UNIVERSITÉ PARIS VI — PIERRE et MARIE CURIE

U.F.R de Mathématiques

THÈSE DE DOCTORAT

Spécialité Mathématiques Appliquées

Présentée par **STEFANIA MAROSO**

Pour obtenir le grade de docteur de l'Université Paris VI

Analyse Numérique de Problèmes de Contrôle Stochastique

Soutenue le 7 Décembre 2006

Directeurs de Thèse :

Frédéric Bonnans
Hasnaa Zidani

Rapporteurs :

Pierre Cardaliaguet
Espen R. Jakobsen
Panagiotis E. Souganidis

Jury :

Frédéric Bonnans
Pierre Cardaliaguet
Rama Cont
Olivier Pironneau
Sylvain Sorin
Denis Talay
Hasnaa Zidani

ENSTA Unité de Mathématiques Appliquées

Cette thèse a été préparée conjointement au Laboratoire de Mathématiques Appliquées de l'ENSTA et au Projet SYDOCO à l'INRIA.

INRIA Projet SYDOCO

Remerciements

Je tiens en premier lieu à remercier sincèrement mes directeurs de thèse, Frédéric Bonnans et Hasnaa Zidani, qui m'ont si bien guidée pendant ces trois années de thèse.

Je suis très reconnaissante à Frédéric Bonnans, de m'avoir donné la chance de travailler avec lui, d'abord pendant mon stage de DEA et ensuite en acceptant de diriger cette thèse. Il m'a donné la grande opportunité de travailler à l'INRIA, un milieu de travail très stimulant et riche intellectuellement. Je le remercie pour le temps qu'il m'a consacré et je lui voue une très grande admiration pour ses connaissances considérables en optimisation et en théorie de viscosité. J'ai beaucoup appris à ses coté sur ces deux domaines qui étaient pour moi totalement nouveaux.

Je ne remercierai jamais assez Hasnaa Zidani, pour m'avoir toujours suivie et aidée. Je lui suis très reconnaissante pour les nombreuses discussions et suggestions qui m'ont permis d'améliorer mes connaissances et la qualité de la thèse. Ses remarques et ses critiques constructives ont énormément contribué à améliorer la clarté de l'exposition, la rigueur des démonstrations, ainsi que la langue. Enfin je voudrais la remercier pour m'avoir toujours encouragée et pour avoir toujours trouvé les bon mots pour me faire retrouver la confiance en moi-même.

C'est un grand honneur pour moi que le Professeur Souganidis ait accepté d'être rapporteur de cette thèse. Je le remercie beaucoup pour le temps qu'il a consacré à la lecture de ce travail et pour ces remarques précieuses.

Je remercie vivement le Professeur Jakobsen d'avoir accepté d'être rapporteur de ma thèse. Ses travaux sur les estimations d'erreurs constituent des résultats très importants sur lesquels cette thèse se fonde. Je voudrais le remercier aussi pour les discussions qu'on a eues lors des conférences sur la théorie de la viscosité.

Ma gratitude va ensuite au Professeur Pierre Cardaliaguet, qui m'a fait l'honneur d'accepter d'être rapporteur. Je le remercie vivement pour le temps qu'il a consacré à la lecture de ce manuscrit et pour la rapidité et l'efficacité dont il a fait preuve dans la correction.

C'est un grand plaisir pour moi que les Professeurs Rama Cont, Olivier Pironneau, Sylvain Sorin et Denis Talay fassent partie du Jury de thèse. Je les remercie en particulier d'avoir accepté cette tâche sans hésitations, en changeant leurs nombreux engagements afin d'être présents à la soutenance.

Je tiens à remercier Olivier Bokanowski et Benjamin Bruder, avec qui j'ai travaillé dans la deuxième partie de la thèse. C'est pour moi un grand plaisir de collaborer avec eux.

Ma gratitude va aussi à Jean-Charles Gilbert, et encore une fois à Hasnaa Zidani, pour m'avoir donné la possiblité de faire de l'enseignement à l'ENSTA. Cette expérience a été pour moi très enrichissante et importante.

Si cette thèse a pu aboutir, le mérite va aussi aux personnes des deux laboratoires dans lesquels j'ai effectué mes recherches. Ces personnes ont contribué a rendre le milieu de travail très riche et très amical. C'est pour cette raison que je remercie les membres (passés et présents) du Projet SYDOCO, avec un merci particulier pour Elisabeth, pour ses nombreux conseils, Audrey pour les pauses café passées à discuter et Hector, pour sa sympathie et pour nos matchs de tennis. Je tiens également à remercier les membres de l'Unité de Mathématiques Appliqués de l'ENSTA, en particulier Pierre Carpentier et Frédéric Jean qui m'ont beaucoup aidée pendant ma première expérience d'enseignement. Je n'oublie certainement pas mes amis thésards, qui sont toujours disponibles et très chaleureux : un grand merci va donc a Nadia, Eve-Marie, Grace, Nicolas et Carlo.

Je voudrais remercier particulièrement le personnel du service d'impression de l'ENSTA, pour leur rapidité et leur efficacité.

Cette thèse est le fruit de trois années de travail. Ces années ont été parfois enthousiasmantes, parfois difficiles. Beaucoup d'amis m'ont aidé, sans hésitations, à faire face aux moments les plus difficiles et ils ont partagé avec moi les moments les plus beaux.

Je voudrais remercier Marta, car son arrivée à Paris m'a fait sentir "à la maison" ! Je la remercie pour les samedis après-midi passés à parler pendant le shopping, comme pendant les années du lycée, pour oublier les semaines difficiles au travail. Un grand merci aussi à Bastien pour sa grande sympathie et bonne humeur. Je n'oublie pas de remercier mes amis turcs, Eda et Burak, pour tous les moments de détente et pour leur grande amitié. Je voudrais remercier Sareh, pour l'amitié qui nous lie, même si on se voit très rarement. Un grand merci de tout mon coeur aux amis de la communauté italienne de Paris, avec qui je partage les avantages et les désavantages d'être à l'étranger : Adriano, Ana, Alessio, Filippo, Ernesto, Concetta, et tous ceux que j'ai oublié de mentionner.

Un grazie dal profondo del cuore ai miei genitori e a mio fratello, che mi hanno sempre e incondizionatamente appoggiato in ogni mia scelta. Li ringrazio per essermi sempre cosi' vicini e per non farmi mai mancare il loro amore.

Ringrazio anche i miei suoceri Elsa ed Antonio, per essere venuti dall'Italia a condividere con me questo momento di gioia.

Enfin, le plus grand merci va à mon mari Andrea, pour son amour, ses encouragements et son soutien sans égaux. Rien de tout cela n'aurait été possible s'il n'était pas avec moi. Je le remercie d'être toujours capable de me faire sourire et de me faire retrouver la sérénité, même dans les moments les plus difficiles de ma vie.

Table des matières

Abstract

Optimal stochastic control problems have a large number of applications in problems of economy, finance (see e.g. the portfolio selection problem in a market with risky assets and non-risky assets, the investment problem and the super-replication price in a model with uncertain volatility), and management of energy. These are typical situations in where we are faced to a dynamical system which evolves under some conditions of uncertainty, and where we have to take a decision at every time, to optimize an economical criterion. In particular, the control variable acts on the state of the system.

Stochastic control problems are historically handled with the Bellman dynamic programming principle, which leads to obtain a characterization of the value function of the optimal control problem as solution of a partial differential equation, said the Hamilton-Jacobi-Bellman equation. In most cases, the value function is not sufficiently smooth to satisfy the HJB-equation in the classical sense. It is for this reason that the notion of viscosity solution, introduced by Crandall and Lions for the deterministic Hamilton-Jacobi-Bellman equation, has been extended to the second order problem by Lions. The theory of viscosity solutions, provided an extremely convenient framework for dealing with the lack of smoothness of the value function of the optimal stochastic control problem.

In some situations, the value function could be smooth : this is the case, for example, of the Merton portfolio selection problem, for which a classical solution of the correspondent HJB-equation can be performed.

However, in the general case, the HJB-equation can not be solved explicitly, hence it is necessary to analyze it numerically. In particular a discretization of the HJB-equation via Markov chain approximation is considered, and an approximate solution is computed.

It is then necessary to guarantee that the numerical solution is a good approximation of the viscosity solution, and for this reason a theory of error estimate has been developed. This theory leads to obtain a theoretical estimate of the differences between the viscosity solution and the discrete solution.

The thesis is divided in two parts. In the first part we give error estimates for a problem on stochastic game theory, and stochastic impulse control problem. Both these problems have some particular difficulties, and classical results on error estimate can not be applied directly.

The second part concerns a study of some algorithms to implement, in particular for two problems : a stochastic impulse control problem, and a problem with unbounded control.

Introduction

L'objet principal de cette thèse est l'étude des approximations numériques de différentes équations Hamilton-Jacobi-Bellman associées à des problèmes de contrôle optimal stochastique.

On considère le problème de contrôle optimal suivant :

$$\begin{cases} u(x) = \inf_{\alpha(\cdot)\in\mathcal{A}} \mathbb{E}\left[\int_0^\infty e^{-\lambda t} f(X(t),\alpha(t))dt\right] \\ dX(t) = b(X(t),\alpha(t))dt + \sigma(X(t),\alpha(t))dW_t \\ X(0) = x \end{cases} \tag{0.1}$$

où $\mathcal{A}$ est l'ensemble des contrôles admissibles :

$$\mathcal{A} := \{\alpha \in L^\infty(0,+\infty) \mid \alpha(t) \in A \text{ a.e.}\},$$

avec A un ensemble de $\mathbb{R}^m$. Les fonctions f, b, σ sont des fonctions de $A \times \mathbb{R}^N$ dans $\mathbb{R}$, $\mathbb{R}^N$, $\mathbb{R}^{N\times P}$ respectivement, et W_t est un mouvement Brownien de dimension N.

On considère les hypothèses suivantes sur les coefficients :

(A1) σ^α est une matrice $N \times P$. Il existe une constante K telle que, pour tout $\alpha \in A$,

$$|\sigma^\alpha|_1 + |b^\alpha|_1 + |f^\alpha|_1 \leq K.$$

(A2) $\lambda > \sup_\alpha\{[\sigma^\alpha]_1^2 + [b^\alpha]_1\}$.

Il est connu (voir [27, 16, 2]) que, sous les hypothèses (A1)-(A2), la fonction valeur du problème (0.1) u est solution de viscosité bornée et Lipschtzienne de l'équation HJB suivante :

$$\sup_{\alpha\in A} L^\alpha(x, u(x), Du(x), D^2u(x)) = 0, \tag{0.2}$$

avec

$$\begin{aligned} L^\alpha(x, u(x), Du(x), D^2u(x)) \;\; &= \;\; -\mathbf{tr}[a(\alpha,x)D^2u(x)] - b(\alpha,x)Du(x) \\ &\quad +\lambda u(x) - f(\alpha,x), \end{aligned}$$

Une solution explicite de (0.2) est très souvent difficile à detérminer. C'est pour cette raison qu'il est nécessaire d'introduire un schéma d'approximation : soit $h \in \mathbb{R}^N$ un pas de discrétisation, considèrons

$$S(h, x, u_h(x), u_h) = 0, \qquad x \in \mathbb{R}^N, \tag{0.3}$$

où $S : \mathbb{R}^N_+ \times \mathbb{R}^N \times \mathbb{R} \times C_b(\mathbb{R}^N) \to \mathbb{R}$ est une approximation de (0.2) sur une grille donnée de pas de discrétisation h. On note $u_h \in C_b(\mathbb{R}^N)$ la solution de (0.3), continue et bornée, qui est une approximation de u. La formulation abstraite du schéma (0.3) a été introduite par Barles et Souganidis [5] pour l'étude de la convergence de la solution du schéma vers la solution de viscosité. Afin de garantir cette convergence, l'opérateur S doit satisfaire les propriétés suivantes :

(S1) Monotonie : pour tout $h > 0$, $r \in \mathbb{R}^N$, $m \geq 0$, $x \in \mathbb{R}^N$ et pour toutes fonctions bornées et continues u, v tells que $u \leq v$ dans $\mathbb{R}^N$,

$$S(h, x, r + m, u + m) \geq \lambda m + S(h, x, r, v).$$

(S2) Regularité : pour tout $h > 0$ et $\phi \in C_b(\mathbb{R}^N)$, $x \mapsto S(h, x, \phi(x), \phi)$ est bornée et continue ; $r \mapsto S(h, x, r, \phi)$ est uniformément continue pour r borné, uniformément en $x \in \mathbb{R}^N$.

(S3) Il existe $n, k_i > 0$, $i \in J \subseteq \{1, \ldots, n\}$ et une constante $K_c > 0$ telle que, pour tout $h > 0$ et x dans $\mathbb{R}^N$, et pour et pour toutes fonctions $\phi \in C^n(\mathbb{R}^N)$ telles que $|D^i \phi|_0$ est bornée, pour tout $i \in J$, on a :

$$|\sup_{\alpha \in A} L^\alpha(x, \mathcal{D}\phi) - S(h, x, \phi(x), \phi)| \leq K_c Q(\phi),$$

où $Q(\phi) := \sum_{i \in J} |D^i \phi|_0 h^{k_i}$.

(S4) Stabilité : pour tout h, l'équation (0.3) admet une unique solution u_h bornée.

Sous (S1)-(S4), la solution discréte u_h converge uniformément, lorsque $h \to 0$, vers la solution de viscosité u de (0.2), [5].

Les hypothèses (S1)-(S4) ne sont pas très restréctives et plusieurs schémas connus les vérifient bien. Citons par exemple, les schémas basés sur les différences finies [17, 8], les méthodes semi-Lagrangiennes [12, 28], et plus généralement les approximations par Chaînes de Markov [26].

Lorsque le problème de contrôle optimal est déterministe, les estimations d'erreur sont (généralement) de l'ordre de $h^{\frac{1}{2}}$ [19]. L'analyse de l'erreur dans le cas du contrôle stochastique est plus délicate. Elle utilise des techniques de perturbations des équations HJB, des résultats précis sur la stabilité des solutions de viscosité. La théorie de l'analyse de l'erreur a été introduite par Krylov [23, 24, 25], Barles et Jakobsen [3, 4].

On donne maintenant un aperçu des techniques existantes pour obtenir l'estimation d'erreur.

Borne supérieure de $u - u_h$ Pour obtenir une borne supérieure, une technique a été introduite par Krylov [23] pour les équations à coefficients constants, et ensuite elle a été étendue au cas des coefficients bornés et Lipschitziens par Barles et Jakobsen [4]. On considère l'équation HJB perturbée, associée à (0.2) :

$$\sup_{\alpha \in A, |e| \leq \varepsilon} L^\alpha(x+e, u^\varepsilon(x), Du^\varepsilon(x), D^2 u^\varepsilon(x)) = 0,$$

et soit u^ε l'unique solution de viscosité bornée et lipschtzienne. Par des techniques de viscosité, il est possible de montrer que $|u - u^\varepsilon| \leq C\varepsilon$, où C est une constante qui dépend de K introduit dans (A1), de λ et des constantes de Lipschitz de u et de u^ε. Ce résultat est l'application d'une estimation plus générale : si on considère deux équations HJB avec coefficients respectivement a, b, f et $\bar{a}, \bar{b}, \bar{f}$, avec solution de viscosité u et $\bar{u}$, respectivement, on arrive à donner une estimation de $u - \bar{u}$ en fonction de $a - \bar{a}$, $b - \bar{b}$, $f - \bar{f}$ et des constantes de Lipschitz de u et $\bar{u}$. On montre ensuite que u^ε est une sous-solution de viscosité de (0.2), et on la régularise par convolution en utilisant un noyau régularisant ρ_ε. On obtient une fonction $u_\varepsilon = u^\varepsilon * \rho_\varepsilon$ régulière. En utilisant la propriété de convexité de l'équation (0.2), on montre qu'une combinaison convexe de sous-solutions de (0.2) est encore une sous-solution de (0.2). On peut alors voir u_ε comme limite de combinaison convexe des u^ε, et donc u_ε est une sous-solution de viscosité de (0.2). De plus, comme elle est régulière, elle est aussi sous-solution au sens classique. En appliquant l'hypothèse de consistance (S3) à u_ε, on arrive à construire une sous-solution du schéma qui dépend de ε. Ensuite, en appliquant un principe de comparaison pour solution discrètes et en optimisant par rapport à ε on arrive au résultat. Avec cette méthode, pour le schéma des Différences Finies une estimation de l'ordre de $h^{1/2}$ a été obtenue, i.e.

$$u - u_h \leq Ch^{1/2},$$

pour tout $x \in \mathbb{R}^N$, où C est une constante qui dépend de K, K_c et λ.

Borné inférieure La borne inférieure est plus difficile à obtenir. La première idée a été introduite par Krylov [23] pour les équations avec coefficients constants, et ensuite elle a été étendue par Barles et Jakobsen pour les équations avec seulement la matrice a^α constante. Cette technique est une approche symétrique à celle utilisée pour obtenir la borne supérieure. On perturbe le schéma et on obtient l'équation

$$\sup_{|e| \leq \varepsilon} S(h, x+e, u_h^\varepsilon(x), u_h^\varepsilon) = 0, \quad x \in \mathbb{R}^N.$$

On régularise u_h^ε par convolution et avec l'hypothèse de consistance on arrive à construire une sous-solution de l'équation (0.2). Finalement, avec le Principe de Comparaison pour solution de viscosité on peut conclure. Cette approche n'aboutit pas pour toutes les équations. Neanmoins, elle permet de traiter le cas où la matrice a est constante, auquel cas on obtient une borne de l'ordre de $h^{1/27}$ pour le schéma des Différences Finies [23].

Récemment Krylov [25] a réussi à étendre cette approche à des équations avec coefficients bornés et lipschitziens, mais seulement pour des schémas de discrétisation particuliers, dans lesquels la diffusion se décompose selon des directions précises avec des coefficients lipschitziens. Pour ce genre de schéma il a obtenu le même résultat que pour la borne supérieure, avec une erreur de l'ordre de $h^{1/2}$. C'est le meilleur résultat obtenu jusqu'à maintenant.

Une autre méthode plus générale a été introduite par Barles et Jakobsen [3], pour les équations avec coefficients bornés et Lipschtziens. Au lieu de perturber le schéma pour construir une sous-solution de (0.2), ils introduisent un "switching system" qui approche l'équation HJB :

$$\min\{\max\{\sup_{\alpha\in\mathcal{A}_i} L^\alpha(x, \mathcal{D}v_i(x)); v_i(x) - \min_{j\neq i}\{v_j(x) + k\}\} = 0,$$

pour $x \in \mathbb{R}^N$ et $i \in \{1, \dots M\}$, $\mathcal{A}_i \subset \mathcal{A}$. Ils considèrent la solution de viscosité de ce système, $v = (v_1, \dots, v_M)$, et ils montrent que pour chaque i, v_i converge vers u, solution de (0.2), quand $k \to 0$. A l'aide de ce système ils construisent une suite de sur-solutions de HJB qu'ils régularisent, et avec la consistance ils arrivent à obtenir une sur-solution du schéma. Avec cette technique, une borne de l'ordre de $h^{1/5}$ a été obtenue, pour le schéma des différences finies.

Dans le chapitre 1 et 2 de la thèse, on a étendu les résultats existants sur les estimations d'erreur à d'autre type de problèmes, en particulier les deux problèmes suivants :

– Un problème à deux joueurs,
– Un problème de contrôle impulsionnel,

Le chapitre 3 est consacré à l'étude numérique d'un problème de contrôle impulsionnel. On présente 2 algorithmes pour le résoudre, et on fait une comparaison numérique entre ces deux méthodes. Dans toute cette partie de la thèse, on a travaillé avec un ensemble de contrôles compact.

On s'est ensuite intéressé à des problèmes de contrôle optimal stochastique, avec contrôle non borné. Cette étude fait l'objet du Chapitre 4. En particulier, on a étudié un problème de contrôle optimal provenant de la finance, associé à un modèle de "pricing" des options dit de sur-couverture. On a considéré l'équation HJB provenant de ce problème et on l'a ensuite discrétisée, avce le schéma des Différences Finies Généralisées. Finalement on a montré l'existence et l'unicité de la solution discrète, et sa convergence vers la solution de viscosité d'une équation HJB.

Chapitre 1. Le problème à deux joueurs

Cette partie de la thèse a fait l'objet de l'article

- J.F. Bonnans, S. Maroso, H. Zidani, *Error estimates for stochastic differential games : the adverse stopping case*, IMA J. Numerical Analysis, 26 : 188-212, 2006. [7]

On considère un jeu à somme nulle, et deux joueurs A et B qui peuvent agir de la fonction suivante :

– A dispose d'un ensemble de contrôle $\mathcal{A}$, et son objectif est de minimiser le gain ;

– B ne dispose pas de contrôles, mais il peut uniquement arrêter le jeu, tout en essayant de maximiser le gain.

En terme de contrôle optimal, le problème s'écrit sous la forme suivante

$$\begin{cases} u(x) = \sup_\tau \inf_{\alpha(\cdot)\in\mathcal{A}} \mathbb{E}\Big[\int_0^\tau e^{-\lambda t} f(X(t),\alpha(t))dt + e^{-\lambda\tau}\psi(\tau)\Big] \\ dX(t) = b(X(t),\alpha(t))dt + \sigma(X(t),\alpha(t))dW_t \\ X(0) = x \end{cases} \tag{0.4}$$

où $\mathcal{A}$ est l'ensemble des contrôles $\alpha(\cdot)$, qui sont à valeurs dans un compact $A \subseteq \mathbb{R}^M$, ψ, f, b, σ, sont des fonctions de $\mathcal{A}\times\mathbb{R}^N$ dans $\mathbb{R}$, $\mathbb{R}^N$, $\mathbb{R}^{N\times P}$ respectivement, bornées et lipschtziennes, et W_t est un mouvement Brownien de dimension N. La fonction ψ représente le paiement que les joueurs reçoivent quand le jeu est arrêté. L'équation HJB de type Isaac associée à ce problème est la suivante :

$$\min\{\sup_{\alpha\in A} L^\alpha(x,u(x),Du(x),D^2u(x)); u(x)-\psi(x)\} = 0. \tag{0.5}$$

Sous les hypothèses (A1)-(A2), (0.5) admet une unique solution de viscosité, qu'on denotera u. On considère maintenant une approximation de (0.5), qui s'écrit :

$$\min\{S(h,x,u_h(x),u_h); u_h(x)-\psi(x)\} = 0, \qquad x\in\mathbb{R}^N, \tag{0.6}$$

où h est le pas de discrétisation, et S est une approximation consistante, monotone et uniformément continue de $\sup_{\alpha\in\mathcal{A}} L^\alpha$. On montre qu'il existe une unique solution de (0.6) u_h, continue et bornée, qui est l'approximation de u. Les schémas typiques qu'on considère sont les Différences Finies [26], les Différences Finies Généralisées [8, 10] et les approximations par Chaînes de Markov [26].

On s'intéresse maintenant à donner une borne supérieure et une borne inférieure de $||u - u_h||$. La difficulté principale ici, consiste dans le fait qu'on n'a plus une équation convexe et donc on ne peut pas appliquer directement les méthodes introduites par Krylov [24], Barles et Jakobsen [4, 3]. On arrive à contourner cette difficulté en combinant ces méthodes avec une séparations des domaines qu'on va expliquer.

Borne supérieure de $u - u_h$. La borne supérieure est plus facile à obtenir que la borné inférieure. En utilisant les idées présentées par Krylov [23], on introduit l'équation perturbée suivante :

$$\min\{\sup_{\alpha,|e|\le\varepsilon} L^\alpha(x+e,u^\varepsilon(x)),Du^\varepsilon(x),D^2u^\varepsilon(x)); u^\varepsilon(x)-\psi(x)\} = 0, \quad x\in\mathbb{R}^N,$$

et on note u^ε l'unique solution de viscosité bornée et Lipschitzienne de cette équation. En régularisant u^ε par convolution, on obtient une fonction u_ε, et on arrive à montrer qu'il existe une constante R qui dépend de ε, telle que

$$\min\{S(h,x,u_\varepsilon(x)-R,u_\varepsilon-R); u_\varepsilon(x)-R-\psi(x)\} \le 0, \quad \forall\, x\in\mathbb{R}^N. \tag{0.7}$$

Donc on peut conclure que $u_\varepsilon - R$ est une sous-solution du schéma, et avec le principe de comparaison discret on conclut.

La partie plus délicate est la démonstration de (0.7), pour laquelle il faut introduire l'ensemble suivant :

$$X(u^\varepsilon) = \{x \in \mathbb{R}^N | u^\varepsilon(x) = \psi(x)\}.$$

Ensuite on étudie les cas séparément :

(a) Pour les $x \in X(u^\varepsilon)$, on a $u^\varepsilon(x) = \psi(x)$, et en régularisant on obtient $u_\varepsilon(x) - C\varepsilon \leq \psi(x)$, où C dépend des constantes de Lipschitz de ψ et de u^ε.

(b) Pour les $x \notin X(u^\varepsilon)$, on applique la méthode standard [4], qui utilise la consistance et les propriétés de la fonction régularisée, et on arrive au résultat.

Le résultat final est le suivant :

Theorem .1. *Sous les hypothèses (A1)-(A2) et (S1)-(S4), on a*

$$u - u_h \leq Ch^\ell,$$

où $\ell = \min_{j \in J}\{k_i/i\}$, *et* C *dépend seulement de* λ, K, K_c *et* C_ψ. □

On rappelle que les k_i et K_c viennent de l'hypothèse de consistance (S3), K et λ viennent de (A1)-(A2), et C_ψ est la constante de Lipschitz de ψ.

La borne inférieure de $u - u_h$ La borne inférieure est plus compliquée à obtenir, et on ne peut pas appliquer une technique symétrique. A la place d'une sur-solution régulière, on construit une suite de sur-solutions locales régulières, comme dans [3]. En particulier on introduit le "switching system" suivant :

$$\min\{\max\{L^{\alpha_i}(x, \mathcal{D}v_i(x)); v_i(x) - \min_{j \neq i}\{v_j(x) + k\}\}; v_i(x) - \psi(x)\} = 0,$$

pour $x \in \mathbb{R}^N$, et $i \in \mathcal{I} = \{1, \ldots, M\}$. On montre que ce système a une unique solution de viscosité $v = (v_1, \ldots, v_M)$ et on donne aussi un taux de convergence de v vers u. En particulier on a le résultat suivant :

Theorem .2. *Pour tout* $i \in \mathcal{I}$ *et pour tout* $x \in \mathbb{R}^N$ *on a*

$$0 \leq v_i - u \leq Ck^{1/3},$$

où C *dépend de* λ *et* K *définies en (A1).* □

La perturbation du "switching-system" conduit à l'équation suivante :

$$\min\{\max\{\inf_{|e| \leq \varepsilon} L^{\alpha_i}(x + e, \mathcal{D}v_i^\varepsilon(x)); v_i^\varepsilon(x) - \min_{j \neq i}\{v_j^\varepsilon(x) + k\}\}; v_i^\varepsilon(x) - \psi(x)\} = 0,$$

pour tout $i \in \mathcal{I}$ et $x \in \mathbb{R}^N$, et on dénote $v^\varepsilon = (v_1^\varepsilon, \ldots, v_M^\varepsilon)$ la solution de viscosité. On régularise v^ε par convolution, on obtient la fonction v_ε qui est une sur-solution locale de $\sup_\alpha L^\alpha(x, v_\varepsilon(x), Dv_\varepsilon(x), D^2v_\varepsilon(x)) \geq 0$.

Pour arriver à l'estimation d'erreur, on doit alors considérer les 2 ensembles suivants :

$$X := \{x \in \mathbb{R}^N | u_h(x) = \psi(x)\}; \quad Y := \{x \in \mathbb{R}^N | S(h, x, u_h(x), u_h) = 0\}.$$

Pour les $x \in Y$, on applique la méthode introduite par Barles et Jakobsen : on applique l'hypothèse de consistance à la sur-solution v_ε et on obtient une sur-solution du schéma. Enfin, le principe de comparaison discret nous permet de conclure. Pour les $x \in X$, il est suffisant de noter que $u_h(x) = \psi(x) \leq u(x)$, pour tout x.

En combinant ces deux résultat on a

Theorem .3. *Sous les hypothèses (A1), (A2) et (S1)-(S3), et en supposant que le schéma a une unique solution u_h dans $C_{b,l}(\mathbb{R}^N)$, on a que*

$$u_h - u \leq Ch^\ell, \qquad \forall\, x \in \mathbb{R}^N,$$

où $\ell = \min_{i \in J}\{k_i/(3i-2)\}$ et C dépend de λ, K et K_c. $\square$

Dans ce chapitre de la thèse on fait aussi une étude des différents schémas numériques et on donne les conditions sous lesquelles un schéma peut donner la meilleure estimation d'erreur. Cette étude reste d'ailleurs valable pour les approximations de toutes les équations HJB.

Chapitre 2. Le problème de contrôle impulsionnel

Cette partie de la thèse fait l'objet de l'article

- J.F. Bonnans, S. Maroso, H. Zidani, *Error estimates for a stochastic impulse control problem*, à paraiitre dans Applied Mathematics and Optimization, version finale acceptée en mars 2006. [6]

Le problème de contrôle optimal impulsionnel est le suivant

$$\begin{cases} u(x) = \inf_{\alpha \in \mathcal{A}} \left\{ \int_0^{+\infty} f^\alpha(X_t) e^{-t} dt + \sum_{i=1}^{+\infty} (k + c(\xi_i)) e^{-\theta_i} \right\} & \\ dX_t = b^\alpha(X_t)dt + \sigma^\alpha(X_t)dW_t, & t \in]\theta_i, \theta_{i+1}[\\ X_{\theta_i^+} = X_{\theta_i^-} + \xi_i, & i \in \mathbb{N} \\ X_0 = x_0 & \end{cases} \tag{0.8}$$

où θ_i est une suite de temps d'arrêt, ξ_i sont les impulsions et c est un coût de transaction. L'équation HJB associée à ce problème est donc

$$\max\{\sup_{\alpha \in A} L^\alpha(x, u(x), Du(x), D^2u(x)); u(x) - \mathcal{M}u(x)\} = 0,\ x \in \mathbb{R}^N, \tag{P}$$

où

$$\begin{cases} \mathcal{M}u(x) := k + \inf_{\xi \in \mathbb{R}_+^N}\{u(x+\xi) + c(\xi)\}, \\ k > 0, \quad c : \mathbb{R}_+^N \to \mathbb{R}_+, \\ c(0) = 0, \quad c(\xi_1 + \xi_2) \leq c(\xi_1) + c(\xi_2). \end{cases} \tag{0.9}$$

Soit u l'unique solution de viscosité bornée et Lipschitzienne de (P).

On a donc que cette équation peut être vue comme un problème d'obstacle, où l'obstacle dépend aussi de la fonction valeur.

On considère un schéma d'approximation monotone de (P), de la forme suivante :

$$\max\{S(h,x,u_h(x),u_h); u_h(x) - \mathcal{M}u_h(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{S}$$

où S est une approximation monotone et consistante de $\sup_\alpha \mathcal{L}^\alpha$.

La difficulté principale pour obtenir l'estimation d'erreur vient du fait que la fonction valeur apparaît aussi dans l'"obstacle" $\mathcal{M}u$. Cela pose un problème aussi pour démontrer l'unicité de la solution de viscosité de (P), qu'on résout en utilisant les techniques introduites par Barles [2], pour le l'équation à l'ordre 1.

En utilisant la méthode introduite par Ishii [22] pour démontrer l'existence d'une solution de viscosité, on approche (P) par une suite de problème à cascade (Pn), $n \geq 0$, où dans chaque problème on admet seulement un nombre fini d'impulsions.

On considère le premier problème, dit le problème sans impulsions :

$$\sup_{\alpha_i} L^{\alpha_i}(x, u_0(x), Du_0(x), D^2u_0(x)) = 0, \quad x \in \mathbb{R}^N. \tag{P_0}$$

Sous les hypothèses (A1-A2), cette équation a une unique solution de viscosité u_0 dans $C_{b,l}(\mathbb{R}^N)$. Ensuite on considère maintenant le problème avec une impulsion, dans lequel l'obstacle dépend de u_0 :

$$\max\{\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}u(x)); u(x) - \mathcal{M}u_0(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{P1}$$

Comme $\mathcal{M}u_0$ est uniformément continue, sous les hypothèses (A1-A2), il existe une unique solution de viscosité u_1 de (P1) dans $C_{b,l}(\mathbb{R}^N)$. De la même façon, pour $n = 2, 3, \cdots$, soit $u_n \in C_{b,l}(\mathbb{R}^N)$ l'unique solution de viscosité du problème avec n impulsions :

$$\max\{\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}u(x)); u(x) - \mathcal{M}u_{n-1}(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{Pn}$$

En utilisant les propriétés de l'opérateur $\mathcal{M}$, et la comparaison entre solutions de viscosité, on obtient que la suite de problèmes $(P_n)_n$ engendre une suite de solutions $(u_n)_n$ telle que

$$0 \leq \cdots \leq u_n \leq \cdots \leq u_2 \leq u_1 \leq u_0. \tag{0.10}$$

On suppose maintenant que $|u_0|_0 > k$, et soit $\mu \in (0,1)$ tel que $\mu|u_0|_0 < k$.

On a alors montré que

$$u_n - u \leq \frac{(1-\mu)^n}{\mu}|u_0|_0,$$

pour tout $n \geq 0$.

De la même façon on approche (S) par une suite à cascade de schémas (Sn), $n \geq 1$,

$$\max\{S(h,x,u_h(x),u_h); u_h(x) - \mathcal{M}u_{h(n-1)}(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{S_n}$$

Soit u_{hn} la solution de (Sn), continue et bornée. Avec la même méthode que pour les solutions de viscosité, on montre que $(u_{hn})_n$ forme une suite décroissante de fonctions positives, et de plus pour tout $n \geq 0$,

$$u_{hn} - u_h \leq \frac{(1-\mu)^n}{\mu} |u_{h0}|_0.$$

En adaptant les méthodes introduites par Barles et Jakobsen [3] à notre problème avec obstacle, on obtient une borne supérieure et une borne inférieure de $u_n - u_{hn}$, pour tout $n < +\infty$. En effet, l'extension de ces méthodes dans le cas d'un problème convexe avec obstacle fixe est facile.

On a alors le résultat suivant.

Proposition .4. *Sous les hypothèses (A1)-(A2) et (S1)-(S4), soient* $u_n \in C_{b,l}(\mathbb{R}^N)$ *l'unique solution de viscosité de (Pn), et* $u_{hn} \in C_b(\mathbb{R}^N)$ *l'unique solution de (Sn-impulse),* $n \geq 1$. *Alors, on a*

$$u_n(x) - u_{hn}(x) \leq \bar{C}_n |h|^{\bar{\gamma}}, \qquad (\bar{E}n)$$

où

$$\bar{C}_n = \bar{C}_{n-1} + C\varepsilon, \qquad (0.11)$$

$\bar{\gamma} = \min_i \{k_i/i\}$, *et* C *est une constante qui dépend de* K *défini dans (A1), de* λ *et des constantes de Lipschitz de* u_n.

En général on obtient que, pour tout $n \geq 1$, $u_n - u_{hn} \leq \bar{C}_n |h|^{\bar{\gamma}}$. On pose

$$\bar{D}_{n-1} := \bar{C}_n - \bar{C}_{n-1}.$$

La suite des $\bar{C}_n$ est croissante, mais on arrive à montrer que $\bar{D}_n \leq \bar{D}_0$, et de plus :

$$\bar{C}_n \leq \bar{C}_0 + n\bar{D}_0.$$

On peut alors donner une estimation de la borne supérieure de $u - u_h$. Comme

$$\sup_x (u_h(x) - u(x)) \leq \sup_x (u_h(x) - u_{hn}(x)) + \sup_x (u_{hn}(x) - u_n(x))$$
$$+ \sup_x (u_n(x) - u(x)),$$

en choisissant le n optimal, on obtient le résultat.

Ce résultat s'applique au schéma des différences finies classiques, aux différences finies généralisées et à tous les schémas du type chaînes de Markov.

La borne inférieure Pour la borne inférieure on analyse d'abord l'erreur entre u_n et u_{hn}, pour tout $n < \infty$, en utilisant la technique du "switching system" introduite dans [3]. Dans le cas d'un nombre fini d'impulsions on a alors le résultat suivant :

Proposition .5. *Sous les hypothèses (A1)-(A2) et (S1)-(S4), soit $u_n \in C_{b,l}(\mathbb{R}^N)$ la solution de viscosité de (Pn) et soit $u_{hn} \in C_{b,l}(\mathbb{R}^N)$ la solution de (Sn), $n \geq 1$. Alors on a*

$$u_{hn}(x) - u_n(x) \leq \underline{C}_n |h|^{\underline{\gamma}}, \quad \forall\, x \in \mathbb{R}^N, \tag{\underline{En}}$$

$$\underline{C}_n = \underline{C}_{n-1} + C\varepsilon, \tag{0.12}$$

$\underline{\gamma} = \min_i\{k_i/(3i-2)\}$*, et C dépend de K, de λ et de la constante de Lipschitz de u_n.*

En général on obtient que, pour tout $n \geq 1$, $u_{hn} - u_n \leq \underline{C}_n |h|^{\underline{\gamma}}$. On pose

$$\underline{D}_{n-1} := \underline{C}_n - \underline{C}_{n-1}.$$

La suite des $\underline{C}_n$ est croissante, mais on arrive à montrer que $\underline{D}_n \leq \underline{D}_0$, et de plus :

$$\underline{C}_n \leq \underline{C}_0 + n\underline{D}_0.$$

Finalement, comme

$$\sup_x (u_h(x) - u(x)) \leq \sup_x (u_h(x) - u_{hn}(x)) + \sup_x (u_{hn}(x) - u_n(x)) + \sup_x (u_n(x) - u(x)),$$

en choisissant le n optimal, on a le résultat.

Theorem .6. *Sous les hypothèses (A1-A2) et (S1-S4), soit $u \in C_{b,l}(\mathbb{R}^N)$ l'unique solution de viscosité de (P), et $u_h \in C_b(\mathbb{R}^N)$ l'unique solution de (S). On a alors la borne suivante :*

$$-C|h|^{\underline{\gamma}} \leq u - u_h \leq C|h|^{\bar{\gamma}}, \tag{0.13}$$

où C est une constante borné qui dépend de K défini dans (A1), et de la vitesse de convergence de u_n et u_{hn}.

Chapitre 3. Un algorithme pour la résolution d'un problème de contrôle impulsionnel

Dans ce chapitre on présente une approche pour la résolution numérique du problème de contrôle impulsionnel présenté dans le chapitre précédent. Cet algorithmes se base sur l'approche théorique de la cascade qu'on a étudié dans le chapitre précédent, et sur l'algorithme de Howard.

On donne aussi quelques résultats sur l'algorithme de Howard (connu aussi sous le nom d'iterations sur les politiques). Plus précisement, on propose une preuve simple et directe pour la convergence surlinéaire, et on prouve que cet algorithme est exactement équivalent à la méthode Primal-Dual étudiée dans [21].

Rappelons ici le problème qu'on souhaite étudier :

$$\max\{\sup_{\alpha\in A}\mathcal{L}^\alpha(x,u(x),Du(x),D^2u(x));u(x)-\mathcal{M}u(x)\}=0,\ x\in\mathbb{R}^N, \tag{0.14}$$

où

$$\begin{cases}\mathcal{M}u(x):=k+\inf_{\xi\in\mathbb{R}^N_+}\{u(x+\xi)+c(\xi)\},\\ k>0, \quad c:\mathbb{R}^N_+\to\mathbb{R}_+,\\ c(0)=0, \quad c(\xi_1+\xi_2)\le c(\xi_1)+c(\xi_2).\end{cases} \tag{0.15}$$

Une méthode possible de calculer la solution de (0.14) est de discrétiser l'espace des états $\mathbb{R}^N$ en introduisant une grille discrète régulière $\mathcal{O}_h\in\mathbb{R}^N$. On denote avec N_{tot} le nombre fini de points dans la grille. La discrétisation de (0.14) sur cette grille, peut être generalement interprétée comme un problème de contrôle impulsionnel pour le contrôle optimal d'une chaîne de Markov ; voir [10, 15, 20, 26] et le références dans ces articles. Soient $\mathcal{L}_h$ et $\mathcal{M}_h$ les discrétisations des operateurs $\mathcal{L}$ et $\mathcal{M}$ respectivement. Alors l'équation discrète s'écrit comme suit :

$$\max\Big\{\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_h-f(\alpha));V_h-\mathcal{M}_hV_h\Big\}=0. \tag{0.16}$$

En particulier $\mathcal{L}_h^\alpha$ est une matrice de dimension $N_{tot}\times N_{tot}$, V_h, $\mathcal{M}_hV_h$ et $f(\alpha)$ sont des vecteurs de dimension N_{tot}. Une politique pour cette équation discrète est une application $\mathcal{O}_h\to\mathcal{A}$; on dénote avec $\mathcal{A}^{N_{tot}}$ l'ensemble des politiques. L'operateur $\mathcal{L}_h^\alpha$ est supposé, pour une politique donnée α, linéaire, et tel que $\mathcal{L}_h^\alpha V_h\ge 0$ implique $V_h\ge 0$. On suppose aussi que les éléments non diagonaux de la matrice $\mathcal{L}_h^\alpha$ sont non positifs, et que si V_h est constant, alors $\mathcal{L}_h^\alpha V_h=\lambda V_h$.

L'étude numérique du problème (0.16) est liée à la résolution du problème de l'obstacle, formulé comme suit :

$$\max\{\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_h-f(\alpha));\lambda V_h-\psi\}=0, \tag{0.17}$$

où l'obstacle ψ est un vecteur de dimension N_{tot}. Le problème (0.16) peut être donc vu comme un problème d'obstacle dans lequel l'obstacle dépend de la solution V_h.

Un problème plus simple à résoudre est l'équation HJB standard (sans obstacle)

$$\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_h-f(\alpha))=0. \tag{0.18}$$

Pour les problèmes en horizon infini, le problème d'obstacle peut être réduit au problème standard. On a alors le résultat suivant :

Lemma .7. *On considère un contrôle* $\alpha_{obs} \notin \mathcal{A}$. *On pose* $\hat{\mathcal{A}} = \mathcal{A} \cup \{\alpha_{obs}\}$, *et on définit* $\hat{\mathcal{L}}$, $\hat{f}$ *par*

$$\begin{array}{lll} (\hat{\mathcal{L}}_h^\alpha V_h)_i = (\mathcal{L}_h^\alpha V_h)_i, & si & \alpha_i \in \mathcal{A}, \\ (\hat{\mathcal{L}}_h^\alpha V_h)_i = (\lambda V_h)_i, & si & \alpha_i = \alpha_{obs} \\ (\hat{f}(\alpha))_i = (f(\alpha))_i & si & \alpha_i \in \mathcal{A}, \\ (\hat{f}(\alpha))_i = \psi_i, & si & \alpha_i = \alpha_{obs}. \end{array}$$

Alors le problème d'obstacle (0.17) *est équivalent au problème standard* (0.18), *avec donnés* $\hat{\mathcal{A}}$, $\hat{\mathcal{L}}_h$ *et* $\hat{f}$. □

Dans la litterature, l'équation de Bellman associée au problème de contrôle optimal d'une chaîne de Markov à été étudié par de nombreux auteurs, (voir, par exemple [26], [20], [15]). (0.18) peut s'écrire aussi sous la forme

$$V_h = \beta \inf_{\alpha \in \mathcal{A}^{N_{tot}}} (V_h + \Delta t(\lambda V_h - \mathcal{L}_h^\alpha V_h + f(\alpha))),$$

où Δt est un pas de temps fictif, et $\beta := (1 + \lambda \Delta t)^{-1}$ est le taux d'actualisation discret. Si $\Delta t > 0$ est suffisament petit, en vue des hypothèses qu'on a fait sur $\mathcal{L}_h^\alpha$, la matrice M_h^α définie by $M_h^\alpha V_h = V_h + \Delta t(\lambda V_h - \mathcal{L}_h^\alpha V_h)$ a des coefficients non-negatifs, et la somme sur chaque ligne est égale à 1. Donc pour chaque politique α, M_h^α est la matrice de transition d'une chaîne de Markov.

La réformulation par point fixe de (0.18) est la base des deux principaux algotithmes, les itérations sur les valeurs

$$V_h^{k+1} = \beta \inf_\alpha \{M_h^\alpha V_h^k + f(\alpha)\},$$

et les itérations sur les politiques (dû à Howard), qui consiste, pour une politique α^k à l'étape k, à résoudre le système linéaire :

$$\mathcal{L}_h^{\alpha^k} V_h^{k+1} - f(\alpha^k) = 0,$$

ce qui revient à résoudre l'équation :

$$V_h^{k+1} = \beta(M_h^{\alpha^k} V_h^{k+1} + f(\alpha^k)).$$

Vient ensuite une étape de mise à jour de la politique en utilisant la formule

$$\alpha^{k+1} \in \text{argmax}\{\mathcal{L}_h^\alpha V_h^k + f(\alpha)\},$$

qui est équivalent à

$$\alpha^{k+1} \in \text{argmin}\{M_h^\alpha V_h^k + f(\alpha)\}.$$

On s'est intéressé, dans ce chapitre, à étudier la vitesse de convergence de l'algorithme de Howard pour le problème (0.18). On montre que, dans le cas où $\mathcal{A}$ est un ensemble fini tel que $card(\mathcal{A}) = p$, l'algorithme de Howard converge en un nombre fini d'itérations, borné par p^N. De plus il est connu (voir par exemple [9]) que si $\mathcal{A}$ est un ensemble dénombrable, alors cet algorithme a une convergence linéaire. On démontre dans ce chapitre la convergcence surlinéaire de l'algorithme de Howard dans le cas où $\mathcal{A}$ est un ensemble compact. Ce résultat a été prouvé dans [30, 31] pour des problèmes particulièrs et sous des hypothèses particulières. On donne dans ce chapitre une preuve qui nous semble claire et simple de ce résultat dans un cadre général.

Si on réecrit (0.18) sous la forme

$$\max_{\alpha \in \mathcal{A}^{N_{tot}}} (A(\alpha)x - f(\alpha)) = 0, \tag{0.19}$$

avec $A(\alpha) = \mathcal{L}_h^\alpha$, on a le résultat suivant.

Theorem .8. *Soit $\mathcal{A}$ un ensemble non vide et compact, et $A : \mathcal{A}^N \to \mathbb{R}^{N\times N}$, $f : \mathcal{A}^N \to \mathbb{R}^N$, des fonctions continue telles que $A(\alpha)$ est monotone pour tout α. Alors (0.19) a une unique solution x^*, et l'algorithme de Howard converge sur-linéairement, i.e. $\lim_{k\to\infty} x^k = x^*$ and*

$$||x^{k+1} - x^*|| = o\Big(||x^k - x^*||\Big), \qquad \textit{quand } k \to \infty.$$

Comme le problème d'ostacle (0.17) peut s'écrire comme un problème standard (0.18), on peut appliquer les mêmes méthodes pour le résoudre. En particulier on considère l'algorithme de Howard appliqué au problème

$$\max(M_1 V_h - b; M_2 V_h - \psi) = 0, \tag{0.20}$$

où les matrices M_i sont monotones de dimension $N_{tot} \times N_{tot}$(i.e. $M_i X \geq 0 \Rightarrow X \geq 0$), $i = 1, 2$, V_h, b et ψ sont vecteurs dans $\mathbb{R}^{N_{tot}}$, et V_h est la solution du problème. Au fai, on a que (0.20) n'est rien d'autre que une discrétisation de (0.17) Dans la litterature, pour résoudre (0.17), des différentes méthodes sont utilisées, et on considère en particulier la méthode des Contraintes Actives (voir [21]). On démontre dans ce chapitre l'équivalence entre la méthode de contraintes actives et l'algorithme de Howard pour le problème d'obstacle.

Le problème

$$\max\{ \sup_{\alpha \in \mathcal{A}^{N_{tot}}} \{\mathcal{L}_h^\alpha V_h - f(\alpha)); V_h - \mathcal{M}_h V_h\} = 0, \tag{0.21}$$

est une sorte de problème d'obstacle dans lequel l'obstacle dépend de la fonction valeur. Tenant compte de l'idée de cascade introduite dans le chapître précédent, on propose l'algorithme de résolution suivant :

Un algorithme de type Cascade implemantable

Data $\mathcal{A}$,$\mathcal{L}_h^\alpha$, $f(\alpha)$, $\mathcal{M}$, une suite m_n de nombre entiers positifs, α, politique initiale ; $k := 0$.

Init Faire m_0 itérations de l'algorithme de Howard pour résoudre (S_0) ; le résultat est une estimation supérieure $\tilde{V}_{h0}$ de V_{h0}.

Loop Pour $k = 1, 2, \ldots$, on définit le problème (avec incounnue V_{hn})

$$\max_{\alpha \in \{\mathcal{A} \cup \{\alpha_{obs}\}\}^{N_{tot}}} \{(\mathcal{L}_h^\alpha V_{hn} + f(\alpha); V_{hn} - \mathcal{M}_h \tilde{V}_{h(n-1)}\} = 0.$$

Faire m_n itérations de l'algorithme de Howard pour résoudre le problème d'obstacle (S_n) ; la politique initiale est la dernière obtenue dans la dernière étape de l'itération précédente. Le résultat est une estimation supérieure $\tilde{V}_{hn}$ de V_{hn}.

End

Si on prends $m_n = +\infty$ dans l'algorithme, on recouvre la cascade. Une autre choix extreme est de choisir $m_n = 1$, pour tout n, c.a.d. mettre à jour la valeur après impulsion après chaque itération. Comme démontré dans le théorème suivant, ce dernier est en effect la cas le plus efficient et moins coûteaux.

Dans le théorème, on utilise un conteur ℓ, pour conter le nombre de fois qu'on passe par l'algorithme de Howard, i.e. il est le nombre totale d'iterations sur les politiques qui a été fait. Par example, à l'iteration i de l'algorithme de Howard pour l'itération k de l'algorithme totale, on aura

$$\ell = i + \sum_{j=0}^{k-1} m_j. \tag{0.22}$$

Theorem .9. *Soit ℓ le conteur total des iterations sur les politiques. La valeur correspondente qu'on denotera V_{hn} est une fonction decroissante de la suite m_n. En d'autres termes, si le coût du calcul d'une impulsion est negligeable par rapport au coût des iterations sur les politiques, la convergence va plus vite quand le calcul de l'impulsion est fait moin souvent.*

On conclut le chapitre par des résultats numériques où on teste l'efficacité des deux approches, avec $m_n = +\infty$ et $m_n = 1$.

Chapitre 4. Approximation numérique d'un problème de sur-replication avec contraintes gamma

Le dernier chapitre de la thèse est consacré à l'étude de l'approximation numérique d'un problème de contrôle stochastique non borné, associé en particulier à un modèle de "pricing" des options, dit problème de sur-couverture [29, 32, 33].

