locally constant on O.

Proof. Let $x_0 \in O$ and $r > 0$ such that $B_r(x_0) \subset O$. Define the function $\tilde{u}$ on $\mathbb{R}^l$ by

$$\tilde{u} \coloneqq u\left(\Pi_{\overline{B_r(x_0)}}(x)\right),$$

where $\Pi_{\overline{B_r(x_0)}}$ is the projection operator on the closure of $B_r(x_0)$. Then $\tilde{u}$ is globally Lipschitz continuous on $\mathbb{R}^l$ and it satisfies $D_x\tilde{u} = 0$, almost everywhere on $B_r(x_0)$. Let $\Phi \in C^\infty(\mathbb{R}^l)$ be a function with compact support such that

$$\Phi \geq 0 \qquad \int_{\mathbb{R}^l} \Phi(x)dx = 1.$$

We define the sequence of functions $(u_k)_k$ as:

$$u_k(x) = k^l \int_{\mathbb{R}^l} \tilde{u}(x-y)\Phi(ky)dy$$

Since $\tilde{u}$ is Lipschitz, we see that $(u_k)_k$ converges uniformly to $\tilde{u}$.

Using a change of variable we also have that

$$u_k(x) = \int_{\mathbb{R}^l} \tilde{u}\left(x - \frac{z}{k}\right) \Phi(z) dz$$

Moreover, for

$$k \geq \frac{\sup\{|x| \ : \ \Phi(x) > 0\}}{{}^r\!/_2},$$

we have that the derivative of the $\tilde{u}(x - {}^z\!/_k)\Phi(z)$ is bounded:

$$\frac{\tilde{u}(x - {}^z\!/_k + h) - \tilde{u}(x - {}^z\!/_k)}{h} \Phi(z) \leq C\Phi(z) \leq C \max_{x \in \operatorname{supp} \Phi} [\Phi(x)],$$

and from the dominated convergence Theorem

$$\begin{aligned} D_x u_k(x) &= D_x \left(k^l \int_{B_{r/2}(0)} \tilde{u}(x-y)\Phi_k(ky)dy \right) \\ &= k^l \int_{B_{r/2}(0)} D_x \tilde{u}(x-y)\Phi_k(ky)dy = 0, \end{aligned}$$

for all $x \in B_{r/2}(x_0)$. Therefore u_k is constant on $x \in B_{r/2}(x_0)$ for k large enough and $\tilde{u}$ and u are constant on $x \in B_{r/2}(x_0)$. □

Bibliography

[1] Idris Kharroubi and Huyên Pham. 'Feynman-Kac representation for Hamilton-Jacobi-Bellman IPDE'. In: *arXiv preprint arXiv:1212.2000, to appear on Annals Of Probability* (2012).

[2] Romuald Elie and Idris Kharroubi. 'Constrained Backward SDEs with Jumps: Application to Optimal Switching'. In: *Preprint* (2009).

[3] Idris Kharroubi et al. 'Backward SDEs with constrained jumps and quasi-variational inequalities'. In: *The Annals of Probability* 38.2 (2010), pp. 794–840.

[4] AB Aries and Nikolaj Vladimirovič Krylov. *Controlled diffusion processes.* Vol. 14. Springer, 2008.

[5] Wendell Helms Fleming et al. *Controlled Markov processes and viscosity solutions.* Vol. 25. Springer, 2006.

[6] Huyên Pham. *Continuous-time stochastic control and optimization with financial applications.* Vol. 1. Springer, 2009.

[7] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations.* Vol. 43. Springer, 1999.

[8] Michael G. Crandall and Pierre-Louis Lions. 'Viscosity solutions of Hamilton-Jacobi equations'. In: *Transactions of the American Mathematical Society* 277.1 (1983), pp. 1–42.

[9] Michael G. Crandall, Hitoshi Ishii and Pierre-Louis Lions. 'User's guide to viscosity solutions of second order partial differential equations'. In: *Bulletin of the American Mathematical Society* 27.1 (1992), pp. 1–67.

[10] Guy Barles. *Solutions de viscosité des équations de Hamilton-Jacobi.* French. Springer Verlag, 1994.

[11] Étienne Pardoux and Shige Peng. 'Adapted solution of a backward stochastic differential equation'. In: *Systems & Control Letters* 14.1 (1990), pp. 55–61.

[12] Jin Ma et al. *Forward-backward stochastic differential equations and their applications*. Vol. 1702. Springer, 1999.

[13] Nicole El Karoui, Shige Peng and Marie Claire Quenez. 'Backward stochastic differential equations in finance'. In: *Mathematical finance* 7.1 (1997), pp. 1–71.

[14] Étienne Pardoux. 'Backward Stochastic Differential Equations and Applications'. English. In: *Proceedings of the International Congress of Mathematicians*. Ed. by S.D. Chatterji. Birkhäuser Basel, 1995, pp. 1502–1510. ISBN: 978-3-0348-9897-3. DOI: 10.1007/978-3-0348-9078-6_83. URL: http://dx.doi.org/10.1007/978-3-0348-9078-6_83.

[15] Shige Peng. 'Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob–Meyers type'. In: *Probability theory and related fields* 113.4 (1999), pp. 473–499.

[16] H Mete Soner, Nizar Touzi and Jianfeng Zhang. 'Wellposedness of second order backward SDEs'. In: *Probability Theory and Related Fields* 153.1-2 (2012), pp. 149–190.

[17] Shige Peng. 'G-expectation, G-Brownian motion and related stochastic calculus of Itô type'. In: *Stochastic analysis and applications*. Springer, 2007, pp. 541–567.

[18] Jakša Cvitanić and Ioannis Karatzas. 'HEDGING AND PORTFOLIO OPTIMIZATION UNDER TRANSACTION COSTS: A MARTINGALE APPROACH12'. In: *Mathematical finance* 6.2 (1996), pp. 133–165.

[19] Paolo Baldi. *Equazioni differenziali stocastiche e applicazioni*. Italian. Vol. 28. Pitagora, 1984.

[20] Bernt Øksendal. *Stochastic differential equations*. Springer, 2003.

[21] Piermarco Cannarsa and Carlo Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Vol. 58. Springer, 2004.

Ringraziamenti

Innanzitutto voglio ringraziare il mio relatore, il Professor Marco Fuhrman, che mi ha seguito e consigliato nella creazione e nella stesura di questo lavoro. Senza la sua guida non sarebbe stato possibile. Ringrazio Alessandra per essere stata al mio fianco in questi ultimi anni, sempre spingendomi ad andare avanti e superare me stesso. Un sentito ringraziamento va alla mia famiglia, per avermi sempre supportato e sopportato.

La mia carriera universitaria è stata piena di incontri ed eventi, vorrei ringraziare tutte quelle persone che hanno contribuito a renderla un'esperienza unica. Diego, Matteo, Bruno, Pier, Rodrigo, Stephan, Enrique, Pablito e tutti gli altri che hanno condiviso con me l'incredibile l'esperienza in Francia. I ragazzi dell'AIM, del 2013 e 2014, insieme ai quali abbiamo lavorato sodo ottenendo ottimi risultati. I miei coinquilini Conny, Silvano, Rémi e tutti coloro con cui ho convissuto finora e con i quali ho tenuto lunghe discussioni notturne. Giorgio e Monica, immancabili compagni di progetto e di pause infra-lezioni. Non può mancare una nomina alla *crew* (i cui membri sono troppi per essere nomitati qui) per tutte le pause, lezioni e serate passate insieme.
Ringrazio di nuovo tutte queste persone, perché hanno reso gli anni di università unici e indimenticabili.