Dans un marché financier, où on a un actif risqué et un actif non risqué, on s'intéresse à étudier le plus petit capital initial nécessaire pour couvrir un certain prix donné. Plusieurs auteurs ont travaillé sur ce problème dans des cas différents et avec différentes contraintes : par exemple [18, 32] pour des problèmes en dimension 1, [11] pour des problèmes en dimension 2, et [33, 14] pour des problèmes en dimension générale d. Dans tous ces articles, les auteurs montrent que le prix de sur-replication est solution de viscosité d'une équation HJB, avec conditions au bord et condition finale.

Dans ce chapitre on étudie numériquement une équation HJB qui provient d'un problème de sur-couverture en dimension 2. On discrétise l'équation HJB en utilisant le schéma des

Différences Finies Généralisées [8, 10], et on étudie l'existence et l'unicité de la solution discrète. Finalement on démontre la convergence de la solution numérique vers la solution de viscosité.

En particulier, on considère l'équation HJB qui provient du problème de sur-couverture en dimension 2 introduit par [11] :

$$\vartheta(t,x,y) = \sup_{(\rho,\xi)\in\mathcal{U}} \mathbf{E}\left[g\left(X^{\rho,\xi}_{t,x,y}(T)\right)\right], \tag{0.23}$$

où

$$\mathcal{U} := \left\{(\rho,\zeta) \text{ à valeurs dans } [-1,1]\times(0,+\infty) \text{ et } \mathcal{F}_t\text{-mesurable} \mid \int_0^T \zeta_t^2\, dt < +\infty\right\},$$

g est le payoff, et le processus $(X^{\rho,\xi}_{t,x,y}, Y^{\rho,\xi}_{t,x,y})$ est un processus positif en dimension 2 qui évolue en suivant le système dynamique stochastique :

$$dX^{\rho,\xi}_{t,x,y}(s) = \sigma(s, Y^{\rho,\xi}_{t,y}(s))X^{\rho,\xi}_{t,x,y}(s)dW^1_s, \qquad s\in(t,T) \tag{0.24a}$$

$$dY^{\rho,\xi}_{t,y}(s) = -\mu(s, Y^{\rho,\xi}_{t,y}(s))ds + \xi(s)Y^{\rho,\xi}_{t,y}(s)dW^2_s, \qquad s\in(t,T) \tag{0.24b}$$

$$\langle dW^1_s, dW^2_s\rangle = \rho(s), \qquad \text{a.e } s\in(t,T) \tag{0.24c}$$

$$X^{\rho,\xi}_{t,x,y}(t) = x\ ,\ Y^{\rho,\xi}_{t,y}(t) = y, \tag{0.24d}$$

où W^1_s et W^2_s sont des mouvements Browiniens, et σ est la volatilité.

Les variables $X^{\rho,\xi}_{t,x,y}$ et $Y^{\rho,\xi}_{t,y}$ décrivent deux actifs dans un marché financier. Le premier actif $X^{\rho,\xi}_{t,x,y}$ est risqué, et le deuxième $Y^{\rho,\xi}_{t,y}$ est un actif tel que son prix est lié à $X^{\rho,\xi}_{t,x,y}$ à travers la volatilité $\sigma(t, Y^{\rho,\xi}_{t,y})$. On suppose que les coefficients μ et σ satisfaient les hypothèses suivantes :

(A1) $\sigma : [0,T]\times\mathbb{R} \to \mathbb{R}^+$ est une fonction positive, telle que σ^2 est Lipschitzienne. Pour tout $t\in[0,T]$, $\sigma(t,0) = 0$ (typiquement $\sigma(t,y) = \sqrt{y}$).

(A2) $\mu : (0,T)\times\mathbb{R}^+ \to \mathbb{R}^+$ est une fonction positive, Lipschtzienne, et telle que $\mu(t,0) = 0$ pour tout $t\in[0,T]$.

Pour obtenir une fonction valeur ϑ bornée et Lipschitzienne on va supposer que g est une fonction telle que

(A3) g est bornée et Lipschtienne. Soit $M_0 > 0$ telle que : $\|g\|_\infty \le M_0$.

(A4) La fonction $f : z \to g(e^z)$ est Lipschtzienne.

(A5) $g \in \mathcal{C}^2(\mathbb{R}^+ \to \mathbb{R})$. Les fonctions $x \to xg'(x)$ et $x \to x^2 g''(x)$ sont bornées.

Si maintenant on considère l'Hamiltonien associé à ce problème, on a, pour $t\in[0,T)$, $x,y\in(0,\infty)^2$, $p=(p_1,p_2)\in\mathbb{R}^2$ et $Q\in\mathcal{M}_2$,

$$H(t,x,y,p,Q) := \inf_{(\xi,\rho)\in\mathbb{R}_+\times[-1,1]} \left\{\mu(t,y)p_2 - \frac{1}{2}\mathrm{tr}\left(a(t,x,y,\xi,\rho)\cdot Q\right)\right\},$$

où la matrice de covariance a est donnée par :

$$a(t,x,y,\xi,\rho) = \begin{pmatrix} \sigma^2(t,y)x^2 & \rho\xi\sigma(t,y)x \\ \rho\xi\sigma(t,y)x & \xi^2 \end{pmatrix}.$$

On voit bien que l'Hamiltonien est une minimization sur un ensemble de contrôles non bornés, et donc il peut ne pas être fini. La première difficulté est alors que le Hamiltonien associé à (0.23) peut ne pas être borné, et on n'est pas capable de traiter numériquement ce type de problèmes. De façon formelle, on a que ϑ satisfait l'équation suivante :

$$-\frac{\partial \vartheta}{\partial t} + H\left(t,x,y,D\vartheta,D^2\vartheta\right) = 0 \quad (t,x,y) \in (0,T) \times (0,+\infty) \times (0,+\infty).$$

Dans la littérature, les problèmes avec contrôle non borné ont été étudiés, par exemple, par [1], [13] et d'autres. Dans tous ces cas, les auteurs ont décidé de tronquer l'ensemble des contrôles pour se ramener à un ensemble de contrôles bornés et donc simplifier l'analyse numérique. Dans ce chapitre on ne tronque pas l'ensemble des contrôles, car on trouve une formulation particulière pour l'équation HJB, et cette formulation nous permet de contourner la difficulté du contrôle non borné avec d'autres techniques. On démontré au fait le résultat suivant :

Theorem .10. *Pour $t \in [0,T)$, $x,y \in \mathbb{R}^+$, sous les hypothèses (A1)-(A2), la fonction valeur ϑ est une solution de viscosité discontinue de :*

$$\Lambda^-(J(t,x,y,\partial_t\vartheta(t,x,y), D\vartheta(t,x,y), D^2\vartheta(t,x,y))) = 0, \tag{0.25}$$

où

$$J(t,x,y,r,p,Q) := \begin{pmatrix} -\frac{\partial\vartheta}{\partial t} + \mu(t,y)\frac{\partial\vartheta}{\partial y} - \frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2} \end{pmatrix},$$

et $\Lambda^-(J)$ représente la plus petite valeur propre de la matrice J. De plus, ϑ est une sur-solution de viscosité discontinue de

$$-\frac{\partial^2\vartheta}{\partial y^2} \geq 0. \tag{0.26}$$

On démontre d'abord que ϑ est une solution de viscosité discontinue, et ensuite, grâçe au principe de comparaison pour solution de viscosité, on démontre que ϑ est Lipschtzienne, quand les hypothèses (A3-5) sont vérifiées.

L'équation (0.25) n'est pas facile à traiter numériquement. Cependant, à partir des calculs standards d'algèbre, on réecrit la plus petite valeur propre sous la forme suivante :

$$\Lambda^-(J) = \min_{\alpha_1^2+\alpha_2^2=1} \alpha^T J \alpha,$$

où $\alpha \in \mathbb{R}^2$, et on a alors :

Corollary .11. *Sous les hypothèses (A1)-(A3), la fonction valeur ϑ est une solution de viscosité de l'équation HJB :*

$$\inf_{\alpha_1^2+\alpha_2^2=1}\left\{\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}^T\begin{pmatrix}-\frac{\partial\vartheta}{\partial t}+\mu(t,y)\frac{\partial\vartheta}{\partial y}-\frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y}\\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}\right\}=0. \tag{0.27}$$

On a donc transformé le problème de départ en un problème équivalent avec contrôle borné et maintenant une analyse numérique est possible. Pour cette équation, des conditions au bord et un principe d'unicité ont été prouvés dans [11]. En particulier, les conditions au bord sont les suivantes :

$$\begin{aligned}\vartheta(T,x,y)=g(x), &\quad \forall\ (x,y)\in(0,\infty)^2\\ \vartheta(t,x,0)=g(x), &\quad \forall\ (t,x)\in[0,T)\times(0,\infty)^2.\end{aligned}$$

On discrétise l'équation HJB (0.27) par le schéma des différences finies généralisées et on obtient :

$$\begin{aligned}S^\rho(t,x,y,r,\phi) \quad &= \min_{\alpha_1^2+\alpha_2^2=1}\Bigg\{-\alpha_1^2\frac{\phi(t+\Delta t,x,y)-r}{\Delta t}+\alpha_1^2\mu\frac{r-\phi(t,x,y-h)}{h}\\ -\frac{1}{2}\sum_{\xi\in\mathcal{S}(x,y)}&\gamma_\xi^{\alpha_1,\alpha_2}(t,x,y)\big[\phi(t,x-\xi_1h,y-\xi_2h)-2r+\phi(t,x+\xi_1h,y+\xi_2h)\big]\Bigg\},\end{aligned} \tag{0.28}$$

pour $(t,x,y)\in[0,T)\times(0,\infty)^2$, où

$$\mathcal{S}(x,y):=\mathcal{S}_p \qquad \text{avec } p=\min(p_{max},\lceil x/h\rceil,\lceil y/h\rceil), \tag{0.29}$$

$$\mathcal{S}_p=\{(\xi_1,\xi_2)\in\mathbb{Z}\times\mathbb{N};\max(|\xi_1|;\xi_2)\le p;(|\xi_1|,\xi_2)\text{ irreductible }\}, \tag{0.30}$$

$$\sum_{\xi\in\mathcal{S}(x,y)}\gamma_\xi^{\alpha_1,\alpha_2}(t,x,y)\xi\xi^{\mathsf{T}}=a_p^h(\alpha_1,\alpha_2,t,x,y), \tag{0.31}$$

et a_p^h est la projection de la matrice a^h sur $\mathcal{C}(\mathcal{S}_p)$, $a^h=a/h^2$.

Finalement, l'équation discrète qu'on va résoudre est la suivante :

$$S^\rho(t,x,y,v_h(t,x,y),v_h)=0, \tag{0.32a}$$

pour $(t,x,y)\in[0,T)\times(0,\infty)^2$, avec conditions au bord :

$$v_h(T,x,y)=g(x),\ \forall\ (x,y)\in[0,\infty)^2,, \tag{0.32b}$$

$$v_h(t,x,0)=g(x),\ \forall\ (t,x)\in[0,T]\times[0,\infty), \tag{0.32c}$$

$$v_h(t,0,y)=g(0),\ \forall\ (t,y)\in[0,T]\times[0,\infty). \tag{0.32d}$$

On démontre que ce schéma est monotone, stable et consistant, au sens des définitions suivantes :

(S1) **Monotonie :** $S^\rho(t,x,y,r,u) \geq S^\rho(t,x,y,r,v)$,
pour tout $r \in \mathbb{R}$, $x, y \in \mathbb{R}_+^*$, $u, v \in C([0,T]\times[0,\infty)^2)$ tels que $u \leq v$ dans $[0,T]\times[0,\infty)^2$.

(S2) **Stabilité :** Pour tout $\rho = (h, \Delta t, p_{\max}) \in (\mathbb{R}_+^*) \times (0,T) \times \mathbb{N}^*$, il existe une solution borné $v_h$de (0.32).

(S3) **Consistance :** Il existe une constante $C_1 > 0$, telle que, pour tout $\phi \in C^n([0,T] \times [0,\infty)^2)$, $n \geq 4$, avec dérivées bornées,

$$\begin{aligned} &\left| \min_{\alpha_1^2+\alpha_2^2} \left\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t,x,y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t,x,y) - \frac{1}{2}\mathrm{tr}[a \cdot D^2\phi(t,x,y)] \right\} \right. \\ &\left. -S^\rho(t,x,y,\phi(t,x,y),\phi) \right| \\ &\leq C_1(|\partial_t^2\phi|_0 \Delta t + \mu |D_y^2 \phi|_0 h) + 16\sqrt{2} p_{max}^2 \|a\| |D^4\phi|_0 h^2 + \varepsilon_p(t,x,y)|D^2\phi|_0, \end{aligned} \tag{0.33}$$

où a_p est la projection de a dans $\mathcal{C}(\mathcal{S}_p)$, pour $p = \min(p_{max}, \lceil x/h \rceil, \lceil y/h \rceil)$, et $\varepsilon_p(t,x,y)$ est l'erreur de projection tel que $\varepsilon_p = ||a - a_p||$ et $p = p_{max}$, et $\varepsilon_p = CK(x,y)h^2$ sinon, où C dépend de la constante de Lipschitz de σ^2, et $K(x,y) \geq 0$, pour tout x, y.

On introduit l'ensemble $\mathcal{A}$ de contrôles associés aux points $(x_i, y_j)_{i\geq 0,\, j\geq 1}$ par : $\bar{\mathcal{B}} := \{\alpha = (\alpha_1, \alpha_2), \alpha_1^2 + \alpha_2^2 = 1\}$ et

$$\mathcal{A} := (\bar{\mathcal{B}})^{\mathbb{N}\times\mathbb{N}^*}.$$

Alors le schéma peut s'écrire sous la forme abstraite suivante :
Trouver $X := v_h(t,.,.) \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}^*}$, borné, tel que

$$\min_{w\in\mathcal{A}} \Big(A(w)X - b(w) \Big) = 0, \tag{0.34}$$

où $A(w)$ est un operateur linéaire dans $\mathbb{R}^{\mathbb{N}\times\mathbb{N}^*}$, et $b(w)$ est un vecteur de $\mathbb{R}^{\mathbb{N}\times\mathbb{N}^*}$, qui sont précisés dans le Chapitre 4.

En appliquant l'algorithme de Howard et des résultats sur les systèmes linéaires en dimension infinie on arrive à montrer le théorème suivant :

Proposition .12. *Il existe une unique solution bornée X au problème*

$$\min_{w\in\mathcal{A}} (A(w)X - b(w)) = 0.$$

On conclut le chapitre en montrant que, sous les hypothèses (A1-5), on peut appliquer le théorème de convergence introduit par Barles et Souganidis [5], et démontrer donc que $v_h \to \vartheta$, pour $h \to 0$.

Bibliographie

[1] M. Akian, J.L. Menaldi, and A. Sulem. On an investement-consumption model with transaction costs. *SIAM J. Control Optim.*, 34 :329–364, 1996.

[2] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume 17 of *Mathématiques et Applications*. Springer, Paris, 1994.

[3] G. Barles and E.R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numerical Analysis*. To appear.

[4] G. Barles and E.R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN. Mathematical Modelling and Numerical Analysis*, 36 :33–54, 2002.

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

[6] J.F. Bonnans, S. Maroso, and H. Zidani. Error estimates for a stochastic impulse control problem. Technical Report 5606, INRIA, 2005. "http ://www.inria.fr/rrrt/rr-5606.html".
To appear in Applied Math. Optimization.

[7] J.F. Bonnans, S. Maroso, and H. Zidani. Error estimates for stochastic differential games : the adverse stopping case. *IMA, J. Numerical Analysis*, 26 :188–212, 2006.

[8] J.F. Bonnans, E. Ottenwaelter, and H. Zidani. Numerical schemes for the two dimensional second-order HJB equation. *ESAIM : M2AN*, 38 :723–735, 2004.

[9] J.F. Bonnans and P. Rouchon. *Commande et optimisation de systèmes dynamiques*. Les Editions de l'Ecole Polytechnique, Palaiseau, France, 2005.

[10] J.F. Bonnans and H. Zidani. Characterization of consistency of some numerical schemes for the stochastic HJB equation. In J.L. Menaldi, E. Rofman, and A. Sulem, editors, *Optimal Control and PDE - Innovations et Applications*. IOS Press, Amsterdam, 2000.

[11] B. Bruder. Superreplication of european options with a derivative asset under constrained finite variation strategies. preprint, 2005.

[12] F. Camilli and M. Falcone. Approximation of control problems involving ordinary and impulsive controls. *ESAIM Control Optim. Calc. Var.*, 4 :159–176, 1999.

[13] J.-Ph. Chancelier, B. Øksendal, and A. Sulem. Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. *Tr. Mat. Inst. Steklova*, 237(Stokhast. Finans. Mat.) :149–172, 2002.

[14] P. Cheridito, H. Soner, and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Annales de l'Institut Henri Poincaré : Analyse Non Linéaire*. To appear.

[15] J. Cochet-Terrasson and S. Gaubert. Policy iteration algorithm for shortest path problems. 2000.

[16] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. American Mathematical Society (New Series)*, 27 :1–67, 1992.

[17] M.G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Mathematics of Computation*, 43 :1–19, 1984.

[18] J. Cvitanic, H. Pham, and N. Touzi. Super-replication in stochastic volatility models with portfolio constraints. *Journal of Applied Probability*, 36 :523–545, 1999.

[19] I. Capuzzo Dolcetta. On a discrete approximation of the Hamilton-Jacobi equation of dynamic programming. *Applied Mathematics and Optimization*, 10 :367–377, 1983.

[20] R. L. V. Gonzalez and C. A. Sagastizabal. Un algorithme pour la résolution rapide d'équations discrètes de Hamilton-Jacobi-Bellman. *C. R. Acad. Sci. PAris*, 311, Série I :45–50, 1990.

[21] M. Hintermuller, K. Ito, and K. Kunish. The primal-dual active set strategy as a semismooth newton method. *SIAM J. Optim.*, 13 :865–888, 2002.

[22] K. Ishii. Viscosity Solutions of Non Linear Second Order Elliptic PDEs associated with Impulse Control Problems II. *Funkcialaj Ekvacioj*, 38 :297–328, 1995.

[23] N.V. Krylov. On the rate of convergence of finite difference approximation for Bellman's equation. *St. Petersburg Math. J.*, 9 :639–650, 1997.

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

[25] N.V. Krylov. The rate of convergence of finite-difference approximations for Bellman's equations with lipschitz coefficients. *Appl Math. Optim.*, 52(3) :365–399, 2005.

[26] H.J. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of mathematics*. Springer, New York, 2001. Second edition.

[27] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 2 : viscosity solutions and uniqueness. *Communications in partial differential equations*, 8 :1229–1276, 1983.

[28] J.-L. Menaldi. Some estimates for finite difference approximations. *SIAM J. Control Optim.*, 27 :579–607, 1989.

[29] H. Pham. On some recent aspects of stochastic control and their applications. *Probability Surveys*, 2 :1–549, 2005.

[30] M.L. Puterman and S.L. Brumelle. On the convergence of policy iteration in stationary dynamic programming. *Math. Oper. Res.*, 4 :60–69, 1979.

[31] M.S. Santos and J. Rust. Convergence properties of policy iteration. *SIAM J. Control Optim.*, 42 :2094–2115, 2004.

[32] M. Soner and N. Touzi. Super replication under gamma constraints. *SIAM J. on Control and Optimization*, 39 :73–96, 2000.

[33] M. Soner and N. Touzi. Stochastic target problems, dynamic programming and viscosity solutions. *SIAM J. on Control and Optimization*, 41 :404–424, 2002.

CHAPITRE I

Error estimate for stochastic differential games : the adverse stopping case

Error estimates for stochastic differential games : the adverse stopping case [1]

Abstract

We obtain error bounds for monotone approximation schemes of a particular Isaacs equation. This is an extension of the theory for estimating errors for the Hamilton-Jacobi-Bellman equation.

To obtain the upper error bound, we consider the "Krylov regularization" of the Isaacs equation to build an approximate sub-solution of the scheme. To get the lower error bound we extend the method of Barles and Jakobsen [1] which consists in introducing a switching system whose solutions are local super-solutions of the Isaacs equation.

keywords : Isaacs equation, Hamilton-Jacobi-Bellman equation, stochastic differential games, finite differences, error estimates.

1 Introduction

The aim of this paper is to give error bounds for approximation schemes of a particular non convex Isaacs equation. More precisely we consider the following equation

$$\min\{\sup_{\alpha\in\mathcal{A}} L^\alpha(x,\mathcal{D}u(x)); u(x)-\psi(x)\}=0, \qquad x\in\mathbb{R}^N, \tag{1.1}$$

where

$$\begin{aligned} L^\alpha(x,\mathcal{D}u(x)) &= L^\alpha(x,u(x),Du(x),D^2u(x)),\\ L^\alpha(x,t,p,X) &= -\operatorname{tr}[a^\alpha(x)X]-b^\alpha(x)p+c^\alpha(x)t-f^\alpha(x). \end{aligned}$$

Here $\mathcal{A}=\{\alpha_1,\dots,\alpha_M\}$ denotes the set of controls, assumed to be finite ; the case of a compact set will be dealt with in Section 4.3. The coefficients $(a^\alpha,b^\alpha,c^\alpha,f^\alpha)$ are, for each $\alpha\in\mathcal{A}$, bounded and Lipschitz functions $\mathbb{R}^N\to\mathcal{S}^N\times\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}$, where $\mathcal{S}^N$ denotes the set of $N\times N$ symmetric matrices ; ψ is a bounded Lipschitz function from $\mathbb{R}^N$ into $\mathbb{R}$. Under classical assumptions, (1.1) has a unique bounded viscosity solution, denoted u. The regularity of u depends on the properties of the coefficients a,b,c,f.
This problem is a particular case of stochastic differential games, called the adverse stopping case. In fact, we can note that in (1.1) we have two players, A and B. Player A has a set of controls and wants to minimize the gain. Player B can only decide to stop the game with the objective of maximizing the gain.
Then we consider monotone approximation schemes of (1.1), of the following form :

$$\min\{S(h,x,u_h(x),u_h);u_h(x)-\psi(x)\}=0, \qquad x\in\mathbb{R}^N, \tag{1.2}$$

[1] Joint work with F. Bonnans and H. Zidani.
IMA Journal of Numerical Analysis, 26 : 188-212,2006.

where S is a consistent, monotonic and uniformly continuous approximation of $\sup_{\alpha\in\mathcal{A}} L^\alpha$. We will note $u_h \in C_b(\mathbb{R}^N)$ the solution of (1.2), which is the approximation of u, and h the mesh size. This abstract notations was introduced by Barles and Souganidis [3] to display clearly the monotonicity of the scheme : $S(h,x,r,v)$ is non decreasing in r and non increasing in v. Typical approximation schemes that we will consider are Classical Finite Differences (Kushner and Dupuis [16]), Generalized Finite Differences [5, 6], and Markov Chain Approximations [16].

Until now, results on convergence rates for monotone approximation schemes of the equation with one player have been obtained ; i.e. for the following equation :

$$\sup_{\alpha\in\mathcal{A}} L^\alpha(x, \mathcal{D}u(x)) = 0, \qquad x \in \mathbb{R}^N, \tag{1.3}$$

and the scheme

$$S(h, x, u_h(x), u_h) = 0, \qquad x \in \mathbb{R}^N. \tag{1.4}$$

Error estimates for this equation have been obtained by Krylov [14, 15], and these results were extended by Barles and Jakobsen [2, 1]. Error estimates for a stochastic impulse control problem was recently obtained by the authors [4]. During the redactions of this paper we learned that also Jakobsen was working on the convergence rate for monotone approximations of (1.1). In [13] he obtained error estimates in the case of finite differences scheme with matrix a independent of x, using a penalization approach, and in [12] he obtained error estimates in the one dimensional case but for general Isaacs equations.

By using the methods introduced by Barles and Jakobsen [1], we give convergence rate for monotonic approximation schemes of the two players equation. The upper estimate of $u - u_h$ is easier to obtain than the lower. The proof involves a "Krylov regularization" of (1.1), i.e. the perturbed equation

$$\min\{\sup_{\alpha, |e|\le\varepsilon} L^\alpha(x+e, \mathcal{D}u^\varepsilon(x)); u^\varepsilon(x) - \psi(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{1.5}$$

and its viscosity solution u^ε. A regularization of u^ε by convolution gives an approximate smooth sub-solution of (1.1), denoted u_ε which is also an approximate sub-solution of (1.2). So, by using the consistency property, we obtain the upper bound, after choosing an optimal parameter of regularization. Unfortunately we can't proceed in a similar way to build a smooth super-solution of (1.1) which would lead to the lower estimate on $u - u_h$. Instead of a smooth super-solution we build a sequence of local smooth super-solution. In particular we introduce the following switching system which approximates (1.1)

$$\min\{\max\{L^{\alpha_i}(x, \mathcal{D}v_i(x)); v_i(x) - \min_{j\neq i}\{v_j(x) + k\}\}; v_i(x) - \psi(x)\} = 0, \tag{1.6}$$

for $x \in \mathbb{R}^N$, and $i \in \mathcal{I} = \{1, \ldots, M\}$. For literature on the switching systems, see Capuzzo-Dolcetta and Evans [7], Evans and Friedman [9], Ishii and Koike [10, 11]. We consider the viscosity solution $v = (v_1, \ldots, v_M)$ of this system, and give an estimate of the rate of convergence of v to u. Then we consider a perturbed system

$$\min\{\max\{\inf_{|e|\le\varepsilon} L^{\alpha_i}(x+e, \mathcal{D}v_i^\varepsilon(x)); v_i^\varepsilon(x) - \min_{j\neq i}\{v_j^\varepsilon(x)+k\}\}; v_i^\varepsilon(x) - \psi(x)\} = 0, \qquad (1.7)$$

for all $i \in \mathcal{I}$ and $x \in \mathbb{R}^N$, and its viscosity solution denoted $v^\varepsilon = (v_1^\varepsilon, \ldots, v_M^\varepsilon)$. We regularize v^ε by convolution obtaining v_ε, and this function allows to build a local super-solution of (1.1). Then by applying the consistency and the monotonicity of the scheme we obtain the desired bound. With our result, we can prove an upper bound of $h^{1/2}$ and a lower bound of $h^{1/5}$ for classical finite differences scheme and for generalized finite differences scheme.

The paper is organized as follows : in Section 2 we introduce the assumptions on equation (1.1) and scheme (1.2). In Section 3 we obtain the rate of convergence of the solution v of (1.6) to u. In the first part of Section 4 we obtain an upper bound of $u - u_h$, and in the second part of this section we use the rate obtained in Section 3 for giving the lower bound of $u-u_h$. In Section 5 we apply our results to the generalized finite difference scheme taken from [6], and studies conditions under which a general Markov chain approximation give better estimates than this scheme. Finally in the Appendix we give some auxiliary theorems which are used throughout the paper.

We conclude this introduction with some notations. In the sequel C is a positive constant independent on parameters ε and h. By $|\cdot|$ we mean the standard Euclidean norm in any $\mathbb{R}^M$ type space. In particular, if $X \in \mathcal{S}^N$, then $|X|^2 = \mathrm{tr}(XX^\top)$, where $X^\top$ is the transpose of X, i.e. $|X|$ is the Frobenius norm. If g is a bounded function from $\mathbb{R}^N$ into either $\mathbb{R}, \mathbb{R}^M$, or the space of $N \times P$ matrices, we set

$$|g|_0 := \sup_{x\in\mathbb{R}^N} |g(x)|.$$

If g is also Lipschitz continuous, we set

$$[g]_1 := \sup_{\substack{x,y\in\mathbb{R}^N\\ x\neq y}} \frac{|g(x)-g(y)|}{|x-y|}, \qquad |g|_1 := |g|_0 + [g]_1.$$

We denote by $\le$ the component wise ordering in $\mathbb{R}^N$, and by $\preceq$ the ordering in the sense of positive semidefinite matrices in $\mathcal{S}(N)$. The space $C_b(\mathbb{R}^N)$ (resp. $C_{b,l}(\mathbb{R}^N)$) will denote the space of continuous and bounded functions (resp. bounded and Lipschitz functions) from $\mathbb{R}^N$ to $\mathbb{R}$.

2 Well-posedness of the Isaacs equation and of the scheme

Throughout this paper, we suppose that the following assumptions are satisfied :

(A1) There exist λ, K such that, for all $x \in \mathbb{R}^N$ and $\alpha \in \mathcal{A}$, we have that $a^\alpha(x) = \frac{1}{2}\sigma^\alpha(x)(\sigma^\alpha(x))^T$, where $\sigma^\alpha(x)$ is, for each x, a $N \times P$ matrix, and

$$c^\alpha \ge \lambda > 0; \qquad |\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1 \le K,$$

(A2) $\lambda > \sup_\alpha\{[\sigma^\alpha]_1 + [b^\alpha]_1\}$.

Definition I.1. *The function $u \in C(\mathbb{R}^N)$ is called a viscosity sub-solution (resp. super-solution) of (1.1) if, for every $x \in \mathbb{R}^N$,*

$$\min\{\sup_\alpha L^\alpha(x, u(x), D\varphi(x), D^2\varphi(x)); u(x) - \psi(x)\} \le 0, \qquad (resp.\ \ge 0),$$

for each $\varphi \in C^2(\mathbb{R}^N)$ such that $u - \varphi$ has a local maximum (resp. a local minimum) at x. Finally we call u a viscosity solution of (1.1) if it is both a sub-solution and a super-solution.

We refer to [13, Lemma A.1] for the proof of the following result.

Proposition I.2. *Assume (A1) and (A2). Then the following statements hold :*

(i) If u_1 and u_2 are respectively viscosity sub-solution and viscosity super-solution of (1.1), $u_1, u_2 \in C_b(\mathbb{R}^N)$, then $u_1 \le u_2$ in $\mathbb{R}^N$.

(ii) There exists an unique viscosity solution u of (1.1), in the space $C_{b,l}(\mathbb{R}^N)$. □

Consider the scheme (1.2), with $h > 0$ and $S : \mathbb{R}^+ \times \mathbb{R}^N \times \mathbb{R} \times C_b(\mathbb{R}^N) \to \mathbb{R}$. We make the following assumptions :

(S1) Monotonicity : for all $h > 0$, $r \in \mathbb{R}^N$, $m \ge 0$, $x \in \mathbb{R}^N$ and bounded and continuous functions u, v such that $u \le v$ in $\mathbb{R}^N$,

$$S(h, x, r + m, u + m) \ge \lambda m + S(h, x, r, v).$$

(S2) Regularity : for all $h > 0$ and $\phi \in C_b(\mathbb{R}^N)$, $x \mapsto S(h, x, \phi(x), \phi)$ is bounded and continuous ; $r \mapsto S(h, x, r, \phi)$ is uniformly continuous for bounded r, uniformly with respect to $x \in \mathbb{R}^N$.

(S3) There exist $n, k_i > 0$, $i \in J \subseteq \{1, \ldots, n\}$ and a constant $K_c > 0$ such that for all $h > 0$ and x in $\mathbb{R}^N$, and for every smooth $\phi \in C^n(\mathbb{R}^N)$ such that $|D^i\phi|_0$ is bounded, for every $i \in J$, the following holds :

$$|\sup_\alpha L^\alpha(x, \mathcal{D}\phi) - S(h, x, \phi(x), \phi)| \le K_c Q(\phi),$$

where $Q(\phi) := \sum_{i \in J} |D^i\phi|_0 h^{k_i}$.

Remark I.3. *(S1) and (S2) imply that S is decreasing w.r.t. v (take $m = 0$), and increasing w.r.t. r (take $v = u + m$).*

In the following, we say that a function $v \in C_b(\mathbb{R}^N)$ is a sub-solution (resp. super-solution) to the scheme (1.2) if

$$\min\{S(h, x, v(x), v); v(x) - \psi(x)\} \le 0, \quad (\text{ resp. } \ge 0), \quad \text{for all } x \in \mathbb{R}^N.$$

Under assumptions (S1) and (S2), we have the existence of a comparison principle for bounded continuous solutions of (1.2) ; i.e.

Theorem I.4. *Let u_h (resp. v_h) be a bounded, continuous sub-solution (resp. super-solution) of (1.2). Then we have $u_h(x) \le v_h(x)$, for all $x \in \mathbb{R}^N$.*

Proof. The proof is an easy extension of [2, Lemma 2.3]. We assume that $m := \sup_x(u_h(x) - v_h(x)) > 0$ and obtain a contradiction. Let $\{x_n\}_n$ be a sequence in $\mathbb{R}^N$ such that $\delta_n := u_h(x_n) - v_h(x_n)$ converges to m as $n \to \infty$. Then $\delta_n > 0$ for large enough n. By using the notion of sub and super-solution, and the fact that $\min(A,B) - \min(C,D) \ge \min(A-C, B-D)$, we get

$$0 \ge \min\{S(h, x_n, u_h(x_n), u_h) - S(h, x_n, v_h(x_n), v_h); u_h(x_n) - v_h(x_n)\}.$$

Since $u_h(x_n) - v_h(x_n) = \delta_n > 0$, by using (S1), we have

$$\begin{aligned} 0 &\ge S(h, x_n, u_h(x_n), u_h) - S(h, x_n, v_h(x_n), v_h) \\ &\ge S(h, x_n, v_h(x_n) + \delta_n, v_h + m) - S(h, x_n, v_h(x_n), v_h) \\ &\ge S(h, x_n, v_h(x_n) + m, v_h + m) + \omega(m - \delta_n) - S(h, x_n, v_h(x_n), v_h) \\ &\ge \lambda m + \omega(m - \delta_n), \end{aligned}$$

where $\omega(t) \to 0$ when $t \to 0^+$ is given by (S2). Letting $n \to \infty$ yields $m \le 0$ which is a contradiction. □

In all the sequel we will use a sequence of mollifiers $(\rho_\varepsilon)_\varepsilon$ defined as follows :

$$\rho_\varepsilon(x) = \varepsilon^{-N}\rho(x/\varepsilon), \tag{2.1}$$

where $\rho \in C^\infty(\mathbb{R}^N)$, $\int_{\mathbb{R}^N} \rho = 1$, $\operatorname{supp}\{\rho\} \subseteq \bar{B}(0,1)$ and $\rho \ge 0$. If g is a continuous function of $\mathbb{R}^N$ to $\mathbb{R}$, then we define the mollification of g as follows :

$$g_\varepsilon(x) := \int_{\mathbb{R}^N} g(x-e)\rho_\varepsilon(e)de. \tag{2.2}$$

Moreover, if g is a Lipschitz function, then

$$|g(x) - g_\varepsilon(x)| \le \varepsilon[g]_1. \tag{2.3}$$

If $g \in C_b^n(\mathbb{R}^N)$ (resp. $C_{b,l}^n(\mathbb{R}^N)$), then

$$|D^i g_\varepsilon(x)| \le C\varepsilon^{-i}|g|_0, \qquad (\text{resp. } C\varepsilon^{1-i}|g|_0), \quad \forall\, i = 1, \dots, n. \tag{2.4}$$

3 Switching system

Consider the following switching system approximation of (1.1) :

$$\min\left\{ \max\left(L^{\alpha_i}(x, \mathcal{D}v_i(x)); v_i(x) - \min_{j\ne i}\{v_j(x) + k\}\right); v_i(x) - \psi(x)\right\} = 0, \tag{3.1}$$

for $i \in \mathcal{I} = \{1, \dots, M\}$ and $x \in \mathbb{R}^N$. In particular we have an equation for every control. A viscosity solution theory for the switching system can be found in [10], [11], [17]. We recall here the definition of viscosity solution for a general switching system of the form :

$$F_i(x, v, Dv_i, D^2 v_i) = 0, \qquad i = 1, \dots, M, \tag{3.2}$$

where $F_i : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N \times \mathcal{S}^N \to \mathbb{R}$.

Definition I.5. *The function $v = (v_1, \dots, v_M) \in C(\mathbb{R}^N)^M$ is called a viscosity sub-solution (resp. super-solution) of (3.2) if, for every $i \in \mathcal{I}$ and $x \in \mathbb{R}^N$,*

$$F_i(x, v(x), D\varphi(x), D^2\varphi(x)) \le 0, \qquad (resp. \ge 0),$$

for each $\varphi \in C^2(\mathbb{R}^N)$ such that $v_i - \varphi$ has a local maximum (resp. a local minimum) at x. Finally we call v a viscosity solution of (3.2) if it is both a sub-solution and a super-solution.

Lemma I.20 implies a comparison principle for (3.1), and then the existence of a unique viscosity solution of (3.1) in $C_{b,l}(\mathbb{R}^N)^M$, denoted $v = (v_1, \dots, v_M)$.

We perturb the system (3.1) and build the following auxiliary system

$$\min\{\max(\sup_{|e| \le \varepsilon} L^{\alpha_i}(x + e, \mathcal{D}v_i^\varepsilon(x)); v_i^\varepsilon(x) - \min_{j \ne i}\{v_j^\varepsilon(x) + k\}); v_i^\varepsilon(x) - \psi(x)\} = 0. \tag{3.3}$$

Lemma I.20 applied to (3.3), implies the existence of a unique viscosity solution of (3.3), denoted $v^\varepsilon = (v_1^\varepsilon, \dots, v_M^\varepsilon)$, in $C_{b,l}(\mathbb{R}^N)^M$. The following lemma is consequence of Theorem I.21.

Lemma I.6. *Under assumptions (A1) and (A2), we have that*

$$|v_i - v_i^\varepsilon| \le C\varepsilon, \tag{3.4}$$

where C only depends on K, λ and $[\psi]_1$. □

For every i, let $v_{i\varepsilon}$ the mollification of v_i^ε, defined as in (2.2). Since v_i^ε is a Lipschitz function, uniformly w.r.t $\varepsilon > 0$ sufficiently small, (2.3) implies

$$|v_i^\varepsilon(x) - v_{i\varepsilon}(x)| \le \max_i [v_i^\varepsilon]_1 \varepsilon; \tag{3.5}$$

Lemma I.20 implies that $\max_i [v_i^\varepsilon]_1$ remains bounded when $\varepsilon \downarrow 0$ (this argument will be used several times in the paper).

Lemma I.7. *The function $v_{i\varepsilon} - R$ is, for all i, sub-solution of equation (1.1), for some*

$$R := C\left(k + \varepsilon + \frac{k}{\varepsilon^2}\right), \tag{3.6}$$

where the constant C depends only on K, λ and $[\psi]_1$.

Proof. Let R satisfy (3.6). We have to prove that, for large enough C,

$$\min\{\sup_{\alpha} L^{\alpha}(x, \mathcal{D}(v_{i\varepsilon}(x) - R); v_{i\varepsilon}(x) - R - \psi(x)\} \leq 0, \quad \forall\, x \in \mathbb{R}^N, \tag{3.7}$$

for all $i \in \mathcal{I}$. Since $v_{i\varepsilon} \in C^{\infty}(\mathbb{R}^N)$, the definition of viscosity sub-solution is equivalent to the notion of classical sub-solution. Therefore we have to prove that one of the following statements holds for all $x \in \mathbb{R}^N$:

$$v_{i\varepsilon}(x) - R \leq \psi(x), \quad \forall\, i \in \mathcal{I}, \tag{3.8a}$$

$$\sup_{\alpha} L^{\alpha}(x, \mathcal{D}(v_{i\varepsilon}(x) - R)) \leq 0, \quad \forall\, i \in \mathcal{I}. \tag{3.8b}$$

For every $x \in \mathbb{R}^N$, set

$$I^{\varepsilon}(x) := \{i \in \mathcal{I} \mid v_i^{\varepsilon}(x) = \psi(x)\}. \tag{3.9}$$

Let $\tilde{x} \in \mathbb{R}^N$. Denote by $B(\tilde{x}, 2\varepsilon)$ the ball centered on $\tilde{x}$ with radius 2ε. Then we have the two following possibilities :

CASE A : There exists $y \in B(\tilde{x}, 2\varepsilon)$ such that $I^{\varepsilon}(y) \neq \varnothing$. We claim that (3.8a) holds. We have $v_{i_0}^{\varepsilon}(y) = \psi(y)$, for some $i_0 \in I^{\varepsilon}(y)$. Let $i \notin I^{\varepsilon}(y)$. The function $v_i^{\varepsilon}(x) - |x-y|^2/\varepsilon_1$ has, for sufficiently small $\varepsilon_1 > 0$, a local maximum at a point x_ε such that $|x_\varepsilon - y| \leq \varepsilon$. Since v^{ε} is the viscosity solution of (3.3), one of the following statements holds :

$$v_i^{\varepsilon}(x_\varepsilon) \leq \psi(x_\varepsilon), \tag{3.10a}$$

$$\max\left\{\sup_{|e|\leq\varepsilon} L^{\alpha_i}(x_\varepsilon + e, v_i^{\varepsilon}(x_\varepsilon), \frac{2}{\varepsilon_1}(x_\varepsilon - y), \frac{2I}{\varepsilon_1}); \; v_i^{\varepsilon}(x_\varepsilon) - \min_{j\neq i}\{v_j^{\varepsilon}(x_\varepsilon) + k\}\right\} \leq 0. \tag{3.10b}$$

If $v_i^{\varepsilon}(x_\varepsilon) \leq \psi(x_\varepsilon)$, since v_i^{ε} and ψ are Lipschitz, we obtain

$$v_i^{\varepsilon}(y) \leq \psi(y) + ([\psi]_1 + \max_i [v_i^{\varepsilon}]_1)\varepsilon. \tag{3.11}$$

Otherwise, with (3.10b), we have

$$v_i^{\varepsilon}(y) \leq \max_i [v_i^{\varepsilon}]_1 \varepsilon + v_{i_0}^{\varepsilon}(x_\varepsilon) + k \leq 2\max_i [v_i^{\varepsilon}]_1 \varepsilon + v_{i_0}^{\varepsilon}(y) + k. \tag{3.12}$$

Since either (3.11) or (3.12) holds, we deduce that

$$v_i^{\varepsilon}(y) - C\varepsilon - k \leq \psi(y), \qquad \forall i \in \mathcal{I}, \tag{3.13}$$

where C depends on $[\psi]_1$ and $\max_i [v_i^{\varepsilon}]_1$. Since $y \in B(\tilde{x}, 2\varepsilon)$, and ψ and v_i^{ε} are Lipschitz, this implies $v_i^{\varepsilon}(\tilde{x}) \leq \psi(\tilde{x}) + k + C\varepsilon$, for all $i \in \mathcal{I}$. Applying (3.5), we obtain $v_{i\varepsilon}(\tilde{x}) \leq \psi(\tilde{x}) + R$, for all $i \in \mathcal{I}$, which implies (3.7).

CASE B : For all $y \in B(\tilde{x}, 2\varepsilon)$, we have that $I^\varepsilon(y) = \varnothing$. We claim that, for all $e \in B(0,\varepsilon)$, $(v_1^\varepsilon(\cdot - e), \dots, v_M^\varepsilon(\cdot - e))$ is a viscosity sub-solution of the following system

$$\max\left\{L^{\alpha_i}(x, \mathcal{D}w_i(x)); w_i(x) - \min_{j\neq i}\{w_j(x) + k\}\right\} = 0, \quad x \in B(\tilde{x}, \varepsilon). \tag{3.14}$$

Fix $e_1 \in B(0,\varepsilon)$ and $i \in \mathcal{I}$. Let $\varphi \in C^2(\mathbb{R}^N)$ be such that $v_i^\varepsilon(\cdot - e_1) - \varphi(\cdot)$ has a local maximum x_{e_1} in the ball $B(\tilde{x}, \varepsilon)$. Then $v_i^\varepsilon(\cdot) - \varphi(\cdot + e_1)$ has a local maximum at $x_{e_1} - e_1$. Since $x_{e_1} - e_1 \in B(\tilde{x}, 2\varepsilon)$, we have that $v_i^\varepsilon(x_{e_1} - e_1) > \psi(x_{e_1} - e_1)$, and since v^ε is the viscosity solution of (3.3), we obtain

$$\max\left\{\sup_{|e|\leq\varepsilon} L^{\alpha_i}(x_{e_1} - e_1 + e, v_i^\varepsilon(x_{e_1} - e_1), D\varphi(x_{e_1}), D^2\varphi(x_{e_1}));\right.$$
$$\left. v_i^\varepsilon(x_{e_1} - e_1) - \min_{j\neq i}\{v_j^\varepsilon(x_{e_1} - e_1) + k\}\right\} \leq 0.$$

Taking $e = e_1$, we obtain

$$\begin{cases} L^{\alpha_i}(x_{e_1}, v_i^\varepsilon(x_{e_1} - e_1), D\varphi(x_{e_1}), D^2\varphi(x_{e_1})) \leq 0, \\ v_i^\varepsilon(x_{e_1} - e_1) - \min_{j\neq i}\{v_j^\varepsilon(x_{e_1} - e_1) + k\} \leq 0. \end{cases}$$

This being true for an arbitrary $e_1 \in B(0,\varepsilon)$ and $i \in \mathcal{I}$, we obtain that, for all $|e| \leq \varepsilon$, $(v_1^\varepsilon(\cdot - e), \dots, v_M^\varepsilon(\cdot - e))$ is a viscosity sub-solution of (3.14). Applying [2, Lemma A.3 and Lemma 2.7], since $v_{i\varepsilon}(\cdot)$ is limit of convex combination of $v_i^\varepsilon(\cdot - e)$, for $e \in B(0,\varepsilon)$, then $(v_{1\varepsilon}(\cdot), \dots, v_{M\varepsilon}(\cdot))$ is a viscosity sub-solution of (3.14). Moreover, since it is a smooth function, it is a sub-solution of (3.14) in the classical sense, and we have

$$L^{\alpha_i}(\tilde{x}, \mathcal{D}v_{i\varepsilon}(\tilde{x})) \leq 0, \quad \forall\, i \in \mathcal{I}. \tag{3.15}$$

We know that $|v_i^\varepsilon(y) - v_j^\varepsilon(y)| \leq k$ for all $i, j \in \mathcal{I}$ and $y \in B(\tilde{x}, \varepsilon)$. Consequently

$$D^n v_{i\varepsilon}(\tilde{x}) - D^n v_{j\varepsilon}(\tilde{x}) \leq \frac{Ck}{\varepsilon^n}, \quad \forall n \geq 1,$$

where C depends only on ρ defined in (2.1). It follows that

$$L^{\alpha_i}(\tilde{x}, \mathcal{D}v_{j\varepsilon}(\tilde{x})) - L^{\alpha_i}(\tilde{x}, \mathcal{D}v_{i\varepsilon}(\tilde{x})) \leq \frac{Ck}{\varepsilon^2}, \quad \forall\, i, j \in \mathcal{I}.$$

Combining with (3.15), we get $L^{\alpha_i}(\tilde{x}, \mathcal{D}v_{j\varepsilon}(\tilde{x})) \leq Ck/\varepsilon^2$, for all i and j in $\mathcal{I}$, for some C depending on ρ and K, and hence, $\sup_\alpha L^\alpha(\tilde{x}, \mathcal{D}v_{i\varepsilon}(\tilde{x})) \leq Ck/\varepsilon^2$ for all i in $\mathcal{I}$. Using assumption (A1), we have that for all i in $\mathcal{I}$, $\sup_\alpha L^\alpha(\tilde{x}, \mathcal{D}v_{i\varepsilon}(\tilde{x}) - Ck/(\lambda\varepsilon^2)) \leq 0$. Therefore (3.7) also holds in this case.□

Theorem I.8. *For every $i \in \mathcal{I}$ and for all $x \in \mathbb{R}^N$ we have*

$$0 \leq v_i - u \leq Ck^{1/3}, \tag{3.16}$$

where C depends only on λ and K from (A1).

Proof. a) We prove the first inequality of (3.16). Let $w = (u, \dots, u)$ be the vector whose M components are equal to u. We claim that w is a viscosity sub-solution of (3.1). Let $\varphi \in C^2(\mathbb{R}^N)$ be such that $u - \varphi$ has a local maximum at $x_0 \in \mathbb{R}^N$. Since u is a viscosity sub-solution of (1.1), either $u(x_0) \leq \psi(x_0)$, or $\sup_\alpha L^\alpha(x_0, u(x_0), D\varphi(x_0), D^2\varphi(x_0)) \leq 0$. If the latter holds, then

$$L^{\alpha_i}(x_0, u(x_0), D\varphi(x_0), D^2\varphi(x_0)) \leq 0, \quad \forall\, i \in \mathcal{I}.$$

Combining both cases, we obtain

$$\min\left\{ \max\left(L^{\alpha_i}(x_0, u(x_0), D\varphi(x_0), D^2\varphi(x_0)); u(x_0) - \min_{j \neq i}\{u(x_0) + k\} \right); \right.$$
$$\left. u(x_0) - \psi(x_0) \right\} \leq 0, \quad \forall\, i \in \mathcal{I}.$$

Therefore w is a viscosity sub-solution of (3.1). By the comparison principle (Lemma I.20), the first inequality of (3.16) holds.
b) We now prove the second inequality in (3.16). By Lemma I.7 and Proposition I.2, we have that $v_{i\varepsilon} - R \leq u$, for all $i \in \mathcal{I}$, and $x \in \mathbb{R}^N$, where R satisfies (3.6). Applying (3.5) and (3.4), we obtain

$$v_i - u \leq |v_i - v_i^\varepsilon| + |v_i^\varepsilon - v_{i\varepsilon}| + |v_{i\varepsilon} - u| \leq C\left(\frac{k}{\varepsilon^2} + \varepsilon + k\right), \quad \forall\, i \in \mathcal{I}, \quad \forall\, x \in \mathbb{R}^N,$$

where C depends on K, λ, $\max_i [v_i^\varepsilon]_1, [\psi]_1$. Minimizing with respect to ε, we obtain the desired upper bound, with $\varepsilon = k^{1/3}$. □

Remark I.9. *We obtain the same estimate that in the case of only one player, (see [1, Theorem 2.3]), by extending the same methods.*

4 Bounds on $u - u_h$

We state in this section our main results : the upper and lower bounds on $u - u_h$.

4.1 Upper bound on $u - u_h$

Perturb the equation (1.1) so as to obtain

$$\min\left\{\sup_{\alpha,|e|\le\varepsilon} L^\alpha(x+e,\mathcal{D}u^\varepsilon(x)); u^\varepsilon(x)-\psi(x)\right\}=0, \quad x\in\mathbb{R}^N. \tag{4.1}$$

Under assumptions (A1) and(A2), by Proposition I.2, (4.1) has a unique viscosity solution $u^\varepsilon \in C_{b,l}(\mathbb{R}^N)$. In view of Theorem I.22, we have that $|u-u^\varepsilon|\le C\varepsilon$, for some C depending on λ, K and $[\psi]_1$. Define the contact domain of u^ε as

$$X(u^\varepsilon) := \{x\in\mathbb{R}^N | u^\varepsilon(x)=\psi(x)\}.$$

Let u_ε be the mollification of u^ε, defined as in (2.2). Since u^ε is a Lipschitz function, uniformly w.r.t $\varepsilon>0$ sufficiently small, (2.3) implies

$$|u^\varepsilon(x)-u_\varepsilon(x)|\le [u^\varepsilon]_1\varepsilon, \tag{4.2}$$

where $[u^\varepsilon]_1$ remains bounded.

Theorem I.10. *Assume that (A1), (A2), (S1)-(S3) hold and let the approximation scheme (1.2) have a unique solution u_h in $C_b(\mathbb{R}^N)$. Then, for sufficiently small $h>0$, we have*

$$u-u_h\le Ch^\ell, \qquad \forall\, x\in\mathbb{R}^N, \tag{4.3}$$

where $\ell := \min_{i\in J}\{k_i/i\}$, C depends only on λ, K, $[\psi]_1$ and K_c, the constants k_i and K_C being defined in (S3).

Proof. We claim that

$$\min\{S(h,x,u_\varepsilon(x)-R_1,u_\varepsilon-R_1); u_\varepsilon(x)-R_1-\psi(x)\}\le 0, \quad \forall\, x\in\mathbb{R}^N, \tag{4.4}$$

for some $R_1>0$ of the form $R_1 := \lambda^{-1}Q(u_\varepsilon)+C\varepsilon$, where $Q(\cdot)$ was defined in (S3) and C depends only on $[\psi]_1$ and $[u^\varepsilon]_1$. Indeed, we will prove a slightly stronger result : for any $x\in\mathbb{R}^N$, one at least of the following two statements holds :

$$u_\varepsilon(x)-C\varepsilon\le\psi(x), \tag{4.5a}$$
$$S(h,x,u_\varepsilon(x)-K_C\lambda^{-1}Q(u_\varepsilon),u_\varepsilon-K_C\lambda^{-1}Q(u_\varepsilon))\le 0. \tag{4.5b}$$

Fix an $\tilde{x}\in\mathbb{R}^N$. We have the following alternative :

CASE A : There exists $y\in B(\tilde{x},2\varepsilon)$ such that $y\in X(u^\varepsilon)$, i.e. $u^\varepsilon(y)=\psi(y)$. Since u^ε and ψ are uniformly Lipschitz, for some C depending only on $[\psi]_1$ and $[u^\varepsilon]_1$, we obtain (4.5a) at point $x=\tilde{x}$.

CASE B : One has $X(u^\varepsilon) \cap B(\tilde{x}, 2\varepsilon) = \varnothing$. We claim that $u^\varepsilon(\cdot - e)$ is, for all $e \in B(0, \varepsilon)$, a viscosity sub-solution of

$$\sup_\alpha L^\alpha(x, \mathcal{D}w(x)) = 0, \quad x \in B(\tilde{x}, \varepsilon). \tag{4.6}$$

Fix $e_1 \in B(\tilde{x}, \varepsilon)$, and let $\varphi \in C^2(\mathbb{R}^N)$ be such that $u^\varepsilon(\cdot - e_1) - \varphi(\cdot)$ has a local maximum at a point $x_{e_1} \in B(\tilde{x}, \varepsilon)$. Then $u^\varepsilon(\cdot) - \varphi(\cdot + e_1)$ has a local maximum at $x_{e_1} - e_1$. Since $x_{e_1} - e_1 \in B(\tilde{x}, 2\varepsilon)$, and hence, $x_{e_1} - e_1 \notin X(u^\varepsilon)$, we have, whenever $|e| \leq \varepsilon$,

$$\sup_{\alpha, |e| \leq \varepsilon} L^\alpha(x_{e_1} - e_1 + e, u^\varepsilon(x_{e_1} - e_1), D\varphi(x_{e_1}), D^2\varphi(x_{e_1})) \leq 0.$$

Taking $e = e_1$, we have

$$\sup_\alpha L^\alpha(x_{e_1}, u^\varepsilon(x_{e_1} - e_1), D\varphi(x_{e_1}), D^2\varphi(x_{e_1})) \leq 0.$$

This proves our claim that $u^\varepsilon(\cdot - e_1)$ is a viscosity sub-solution of (4.6). Since e_1 is an arbitrary point of $B(0, \varepsilon)$, $u^\varepsilon(\cdot - e)$ is a viscosity sub-solution of (4.6), for all $|e| \leq \varepsilon$. Since $u_\varepsilon(\cdot)$ is a C^∞ function, and it is limit of convex combination of $u^\varepsilon(\cdot - e)$ (see [2, Lemma A.3 and Lemma 2.7]), hence, applying ([2, Lemma 2.7]), we can say that $u_\varepsilon(\cdot)$ is a sub-solution of (4.6) in the classical sense. This implies

$$\sup_\alpha L^\alpha(\tilde{x}, \mathcal{D}u_\varepsilon(\tilde{x})) \leq 0. \tag{4.7}$$

By consistency, $S(h, \tilde{x}, u_\varepsilon(\tilde{x}), u_\varepsilon) \leq K_C Q(u_\varepsilon)$. Applying (S1) with $u = v = u^\varepsilon(\tilde{x}) - K_C \lambda^{-1} Q(u^\varepsilon)$ and $m = K_C \lambda^{-1} Q(u^\varepsilon)$, we obtain (4.5b) at point $x = \tilde{x}$.
Combining cases A and B, we obtain (4.4). So $u_\varepsilon - R_1$ is a sub-solution of (1.2). By Theorem I.4, $u_\varepsilon - R_1 \leq u_h$, i.e., $u - u_h \leq K_C \lambda^{-1} Q(u_\varepsilon) + C\varepsilon$, where C depends on $[u^\varepsilon]_1$ and $[\psi]_1$. Using $Q(u_\varepsilon) = \sum_{i \in J} |D^i u_\varepsilon| h^{k_i}$, and (2.4), we obtain $Q(u_\varepsilon) \leq C \sum_{i \in J} \varepsilon^{1-i} h^{k_i}$. The result follows by setting $\varepsilon = \max_{i \in J} h^{k_i/i}$. □

4.2 Lower bound on $u - u_h$

We perturb the switching system (3.1) as follows

$$\min\{\max(\inf_{|e| \leq \varepsilon} L^{\alpha_i}(x + e, \mathcal{D}v_i^\varepsilon(x)); v_i^\varepsilon(x) - \min_{j \neq i}\{v_j^\varepsilon(x) + k\}); v_i^\varepsilon(x) - \psi(x)\} = 0, \tag{4.8}$$

for all $i \in \mathcal{I}$ and $x \in \mathbb{R}^N$. By Lemma I.20, this system has a unique viscosity solution $v^\varepsilon = (v_1^\varepsilon, \ldots, v_M^\varepsilon)$ in $C_{b,l}(\mathbb{R}^N)^M$. Consider v_ε, the mollification of v^ε, defined as in (2.2). Since v^ε is a Lipschitz function, uniformly w.r.t. $\varepsilon > 0$ sufficiently small, applying Theorem I.21 and (2.3) we have

$$|v_i - v_i^\varepsilon| \leq C\varepsilon, \quad |v_i^\varepsilon - v_{i\varepsilon}| \leq \max_i [v_i^\varepsilon]_1 \varepsilon, \tag{4.9}$$

where C depends on λ, K and $[\psi]_1$, and $\max_i [v_i^\varepsilon]_1$ remains bounded.

Lemma I.11. *Let $x_0 \in \mathbb{R}^N$, $i_0 \in \arg\min_{j\in\mathcal{I}} v_{j\varepsilon}(x_0)$, and assume that*

$$\varepsilon \leq (12\sup_i [v_i^\varepsilon]_1)^{-1} k. \tag{4.10}$$

Then the following statements hold

$$v_{i_0}^\varepsilon(y) < v_j^\varepsilon(y) + k, \quad \textit{for all } j \in \mathcal{I}, \textit{ and } y \in B(x_0, 2\varepsilon), \tag{4.11}$$

$$\sup_\alpha L^\alpha(x_0, \mathcal{D}v_{i_0\varepsilon}(x_0)) \geq 0. \tag{4.12}$$

Proof. We follow the method of ([1, Lemma 3.4]).
a) Let us prove (4.11). Since $i_0 \in \arg\min_{j\in\mathcal{I}} v_{j\varepsilon}(x_0)$,

$$v_{i_0\varepsilon}(x_0) - \min_{j\neq i_0}\{v_{j\varepsilon}(x_0) + k\} = \max_{j\neq i_0}\{v_{i_0\varepsilon}(x_0) - v_{j\varepsilon}(x_0) - k\} \leq -k. \tag{4.13}$$

Since for every i, v_i^ε is Lipschitz, we apply (2.3) and we have that

$$v_{i_0}^\varepsilon(x_0) - \min_{j\neq i_0}\{v_j^\varepsilon(x_0) + k\} \leq -k + 2\varepsilon\max_i [v_i^\varepsilon]_1,$$

and, for all $y \in B(x_0, 2\varepsilon)$,

$$v_{i_0}^\varepsilon(y) - \min_{j\neq i_0}\{v_j^\varepsilon(y) + k\} \leq -k + 4\max_i [v_i^\varepsilon]_1(\varepsilon + |x_0 - y|) < -k + 12\varepsilon\max_i [v_i^\varepsilon]_1.$$

Taking $\varepsilon \leq (12\max_i [v_i^\varepsilon]_1)^{-1}k$, we obtain (4.11).
b) We prove (4.12). We claim that $v_{i_0}^\varepsilon(\cdot - e)$ is, for all $|e| \leq \varepsilon$, a viscosity super-solution of

$$L^{\alpha_{i_0}}(x, Dw(x)) = 0, \quad x \in B(x_0, \varepsilon). \tag{4.14}$$

Fix $e_1 \in B(0, \varepsilon)$, and let $\varphi \in C^2(\mathbb{R}^N)$ be such that $v_{i_0}^\varepsilon(\cdot - e_1) - \varphi(\cdot)$ has a local minimum at $x_{e_1} \in B(x_0, \varepsilon)$. Then $v_{i_0}^\varepsilon(\cdot) - \varphi(\cdot + e_1)$ has a local minimum at $x_{e_1} - e_1 \in B(x_0, 2\varepsilon)$. Since $v^\varepsilon(\cdot)$ is a viscosity solution of (4.8), and $v_{i_0}^\varepsilon(x_0) - \min_{j\neq i_0}\{v_j^\varepsilon(x_0) + k\} \leq 0$ by (4.11), we have that

$$\inf_{|e|\leq\varepsilon} L^{\alpha_{i_0}}(x_{e_1} - e_1 + e, v_{i_0}^\varepsilon(x_{e_1} - e_1), D\varphi(x_{e_1}), D^2\varphi(x_{e_1})) \geq 0, \quad \forall\, |e| \leq \varepsilon.$$

In particular, for $e = e_1$, we obtain that $v_{i_0}^\varepsilon(\cdot - e_1)$ is a viscosity super-solution of (4.14). Since e_1 is an arbitrary point in $B(0, \varepsilon)$, we obtain that $v_{i_0}^\varepsilon(\cdot - e)$ is, for all $e \in B(0, \varepsilon)$, a viscosity super-solution of (4.14). Being a limit of convex combinations of $v_{i_0}^\varepsilon(\cdot - e)$, and a smooth function, $v_{i_0\varepsilon}(\cdot)$ is a classical super-solution on (4.14), and hence $L^{\alpha_{i_0}}(x_0, \mathcal{D}v_{i_0\varepsilon}(x_0)) \geq 0$; relation (4.12) follows.□

Define the two following sets :

$$X := \{x \in \mathbb{R}^N | u_h(x) = \psi(x)\}; \quad Y := \{x \in \mathbb{R}^N | S(h, x, u_h, [u_h]_x) = 0\}.$$

Proposition I.12. *Under assumptions (A1), (A2) and (S1)-(S3), and assuming that (1.2) has a unique solution u_h in $C_b(\mathbb{R}^N)$, we have that, if $x \in Y$, the following holds :*

$$u_h(x) - u(x) \leq Ch^{\ell}, \tag{4.15}$$

where $\ell := \min_{i\in J}\{k_i/(3i-2)\}$ and C depends only on λ, K and K_c.

Proof. Consider the switching system (4.8), its solution $v^{\varepsilon} = (v_1^{\varepsilon}, \ldots, v_M^{\varepsilon})$ and mollification $v_{\varepsilon} = (v_{1\varepsilon}, \ldots, v_{M\varepsilon})$. Let $w(y) := \min_i v_{i\varepsilon}(y)$. Define

$$m := \sup_{y\in Y}\{u_h(y) - w(y)\} = \sup_{i\in\mathcal{I}, y\in Y}\{u_h(y) - v_{i\varepsilon}(y)\}. \tag{4.16}$$

Let $\phi(y) := (1 + |y|^2)^{1/2}$. An approximation of m is, for $k > 0$, given by

$$m_k := \sup_{y\in Y}\{u_h(y) - w(y) - k\phi(y)\}. \tag{4.17}$$

Since u_h and w are bounded, ϕ is coercive and Y is a closed set, the supremum in (4.17) is attained at some point $x_0 \in Y$. By the definition of w, we also have

$$x_0 \in \arg\max_{y\in Y}\{u_h(y) - v_{i_0\varepsilon}(y) - k\phi(y)\}, \tag{4.18}$$

when $i_0 \in \arg\min_{j\in\mathcal{I}} v_{j\varepsilon}(x_0)$. In particular,

$$m_k \geq u_h(y) - v_{i_0\varepsilon}(y) - k\phi(y), \qquad \text{for all } \ y \in Y. \tag{4.19}$$

Let ε be such that (4.10) holds. Applying Lemma I.11, and since the first and the second order derivatives of ϕ are bounded, we have $\sup_\alpha \ L^{\alpha}(x_0, \mathcal{D}(v_{i_0\varepsilon} + k\phi)(x_0)) \geq -Ck$. Combining with (S1), (S3), (4.17) and $x_0 \in Y$, we get

$$\begin{aligned} -Ck - K_C Q(v_{i_0\varepsilon} + k\phi) &\leq S(h, x_0, (v_{i_0\varepsilon} + k\phi)(x_0), v_{i_0\varepsilon} + k\phi) \\ &\leq S(h, x_0, u_h(x_0) - m_k, u_h - m_k) \\ &\leq -\lambda m_k + S(h, x_0, u_h(x_0), u_h) = -\lambda m_k. \end{aligned}$$

We obtain $\lambda m_k \leq K_C Q(v_{i_0\varepsilon} + k\phi) + Ck$. Passing to the limit in k, we get

$$m \leq K_C Q(v_{i_0\varepsilon}). \tag{4.20}$$

In conclusion, we can say that for $x \in Y$ and for every $i \in \mathcal{I}$,

$$\begin{aligned} \sup_{y\in Y}\{u_h(y) - u(y)\} &\leq m + \sup_{y\in Y}\{w(y) - u(y)\} \\ &\leq m + \sup_{y\in Y}\{w(y) - v_{i\varepsilon}(y)\} + \sup_{y\in Y}\{v_{i\varepsilon}(y) - v_i^{\varepsilon}(y)\} \\ &\quad + \sup_{y\in Y}\{v_i^{\varepsilon}(y) - v_i(y)\} + \sup_{y\in Y}\{v_i(y) - u(y)\}. \end{aligned} \tag{4.21}$$

Applying (4.9), (3.16), and the fact that $w(y) \le v_{i\varepsilon}(y)$ for all $i \in \mathcal{I}$, we obtain

$$\sup_{y\in Y}\{u_h(y) - u(y)\} \le m + C\varepsilon + Ck^{1/3}. \tag{4.22}$$

where C depends on $K, \lambda, [\psi]_1$ and $\max_i [v_i^\varepsilon]_1$. Using (4.20), we obtain

$$u - u_h \le K_C Q(v_{i_0\varepsilon}) + C\varepsilon + Ck^{1/3}, \quad \forall\, x \in Y.$$

The result follows by setting $\varepsilon = \max_{i\in J} h^{\frac{3k_i}{3i-2}}$ and $k = (12 \sup_i [v_i^\epsilon]_1)\epsilon$. □

Theorem I.13. *Under assumptions (A1), (A2) and (S1)-(S3), and assuming that (1.2) has a unique solution u_h in $C_{b,l}(\mathbb{R}^N)$, we have that*

$$u_h - u \le Ch^\ell, \qquad \forall\, x \in \mathbb{R}^N, \tag{4.23}$$

where $\ell = \min_{i\in J}\{k_i/(3i-2)\}$ and C depends only on λ, K and K_c.

Proof. If $x \in X$ we have that $u_h(x) = \psi(x) \le u(x)$, therefore (4.23) holds. If $x \in Y$, then by Theorem I.12, we have that $u_h(x) - u(x) \le Ch^\ell$. Since $X \cup Y = \mathbb{R}^N$, the conclusion follows. □

4.3 Extension to the case of a compact control set

In this section we show that our results extend to the case of a precompact set of controls. We endow the set of controls with the distance $\mathrm{d}(\alpha, \alpha') := |\Phi^\alpha - \Phi^{\alpha'}|_0$, where $\Phi^\alpha := (a^\alpha, b^\alpha, c^\alpha, f^\alpha)$. We suppose that $\sup_{\alpha\in\mathcal{A}} |\Phi^\alpha|_1 < +\infty$. Precompactness of $\mathcal{A}$ is equivalent to the following condition :

(A3) for every $\delta > 0$, there are $M \in \mathbb{N}$ and $\{\alpha_i\}_{i=1}^M \subset \mathcal{A}$, such that

$$\sup_{\alpha\in\mathcal{A}} \inf_{1\le i\le M} (|\sigma^\alpha - \sigma^{\alpha_i}|_0 + |b^\alpha - b^{\alpha_i}|_0 + |c^\alpha - c^{\alpha_i}|_0 + |f^\alpha - f^{\alpha_i}|_0) \le \delta.$$

Consider the viscosity solution u of

$$\min\{\sup_{\alpha\in\mathcal{A}} L^\alpha(x, \mathcal{D}u(x)); u(x) - \psi(x)\} = 0, \qquad x \in \mathbb{R}^N.$$

Existence, unicity and Lipschitzness of u are proved in [13, Lemma A.1]. Fix δ and consider w_δ the viscosity solution of

$$\min\{\sup_{i\in\mathcal{I}_M} L^{\alpha_i}(x, \mathcal{D}w_\delta(x)); w_\delta(x) - \psi(x)\} = 0, \qquad x \in \mathbb{R}^N,$$

where $\mathcal{I}_M := \{1, \dots, M\}$, M given by (A3). As in ([1, Lemma 3.3]), we can show, by adapting the methods, that

$$|u - w_\delta|_0 \le C\delta, \tag{4.24}$$

where C depends only on K and λ. If we note u_h the approximation of u and $w_{h,\delta}$ the approximation of w_δ, then we have $u_h \leq w_{h,\delta}$, and $w_{h,\delta} - w_\delta \leq Ch^{\bar{\gamma}}$, where $\bar{\gamma} = \min_{i \in J}\{k_i/(3i-2)\}$, k_i given by (S3). From the proof of Proposition I.12, we can see that C is independent of the size of $\mathcal{I}_M$. Then we have that

$$-Ch^{\gamma} \leq u_h - u \leq u_h - w_{h,\delta} + w_{h,\delta} - w + w - u \leq Ch^{\bar{\gamma}} + C_1\delta, \tag{4.25}$$

where $\gamma = \min_{i\in J}\{k_i/i\}$, $\bar{\gamma} = \min_{i\in J}\{k_i/(3i-2)\}$, k_i given by (S3). All constants being independent of the size of $\mathcal{I}_M$, then we can choose δ of the order of $h^{\bar{\gamma}}$ and we obtain the same result as in Theorem I.13.

Remark I.14. *It may happen that only $w_{h,\delta}$ is actually computed, and in that case it is useful to estimate $u - w_{h,\delta}$. Since $|u - w_\delta| \leq C\delta$, it follows from previous discussion that*

$$-C(\delta + h^{\gamma}) \leq w_{h,\delta} - u \leq C(\delta + h^{\bar{\gamma}}).$$

5 Specific approximation schemes

In this section we apply our previous results to some specific monotone discretization schemes.

5.1 Finite differences, one dimensional problem

Let x be in $\mathbb{R}$, ϕ in $C^n(\mathbb{R})$, h in $\mathbb{R}$ and define

$$\delta_{\pm}\phi(x) = \frac{\phi(x \pm h) - \phi(x)}{h}, \quad \Delta\phi(x) = \frac{\phi(x+h) - 2\phi(x) + \phi(x-h)}{h^2}.$$

In particular, by a Taylor expansion, we obtain

$$|\delta_{\pm}\phi(x) - D\phi(x)| \leq \frac{1}{2}h|D^2\phi|, \quad |\Delta_h\phi(x) - D^2\phi(x)| \leq \frac{1}{12}h^2|D^4\phi|.$$

Consider the finite difference scheme in $\mathbb{R}$:

$$S(h,x,r,\phi) :=$$
$$= \sup_{\alpha}\{-a^{\alpha}(x)\Delta\phi(x) - b_+^{\alpha}(x)\delta_+\phi(x) + b_-^{\alpha}(x)\delta_-\phi(x) + c^{\alpha}(x)r - f^{\alpha}(x)\}, \tag{5.1}$$

where $b_+^{\alpha}(x) = \max(b^{\alpha}(x), 0)$, and $b_-^{\alpha}(x) = \max(-b^{\alpha}(x), 0)$. For the consistency hypothesis (S3), we obtain, from the above Taylor expansion, $Q(\phi) = |D^2\phi|h + |D^4\phi|h^2$, i.e. $k_2 = 1$ and $k_4 = 2$. Then, by (4.3) and (4.23), we have

$$-Ch^{1/5} \leq u - u_h \leq Ch^{1/2}. \tag{5.2}$$

Remark I.15. *Consider a general scheme $S : \mathbb{R}^+ \times \mathbb{R}^N \times \mathbb{R} \times C_b(\mathbb{R}^N) \to \mathbb{R}$, which satisfies (S1), (S2) and (S3), for some $k_i > 0$, $i \in J$. To obtain equal or better estimate than (5.2) we must have :*

$$\min_{i\in J} \frac{k_i}{i} \geq \frac{1}{2}, \quad \min_{i\in J} \frac{k_i}{3i-2} \geq \frac{1}{5}. \tag{5.3}$$

In particular, the k_i which give an equal or better estimate than (5.2) are

$$\text{(i) } \mathrm{k_i} \geq \mathrm{i}/2,\ \textit{for } \mathrm{i} \leq 4; \quad \text{(ii) } \mathrm{k_i} \geq (3\mathrm{i}-2)/5,\ \textit{for } \mathrm{i} \geq 4. \tag{5.4}$$

Indeed, let $i \leq 4$. If $k_i \geq i/2$, then we have also $k_i \geq (3i-2)/5$. Moreover, if $i > 4$, we have $k_i \geq (3i-2)/5$ and also $k_i \geq i/2$. Hence we obtain (5.3).
If the inequalities in (5.4) are strictly satisfied, then also the inequalities in (5.3) are strictly satisfied and we obtain a better estimate.

5.2 Markov chain approximation

The scheme (5.1) may be viewed as a particular Markov chain approximation of (1.1). We consider now a general Markov chain approximation of (1.1) in a regular grid, and we want to find conditions on the probabilities of transition to obtain estimate as in (5.2). We consider a discretization step $h \in \mathbb{R}$ and a regular grid of discretization G^h. With the coordinate $k = (k_1, \ldots, k_N)$ in $\mathbb{Z}^N$, is associated the point $x_k \in \mathbb{R}^N$ of the form $x_k := (k_1 h, \ldots, k_N h)$. Let $\{X_q^h, q \geq 0\}$ the states of the Markov chain at time q, with transition probabilities $p(x_k, y|\alpha)$, α being the control value. Let Δt^h be an interpolation interval satisfying $\Delta t^h \to 0$ as $h \to 0$, and let $\mathbb{E}_{k,q}^{h,\alpha}$ be the conditional expectation of X_{q+1}^h, given $\{X_q^h = x_k\}$ and the control value α. A possible adaptation for the cost function to this Markov chain is the following :

$$W^h(x,\alpha) = \Delta t^h \Big[\sum_{q \geq 0} f^\alpha(X_q^h)(1 + c^\alpha(x)\Delta t^h)^{-q-1} \Big].$$

Applying the dynamic programmic principle for the controlled chain $\{X_q^h, q \geq 0\}$, at state $x_k \in G_h$, we obtain the following relation :

$$u_h(x_k) = \max\{\inf_\alpha (\frac{1}{1 + c^\alpha(x_k)\Delta t^h} (\sum_y p(x_k, y|\alpha) u_h(y) + f^\alpha(x_k)\Delta t^h)); \psi(x_k)\}. \tag{5.5}$$

Since $1 + c^\alpha(x_k)\Delta t^h \geq 0$ for all α, (5.5) may be written in the form (1.2), with

$$S(h, x_k, r, \phi) = \sup_\alpha \Big\{ -\frac{1}{\Delta t^h} \sum_y p(x_k, y|\alpha)\phi(y) - f^\alpha(x_k) + \frac{1}{\Delta t^h} r + c^\alpha(x_k) r \Big\}. \tag{5.6}$$

With the above definition for S, the assumptions (S1) and (S2) are satisfied. Suppose that (S3) is satisfied and we want to look for simple conditions on the probabilities $p(x_k, y)$ and on the k_i defined in (S3), under which we obtain equal or better estimate than (5.2). We note $\mathbb{P}_{x,y} = \sum_y p(x, y|\alpha)$. Using remark I.15, we obtain the following :

Theorem I.16. *Let S defined as in (5.6). Suppose that (S3) is satisfied for some $k_i, i \in J$.*

(i) We have an equal or better estimate than (5.2) if and only if
(a1) $\|\frac{1}{\Delta t^h}\mathbb{P}_{x,y}(x-y) - b^\alpha(x)\| = K_C h^{k_1}$,
(b1) $\|\frac{1}{2\Delta t^h}\mathbb{P}_{x,y}(y-x)^2 - a^\alpha(x)\| = K_C h^{k_2}$,
(c1) $\|\frac{1}{i!\Delta t^h}\mathbb{P}_{x,y}(y-x)^i\| = K_C h^{k_i}$, *for* $i = 3, 4$,
with

$$k_1 \geq \frac{1}{2},\; k_2 \geq 1,\; k_3 \geq \frac{3}{2},\; k_4 \geq 2. \tag{5.7}$$

(ii) Moreover we have a better lower bound if and only if, in addition, k_4 satisfies strictly (5.7).
(iii) We have a better upper bound if and only if all the inequalities in (5.7) are strictly satisfied.

Proof. We give the proof for $N = 1$, the general case follows immediately. Fix $x \in \mathbb{R}$, and let $\phi \in C^n(\mathbb{R})$, such that $D^i\phi$ is bounded for $i = 1, \dots, n$. Set $\Delta^\phi := |\sup_\alpha L^\alpha(x, \mathcal{D}\phi(x)) - S(h, x, \phi(x), \phi)|$. An upper bound of Δ^ϕ is

$$|\sup_\alpha\Big(-\operatorname{tr}[a^\alpha(x)D^2\phi(x)] - b^\alpha(x)D\phi(x) + \frac{1}{\Delta t^h}\sum_y p(x,y|\alpha)(\phi(y)-\phi(x))\Big)|$$

From the Taylor expansion of $\phi(y)$ up to order 4, we deduce that

$$\Delta^\phi \leq \Delta_1^\phi + \Delta_2^\phi + \Delta_3^\phi + \Delta_4^\phi, \tag{5.8}$$

where

$$\begin{aligned}
\Delta_1^\phi &:= \sup_\alpha |-b^\alpha(x)D\phi(x) + \frac{1}{\Delta t^h}\mathbb{P}_{x,y}D\phi(x)(y-x)|,\\
\Delta_2^\phi &:= \sup_\alpha |-tr[a^\alpha(x)D^2\phi(x)] + \frac{1}{2\Delta t^h}\mathbb{P}_{x,y}D^2\phi(x)(y-x)^2|,\\
\Delta_3^\phi &:= \sup_\alpha |\frac{1}{3!\Delta t^h}\mathbb{P}_{x,y}D^3\phi(x)(y-x)^3|,\\
\Delta_4^\phi &:= \sup_\alpha |\frac{1}{4!\Delta t^h}\mathbb{P}_{x,y}D^4\phi(c)(y-x)^3|,
\end{aligned}$$

where $c \in [x, y]$ if $y \geq x$, $c \in [y, x]$ otherwise. Suppose now that conditions (a1)-(d1) and (5.7) are satisfied. Then $J = \{1, 2, 3, 4\}$, and applying remark I.15, we obtain the result. Moreover, if $k_4 > 2$, then $k_i/(3i-2) > 1/5$ for all i. Hence we obtain a strictly better lower bound. Since $k_i/i \geq 1/2$ for all i in J, if all k_i satisfy strictly (5.7), we have a better upper bound. Suppose now that we have a better or equal estimate than (5.2). Then we have

$$\Delta^\phi \leq K_C \sum_{i\in J} |D^i\phi| h^{k_i}, \tag{5.9}$$

with $\min_{i\in J} k_i/i \geq 1/2$ and $\min_{i\in J} k_i/(3i-2) \geq 1/5$, and here $J = \{1,2,3,4\}$. From (5.9), (5.8) and remark I.17, we have that (a1)-(d1) are satisfied with k_i as in (5.7). If the lower bound is strictly bigger that $1/5$, since $k_i/(3i-2) > 1/5$ for $i = 1,2,3$, then we must have $k_4 > 2$. If the upper bound is strictly bigger than $1/2$, since $k_i \geq i/2$ for all i, then we must have $k_i > i/2$ for all i.□

Remark I.17. *(i) We have that conditions (a1) and (b1) imply the consistence in the sense of Kushner (see [16]), i.e.*

$$\|\mathbb{E}(y-x) - b^\alpha(x)\Delta t^h\| \leq \Delta t^h r_1, \quad \|Cov(y) - 2a^\alpha(x)\| \leq \Delta t^h r_2.$$

In [16] we have $r_i = o(1)$, for $i = 1,2$. Our error estimate need the more restrictive conditions $r_1 = h^{k_1}$ and $r_2 = h^{k_2} + \Delta t^h h^{2k_1} + \Delta t^h$.
(ii) We remark that to obtain (a1)-(d1), we use the inequality

$$\Delta_i^\phi \leq |D^i\phi| \cdot \|\mathbb{E}(y-x)^i - a^i\|, \tag{5.10}$$

for some a^i and for all ϕ. This inequality is sharp, since $\|\mathbb{E}(y-x)^i - a^i\|$ is the optimal constant for which we have this upper bound (for any function ϕ). Indeed, let B an i-linear symmetric form. The optimal constant C for which

$$|D^i\phi(x)B| \leq C|D^i\phi|, \ \forall \phi \in C^n(\mathbb{R}^N) \text{ such that } D^i\phi \text{ is bounded } \forall\, i,$$

is $C = |B|$. Indeed, we may identify a^i and $\mathbb{E}(y-x)^i$ with i-linear symmetric forms, and the above display reduces to the Cauchy-Schwarz inequality for i-linear symmetric forms. (iii) Let $\Delta^\phi := |\sum_{i=1}^4 D^i\phi(x)\mathbb{E}(y-x)^i - a^i D^i\phi(x)|$, for some a^i. We have that the optimal constants C_i such that

$$\Delta \leq \sum_{i=1}^4 C_i|D^i\phi(\mathbb{E}(y-x)^i - a^i)|, \ \forall \phi \in C^n(\mathbb{R}^N) \text{ such that } D^i\phi \text{ is bounded } \forall\, i,$$

are $C_i = 1$, for all i.

5.3 A counter example

We give here an example of a finite difference scheme for which the k_i do not satisfy conditions given in remark I.15, and we will show that we obtain a estimate worse than (5.2). Consider the following equation

$$\sup_\alpha\{-\operatorname{tr}[a^\alpha(x)D^2u(x)] + c^\alpha(x)u(x) - f^\alpha(x)\} = 0, \quad x \in \mathbb{R}^2, \tag{5.11}$$

with

$$b^\alpha(x) = 0, a^\alpha(x) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \forall\, x,\ \alpha.$$

Let h the discretization step and Δt^h the interpolation interval. We consider the following probabilities of transition :

$$p(x, x - he_2|\alpha) = \frac{1}{2}; \quad p(x, x \pm he_1 + he_2|\alpha) = \frac{1}{4}.$$

In particular, if we choose $\Delta t^h = \frac{1}{4}h^2$, we have that these probabilities verify

$$\mathbb{E}(y - x) = \frac{1}{2}\begin{pmatrix} 0 \\ -h \end{pmatrix} + \frac{1}{4}\begin{pmatrix} h \\ h \end{pmatrix} + \frac{1}{4}\begin{pmatrix} -h \\ h \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$\begin{aligned} \mathbb{E}(y - x)^2 &= \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & h^2 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} h^2 & h^2 \\ h^2 & h^2 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} h^2 & -h^2 \\ -h^2 & h^2 \end{pmatrix} \\ &= \Delta t^h \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

We have that

$$\begin{aligned} &\sum_y \frac{p(x, y)}{h^2} D^3\phi(x)(y - x)^3 \\ &= \sum_y \frac{p(x, y)}{h^2}\Big[\sum_{i=0}^{3} \frac{\partial^3\phi}{\partial x_1^i \partial x_2^{3-i}}(x)(y_1 - x_1)^i (y_2 - x_2)^{3-i}\Big] \\ &= -\frac{1}{2}\frac{\partial^3\phi}{\partial x_2^3}(x)h + \frac{1}{4}\Big(\frac{\partial^3\phi}{\partial x_1^3}(x)h + \frac{\partial^3\phi}{\partial x_1^2 x_2}(x)h + \frac{\partial^3\phi}{\partial x_1 x_2^2}(x)h + \frac{\partial^3\phi}{\partial x_2^3}(x)h\Big) \\ &\quad + \frac{1}{4}\Big(-\frac{\partial^3\phi}{\partial x_1^3}(x)h + \frac{\partial^3\phi}{\partial x_1^2 x_2}(x)h - \frac{\partial^3\phi}{\partial x_1 x_2^2}(x)h + \frac{\partial^3\phi}{\partial x_2^3}(x)h\Big) \\ &= \frac{1}{2}\frac{\partial^3\phi}{\partial x_1^2 x_2}(x)h. \end{aligned}$$

Hence we can write (S3) in the following way

$$|\sup_\alpha L^\alpha(x, \phi(x), D\phi(x), D^2\phi(x)) - S(x, h, \phi(x), \phi)| \leq$$

$$\frac{1}{2}|\frac{\partial^3\phi}{\partial x_1^2 \partial x_2}(x)|h + \left(\frac{1}{2}|\frac{\partial^4\phi}{\partial x_1^4}(x)| + \frac{1}{2}|\frac{\partial^4\phi}{\partial x_1^2 \partial x_2^2}|(x) + |\frac{\partial^4\phi}{\partial x_2^4}|(x)\right)h^2.$$

So, we have $k_3 = 1$ and $k_4 = 2$, and by applying Theorem I.10 and Theorem I.13 we obtain

$$-Ch^{1/7} \leq u - u_h \leq Ch^{1/3}. \tag{5.12}$$

5.4 The generalized finite differences scheme

We consider the generalized finite differences scheme defined in [6]. Let $\varphi = \{\varphi_k\}$ be a real valued function over $\mathbb{Z}^N$. Let $\xi \in \mathbb{Z}^N$ and consider the finite difference operator

$$\Delta_\xi \varphi_k := \varphi_{k+\xi} + \varphi_{k-\xi} - 2\varphi_k.$$

In ϕ belongs to $C^2(\mathbb{R}^N)$, and $\varphi_k := \phi(x_k)$ for all k, then we have

$$\Delta_\xi \varphi_k := \phi(x_{k+\xi}) + \phi(x_{k-\xi}) - 2\phi(x_k).$$

Then we consider

$$(D_k u_h(x_k))_i = \begin{cases} \dfrac{u_h(x_{k+e_i}) - u_h(x_k)}{h} & \text{if } \; b_i^\alpha(x_k) \geq 0, \\ \dfrac{u_h(x_k) - u_h(x_{k-e_i})}{h} & \text{if } \; b_i^\alpha(x_k) \leq 0. \end{cases}$$

Let $\mathcal{S}$ be a finite set of $\mathbb{Z}\backslash\{0\}$ containing $\{e_1, \dots, e_N\}$. We consider the following probabilities of transition

$$\begin{aligned} & p^\alpha(x_k, x_k|\alpha) = 1 - \Delta t^h \sum_{i=1}^N \left(\frac{|b_i^\alpha(x_k)|}{h} + 2\sum_{\xi\in\mathcal{S}} \alpha_{k,\xi} \right), \\ & p^\alpha(x_k, x_k \pm e_i h|\alpha) = \Delta t^h \left(\frac{b_i^{\alpha\pm}(x_k)}{h} + \alpha_{k,e_i} \right), \\ & p^\alpha(x_k, x_k + \xi h|\alpha) = \Delta t^h \alpha_{k,\xi} \quad \text{ for } \xi \in \mathcal{S},\ \xi \neq e_i, \\ & p^\alpha(x_k, y|\alpha) = 0 \quad \text{ for } y \notin x_{k+\mathcal{S}}. \end{aligned}$$

Then we can write (5.6) in the following way :

$$S(h, x_k, r, \phi) = \sup_\alpha \Big\{ -\sum_{\xi\in\mathcal{S}} \alpha_{k,\xi} \Delta_\xi \phi(x_k) - b^\alpha(x_k) D_k \phi(x_k) + c^\alpha(x_k) r - f^\alpha(x_k) \Big\}. \qquad (5.13)$$

We make the strong consistency hypothesis on the matrix

$$a^\alpha(x) = \sum_{i,j} h^2 \xi_i \xi_j \alpha_{k,\xi} e_i e_j^T, \quad \forall\, k \in \mathbb{Z}^N.$$

The scheme defined in (5.13) satisfies (S1) et (S2). We consider a function $\phi \in C^2(\mathbb{R}^N)$. By applying a Taylor expansion, we obtain

(S3) $|\sup_{\alpha\in\mathcal{A}} L^\alpha(x, \phi, D\phi, D^2\phi) - S(x, h, \phi(x), \phi)| \leq$
$\sup_{\alpha\in\mathcal{A}} |b^\alpha|_0 |D^2\phi|_0 h + \sup_{\alpha\in\mathcal{A}} |\sigma^\alpha|_0^2 |D^4\phi|_0 h^2.$

So we can say that $k_2 = 1$ and $k_4 = 2$. Applying Theorems I.10 and I.13, we obtain the same estimate as in the case of one player (see [1]).

Theorem I.18. *Assume (A1)-(A4), (S1)-(S3). If u and u_h are solution of (1.1) and (1.2), with S defined as in (5.13), then for h sufficiently small we obtain*

$$-Ch^{1/5} \leq u - u_h \leq Ch^{1/2}. \quad \square$$

Appendix

6 Well-posedness of the switching system

In this appendix we prove the well-posedness of the switching system (3.1), for $k \geq 0$, under assumptions (A1) and (A2) on the coefficients (stated in Section 2). Well-posedness of the original equation (1.1) is given in [13, lemma A.1]. Let us start by stating a technical lemma which is an easy extension of ([1, Lemma A.2]) :

Lemma I.19. *Let v be a bounded and continuous sub-solution on (3.1) and $\bar{v}$ be a bounded and continuous super-solution of another equation (3.1), where L^α is replaced by $\bar{L}^\alpha$, satisfying the same assumptions with coefficients $(\bar{\sigma}^\alpha, \bar{b}^\alpha, \bar{c}^\alpha, \bar{f}^\alpha)$. Let $g \in C^2(\mathbb{R}^N \times \mathbb{R}^N)$. Consider*

$$m := \sup_{i,x,y}\{v_i(x) - \bar{v}_i(y) - g(x,y)\},$$

and suppose that the "sup" is attained at some point (i_0, x_0, y_0). Set

$$A := \{i \in \mathcal{I} | (i, x_0, y_0) \text{ realizes the sup}\}; \quad I(x_0) := \{i \in \mathcal{I} | v_i(x_0) \leq \psi(x_0)\}.$$

If $A \cap I(x_0) = \varnothing$, then there exists $i \in A$ such that

$$\bar{v}_i(y_0) < \min_{j \neq i}\{\bar{v}_j(y_0) + k\}. \tag{6.1}$$

Proof. We proceed by contradiction. Let j in A. If (6.1) does not hold, there exists $\ell \in \mathcal{I}$ such that

$$\bar{v}_j(y_0) \geq \bar{v}_l(y_0) + k. \tag{6.2}$$

Since $A \cap I(x_0) = \varnothing$, then for all $i \in A$,

$$\max\{L^{\alpha_i}(x_0, v_i(x_0), D_x g(x_0,y_0), D^2_{xx} g(x_0,y_0)); v_i(x_0) - \min_{j \neq i}\{v_j(x_0) + k\}\} \leq 0.$$

Hence we obtain $v_j(x_0) \leq v_l(x_0) + k$, and then with (6.2),

$$v_j(x_0) - \bar{v}_j(y_0) \leq v_l(x_0) - \bar{v}_l(y_0). \tag{6.3}$$

Therefore $l \in A$, and equality holds in (6.3). Then $\bar{v}_j(y_0) = \bar{v}_l(y_0) + k$. Since A is a finite set, this proves that there exists a finite sequence $j_1, \ldots, j_K \in A$ such that $\bar{v}_{j_i}(y_0) = \bar{v}_{j_{i+1}}(y_0) + k$ for $i = 1, \ldots, K-1$, and $j_1 = j_K$. So we obtain

$$0 = \sum_{i=1}^{K-1} (\bar{v}_{j_i}(y_0) - \bar{v}_{j_{i+1}}(y_0)) = (K-1)k > 0,$$

and we have a contradiction. Therefore (6.1) holds. □

Now we can state the following lemma about comparison, existence, uniqueness and the bounds on the solution $v = (v_1, \ldots, v_M)$ of (3.1). This is an easy extension of ([1, Theorem A.1]).

Lemma I.20. *Under assumptions (A1) and (A2), the following statements hold :*

(a) If v and w are respectively sub-solution and super-solution of (3.1), such that v_i, $w_i \in C_b(\mathbb{R}^N)$ for all $i \in \mathcal{I}$, then $v \leq w$ in $\mathbb{R}^N$.

(b) There exists a unique viscosity solution v of (3.1), such that $v_i \in C_{b,l}(\mathbb{R}^N)$ for all $i \in \mathcal{I}$. This solution satisfies

$$\max_i |v_i|_0 \leq \max\left\{\lambda^{-1}\sup_\alpha |f^\alpha|_0; |\psi|_0\right\}, \tag{6.4}$$

$$\max_i [v_i]_1 \leq \max\left\{\sup_{i,\alpha}\frac{[c^\alpha]_1|v_i|_0 + [f^\alpha]_1}{\lambda - [\sigma^\alpha]_1^2 - [b^\alpha]_1}; [\psi]_1\right\}. \tag{6.5}$$

Proof. (a) This is a consequence of the comparison principle [11, Theorem 3.1]. Indeed, in [11], the comparison principle is proved for the Dirichlet problem on a bounded domain. To extend the result to an unbounded domain, we have only to modify the test functions of [11] in the standard way. (b) Existence and uniqueness follow from the comparison principle (a). Let $M := \max\{\sup_\alpha \lambda^{-1}|f^\alpha|_0; |\psi|_0\}$. It is easy to see that M and $-M$ are respectively super and sub-solution of (3.1). Hence, by the comparison principle (a) we obtain the bound on $\max_i |v_i|_0$.

To obtain the bound on $\max_i [v_i]_1$, we set

$$m := \sup_{i,x,y} \phi_i(x,y) := \sup_{i,x,y}\{v_i(x) - v_i(y) - L|x-y| - \epsilon(|x|^2 + |y|^2)\},$$

where $L > 0$. If, by setting

$$L := \max\left\{\sup_{i,\alpha}\left\{\frac{[c^\alpha]_1|v_i|_0 + [f^\alpha]_1}{\lambda - [\sigma^\alpha]_1^2 - [b^\alpha]_1}\right\}; [\psi]_1\right\},$$

we can conclude that $m \leq 0$, then, sending ϵ to 0, we have done. Assume that the supremum is attained at a point (i_0, x_0, y_0). If $x_0 = y_0$, then $m \leq 0$, and, sending ϵ to 0 we have the result. If not, since $L|x-y|$ is smooth at x_0, y_0, we can apply a doubling of variables argument. Set

$$P_0 := \left[I - \frac{(x_0 - y_0)}{|x_0 - y_0|}\left(\frac{x_0 - y_0}{|x_0 - y_0|}\right)^T\right], \quad \Lambda := \begin{pmatrix} P_0 & -P_0 \\ -P_0 & P_0 \end{pmatrix}.$$

Define the following sets :

$$A := \{i \in \mathcal{I} | (i, x_0, y_0) \text{ realizes the sup}\}, \; I(x_0) := \{i \in \mathcal{I} | v_i(x_0) \leq \psi(x_0)\}.$$

The maximum principle for semi-continuous functions (see [8]), and the definition of viscosity solutions imply that, for $i \in A$, there exist $X, Y \in \mathcal{S}^N$ such that

$$\min\{\max(L^{\alpha_i}(x_0, v_i(x_0), p_x, X); v_i(x_0) - \min_{j\neq i}\{v_j(x_0) + k\}); v_i(x_0) - \psi(x_0)\} \leq 0,$$

$$\min\{\max(L^{\alpha_i}(y_0, v_i(y_0), p_y, Y); v_i(y_0) - \min_{j\neq i}\{v_j(y_0) + k\}); v_i(y_0) - \psi(y_0)\} \geq 0,$$

where $p_x = L\frac{(x_0-y_0)}{|x_0-y_0|} + 2\varepsilon x_0$, $p_y = L\frac{(x_0-y_0)}{|x_0-y_0|} - 2\varepsilon y_0$, and

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{L}{|x_0 - y_0|}\Lambda + 2\varepsilon \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

<u>CASE 1</u> : There exists $i \in A \cap I(x_0)$, i.e., $v_i(x_0) \leq \psi(x_0)$. Since $\bar{v}_i(y_0) \geq \psi(y_0)$, for all $i \in A$, we have

$$v_i(x_0) - \bar{v}_i(y_0) \leq \psi(x_0) - \psi(y_0) \leq [\psi]_1|x_0 - y_0|.$$

<u>CASE 2</u> : The set $A \cap I(x_0)$ is empty. Then

$$\max\{L^{\alpha_i}(x_0, v_i(x_0), p_x, X); v_i(x_0) - \min_{j\neq i}\{v_j(x_0) + k\}\} \leq 0, \quad \forall\, i \in A. \tag{6.6}$$

Since $\max\{L^{\alpha_{i_0}}(y_0, v_{i_0}(y_0), p_y, Y); v_{i_0}(y_0) - \min_{j\neq i_0}\{v_j(y_0) + k\}\} \geq 0$, applying Lemma I.19, we obtain

$$L^{\alpha_{i_0}}(x_0, v_{i_0}(x_0), p_x, X) \leq 0 \leq L^{\alpha_{i_0}}(y_0, v_{i_0}(y_0), p_y, Y).$$

Since P_0 is a projection, and hence, is nonexpansive, we have that

$$\begin{aligned} \begin{pmatrix} \sigma^\alpha(x_0) \\ \sigma^\alpha(y_0) \end{pmatrix}^T \Lambda \begin{pmatrix} \sigma^\alpha(x_0) \\ \sigma^\alpha(y_0) \end{pmatrix} &= (\sigma^\alpha(x_0) - \sigma^\alpha(y_0))\, P_0\, (\sigma^\alpha(x_0) - \sigma^\alpha(y_0)) \\ &\leq |\sigma^\alpha(x_0) - \sigma^\alpha(y_0)|^2. \end{aligned}$$

Now we can proceed as in the standard situation (see [1, Theorem A.1]).
Combining cases 1 and 2 we obtain the result.□

By using ([2, Theorem A.1]), we prove the following theorem.

Theorem I.21. *Let v and $\bar{v}$ be solutions of (3.1) with coefficients σ, b, c, f and $\bar{\sigma}$, $\bar{b}$, $\bar{c}$, $\bar{f}$ respectively. Suppose that assumptions (A1), (A2) are satisfied for both sets of coefficients with the same λ, and $\max_i |v_i|_1 + \max_i |\bar{v}_i|_1 \leq M < \infty$. Then*

$$\lambda \max_i |v_i - \bar{v}_i|_0 \leq M\big(\sup_\alpha\{|\sigma^\alpha - \bar{\sigma}^\alpha|_0 + |b^\alpha - \bar{b}^\alpha|_0 + |c^\alpha - \bar{c}^\alpha|_0 + |f^\alpha - \bar{f}^\alpha|_0\}\big),$$

where M depends on K, $\sup_i |v_i|_0$, $\sup_i |\bar{v}_i|_0$.

Proof. We set

$$m := \sup_{i,x,y} \phi_i(x, y) := \sup_{i,x,y}\{v_i(x) - \bar{v}_i(y) - \delta|x - y|^2 - \varepsilon(|x|^2 + |y|^2)\},$$

where $\delta > 0$ and $\varepsilon > 0$. The "sup" is attained at a point (i, x_0, y_0), so

$$m = v_i(x_0) - \bar{v}_i(y_0) - \delta|x_0 - y_0|^2 - \varepsilon(|x_0|^2 + |y_0|^2).$$

Let

$$A := \{i \in \mathcal{I} | (i, x_0, y_0) \text{ realize the } \sup\}, \quad I(x_0) := \{i \in \mathcal{I} | v_i(x_0) \le \psi(x_0)\}.$$

The maximum principle for semi-continuous functions (see [8]), and the definition of viscosity solutions imply that, for $i \in A$, there exist $X, Y \in \mathcal{S}^N$ such that

$$\min\{\max(L^{\alpha_i}(x_0, v_i(x_0), p_x, X); v_i(x_0) - \min_{j \neq i}\{v_j(x_0) + k\}); v_i(x_0) - \psi(x_0)\} \le 0,$$

$$\min\{\max(L^{\alpha_i}(y_0, \bar{v}_i(y_0), p_y, Y); \bar{v}_i(y_0) - \min_{j \neq i}\{\bar{v}_j(y_0) + k\}); \bar{v}_i(y_0) - \psi(y_0)\} \ge 0,$$

where $p_x = 2\delta(x_0 - y_0) + 2\varepsilon x_0$, $p_y = 2\delta(x_0 - y_0) - 2\varepsilon y_0$, and there exists $\ell > 0$ such that

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le \ell\delta \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \ell\varepsilon \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \mathcal{O}(k).$$

We have to study two different cases.
CASE 1 : If there exists $i \in A \cap I(x_0)$, then $v_i(x_0) \le \psi(x_0)$. Since $\bar{v}_i(y_0) \ge \psi(y_0)$, then we have

$$v_i(x_0) - \bar{v}_i(y_0) \le \psi(x_0) - \psi(y_0) \le [\psi]_1 |x_0 - y_0|.$$

CASE 2 : If $A \cap I(x_0) = \varnothing$, then

$$\max\{L^{\alpha_i}(x_0, v_i(x_0), p_x, X); v_i(x_0) - \min_{j \neq i}\{v_j(x_0) + k\}\} \le 0, \quad \forall\, i \in A. \tag{6.7}$$

Since $\max\{L^{\alpha_{i_0}}(y_0, \bar{v}_{i_0}(y_0), p_y, Y); \bar{v}_{i_0}(y_0) - \min_{j \neq i_0}\{\bar{v}_j(y_0) + k\}\} \ge 0$, applying Lemma I.19, we obtain

$$L^{\alpha_{i_0}}(x_0, v_{i_0}(x_0), p_x, X) \le 0 \le \bar{L}^{\alpha_{i_0}}(y_0, \bar{v}_{i_0}(y_0), p_y, Y),$$

and then

$$\begin{aligned} 0 \le & -\operatorname{tr}[\bar{a}^{\alpha_{i_0}}(y_0)Y] + \operatorname{tr}[a^{\alpha_{i_0}}(x_0)X] - \bar{b}^{\alpha_{i_0}}(y_0)p_y + b^{\alpha_{i_0}}(x_0)p_x + \\ & \bar{c}^{\alpha_{i_0}}(y_0)\bar{v}_{i_0}(y_0) - c^{\alpha_{i_0}}(x_0)v_{i_0}(x_0) - \bar{f}^{\alpha_{i_0}}(y_0) + f^{\alpha_{i_0}}(x_0) \\ & =: (I) + (II) + (III) + (IV). \end{aligned}$$

As in [2], we analyze each term separately :

$$\begin{aligned} (I) &= \operatorname{tr}[a^{\alpha_{i_0}}(x_0)X] - \operatorname{tr}[\bar{a}^{\alpha_{i_0}}(y_0)Y] \\ &\le 2\ell\delta\{|\sigma^{\alpha_{i_0}}(x_0) - \bar{\sigma}^{\alpha_{i_0}}(x_0)|^2 + |\bar{\sigma}^{\alpha_{i_0}}(x_0) - \bar{\sigma}^{\alpha_{i_0}}(y_0)|^2\} \\ &\quad + \ell\varepsilon\{|\sigma^{\alpha_{i_0}}(x_0)|^2 + |\bar{\sigma}^{\alpha_{i_0}}(y_0)|^2\}, \\ (II) &= (b^{\alpha_{i_0}}(x_0) - \bar{b}^{\alpha_{i_0}}(y_0))(x_0 - y_0) \\ &\le 2|b^{\alpha_{i_0}}(x_0) - \bar{b}^{\alpha_{i_0}}(x_0)|^2 + 2|x_0 - y_0|^2 \\ &\quad + |\bar{b}^{\alpha_{i_0}}(x_0) - \bar{b}^{\alpha_{i_0}}(y_0)|\,|x_0 - y_0|, \\ (III) &= \bar{c}^{\alpha_{i_0}}(y_0)\bar{u}(y_0) - c^{\alpha_{i_0}}(x_0)u(x_0) \\ &\le |u(x_0)|\,|c^{\alpha_{i_0}}(x_0) - \bar{c}^{\alpha_{i_0}}(x_0)| + |\bar{u}(y_0)|\,|\bar{c}^{\alpha_{i_0}}(x_0) - \bar{c}^{\alpha_{i_0}}(y_0)| \\ &\quad - \lambda m, \\ (IV) &= f^{\alpha_{i_0}}(x_0) - \bar{f}^{\alpha_{i_0}}(y_0) \\ &\le |f^{\alpha_{i_0}}(x_0) - \bar{f}^{\alpha_{i_0}}(x_0)| + |\bar{f}^{\alpha_{i_0}}(x_0) - \bar{f}^{\alpha_{i_0}}(y_0)|\,|x_0 - y_0|. \end{aligned}$$

Hence we obtain

$$\lambda m \le 2\ell\delta\{|\sigma^{\alpha_{i_0}} - \bar\sigma^{\alpha_{i_0}}|_0^2 + |b^{\alpha_{i_0}} - \bar b^{\alpha_{i_0}}|_0^2\} +$$
$$+\{|v_{i_0}|_0|\bar c^{\alpha_{i_0}} - c^{\alpha_{i_0}}|_0 + |f^{\alpha_{i_0}} - \bar f^{\alpha_{i_0}}|_0\} +$$
$$+2\delta\{\ell[\bar\sigma^{\alpha_{i_0}}]_1^2 + [\bar b^{\alpha_{i_0}}]_1 + 2\}|x_0 - y_0|^2 +$$
$$\{|\bar v_{i_0}|_0[\bar c^{\alpha_{i_0}}]_1 + [\bar f^{\alpha_{i_0}}]_1\}|x_0 - y_0| + C\varepsilon(1 + |x_0| + |y_0|).$$

We sum the bounds obtained in the two cases to have a general bound of m. So we obtain

$$\lambda m \le 2\ell\delta\{|\sigma^{\alpha_{i_0}} - \bar\sigma^{\alpha_{i_0}}|_0^2 + |b^{\alpha_{i_0}} - \bar b^{\alpha_{i_0}}|_0^2\} +$$
$$+\{|v_{i_0}|_0|\bar c^{\alpha_{i_0}} - c^{\alpha_{i_0}}|_0 + |f^{\alpha_{i_0}} - \bar f^{\alpha_{i_0}}|_0\} +$$
$$+2\delta\{\ell[\bar\sigma^{\alpha_{i_0}}]_1^2 + [\bar b^{\alpha_{i_0}}]_1 + 2\}|x_0 - y_0|^2 +$$
$$\{|\bar v_{i_0}|_0[\bar c^{\alpha_{i_0}}]_1 + [\bar f^{\alpha_{i_0}}]_1 + \lambda[\psi]_1\}|x_0 - y_0| + C\varepsilon(1 + |x_0|^2 + |y_0|^2).$$

We set $k_1 := \{2\ell[\bar\sigma^{\alpha_{i_0}}]_1^2 + 2[\bar b^{\alpha_{i_0}}]_1 + 4\}$, $k_2 := \{|\bar v_{i_0}|_0[\bar c^{\alpha_{i_0}}]_1 + [\bar f^{\alpha_{i_0}}]_1 + \lambda[\psi]_1\}$. From now on we proceed as in ([2, Theorem A.1]). Noting that $2\phi(x_0, y_0) \ge \phi(x_0, x_0) + \phi(y_0, y_0)$, we have

$$|x_0 - y_0| \le C\delta^{-1}, \tag{6.8}$$

where C depends K. The inequality (6.8) implies that

$$|x_0 - y_0|^2 \le C\delta^{-2}, \tag{6.9}$$

where C depends on K. So we obtain

$$\lambda m \le \{|v_{i_0}|_0|\bar c^{\alpha_{i_0}} - c^{\alpha_{i_0}}|_0 + |f^{\alpha_{i_0}} - \bar f^{\alpha_{i_0}}|_0\} +$$
$$+2\ell\delta\{|\sigma^{\alpha_{i_0}} - \bar\sigma^{\alpha_{i_0}}|_0^2 + |b^{\alpha_{i_0}} - \bar b^{\alpha_{i_0}}|_0^2\} + C(k_1 + k_2)\delta^{-1} + C\varepsilon(1 + |x_0|^2 + |y_0|^2).$$

We know that $v_{i_0}(x) - \bar v_{i_0}(x) - 2\varepsilon|x|^2 \le m$, and so

$$v_{i_0}(x) - \bar v_{i_0}(x) \le \{|v_{i_0}|_0|\bar c^{\alpha_{i_0}} - c^{\alpha_{i_0}}|_0 + |f^{\alpha_{i_0}} - \bar f^{\alpha_{i_0}}|_0\} +$$
$$+2\ell\delta\{|\sigma^{\alpha_{i_0}} - \bar\sigma^{\alpha_{i_0}}|_0^2 + |b^{\alpha_{i_0}} - \bar b^{\alpha_{i_0}}|_0^2\} + C(k_1 + k_2)\delta^{-1} + 2\varepsilon|x|^2 + C\varepsilon(1 + |x_0|^2 + |y_0|^2).$$

This inequality holds for all δ, and hence we minimize with respect to δ, by noting that for $l > 0$,

$$\min_{\delta > 0}(l\delta + C\delta^{-1}) = Cl^{1/2}.$$

Hence, by sending ε to zero, we obtain

$$v_{i_0}(x) - \bar v_{i_0}(x) \le \{|v_{i_0}|_0|\bar c^{\alpha_{i_0}} - c^{\alpha_{i_0}}|_0 + |f^{\alpha_{i_0}} - \bar f^{\alpha_{i_0}}|_0\} +$$
$$C\{|\sigma^{\alpha_{i_0}} - \bar\sigma^{\alpha_{i_0}}|_0^2 + |b^{\alpha_{i_0}} - \bar b^{\alpha_{i_0}}|_0^2\}^{1/2},$$

where C depends on K, $|v_{i_0}|_0, |\bar v_{i_0}|_0$ and $[\psi]_1$. Since $(s^2 + t^2)^{1/2} \le |s| + |t|$, we can conclude. □

Similarly we have the following result.

Theorem I.22. *Let u and $\bar{u}$ be solutions of (1.1) with coefficients σ, b, c, f and $\bar{\sigma}$, $\bar{b}$, $\bar{c}$, $\bar{f}$ respectively. Suppose that assumptions (A1), (A2) are satisfied for both sets of coefficients with the same λ, and $\max|u|_1 + \max|\bar{u}|_1 \leq M < \infty$. Then*

$$\lambda \max|u-\bar{u}|_0 \leq M\big(\sup_\alpha\{|\sigma^\alpha - \bar{\sigma}^\alpha|_0 + |b^\alpha - \bar{b}^\alpha|_0 + |c^\alpha - \bar{c}^\alpha|_0 + |f^\alpha - \bar{f}^\alpha|_0\}\big),$$

where M depends on K, $\sup|u|_0$, $\sup|\bar{u}|_0$. □

Acknowledgements We thank a referee for his numerous useful remarks, in particular the one improving the proof of Lemma I.20.

Bibliographie

[1] G. Barles and E.R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numerical Analysis*. To appear.

[2] G. Barles and E.R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN. Mathematical Modelling and Numerical Analysis*, 36 :33–54, 2002.

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

[4] J.F. Bonnans, S. Maroso, and H. Zidani. Error estimates for a stochastic impulse control problem. Technical Report 5606, INRIA, 2005. "http ://www.inria.fr/rrrt/rr-5606.html".
To appear in Applied Math. Optimization.

[5] J.F. Bonnans, E. Ottenwaelter, and H. Zidani. Numerical schemes for the two dimensional second-order HJB equation. *ESAIM : M2AN*, 38 :723–735, 2004.

[6] J.F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numerical Analysis*, 41 :1008–1021, 2003.

[7] I. Capuzzo-Dolcetta and L.C. Evans. Optimal switching for ordinary differential equations. *SIAM J. Control Optim.*, 22 :143–161, 1984.

[8] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. American Mathematical Society (New Series)*, 27 :1–67, 1992.

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

[10] H. Ishii and S. Koike. Viscosity solutions for monotone systems of second order elliptic PDEs. *Comm. Partial Differential Equation*, 16 :1095–1128, 1991.

[11] H. Ishii and S. Koike. Viscosity solutions of a system of nonlinear second order elliptic PDEs arising in switching games. *Funkcial. Ekvac.*, 34 :143–155, 1991.

[12] E. R. Jakobsen. On error bounds for approximation schemes for non-convex degenerate elliptic equations. *BIT*, 44(2) :269–285, 2004.

[13] E. R. Jakobsen. On error bounds for monotone approximation schemes for multidimensional Isaac equations. To appear, 2004.

[14] N.V. Krylov. On the rate of convergence of finite difference approximation for Bellman's equation. *St. Petersburg Math. J.*, 9 :639–650, 1997.

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

[16] H.J. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of mathematics*. Springer, New York, 2001. Second edition.

[17] N. Yamada. Viscosity solutions for a system of elliptic inequalities with bilateral obstacles. *Funkcial. Ekvac.*, 30 :417–425, 1987.

CHAPITRE II

Error estimate for a stochastic impulse control problem

Error estimates for a stochastic impulse control problem [1]

Abstract

We obtain error bounds for monotone approximation schemes of a stochastic impulse control problem. This is an extension of the theory for error estimates for the Hamilton-Jacobi-Bellman equation.

We obtain almost the same estimate on the rate of convergence as in the equation without impulsions [3], [2].

1 Introduction

The aim of this paper is to give error bounds for approximation schemes of the impulse control problem. More precisely we consider the following equation

$$\max\{\sup_{\alpha_i\in\mathcal{A}} L^{\alpha_i}(x,\mathcal{D}u); u(x)-\mathcal{M}u(x)\} = 0,\ x\in\mathbb{R}^N, \qquad (P)$$

where

$$\begin{aligned} L^{\alpha_i}(x,\mathcal{D}u(x)) &= L^{\alpha_i}(x,u(x),Du(x),D^2u(x)),\\ L^{\alpha_i}(x,r,p,X) &= -tr[a^{\alpha_i}(x)X]-b^{\alpha_i}(x)p+c^{\alpha_i}(x)r-f^{\alpha_i}(x). \end{aligned}$$

and

$$\begin{cases} \mathcal{M}u(x) := k+\inf_{\xi\in\mathbb{R}^N_+}\{u(x+\xi)+c(\xi)\},\\ k>0, \quad c:\mathbb{R}^N_+\to\mathbb{R}_+,\\ c(0)=0, \quad c(\xi_1+\xi_2)\le c(\xi_1)+c(\xi_2). \end{cases} \qquad (1.1)$$

Here $\mathcal{A}=\{\alpha_1,\dots,\alpha_M\}$ denotes the set of controls, assumed to be finite. The coefficients $(a^{\alpha_i},b^{\alpha_i},c^{\alpha_i},f^{\alpha_i})$ are, for each $\alpha_i\in\mathcal{A}$, bounded and Lipschitz functions $\mathbb{R}^N\to\mathcal{S}^N\times\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}$, where $\mathcal{S}^N$ denotes the set of $N\times N$ symmetric matrices. Under classical assumptions, (P) has a unique bounded viscosity solution, denoted u. The regularity of u depends on the properties of the coefficients a,b,c,f. We refer to [14, 15], for existence, uniqueness and regularity of u.

Then we consider monotone approximation schemes of (P), of the following form :

$$\max\{S(h,x,u_h(x),u_h); u_h(x)-\mathcal{M}u_h(x)\}=0, \quad x\in\mathbb{R}^N, \qquad (S)$$

where $S:\mathbb{R}^N_+\times\mathbb{R}^N\times\mathbb{R}\times C_b(\mathbb{R}^N)\to\mathbb{R}$ is a consistent, monotonic and uniformly continuous approximation of $\sup_{\alpha_i\in\mathcal{A}} L^{\alpha_i}(x,\mathcal{D}u(x))$ (see section 2). We will denote $u_h\in C_b(\mathbb{R}^N)$ the solution of (S), which is the approximation of u, and $h\in\mathbb{R}^N$ the mesh size. This abstract notations was introduced by Barles and Souganidis [4] to display clearly the monotonicity of

[1] Joint work with F. Bonnans and H. Zidani. To appear in Applied Math. Optimiz. Revised version accepted in March 2006.

the scheme : $S(h,x,r,v)$ is non decreasing in r and non increasing in v. Typical approximation schemes that we will consider are Classical Finite Differences [21], Generalized Finite Differences [6, 7], and Markov Chain Approximations [21].

Results on convergence rates for monotone approximation schemes of the corresponding equation without impulses are known ; i.e., for the following equation :

$$\sup_{\alpha_i \in \mathcal{A}} L^{\alpha_i}(x, \mathcal{D}u(x)) = 0, \qquad x \in \mathbb{R}^N, \tag{1.2}$$

and the related scheme

$$S(h, x, u_h(x), u_h) = 0, \qquad x \in \mathbb{R}^N. \tag{1.2 bis}$$

Error estimates for this equation have been obtained by Krylov [19, 20], and these results were extended by Barles and Jakobsen [3, 2]. Moreover, results on convergence rate for monotone approximation schemes of a particular Isaac equation have been obtained by the authors [5], and by Jakobsen [17, 16].

Using the method introduced by Ishii [15], to prove the existence of a unique viscosity solution of (P), we approach (P) by a sequence of cascade problems (P_n), $n \geq 1$,

$$\max\{\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}u(x)); u(x) - \mathcal{M}u_{n-1}(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{P_n}$$

where u_0 is solution of (1.2). Let u_n be the viscosity solution of (P_n). In the same way we approach (S) by a sequence of cascade schemes (S_n), $n \geq 1$,

$$\max\{S(h, x, u_h(x), u_h); u_h(x) - \mathcal{M}u_{h(n-1)}(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{S_n}$$

where u_{h0} is solution of (1.2 bis). Let u_{hn} denote the solution of (S_n).

Using the methods introduced by Barles and Jakobsen [2], upper and lower bounds of $u_n - u_{hn}$, for all $n < +\infty$, are obtained. The upper estimate of $u_n - u_{hn}$ is easier to obtain than the lower. The proof involves a "Krylov regularization" of (P_n), i.e. the perturbed equation

$$\max\left\{\sup_{\alpha_i, |e| \leq \epsilon} L^{\alpha_i}(x+e), \mathcal{D}u_n^\epsilon(x)); u_n^\epsilon(x) - \mathcal{M}u_{n-1}(x)\right\} = 0,$$

and its viscosity solution u_n^ϵ. A regularization of u_n^ϵ by convolution gives an approximate smooth sub-solution of (P_n), denoted $u_{n\epsilon}$, which is also an approximate sub-solution of (S_n). So, by using the consistency property, we obtain the upper bound of $u_n - u_{hn}$, after choosing an optimal parameter of regularization. Then, we consider $u - u_h$ and we do the following decomposition

$$\sup_x (u(x) - u_h(x)) \leq \sup_x (u(x) - u_n(x)) + \sup_x (u_n(x) - u_{hn}(x))$$
$$+ \sup_x (u_{hn}(x) - u_h(x)),$$

for all n in $\mathbb{N}$. Choosing the optimal n, we obtain the result. In particular we have that $n \sim |\ln h|$.

To obtain the lower estimate, we start by giving lower bound of $u_n - u_{hn}$, for $n \in \mathbb{N}$. We introduce the following switching system approximating (P_n) :

$$\max\{L^{\alpha_i}(x, \mathcal{D}v_i^n(x)); v_i^n(x) - \min_{j \neq i}\{v_j^n(x) + \ell\}; v_i^n(x) - \mathcal{M}u_{n-1}(x)\} = 0, \tag{1.3}$$

for $x \in \mathbb{R}^N$, and $i \in \mathcal{I} = \{1, \ldots, M\}$, $\ell \geq 0$. For literature on the switching systems, see [8, 11, 12, 13]. We consider the viscosity solution $v^n = (v_1^n, \ldots, v_M^n)$ of this system, and give an estimate of the rate of convergence of v^n to u_n. Then we consider a perturbed system

$$\begin{aligned} \max\{\inf_{|e| \leq \epsilon} L^{\alpha_i}(x + e, \mathcal{D}w_i^{n\epsilon}(x)); w_i^{n\epsilon}(x) - \min_{j \neq i}\{w_j^{n\epsilon}(x) + \ell\}; \\ w_i^{n\epsilon}(x) - \mathcal{M}u_{n-1}(x)\} = 0, \end{aligned} \tag{1.4}$$

for all $i \in \mathcal{I}$ and $x \in \mathbb{R}^N$, and its viscosity solution $w^{n\epsilon} = (w_1^{n\epsilon}, \ldots, w_M^{n\epsilon})$. We regularize $w^{n\epsilon}$ by convolution obtaining $w_{n\epsilon}$, and this function allows to build a local super-solution of (P_n). Then, by applying the consistency and the monotonicity of the scheme, we obtain the lower bound of $u_n - u_{hn}$. Finally, since

$$\begin{aligned} \sup_x (u_h(x) - u(x)) \leq \sup_x (u_h(x) - u_{hn}(x)) + \sup_x (u_{hn}(x) - u_n(x)) \\ + \sup_x (u_n(x) - u(x)), \end{aligned}$$

choosing the optimal n, we obtain the result. With our result, we can prove an upper bound of $|h|^{1/2}|\ln h|$ and a lower bound of $|h|^{1/5}|\ln h|$ for classical finite differences scheme and for generalized finite differences scheme. Observe that, in the case without impulsions, the results of [3, 2] give an upper bound of $|h|^{1/2}$ and a lower bound of $|h|^{1/5}$ for classical finite differences scheme and for generalized finite differences scheme. Therefore for impulse control problems we obtain very close estimates.

The paper is organized as follows : Section 2 introduces notations and states the main result. Section 3 introduces the cascade approximations of (P) and (S). Section 4 obtains upper bound of $u_n - u_{hn}$, for all $n < +\infty$, whereas Section 5 gives lower bound of $u_n - u_{hn}$, for all $n < +\infty$. Section 6 is devoted to the proof of the main theorem. Finally the Appendix gives some auxiliary theorems which are used throughout the paper.

2 Notations and main result

We start by introducing some notations used in the article. By $|\cdot|$ we mean the standard Euclidean norm in any $\mathbb{R}^M$ type space. In particular, if $X \in \mathcal{S}^N$, then $|X|^2 = tr(XX^\top)$, where $X^\top$ is the transpose of X, i.e. $|X|$ is the Frobenius norm. If g is a bounded function from $\mathbb{R}^N$ into either $\mathbb{R}, \mathbb{R}^M$, or the space of $N \times P$ matrices, we set $|g|_0 := \sup_{x\in\mathbb{R}^N} |g(x)|$. If g is also Lipschitz continuous, we set

$$[g]_1 := \sup_{\substack{x,y\in\mathbb{R}^N \\ x\neq y}} \frac{|g(x)-g(y)|}{|x-y|}, \qquad |g|_1 := |g|_0 + [g]_1.$$

We denote by $\leq$ the component wise ordering in $\mathbb{R}^N$, and by $\preceq$ the ordering in the sense of positive semidefinite matrices in $\mathcal{S}(N)$. The space $C_b(\mathbb{R}^N)$ (resp. $C_{b,l}(\mathbb{R}^N)$) will denote the space of continuous and bounded functions (resp. bounded and Lipschitz functions) from $\mathbb{R}^N$ to $\mathbb{R}$.

Given $g \in C_{b,l}(\mathbb{R}^N)^M$, $M \geq 1$, we denote by L_g an *upper bound* of the Lipschitz constant of g, $L_g \geq \max_{i=1,\dots,M}[g_i]_1$.

We will use a sequence of mollifiers $(\rho_\epsilon)_\epsilon$ defined as follows :

$$\rho_\epsilon(x) = \epsilon^{-N}\rho(x/\epsilon), \tag{2.1}$$

where $\rho \in C^\infty(\mathbb{R}^N)$, $\int_{\mathbb{R}^N}\rho = 1$, $\operatorname{supp}\{\rho\} \subseteq \bar{B}(0,1)$ and $\rho \geq 0$. We define the mollification of $g \in C_b(\mathbb{R}^N)$ as follows :

$$g_\epsilon(x) := \int_{\mathbb{R}^N} g(x-e)\rho_\epsilon(e)\mathrm{d}e. \tag{2.2}$$

If g is Lipschitz continuous, then

$$|g(x) - g_\epsilon(x)| \leq L_g\epsilon. \tag{2.3}$$

If $g \in C_b(\mathbb{R}^N)$ (resp. $C_{b,l}(\mathbb{R}^N)$), then we have

$$|D^i g_\epsilon(x)| \leq C\epsilon^{-i}|g|_0, \qquad \big(\text{resp. } C\epsilon^{1-i}|g|_1\big) \quad \forall\, i = 1,\dots,n. \tag{2.4}$$

From [15], we have the following properties on $\mathcal{M}$, defined in (1.1).

Proposition II.1. *Let $u, v : \mathbb{R}^N \to \mathbb{R}$. We have :*
(1) If $u \leq v$ in $\mathbb{R}$, then $\mathcal{M}u \leq \mathcal{M}v$ in $\mathbb{R}^N$.
(2) $\mathcal{M}(tu + (1-t)v) \geq t\mathcal{M}u + (1-t)\mathcal{M}v$; $t \in [0,1]$.
(3) $\mathcal{M}(u + c) = \mathcal{M}u + c$, for $c \in \mathbb{R}$.
(4) $|\mathcal{M}u - \mathcal{M}v|_0 \leq |u - v|_0$ for all $u, v \in C(\mathbb{R}^N)$. □

The assumption we use on equation (P) are as follows :

(A1) For all $\alpha_i \in \mathcal{A}$, we have $a^{\alpha_i} = \frac{1}{2}\sigma^{\alpha_i}\sigma^{\alpha_i T}$, where σ^{α_i} is a $N \times P$ measurable function of x. There exists a constant K such that, for all $\alpha_i \in \mathcal{A}$,

$$c^{\alpha_i} \geq 1 \quad \text{ and } \quad |\sigma^{\alpha_i}|_1 + |b^{\alpha_i}|_1 + |c^{\alpha_i}|_1 + |f^{\alpha_i}|_1 \leq K.$$

(A2) $1 > \sup_{\alpha_i}\{[\sigma^{\alpha_i}]_1^2 + [b^{\alpha}]_1\}$.

Assumption (A1) ensures that all equations we will use are well-posed ; assumption (A2) ensures that all solutions are Lipschitz and bounded. Without assumption (A2), we have that all solutions are Hölder and bounded. We conjecture that the techniques used in this paper can be extended to this case. In assumption (A1), we have assumed $c^{\alpha_i} \geq 1$ for simplicity of future computations. All our results can be extended to the general case $c^{\alpha_i} \geq \lambda$, where $\lambda > 0$. In this case, in assumption (A2) and in all estimates of Lipschitz constants obtained in the appendix, we have to write $\min(\lambda, 1)$ instead of 1.

Result of [15, Theorem 4.2] gives the existence of a viscosity solution of (P). Moreover, generalizing, in the obvious way, the proof of [1, Theorem 3.5], which involves only first order impulse control problem, we obtain the uniqueness of this viscosity solution. We can then give the following proposition.

Proposition II.2. *Under assumptions (A1-A2), (P) has a unique viscosity solution u in $C_{b,l}(\mathbb{R}^N)$. In particular we have*

$$|u|_0 \leq \sup_{\alpha_i} |f^{\alpha_i}|_0. \qquad \square$$

Let $C \geq 0$ a constant, and consider the following equation :

$$\max\{\sup_{\alpha_i} L_C^{\alpha_i}(x, \mathcal{D}u); u(x) - \mathcal{M}u(x)\} = 0, \; x \in \mathbb{R}^N, \tag{P_C}$$

where $L_C^{\alpha_i}(x, r, p, X) = L^{\alpha_i}(x, r, p, X) - Cc^{\alpha_i}(x)$. We have then the following lemma, which is given without proof.

Lemma II.3. *Under assumptions (A1-A2), u is a viscosity solution of (P) if and only if $u + C$ is a viscosity solution of (P_C).*$\square$

Remark II.4. *In the sequel we assume that $f^{\alpha_i}(x) \geq 0$, for all x and α_i, since that slightly simplifies the proofs ; however, using lemma II.3, all our results are easily extended to the case when f is not nonnegative.*

We now state assumptions on the discrete scheme (S), which is an approximation of the equation (P) :

(S1) Monotonicity : $S(h, x, r + m, u + m) \geq m + S(h, x, r, v)$
for all $h \in \mathbb{R}_+^N$, $r \in \mathbb{R}$, $m \geq 0$, $x \in \mathbb{R}^N$ and u, v in $C_b(\mathbb{R}^N)$ such that $u \leq v$ in $\mathbb{R}^N$.

(S2) Regularity : for all $h \in \mathbb{R}_+^N$ and $\phi \in C_b(\mathbb{R}^N)$, $x \mapsto S(h, x, \phi(x), \phi)$ is bounded and continuous ; $r \mapsto S(h, x, r, \phi)$ is uniformly continuous for bounded r, uniformly with respect to $x \in \mathbb{R}^N$.

(S3) There exist $n, k_i > 0$, $i \in J \subseteq \{1, \ldots, n\}$ and a constant $K_c > 0$ such that for all $h \in \mathbb{R}_+^N$ and x in $\mathbb{R}^N$, and for every smooth $\phi \in C^n(\mathbb{R}^N)$ such that $|D^i\phi|_0$ is bounded, for every $i \in J$, the following holds :

$$\left| \sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}\phi) - S(h, x, \phi(x), \phi) \right| \leq K_c Q(\phi),$$

where $Q(\phi) := \sum_{i \in J} |D^i\phi|_0 |h|^{k_i}$.

(S4) Suppose now that the scheme $S(h,x,u_h(x),u_h)$ can be written in the following form

$$\sup_{\alpha_i} S^{\alpha_i}(h,x,u_h(x),u_h), \tag{2.5}$$

as is the case for Finite Differences scheme and Generalized Finite Differences scheme.

(i) Let $C \geq 0$ a constant.
If v is solution of $\max\{\sup_{\alpha_i} S^{\alpha_i}(h,x,v(x),v); v(x) - \mathcal{M}v(x)\} = 0$, then $v + C$ is solution of $\max\{\sup_{\alpha_i}(S^{\alpha_i}(h,x,v(x),v) - Cc^{\alpha_i}(x)); v(x) - \mathcal{M}v(x)\} = 0$.

(ii) If v is solution of

$$\max\{\sup_{\alpha_i} S^{\alpha_i}(h,x,v(x),v); v(x) - \mathcal{M}v(x)\} = 0, \tag{2.6}$$

then νv is solution of

$$\max\{\sup_{\alpha_i}(S^{\alpha_i}(h,x,\nu v(x),\nu v) + (\nu-1)f^{\alpha_i}(x)); \nu v(x) - \nu\mathcal{M}v(x)\} = 0, \tag{2.7}$$

where $\nu \in (0,1)$, and f^{α_i} defined after equation (P).

Remark II.5. *Assumption (S4(i)) leads us to have* 0 *as lower bound for every solution of the cascade schemes. If we don't assume (S4(i)), we obtain a negative constant as lower bound ; all our results can be extended to this case, but computations become more complicated. Assumption (S4(ii)) is useful to prove the uniqueness of the solution of* (S).

Example II.6. *An example of a numerical scheme which satisfies these assumptions is the standard Finite Difference Scheme when* $N = 1$*, defined as :*

$$S(h,x,r,\phi) := \sup_{\alpha_i \in \mathcal{A}} \Bigg\{ -a^{\alpha_i}(x)\left[\frac{\phi(x+h) - 2r + \phi(x-h)}{h^2}\right]$$
$$-b_+^{\alpha_i}(x)\left[\frac{\phi(x+h)-r}{h}\right] + b_-^{\alpha_i}(x)\left[\frac{\phi(x-h)-r}{h}\right] + c^{\alpha_i}(x)r - f^{\alpha_i}(x)\Bigg\}, \tag{2.8}$$

where $b_+^{\alpha_i}(x) := \max(b^{\alpha_i}(x),0)$*,* $b_-^{\alpha_i}(x) := \max(-b^{\alpha_i}(x),0)$*. Clearly assumptions (S1), (S2), and (S4) are satisfied. For the consistency hypothesis (S3), we obtain, from a Taylor expansion,*

$$Q(\phi) = |D^2\phi|h + |D^4\phi|h^2, \tag{2.9}$$

i.e. $n = 4$*,* $J = \{2,4\}$*,* $k_2 = 1$ *and* $k_4 = 2$*.*

We set, J being defined in (S3) :

$$\bar{\gamma} := \min_{i\in J}\left\{\frac{k_i}{i}\right\}, \qquad \underline{\gamma} := \min_{i\in J}\left\{\frac{k_i}{3i-2}\right\}. \tag{2.10}$$

We explain briefly how we obtain our main result. In the following we build sequences P_n and S_n, $n \geq 0$, of equation of type (P_n) and (S_n) respectively, which approximate (P) and

(S). Then we have that the sequence of viscosity solutions u_n of (P_n), $n \geq 0$, converges to u, and the sequence of solution u_{hn} of (S_n), $n \geq 0$, converges to u_h. We will give these rates of convergence. Finally, for each n we give an upper and a lower bound of $u_n - u_{hn}$, and we use these bounds to obtain (2.11).

We state now our main result.

Theorem II.7. *Assume that (A1-A2) and (S1-S4) hold, and let $u \in C_{b,l}(\mathbb{R}^N)$ be the unique viscosity solution of (P). Then (S) has a unique solution $u_h \in C_b(\mathbb{R}^N)$. Moreover, for h small enough, the following two bounds hold :*

$$-C|h|^{\underline{\gamma}}|\ln h| \leq u - u_h \leq C|h|^{\bar{\gamma}}|\ln h|, \tag{2.11}$$

where C is a bounded constant, which depends on K defined in (A1), on k, and on the rates of convergence of u_n and u_{hn}.

Consider now the finite difference scheme given in example II.6. We have the following result :

Corollary II.8. *Let u the solution of (P), for $N = 1$, and let u_h the solution of (S), with S defined as in (2.8). The following two bounds hold :*

$$-C|h|^{1/5}|\ln h| \leq u - u_h \leq C|h|^{1/2}|\ln h|, \tag{2.12}$$

where C is a bounded constant, which depends on K defined in (A1), and on the rates of convergence of u_n and u_{hn}.

Proof. Applying (2.9), we obtain $\underline{\gamma} = 1/5$ and $\bar{\gamma} = 1/2$. Then we can use the precedent theorem to obtain the result. □

Remark II.9. *Corollary II.8 can be extended to the Finite Differences scheme in dimension $N > 1$ [21], and to the Generalized Finite Differences scheme in dimension $N \geq 1$ [6, 7]. The bounds that we obtain are the same as (2.12), where now h is the vector of space steps along each component of x.*

3 The cascade approximations

In this section we present the approximations of (P) and (S), and we study their main properties.

3.1 Cascade for the HJB equation

We approach equation (P) by a sequence of obstacle problems, and use the same methods as in [15, Proof of theorem 4.2], to prove that the solutions of the sequence of equations converge to the solution of (P). By remark II.4, we have that $u \equiv 0$ is a viscosity sub-solution of (P). Consider the following problem :

$$\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}u(x)) = 0, \quad x \in \mathbb{R}^N. \tag{P_0}$$

Under assumptions (A1-A2), this equation has a unique viscosity solution u_0 in $C_{b,l}(\mathbb{R}^N)$. Since $u \equiv 0$ is a viscosity sub-solution of (P_0), the comparison principle (see [15, Theorem 3.3]) implies $0 \le u_0$. Consider the following problem :

$$\max\{\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}u(x)); u(x) - \mathcal{M}u_0(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{P_1}$$

Since $\mathcal{M}u_0$ is uniformly continuous, under assumptions (A1-A2), there exists a unique viscosity solution u_1 of (P_1) in $C_{b,l}(\mathbb{R}^N)$. Similarly, for $n = 2, 3, \cdots$, let $u_n \in C_{b,l}(\mathbb{R}^N)$ be the unique viscosity solution of

$$\max\{\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}u(x)); u(x) - \mathcal{M}u_{n-1}(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{P_n}$$

It is easy to check that u_1 is a viscosity sub-solution of (P_0). By the comparison principle, $u_1 \le u_0$. Moreover, $u \equiv 0$ is a sub-solution of (P_1) in $\mathbb{R}^N$, and then $0 \le u_1 \le u_0$ in $\mathbb{R}^N$. By point (1) of Proposition II.1, $\mathcal{M}u_1 \le \mathcal{M}u_0$, then we can say that u_2 is a viscosity sub-solution of (P_1), and also $u_2 \le u_1$ in $\mathbb{R}^N$. By induction over n, we obtain :

$$0 \le \cdots \le u_n \le \cdots \le u_2 \le u_1 \le u_0. \tag{3.1}$$

We can see that, if $|u_0|_0 \le k$, then u is a viscosity solution of (P) and then we refer to [3], [2] for error estimates. Suppose now that $|u_0|_0 > k$, and let $\mu \in (0, 1)$ such that $\mu |u_0|_0 < k$.

Theorem II.10. *We have that, for all n,*

$$u_n - u_{n+1} \le (1 - \mu)^n |u_0|_0. \tag{3.2}$$

Proof : Let $n \in \mathbb{N}$, and $\theta_n \in (0, 1]$ be such that

$$u_n - u_{n+1} \le \theta_n u_n, \ \text{ in } \mathbb{R}^N. \tag{3.3}$$

By (3.1), this holds at least for $\theta_n = 1$. Rewriting (3.3) as $(1 - \theta_n)u_n \le u_{n+1}$, and using proposition II.1, get

$$(1 - \theta_n)\mathcal{M}u_n + \theta_n k \le (1 - \theta_n)\mathcal{M}u_n + \theta_n \mathcal{M}0 \le \mathcal{M}[(1 - \theta_n)u_n] \le \mathcal{M}u_{n+1}. \tag{3.4}$$

We now prove that

$$(1 - \theta_n + \mu\theta_n)u_{n+1} \le u_{n+2}, \tag{3.4a}$$

where u_{n+2} is the viscosity solution of (P_{n+2}). Since u_{n+1} is the viscosity solution of (P_{n+1}), and $f^{\alpha_i}(x) \geq 0$, for all x and for all α_i, we have that $(1-\theta_n+\mu\theta_n)u_{n+1}$ is a viscosity sub-solution of $\sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}v(x)) = 0$. Moreover, by the construction of the sequence (3.1), and by (3.4), we have

$$(1-\theta_n+\mu\theta_n)u_{n+1} \leq (1-\theta_n)u_{n+1} + \mu\theta_n|u_0|_0, \tag{3.5a}$$

$$\mathcal{M}u_{n+1} \geq (1-\theta_n)\mathcal{M}u_n + \theta_n k. \tag{3.5b}$$

Taking the difference between (3.5a) and (3.5b), and knowing that u_{n+1} is the viscosity solution of (P_n), we have

$$\begin{aligned}
&(1-\theta_n+\mu\theta_n)u_{n+1} - \mathcal{M}u_{n+1} \\
&\leq (1-\theta_n)u_{n+1} + \mu\theta_n|u_0|_0 - (1-\theta_n)\mathcal{M}u_n - \theta_n k \\
&\leq (1-\theta_n)u_{n+1} + \theta_n k - (1-\theta_n)\mathcal{M}u_n - \theta_n k \leq 0.
\end{aligned}$$

So we can say that $(1-\theta_n+\mu\theta_n)u_{n+1}$ is a viscosity sub-solution of (P_{n+2}). The comparison principle implies (3.4a), or equivalently

$$u_{n+1} - u_{n+2} \leq \theta_n(1-\mu)u_{n+1}. \tag{3.6}$$

As in [15, Proof of theorem 4.2], by the inequalities $u_0 - u_1 \leq u_0$ in $\mathbb{R}^N$, we obtain $u_1 - u_2 \leq (1-\mu)u_1$ in $\mathbb{R}^N$. Then we can take $\theta_1 = 1-\mu$ and we obtain $u_2 - u_3 \leq (1-\mu)^2 u_2$, and by induction we have

$$u_{n+1} - u_{n+2} \leq (1-\mu)^{n+1}u_{n+1} \leq (1-\mu)^{n+1}|u_0|_0. \qquad \square \tag{3.7}$$

By (3.1) and (3.2), we can find a function $u \in C(\mathbb{R}^N)$, such that $|u_n - u|_0 \to 0$, when $n \to +\infty$. Proposition II.1 and the stability of solutions imply that u is a viscosity solution of (P). Then we can say that u_n converges to u, the unique viscosity solution of (P), when $n \to +\infty$. We want to estimate an upper bound of $u_n - u$ for an arbitrary n. By (3.2) and since $(1-\mu) < 1$, we obtain that, for all $n \geq 0$,

$$u_n - u \leq \sum_{i=n}^{+\infty}(1-\mu)^i|u_0|_0 = \frac{(1-\mu)^n}{1-(1-\mu)}|u_0|_0 = \frac{(1-\mu)^n}{\mu}|u_0|_0. \tag{3.8}$$

3.2 Cascade for the numerical scheme

As we have done for the equation (P), we will approach (S) by a sequence of equations (S_n). This kind of approach has been already used for numerical study of the impulse control problem, see in particular [9].

In all the paper we suppose that every equation (S_n) has at least one solution $u_{hn} \in C_b\mathbb{R}^N$. We give now a useful lemma to obtain the uniqueness of u_{hn}, for all n. Consider the equation

$$\max\{S(h, x, \phi(x), \phi); \phi(x) - \psi(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{3.9}$$

where $\psi \in C_b(\mathbb{R}^N)$. We say that a function $\phi \in C_b(\mathbb{R}^N)$ is a sub-solution (resp. super-solution) of (3.9) if

$$\max\{S(h, x, \phi(x), \phi); \phi(x) - \psi(x)\} \le 0, \quad (\text{resp } \ge 0), \text{ for all } \quad x \in \mathbb{R}^N.$$

Lemma II.11. *Let S satisfy (S1)-(S4), and u and v respectively sub and super-solution of (3.9). Then*

$$u(x) \le v(x), \quad \textit{for all } x \in \mathbb{R}^N. \quad \square$$

The proof is a particular case of proof of Proposition II.14 where we take $g = 0$.

Let $u_{h0} \in C_b(\mathbb{R}^N)$ be a solution of

$$S(h, x, u_h(x), u_h) = 0, \quad x \in \mathbb{R}^N. \tag{S_0}$$

By Lemma II.11, u_{h0} is unique. Since $\mathcal{M}u_{h0}$ is bounded and continuous, by the same reason, there exists a unique solution $u_{h1} \in C_b(\mathbb{R}^N)$ of

$$\max\{S(h, x, u_h(x), u_h); u_h(x) - \mathcal{M}u_{h0}(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{3.10}$$

For $n = 2, 3, \cdots$, we denote u_{hn} the unique continuous and bounded solution of

$$\max\{S(h, x, u_h(x), u_h); u_h(x) - \mathcal{M}u_{h(n-1)}(x)\} = 0, \quad x \in \mathbb{R}^N. \tag{S_n}$$

The function u_{h1} is a sub-solution of (S_0), and then $u_{h1} \le u_{h0}$ in $\mathbb{R}^N$. Using remark II.4 and assumption (S4), we verify that $u_h \equiv 0$ is a sub-solution of (3.10) in $\mathbb{R}^N$, and then we have $0 \le u_{h1} \le u_{h0}$ in $\mathbb{R}^N$. Proposition II.1 implies that $0 \le \mathcal{M}u_{h1} \le \mathcal{M}u_{h0}$, then u_{h2} is a sub-solution of (3.10), and hence $u_{h2} \le u_{h1}$ in $\mathbb{R}^N$. By induction on n, we obtain

$$0 \le \cdots \le u_{hn} \le \cdots \le u_{h2} \le u_{h1} \le u_{h0}. \tag{3.11}$$

As in subsection 3.1, we suppose that $|u_0|_0 > k$. Then, since $u_{h0} \to u_0$ uniformly, we have also $|u_{h0}|_0 > k$ and we can choose $\mu \in (0, 1)$ such that $\mu|u_0|_0 < k$, and $\mu|u_{h0}|_0 < k$.

Theorem II.12. *For all n and for h small enough, we have*

$$u_{hn} - u_{h(n+1)} \le (1 - \mu)^n |u_{h0}|_0. \tag{3.12}$$

Proof : We use the same methods as in theorem II.10, taking some θ_n. The unique difference is that we have to show that $(1 - \theta_n - \mu\theta_n)u_{h(n+1)}$ is a sub-solution of (S_{n+2}), which can be written

$$\max\{S(h, x, u_h(x), u_h); u_h(x) - \mathcal{M}u_{h(n+1)}(x)\} = 0, \quad x \in \mathbb{R}^N.$$

With the monotonicity of S, we obtain the result. $\square$

Proposition II.13. *Under assumptions (S1)-(S4), there exists a solution of the equation (S).*

Proof : By (3.11) and (3.12), we can find a function $u_h \in C_b(\mathbb{R}^N)$, such that $|u_{hn} - u_h|_0 \to +\infty$, when $n \to +\infty$. Proposition II.1 and the stability of solutions implies that u_h is a solution of (S). □

We can prove a comparison principle for (S), with S written as in (2.5), and hence the uniqueness of its solution.

Proposition II.14. *Let S satisfy (S1)-(S4). Let u, v be the solutions of*

$$\max\{\sup_{\alpha_i} S^{\alpha_i}(h, x, u(x), u); u(x) - \psi_1(x)\} = 0, \; x \in \mathbb{R}^N, \tag{3.13}$$

and

$$\max\{\sup_{\alpha_i}(S^{\alpha_i}(h, x, v(x), v) + g(x)); v(x) - \psi_2(x)\} = 0, \; x \in \mathbb{R}^N, \tag{3.14}$$

where ψ_1, ψ_2 and g are elements of $C_b(\mathbb{R}^N)$. Then,

$$|u - v|_0 \leq \max\{|g|_0; |\psi_1 - \psi_2|_0\}. \tag{3.15}$$

Proof. Since u and v are solutions of (3.13) and (3.14) respectively, we have that

$$\max\{\sup_{\alpha_i} S^{\alpha_i}(h, x, u(x), u); u(x) - \psi_1(x)\} \leq 0,$$

$$\max\{\sup_{\alpha_i}(S^{\alpha_i}(h, x, v(x), v) + g(x)); v(x) - \psi_2(x)\} \geq 0,$$

for all x in $\mathbb{R}^N$. Since $\max\{A; B\} - \max\{C; D\} \leq \max\{A - C; B - D\}$, (3.13) and (3.14) imply

$$0 \leq \max\{\sup_{\alpha_i}(S^{\alpha_i}(h, x, v(x), v) + g(x)) - \sup_{\alpha_i} S(h, x, u(x), u);$$
$$v(x) - \psi_2(x) - (u(x) - \psi_1(x))\}.$$

Hence we have the two following cases.
a) $u(x) - v(x) \leq \psi_1(x) - \psi_2(x)$, which implies $u(x) - v(x) \leq |\psi_1 - \psi_2|_0$.
b) $\sup_{\alpha_i} S^{\alpha_i}(h, x, u(x), u) \leq 0$, and $\sup_{\alpha_i}(S^{\alpha_i}(h, x, v(x), v) + g(x)) \geq 0$.
Then $\sup_{\alpha_i} S^{\alpha_i}(h, x, v(x), v) + |g|_0 \geq 0$, and applying the monotonicity, $\sup_{\alpha_i} S(h, x, v(x) + |g|_0, v + |g|_0) \geq 0$. By [3, Theorem 2.1], obtain $u(x) - v(x) \leq |g|_0$.

Combining the two cases we have

$$\sup_x (u(x) - v(x)) \leq \max\{\sup_x g(x); \sup_x |\psi_1(x) - \psi_2(x)|\}.$$

On the other hand, we have

$$\max\{\sup_{\alpha_i} S^{\alpha_i}(h, x, u(x), u); u(x) - \psi_1(x)\} \geq 0,$$

$$\max\{\sup_{\alpha_i}(S^{\alpha_i}(h, x, v(x), v) + g(x)); v(x) - \psi_2(x)\} \leq 0,$$

for all x in $\mathbb{R}^N$. We have the two following cases.
a) $v(x) - u(x) \leq \psi_2(x) - \psi_1(x)$, which implies $u(x) - v(x) \geq -|\psi_1 - \psi_2|_0$.
b) $\sup_{\alpha_i} S^{\alpha_i}(h, x, u(x), u) \geq 0$, and $\sup_{\alpha_i}(S^{\alpha_i}(h, x, v(x), v) + g(x)) \leq 0$. Since g is a positive function, we have $\sup_{\alpha_i} S(h, x, v(x), v) \leq 0$, and by [3, Theorem 2.1], $u(x) - v(x) \geq 0$, which implies $u(x) - v(x) \geq -|g|_0$. □

We can give now the uniqueness result.

Proposition II.15. *Under assumptions (S1)-(S4), (S) has a unique solution* $u_h \in C_b(\mathbb{R}^N)$.

Proof. Let u_h and v_h be solutions of (S). By (S5), νu_h is a solution of

$$\max\{\sup_{\alpha_i}(S^{\alpha_i}(h, x, \nu u_h(x), u_h) + (\nu - 1)f^{\alpha_i}(x)); \nu u_h(x) - \nu \mathcal{M} u_h(x)\} = 0, \ x \in \mathbb{R}^N,$$

for $\nu \in (0, 1)$. Apply Proposition II.14 to obtain

$$|\nu u_h - v_h|_0 \leq \max\{|(\nu - 1)f^{\alpha_i}|_0; |\nu \mathcal{M} u_h - \mathcal{M} v_h|_0\}.$$

By [1, Theorem 3.5], we know that $|\nu \mathcal{M} u_h - \mathcal{M} v_h|_0 < |\nu u_h - v_h|_0$, and hence $|\nu u_h - v_h|_0 \leq |(\nu - 1)f|_0$. Letting ν go to 1, we have the result. □

We have proved that u_{hn} converges to the solution u_h of (S), for $n \to +\infty$, and by Proposition II.15 this solution is unique. Moreover we have

$$u_{hn} - u_h \leq \sum_{i=n}^{+\infty} (1 - \mu)^i |u_{h0}|_0 = \frac{(1 - \mu)^n}{\mu} |u_{h0}|_0. \tag{3.16}$$

4 The upper bound for the cascade problems

In this section we will use the methods of [3], [2], to obtain an upper bound of $u_n - u_{hn}$, for all n. We start with the case $n = 0$, and then we will study the general case $n \geq 1$. Finally, we will use these estimates to obtain the upper bound of $u - u_h$.

4.1 Problem without impulses

Consider the problem (P_0) and its viscosity solution $u_0 \in C_{b,l}(\mathbb{R}^N)$. Let

$$L_{u_0} := \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_0|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1}.$$

We recall here the result of [18, Lemma A.1].

Lemma II.16. L_{u_0} *is an upper bound of the Lipschitz constant of* u_0. □

Consider the scheme (S_0) and its solution $u_{h0} \in C_b(\mathbb{R}^N)$. We recall that L^{α_i} and S satisfy assumptions (A1-A2) and (S1-S4). An upper bound of $u_0 - u_{h0}$ has been obtained in [3]. Here we need to rewrite some ideas of this paper, in order to detail constants which appear in various proofs. The auxiliary equation (see [19])

$$\sup_{\alpha_i \in \mathcal{A}, |e| \leq \epsilon} L^{\alpha_i}(x+e, \mathcal{D}u_0^\epsilon(x)) = 0, \qquad x \in \mathbb{R}^N, \tag{P_0P}$$

has a unique viscosity solution $u_0^\epsilon \in C_{b,l}(\mathbb{R}^N)$. Let $u_{0\epsilon}$ the mollification of u_0^ϵ, defined as in (2.2). We give now a lemma useful in the sequel. We recall that $\bar{\gamma}$ is defined in (2.10).

Lemma II.17. *Let $g \in C_{b,l}(\mathbb{R}^N)$, and its mollification g_ϵ. Set $\epsilon = |h|^{\bar{\gamma}}$. Then, J being defined in (S3),*

$$Q(g_\epsilon) \leq |J| K_c |g|_1 |h|^{\bar{\gamma}}. \tag{4.1}$$

Proof : Using (2.4), get

$$Q(g_\epsilon) = K_c |g|_1 \sum_{i \in J} \epsilon^{1-i} |h|^{k_i} = K_c |g|_1 \sum_{i \in J} |h|^{\bar{\gamma}(1-i)+k_i}.$$

Since, by (2.10) $\bar{\gamma}(1-i) + k_i \geq \bar{\gamma}$, for all $i \in J$, we obtain the result. □

We recall here the result of [2, Proposition 3.2], where we detail some constants.

Proposition II.18. *Let $u_0 \in C_{b,l}(\mathbb{R}^N)$ the viscosity solution of (P_0), and $u_{h0} \in C_{b,l}(\mathbb{R}^N)$ the solution of (S_0). Then we have*

$$u_0(x) - u_{h0}(x) \leq \bar{C}_0 |h|^{\bar{\gamma}}, \quad \forall\, x \in \mathbb{R}^N, \tag{$\bar{E}_0$}$$

$$\bar{C}_0 := |J| K_c |u_0^\epsilon|_1 + R, \tag{4.2}$$

where R depends only on the constant K of assumption (A1).

Proof : In [3] the authors verify that $u_{0\epsilon}$ is a classical sub-solution of (P_0). By the consistency hypothesis (S3), (2.4) and lemma II.17, for $\epsilon = |h|^{\bar{\gamma}}$,

$$S(h, x, u_{0\epsilon}(x), u_{0\epsilon}) \leq Q(u_{0\epsilon}) \leq |J| K_c |u_0^\epsilon|_1 |h|^{\bar{\gamma}}, \quad x \in \mathbb{R}^N.$$

Monotonicity implies that $u_{0\epsilon} - |J| K_c |u_0^\epsilon|_1 |h|^{\bar{\gamma}} \leq u_{h0}$. By [3, Lemma A.1], we have that $|u_0 - u_{0\epsilon}| \leq R\epsilon$, where R depends only on K defined in (A1). So we have the result. □

4.2 Problem with n impulses, $n \geq 1$

Consider now the problem with n impulses (P_n), for $n \geq 1$, and its viscosity solution $u_n \in C_{b,l}(\mathbb{R}^N)$. We generalize here the method of [3], by introducing the perturbed equation

$$\max \left\{ \sup_{\alpha_i, |e| \leq \epsilon} L^{\alpha_i}(x+e), \mathcal{D}u_n^\epsilon(x)); u_n^\epsilon(x) - \mathcal{M}u_{n-1}(x) \right\} = 0, \tag{P_nP}$$

whose unique viscosity solution in $C_{b,l}(\mathbb{R}^N)$ is denoted u_n^ϵ. We recall that, for the problem without impulses, u_0^ϵ is the solution of (P_0P). The next result, proved in the appendix, gives upper bounds of Lipschitz constants of u_n and u_n^ϵ.

Lemma II.19. *Let u_n and u_n^ϵ denote the viscosity solutions of (P_n) and (P_nP) respectively, for $n \geq 1$. Then, upper bounds of Lipschitz constants of u_n and u_n^ϵ are*

$$L_{u_n} = L_{u_0}, \tag{4.3}$$

$$L_{u_n^\epsilon} = \max\big(L_{u_0}; \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_n^\epsilon|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1}\big). \qquad \Box \tag{4.4}$$

Using the same methods as for sequence (3.1), we can show that

$$0 \leq \cdots \leq u_n^\epsilon \leq \cdots \leq u_2^\epsilon \leq u_1^\epsilon. \tag{4.5}$$

Combining with (4.4), get

$$0 \leq \cdots \leq L_{u_n^\epsilon} \leq \cdots \leq L_{u_2^\epsilon} \leq L_{u_1^\epsilon}. \tag{4.6}$$

The following result is proved in the appendix.

Proposition II.20. *Let u_n and u_n^ϵ be the viscosity solutions of (P_n) and (P_nP) respectively, and A_{u_n, u_n^ϵ} be defined in (8.1). Then*

$$|u_n - u_n^\epsilon|_0 \leq A_{u_n, u_n^\epsilon} \epsilon.$$

Relations (4.6), (8.1), (3.1) and (4.4) imply the following result.

Lemma II.21. $0 \leq \cdots \leq A_{u_n, u_n^\epsilon} \leq \cdots \leq A_{u_2, u_2^\epsilon} \leq A_{u_1, u_1^\epsilon}$. $\Box$

Proof : This follows from the expression of coefficients A_{u_i, u_i^ϵ}, $i = 1, \ldots, n$, given in (8.1), combined with lemma II.19 and relation (4.6). $\Box$

We can give now the error estimate of the upper bound. We recall that $\bar{C}_0$ was defined in (4.2).

Proposition II.22. *Let $u_n \in C_{b,l}(\mathbb{R}^N)$ be the unique viscosity solution of (P_n), and $u_{hn} \in C_b(\mathbb{R}^N)$ the unique solution of (S_n), $n \geq 1$. Then we have*

$$u_n(x) - u_{hn}(x) \leq \bar{C}_n |h|^{\bar{\gamma}}, \tag{$\bar{E}_n$}$$

$$\bar{C}_n = \bar{C}_{n-1} + A_{u_n, u_n^\epsilon} + L_{u_n^\epsilon} + L_{u_0}. \tag{4.7}$$

Proof : For all $n \in \mathbb{N}$ and $\epsilon > 0$, we denote by $u_{n\epsilon}$ the mollification of u_n^ϵ. We prove the proposition by induction over n. Take $n = 1$. We show that $u_{1\epsilon} - \bar{C}_0 |h|^{\bar{\gamma}} - L_{u_0}\epsilon$ is a sub-solution of (3.10). Applying the classical methods (see [3, 2, 5]), we have that $u_{1\epsilon} - L_{u_0}\epsilon$ is a classical sub-solution of (P_1). Using the consistency hypothesis (S3), proposition II.18, the equality $Q(u_{1\epsilon} - L_{u_0}\epsilon) = Q(u_{1\epsilon})$, and the monotonicity of S, obtain

$$\begin{cases} S(h, x, u_{1\epsilon}(x) - L_{u_0}\epsilon - Q(u_{1\epsilon}), u_{1\epsilon} - L_{u_0}\epsilon - Q(u_{1\epsilon})) \leq 0 \\ u_{1\epsilon}(x) - L_{u_0}\epsilon - \bar{C}_0 |h|^{\bar{\gamma}} \leq \mathcal{M} u_{h0}(x). \end{cases}$$

We deduce that $u_{1\epsilon}(x) - L_{u_0}\epsilon - \max\{\bar{C}_0|h|^{\bar{\gamma}}, Q(u_{1\epsilon})\}$ is sub-solution of (S_1). We now set $\epsilon = |h|^{\bar{\gamma}}$. By lemma II.17, and by (4.5), (4.6) and (4.2), we obtain,

$$Q(u_1^\epsilon) \leq |J|K_c|u_1^\epsilon|_1|h|^{\bar{\gamma}} \leq |J|K_c|u_0^\epsilon|_1|h|^{\bar{\gamma}} \leq \bar{C}_0|h|^{\bar{\gamma}}.$$

Then $\max\{\bar{C}_0|h|^{\bar{\gamma}}, Q(u_{1\epsilon})\} = \bar{C}_0|h|^{\bar{\gamma}}$, which implies $u_{1\epsilon}(x) - \bar{C}_0|h|^{\bar{\gamma}} - L_{u_0}\epsilon \leq u_{h1}(x)$, for all x. Hence, with (2.3) and proposition II.20,

$$\begin{aligned} u_1(x) - u_{h1}(x) &= u_1(x) - u_1^\epsilon(x) + u_1^\epsilon(x) - u_{1\epsilon}(x) + u_{1\epsilon}(x) - u_{h1}(x) \\ &\leq A_{u_1,u_1^\epsilon}\epsilon + L_{u_1^\epsilon}\epsilon + L_{u_0}\epsilon + \bar{C}_0|h|^{\bar{\gamma}} \\ &= (A_{u_1,u_1^\epsilon} + L_{u_1^\epsilon} + L_{u_0} + \bar{C}_0)|h|^{\bar{\gamma}}. \end{aligned}$$

We obtain that (4.7) holds for $n = 1$.

Now we suppose the proposition true for $n-1$. The same methods as before, the assumption of induction and lemma II.19 give us the result. □

So we have obtained that, for all $n \geq 1$, $u_n - u_{hn} \leq \bar{C}_n|h|^{\bar{\gamma}}$. We set

$$\bar{D}_{n-1} := \bar{C}_n - \bar{C}_{n-1} = A_{u_n,u_n^\epsilon} + L_{u_n^\epsilon} + L_{u_0}.$$

Lemma II.21 and relation (4.6) imply that $\bar{D}_n \leq \bar{D}_0$, and hence, by (4.7) :

$$\bar{C}_n \leq \bar{C}_0 + n\bar{D}_0. \tag{4.8}$$

5 The lower bound for the cascade problems

In this section we will use the methods of [3, 2], to obtain a lower bound of $u_n - u_{hn}$, for all n. We start with the case $n = 0$, and then we will study the general case $n \geq 1$. Finally, we will use these estimates to obtain the lower bound of $u - u_h$.

5.1 Problem without impulses

Consider problem (P_0) of solution $u_0 \in C_{b,l}(\mathbb{R}^N)$, and the scheme (S_0) of solution $u_{h0} \in C_b(\mathbb{R}^N)$. We recall that L^{α_i} and S satisfy assumptions (A1-A2) and (S1-S4). A lower bound of $u_0 - u_{h0}$ has been obtained in [3]. Here we need to rewrite some parts of this paper, in order to give explicit bounds of constants appearing in various proofs. Consider the following switching system, which approaches (P_0),

$$\max\{L^{\alpha_i}(x, \mathcal{D}v_i^0(x)); v_i^0(x) - \min_{j\neq i}\{v_j^0(x) + \ell\}\} = 0, \tag{SS_0}$$

for $x \in \mathbb{R}^N$, $i \in \mathcal{I} = \{1, \dots, M\}$, and $\ell > 0$. Let $v^0 = (v_1^0, \dots, v_M^0)$ be the unique viscosity solution of (SS_0), $v^0 \in C_{b,l}(\mathbb{R}^N)^M$. By remark II.4, we have that $(0, \dots, 0)$ is a viscosity sub-solution of (SS_0), hence $0 \leq v_i^0(x)$, for all $i \in \mathcal{I}$ and $x \in \mathbb{R}^N$.

For every i, v_i^0 converges to u_0, when $\ell \to 0$. We rewrite here the result of [2, Theorem 2.3], which give this rate of convergence.

Lemma II.23. *Let u_0 and v^0 be the viscosity solutions of (P_0) and (SS_0) respectively. Then, for all i, we have*

$$0 \leq v_i^0 - u_0 \leq C\ell^{1/3}, \tag{H_0}$$

where C depends only on K, defined in (A1). □

Error Estimate

Consider the following perturbation of the switching system (SS_0) :

$$\max\left\{ \inf_{|e|\leq\epsilon} L^{\alpha_i}(x+e, \mathcal{D}w_i^{\epsilon 0}(x)); w_i^{\epsilon 0}(x) - \min_{j\neq i}\{w_j^{\epsilon 0}(x) + \ell\} \right\} = 0. \tag{SS_0P}$$

We denote by $w^{0\epsilon} = (w_1^{0\epsilon}, \ldots, w_M^{0\epsilon})$ the unique viscosity solution of (SS_0P) in $C_{b,l}(\mathbb{R}^N)^M$. We have $0 \leq w_i^{0\epsilon}(x)$, for all i and for all x.

The following result is proved in the appendix.

Lemma II.24. *Let v^0 and $w^{0\epsilon}$ be the viscosity solutions of (SS_0) and (SS_0P) respectively. Then, $\max_i |v_i^0 - w_i^{0\epsilon}|_0 \leq \epsilon A_{v^0,w^{0\epsilon}}$, where $A_{v^0,w^{0\epsilon}}$ is defined in (8.1).* □

Consider $\underline{\gamma}$ defined in (2.10). We have the following result.

Lemma II.25. *Given $g \in C_{b,l}(\mathbb{R}^N)$, its mollification g_ϵ, and $\epsilon = |h|^{3\underline{\gamma}}$, we have that, for J defined in (S3),*

$$Q(g_\epsilon) \leq |J|K_c|g|_1|h|^{\underline{\gamma}}. \tag{5.1}$$

Proof : By (2.4), we know that

$$Q(g_\epsilon) = K_c|g|_1 \sum_{i\in J} \epsilon^{1-i}|h|^{k_i} = K_c|g|_1 \sum_{i\in J} |h|^{3(1-i)\underline{\gamma}+k_i}.$$

Since $3(1-i)\underline{\gamma} + k_i \geq \underline{\gamma}$, for all $i \in J$, we obtain the result. □

We recall here [2, Lemma 3.4], which gives some auxiliary results to obtain the error estimate.

Lemma II.26. *Assume (A1), (A2), let $w_{\epsilon i}^0 = \rho_\epsilon * w_i^{0\epsilon}$, for $i \in \mathcal{I}$. Moreover assume that $\epsilon \leq (4\sup_i[w_i^{0\epsilon}]_1)^{-1}\ell$. Then, for every $x \in \mathbb{R}^N$, if $j = argmin_{i\in\mathcal{I}} w_{\epsilon i}^0(x)$, then*

$$L^{\alpha_j}(x, w_{\epsilon j}^0(x), Dw_{\epsilon j}^0(x), D^2 w_{\epsilon j}^0(x)) \geq 0. \quad \square$$

We recall now the result of [2, Theorem 3.5], where we detail some constants.

Proposition II.27. *Let $u_0 \in C_{b,l}(\mathbb{R}^N)$ be the viscosity solution of (P_0) and $u_{h0} \in C_{b,l}(\mathbb{R}^N)$ the solution of (S_0). Then, we have*

$$u_{h0}(x) - u_0(x) \leq \underline{C}_0|h|^{\underline{\gamma}}, \quad \forall\, x \in \mathbb{R}^N, \tag{$\underline{E}_0$}$$

$$\underline{C}_0 = |J|K_c|w^{0\epsilon}|_1 + R, \tag{5.2}$$

where R depends only on K defined in (A1), and J is defined in (S3).

Proof : We recall the ideas of [2, Theorem 3.5]. We set

$$m := \sup_{y\in\mathbb{R}^N} \{u_{h0}(y) - g_0(y)\},$$

where $g_0 = \min_{i\in\mathcal{I}} w^0_{\epsilon i}$. We now set $\epsilon = |h|^{3\underline{\gamma}}$. Computations of [2, Theorem 3.5], combined with lemma II.25 and lemma II.26, give

$$m \leq |J| K_c |w_i^{\epsilon 0}|_1 |h|^{\underline{\gamma}}, \tag{5.3}$$

where J is defined in (S3). Applying lemma II.23, lemma II.24, and (5.3), we have that, for all $i \in \mathcal{I}$,

$$\begin{aligned}
\sup_x (u_{h0}(x) - u_0(x)) &\leq m + \sup_x (w^0_{\epsilon i}(x) - w_i^{0\epsilon}(x)) + \sup_x (w_i^{0\epsilon}(x) - v_i^0(x)) \\
&\quad + \sup_x (v_i^0(x) - u_0(x)) \\
&\leq |J| K_c |w_i^{\epsilon 0}|_1 \sum_{i\in J} \epsilon^{1-i} |h|^{k_i} + C\epsilon + A_{v^0, w^{0\epsilon}}\epsilon + C\ell^{1/3},
\end{aligned}$$

where C depends only on K defined in (A1). In agreement with lemma II.26, $\ell = 4\epsilon L_{w^{0\epsilon}} = 4|h|^{3\underline{\gamma}} L_{w^{0\epsilon}}$, where $L_{w^{0\epsilon}}$ is an upper bound of the Lipschitz constant of $w^{0\epsilon}$. By lemma II.25, we have

$$\sup_x (u_{h0}(x) - u_0(x)) \leq R_0 [2|h|^{3\underline{\gamma}} + |h|^{\underline{\gamma}}] + |J| K_c |w_0^{0\epsilon}|_1 |h|^{\underline{\gamma}},$$

where R_0 depends only on K defined in (A1). Setting $R = 3R_0$, we obtain the result. □

5.2 Problem with n impulses, $n \geq 1$

We generalize here the methods of [3]. Consider problem (P_n) and its solution $u_n \in C_{b,l}(\mathbb{R}^N)$, defined in section 3.1. We know that L_{u^0} is an upper bound of the Lipschitz constant of u_n, for all n.

Then consider the scheme (S_n) of solution $u_{hn} \in C_b(\mathbb{R}^N)$, defined in section 3.2. We recall that L^{α_i} and S satisfy assumptions (A1-A2), (S1-S4). Consider the following switching system which approach (P_n) :

$$\max\{L^{\alpha_i}(x, \mathcal{D}v_i^n(x)); v_i^n(x) - \min_{j\neq i}\{v_j^n(x) + \ell\}; v_i^n(x) - \mathcal{M}u_{n-1}(x)\} = 0, \tag{SS_n}$$

for $x \in \mathbb{R}^N$ and $i \in \mathcal{I} = \{1, \dots, M\}$. Under assumptions (A1-A2), (SS_n) has a unique viscosity solution $v^n = (v_1^n, \dots, v_M^n) \in C_{b,l}(\mathbb{R}^N)^M$. By remark II.4, it is easy to see that $(0, \dots, 0)$ is a viscosity sub-solution of (SS_n), and that v^n is a viscosity sub-solution of ($SS_{(n-1)}$), for all n. We can build, then the following sequence

$$0 \leq \cdots \leq v_i^n(x) \leq \cdots \leq v_i^1(x) \leq v_i^0(x),$$

for all i, and for all x.

Convergence of the switching system

Using the same methods as in [2, Theorem 2.3], we introduce an auxiliary switching system

$$\max\{\sup_{|e|\leq\epsilon} L^{\alpha_i}(x+e,\mathcal{D}v_i^{n\epsilon}(x)); v_i^{n\epsilon}(x)-\min_{j\neq i}\{v_j^{n\epsilon}(x)+\ell\}; v_i^{n\epsilon}(x)-\mathcal{M}u_{n-1}(x)\}, \tag{5.4}$$

and denote by $v^{n\epsilon}=(v_1^{n\epsilon},\ldots,v_M^{n\epsilon})$ its viscosity solution in $C_{b,l}(\mathbb{R}^N)^M$. As before, we have that $n\mapsto v_i^{n\epsilon}(x)$ is non-increasing, for all i and for all x. Let $v_{\epsilon i}^n=\rho_\epsilon * v_{\epsilon i}^n$, for all $i\in\mathcal{I}$.

We can give now the following result about the convergence.

Proposition II.28. *Let u_n and v^n be the solutions of (P_n) and (SS_n) respectively. Then, for all i, we have*

$$0\leq v_i^n-u_n\leq H_{v^n,v^{n\epsilon}}\ell^{1/3}, \tag{H_n}$$

where $H_{v^n,v^{n\epsilon}}$ is defined in (8.1).

Proof : We start by giving the proof for $n=1$. Consider $w=(u_1,\ldots,u_1)$ (a vector with M components equal to u_1). Then, for every i, we have :

$$\begin{cases} L^{\alpha_i}(x,\mathcal{D}u_1(x))\leq 0 \\ u_1(x)\leq u_1(x)+\ell \\ u_1(x)\leq \mathcal{M}u_0(x)\end{cases} \Rightarrow u_1 \text{ is a sub-solution of (SS1) } \Rightarrow u_1(x)\leq v_i^1(x),$$

for all $x\in\mathbb{R}^N$, $i\in\mathcal{I}$. We show that, for all i, $v_{i\epsilon}^1-C\ell\epsilon^{-2}-L_{u_0}\epsilon$ is a sub-solution of (P_1), where

$$C=C_\rho\ell\sup_{\alpha_i}(|\sigma^{\alpha_i}|_0+|b^{\alpha_i}|_0+|c^{\alpha_i}|_0). \tag{5.5}$$

With classical methods (see [3], [2], [5]), we have that $v_{i\epsilon}^1$ is, for all i, a sub-solution, in the classical sense, of

$$L^{\alpha_i}(x,\mathcal{D}v(x))=0,\ \forall\, x\in\mathbb{R}^N. \tag{5.6}$$

The definition of switching system implies that $|v_i^{1\epsilon}-v_j^{1\epsilon}|\leq\ell$, for all i,j. Combining with (2.4), we obtain

$$|L^{\alpha_i}(x,\mathcal{D}v_{\epsilon j}^1(x))-L^{\alpha_i}(x,\mathcal{D}v_{\epsilon i}^1(x))|\leq\frac{C\ell}{\epsilon^2},\forall\, i,j\in\mathcal{I},\text{ et }\forall\, x\in\mathbb{R}^N.$$

Since $v_{\epsilon i}^1$ is a sub-solution of (5.6), this implies

$$L^{\alpha_i}(x,\mathcal{D}v_{\epsilon j}^1(x))\leq\frac{C\ell}{\epsilon^2},\quad \forall i,j,\text{ and }\forall\, x\in\mathbb{R}^N. \tag{5.6a}$$

Consequently $v_{\epsilon i}^1-C\ell\epsilon^{-2}$ is a classical sub-solution of $\sup_{\alpha_i}L^{\alpha_i}(x,\mathcal{D}w(x))=0$. Moreover, by the definition of the auxiliary system, we have that $v_i^{1\epsilon}(x)-\mathcal{M}u_0(x)\leq 0$, for all $i\in\mathcal{I}$, and for all $x \in \mathbb{R}^N$. Let $u_{\epsilon 0}$ be the mollification of u_0, defined as in (2.2). Then, we have $v^1_{\epsilon i}(x) - \mathcal{M}u_{0\epsilon}(x) \leq 0$, which implies $v^1_{\epsilon i}(x) - \mathcal{M}u_0(x) \leq L_{u_0}\epsilon$, and also

$$v^1_{\epsilon i}(x) - L_{u_0}\epsilon - C\ell\epsilon^{-2} - \mathcal{M}u_0(x) \leq 0, \; \forall\, x \in \mathbb{R}^N.$$

Hence, for all $x \in \mathbb{R}^N$, we have

$$\begin{cases} \sup_{\alpha_i} L^{\alpha_i}(x, \mathcal{D}(v^1_{\epsilon i} - C\ell\epsilon^{-2})(x)) \leq L_{u_0}\epsilon, \\ v^1_{\epsilon i}(x) - \mathcal{M}u_0(x) \leq L_{u_0}\epsilon + C\ell\epsilon^{-2}. \end{cases}$$

So $v^1_{\epsilon i} - L_{u_0}\epsilon - C\ell\epsilon^{-2}$ is a viscosity sub-solution of (P1), and we have $v^1_{\epsilon i}(x) - L_{u_0}\epsilon - C\ell\epsilon^{-2} \leq u_1(x)$, for all $x \in \mathbb{R}^N$. Finally we obtain

$$v^1_i(x) - u_1(x) \leq \frac{C_\rho \ell}{\epsilon^2} \sup_{\alpha_i}(|\sigma^{\alpha_i}|_0 + |b^{\alpha_i}|_0 + |c^{\alpha_i}|_0) + (L_{u_0} + L_{v^{1\epsilon}} + A_{v^1, v^{1\epsilon}})\epsilon,$$

for all x in $\mathbb{R}^N$. Minimizing with respect to ϵ, obtain

$$v^1_i(x) - u_1(x) \leq H_{v^1, v^{1\epsilon}} \ell^{1/3}.$$

The result for $n > 1$ can be proved similarly, using $L_{u_{n-1}} = L_{u_0}$ as an upper bound of the Lipschitz constant of u_{n-1}. □

Error Estimates

Consider the following perturbed switching system which approaches (P_n),

$$\max\{\inf_{|e| \leq \epsilon} L^{\alpha_i}(x + e, \mathcal{D}w^{n\epsilon}_i(x)); w^{n\epsilon}_i(x) - \min_{j \neq i}\{w^{n\epsilon}_j(x) + \ell\};$$
$$w^{n\epsilon}_i(x) - \mathcal{M}u_{n-1}(x)\} = 0, \qquad (SS_nP)$$

and its unique viscosity solution $w^{n\epsilon} \in C_{b,l}(\mathbb{R}^N)^M$. As before, we can prove that $0 \leq \cdots \leq w^{n\epsilon}_i(x) \leq \cdots \leq w^{1\epsilon}_i(x) \leq w^{0\epsilon}_i(x)$, for all i and x. Let $g^n := v^n,\ v^{n\epsilon},\ w^{n\epsilon}$. Then, we set

$$L_{g^n} := \max\left(\sup_i \frac{[c^{\alpha_i}]_1 |g^n_i|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1}; L_{u_0} \right). \qquad (5.7)$$

We have the following results, which are showed in the appendix.

Lemma II.29. *Let $g^n := v^n,\ v^{n\epsilon},\ w^{n\epsilon}$. Then $\max_i [g^n_i]_1 \leq L_{g^n}$.* □

Lemma II.30. *Let v^n, $v^{n\epsilon}$ and $w^{n\epsilon}$ be the viscosity solutions of (SS_n), (5.4) and (SS_nP) respectively. Then, we have*

$$\max_i |v^n_i - v^{n\epsilon}_i|_0 \leq A_{v^n, v^{n\epsilon}}\epsilon, \quad \max_i |v^n_i - w^{n\epsilon}_i|_0 \leq A_{v^n, w^{n\epsilon}}\epsilon,$$

where $A_{v^n, v^{n\epsilon}}$ and $A_{v^n, w^{n\epsilon}}$ are defined in (8.1). □

The following result in proved in theorems II.34 and II.37.

Lemma II.31. *Let $g^i := v^i,\ v^{i\epsilon},\ w^{i\epsilon}$, and let L_{g^i} be defined as in (5.7). Then*

$$L_{g^n} \leq \cdots \leq L_{g^2} \leq L_{g^1},$$

$$A_{v^n,w^{n\epsilon}} \leq \cdots \leq A_{v^1,w^{1\epsilon}} \leq A_{v^0,w^{0\epsilon}},$$

$$A_{v^n,v^{n\epsilon}} \leq \cdots \leq A_{v^1,v^{1\epsilon}} \leq A_{v^0,v^{0\epsilon}}. \qquad \square.$$

We can give now the lower bound.

Proposition II.32. *Let $u_n \in C_{b,l}(\mathbb{R}^N)$ the viscosity solution of (P_n) and let $u_{hn} \in C_{b,l}(\mathbb{R}^N)$ the solution of (S_n), $n \geq 1$. Then we have*

$$u_{hn}(x) - u_n(x) \leq \underline{C}_n|h|^{\underline{\gamma}}, \quad \forall\, x \in \mathbb{R}^N, \tag{$\underline{E}_n$}$$

$$\underline{C}_n = \underline{C}_{n-1} + 12L_{w^{n\epsilon}} + 4L_{u_0} + A_{v^n,w^{n\epsilon}} + H_{v^n,v^{n\epsilon}}(6L_{w^{n\epsilon}})^{1/3}. \tag{5.8}$$

Proof : The proof is by induction over n. Let $n = 1$, and let

$$m := \sup_{y\in\mathbb{R}^N}\{u_{h1}(y) - g(y)\}, \tag{5.9}$$

where $g = \min_{i\in\mathcal{I}} w^1_{\epsilon i}$. For $\eta \geq 0$, let

$$m_\eta := \sup_{y\in\mathbb{R}^N}\{u_{h1}(y) - g(y) - \eta\phi(y)\},$$

where $\eta > 0$ is a small constant, and $\phi(x) = (1 + |x|^2)^{1/2}$. Let x_0 be such that $m_\eta = u_{h1}(x_0) - g(x_0) - \eta\phi(x_0)$. Then we have also $m_\eta = u_{h1}(x_0) - w^1_{\epsilon i_0}(x_0) - \eta\phi(x_0)$, where $w^1_{\epsilon i_0}(x_0) = \min_{j\in\mathcal{I}} w^1_{\epsilon j}(x_0)$. After some computations (see [2, Theorem 3.4]), we can say that, if $\epsilon \leq (6L_{w^{1\epsilon}})^{-1}\ell$, then

$$w^{1\epsilon}_{i_0}(y) - \min_{j\neq i_0}\{w^{1\epsilon}_j(y) + \ell\} < 0, \quad \forall\, y \in B(x_0, 2\epsilon). \tag{5.10}$$

Then, equation i_0 in the system (SS_1P) becomes

$$\max\{\inf_{|e|\leq\epsilon} L^{\alpha_{i_0}}(y + e, \mathcal{D}w^{1\epsilon}_{i_0}(y)); w^{1\epsilon}_{i_0}(y) - \mathcal{M}u_0(y)\} = 0, \quad y \in B(x_0, 2\epsilon). \tag{5.11}$$

We have to study two cases.
<u>CASE 1 :</u> There exists $\bar{x} \in B(x_0, 2\epsilon)$ such that

$$w^{1\epsilon}_{i_0}(\bar{x}) = \mathcal{M}u_0(\bar{x}), \quad \text{i.e.} \quad w^{1\epsilon}_{i_0}(\bar{x}) = k + \inf_\xi\{u_0(\bar{x} + \xi) + c(\xi)\}.$$

Then, for all $y \in B(x_0, 2\epsilon)$,

$$w^{\epsilon 1}_{i_0}(y) + 4(L_{w^{1\epsilon}} + L_{u_0})\epsilon \geq k + \inf_\xi\{u_0(y + \xi) + c(\xi)\}.$$

Consider now $\mathcal{M}u_{h0}(y) - \mathcal{M}u_0(y)$. By proposition II.27, we have that $\mathcal{M}u_0(y) \geq \mathcal{M}u_{h0}(y) - \underline{C}_0|h|^{\underline{\gamma}}$. Then, we obtain

$$w_{i_0}^{\epsilon 1}(y) + 4(L_{w^{1\epsilon}} + L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}} \geq k + \inf_{\xi}\{u_{h0}(y+\xi) + c(\xi)\}, \quad \forall\, y \in B(x_0, 2\epsilon).$$

Since $u_{h1}(y) \leq k + \inf_\xi\{u_{h0}(y+\xi) + c(\xi)\}$, for all $y \in B(x_0, 2\epsilon)$, hence

$$u_{h1}(x_0) - w_{\epsilon i_0}^{1}(x_0) \leq 4(L_{w^{1\epsilon}} + L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}} + L_{w^{1\epsilon}}\epsilon = (5L_{w^{1\epsilon}} + 4L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}},$$

which implies

$$m_\eta \leq (5L_{w^{1\epsilon}} + 4L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}} - \eta\phi(x). \tag{5.12}$$

<u>CASE 2 :</u> For all $y \in B(x_0, 2\epsilon)$, we have

$$w_{i_0}^{1\epsilon}(y) < \mathcal{M}u_0(y).$$

The classical methods (see [2], [5]) imply that

$$\sup_{\alpha_i} L^{\alpha_i}(x_0, \mathcal{D}w_{\epsilon i_0}^{1}(x_0)) \geq 0.$$

We can apply the consistency hypothesis, to obtain

$$\begin{aligned} &-C\eta \leq S(h, x_0, (w_{\epsilon i_0}^1 + \eta\phi)(x_0)), w_{\epsilon i_0}^1 + \eta\phi) + Q(w_{\epsilon i_0}^1 + \eta\phi) \\ &\Rightarrow S(h, x_0, (w_{\epsilon i_0}^1 + \eta\phi)(x_0), w_{\epsilon i_0}^1 + \eta\phi) \geq -Q(w_{\epsilon i_0}^1) + O(l\eta). \end{aligned}$$

Monotonicity implies that

$$\begin{array}{rcl} S(h, x_0, (w_{\epsilon i_0}^1 + \eta\phi)(x_0), w_{\epsilon i_0}^1 + \eta\phi) & \leq & S(h, x_0, u_{h1}(x_0) - m_\eta, u_{h1} - m_\eta) \\ & \leq & -m_\eta + S(h, x_0, u_{h1}(x_0), u_{h1}) \\ & \leq & -m_\eta. \end{array}$$

The last inequality follows from the definition of (S1). Then, we have

$$m_\eta \leq Q(w_{\epsilon i_0}^1) + O(\eta). \tag{5.13}$$

<u>CONCLUSION :</u>
By (5.12) and (5.13), we obtain that

$$m_\eta \leq \max\left\{(5L_{w^{1\epsilon}} + 4L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}} - \eta\phi(x); Q(w_{\epsilon i_0}^1) + O(\eta)\right\}.$$

We set $\epsilon = |h|^{3\underline{\gamma}}$. Then, if η goes to 0, we can conclude that

$$m \leq \max\left\{(5L_{w^{1\epsilon}} + 4L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}}; K_c|w^{1\epsilon}|_1 \sum_{i\in J} \epsilon^{1-i}|h|^{k_i}\right\}.$$

Hence

$$\begin{aligned} u_{h1} - u_1 &= u_{h1} - w^1_{\epsilon i} + w^1_{\epsilon i} - u_1 \\ &\leq m + w^1_{\epsilon i} - w^{1\epsilon}_i + w^{1\epsilon}_i - v^1_i + v^1_i - u_1 \\ &\leq \max\left\{(5L_{w^{1\epsilon}} + 4L_{u_0})\epsilon + \underline{C}_0|h|^{\underline{\gamma}}; K_c|w^{1\epsilon}|_1 \sum_{i\in J} \epsilon^{1-i}|h|^{k_i}\right\} \\ &\quad + \ell + L_{w^{1\epsilon}}\epsilon + A_{v^1,w^{1\epsilon}}\epsilon + H_{v^1,v^{1\epsilon}}\ell^{1/3}. \end{aligned}$$

Setting $\ell = (6Lw^{1\epsilon})$, we obtain

$$\begin{aligned} u_{h1} - u_1 \leq \max\{&(12L_{w^{1\epsilon}} + 4L_{u_0})|h|^{3\underline{\gamma}} + (\underline{C}_0 + H_{v^1,v^{1\epsilon}}(6L_{w^{1\epsilon}})^{1/3})|h|^{\underline{\gamma}}; \\ &(7L_{w^{1\epsilon}})|h|^{3\underline{\gamma}} + (|J|K_c|w^{1\epsilon}|_1 + H_{v^1,v^{1\epsilon}}(6L_{w^{1\epsilon}})^{1/3})|h|^{\underline{\gamma}}\}. \end{aligned}$$

Since $|J|K_c|w^{1\epsilon}|_1 \leq |J|K_c|w^{0\epsilon}|_1 \leq \underline{C}_0$, the maximum is attained by the first term. Then we have the result.

Suppose now that ($\underline{E}_n$) and (5.8) hold for $n-1$. The same methods as before, the induction and the fact that $L_{u_{n-1}} = L_{u_0}$ give the result. □

We set

$$\underline{D}_{n-1} := \underline{C}_n - \underline{C}_{n-1} = 12L_{w^{n\epsilon}} + 4L_{u_0} + A_{w^n} + H_n(6L_{w^{n\epsilon}})^{1/3}. \tag{$\underline{D}_n$}$$

Lemma II.31 implies that

$$\underline{C}_n \leq \underline{C}_0 + n\underline{D}_0. \tag{5.14}$$

6 Proof of the main result

Before giving the proof of theorem II.7, consider the following result. Let $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(x) = \nu a^x + bx + c$, where $0 < a < 1$, $b \in \mathbb{R}^+$, $\nu > 0$ and $c \geq 0$. Let $m := \min_{n\in\mathbb{N}} \phi(n)$. Then we have the following elementary lemma that we state without proof.

Lemma II.33. *(i) ϕ attains its minimum over $\mathbb{R}$ at $r := \log_a\left(-\frac{b}{\nu \ln a}\right)$, where $-b/\nu \ln a > 0$, since $a < 1$.*

(ii) If $-\frac{b}{\nu \ln a} \geq 1$, then $r \leq 0$, and hence $m = \phi(0) = \nu + c$.

(iii) If $-\frac{b}{\nu \ln a} < 1$, then

$$\begin{aligned} m &\leq \phi(\lceil r\rceil) = a^{\lceil r\rceil} + b\lceil r\rceil \leq a^{r+1} + b(r+1) \\ &= -\frac{ab}{\ln a} + b\left(\log_a\left(\frac{b}{\nu \ln a}\right) + 1\right) + c. \quad \square \end{aligned}$$

Proof of theorem II.7, page 7 We already proved in Propositions II.13 and II.15 that (S) has a unique solution.

We start by proving the upper bound of (2.11). Consider the following decomposition :

$$\sup_x (u(x) - u_h(x)) \leq \sup_x (u(x) - u_n(x)) + \sup_x (u_n(x) - u_{hn}(x)) + \sup_x (u_{hn}(x) - u_h(x)), \tag{6.1}$$

for all $n < +\infty$. Using $u - u_n \leq 0$, $u_n - u_{hn} \leq \bar{C}_n |h|^{\bar{\gamma}}$, $u_{hn} - u_{h\infty} \leq \frac{(1-\mu)^n}{\mu} |u_{h0}|_0$, and (4.8), obtain

$$\sup_x (u(x) - u_h(x)) \leq (\bar{C}_0 + n\bar{D}_0)|h|^{\bar{\gamma}} + \frac{(1-\mu)^n}{\mu}|u_{h0}|_0. \tag{6.1b}$$

Let $\phi(n) = (\bar{C}_0 + n\bar{D}_0)|h|^{\bar{\gamma}} + \frac{(1-\mu)^n}{\mu}|u_{h0}|_0$, and let $m := \min_{n\in\mathbb{N}} \phi(n)$. Applying lemma II.33 and the fact that $r \leq \lceil r \rceil \leq r + 1$, we obtain that

- $u - u_h \leq \left(\bar{C}_0 + \frac{\bar{D}_0}{(-\ln(1-\mu))}\right)|h|^{\bar{\gamma}}$, if $-\frac{\bar{D}_0 \mu |h|^{\bar{\gamma}}}{|u_{h0}|_0 \ln(1-\mu)} \geq 1$;
- $u - u_h \leq \left[-\frac{(1-\mu)\bar{D}_0}{\ln(1-\mu)} + \bar{C}_0 + \bar{D}_0 \left(\log_{(1-\mu)} \left(-\frac{\mu \bar{D}_0 |h|^{\bar{\gamma}}}{|u_{h0}|_0 \ln(1-\mu)} \right) + 1 \right) \right] |h|^{\bar{\gamma}}$, otherwise.

Hence we have the result. We prove now the lower bound. Consider the following decomposition :

$$\sup_x (u_h(x) - u(x)) \leq \sup_x (u_h(x) - u_{hn}(x)) + \sup_x (u_{hn}(x) - u_n(x)) + \sup_x (u_n(x) - u(x)), \tag{6.2}$$

for all $n < +\infty$. Since $u_h - u_{hn} \leq 0$, $u_{hn} - u_n \leq \underline{C}_n |h|^{\underline{\gamma}}$, $u_n - u \leq \frac{(1-\mu)^n}{\mu}|u_0|_0$, and (5.14), we obtain

$$u_h - u \leq \frac{(1-\mu)^n}{\mu}|u_0|_0 + \underline{C}_0 |h|^{\underline{\gamma}} + n\underline{D}_0 |h|^{\underline{\gamma}}. \tag{6.2b}$$

Applying lemma II.33, we obtain that

- $u_h - u \leq \left(\underline{C}_0 + \frac{\underline{D}_0}{(-\ln(1-\mu))}\right)|h|^{\underline{\gamma}}$, if $-\frac{\underline{D}_0 \mu |h|^{\underline{\gamma}}}{|u_0|_0 \ln(1-\mu)} \geq 1$;
- $u_h - u \leq \left[-\frac{(1-\mu)\underline{D}_0}{\ln(1-\mu)} + \underline{C}_0 + \underline{D}_0 \left(\log_{(1-\mu)} \left(-\frac{\mu \underline{D}_0 |h|^{\underline{\gamma}}}{|u_0|_0 \ln(1-\mu)} \right) + 1 \right) \right] |h|^{\underline{\gamma}}$, otherwise.

Hence we have the result. □

Appendix

7 The upper bounds of Lipschitz constants

Proof of lemma II.19. We prove this lemma by induction. Let $n = 1$, and set

$$m_{\epsilon_1} := \sup_{x,y} \phi(x,y) := \sup_{x,y\in\mathbb{R}^N} \{u_1(x) - u_1(y) - \frac{\delta}{2}|x-y|^2 - \frac{\epsilon_1}{2}(|x|^2 + |y|^2)\}.$$

Let $m_{\epsilon_1} = \phi(x_0, y_0)$. By Ishii's lemma (see [10]), there exist $X,\ Y \in \mathcal{S}^N$ such that

$$0 \le \max\{\sup_{\alpha_i} L^{\alpha_i}(y_0, u_1(y_0), p_y, Y); u_1(y_0) - \mathcal{M}u_0(y_0)\}$$
$$- \max\{\sup_{\alpha_i} L^{\alpha_i}(x_0, u_1(x_0), p_x, X); u_1(x_0) - \mathcal{M}u_0(x_0)\}, \tag{7.1}$$

where

$$p_x = \delta(x_0 - y_0) + \epsilon_1 x_0,\ \ p_y = \delta(x_0 - y_0) - \epsilon_1 y_0, \tag{7.2}$$
$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le \delta \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \epsilon_1 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{7.3}$$

Then, (7.1) implies

$$0 \le \max\{\sup_{\alpha_i}[L^{\alpha_i}(y_0, u_1(y_0), p_y, Y) - L^{\alpha_i}(x_0, u_1(x_0), p_x, X)];$$
$$u_1(y_0) - \mathcal{M}u_0(y_0) - u_1(x_0) - \mathcal{M}u_0(x_0)\}.$$

We can reduce us to study two different cases.
CASE 1 : $\quad u_1(y_0) - \mathcal{M}u_0(y_0) - (u_1(x_0) - \mathcal{M}u_0(x_0)) \ge 0.$
This last inequality implies that $u_1(x_0) - u_1(y_0) \le L_{u_0}|x_0 - y_0|$. Then we deduce that

$$m_{\epsilon_1} \le L_{u_0}|x_0 - y_0| - \delta|x_0 - y_0|^2. \tag{7.4}$$

Setting $r := |x_0 - y_0|$, and noting that $\max_r(L_{u_0} r - \delta r^2) = L_{u_0}^2/4\delta$, we obtain

$$m_{\epsilon_1} \le \frac{L_{u_0}^2}{4\delta}.$$

Applying [18, Lemma 2.3], for fixed δ, we have that

$$\lim_{\epsilon_1 \to 0} m_{\epsilon_1} = \sup_{x,y \in \mathbb{R}^N} \{u_1(x) - u_1(y) - \delta|x - y|^2\} := m,$$

and hence $m \le L_{u_0}^2/4\delta$.
Then we have, by definition of m,

$$u_1(x) - u_1(y) \le \frac{L_{u_0}^2}{4\delta} + \delta|x - y|^2,\ \forall\, x, y \in \mathbb{R}^N.$$

Use $\min_\delta \left(\frac{L_{u_0}^2}{4\delta} + \delta|x - y|^2\right) = L_{u_0}|x - y|$, to obtain

$$u_1(x) - u_1(y) \le L_{u_0}|x - y|,\ \forall\, x, y \in \mathbb{R}^N.$$

<u>CASE 2</u> : $\sup_{\alpha_i} L^{\alpha_i}(y_0, u_1(y_0), p_y, Y) - \sup_{\alpha_i} L^{\alpha_i}(x_0, u_1(x_0), p_x, X) \geq 0$.
This is the standard case (see [18, Lemma A.1]), and we have that

$$u_1(x) - u_1(y) \leq \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_1|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1} |x - y|, \ \forall\, x, y, \in \mathbb{R}^N.$$

In conclusion, we obtain

$$L_{u_1} = \max\left\{ L_{u_0}; \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_1|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1} \right\}.$$

Since by (3.1) $|u_1|_0 \leq |u_0|_0$, using the definition of L_{u_0}, we have $L_{u_1} = L_{u_0}$.

We compute now $L_{u_1^\epsilon}$. With the same methods as before, we obtain

$$L_{u_1^\epsilon} = \max\big(L_{u_0}; \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_1^\epsilon|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1} \big).$$

In this case we have not estimate between $|u_0|_0$ and $|u_1^\epsilon|_0$, hence we must give the result in this form.

Suppose now that lemma is true for $n-1$, i.e.

$$L_{u_{n-1}} = L_{u_0}, \ L_{u_{n-1}^\epsilon} = \max\big(L_{u_0}; \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_{n-1}^\epsilon|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1} \big).$$

Applying the same method as before, we can show that

$$L_{u_n} = \max\big(L_{u_{n-1}}; \sup_{\alpha_i} \frac{[c^{\alpha_i}]_1 |u_{n-1}|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1} \big).$$

Induction and definition of (3.1) give the result. The same for $L_{u_n^\epsilon}$. □

Proof of lemma II.29 : We start by computing L_{v^1}. We set

$$m_{\epsilon_1} := \sup_{i,x,y} \phi_i(x,y) := \sup_{x,y \in \mathbb{R}^N, i \in \mathcal{I}} \{ v_i^1(x) - v_i^1(y) - \frac{\delta}{2}|x-y|^2 + \frac{\epsilon_1}{2}(|x|^2 + |y|^2) \}.$$

Let $m = \phi_j(x_0, y_0)$, i.e. (j, x_0, y_0) attains the supremum.
Let $A := \{ i \in \mathcal{I}, \ (i, x_0, y_0) \text{ attains the supremum} \}$. Then, by [2, Lemma A.2], there exists $i_0 \in A$, such that $v_{i_0}^1(y_0) < \min_{j \neq i_0} \{ v_j^1(y_0) + l \}$. Hence we have $m = \phi_{i_0}(x_0, y_0)$. The definition of viscosity solution, and Ishii's lemma imply the existence of $X, \ Y \in \mathcal{S}^N$ such that

$$\max\{ L^{\alpha_{i_0}}(x_0, v_{i_0}^1(x_0), p_x, X); v_{i_0}^1(x_0) - \min_{j \neq i} \{ v_j^1(x_0) + l \};$$

$$v_{i_0}^1(x_0) - \mathcal{M}u_0(x_0) \} \leq 0,$$

$$\max\{ L^{\alpha_{i_0}}(y_0, v_{i_0}^1(y_0), p_y, X); v_{i_0}^1(y_0) - \mathcal{M}u_0(y_0) \} \geq 0,$$

where $p_x,\ p_y,\ X\ Y$ satisfy (7.2) and(7.3). Then we can reduce us to study two cases.
CASE 1 : $\quad v^1_{i_0}(y_0) - \mathcal{M}u_0(y_0) - (v^1_{i_0}(x_0) - \mathcal{M}u_0(x_0)) \geq 0.$
This last inequality implies that $v^1_{i_0}(x_0) - v^1_{i_0}(y_0) \leq L_{u_0}|x_0 - y_0|$. From now on, we continue as the case 1 of the precedent proof, and we have

$$v^1_i(x) - v^1_i(y) \leq L_{u_0}|x-y|,\ \forall\, x, y \in \mathbb{R}^N,\ \forall\, i \in \mathcal{I}.$$

CASE 2 : $\quad L^{\alpha_{i_0}}(y_0, v^1_{i_0}(y_0), p_y, Y) - L^{\alpha_{i_0}}(x_0, v^1_{i_0}(x_0), p_x, X) \geq 0.$
This is the standard case (see [2, Lemma A.2]), and we have

$$v^1_i(x) - v^1_i(y) \leq \sup_{\alpha_i, i} \frac{[c^{\alpha_i}]_1 |v^1_i|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1}|x-y|,\ \forall\, x, y, \in \mathbb{R}^N,\ \forall\, i \in \mathcal{I}.$$

Then we obtain

$$L_{v^1} = \max\big(L_{u_0}; \sup_{\alpha_i, i} \frac{[c^{\alpha_i}]_1 |v^1_i|_0 + [f^{\alpha_i}]_1}{1 - [\sigma^{\alpha_i}]_1^2 - [b^{\alpha_i}]_1}\big).$$

The same computations lead us to obtain $L_{v^{1\epsilon}}$, and $L_{w^{1\epsilon}}$. For $n > 1$, we apply exactly the same method. We only need to recall that $L_{u_{n-1}} = L_{u_0}$. □

Theorem II.34. *The sequences $(L_{v^n})_n$, $(L_{v^{n\epsilon}})_n$, $(L_{w^{n\epsilon}})_n$ are non increasing.*

Proof : We prove this theorem for $(L_{v^n})_n$, the other cases are similar. Using lemma II.29, and since $(v^n_i)_n$ is a decreasing sequence, we obtain that $(L_{v^n})_n$ is decreasing, and then we have the result. □

8 Constants A_i

We begin this section by introducing the following notation. Let $\psi, \varphi \in C_{b,l}(\mathbb{R}^N)^M$, $M \geq 1$. We define constants $A_{\psi,\varphi}$ and $H_{\psi,\varphi}$ as follows

$$A_{\psi,\varphi} := \sqrt{2k_1 k_2^{\psi,\varphi}} + k_3^{\psi,\varphi}, \quad H_{\psi,\varphi} := \frac{3}{2^{2/3}} h_1 h_2^{\psi,\varphi}, \tag{8.1}$$

where
$k_1 = \sup_{\alpha_i}\{[\sigma^{\alpha_i}]_1^2 + [b^{\alpha_i}]_1\}$,
$k_2^{\psi,\varphi} = \sup_{\alpha_i}\{\frac{1}{4}(L_\psi + L_\varphi)^2(2[\sigma^{\alpha_i}]_1^2 + 4 + 2[b^{\alpha_i}]_1) + \frac{1}{2}(L_\psi + L_\varphi)(|\psi|_0 \wedge |\varphi|_0 [c^{\alpha_i}]_1 + [f^{\alpha_i}]_1 + L_{u_0})\}$,
$k_3^{\psi,\varphi} = \sup_{\alpha_i}\{|\psi|_0 \wedge |\varphi|_0 [c^{\alpha_i}]_1 + [f^{\alpha_i}]_1\}$.
$h_1 := \big(C_\rho \sup_{\alpha_i}(|\sigma^{\alpha_i}|_0 + |b^{\alpha_i}|_0 + |c^{\alpha_i}|_0)\big)^{1/3}$, C_ρ depends only on ρ.
$h_2^{\psi,\varphi} := (L_\varphi + A_{\psi,\varphi} + L_{u_0})^{2/3}$.

We give here an extension of the comparison principle of [3, Lemma A.1].

Proposition II.35. *Let u_n and v_n the viscosity solutions of two equations like (P_n), for $n \geq 1$, with coefficients σ, b, c, f and $\bar{\sigma}, \bar{b}$, $\bar{c}$, $\bar{f}$, respectively. Then, we have*

$$\sup_x \{u_n(x) - v_n(x)\} \leq (2k_1 k_2^{u_n,v_n})^{1/2} + k_3^{u_n,v_n},$$

where

- $k_1 = \sup_{\alpha_i} \{|\bar{\sigma}^{\alpha_i} - \sigma^{\alpha_i}|_0^2 + |\bar{b}^{\alpha_i} - b^{\alpha_i}|_0^2\}$,
- $k_2^{u_n,v_n} = \sup_{\alpha_i} \{\frac{(L_{u_n}+L_{v_n})^2}{4}(2[\sigma^{\alpha_i}]_1^2 + 4 + 2[b^{\alpha_i}]_1) + \frac{(L_{u_n}+L_{v_n})}{2}(|u_n|_0 \wedge |v_n|_0 [c^{\alpha_i}]_1 + [f^{\alpha_i}]_1 + L_{u_0})\}$,
- $k_3^{u_n,v_n} = \sup_{\alpha_i} \{|u_n|_0 \wedge |v_n|_0 |\bar{c}^{\alpha_i} - c^{\alpha_i}|_0 + |\bar{f}^{\alpha_i} - f^{\alpha_i}|_0\}$.

Proof. We prove the proposition for $n = 1$. We apply the same methods as in [3, Theorem A.1] ; we set

$$m := \sup_{x,y} \phi(x, y) := \sup_{x,y} \{u_1(x) - v_1(y) - \delta|x - y|^2 - \epsilon_1(|x|^2 + |y|^2)\}.$$

Let $m = \phi(x_0, y_0)$. Applying the notion of viscosity solution and Ishii's lemma, there exist $X, Y \in \mathcal{S}^N$ such that

$$\begin{aligned} 0 \quad \leq \quad & \max\{\sup_{\alpha_i} \bar{L}^{\alpha_i}(y_0, v_1(y_0), p_y, Y); v_1(y_0) - \mathcal{M}u_0(y_0)\} \\ & - \max\{\sup_{\alpha_i} L^{\alpha_i}(x_0, u_1(x_0), p_x, X); u_1(x_0) - \mathcal{M}u_0(x_0)\}, \end{aligned} \tag{8.2}$$

where $(p_x,\, p_y,\, X,\, Y)$ satisfy (7.2)-(7.3). Using $2\phi(x_0, y_0) \geq \phi(x_0, x_0) + \phi(y_0, y_0)$, obtain

$$|x_0 - y_0| \leq \frac{L_{u_1} + L_{v_1}}{2} \delta^{-1}. \tag{8.3}$$

Now we have to study two different cases.
CASE 1 : $v_1(y_0) - \mathcal{M}u_0(y_0) - (u_1(x_0) - \mathcal{M}u_0(x_0)) \geq 0$.
This last inequality implies that $u_1(x_0) - v_1(y_0) \leq L_{u_0}|x_0 - y_0|$, and, using (8.3), we have $u_1(x_0) - v_1(y_0) \leq L_{u_0}(L_{u_1} + L_{v_1})(2\delta)^{-1}$, which implies

$$m \leq \frac{1}{2}(L_{u_1} + L_{v_1}) L_{u_0} \delta^{-1}. \tag{8.4}$$

CASE 2 : $\sup_{\alpha_i} L^{\alpha_i}(y_0, v_1(y_0), p_y, Y) - \sup_{\alpha_i} L^{\alpha_i}(x_0, u_1(x_0), p_x, X) \geq 0$.
This is the standard case, and we use the same computations as in [3, Theorem A.1], detailing all constants. For the bounds of $-tr[\bar{a}^{\alpha_i}(y_0)Y - a^{\alpha_i}(x_0)X]$, $(b^{\alpha_i}(x_0)p_x - \bar{b}^{\alpha_i}(y_0)p_y)$, $(\bar{c}^{\alpha_i}(y_0)v_1(y_0) - c^{\alpha_i}(x_0)u_1(x_0))$, $(f^{\alpha_i}(x_0) - \bar{f}^{\alpha_i}(y_0))$, we use the estimates given in [3, Theorem A.1]. Finally we obtain

$$m \leq 2\delta \sup_{\alpha_i} \{|\bar{\sigma}^{\alpha_i} - \sigma^{\alpha_i}|_0^2 + |\bar{b}^{\alpha_i} - b^{\alpha_i}|_0^2\} + \frac{1}{\delta} \sup_{\alpha_i} \{(2[\sigma^{\alpha_i}]_1^2 + 4 + 2[b^{\alpha_i}]_1)\big(\frac{L_{v_1} + L_{u_1}}{2}\big)^2$$

$+(|u_1|_0[c^{\alpha_i}]_1+[f^{\alpha_i}]_1)\big(\frac{L_{v_1}+L_{u_1}}{2}\big)\}+\sup_{\alpha_i}\{|v_1|_0|\bar{c}^{\alpha_i}-c^{\alpha_i}|_0+|\bar{f}^{\alpha_i}-f^{\alpha_i}|_0\}+\epsilon_1(1+|x_0|^2+|y_0|^2)$.
If we add the two cases, we have

$$m \leq 2k_1\delta + \frac{k_2}{\delta} + k_3 + \epsilon_1 k_4,$$

where

- $k_1 = \sup_{\alpha_i}\{|\bar{\sigma}^{\alpha_i} - \sigma^{\alpha_i}|_0^2 + |\bar{b}^{\alpha_i} - b^{\alpha_i}|_0^2\}$,
- $k_2 = \sup_{\alpha_i}\{\frac{(L_{u_1}+L_{v_1})^2}{4}(2[\sigma^{\alpha_i}]_1^2 + 4 + 2[b^{\alpha_i}]_1) + \frac{(L_{u_1}+L_{v_1})}{2}(|u_1|_0[c^{\alpha_i}]_1 + [f^{\alpha_i}]_1 + L_{u_0})\}$,
- $k_3 = \sup_{\alpha_i}\{|v_1|_0|\bar{c}^{\alpha_i} - c|_0 + |\bar{f}^{\alpha_i} - f^{\alpha_i}|_0\}$,
- $k_4 = (1 + |x_0|^2 + |y_0|^2)$.

Since $\min_\delta\{2k_1\delta + \frac{k_2}{\delta}\} = \sqrt{2k_1k_2}$, letting ϵ_1 go to 0, we obtain

$$m \leq \sqrt{2k_1k_2} + k_3.$$

Reversing $|u_1|_0$ and $|v_1|_0$, we have also the symmetric inequality, hence we have the result, with $k_i^{u_1,v_1}$ defined as before. For the general case, we have only to recall that $L_{u_{n-1}} = L_{u_0}$, for all n.□

Proof of proposition II.20. We apply the precedent proposition, using that $|\bar{g}-g| \leq [g]_1\epsilon$, for $g = \sigma, b, c, f$. Then we have the result.□

Consider now the switching systems. We give here an extension of [3, Lemma A.1].

Proposition II.36. *Let v^n and w^n be solutions of two equations (SS_n), for $n \geq 1$, with coefficients σ, b, c, f and $\bar{\sigma}, \bar{b}$, $\bar{c}$, $\bar{f}$, respectively. Then, we have*

$$\sup_{x,i}\{v_i^n(x) - w_i^n(x)\} \leq (2k_1k_2^{v^n,w^n})^{1/2} + k_3^{v^n,w^n},$$

where

- $k_1 = \sup_{\alpha_i}\{|\bar{\sigma}^{\alpha_i} - \sigma^{\alpha_i}|_0^2 + |\bar{b}^{\alpha_i} - b^{\alpha_i}|_0^2\}$,
- $k_2^{v^n,v^n} = \sup_{\alpha_i}\{\frac{(L_{v^n}+L_{w^n})^2}{4}(2[\sigma^{\alpha_i}]_1^2 + 4 + 2[b^{\alpha_i}]_1)$
  $+\frac{(L_{u_n}+L_{v_n})}{2}(|v^n|_0 \wedge |w^n|_0[c^{\alpha_i}]_1 + [f^{\alpha_i}]_1 + L_{u_0})\}$,
- $k_3^{v^n,w^n} = \sup_{\alpha_i}\{|v^n|_0 \wedge |w^n|_0|\bar{c}^{\alpha_i} - c^{\alpha_i}|_0 + |\bar{f}_i^{\alpha} - f^{\alpha_i}|_0\}$.

Proof. We prove the proposition for $n = 1$. We apply the same methods as in [3, Theorem A.1] ; we set

$$m := \sup_{x,y,i}\phi_i(x,y) := \sup_{x,y,i}\{v_i^1(x) - w_i^1(y) - \delta|x-y|^2 - \epsilon_1(|x|^2 + |y|^2)\}.$$

Let $m = \phi_j(x_0, y_0)$, i.e. (j, x_0, y_0) attains the supremum.
Let $A := \{i \in \mathcal{I}, (i, x_0, y_0)$ attains the supremum$\}$. Then, by [2, Lemma A.2], there exists

$i_0 \in A$, such that $w^1_{i_0}(y_0) < \min_{j\neq i_0}\{w^1_j(y_0) + l\}$. Applying the notion of viscosity solution, and Ishii's lemma, there exist $X, Y \in \mathcal{S}^N$ such that

$$\begin{aligned} 0 \le \max\{&\bar{L}^{\alpha_{i_0}}(y_0, w^1_{i_0}(y_0), p_y, Y); w^1_{i_0}(y_0) - \mathcal{M}u_0(y_0)\} \\ &- \max\{L^{\alpha_{i_0}}(x_0, v^1_{i_0}(x_0), p_x, X); v^1_{i_0}(x_0) - \min_{j\neq i_0}\{v^1_j(x_0) + \ell\}; v^1_{i_0}(x_0) - \mathcal{M}u_0(x_0)\}, \end{aligned} \tag{8.5}$$

where p_x, p_y, X and Y satisfy (7.2) and (7.3). Continuing as in proposition II.35, we obtain the result. □

Proof of lemma II.30. We apply the precedent theorem, using that $|\bar{g} - g| \le [g]_1\epsilon$, for $g = \sigma,\ b,\ c,\ f$. □

Theorem II.37. *We have that*

$$A_{v^n} \le \cdots \le A_{v^2} \le A_{v^1},$$

$$A_{w^n} \le \cdots \le A_{w^2} \le A_{w^1}.$$

Proof : The form of A_g and L_g, $g = v^i, w^i$, $i \ge 0$, defined in (8.1) and (5.7) respectively, imply the result. □

Acknowledgments We thank the referees for their numerous useful remarks that helped to improve the paper.

Bibliographie

[1] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume 17 of *Mathématiques et Applications*. Springer, Paris, 1994.

[2] G. Barles and E.R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numerical Analysis*. To appear.

[3] G. Barles and E.R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN. Mathematical Modelling and Numerical Analysis*, 36 :33–54, 2002.

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

[5] J.F. Bonnans, S. Maroso, and H. Zidani. Error estimates for stochastic differential games : the adverse stopping case. *IMA, J. Numerical Analysis*, 26 :188–212, 2006.

[6] J.F. Bonnans, E. Ottenwaelter, and H. Zidani. Numerical schemes for the two dimensional second-order HJB equation. *ESAIM : M2AN*, 38 :723–735, 2004.

[7] J.F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numerical Analysis*, 41 :1008–1021, 2003.

[8] I. Capuzzo-Dolcetta and L.C. Evans. Optimal switching for ordinary differential equations. *SIAM J. Control Optim.*, 22 :143–161, 1984.

[9] J.-Ph. Chancelier, B. Øksendal, and A. Sulem. Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. *Tr. Mat. Inst. Steklova*, 237(Stokhast. Finans. Mat.) :149–172, 2002.

[10] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. American Mathematical Society (New Series)*, 27 :1–67, 1992.

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

[12] H. Ishii and S. Koike. Viscosity solutions for monotone systems of second order elliptic PDEs. *Comm. Partial Differential Equation*, 16 :1095–1128, 1991.

[13] H. Ishii and S. Koike. Viscosity solutions of a system of nonlinear second order elliptic PDEs arising in switching games. *Funkcial. Ekvac.*, 34 :143–155, 1991.

[14] K. Ishii. Viscosity Solutions of Non Linear Second Order Elliptic PDEs associated with Impulse Control Problems. *Funkcialaj Ekvacioj*, 36 :123–141, 1993.

[15] K. Ishii. Viscosity Solutions of Non Linear Second Order Elliptic PDEs associated with Impulse Control Problems II. *Funkcialaj Ekvacioj*, 38 :297–328, 1995.

[16] E. R. Jakobsen. On error bounds for approximation schemes for non-convex degenerate elliptic equations. *BIT*, 44(2) :269–285, 2004.

[17] E. R. Jakobsen. On error bounds for monotone approximation schemes for multidimensional Isaac equations. To appear, 2004.

[18] E.R. Jakobsen and K.H. Karlsen. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate elliptic equations. *Electronic J. Differential Equations*, pages 1–10, 2002.

[19] N.V. Krylov. On the rate of convergence of finite difference approximation for Bellman's equation. *St. Petersburg Math. J.*, 9 :639–650, 1997.

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

[21] H.J. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of mathematics*. Springer, New York, 2001. Second edition.

CHAPITRE III

An algorithm for a stochastic impulse control problem

An algorithm for a stochastic impulse control problem

1 Introduction

The aim of this paper is to study an algorithm for solving numerically stochastic control problems with obstacle, and in particular stochastic impulse control problem with infinite horizon. For stochastic differential equation with Brownian motion, this amounts to solve a discounted Hamilton-Jacobi-Bellman equation of the form :

$$\max\{\sup_{\alpha\in\mathcal{A}}(\mathcal{L}^\alpha V(x) - f^\alpha(x)); V(x) - \mathcal{M}V(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{1.1}$$

where $\mathcal{L}^\alpha$ is the HJB operator

$$\mathcal{L}^\alpha(V)(x) = \Big[- tr[a^\alpha(x) D^2 V(x)] - b^\alpha(x) DV(x) + \lambda V(x) \Big]. \tag{1.2}$$

Here λ is the actualization coefficient. The mapping $b^\alpha(x)$, $a^\alpha(x)$ and $f^\alpha(x)$ represent the drift, diffusion matrix and running cost, respectively. They are continuous mapping from $\mathcal{A} \times \mathbb{R}^N$ into $\mathbb{R}^N$, the space $\mathcal{S}_n$ of symmetric $N \times N$ matrices, and $\mathbb{R}$ respectively. The operator $\mathcal{M}$ takes into account the impulse control itself. We assume it to be of the form :

$$\mathcal{M}V(x) = k + \inf_\xi \{c(\xi) + V(x+\xi)\}, \tag{1.3}$$

with

$$\begin{cases} k > 0, \quad c : \mathbb{R}^N_+ \to \mathbb{R}_+, \\ c(0) = 0, \quad c(\xi_1 + \xi_2) \le c(\xi_1) + c(\xi_2). \end{cases}$$

The function c may have value in $\mathbb{R}\cup\{+\infty\}$; then at the point $x \in \mathbb{R}^N$, only values of impulse in the domain of c (set of ξ such that $c(\xi) < +\infty$) are allowed. As proved in [14], the operator $\mathcal{M}$ has the following properties :

(i) If $u \le v$ in $\mathbb{R}$, then $\mathcal{M}u \le \mathcal{M}v$ in $\mathbb{R}^N$.

(ii) $\mathcal{M}(tu + (1-t)v) \ge t\mathcal{M}u + (1-t)\mathcal{M}v$, $t \in [0,1]$.

(iii) $\mathcal{M}(u + c) = \mathcal{M}u + c$, for all $c \in \mathbb{R}$.

(iv) $|\mathcal{M}u - \mathcal{M}v|_0 \le |u - v|_0$, for all $u, v \in C(\mathbb{R}^N)$.

A related problem is the obstacle one, in which the equation to be solved is

$$\max\{\sup_{\alpha\in\mathcal{A}}(\mathcal{L}^\alpha V(x) - f^\alpha(x)); \lambda V(x) - \psi(x)\} = 0, \tag{1.4}$$

where the obstacle ψ is a real function over $\mathbb{R}^N$. The impulse problem may be viewed as an obstacle problem where the obstacle depends on the solution V.

An even simpler problem is the standard Hamilton-Jacobi-Bellman (HJB) equation (without obstacle)

$$\sup_{\alpha\in\mathcal{A}}(\mathcal{L}^\alpha V(x) - f^\alpha(x)) = 0, \quad x \in \mathbb{R}^N. \tag{1.5}$$

For discounted infinite horizon problems, the obstacle problem can be reduced to the standard one. The following reduction is obviously well known, altought we have no reference for it. We skeep the proof since it is elementary.

Lemma III.1. *Consider a control $\alpha_{obs} \notin \mathcal{A}$. Set $\hat{\mathcal{A}} = \mathcal{A} \cup \{\alpha_{obs}\}$, and define $\hat{\mathcal{L}^\alpha}$, $\hat{f}^\alpha$ by*

$$\begin{aligned}
&\hat{\mathcal{L}^\alpha} V(x) = \mathcal{L}^\alpha V(x), && \text{if} \quad \alpha \in \mathcal{A}, \\
&\mathcal{L}^{\alpha_{obs}} V(x) = \lambda V(x), && \\
&\hat{f}^\alpha(x) = f^\alpha(x) && \text{if} \quad \alpha \in \mathcal{A}, \\
&\hat{f}^{\alpha_{obs}}(x) = \psi(x). &&
\end{aligned}$$

Then the obstacle problem (1.4) is equivalent to the standard problem (1.5), with data $\hat{\mathcal{A}}$, $\hat{\mathcal{L}^\alpha}$ and $\hat{f}^\alpha$. □

In view of this reduction procedure, we see that any algorithm for solving the standard HJB equation has an immediate extension for obstacle problems. We will look in particular at the Howard policy iteration algorithm.

2 Discretized problems

A possible way of computing solution of (1.1) is to discretize the state space $\mathbb{R}^N$ by introducing a regular discrete grid $\mathcal{O}_h \in \mathbb{R}^N$. Denote N_{tot} the (finite) number of points in the grid. Discretization of (1.1) over such a grid, may generally be interpreted as an impulse control problem for the optimal control of a Markov chain ; see e.g. [7, 10, 11, 15] and references therein. Let $\mathcal{L}_h^\alpha$ and $\mathcal{M}_h$ the discretizations of operators $\mathcal{L}^\alpha$ and $\mathcal{M}$ respectively. So, the discrete equation reads as follows :

$$\max\Big\{ \sup_{\alpha \in \mathcal{A}^{N_{tot}}} (\mathcal{L}_h^\alpha V_h - f(\alpha)); V_h - \mathcal{M}_h V_h \Big\} = 0. \tag{2.1}$$

In particular $\mathcal{L}_h^\alpha$ is a matrix of dimension $N_{tot} \times N_{tot}$, V_h, $\mathcal{M}_h V_h$ and $f(\alpha)$ are vectors of dimension N_{tot}. A policy for this discrete equation is a mapping $\mathcal{O}_h \to \mathcal{A}$; we denote by $\mathcal{A}^{N_{tot}}$ the set of policies. The operator $\mathcal{L}_h^\alpha$ is assumed to be, for a given policy α, linear, and to satisfy the maximum principle, i.e. $\mathcal{L}_h^\alpha V_h \geq 0$ implies $V_h \geq 0$. We assume also that non diagonal elements of the matrix $\mathcal{L}_h^\alpha$ are non positive, and that if V_h is constant, then $\mathcal{L}_h^\alpha V_h = \lambda V_h$. The operator $\mathcal{M}_h$ will be detailed later. The main problem in solving equation (2.1) is that the value function appears in the obstacle, so we present here an algorithm in which this problem has been avoided.

In the literature, Bellman's equations associated to optimal control problem of Markov chain on infinite horizon, with discount factor $\lambda > 0$, have been studied for a long time by many authors (see, for example [15], [11], [10] and the references therein). Typically the continuous Bellman equation can be written in the form

$$\sup_{\alpha \in \mathcal{A}} \{\mathcal{L}^\alpha V(x) - f^\alpha(x)\} = 0, \quad x \in \mathbb{R}^N, \tag{2.2}$$

and the correspondent Markov chain approximation becomes

$$\sup_{\alpha\in\mathcal{A}^{N_{tot}}}\left\{\mathcal{L}_h^\alpha V_h - f(\alpha)\right\}=0, \tag{2.3}$$

in the grid $\mathcal{O}_h$, which can also be written as

$$V_h=\beta\inf_{\alpha\in\mathcal{A}^{N_{tot}}}(V_h+\Delta t(\lambda V_h-\mathcal{L}_h^\alpha V_h+f(\alpha))),$$

where Δt is a so called fictitious time step, and $\beta := (1+\lambda\Delta t)^{-1}$ is the discrete actualization coefficient. If $\Delta t > 0$ is small enough, then in view of the hypothesis made on $\mathcal{L}_h^\alpha$, the matrix M_h^α defined by $M_h^\alpha V_h = V_h + \Delta t(\lambda V_h - \mathcal{L}_h^\alpha V_h)$ has non-negative coefficients whose sum over a row equals to 1. We see that for each policy α, M_h^α is the transition matrix of a Markov chain.

The fixed point reformulation of (2.2) is the basis of the two main algorithms, the value iteration

$$V_h^{k+1}=\beta\inf_\alpha\{M_h^\alpha V_h^k+f(\alpha)\},$$

and the policy iteration (due to Howard), which consists for a policy α^k at step k to solve the linear system

$$\mathcal{L}_h^{\alpha^k}V_h^{k+1}-f(\alpha^k)=0,$$

which is equivalent to solve

$$V_h^{k+1}=\beta(M_h^{\alpha^k}V_h^{k+1}+f(\alpha^k)),$$

and then to update the policy by the formula (where the minimum and the maximum are taken componentwise)

$$\alpha^{k+1}\in\operatorname{argmax}\{\mathcal{L}_h^\alpha V_h^k+f(\alpha)\},$$

which is equivalent to

$$\alpha^{k+1}\in\operatorname{argmin}\{M_h^\alpha V_h^k+f(\alpha)\}.$$

It is known (see e.g. [6]) that the Howard's policy iteration algorithm converges faster than the Value Iteration algorithm. Moreover, in [11], the authors accelerate the value iteration algorithm, by using Howard's policy iterations.

Since the obstacle problem (1.4) can be written as a standard problem (1.5), we can applied the same methods to solve it. In particular we consider Howard policy iteration algorithm applied to the following problem :

$$\max(M_1V_h-b;M_2V_h-\psi)=0, \tag{2.4}$$

where M_i are monotone matrices of dimension $N_{tot}\times N_{tot}$(i.e. $M_iX\geq 0\Rightarrow X\geq 0$), $i=1,2$, V_h, b and ψ are vectors in $\mathbb{R}^{N_{tot}}$, and V_h is the solution of the problem. In the litterature, to solve (2.4), different methods are used, and we look in particular to Primal-Dual active set (see [12]). We prove in this paper the equivalence between the Primal-Dual Active Set algorithm and the Howard policy iteration algorithm.

The problem

$$\max\{\sup_{\alpha\in\mathcal{A}}\{\mathcal{L}_h^\alpha V_h - f(\alpha)); V_h - \mathcal{M}_h V_h\} = 0,$$

is a kind of obstacle problem, in which the obstacle depends on the value function. In this paper we present an approach to solve it numerically, and we give numerical examples.

The approach consists in building a sequence of problems which approximate (2.1), and solve each problem of this sequence using an Howard policy iteration algorithm. In particular, we will call this sequence the cascade approximation of (2.1), and we can write it in the following way : the first problem is

$$\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_{h0} - f(\alpha)) = 0,$$

which gives a solution V_{h0}, and for $n > 0$, the n-th problem is written as follows

$$\max\{\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_{hn} - f(\alpha)); V_{hn} - \mathcal{M}_h V_{h(n-1)}\} = 0, \tag{2.5}$$

where V_{hn} is the solution. We note that, we do not have the value function in the obstacle part, but we have that the obstacle of the n-th problem depends on the solution of the $(n-1)$-th problem. We have then a standard obstacle problem and we solve it by Howard algorithm. In the algorithm that we propose, we consider two possibilities. In the first one, we solve every problem of the cascade using Howard algorithm, until its convergence, i.e. for the n-th problem we build a sequence $(V_{hn}^k)_k$, and we iterate Howard algorithm until the convergence of this sequence. The second possibility consists in doing only one iteration of the Howard algorithm for every problem of the cascade, until the final convergence. We will see that the second approach is better than the first, because it gives the solution faster.

The paper is organized as follows. In Section 3 we present the Howard policy iteration algorithm for control problem, and we give its main properties. In Section 4 we introduce the obstacle problem, we give Howard algorithm and Primal-Dual Active set algorithm to solve it, and finally we show the equivalence between the two approaches. In Section 5 we present the algorithm to solve impulse control problem. Finally, in section 6 we give a numerical example for which we compare the two ways to apply the algorithm.

Notations We give here some definition we will use in the paper.

Definition III.2. *A $n \times n$ matrix M is called a **P**-matrix if all its principal minors are positive.*

Definition III.3. *A $n \times n$ matrix M is called a M-matrix if it is nonsingular, $(M_{ij}) \leq 0$, for $i \neq j$, and $M^{-1} \geq 0$.*

Definition III.4. *(i) A $n \times n$ matrix M is monotone if $My \geq 0 \Rightarrow y \geq 0$, for all $y \in \mathbb{R}^n$.*
(ii) A $n \times n$ matrix M is anti-monotone if $My \geq 0 \Rightarrow y \leq 0$, for all $y \in \mathbb{R}^n$.

3 Howard algorithm

We recall here the classical Howard policy iteration algorithm to solve

$$\max_{\alpha\in\mathcal{A}^N}(A(\alpha)y - f(\alpha)) = 0, \tag{3.1}$$

where $y \in \mathbb{R}^N$, $\mathcal{A}$ is a compact set and it is the set of the controls, $A(\alpha)$ is a matrix $N \times N$, and $f(\alpha) \in \mathbb{R}^N$. To apply Howard algorithm, we suppose the following assumption

(A1) For all $\alpha \in \mathcal{A}$, the matrix $A(\alpha)$ is monotone, i.e.

$$\forall X \in \mathbb{R}^n,\ A(\alpha)X \geq 0 \Rightarrow X \geq 0.$$

(A2) The functions $\alpha \to A(\alpha)$, and $\alpha \to f(\alpha)$ are continuous.

We give now the Howard policy iteration algorithm.

Howard algorithm

(i) Initialize $\alpha^0 \in \mathcal{A}^N$. Set $k = 0$.

(ii) **Iteration k+1 :**

- Solve in y^{k+1} : $A(\alpha^k)y^{k+1} - f(\alpha^k) = 0$.
- Set $\alpha^{k+1} := \text{argmax}_{\alpha\in\mathcal{A}^N}(A(\alpha)y^{k+1} - f(\alpha))$

(iii) If $y^k = y^{k+1}$ the stop ; else set $k = k + 1$ and return to (ii).

Remark III.5. Clearly, since the obstacle problem

$$\max\{\max_{\alpha\in\mathcal{A}^N}(A(\alpha)y - f(\alpha); y - \psi) = 0, \tag{3.2}$$

can be reformulated in the standard form

$$\max_{\alpha\in\{\mathcal{A}\cup\{\alpha_{obs}\}\}^N}(\hat{A}(\alpha)y - \hat{f}(\alpha)) = 0,$$

this algorithm has an immediate extension to problems as (3.2).

3.1 Convergence of the algorihtm

This algorithm satisfies the following proposition.

Proposition III.6. *Under assumptions (A1-2), we have the following properties :*

(a) $y^{k+1} \leq y^k$*, for all* $k \geq 0$.

(b) The sequence $(y^k)_k$ *converges to the solution* y *of* (3.1).

(c) For the problem without obstacle, if $\mathcal{A}$ *is a finite set, card*$(\mathcal{A}) = p$*, then the sequence* $(y^k)_k$ *converges in no more than* p^N *iterations.*

(d) If at an iteration k *we have* $y_i^k < \psi_i$*, for a component* i*, then we will have* $y_i^{\bar{k}} < \psi_i$*, for all* $\bar{k} \geq k$.

(e) For the obstacle problem (3.2)*, suppose that* $\mathcal{A}$ *is a finite set, card*$(\mathcal{A}) = p$*, and that* $\alpha^{k+1} \in A$ *if* $\max_{\alpha\in\mathcal{A}}(A(\alpha)y^{k+1} - f(\alpha)) = y^{k+1} - \psi$*. Then the sequence* $(y^k)_k$ *converges in at most* Np^N *iterations.*

We recall here the proof of these properties.

Proof.
(a) We have that,

$$\begin{array}{rcl} A(\alpha^k)y^{k+1} - f(\alpha^k) & = & 0 \\ & = & A(\alpha^{k-1})y^k - f(\alpha^{k-1}) \\ & \leq & \max_{\alpha\in\mathcal{A}}\{A(\alpha)y^k - f(\alpha)\} \\ & \leq & A(\alpha^k)y^k - f(\alpha^k). \end{array}$$

The monotonicity of the matrix $A(\alpha)$ implies the result.

(b) We note that $y^k = (A(\alpha^{k-1}))^{-1}f(\alpha^{k-1})$, with $A(\alpha)$ always invertible since $A(\alpha)$ is monotone. Then, we have

$$||y^k|| \leq \max_{\alpha\in\mathcal{A}} ||A^{-1}(\alpha)f(\alpha)||.$$

We have that $||A^{-1}(\alpha)f(\alpha)||$ is a continuous function on $\mathcal{A}$ compact, and then it is bounded. Then we can say that every component of the sequence y^k is a decreasing and bounded sequence, and hence $(y^k)_k$ converges to some $y \in \mathbb{R}^N$.
We define $F : \mathbb{R}^N \to \mathbb{R}^N$ as follows

$$F(y) = \max_{\alpha\in\mathcal{A}}(A(\alpha)y - f(\alpha)). \tag{3.3}$$

We note that $||F(y)-F(z)||_\infty \leq \max_{\alpha\in\mathcal{A}} ||A(\alpha)y-f(\alpha)-(A(\alpha)z-f(\alpha))|| \leq C||y-z||_\infty$, where C is a constant. Then F is continuous and $F(y) = \lim_{k\to\infty} F(y_k)$. Moreover (α^k) is a sequence in $\mathcal{A}^N$ compact, then we can extract a sub-sequence $(\alpha^{k'})_{k'}$ which converges to a $\bar{\alpha}$. Then we have $A(\alpha^{k'-1})y^{k'} - f(\alpha^{k'-1}) = 0$, for all k', and passing to the limit we have $A(\bar{\alpha})y - f(\bar{\alpha}) = 0$. On the other hand,

$$\begin{array}{rcl} F(y) & = & \lim_{k\to+\infty} F(y^{k'}) \\ & = & \lim_{k\to+\infty}(A(\alpha^{k'-1})y^{k'} - f(\alpha^{k'-1})) \\ & = & A(\bar{\alpha})y - f(\bar{\alpha}). \end{array}$$

Then $F(y) = 0$, and y is the solution of (4.1).

(c) Let $q = \text{card}(\mathcal{A}^N) = p^N$. The set of admissible control is finite and of cardinality p, and this implies that after p iterations of the Howard algorithm we have to use a control that we have already used before. Then, there exists k, $0 \leq k \leq q-1$, such that $\alpha^q = \alpha^k$. Then we have also $y^q = y^k$. Since $y^k \geq y^{k+1} \geq \cdots \geq y^q$, we must have $y^k = y^{k+1} = \cdots = y^q$. Finally $y^{q-1} = y^q$ and $\alpha = \alpha^q$ gives a solution of problem (3.1).

(d) We have proved that, for every component i, $(y_i^k)_k$ is a decreasing sequence. So we can say that, if at the iteration k, we have $y^k < \psi$, hence $y_i^{\bar{k}} \leq y_i^k < \psi_i$, for all $\bar{k} \geq k$.

(e) Form (d) we deduce that, once for a given k the obstacle is not active at some point of the discretized state space, it will remain inactive. Moreover, when both member are active, we choose the obstacle to be inactive. So there are at most N iterations having the same number of states where the obstacle constraint is active. The results follows. □

Remark III.7. It is known (see e.g. [6]) that if $\mathcal{A}$ is a numerable set, and matrix M_h defined in the previous section is a transition matrix of a Markov chain, then $||y^{k+1} - y|| \leq \beta||y^k - y||$, for $k \to +\infty$, where $0 < \beta < 1$. The proof of this result is based on a comparison of the Howard policy iteration algorithm with the Value Iteration algorithm.

Remark III.8. The difficult step for implementing the policy algorithm is to solve linear system $A(\alpha^k)y^{k+1} - f(\alpha^k) = 0$, especially for large scale problems. In this respect let us mention the fast multi-grid approach due to [1].

Remark III.9. To obtain the convergence of the Howard algorithm it is important that the matrix A satisfies monotonicity assumption (or anti-monotonicity assumption). In particular, for the obstacle problem, we have to ensure that the matrix $\hat{A}(\alpha)$ is monotone, for all α.

Remark III.10. Policy iteration algorithm can be applied to solve for shortest path problems (see e.g. [10]). Given a graph $\mathcal{G}$, and given a node r on this graph, the shortest paths tree problem consists in finding a spanning tree of $\mathcal{G}$, rooted on r, such that, for all nodes i, the path from i to r in the spanning tree is a path of minimal weights from i to r. It can be shown [10] that optimal path problems are nothing but special discrete deterministic Optimal Control problems.

3.2 Superlinear convergence of the Howard algorithm when $\mathcal{A}$ is a compact set

We have seen in Proposition III.6 that when $\mathcal{A}$ is a finite set such that $card(\mathcal{A}) = p$, Howard algorithm converges in a finite number of iterations, bounded by p^N. Moreover, it is known (see e.g. [6]) that when $\mathcal{A}$ is a numerable set, then Howard algorithm has a linear convergence. We prove in this section the superlinear convergence of the Howard algorithm when $\mathcal{A}$ is a compact set.

This result has been proved in [16], and [17] for particular problems and under particular assumptions (see Remark III.16 below). We give here a proof, which seems to us clear and simple, of the superlinear convergence for the generale case.

The key step of this proof is to prove that Howard's algorithm for problem (3.1) is a semi-smooth Newton's method applied to find the zero of the function F defined in (3.3),

$$F(x) := \max_{\alpha \in \mathcal{A}^N} (A(\alpha)x - f(\alpha)).$$

For $k \geq 0$, by definitions of α^{k+1} and x^k, we have

$$A(\alpha^{k+1})x^k - f(\alpha^{k+1}) = F(x^k), \text{ and } A(\alpha^{k+1})x^{k+1} - f(\alpha^{k+1}) = 0.$$

Therefore, $A(\alpha^{k+1})(x^k - x^{k+1}) = F(x^k)$, and thus

$$x^{k+1} = x^k - A(\alpha^{k+1})^{-1}F(x^k).$$

This is as in a semi-smooth Newton's method, where $A(\alpha^{k+1})$ plays the role of a derivative of F at point x^k.

In order to prove the superlinear convergence, we shall prove that F is slantly differentiable in the sense of [12, Definition 1].

Definition III.11 (Slantly differentiability)**.** *A function $F : \mathbb{R}^N \to \mathbb{R}^N$ is said* slantly differentiable *in an open set $U \subset \mathbb{R}^N$ if there exists a family of mappings $G : U \to \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$ such that*

$$F(x + h) = F(x) + G(x + h)h + o(h)$$

as $h \to 0$, $\forall x \in U$. G is called a slanting function.

It is then easy to show the local super-linear convergence of the semi-smooth Newton's method defined by $x^{k+1} = x^k - G(x^k)^{-1}F(x^k)$, (see [12, Theorem 1.1]).

In our case, let

$$\alpha(x) := \arg\max_{\alpha\in\mathcal{A}^N} (A(\alpha)x - f(\alpha))$$

Theorem III.12. *We assume (A1)-(A2). Then the function F defined by (3.3) is slantly differentiable, with slanting function $G(x) = A(\alpha(x))$.*

Proof. Consider first the case when $\mathcal{A}$ is finite. Let $\mathcal{A}^x$ be the set of optimal controls α associated to x, i.e.

$$\mathcal{A}^x := \left\{\alpha \in \mathcal{A}^N,\ A(\alpha)x - f(\alpha) = A(\alpha(x))x - f(\alpha(x))\right\}.$$

We note that, for h sufficiently small,

$$\alpha(x + h) \in \mathcal{A}^x. \tag{3.4}$$

Indeed, if let $1 \le i \le N$ and let $\alpha_i \in \mathcal{A}$ be such that

$$\big(A(\alpha)x - f(\alpha)\big)_i < \big(A(\alpha(x))x - f(\alpha(x))\big)_i, \tag{3.5}$$

then for $||h|| \le \eta$ small (for some $\eta > 0$ independent of i), we still have

$$\big(A(\alpha)(x + h) - f(\alpha)\big)_i < \big(A(\alpha(x))(x + h) - f(\alpha(x))\big)_i. \tag{3.6}$$

With $\alpha_i = \alpha_i(x)$ we obtain an equality in (3.6). Hence an optimal maximizer $\alpha_i = \alpha_i(x + h)$ can not satisfy (3.5), which means that it will satisfy (3.5) with the equality sign and thus that $\alpha(x + h) \in \mathcal{A}^x$.

From this we deduce that, for h sufficiently small, $A(\alpha(x+h))x - f(\alpha(x+h)) = A(\alpha(x))x - f(\alpha(x))$, and thus,

$$F(x + h) - F(x) - A(\alpha(x + h))h = 0.$$

This means that F is slantly differentiable with slanting function $G(x) := A(\alpha(x))$.

When $\mathcal{A}$ is infinite (and compact), (3.4) is not necesserly satisfied. However we have the following Lemma that we prove later on the paper.

Lemma III.13. *For any $x \in \mathbb{R}^N$,*

$$d(\alpha(x+h),\ \mathcal{A}^x) \to 0 \quad \textit{as} \quad h \to 0.$$

Now, for any $\alpha \in \mathcal{A}^x$, we have, by using the convexity argument,

$$F(x) + A(\alpha)h \le F(x+h) \le F(x) + A(\alpha(x+h))h,$$

and thus

$$0 \ge F(x+h) - F(x) - A(\alpha(x+h))h \ge (A(\alpha) - A(\alpha(x+h)))h. \tag{3.7}$$

Then by Lemma III.13 there exists a sequence $\alpha^h \in \mathcal{A}^x$ such that $d(\alpha^h, \alpha(x+h)) \to 0$ as $h \to 0$. Thus, by continuity, $A(\alpha^h) - A(\alpha(x+h)) = o(1)$ and using $\alpha = \alpha^h$ in (3.7) we obtain the result. □

Proof of Lemma III.13 First we note that $\mathcal{A}^x = \mathcal{A}_1^x \times \cdots \times \mathcal{A}_N^x$ where

$$\mathcal{A}_i^x := \left\{\alpha_i \in \mathcal{A},\ \big(A(\alpha)x - f(\alpha)\big)_i = \big(A(\alpha(x))x - f(\alpha(x))\big)_i\right\}.$$

Hence it suffices to prove that $d(\alpha_i(x+h), \mathcal{A}_i^x) \to 0$. Suppose on the contrary that there exists some $\delta > 0$ and a subsequence $h_n > 0$, $h_n \to 0$ such that $d(\alpha_i(x+h_n), \mathcal{A}_i^x) \ge \delta\ \forall n \ge 0$. Let $K_\delta := \{\alpha_i \in \mathcal{A},\ d(\alpha_i, \mathcal{A}_i^x) \ge \delta\}$, $\ell(\alpha_i) := (A(\alpha)x - f(\alpha))_i$, and

$$m_\delta := \sup_{\alpha_i \in K_\delta} \ell(\alpha_i).$$

We note that K_δ is a compact set, hence $m_\delta = \ell(\bar{\alpha}_i)$ for some $\bar{\alpha}_i \in K_\delta$. In particular $\bar{\alpha}_i \notin \mathcal{A}_i^x$ and thus $m_\delta = \ell(\bar{\alpha}_i) < \ell(\alpha_i(x))$. On the other hand, $\alpha_i(x+h_n) \in K_\delta$ thus $\ell(\alpha_i(x+h_n)) - \ell(\alpha_i(x)) \le \mu$ where $\mu := \ell(\bar{\alpha}_i) - \ell(\alpha_i^x) < 0$.

We noticed that F is C-lipschitz, and also we have $(F(x+h) - F(x))_i = \ell(\alpha_i(x+h)) - \ell(\alpha_i(x)) + (A(\alpha(x+h))h)_i$. Hence $\ell(\alpha_i(x+h)) - \ell(\alpha_i(x)) \ge 0$, and we obtain a contradiction.□

Theorem III.14. *Let $\mathcal{A}$ be a non empty compact set, and $A : \mathcal{A}^N \to \mathbb{R}^{N\times N}$, $f : \mathcal{A}^N \to \mathbb{R}^N$, be continuous functions satisfying (A1) and (A2). Then (3.1) has a unique solution x^*, and Howard's algorithm converges globaly super-linearly, i.e., $\lim_{k\to\infty} x^k = x^*$ and*

$$||x^{k+1} - x^*|| = o\Big(||x^k - x^*||\Big), \qquad \textit{as } k \to \infty.$$

Proof. Unicity of the solution and the fact that $x^k \ge x^{k+1}$ do hold as in the case when $\mathcal{A}$ is finite. Since $\alpha \to A(\alpha)^{-1}$ is continous, $A^{-1}(\alpha)f(\alpha)$ is also continous on $\mathcal{A}^N$ compact, and thus is bounded. Hence x^k is bounded, and we deduce that x^k converges to some element x^* of $\mathbb{R}^N$. Consider now F defined as in (3.3). We remark that F is C-lipschitz in the $||.||_\infty$ norm, where $C := \max_\alpha ||A(\alpha)||_\infty$. In particular, F is continuous and $F(x^*) = \lim_{k\to+\infty} F(x^k)$.

Using exactely the same arguments as in the proof of Proposition III.6 (b), we can prove that $F(x^*) = 0$.

There remains to show the super-linear convergence. Let us denote $\alpha(x)$ an optimal control associated to $F(x)$, i.e.

$$\alpha(x) := \arg\max_{\alpha\in\mathcal{A}^N}(A(\alpha)x - f(\alpha)),$$

$$\mathcal{A}^x := \left\{\alpha \in \mathcal{A}^N,\ A(\alpha)x - f(\alpha) = A(\alpha(x))x - f(\alpha(x))\right\}$$

(i.e the set of optimal controls associated to $F(x)$).

Now we consider $h_k := x^k - x^*$ and note that $\alpha^{k+1} := \alpha(x^k) = \alpha(x^* + h_k)$. By the previous Lemma, for $k \geq 0$, there exists $\alpha^{k,*} \in \mathcal{A}^{x^*}$ such that

$$d(\alpha^{k+1}, \alpha^{k,*}) = d(\alpha^{x^*+h_k}, \alpha^{k,*}) \to 0 \text{ as } k \to \infty. \tag{3.8}$$

Using the convexity of $F(x)$ and the fact that $F(x^*) = 0$ we obtain $F(x^k) \geq A(\alpha^{k,*})(x^k - x^*)$. (Hence, by monotonicity of $A(\alpha^{k,*})$,

$$\begin{aligned} x^{k+1} &= x^k - A(\alpha^{k+1})^{-1}F(x^k) \\ &\leq x^k - A(\alpha^{k+1})^{-1}A(\alpha^{k,*})(x^k - x^*) \end{aligned}$$

and thus

$$0 \leq x^{k+1} - x^* \leq (I - A(\alpha^{k+1})^{-1}A(\alpha^{k,*}))(x^k - x^*).$$

By continuity of $A(\alpha)$ and by (3.8) we obtain $I - A(\alpha^{k+1})^{-1}A(\alpha^{k,*}) \to 0$, hence

$$0 \leq x^{k+1} - x^* \leq o(x^k - x^*).$$

This concludes the proof. □

Remark III.15. *Note that the proof is strongly dependent on the fact that F is convex.*

Remark III.16. *In [17] authors analyze convergence of Howard algorithm in a class of stationary, infinite-horizon Markov decision problem, which are discretized to compute an approximate solution. Under regularity assumptions on the value function, they proves that for piecewise linear interpolation, policy iteration algorithm converges superlinearly. The proof is essentially based on the equivalence between the Howard algorithm and a semi-smooth Newton method. Moreover, they obtain that the constants involved in this convergence odrer depend on the grid size of discretization. In [16], a superlinear convergence is proved for Howard algorithm, but under particular assumptions : it is supposed that the exact value functions are computed at each policy evaluation step. Furthermore the authors impose a Lipschitz order condition which is not easily verifiable.*

Remark III.17. *The result can be generalized to the problem*

$$\max_{\alpha\in\mathcal{A}_1\times\cdots\times\mathcal{A}_N}(A(\alpha)x - f(\alpha)) = 0.$$

where $\mathcal{A}_i$ are different compact set.

The extension work of Theorem III.14 in infinite dimension is in progress, [4].

4 Obstacle problems

In this section, we consider a general obstacle problem of the form

$$\max\{My - b; y - \psi\} = 0, \tag{4.1}$$

where M is a monotone matrix of dimension $N \times N$, $y \in \mathbb{R}^N$ is the solution, b and ψ in $\mathbb{R}^N$. Obstacle problems come from different problems : for example, in [12] (4.1) is the optimality system of a linearly constrained quadratic problem.

Following Lemma III.1, we can interpret (4.1) as a control problem (3.1), with the following notations :

(i) $\mathcal{A} = \{0, 1\}$,

(ii) $A_{ij}(\alpha) = M_{ij}$ if $\alpha_i = 0$, $A_{ij}(\alpha) = \delta_{ij}$ if $\alpha_i = 1$;

(iii) $f_i(\alpha) = b_i$ if $\alpha_i = 0$, $f_i(\alpha) = \psi_i$, if $\alpha_i = 1$.

Remark III.18. Applying Proposition III.6, we obtain the following results :

(a) In the case that $\mathcal{A} = \{0, 1\}$, if we will chose $\alpha_i^{k+1} = 0$ when $(A(0)y^{k+1} - f(0))_i = (A(1)y^{k+1} - f(1))_i$, then the control sequence $(\alpha^k)_k$ is a decreasing sequence.

(b) In the case that $\mathcal{A} = \{0, 1\}$, if we will chose $\alpha_i^{k+1} = 0$ when $(A(0)y^{k+1} - f(0))_i = (A(1)y^{k+1} - f(1))_i$, then the algorithm converges in no more than N iterations.

(c) If at an iteration k we have $y_i^k < \psi_i$, for a component i, then we will have $y_i^{\bar{k}} < \psi_i$, for all $\bar{k} \geq k$.

Remark III.19. Remark III.18(a) implies that if for a component i, $\alpha_i^k = 0$, then $\alpha_i^{k+1} = 0$. We can interpret this fact as follows : if $\alpha_k^i = 0$, then the obstacle term is not active for the component i at iteration k. The fact that also $\alpha_i^{k+1} = 0$ means that the obstacle term remains not active a iteration $k + 1$ for the component i. More in general we can say that if the obstacle becomes inactive at an iteration k, then it will remain always inactive.

4.1 Primal-Dual Active set strategy for obstacle problem

We recall here the obstacle problem

$$\max\{My - b; y - \psi\} = 0,$$

where M is a monotone matrix of dimension $N \times N$, $y \in \mathbb{R}^N$ is the solution, b and ψ in $\mathbb{R}^N$.

It it well known that, y is solution of (4.1) if and only if there exists $\lambda \in \mathbb{R}^N$ such that

$$\begin{cases} My + \lambda = b \\ y \leq \psi, \ \lambda \geq 0 \\ (\lambda, y - \psi) = 0 \end{cases} \tag{4.2}$$

The system (4.2) is equivalent to the following system :

$$\begin{cases} My + \lambda &= b \\ \mathcal{C}(y, \lambda) &= 0, \end{cases} \tag{4.3}$$

where

$$\mathcal{C}(y,\lambda) = \lambda - \max(0, \lambda + c(y-\psi)),$$

for $c > 0$, and we can also write that $\lambda = \mathbb{P}_{[0,+\infty)}(\lambda + c(y-\psi))$.

Following the ideas developed in [12], we can use $\mathcal{C}(y,\lambda) = 0$, as a prediction strategy ; i.e., given a current primal-dual pair (y,λ), the choice for the next active and inactive sets is given by

$$\mathcal{I} = \{i : \lambda_i + c(y-\psi)_i \le 0\} \quad \text{and} \quad \mathcal{AC} = \{i : \lambda_i + c(y-\psi)_i > 0\}.$$

This leads to the following Primal-Dual Active Set Algorithm (see [12]) :

Primal-Dual Active set algorithm
(i) Initialize y^0, λ^0. Set $k = 0$.
(ii) Set $\mathcal{I}_k = \{i : \lambda_i^k + c(y^k-\psi)_i \le 0\}$, $\mathcal{AC}_k = \{i : \lambda_i^k + c(y^k-\psi)_i > 0\}$.
(iii) Solve

$$\begin{array}{l} My^{k+1} + \lambda^{k+1} = b, \\ y^{k+1} = \psi \text{ on } \mathcal{AC}_k, \quad \lambda^{k+1} = 0 \text{ on } \mathcal{I}_k. \end{array}$$

(iv) Stop, or set $k = k+1$ and return to (ii).

Remark III.20. *In [12], authors prove that the above algorithm can be interpreted as a semismooth Newton method, and using this property, in [12, Theorem 3.1] they show that this algorithm converges superlinearly.*

Remark III.21. We refer to [2, 3] for application of Primal-Dual active set strategy and interior point method to solve constrained optimal control problem.

Remark III.18(a) implies that, when we apply Howard algorithm, if at an iteration k the control which is active is $\alpha_i^k = 0$, then for the next iterations we do not need to recompute α_i^{k+1}, because it takes always the value 0. With this property, we can establish the equivalence between Howard policy iteration algorithm and Primal-Dual Active Set Algorithm.

Moreover, the sets $\mathcal{AC}_k$ and $\mathcal{I}_k$ defined on the Primal-Dual active set algorithm, satisfy :

$$\mathcal{AC}_k = \{i : \lambda_i^k + c(y^k-\psi)_i > 0\}, \quad \text{and} \quad \mathcal{I}_k = \{i : \lambda_i^k + c(y^k-\psi)_i \le 0\}.$$

Proposition III.22. *The following statements are true :*
(a) $i \in \mathcal{AC}_k \Rightarrow \alpha_i^k = 1$.
(b) $i \in \mathcal{I}_k \Rightarrow \alpha_i^k = 0$.

Proof. (a) Let $i \in \mathcal{AC}_k$, then $\lambda_i^k > -c(y_i^k - \psi_i)$. By the definition of the primal dual active set algorithm, we have

$$\begin{array}{l} My^k + \lambda^k = b, \\ y^k = \psi \text{ on } \mathcal{AC}_{k-1}, \\ \lambda^k = 0 \text{ on } \mathcal{I}_{k-1}. \end{array}$$

The first equality implies $\lambda^k = b - My^k$ for all i, and then, for $i \in \mathcal{AC}_k$ we obtain

$$(My^k - b)_i < c(y_i^k - \psi_i).$$

We can choose $c = 1$, and then $(My^k - b)_i < y_i^k - \psi_i$, which implies $\alpha_i^k = 1$.

(b) Let $i \in \mathcal{I}_k$, then $\lambda_i^k \leq -c(y_i^k - \psi_i)$. By the definition of the primal dual active set algorithm, we have

$$\begin{aligned} &My^k + \lambda^k = b, \\ &y^k = \psi \ \text{ on } \mathcal{AC}_{k-1}, \\ &\lambda^k = 0 \ \text{ on } \mathcal{I}_{k-1}. \end{aligned}$$

The first equality implies $\lambda^k = b - My^k$ for all i, and then, for $i \in \mathcal{I}_k$ we obtain

$$(My^k - b)_i \geq c(y_i^k - \psi_i).$$

We can choose $c = 1$, and then $(My^k - b)_i \geq y_i^k - \psi_i$, which implies $\alpha_i^k = 0$. □

In Howard algorithm, we start with a control α_0, and by proposition III.22 it is equivalent to give sets $\mathcal{I}_0 = \{i : \alpha_i^0 = 0\}$, $\mathcal{AC}_0 = \{i : \alpha_i^0 = 1\}$. Then we compute y^1 such that

$$A(\alpha^0)y^1 - f(\alpha^0) = 0,$$

and this is equivalent to compute

$$\begin{aligned} &My^1 + \lambda^1 = b \\ &y^1 = \psi, \ \text{ on } \mathcal{AC}_0, \quad \lambda^1 = 0 \text{ on } \mathcal{I}_0. \end{aligned}$$

Then, the next step of the Howard algorithm is to compute

$$\alpha^1 = \text{argmax}_{\alpha \in \mathcal{A}^N}\{A(\alpha)y^1 - f(\alpha)\},$$

and this is equivalent to give sets $\mathcal{I}_1 = \{i : \alpha_i^1 = 0\}$, $\mathcal{AC}_1 = \{i : \alpha_i^1 = 1\}$.

Then, in a general step $k+1$, we have the following equivalences :

- Solve $A(\alpha^k)y^{k+1} - f(\alpha^k) = 0$ in Howard algorithm is equivalent to solve $My^{k+1} + \lambda^{k+1} = b$, $y_i^{k+1} = \psi_i$ on $\mathcal{AC}_k$, $\lambda_i^{k+1} = 0$ on $\mathcal{I}_k$, on the Primal-Dual Active set Strategy.
- Compute $\alpha^{k+1} = \text{argmax}\{A(\alpha)y^{k+1} - f(\alpha)\}$ in the Howard algorithm is equivalent to give sets $\mathcal{AC}_{k+1}$ and $\mathcal{I}_{k+1}$ in the Primal-Dual Active set Algorithm.

Remark III.23. *(a) In Howard algorithm, if i is such that $\alpha_i^k = 0$, for a k, then $\alpha_i^{\bar{k}} = 0$, for all $\bar{k} \geq k$.*

(b) In the Primal-Dual Active set algorithm, if $i \in \mathcal{I}_k$, then $i \in \mathcal{I}_{\bar{k}}$, for all $\bar{k} \geq k$.

In the previous remark, (a) and (b) give the same monotonicity property for the two algorithm, and then we have the equivalence.

Remark III.24. We have shown that for obstacle problem, with monotone matrix M, Primal-Dual Active set strategy and Howard algorithm are equivalent. Since Howard algorithm for obstacle problem converges in no more than N iterations, then we can say also that the Primal Dual Active-Set strategy converges in no more that N iterations.

Remark III.25. In [12], the Primal-Dual active set algorithm has been written when M is a P-matrix, i.e. for matrix M such that all its principal minors are positive. Moreover the authors prove the convergence of the sequence $(y^k, \lambda^k)_k$, and the monotonicity property of the algorithm, when M is an M-matrix, i.e. M is non-singular, $M_{ij} \leq 0$ for all $i \neq j$, and $M^{-1} \geq 0$.

Remark III.26. Until know we do not know if an active set algorithm can be written for problem as (3.1), and if there is any equivalence between this kind of algorithm and Howard algorithm for (3.1).

Remark III.27. In [12] Primal-Dual Active set method has been generalized to infinite dimension, i.e. $y \in L^2(\Omega)$, $\Omega \in \mathbb{R}^N$, a bounded domain or manifold, with Lipschitz boundary. Also for this case the authors prove the superlinear convergence.

5 The cascade approach

We now come back to the impulse problem :

$$\max\{\sup_{\alpha\in\mathcal{A}^{N_{tot}}}\{\mathcal{L}_h^\alpha V_h - f(\alpha)); V_h - \mathcal{M}_h V_h\} = 0. \tag{5.1}$$

We may define impulse control problems where the number n of allowed impulse times is limited. Let us denote by $(S_n)_n$ this sequence of problems, and by $(V_{hn})_n$ the sequence of associated values. Problem (S_0) is the one without impulse control, whereas for problem (S_n), if an impulse control is chosen, then the relevant value at the new point of the state space is $V_{h(n-1)}$. In other words, we may write the sequence of problems as follows :

(S_0) $\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_{h0} - f(\alpha)) = 0$, V_{h0} is the solution ;

(S_1) $\max\{\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_{h1} - f(\alpha)); V_{h1} - \mathcal{M}_h V_{h0}\} = 0$, V_{h1} is the solution ;

...

(S_n) $\max\{\sup_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha V_{hn} - f(\alpha)); V_{hn} - \mathcal{M}_h V_{h(n-1)}\} = 0$, V_{hn} is the solution,

where $\mathcal{L}_h^\alpha$ is the discretization of the $\mathcal{L}^\alpha$ operator, and it is a monotone matrix of dimension $N_{tot} \times N_{tot}$, N_{tot} is the number of points in the grid $\mathcal{O}_h$. We have that $f(\alpha)$ is a vector in $\mathbb{R}^{N_{tot}}$, and in particular $f_i(\alpha) = f^\alpha(i)$, $i \in \mathcal{O}_h$. Moreover, V_{hn}, and $\mathcal{M}_h V_{hn}$ are vectors in $\mathbb{R}^{N_{tot}}$, for all n.

We can say that the first equation is the discretization of the HJB-equation without obstacle, and for a general problem S_n, the obstacle depends on the solution of the previous problem S_{n-1}. Then, we can say that, for each problem S_n, the obstacle is "fixed", in the sense that it does not depend on V_{hn}.

Since we want to solve it using Howard algorithm, we use the same method as Lemma III.1 to write the obstacle problem as a control problem. Let α_{obs} a fictitious control, and we formulate S_n as follows :

$$\max\{\max_{\alpha\in\mathcal{A}^{N_{tot}}}(\mathcal{L}_h^\alpha(V_{hn}) - f(\alpha)); V_{hn} - f^{(n-1)}(\alpha_{obs})\} = 0, \tag{5.2}$$

where $f_i^{(n-1)}(\alpha_{obs}) = \mathcal{M}_h(V_{h(n-1)})_i$, for every i. Finally, this problem can be written in the form

$$\max_{\alpha\in\{\mathcal{A}\cup\alpha_{obs}\}^{N_{tot}}}(A_h(\alpha)(V_{hn}) - b^{n-1}(\alpha)) = 0, \tag{5.3}$$

where $(A_h(\alpha))_{ij} = \delta_{ij}$ if $\alpha_i = \alpha_{obs}$, $(A_h(\alpha))_{ij} = (\mathcal{L}_h^\alpha)_{ij}$ if $\alpha_i \in \mathcal{A}$, $b_i^{n-1}(\alpha_{obs}) = f_i^{n-1}(\alpha_{obs})$ and $b^{n-1}(\alpha) = f_i(\alpha)$ for all $\alpha \in \mathcal{A}$.

Then we can conclude that the matrix associated to the obstacle is the identity matrix, of dimension $N_{tot} \times N_{tot}$. So this matrix is monotone, invertible and bounded. Look now at the sequence of solutions $(V_{hn})_n$. We will see later on section that every V_{hn} is a bounded function.

The idea of the cascade is classical. The following estimates have been obtained in Ishii [13, 14] for HJB equations with state space equal to $\mathbb{R}^N$, and it turns out that these estimates also hold in a discrete state space setting, see [5, Theorem 3.2].

Theorem III.28. *The sequence $(V_{hn})_n$ is nonincreasing and bounded, and converges to V_h, solution of (5.1). Moreover, assume that f is nonnegative. If $\|V_{h0}\|_\infty \leq k$, then V_{h0} is solution of (5.1). Otherwise, for all $\mu \in (0,1)$ such that $\mu \leq k/\|V_{h0}\|_\infty$, we have that*

$$\max(V_{hn} - V_h) \leq \frac{(1-\mu)^n}{\mu}|V_{h0}|_\infty.\Box \tag{5.4}$$

Remark III.29. (i) Since adding a constant c to f amounts to add c/λ to V_h, assuming that f is nonnegative is not restrictive. The best estimate is of course obtained for $\mu = k/\|V_{h0}\|_\infty$.
(ii) The above estimate is constructive in the sense that the policy algorithm computes an upper estimate c_0 of V_{h0}, and hence, of $\|V_{h0}\|_\infty$ whenever f is nonnegative.

So we may assume that such a μ is known when designing of algorithms. In that case (5.4) guarantees that, for computing V_h with uniform precision $\varepsilon > 0$, it is enough to perform n iterations, where n is the smallest integer such that $(1-\mu)^n c_0 \leq \mu\varepsilon$, i.e., $n \geq \log(\mu\varepsilon/c-0)/(\log(1-\mu))$. This estimate, however, does not take into account the fact that in practice functions V_{hn} are only approximately computed.

So let us present an *implementable* algorithm based on the idea of the cascade. The inner iterations consists in the policy iterations :

An implementable Cascade type Algorithm

Data $\mathcal{A}$,$A_h(\alpha)$, $f(\alpha)$, $\mathcal{M}$, a sequence m_n of positive integer numbers, α, initial policy ; $k := 0$.

Init Perform m_0 iterations of the policy algorithm for solving (S_0) ; the output is the upper estimate $\tilde{V}_{h0}$ of V_{h0}.

Loop For $k = 1, 2, \ldots$, define the problem (with unknown V_{hn})

$$\max_{\alpha\in\{\mathcal{A}\cup\{\alpha_{obs}\}\}^{N_{tot}}}\{(A_h(\alpha_h)V_{hn} + f^\alpha); V_{hn} - \mathcal{M}_h\tilde{V}_{h(n-1)}\} = 0.$$

perform m_n iterations of the policy algorithm for solving the obstacle problem (S_n) ; the initial policy is the last one obtained in the last step of the previous iteration ; the output is the upper estimate $\tilde{V}_{hn}$ of V_{hn}.

End

If we take $m_n = +\infty$ in the above algorithm (which then is no more implementable) we recover the cascade. Another extreme choice is to take $m_n = 1$ for all n, i.e., to update the values after impulsion after each inner iteration. As shows the following theorem, this is in fact the most efficient case, if updating the impulse term is cheap.

In the theorem below we use a counter ℓ be the counter of the inner iterations of the algorithm. This is the total number of inner iterations (i.e., of policy iterations) that have been done, i.e., at the inner step i of iteration k, we have

$$\ell = i + \sum_{j=0}^{k-1} m_j. \tag{5.5}$$

Theorem III.30. *Let ℓ be the counter of the inner iterations of the algorithm. Then the corresponding value that we will denote V_{hn} is a nondecreasing function of the sequence m_n. In other words, if the cost of computing the impulse control is negligible w.r.t. the cost of policy iterations, the convergence is faster when the computation of the impulse control occurs more often.*

Proof. We observe that the solution of a policy iteration for the obstacle problem is a nonincreasing function of the obstacle itself, i.e., if for a given extended policy $\alpha \in \mathcal{A} \cup \{\alpha_{obs}\}$, we denote by V the solution of the equation

$$A_h(\alpha)V - f(\alpha) = 0, \tag{5.6}$$

with (in the spirit of lemma III.1)

$$\begin{aligned} &(A_h(\alpha)V)_i = (\mathcal{L}_h^{\alpha}V)_i \text{ if } \quad \alpha_i \in \mathcal{A}, \\ &(A_h(\alpha_{obs})V)_i = V_i, \quad (f(\alpha_{obs}))_i := (\mathcal{M}_h V)_i, \end{aligned}$$

then (as is well-known) V is a nondecreasing function of f, and hence, of the obstacle.

On the other hand, if we reduce m_n then we update earlier the impulse term, and so (by induction) it follows that the values computed in the next iterations will be lower. □

Remark III.31. The cascade approach has been used also in [9] to solve discrete impulse control problem. In this paper, every problem of the cascade is solved by applying Howard algorithm only for the points of the domain in which the obstacle is inactive, and to set the value function equal to the obstacle in the rest of the domain.

Remark III.32. We refer to [8] for an algorithm to solve general obstacle problems, under the assumption that the first operator is contractive and the second non-expansive. From the two operators, a unique operator is built using a partition of the domain. Finally the problem can be interpreted as a fixed point problem, and Howard policy iteration algorithm is used to solve it. In this case the policies are the partitions of the domain.

Control of the precision of the solution We consider the following question : given $\varepsilon > 0$, is it possible to guarantee that the previous implementable cascade algorithm computes an approximation of the true solution V_h with precision ε, in the ℓ^∞ norm ?

Remember that we denote by V_{hi}, $i \in \mathbb{N}$, the solution of the cascade problem. We may write $V_{h,i+1} = obs(V_{hi})$, for all $i \in \mathbb{N}$, where by $obs(f)$ we denote the solution of the obstacle problem, the obstacle being f. The first element of the sequence is V_{h0}, solution of the problem without obstacle.

We know (see e.g. [6, Prop. 5.6]) that the Howard algorithm, applied to the problem of computing V_{h0}, provides a sequence V_{h0}^k of nonincreasing upper estimates of V_{h0}, and that these estimates satisfy (using the ℓ^∞ norm ; the contraction coefficient β was introduced in section 2)

$$\|V_{h0}^{k+1} - V_{h0}\| \leq \beta \|V_{h0}^k - V_{h0}\|. \tag{5.7}$$

Since $V_{h0}^k \geq V_{h0}$, it follows that

$$\|V_{h0}^{k+1} - V_{h0}\| \leq \beta^k \|V_{h0}^0 - V_{h0}\| \leq \beta^k \max V_{h0}^0. \tag{5.8}$$

We say that V' is an ε-upper estimate of V if $V \leq V'$ and $\sup(V' - V) \leq \varepsilon$. By (5.8), for any $\varepsilon_0 > 0$, an ε_0-upper estimate $\tilde{V}_{h0}$ of V_{h0} can be computed in K_0 iterations, with

$$k_0 \sim \log_\beta \left(\frac{V_{h0}^0}{\varepsilon_0} \right) \tag{5.9}$$

(here by $k \sim \alpha$ we mean that k is the smallest integer not less than α).

We can now proceeed by induction. Let ε_i be a sequence of positive number, and set $e_n := \sum_{\ell \leq n} \varepsilon_\ell$. Assume that we have computed an e_n-upper bound of V_{hn}, denoted $\tilde{V}_{hn}$. Let us put $\hat{V}_{h,n+1} := obs(\tilde{V}_{hn})$. Since the mapping $obs(\cdot)$ is non expansive and non decreasing, we have that

$$V_{h,n+1} = obs(V_{hn}) \leq \hat{V}_{h,n+1} = obs(\tilde{V}_{hn}) \leq V_{h,n+1} + e_n. \tag{5.10}$$

In other words, $\hat{V}_{h,n+1}$ is a e_n-upper bound of $V_{h,n+1}$. After k steps of the Howard algorithm, we obtain (using the estimate of [14] for $V_{hn} - V_{h,n+1}$)

$$\begin{aligned} \|V_{h,n+1}^{k+1} - \hat{V}_{h,n+1}\| &\leq \beta^k \max(\tilde{V}_{hn} - \hat{V}_{h,n+1}) \\ &\leq \beta^k (e_n + \max(V_{hn} - V_{h,n+1})) \leq \beta^k (e_n + (1-\mu)^n |V_{h0}|). \end{aligned} \tag{5.11}$$

Therefore, we obtain $\tilde{V}_{h,n+1} - \hat{V}_{h,n+1} \leq \varepsilon_{n+1}$ is after at most k_{n+1} Howard iterations, where

$$k_{n+1} \sim \log_\beta \left(\frac{e_n + (1-\mu)^n |V_{h0}|}{\varepsilon_{n+1}} \right), \tag{5.12}$$

and then $\tilde{V}_{h,n+1} \leq V_{h,n+1} + e_{k+1}$.

On the other hand, given we know by Ishii's result [13, 14] that

$$V_{h,n} - V_h \leq a_n := \frac{(1-\mu)^n}{\mu}. \tag{5.13}$$

We deduce the following

Theorem III.33. *Given $\varepsilon > 0$, let $n \in, \mathbb{N}$ and ε_i, $i \leq n$, be positive numbers such that*

$$\varepsilon_0 + \cdots + \varepsilon_n + a_n \leq \varepsilon. \tag{5.14}$$

Then by performing k_i iterations of the Howard algorithm at step $1 \leq i \leq n$, where k_i is given by (5.9) for $i = 0$, and by (5.12) for $1 \leq i \leq n$, we obtain an ε-upper bound of V_h.

Remark III.34. *Let L be the ratio between the costs of computing the impulse term and a Howard iteration. It can be greater or less than one, depending on the examples. The total cost for computing the ε-upper bound of V_h, by the implementable Cascade algorithm, is at most*

$$C_\varepsilon := nL + \log_\beta\left(\frac{V_{h0}^0}{\varepsilon_0}\right) + \sum_{i=0}^{n-1} \log_\beta\left(\frac{e_i + (1-\mu)^i |V_{h0}|}{\varepsilon_{i+1}}\right) \tag{5.15}$$

For given n and ε we may minimize this expression w.r.t. $\varepsilon_0, \cdots, \varepsilon_n$, subject to the equality constraint $\varepsilon_0 + \cdots + \varepsilon_n \leq \varepsilon - a_n$. Setting $\gamma := |V_{h0}|$, we see that

$$C_\varepsilon = -\sum_{i\leq n} \log \varepsilon_i + \sum_{i \leq n} \log(e_i + (1-\mu)^i \gamma) + \text{ constant term} \tag{5.16}$$

By Lagrange's rule we have that for some Lagrange multiplier $\lambda \geq 0$

$$-\lambda = \frac{\partial C_\varepsilon}{\partial \varepsilon_i} = -\frac{1}{\varepsilon_i} + \sum_{j=i}^{n} \frac{1}{e_i + (1-\mu)^i \gamma}. \tag{5.17}$$

This shows in particular that ε_i should be a decreasing function of i. It is possible to solve numerically these equations in order to compute the best estimates od the ε_i for given n, and then to compute the best n.

6 Numerical Examples

We consider now some numerical examples that we have studied to make a comparison between the case $m_n = +\infty$ for all n, and the case $m_n = 1$ for all n.

Consider the following equation in $\mathbb{R}^2$:

$$\min(-\Delta V(x,y) - f(x,y); V(x,y) - \gamma \mathcal{M} V(x,y)) = 0,$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ and $\mathcal{M}$ will be detailed later.

We localize the problem on the domain $\mathcal{O} = (0, x_{max}) \times (0, y_{max})$, assuming the following boundary conditions :

$$V(0,0) = 0, \quad V(x,0) = 0, \quad V(0,y) = 0, \tag{6.1}$$
$$V(x_{max}, y) = 0, \quad V(x, y_{max}) = 0, \quad V(x_{max}, y_{max}) = 0. \tag{6.2}$$

We want to solve the localized problem by using the two algorithms that we have presented in the previous section.

6.1 The discretization

Let $\mathcal{O}_h$ a discretization grid on $\mathcal{O}$. Let N_x, N_y the number of discretization points on the x and the y axis respectively, and let $h_x = x_{max}/(N_x+1)$, $h_y = y_{max}/(N_y+1)$ the respective discretization steps. The total number of points in the grid is $N_{tot} = N_x \times N_y$. We use a Finite Differences Scheme to approximate Laplacian term :

$$\begin{aligned} \Delta^h V_h(x,y) &= \frac{V_h(x+h_x,y) - 2V_h(x,y) + V_h(x-h_x,y)}{h_x^2}, \\ &+ \frac{V_h(x,y_{h+y}) - 2V_h(x,y) + V_h(x,y-h_y)}{h_y^2}. \end{aligned}$$

For the points on the boundary we use conditions (6.1), (6.2). For simplicity of notations, we will denote $\mathcal{O}_h = \{1, \dots, N_{tot}\}$.

First Example In this first example we consider $\gamma = 0.5$,

$$f(x,y) = 1, \quad \forall\, x, y \in \mathcal{O}$$
$$\mathcal{M}V(x,y) = \int_{(0,x_{max})\times(0,y_{max})} u(x,y)dxdy. \tag{6.3}$$

We have that $\mathcal{M}$ verifies the following properties :

- $u \le v \Rightarrow \mathcal{M}u \le \mathcal{M}v$,
- $||\mathcal{M}u - \mathcal{M}v|| \le ||u - v||$.

For this example, in the case $x_{max} = y_{max} = 1$, we can compute the exact solution which is :

$$V(x,y) = \frac{\gamma}{12(1-\gamma)} + \frac{1}{4}(x(1-x) + y(1-y)).$$

Consider $\mathcal{M}_h V_h$ the discretization of $\mathcal{M}V$:

$$\mathcal{M}_h V_h = \sum_{i \in \mathcal{O}_h} (V_h)_i h_x h_y. \tag{6.4}$$

For the discretization of f, we have $f_i = 1$, for all $i = 1, \dots, N_{tot}$.

Then, the discrete equation can be written as

$$\min\{A_h V_h - f; V_h - \mathcal{M}_h V_h\} = 0, \tag{6.5}$$

with the boundary conditions, where $A_h V_h = -\Delta^h V_h$.

In all the sequel, we will note by Howard-1 the algorithm which solves every problem of the cascade until the convergence of the Howard algorithm, i.e. it solves each (S_n) generating a sequence $(V_{hn}^k)_k$, and it stops when $||V_{hn}^k - V_{hn}^{k+1}|| \leq 10^{-10}$. Moreover we have a condition to stop the cascade : $||V_{hn} - V_{h(n+1)}|| \leq 10^{-8}$. We denote by Howard-2 the algorithm which does only one iteration of the Howard algorithm, for every problem of the cascade.

Numerical Results Since we have set some boundary conditions, our solution has a discontinuity on the first points of the grid near to the boundary. In fact, we have that on the boundary $V_h = 0$, but in the first point of the grid which is at a distance h_x from $x = 0$, or $x = x_{max}$, or at distance h_y from $y = 0$ or $y = y_{max}$ we have that the value of V_h is different from zero, and the jump depends on h.

$N_{tot} = N_x \times N_y$	$1600 = 40 \times 40$		$4900 = 70 \times 70$	
	# It.	CPU (sec)	# It.	CPU (sec)
Howard-1	26×3	21.93	27×3	505.15
Howard-2	32	8.34	36	211.20

$N_{tot} = N_x \times N_y$	$1600 = 40 \times 40$	$4900 = 70 \times 70$
	$V_h - V$	$V_h - V$
Howard-1	$1.24e-1$	$1.18e1$
Howard-2	$1.24e-1$	$1.18e-1$

In Figure III.1 we have the difference between the exact solution and the numerical solution in the case $N_{tot} = 4900$.

Second Example In this second example we consider $\gamma = 0.5$,

$$f(x,y) = 1, \quad \forall\, x, y \in \mathcal{O},$$
$$\mathcal{M}u(x,y) = \sup_{(x,y)\in\mathcal{O}} u(x,y). \tag{6.6}$$

Also in this case we have an exact solution :

$$V(x,y) = \frac{\gamma}{8(1-\gamma)} + \frac{1}{4}(x(1-x) + y(1-y)).$$

$N_{tot} = N_x \times N_y$	$1600 = 40 \times 40$		$4900 = 70 \times 70$	
	# It.	CPU (sec)	# It.	CPU (sec)
Howard-1	30×3	22.51	30×3	484.48
Howard-2	37	9.25	39	224.59

$N_{tot} = N_x \times N_y$	$1600 = 40 \times 40$	$4900 = 70 \times 70$
	$V_h - V$	$V_h - V$
Howard-1	$1.26e-1$	$1.21e-1$
Howard-2	$1.26e-1$	$1.21e-1$

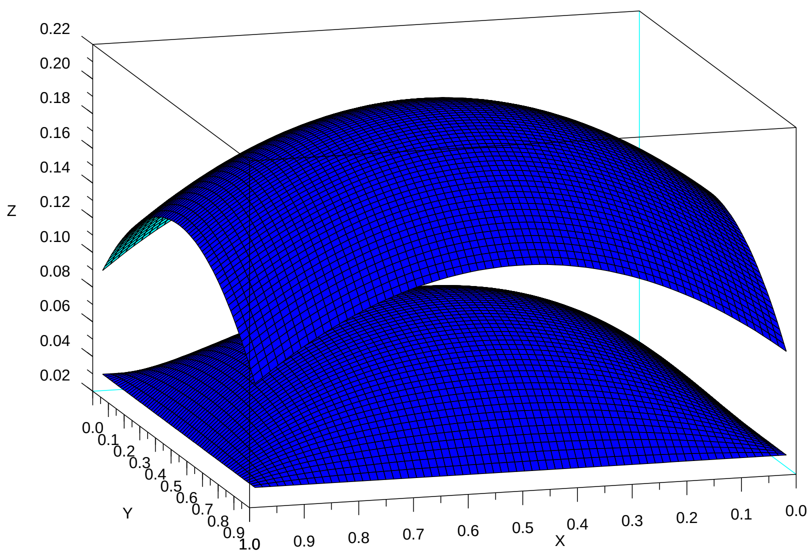


FIG. III.1 – Value function on the first example, for $N_{tot} = 4900$.

In the case of $N_{tot} = 4900$, the value function is represented by the picture in Figure III.2. Moreover, in Figure III.3 we have the difference between the exact solution and the numerical solution in the case $N_{tot} = 4900$.

Third example In this third example we consider $\gamma = 0.5$

$$f(x,y) = 3x^2 + 5y^2, \quad \forall\, x, y \in \mathcal{O},$$
$$\mathcal{M}u(x,y) = \sup_{(x,y)\in\mathcal{O}} u(x,y). \tag{6.7}$$

$N_{tot} = N_x \times N_y$	$1600 = 40 \times 40$		$4900 = 70 \times 70$	
	# It.	CPU (sec)	# It.	CPU (sec)
Howard-1	31×3	23.53	31×3	502.45
Howard-2	39	9.48	43	244.06

In the case of $N_{tot} = 4900$, the value function is represented by the picture in Figure III.4.

Bibliographie

[1] M. Akian. Analyse de l'algorithme multigrille FMGH de résolution d'équations d'Hamilton-Jacobi-Bellman. In A. Bensoussan and J.-L. Lions, editors, *Analysis and*

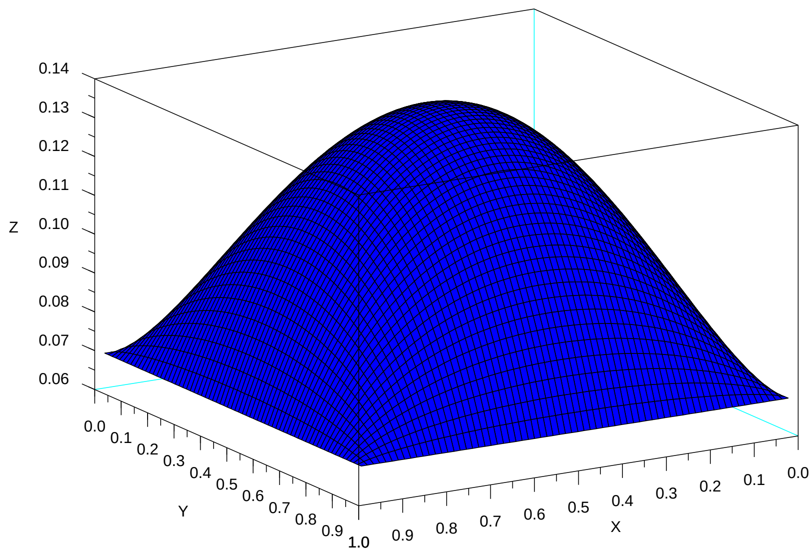


FIG. III.2 – Value function on the second example, for $N_{tot} = 4900$.

optimization of systems (Antibes, 1990), volume 144 of *Lecture Notes in Control and Information Sciences*, pages 113–122. Springer Verlag, Berlin, 1990.

[2] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch. A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems. *SIAM Journal on Optimization*, 11 :495–521 (electronic), 2000.

[3] M. Bergounioux, M. Hintermüller, and K. Kunisch. Primal dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, 37 :1176–1194, 1999.

[4] O. Bokanowski, S. Maroso, and H. Zidani. A superlinear convergence result for howard algorithm. Work in progress.

[5] J.F. Bonnans, S. Maroso, and H. Zidani. Error estimates for a stochastic impulse control problem. Technical Report 5606, INRIA, 2005. "http ://www.inria.fr/rrrt/rr-5606.html".
To appear in Applied Math. Optimization.

[6] J.F. Bonnans and P. Rouchon. *Commande et optimisation de systèmes dynamiques*. Les Editions de l'Ecole Polytechnique, Palaiseau, France, 2005.

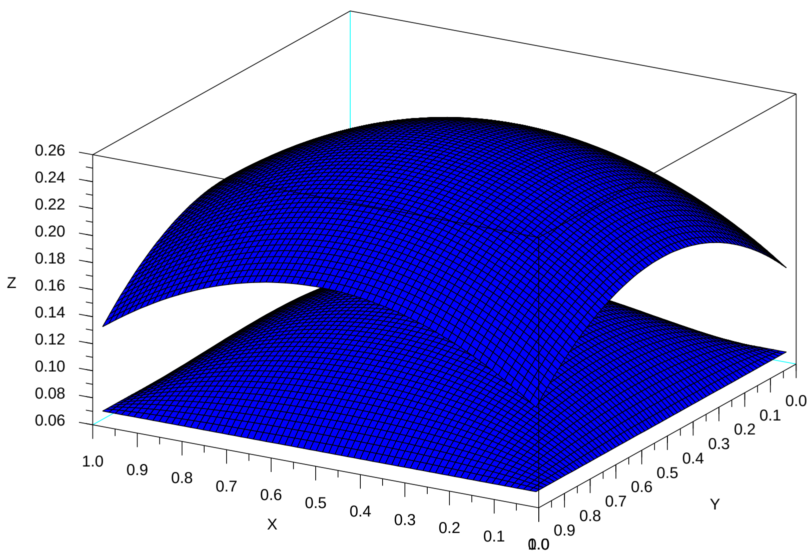


Fig. III.3 – On the top we have the exact solution of the second example, and on the bottom the numerical solution. The error depends on the steps of discretization h_x and h_y.

[7] J.F. Bonnans and H. Zidani. Characterization of consistency of some numerical schemes for the stochastic HJB equation. In J.L. Menaldi, E. Rofman, and A. Sulem, editors, *Optimal Control and PDE - Innovations et Applications*. IOS Press, Amsterdam, 2000.

[8] J.-Ph. Chancelier, B. Messaoud M., and A. Sulem. A policy iteration algorithm for fixed point problems with nonexpansive operators. *Preprint Cermics*, 2004.

[9] J.-Ph. Chancelier, B. Øksendal, and A. Sulem. Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. *Tr. Mat. Inst. Steklova*, 237(Stokhast. Finans. Mat.) :149–172, 2002.

[10] J. Cochet-Terrasson and S. Gaubert. Policy iteration algorithm for shortest path problems. 2000.

[11] R. L. V. Gonzalez and C. A. Sagastizabal. Un algorithme pour la résolution rapide d'équations discrètes de Hamilton-Jacobi-Bellman. *C. R. Acad. Sci. PAris*, 311, Série I :45–50, 1990.

[12] M. Hintermuller, K. Ito, and K. Kunish. The primal-dual active set strategy as a semismooth newton method. *SIAM J. Optim.*, 13 :865–888, 2002.

[13] K. Ishii. Viscosity Solutions of Non Linear Second Order Elliptic PDEs associated with Impulse Control Problems. *Funkcialaj Ekvacioj*, 36 :123–141, 1993.

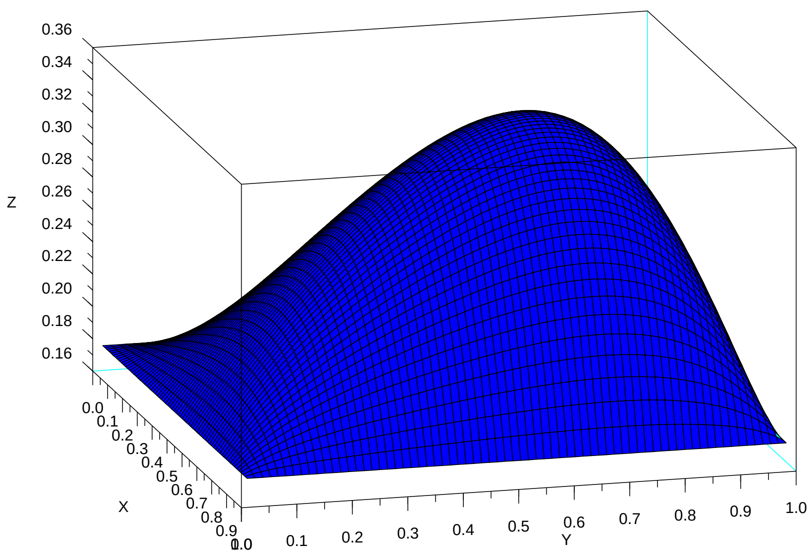


FIG. III.4 – Value function in the third example for $N_{tot} = 4900$.

[14] K. Ishii. Viscosity Solutions of Non Linear Second Order Elliptic PDEs associated with Impulse Control Problems II. *Funkcialaj Ekvacioj*, 38 :297–328, 1995.

[15] H.J. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of mathematics*. Springer, New York, 2001. Second edition.

[16] M.L. Puterman and S.L. Brumelle. On the convergence of policy iteration in stationary dynamic programming. *Math. Oper. Res.*, 4 :60–69, 1979.

[17] M.S. Santos and J. Rust. Convergence properties of policy iteration. *SIAM J. Control Optim.*, 42 :2094–2115, 2004.

CHAPITRE IV

Numerical approximation for super-replication problems

Numerical approximation for super-replication problem under gamma constraints [1]

1 Introduction

In a financial market, consisting in a non-risky asset and some risky assets, people are interested to study the minimal initial capital needed in order to super-replicate a given contingent claim, under gamma constraints. Many authors have studied this problem in different cases and with different constraints : for example, see [12, 17], for problems in dimension 1, [8] for problems in dimension 2, and [18, 10] for problems in a general dimension d. In all these papers, the authors characterize the super-replication price as the viscosity solution of an HJB-equation with terminal and boundary conditions. In a particular case, the dual formulation of the super-replication problem leads to a standard form of optimal stochastic control problem [8].

In this paper we study numerically an HJB-equation coming from the super-replication problem in dimension 2. We discretize the HJB equation using the Generalized Finite Differences scheme [6, 7], then we study existence and uniqueness of the discrete solution. Finally we prove the convergence of the numerical solution to the viscosity solution. In particular, we are interested on the HJB equation which comes from the two dimensional dual problem introduced in [8] :

$$\vartheta(t,x,y) = \sup_{(\rho,\xi)\in\mathcal{U}} \mathbb{E}\left[g\left(X^{\rho,\xi}_{t,x,y}(T)\right)\right], \tag{1.1}$$

where (ρ,ξ) are valued in $[-1,1]\times(0,\infty)$, the process $(X^{\rho,\xi}_{t,x,y}, Y^{\rho,\xi}_{t,y})$ is a 2-dimensional positive process which evolves according to the stochastic dynamics (2.1), and g is a payoff function. The main difficulty of the above problem is due to the non-boundness of the control set, this fact implies that the Hamiltonian associated to (1.1) is not bounded, and numerical approximation for such a problem becomes more complicate.

In the literature, problems with unbounded control have been studied by many authors (for example, [1, 9]). In all these cases, the authors decide to truncate the set of controls to make it bounded. This truncation simplifies the numerical analysis of the problem. However, there is no theoretical result justifying this truncation.

In this paper we do not truncate the set of controls, because we find a particular form of our HJB equation which leads us to avoid the difficulty of unbounded control. In fact, our HJB equation can be reformulated in the following way

$$\Lambda^-(J(t,x,y,D\vartheta(t,x,y),D^2\vartheta(t,x,y))) = 0,$$

where J is a symetric matrix differential operator associated to the Hamiltonian, and where $\lambda^-(J)$ means the smallest eigenvalue of the matrix operator J. J does not depend on the

[1] Joint work with O. Bokanowski, B. Bruder and H. Zidani

control, but when we look for the first time at this equation, it seems that it is very difficult to treat. From standard computations on algebra, we rewrite the smallest eigenvalue as follows :

$$\Lambda^{-}(J) = \min_{\|\alpha\|=1} \alpha^T J \alpha,$$

where $\alpha \in \mathbb{R}^2$. Then we have transformed our problem into a bounded control problem, and now the numerical analysis is possible.

The structure of the paper is the following : in Section 2 we present the problem and the associated HJB-equation. We prove boundary conditions satisfied by the value function, then the existence, uniqueness and Lipschitz property of the viscosity solution. In Section 3 we consider the discretization of the HJB equation, and recall the main properties of the Generalized Finite Differences Scheme and we prove the consistency of this scheme. In section 4, we prove existence and uniqueness of a bounded discrete solution, and finally in Section 5 we prove the convergence of the numerical approximation.

2 Problem formulation and PDE

Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a probability space, and $T > 0$ be a fixed finite time horizon. Let $\mathcal{U}$ denotes the set of all $\mathcal{F}_t$-measurable processes $(\rho, \zeta) := \{(\rho(t), \zeta(t)); 0 \leq t \leq T\}$ with values in $[-1, 1] \times \mathbb{R}_+$:

$$\mathcal{U} := \Big\{(\rho, \zeta) \text{ valued in } [-1, 1] \times (0, +\infty) \text{ and } \mathcal{F}_t\text{-measurable} \mid \int_0^T \zeta_t^2 \, dt < +\infty\Big\}.$$

For a given control process (ρ, ζ), and an initial data $(t, x, y) \in (0, T) \times \mathbb{R}^+ \times \mathbb{R}^+$, we consider the controlled 2-dimensional positive process $(X^{\rho,\zeta}_{t,x,y}, Y^{\rho,\zeta}_{t,y})$ evolving according to the stochastic dynamics :

$$dX^{\rho,\zeta}_{t,x,y}(s) = \sigma(s, Y^{\rho,\zeta}_{t,y}(s)) X^{\rho,\zeta}_{t,x,y}(s) dW^1_s, \qquad s \in (t, T) \tag{2.1a}$$

$$dY^{\rho,\zeta}_{t,y}(s) = -\mu(s, Y^{\rho,\zeta}_{t,y}(s))ds + \zeta(s) Y^{\rho,\zeta}_{t,y}(s) dW^2_s, \quad s \in (t, T) \tag{2.1b}$$

$$\langle dW^1_s, dW^2_s \rangle = \rho(s), \qquad \text{a.e } s \in (t, T) \tag{2.1c}$$

$$X^{\rho,\zeta}_{t,x,y}(t) = x \ , \ Y^{\rho,\zeta}_{t,y}(t) = y, \tag{2.1d}$$

where W^1_s and W^2_s denote the standard Brownian motion defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The volatility σ and the cash flow μ satisfy the following assumptions :

(A1) $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}^+$ is a positive function, such that σ^2 is Lipschitz. For every $t \in [0, T]$, $\sigma(t, 0) = 0$ (typically $\sigma(t, y) = \sqrt{y}$).

(A2) $\mu : (0, T) \times \mathbb{R}^+ \to \mathbb{R}^+$ is a positive Lipschitz function, with $\mu(t, 0) = 0$ for every $t \in [0, T]$.

Assumptions (A1) and (A2) ensure that the stochastic dynamic system (2.1) has a unique strong solution.

The variables $X^{\rho,\zeta}_{t,x,y}$ and $Y^{\rho,\zeta}_{t,y}$ describe two different assets from a financial market. The first asset $X^{\rho,\zeta}_{t,x,y}$ is risky, while the second one $Y^{\rho,\zeta}_{t,y}$ distributes an instaneous cash flow $\mu(s, Y^{\rho,\zeta}_{t,y}(s))$, and its price is linked to the asset $X^{\rho,\zeta}_{t,x,y}$ by the means of volatility $\sigma(s, Y^{\rho,\zeta}_{t,y}(s))$.

Remark IV.1. *It is important to remark that the evolution of the variable $Y^{\rho,\zeta}_{t,y}$ does not depend on $X^{\rho,\zeta}_{t,x,y}$.*

Now consider a function $g : \mathbb{R}^+ \to \mathbb{R}$. Different assumptions will be made on g :

- **(A3)** g is a bounded Lipschitz function. Let $M_0 > 0$ such that : $\|g\|_\infty \leq M_0$.
- **(A4)** The function $f : z \to g(e^z)$ is Lipschitz continuous.
- **(A5)** $g \in \mathcal{C}^2(\mathbb{R}^+ \to \mathbb{R})$. The functions $x \to xg'(x)$ and $x \to x^2 g''(x)$ are bounded.

Consider the following stochastic control problem $(\mathcal{P}_{t,x,y})$ with its associated value function ϑ defined by :

$$\vartheta(t, x, y) := \sup_{(\rho,\zeta)\in\mathcal{U}} \mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right]. \tag{2.2}$$

Assumption (A3) leads us to obtain a bounded and Lipschitz value function ϑ of (2.2). Assumption (A4) will be usefull to prove some boundary conditions satisfied by ϑ (see section 2.1).

This control problem can be interpreted in [8] in the following sense : A trader wants to sell an European option of terminal payoff $g(X_T)$ without taking any risk. Hence we use a superreplication framework. The underlying X of the option is a risky asset, for axample a stock, an index or a mutual fund. Unfortunately, in several cases, the volatility σ of the underlying X exhibits large random changes across time. Therefore, the Black-Scholes model fails to capture the risks of the trader. One must then use a model that features stochastic volatility. It is known that in this framework, the superreplication problem has a trivial solution (see [12]). For example, if the volatility has no a priori bound, the superreplication price is the concave envelope of the payoff $g(X(T))$, and the hedging strategy is static. To obtain more accurate prices, we introduce another financial asset Y whose price is linked to the volatility of the underlying X. For example, we can consider a variance swap which continuously pays the instantaneous variance of X (hence $\mu(t, Y) = \sigma^2$). For the sake of simplicity we assume that the price of Y and the volatility of X are driven by a single common factor (hence $\sigma = \sigma(t, Y)$). If the parameters ζ and ρ of the dynamics of the price Y were known, and if there were no transaction costs for Y, the super-replication price would simply be $\mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right]$. But we face two problems :

- The parameters (ζ, ρ) of the dynamics of Y are likely to be random and difficult to measure. As there is no a priori bound to these parameters, the super-replication price is given be the supremum of $\mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right]$ over all adapted processes ζ, ρ (see [14]).
- The asset Y is likely to introduce transaction costs, and hence the trader cannot buy and sell an infinite amount of asset Y during the period $[0, T]$. It is proved in [8] that the super-replication price of $g(X(T))$ under the constraint of a finite amount of transactions involving Y during $[0, T]$ is given by the value function of problem (2.2). See also [17, 18] for a similar approach.

Denote by $\mathcal{M}_2$ the set of symmetric 2×2 matrices. The Hamiltonian function is defined by : for $t \in (0, T)$, $x, y \in \mathbb{R}^+$, $p = (p_1, p_2)^\intercal \in \mathbb{R}^2$, and $Q \in \mathcal{M}_2$:

$$H\left(t, x, y, p, Q\right) := \inf_{(\zeta,\rho)\in\mathbb{R}_+\times[-1,1]} \left\{ \mu(t, y)p_2 - \frac{1}{2}\mathrm{tr}\left(a(t, x, y, \zeta, \rho) \cdot Q\right) \right\}, \tag{2.3}$$

and the covariance matrix a is given by :

$$a(t, x, y, \zeta, \rho) := \begin{pmatrix} \sigma^2(t, y)x^2 & \rho\zeta\sigma(t, y)x \\ \rho\zeta\sigma(t, y)x & \zeta^2 \end{pmatrix}.$$

Now we look for a characterization of ϑ as a viscosity solution of an HJB equation. In a formal way, we get that ϑ satisfies the following PDE :

$$-\frac{\partial\vartheta}{\partial t} + H\left(t, x, y, D\vartheta, D^2\vartheta\right) = 0 \quad (t, x, y) \in (0, T) \times (0, +\infty) \times (0, +\infty). \tag{2.4}$$

However, we will prove in Proposition 2.3 that the precise HJB equation satisfied by ϑ in the viscosity sens is

$$\Lambda^-\begin{pmatrix} -\frac{\partial\vartheta}{\partial t} + \mu(t, y)\frac{\partial\vartheta}{\partial y} - \frac{1}{2}\sigma^2(t, y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t, y)x\frac{\partial^2\vartheta}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t, y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2} \end{pmatrix} = 0, \tag{2.5}$$

where $\Lambda^-(A)$ denotes the smallest eigenvalue of a given symmetric matrix A. We first prove that ϑ is a discontinuous viscosity solution of (2.5). We will see later on that, under (A1), ϑ is continuous thanks to a comparison principle, and even Lipschitz continuous when assumptions (A3)-(A5) hold.

First, it is easy to see that the infimum in (2.3) can only be achieved for $\rho = \pm 1$. Hence denoting ζ as $\rho\zeta$, one can see that the Hamiltonian can be rewritten as :

$$H\left(t, x, y, p, Q\right) = \inf_{\zeta\in\mathbb{R}} \left\{ \mu(t, y)p_2 - \frac{1}{2}\mathrm{tr}\left(a(t, x, y, \zeta) \cdot Q\right) \right\}, \tag{2.6}$$

where, this time, there is only one control variable ζ taking values on the whole real line, and the covariance matrix a is defined by :

$$a(t, x, y, \zeta) = \begin{pmatrix} \sigma^2(t, y)x^2 & \zeta\sigma(t, y)x \\ \zeta\sigma(t, y)x & \zeta^2 \end{pmatrix}.$$

By elementary techniques, the minimization over ζ, in (2.6) gives :

$$H(t,x,y,p,Q) = -\infty \qquad \text{if } Q_{22} > 0, \tag{2.7a}$$
$$\text{or} \quad Q_{22} = 0 \text{ and } \sigma(t,y)xQ_{12} \neq 0, \tag{2.7b}$$
$$H(t,x,y,p,Q) \in \mathbb{R}, \qquad \text{otherwise.} \tag{2.7c}$$

Remark IV.2. *For this particular problem, it is not possible to find a continuous function* $G : [0,T] \times \mathbb{R}^2 \times \mathbb{R}_+^2 \times \mathcal{M}_2 \to \mathbb{R}$ *such that*

$$H(t,x,y,p,Q) > -\infty \Leftrightarrow G(t,x,y,p,Q) \geq 0.$$

Hence we can not use arguments introduced in [16] to define a precise notion of solution for equation (2.4).

For $t \in (0,T)$, $x, y \in \mathbb{R}^+$, $r \in \mathbb{R}$, $p = (p_1,p_2)^T \in \mathbb{R}^2$ and $Q \in \mathcal{M}_2$, introduce

$$J(t,x,y,r,p,Q) := \begin{pmatrix} -r + \mu(t,y)p_2 - \frac{1}{2}\sigma^2(t,y)x^2 Q_{11} & -\frac{1}{2}\sigma(t,y)xQ_{12} \\ -\frac{1}{2}\sigma(t,y)xQ_{12} & -\frac{1}{2}Q_{22} \end{pmatrix}.$$

With straightforward computations we obtain the following result.

Lemma 2.1. *For* $t \in (0,T)$, $x, y \in \mathbb{R}^+$, $r \in \mathbb{R}$, $p = (p_1,p_2)^T \in \mathbb{R}^2$ *and* $Q \in \mathcal{M}_2$, *the following assertions hold :*

(i) $-r + H(t,x,y,p,Q) \geq 0 \Leftrightarrow \Lambda^-(J(t,x,y,r,p,Q)) \geq 0.$
(ii) $-r + H(t,x,y,p,Q) \geq 0 \Rightarrow -Q_{22} \geq 0.$
(iii) $-r + H(t,x,y,p,Q) = 0 \Rightarrow \Lambda^-(J(t,x,y,r,p,Q)) = 0.$
(iv) $\Lambda^-(J(t,x,y,r,p,Q)) > 0 \Rightarrow -r + H(t,x,y,p,Q) > 0.$

Now, for a function $u : [0,T] \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$, we define the upper (resp. lower) semicontinuous envelope u^* (resp. u_*) of u by : for $t \in [0,T), x, y \in (0,+\infty)$,

$$u^*(t,x,y) = \limsup_{\substack{(s,w,z)\to(t,x,y) \\ s\geq 0, w,z\in(0,+\infty)}} u(s,w,z),$$
$$u_*(t,x,y) = \liminf_{\substack{(s,w,z)\to(t,x,y) \\ s\geq 0, w,z\in(0,+\infty)}} u(s,w,z).$$

With these definitions, we can give the sens of viscosity solution of (2.5), according to [2, 3, 11].

Definition 2.2. *(i) u is a discontinuous viscosity subsolution of* (2.5) *iff for any* $(\hat{t},\hat{x},\hat{y}) \in [0,T) \times (0,+\infty)^2$, *and any* $\phi \in C^2\left([0,T) \times (0,+\infty)^2\right)$, *such that* $(\hat{t},\hat{x},\hat{y})$ *is a local maximum of* $u^* - \phi$ *:*

$$\Lambda^-(J(\hat{t},\hat{x},\hat{y}), \partial_t\phi(\hat{t},\hat{x},\hat{y}), D\phi(\hat{t},\hat{x},\hat{y}), D^2\phi(\hat{t},\hat{x},\hat{y}))) \leq 0.$$

(ii) u is a discontinuous viscosity super-solution of (2.5) *iff for any* $(\hat{t},\hat{x},\hat{y}) \in [0,T) \times (0,+\infty)^2$, *and any* $\phi \in C^2\left([0,T) \times (0,+\infty)^2\right)$, *such that* $(\hat{t},\hat{x},\hat{y})$ *is a local minimum of* $u_* - \phi$ *:*

$$\Lambda^-(J(\hat{t},\hat{x},\hat{y}), \partial_t\phi(\hat{t},\hat{x},\hat{y}), D\phi(\hat{t},\hat{x},\hat{y}), D^2\phi(\hat{t},\hat{x},\hat{y}))) \geq 0.$$

(iii) u is a discontinuous viscosity solution of (2.5) *iff it is both sub and a super solution.*

Theorem 2.3. *Under assumptions (A1)-(A2), the value function ϑ is a viscosity discontinuous solution of* (2.5) *:*

$$\Lambda^-\begin{pmatrix} -\frac{\partial\vartheta}{\partial t}+\mu(t,y)\frac{\partial\vartheta}{\partial y}-\frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2}\end{pmatrix}=0.$$

Moreover ϑ is a discontinuous viscosity super-solution of

$$-\frac{\partial^2\vartheta}{\partial y^2}\geq 0. \tag{2.8}$$

Proof. The proof is splitted on two parts : the super-solution property and the sub-solution property.
(a) Super-solution property. By a classical application of the Dynamic Programming Principle, as done in [15], we obtain that $\vartheta(t,x,y)$ is a viscosity super-solution of

$$-\frac{\partial\vartheta}{\partial t}+H(t,x,y,D\vartheta,D^2\vartheta)\geq 0.$$

Then, Lemma 2.1(i) implies that also

$$\Lambda^-(J(t,x,y,\partial_t\vartheta(t,x,y),D\vartheta(t,x,y),D^2\vartheta(t,x,y)))\geq 0,$$

and then ϑ is a viscosity super-solution of (2.5). Moreover, from Lemma 2.1, we have $-\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2}\geq 0$, and hence (2.8) is verified.
(b) Sub-solution property. Let φ be a smooth function, and let $(\bar t,\bar x,\bar y)$ be a strict maximizer of $\vartheta^*-\varphi$, such that

$$0=(\vartheta^*-\varphi)(\bar t,\bar x,\bar y).$$

Suppose that $(\bar t,\bar x,\bar y)$ belongs to the set $\mathcal{M}(\varphi)$ defined by :

$$\mathcal{M}(\varphi)=\{(t,x,y)\in[0,T)\times(0,+\infty)^2:\Lambda^-(J(t,x,y,\partial_t\varphi(t,x,y),D\varphi(t,x,y),D^2\varphi(t,x,y)))>0\}$$

Since $\mathcal{M}(\varphi)$ is an open set, then there exists $\eta>0$ such that

$$[0\wedge(\bar t-\eta),\bar t+\eta]\times\overline{B}_\eta(\bar x,\bar y)\subset\mathcal{M}(\varphi),$$

where $\overline{B}_\eta(\bar x,\bar y)$ denotes the closed ball centered in $(\bar x,\bar y)$ and with radius η. From Lemma 2.1(iii), if $(t,x,y)\in\mathcal{M}(\varphi)$, then

$$-\frac{\partial\varphi}{\partial t}(t,x,y)+H(t,x,y,D\varphi(t,x,y),D^2\varphi(t,x,y))>0.$$

Using the Dynamic Programming Principle and the same arguments that in [16, Lemma 3.1], we get that :

$$\sup_{\partial_p([0\wedge(\bar t-\eta),\bar t+\eta]\times\overline{B}_\eta(\bar x,\bar y))}(\vartheta-\varphi)=\max_{[0\wedge(\bar t-\eta),\bar t+\eta]\times\overline{B}_\eta(\bar x,\bar y)}(\vartheta^*-\varphi), \tag{2.9}$$

where $\partial_p([t_1,t_2]\times\overline{B}_\eta(\bar{x},\bar{y}))$ is the forward parabolic boundary of $[t_1,t_2]\times\bar{B}_\eta(\bar{x},\bar{y})$, i.e. $\partial_p([t_1,t_2]\times\overline{B}_\eta(\bar{x},\bar{y}))=[t_1,t_2]\times\partial\overline{B}_\eta(\bar{x},\bar{y})\cup\{t_2\}\times\overline{B}_\eta(\bar{x},\bar{y})$. However, since $(\bar{t},\bar{x},\bar{y})$ is a strict maximizer of $\vartheta^*-\varphi$, equality (2.9) leads to a contradiction. Therefore, $(\bar{t},\bar{x},\bar{y})\notin\mathcal{M}(\varphi)$, and the result follows. □

In our paper, we are interested by the numerical approximation of the value function ϑ. Although equation (2.5) has a rigorous meaning, the formulation with the smallest eigenvalue might seem to be more complicated than the setting of (2.4). Of course, one can be tempted to modify the hamiltonian in the following way : for $\zeta_{\max}>0$, replace H by

$$\tilde{H}(t,x,y,p,Q)=\min_{\zeta\in[-\zeta_{\max},\zeta_{\max}]}\left\{\mu(t,y)p_2-\frac{1}{2}\mathrm{tr}(a(t,x,y,\zeta)\cdot Q)\right\},$$

and then deal with (2.4) with $\tilde{H}$ instead of H. However, the choice of $\zeta_{\max}$, guaranteeing a good approximation of H, does not appear obvious to us. To avoid these difficulties, we first give an equivalent HJB equation satisfied by ϑ and which is formulated with bounded controls. More precisely, we have :

Corollary IV.3. *Under assumptions (A1)-(A3), the value function ϑ is a viscosity solution of the HJB equation :*

$$\inf_{\alpha_1^2+\alpha_2^2=1}\left\{\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}^T\begin{pmatrix}-\frac{\partial\vartheta}{\partial t}+\mu(t,y)\frac{\partial\vartheta}{\partial y}-\frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y}\\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}\right\}=0. \tag{2.10}$$

Proof. Simple to obtain from Theorem 2.3. □

Remark IV.4. *Equation (2.5) can be reformulated as follows,*

$$\Lambda^-\begin{pmatrix}-\frac{\partial\vartheta}{\partial t}+\mu(t,y)\frac{\partial\vartheta}{\partial y}-\frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\eta(t,y)\frac{\partial^2\vartheta}{\partial x\partial y}\\ -\frac{1}{2}\sigma(t,y)x\eta(t,y)\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\eta^2(t,y)\frac{\partial^2\vartheta}{\partial y^2}\end{pmatrix}=0,$$

where $\eta(t,y):[0,T]\times(0,+\infty)\to(0,+\infty)$ is any strictly positive function. It is easy to see that changing the positive function $\eta(t,y)$ into another positive function, does not change the sign of the operator in (2.5), for fixed $(t,x,y,D\vartheta,D^2\vartheta)$.

In particular, when we will deal with the discretization of (2.10), we will use $\eta(t,y)=\min(1;y)$.

2.1 Boundary conditions. Uniqueness result

Unlike in most similar parabolic problems, here we do not only need a terminal condition to obtain the uniqueness, but also a border conditions when y tends to zero. Another boundary condition is hidden by the fact that we only consider bounded solutions, which is, intuitively, equivalent to Neumann conditions near infinity.

Lemma 2.4. *Under assumptions (A1)-(A3), the value function ϑ is bounded and satisfies the following conditions on the boundaries $x = 0$ and $y = 0$:*

$$\lim_{(t',x',y')\to(t,x,0)} \vartheta(t',x',y') = \vartheta(t,x,0) = g(x), \forall\,(t,x) \in [0,T]\times\mathbb{R}_+^* \tag{2.11a}$$

$$\lim_{(t',x',y')\to(t,0,y)} \vartheta(t',x',y') = \vartheta(t,0,y) = g(0), \forall\,(t,y) \in [0,T]\times\mathbb{R}_+^* \tag{2.11b}$$

and the terminal condition of the equation for $t = T$ is :

$$\lim_{(t',x',y')\to(T,x,y)} \vartheta(t',x',y') = \vartheta(T,x,y) = g(x) \text{ for all } (x,y) \in (\mathbb{R}_+^*)^2. \tag{2.11c}$$

Proof. The statements (2.11a)-(2.11c) are proved in [8, lemma 5.6]. The proof is based on the assumptions (A1) and (A2) of σ and μ, and on the continuity and boundedness of g (see (A3)).

Now to prove statement (2.11b), we first give a representation of $\vartheta(t,x,y)$ using Doleans integral. Indeed, for every (t,x,y), we have :

$$X_{t,x,y}^{\rho,\zeta} = xZ_y^{\zeta,\rho}, \quad \text{where} \quad Z_y^{\zeta,\rho} := e^{\int_t^T \sigma(s,Y_{t,y}^{\rho,\zeta}(s)dW_s^1 + \frac{1}{2}\int_t^T(\sigma(s,Y_{t,y}^{\rho,\zeta}(s)))^2 ds}.$$

Therefore,

$$\vartheta(t,x,y) = \mathbb{E}\Big[g(X_{t,x,y}^{\rho,\zeta})(T)\Big] = \mathbb{E}\Big[g\Big(xZ_y^{\zeta,\rho}\Big)\Big]. \tag{2.12}$$

We conclude that statements (2.11b) holds. □

We recall here the uniqueness result, proved in [8, Lemma 4.3, Proposition 4.4, Proposition 4.6].

Theorem 2.5. *[8, Proposition 4.4] Assume (A1)-(A3). Suppose that u is an upper semi-continuous viscosity sub-solution of (2.5) bounded from above, and w a lower semi-continuous viscosity super-solution of (2.5) bounded from below. If, furthermore,*

$$\begin{array}{lllll} u(T,x,y) & \leq & g(x) & \leq & w(T,x,y), \\ u(t,x,0) & \leq & g(x) & \leq & w(t,x,0), \end{array} \tag{2.13}$$

then $u(t,x,y) \leq w(t,x,y)$, for all $(t,x,y) \in [0,T]\times\mathbb{R}_+^2$. In particular, the solution of (2.5) in the viscosity sense with boundary conditions (2.11a) and (2.11c) is unique.

We recall here the main ideas of the proof.

Proof. Suppose that u and w are respectively sub- and super-solution of (2.5), and that they both satisfy the limit conditions (2.11a), (2.11b) and (2.11c). A classical argument (see [4]) to prove uniqueness for equation as (2.5), consists in building a strict viscosity super-solution of (2.5) w_ε, depending on the super-solution and on a parameter ε. Moreover w_ε must to be such that, when the parameter ε goes to zero, w_ε tends to w. Then with classical arguments

[11], a comparison principle between the strict super-solution and the sub-solution can be obtained, and sending ε to zero we have the desired comparison principle.

In our particular case, for any $\varepsilon > 0$, we build

$$w_\varepsilon = w + \varepsilon((T-t) + \ln(1+y)).$$

From [8, Lemma 4.3], w_ε is a strict viscosity super-solution of (2.5), bounded from below and such that conditions (2.13) are satisfied. Then we can apply [8, Proposition 4.6] which is a comparison principle between a strict viscosity super-solution and a viscosity sub-solution, and we obtain

$$w_\varepsilon \geq u,$$

for all $(t,x,y) \in [0,T] \times \mathbb{R}^2_+$. Sending ε to zero, we have the result. □

Since the boundedness property of ϑ would be tricky to manipulate numerically, in the following proposition we give some growth properties of the value function which are a sort of Neumann conditions at infinity. These conditions will guide us toward an implementable scheme.

Proposition IV.5. *Assume that (A1)-(A4) are satisfied. Then the following holds :*
(i) For any $a > 0$, the function :

$$h^1_{t,y} : x \to \vartheta(t, x+a, y) - \vartheta(t,x,y)$$

converges to zero, uniformly in (t,y), when $x \to +\infty$.
(ii) The function;

$$h^2_{t,x} : y \to \vartheta(t,x,y+a) - \vartheta(t,x,y)$$

converges to zero, uniformly in (t,x), when $y \to +\infty$.

Proof. (i) Let $(t,x,y) \in (0,T) \times \mathbb{R}^+ \times \mathbb{R}^+$. As in (2.12), we have :

$$\vartheta(t,x,y) = \sup_{\zeta,\rho} \mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right] = \sup_{\zeta,\rho} \mathbb{E}\left[g\left(xZ^{\zeta,\rho}_y\right)\right]. \tag{2.14}$$

By assumption (A3), the function $f : z \to g(e^z)$ is Lipschitz continuous on $\mathbb{R}$. Then, for $x' \in \mathbb{R}^+$, we get :

$$\begin{aligned}
\vartheta(t,x,y) - \vartheta(t,x',y) &= \sup_{\zeta,\rho} \mathbb{E}\left(g\left(xZ^{\zeta,\rho}_y\right)\right) - \sup_{\zeta,\rho} \mathbb{E}\left(g\left(x'Z^{\zeta,\rho}_y\right)\right) \\
&\leq \sup_{\zeta,\rho}\left\{\mathbb{E}\left(g\left(xZ^{\zeta,\rho}_y\right)\right) - \mathbb{E}\left(g\left(x'Z^{\zeta,\rho}_y\right)\right)\right\} \\
&\leq \sup_{\zeta,\rho}\left\{\mathbb{E}\left[f\left(\ln(x) + \ln\left(Z^{\zeta,\rho}_y\right)\right) - f\left(\ln(x') + \ln\left(Z^{\zeta,\rho}_y\right)\right)\right]\right\},
\end{aligned}$$

and using the Lipschitz property of f, it yields to :

$$\vartheta(t,x,y) - \vartheta(t,x',y) \leq K\left|\ln(x) - \ln(x')\right|.$$

Therefore we get that

$$\left|h^1_{t,y}(x)\right| \le K \left|\ln\left(\frac{x+a}{x}\right)\right| \to 0 \text{ as } x \to +\infty \text{ uniformly in } (t,y). \tag{2.15}$$

To prove assertion (ii), we recall that by Theorem 2.3, ϑ is a supersolution of

$$-\frac{\partial^2 \vartheta}{\partial y^2} = 0.$$

Then, from [12], we deduce that the function ϑ is concave w.r.t. y. That is, for each (t,x), $\vartheta(t,x,\cdot)$ is a concave function. Moreover, from (A3), ϑ is bounded and $\|\vartheta\|_\infty \le M_0$ (where the constant $M_0 > 0$ is the same as in (A3)). Therefore, for any λ, the function

$$h^2_{t,x} : y \to \vartheta(t,x,y+\lambda) - \vartheta(t,x,y)$$

is decreasing. Considering that $\vartheta(t,x,n\lambda+y_0) = \vartheta(t,x,y_0) + \sum_{i=1}^n h^2_{t,x}(i\lambda+y_0)$. Hence, it follows that :

$$\vartheta(t,x,n\lambda+y_0) \ge \vartheta(t,x,y_0) + \sum_{i=1}^n h^2_{t,x}(n\lambda+y_0)$$

which gives :

$$h^2_{t,x}(n\lambda+y_0) \le \frac{2M}{n}$$

and we get convergence of $h^2_{t,x}(y)$ to 0, which is uniform in (t,x). □

2.2 Lipschitz property

Here we establish the Lipschitz property of the value function ϑ.

Proposition IV.6. *Under assumptions (A1)-(A4), we have :*
(i) The value function ϑ is Lipschitz w.r.t. x.
(ii) ϑ is Lipschitz w.r.t. y.

Proof. (i)As in the proof of proposition IV.5, we consider the representation of ϑ using Doleans exponential :

$$\vartheta(t,x,y) = \sup_{\zeta,\rho} \mathbb{E}\left(g(X^{\zeta,\rho}_{t,x,y})\right) = \sup_{\zeta,\rho} \mathbb{E}\left[g\left(xZ^{\zeta,\rho}_y\right)\right] \qquad \forall t \in (0,T), x, y \in \mathbb{R}^+, \tag{2.16}$$

where $Z^{\zeta,\rho}_y = e^{\int_t^T \sigma(s,Y^{\rho,\zeta}_{t,y}(s)dW^1_s + \frac{1}{2}\int_t^T (\sigma(s,Y^{\rho,\zeta}_{t,y}(s)))^2 ds}$.

Then, for $t \in (0,T), x, x', y \in \mathbb{R}^+$ we have :

$$\left|\vartheta(t,x,y) - \vartheta(t,x',y)\right| \le \sup_{\zeta,\rho} \mathbb{E}\left[g\left(xZ^{\zeta,\rho}_y\right) - g\left(xZ^{\zeta,\rho}_y\right)\right].$$

As g is Lipschitz of constant K, we get :

$$\left|\vartheta(t,x,y)-\vartheta(t,x',y)\right| \leq \sup_{\zeta,\rho} \mathbb{E}\left|K(x-x')Z_y^{\zeta,\rho}\right| \leq K|x-x'|\sup_{\zeta,\rho}\mathbb{E}\left(Z_y^{\zeta,\rho}\right).$$

Therefore, using the fact that the Doleans exponential is a positive local martingale, and hence a super-martingale, which implies that for any control $(\zeta,\rho)\in\mathcal{U}$:

$$\mathbb{E}\left(e^{\int_t^T \sigma_u^{\zeta,\rho}dW_u+\frac{1}{2}\int_t^T(\sigma_u)^{\zeta,\rho})^2du}\right)\leq 1,$$

and then taking the supremum leads to :

$$\left|\vartheta(t,x,y)-\vartheta(t,x',y)\right| \leq K|x-x'|$$

Which proves that ϑ is Lipschitz w.r.t. x with the same constant as g.

(ii) Now we treat the Lipschitz property of ϑ w.r.t. y.

First,we recall that ϑ is concave w.r.t. y. Furthermore, as g is bounded, we immediately get that ϑ shares the same bound. Hence, it is sufficient to prove that ϑ is Lipschitz near the boundary $y=0$.

Recall that by (2.11a), we know that $\vartheta(t,x,0)=g(x)$ for all $(t,x)\in(0,T)\times(0,+\infty)$.

Let $(t,x,y)\in[0,T]\times(0,+\infty)^2$, with $y>0$. For any control $(\zeta,\rho)\in\mathcal{U}$, we have :

$$Y_{t,y}^{\rho,\zeta}(s)=y+\int_t^s -\mu(\tau,Y_{t,y}^{\rho,\zeta}(\tau))d\tau+\int_t^s \zeta(\tau)Y_{t,y}^{\rho,\zeta}\tau\, dW_\tau^2.$$

Furthermore, by a comparison argument for SDEs, we get, for any $\tau\in[t,T]$:

$$Y_{t,y}^{\rho,\zeta}(\tau)\geq 0.$$

Using the positivity of μ, we get :

$$0\leq Y_{t,y}^{\rho,\zeta}(s)\leq y+\int_t^s Y_{t,y}^{\rho,\zeta}(\tau)dW_\tau^2.$$

Hence, the quantity above is a super-martingale and we get :

$$\mathbb{E}\left[Y_{t,y}^{\rho,\zeta}(s)\right]\leq y. \tag{2.17}$$

Now, applying Itô's formula on $g(X_{t,x,y}^{\rho,\zeta})$:

$$\begin{aligned} g(X_{t,x,y}^{\rho,\zeta}(s)) &= g(x)+\int_t^s g'(X_{t,x,y}^{\rho,\zeta}(\tau))dX_{t,x,y}^{\rho,\zeta}(\tau)+ \\ &\quad \frac{1}{2}\int_t^s g''(X_{t,x,y}^{\rho,\zeta}(\tau))\left\langle dX_{t,x,y}^{\rho,\zeta}(\tau),dX_{t,x,y}^{\rho,\zeta}(\tau)\right\rangle \\ &= g(x)+\int_t^s g'(X_{t,x,y}^{\rho,\zeta}(\tau))dX_{t,x,y}^{\rho,\zeta}(\tau)+ \\ &\quad \frac{1}{2}\int_t^s \left(X_{t,x,y}^{\rho,\zeta}(\tau)\right)^2 g''(X_{t,x,y}^{\rho,\zeta}(\tau))\sigma^2(Y_{t,y}^{\rho,\zeta}(\tau))d\tau. \end{aligned}$$

Since $X^{\rho,\zeta}_{t,x,y}$ is a locale martingale, there exists a sequence $(s_n)_n$, with $s_n \to \infty$ such that :

$$\mathbb{E}\bigg(\int_t^{s_n\wedge T} g'(X^{\rho,\zeta}_{t,x,y}(u))dX^{\rho,\zeta}_{t,x,y}(u)\bigg) = 0.$$

Using (2.17), the Lipschitz property of σ^2, and the boundedness of $x \longmapsto x^2 g''(x)$, it yields : there exists a constant $C > 0$, such that :

$$\left|\mathbb{E}\left(g(X^{\rho,\zeta}_{t,x,y}(s_n\wedge T)) - g(x)\right)\right| \le \int_t^{s_n\wedge T} Cy\, d\tau$$

Finally, as g is bounded, we conclude with Fatou's lemma that :

$$\mathbb{E}\left(g(X^{\rho,\zeta}_{t,x,y}(T)) - g(x)\right) \le C(T-t)y,$$

and since the constant C is independent of ρ, ζ, we obtain :

$$\begin{aligned}|\vartheta(t,x,y) - \vartheta(t,x,0)| &\le \sup_{(\rho,\zeta)\in\mathcal{U}} \left\{|\mathbb{E}\left(g(X^{\rho,\zeta}_{t,x,y}(T)) - g(x)\right)|\right\}\\ &\le CTy.\end{aligned}$$

Hence, as ϑ is concave w.r.t. y and bounded, it is Lipschitz with respect to y. □

3 Approximation Scheme

Now we want to approximate the (unique) bounded solution of the following Hamilton-Jacobi-Bellman equation :

$$\min_{\alpha_1^2+\alpha_2^2=1}\left\{ -\alpha_1^2\frac{\partial\vartheta}{\partial t}(t,x,y) + \mu(t,y)\alpha_1^2\frac{\partial\vartheta}{\partial y}(t,x,y) - \frac{1}{2}\mathrm{tr}[a(\alpha_1,\alpha_2,t,x,y)D^2\vartheta(t,x,y)]\right\} = 0, \tag{3.1}$$

with boundary conditions (2.11a), (2.11b), (2.11c), where μ is a positive Lipschitz function, and the diffusion matrix a is defined as follows :

$$\begin{aligned} a(\alpha_1,\alpha_2,t,x,y) &:= \begin{pmatrix} \alpha_1^2\sigma^2(t,y)x^2 & \alpha_1\alpha_2\sigma(t,y)\eta(t,y)x \\ \alpha_1\alpha_2\eta(t,y)\sigma(t,y)x & \eta^2(t,y)\alpha_2^2 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1\sigma(t,y)x \\ \eta(t,y)\alpha_2 \end{pmatrix}\begin{pmatrix} \alpha_1\sigma(t,y)x \\ \alpha_2\eta(t,y) \end{pmatrix}^{\mathsf{T}}, \end{aligned} \tag{3.2}$$

where we will use $\eta(t,y) = \min(1;y)$, in agreement with remark IV.4. From now on we will write only a instead of $a(\alpha_1,\alpha_2,t,x,y)$, and μ instead of $\mu(t,y)$, we omit all the dependences.

We can easily see that a is not a dominant diagonal matrix[2], in fact we can not ensure that

$$\alpha_2\eta \geq \alpha_1\sigma x, \quad \forall\,(t,x,y)\in[0,T)\times[0,+\infty)^2, \text{ and } \forall\,\alpha_1^2+\alpha_2^2=1.$$

This fact implies that we can not choose the classical finite differences (FD) scheme to approximate equation (3.1), we shall use the generalized finite differences (GFD) scheme introduced in [6].

Consider a regular grid G_h of discretization of $\mathbb{R}^2_+$, with discretization steps $h=(h_1,h_2)$:

$$G_h := \Big\{(x_i,y_j),\ x_i := ih_1,\ y_j := jh_2,\ i,j\in\mathbb{N}\times\mathbb{N}\Big\},$$

and consider a discretization time step Δt. On the grid G_h, the derivative on time is approximated by an implicit Euler scheme, and for the first derivative in y we use a finite difference approximation. The main idea of the Generalized Finite Differences scheme is to approximate the diffusion term $a\cdot D^2\phi$ by a linear combination of elementary diffusions pointing towards grid points. More precisely, for $\xi=(\xi_1,\xi_2)\in\mathbb{Z}^2$, associate the second order finite difference operator (for $x,y\in\mathbb{R}$) :

$$\Delta_\xi\phi(t,x,y)=\phi(t,x+\xi_1h_1,y+\xi_2h_2)+\phi(t,x-\xi_1h_1,y-\xi_2h_2)-2\phi(t,x,y),$$

where Δ_ξ is an elementary diffusion in the direction ξ. By a Taylor expansion, we know that

$$\Delta_\xi\phi(t,x,y)=\sum_{i,j=1}^{2}h_ih_j\xi_i\xi_j\phi_{x_ix_j}+o(||h^2||),$$

where $x_1=x$ and $x_2=y$.

Following ([6, 7]), we introduce a set $\mathcal{S}\subseteq\mathbb{Z}^2\setminus 0$, which contains $\{e_1,e_2\}$. We will specify later how we choose this set. We approximate the second order term $a\cdot D^2\phi$ by a linear combination of elementary diffusions along ξ, with $\xi\in\mathcal{S}$:

$$a\cdot D^2\phi \cong \sum_{\xi\in\mathcal{S}}\gamma_\xi^{\alpha_1,\alpha_2}\Delta_\xi\phi,$$

where the $\gamma_\xi^{\alpha_1,\alpha_2}$ are coefficients which will be specified later.

For a given set $\mathcal{S}$, the scheme takes the following form :

$$v_h(T,x,y)=g(x)=v_h(t,x,0), \quad v_h(t,0,y)=g(0), \tag{3.3}$$

$$\min_{\alpha_1^2+\alpha_2^2=1}\{-\alpha_1^2\delta_tv_h(t,x,y)-\alpha_1^2\mu\delta_yv_h(t,x,y)-\frac{1}{2}\sum_{\xi\in\mathcal{S}}\gamma_\xi^{\alpha_1,\alpha_2}\Delta_\xi v_h(t,x,y)\}=0, \tag{3.4}$$

[2] We recall that a matrix X of dimension $N\times N$ is diagonal dominant if

$$X_{ii}\geq\sum_{i\neq j}|X_{ij}|, \quad \forall\,i=1,\dots,N.$$

for $t < T - \Delta t$, with

$$\begin{aligned} \delta_t v_h(t,x,y) &= \frac{v_h(t+\Delta t,x,y) - v_h(t,x,y)}{\Delta t}, \\ \delta_y v_h(t,x,y) &= \frac{v_h(t,x,y-h_2) - v_h(t,x,y)}{h_2}. \end{aligned}$$

It is shown in [6, 7] that the above scheme is consistent if we choose a set $\mathcal{S}$ and variables $\gamma_\xi^{\alpha_1,\alpha_2}$ such that : for all α_1, α_2, t, x, y

$$\gamma_\xi^{\alpha_1,\alpha_2} \geq 0, \, \forall\, \xi \in \mathcal{S}, \qquad \text{and} \qquad \sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \xi \xi^\mathsf{T} = a^h, \tag{3.5}$$

where a^h denotes the scaled matrix, $a^h = \{a_{ij}/(h_i h_j)\}$. Assertion (3.5) means that a^h belongs to the cone generated by $\{\xi\xi^\mathsf{T}; \xi \in \mathcal{S}\}$,

$$\mathcal{C}(\mathcal{S}) = \left\{ \sum_{\xi \in \mathcal{S}} \gamma_\xi \xi \xi^\mathsf{T}, \gamma \in \mathbb{R}_+^{|\mathcal{S}|} \right\}.$$

A natural choice for $\mathcal{S}$ is the following :

$$\mathcal{S} = \mathcal{S}_p = \{(\xi_1, \xi_2) \in \mathbb{Z} \times \mathbb{N}; \max(|\xi_1|; \xi_2) \leq p; (|\xi_1|, \xi_2) \text{ irreducible}\},$$

for $p \geq 1$, and the correspondent cones $\mathcal{C}(\mathcal{S}_p)$. These cones have the following property :

$$\mathcal{C}(\mathcal{S}_1) \subset \mathcal{C}(\mathcal{S}_2) \subset \cdots \subset \mathcal{C}(\mathcal{S}_p) \subset \cdots \subset \mathcal{M}_+^{\#},$$

where $\mathcal{M}_+^{\#}$ denotes the set of symmetric positive matrices. Unfortunately, even for a big order $p >> 1$, the matrix a^h does not satisfy necessarily the strong consistency (3.5).

Moreover a^h is a rank one matrix and it is degenerated. This fact implies two possibilities :

• The direction of diffusion $\begin{pmatrix} \alpha_1 \sigma x \\ \alpha_2 \eta \end{pmatrix}$ points toward a point of the grid. This situation happens if the slope is a rational number r/q (with $r \in \mathbb{Z}$ and $q \in \mathbb{N}^*$). Then we consider the vector $\xi_{r,q} = (r\ q)^\mathsf{T}$, and we can write

$$a^h = \gamma_{\xi_{r,q}}^{\alpha_1,\alpha_2} \xi_{r,q} \xi_{r,q}^\mathsf{T}.$$

• The second possibility is that the direction of the diffusion $\begin{pmatrix} \alpha_1 \sigma x \\ \alpha_2 \eta \end{pmatrix}$ has a real slope. In this case, we approximate a^h by its projection in one of the cones $\mathcal{C}(\mathcal{S}_p)$, the order p being the order of neighbouring points allowed to enter in the scheme (of course, this order depends on where we are situated on the grid and on the direction of the diffusion).

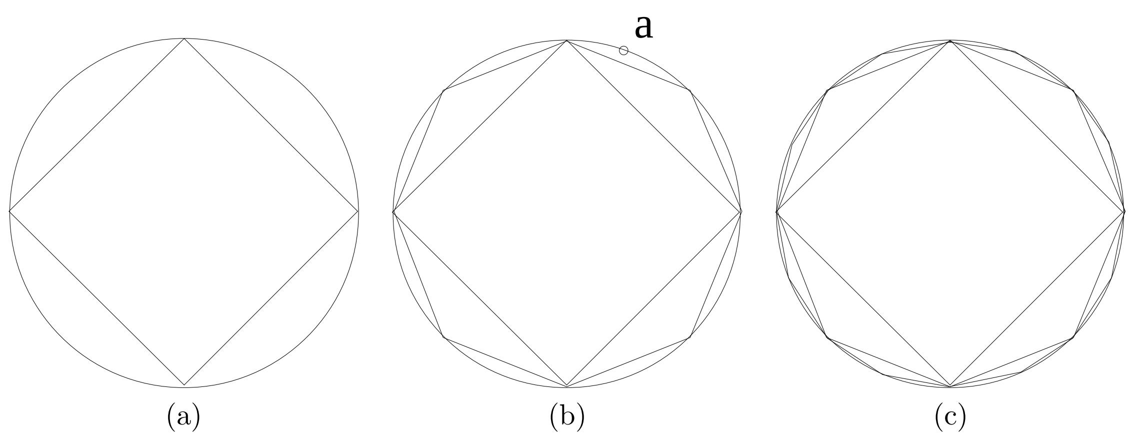


FIG. IV.1 – (a) Symmetric semi-definite positive matrix with trace equal to 1 and cone of diagonal dominant matrix. (b) Cone $\mathcal{C}(\mathcal{S}_1)$, a is on the border of the semi-definite positive matrix. (c) Cone $\mathcal{C}(\mathcal{S}_2)$.

Remark IV.7. *As we can see in Figure 1(b), matrix a^h belongs to the border of the cone $\mathcal{M}_+^{\#}$ (the cone of symmetric semi-definite positive matrices), and then there exist two vectors $\xi_{p',q'}$ and $\xi_{p",q"}$ on $\mathcal{S}_p$, such that we can project a^h on the hyperplane generated by $\xi_{p',q'}\xi_{p',q'}^{\mathsf{T}}$ and $\xi_{p",q"}\xi_{p",q"}^{\mathsf{T}}$. Then, we can write the projection of a^h as follows :*

$$a_p^h = \gamma_{\xi_{p',q'}}^{\alpha_1,\alpha_2}\xi_{p',q'}\xi_{p',q'}^{\mathsf{T}} + \gamma_{\xi_{p",q"}}^{\alpha_1,\alpha_2}\xi_{p",q"}\xi_{p",q"}^{\mathsf{T}}, \tag{3.6}$$

where $\gamma_\xi^{\alpha_1,\alpha_2}$ are positive coefficients, and moreover

$$\gamma_{\xi_{p',q'}}^{\alpha_1,\alpha_2} + \gamma_{\xi_{p",q"}}^{\alpha_1,\alpha_2} \leq \operatorname{tr}(a_p^h).$$

As studied in [6], the generation of the directions $\xi_{p',q'}$ and $\xi_{p",q"}$, can be performed (in effective way) in $O(p)$ operations, by using Stern-Brocot algorithm [13].

Remark IV.8. *The choice of the order p depends on where we are situated on the grid. For instance, if we consider a point (x, y) in the middle of the grid, and we want to discretize $a \cdot D^2\phi(t, x, y)$, we can follow the direction of diffusion and choose the biggest order of discretization p, because more p is bigger and better is the approximation of the scaled covariance matrix a^h. On the other hand, if we consider a point (x, y) near to the boundary, it can often happen that following the direction of the diffusion, we involve in the discretization some points which are out of the grid. In this case the choice of p is not free, and we refer to the Appendix for a detailed discussion of this case.*

Remark IV.9. *In all the decompositions, the coefficients $\gamma_\xi^{\alpha_1,\alpha_2}$ and also the vectors ξ are in terms of (t, x, y). Sometimes, for simplicity of notations we do no specify this dependence.*

Error projection for scaled covariance matrices. From now on, by $\lceil r \rceil$ we denote the smallest integer greater than the real r, and for a symmetric matrix b of dimension 2 we use the Frobenius norm $\|b\| = (\sum_{i,j=1,2} b_{ij}^2)^{1/2}$. Let $p_{\max}$ the maximum order that we can consider for the discretization, and let us consider the projection b' of a general matrix $b \in \mathcal{M}_+^{\#}$ on a hyperplane of $\mathcal{C}(\mathcal{S}_{p_{\max}})$ spanned by $\xi\xi^T$ and $\xi'(\xi')^T$. As proved in [6, (17)], we have

$$\|b - b'\| \leq \frac{(1 - \cos(\widehat{\xi, \xi'}))}{\sqrt{2}\sqrt{1 + \cos^2(\widehat{\xi, \xi'})}} \|b\|. \tag{3.7}$$

From this inequality and [6, Lemma 6.1], the projection error will be $\epsilon_p ||b||$, where

$$\epsilon_p = \frac{\sqrt{p_{\max}^2 + 1} - p_{\max}}{\sqrt{2}\sqrt{2p_{\max}^2 + 1}} \leq \frac{1}{4} p_{\max}^{-2}. \tag{3.8}$$

Moreover, the error projection is guaranteed to be at most equal to ε (for any $\varepsilon > 0$), if we choose $p_{\max} \geq p_\varepsilon$, where

$$p_\varepsilon := \left\lceil \frac{\sqrt{1 - \varepsilon^2} - \varepsilon}{2\sqrt{\varepsilon\sqrt{1 - \varepsilon^2}}} \right\rceil. \tag{3.9}$$

In particular, if we aim at having a projection error of the order of h, then we have to choose $p_{\max}$ such that

$$p_{\max} \geq \left\lceil \frac{\sqrt{1 - h^2} - h}{2\sqrt{h\sqrt{1 - h^2}}} \right\rceil.$$

(Some examples : If $h \leq 10^{-1}$ then $p_{\max} \geq 2$. If $h \leq 10^{-3}$ then $p_{\max} \geq 16$, and if $h \leq 10^{-5}$ then $p_{\max} \geq 159$.)

3.1 The discrete equation

From now on, we fix $h_1 = h_2 = h$, the space step size[3]. Let $p_{\max} \in \mathbb{N}$ be the maximal order of grid points allowed to enter in the scheme, and Δt be the time step size. Set $\rho = (p_{\max}, h, \Delta t)$, and define the scheme S^ρ (given in a general setting) as follows : let $\phi : [0, T] \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$:

$$\begin{aligned} S^\rho(t, x, y, r, \phi) \quad &= \min_{\alpha_1^2 + \alpha_2^2 = 1} \Big\{ -\alpha_1^2 \frac{\phi(t + \Delta t, x, y) - r}{\Delta t} + \alpha_1^2 \mu \frac{r - \phi(t, x, y - h)}{h} \\ &- \frac{1}{2} \sum_{\xi \in \mathcal{S}(x,y)} \gamma_\xi^{\alpha_1, \alpha_2}(t, x, y) \big[\phi(t, x - \xi_1 h, y - \xi_2 h) - 2r + \phi(t, x + \xi_1 h, y + \xi_2 h)\big] \Big\}, \end{aligned} \tag{3.10a}$$

for $(t, x, y) \in [0, T) \times (0, \infty)^2$, where

$$\mathcal{S}(x, y) := \mathcal{S}_p \quad \text{with } p = \min(p_{max}, \lceil x/h \rceil, \lceil y/h \rceil), \tag{3.10b}$$

$$\sum_{\xi \in \mathcal{S}(x,y)} \gamma_\xi^{\alpha_1, \alpha_2}(t, x, y) \xi \xi^\mathsf{T} = a_p^h(\alpha_1, \alpha_2, t, x, y), \tag{3.10c}$$

[3] We set $h_1 = h_2 = h$ to simplify the analysis.

the projection of the scaled covariance matrix a^h on $\mathcal{C}(\mathcal{S}_p)$ ($a^h = a/h^2$). In particular, $p = p_{max}$ if $x - p_{max}h \geq 0$ and $y - p_{max}h \geq 0$ (points in the interior of the domain), otherwise $p = \min(\lceil x/h \rceil, \lceil y/h \rceil)$ (points near to the boundary).

Now the discrete scheme for (3.1) is :

$$S^\rho(t, x, y, v_h(t, x, y), v_h) = 0, \tag{3.11a}$$

for $(t, x, y) \in [0, T) \times (0, \infty)^2$, and with the boundary conditions :

$$v_h(T, x, y) = g(x), \ \forall\ (x, y) \in [0, \infty)^2, \tag{3.11b}$$
$$v_h(t, x, 0) = g(x), \ \forall\ (t, x) \in [0, T] \times [0, \infty), \tag{3.11c}$$
$$v_h(t, 0, y) = g(0), \ \forall\ (t, y) \in [0, T] \times [0, \infty). \tag{3.11d}$$

(the solution v_h will stand for an approximation of the value function ϑ).

Remark IV.10. *It is clear that if $p_{\max}$ is not linked to the step size h, then (3.11) is a discrete scheme for the HJB equation with the covariance matrix a_p instead of a, where a_p is the projection of a on the cone $C(\mathcal{S}_p)$.*

$$\min_{\alpha_1^2+\alpha_2^2} \left\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t, x, y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t, x, y) - \frac{1}{2}\mathrm{tr}[a_p D^2\phi(t, x, y)] \right\} = 0. \tag{3.12}$$

In what follows (subsection 3.2, and section 4), we will prove that the scheme (3.11) satisfies the following properties :

(S1) **Monotonicity :** $S^\rho(t, x, y, r, u) \geq S^\rho(t, x, y, r, v)$,
for all $r \in \mathbb{R}$, $x, y \in \mathbb{R}_+^*$, $u, v \in C([0, T] \times [0, \infty)^2)$ such that $u \leq v$ in $[0, T] \times [0, \infty)^2$.

(S2) **Stability :** For all $\rho = (h, \Delta t, p_{\max}) \in (\mathbb{R}_+^*) \times (0, T) \times \mathbb{N}^*$, there exists a bounded solution v_h of (3.11).

(S3) **Consistency :** There exists a constant $C_1 > 0$ such that, for every $\phi \in C^n([0, T] \times [0, \infty)^2)$, $n \geq 4$, with bounded derivatives,

$$\begin{aligned} &\Bigg| \min_{\alpha_1^2+\alpha_2^2} \left\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t, x, y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t, x, y) - \frac{1}{2}\mathrm{tr}[a \cdot D^2\phi(t, x, y)] \right\} \\ &-S^\rho(t, x, y, \phi(t, x, y), \phi) \Bigg| \\ &\leq C_1(|\partial_t^2 \phi|_0 \Delta t + \mu |D_y^2 \phi|_0 h) + 16\sqrt{2} p_{max}^2 \|a\| |D^4\phi|_0 h^2 + \varepsilon_p(t, x, y)|D^2\phi|_0, \end{aligned} \tag{3.13}$$

where a_p is the projection of a on $\mathcal{C}(\mathcal{S}_p)$, for $p = \min(p_{max}, \lceil x/h \rceil, \lceil y/h \rceil)$, and $\varepsilon_p(t, x, y)$ is the projection error such that $\varepsilon_p = ||a - a_p||$ if $p = p_{max}$, and $\varepsilon_p = CK(x, y)h^2$ otherwise, where C depends on the Lipschitz constant of σ^2, and $K(x, y) \geq 0$, for all x, y.

3.2 The consistency property

We start by proving the consistency property (S3). Consider a function $\phi \in C^n([0,T] \times [0,+\infty)^2)$, with bounded derivatives and compute (first) the difference term :

$$\begin{aligned} &\Bigg| \min_{\alpha_1^2+\alpha_2^2} \Bigg\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t,x,y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t,x,y) - \frac{1}{2}\mathrm{tr}[a_p \cdot D^2\phi(t,x,y)] \Bigg\} \\ &-S^\rho(t,x,y,\phi(t,x,y),\phi) \Bigg|, \end{aligned} \tag{3.14}$$

for the HJB-equation with the matrix a_p instead of a. For the derivatives on t and on y we just apply a Taylor development to obtain the bound terms $|\partial_t^2\phi|_0 \Delta t$ and $\mu |D_y^2 \phi|_0 h$. Consider now the diffusion term : by a Taylor development, we get (for $\xi \in \mathcal{S}$) :

$$\begin{aligned} D^2\phi(h\xi, h\xi) - \Delta_\xi \phi &\leq 2h^4 \textstyle\sum_{k=0}^4 \xi_1^k \xi_2^{4-k} \frac{\partial^4 \phi}{\partial x^k \partial y^{4-k}}, \\ &\leq 4h^4 ||\xi||^4 |D^4\phi|_0, \end{aligned}$$

where $D^4\phi = \sum_{k=0}^4 \frac{\partial^4 \phi}{\partial x_1^k \partial x_2^{4-k}}$, and the last inequality follows from the fact that $\sum_{k=0}^4 \xi_1^k \xi_2^{4-k} \leq 2||\xi||^4$. Moreover, from (3.6), we can deduce that

$$0 \leq \gamma_\xi^{\alpha_1,\alpha_2} \leq \frac{\mathrm{tr}(a_p^h)}{\|\xi\|^2},$$

for every ξ which appear in the decomposition of a_p^h. Then, for the global diffusion term we obtain

$$\begin{aligned} \mathrm{tr}[a_p D^2\phi(t,x,y)] - \sum_{\xi\in\mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi \phi(t,x,y) &\leq 2\mathrm{tr}(a_p^h)|D^4\phi|_0 h^4 \textstyle\sum_{\xi\in\mathcal{S}} \|\xi\|^2 \\ &\leq 8\mathrm{tr}(a_p^h)|D^4\phi|_0 h^4 p_{max}^2 \\ &\leq 8\mathrm{tr}(a_p)|D^4\phi|_0 h^2 p_{max}^2 \end{aligned}$$

where the last inequality follows from the fact that $\xi_i \leq p_{max}$, for $i = 1, 2$. We are now looking for a bound of $\mathrm{tr}(a_p)$ which depends on (t,x,y). It is easy to see that $\mathrm{tr}(a_p) \leq \sqrt{2}\|a_p\|$, and moreover, by (3.7), we can show that

$$\|a_p\| \leq 2\|a\|, \tag{3.15}$$

where $\|a\|$ depends on t,x,y.

Therefore, we obtain

$$\mathrm{tr}[a_p D^2\phi(t,x,y)] - \sum_{\xi\in\mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi \phi(t,x,y) \leq 16\sqrt{2}\|a\||D^4\phi|_0 h^2 p_{max}^2.$$

Then we can conclude that

$$(3.14) \quad \leq \quad C_1(|\partial_t^2\phi|_0 \Delta t + \mu |D_y^2\phi|_0 h) + 16\sqrt{2} p_{max}^2 \|a\| |D^4\phi|_0 h^2. \tag{3.16}$$

On the other hand,

- For the points (x, y) such that $x - p_{max}h \geq 0$, and $y - p_{max}h \geq 0$,

$$\left| \mathrm{tr}[a \cdot D^2\phi(t, x, y)] - \mathrm{tr}[a_p \cdot D^2\phi(t, x, y)] \right| \leq \sqrt{2}\|a - a_p\||D^2\phi|_0.$$

  Let us note that to obtain an error $\|a - a_p\|$ depending on the space step h, we should choose a suitable value p_{max} as in (3.9).
- For the points such that $x < p_{max}h$ or $y < p_{max}h$, using the definition (3.2) of a, and the Lipschitz property of σ^2, one gets

$$||a|| \leq C(x^2y + y^2 + xy),$$

  where C depends on α_1,α_2 and on the Lipschitz constant of σ^2. Moreover, since we know that $p = \min(\lceil x/h \rceil, \lceil y/h \rceil)$, then $x, y \leq K(x, y)p \cdot h$, for a convenient integer $K(x, y) = \max(\lceil x/h \rceil, \lceil y/h \rceil) + 1$. Hence $||a|| \leq CK(x, y)p^2h^2$. Then this estimate with (3.8) yields to :

$$\begin{aligned} \left| \mathrm{tr}[a \cdot D^2\phi(t, x, y)] - \mathrm{tr}[a_p \cdot D^2\phi(t, x, y)] \right| &\leq \|a - a_p\||D^2\phi|_0 \\ &\leq \tfrac{1}{4p^2}||a|| \cdot |D^2\phi|_0 \\ &\leq CK(x, y)|D^2\phi|_0h^2. \end{aligned}$$

This concludes the consistency property (S3), with $\varepsilon_p(t, x, y) = ||a - a_p||$ if $p = p_{max}$, and $\varepsilon_p(t, x, y) = CK(x, y)h^2$ if $p = \min(\lceil x/h \rceil, \lceil y/h \rceil)$.

Proposition IV.11. *Suppose that*

(a) $p_{max} = o(\frac{1}{h})$,

(b) there exists $C > 0$ *such that* $p_{max} \geq \frac{C}{\sqrt{h}}$,

then

$$\left| \min_{\alpha_1^2+\alpha_2^2} \left\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t, x, y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t, x, y) - \frac{1}{2}\mathrm{tr}[a \cdot D^2\phi(t, x, y)] \right\} -S^\rho(t, x, y, \phi(t, x, y), \phi) \right| = O(h) + O(\Delta t).$$

for all $(t, x, y) \in [0, T] \times (\mathbb{R}_+^*)^2$.

Proof. The proof follows from the explicit form of the consistency property (S3).□

Remark IV.12. *In the case when the direction of diffusion points toward a point of the grid, the consistency remains the same, except for the error of projection which will be zero.*

4 Existence of the numerical solution

In this section we prove the well-posedness of the implicit scheme (3.11a) with boundary conditions (3.11b), (3.11c) and (3.11d), and show that it satisfies the required monotony and stability properties (S1)-(S2).

We recall that the grid is $G_h = \{(x_i, y_j), i, j \geq 0\} \subset \mathbb{R}_+^2$, where $x_i := ih$, $y_j := jh$.
We first start to initialize the scheme by

$$v_h(T, x, y) := g(x), \qquad (x, y) \in G_h.$$

Then, given $v_h(t + \Delta t, .)$ for some time t, we need to find $v_h(t, x, y)$ for $(x, y) \in G_h$ such that

$$\min_{\alpha_1^2+\alpha_2^2=1} \Bigg\{ \alpha_1^2 \frac{v_h(t,x,y) - v_h(t+\Delta t, x, y)}{\Delta t} + \alpha_1^2 \mu(t,y) \, \frac{v_h(t,x,y) - v_h(t,x,y-h)}{h} - \frac{1}{2} \sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi v_h(t,x,y) \Bigg\} = 0, \quad \forall (x,y) \in G_h, \ y > 0, \tag{4.1}$$

and with the following "boundary conditions" :

$$v_h(t, x, 0) = g(x), \ \forall \, x \in h\mathbb{N} \tag{4.2}$$

$$v_h(t, ., .) \text{ bounded} \tag{4.3}$$

The scheme in abstract form. Since for all $(x, y) \in G_h$ with $y > 0$, an optimal control (α_1, α_2) must be found, we introduce $\bar{\mathcal{B}} := \{\alpha = (\alpha_1, \alpha_2), \alpha_1^2 + \alpha_2^2 = 1\}$ and

$$\mathcal{A} := (\bar{\mathcal{B}})^{\mathbb{N} \times \mathbb{N}^*}$$

the set of controls associated to the grid mesh $(x_i, y_j)_{i \geq 0, \, j \geq 1}$.
The scheme can then be expressed in the following abstract form :
find $X := v_h(t, ., .) \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}^*}$, bounded, such that

$$\min_{w \in \mathcal{A}} \Big(A(w)X - b(w) \Big) = 0, \tag{4.4}$$

where $A(w)$ is a linear operator on $\mathbb{R}^{\mathbb{N} \times \mathbb{N}^*}$, and $b(w)$ is a vector of $\mathbb{R}^{\mathbb{N} \times \mathbb{N}^*}$, and are made precise below.

Definition of the matrix $A(w)$ **and vector** $b(w)$ **:** Let $X = (X_{ij})_{i \geq 0, \, j \geq 1}$, (resp. $w = (\alpha_{ij})_{i \geq 0, \, j \geq 1}$, with $\alpha_{ij} = (\alpha_{ij,1}, \alpha_{ij,2})$) be values (resp. controls) corresponding to the mesh points (x_i, y_j) of G_h. Then

- $A(w)$ is an infinite matrix determined by $\forall X$, $\forall i \geq 0$, $\forall j \geq 1$,

$$\begin{aligned} (A(w)X)_{ij} \quad := \quad & \frac{\alpha_{ij,1}^2}{\Delta t} X_{ij} + \alpha_{ij,1}^2 \mu(t, y_j) \frac{1}{h} (X_{ij} - (1 - \kappa_{j-1}) X_{i,j-1}) \\ & + \frac{1}{2} \sum_{\xi = (\xi_1, \xi_2) \in \mathcal{S}} \gamma_\xi^{\alpha_{ij}} (-(1 - \kappa_{j-\xi_2}) X_{i-\xi_1, j-\xi_2} + 2X_{ij} - X_{i+\xi_1, j+\xi_2}) \end{aligned}$$

where $\kappa_k := 1$ if $k = 0$ and $\kappa_k := 0$ if $k \neq 0$.

- $b(w)$ is defined by

$$b_{i,j}(w) \;:=\; \frac{\alpha_{ij,1}^2}{\Delta t} v_h(t+\Delta t, x_i, y_j) + \alpha_{ij,1}^2 \frac{\mu(t,y_j)}{h} \kappa_{j-1} g(x_i) + \frac{1}{2} \sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}} \kappa_{j-\xi_2} g(x_{i-\xi_1}) \tag{4.5}$$

where $v_h(t+\Delta t, x, y)$ is the solution at the previous time step and is assumed to be bounded.

We shall also denote

$$\delta_{ij}(w) := \frac{\alpha_{ij,1}^2}{\Delta t} + \frac{\alpha_{ij,1}^2}{h}\mu(t,y_j)\kappa_{j-1} + \frac{1}{2}\sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}} \kappa_{j-\xi_2}.$$

Remark IV.13. *The matrix $A(w)$ is $\delta(w)$-diagonal dominant in the following sense :*

$$A_{(i,j),(i,j)}(w) = \delta_{ij}(w) + \sum_{(k,\ell)\neq(i,j)} |A_{(i,j),(k,\ell)}(w)|.$$

Remark IV.14. *Note that in the case no border points $y=0$ are involved (i.e. when $j > p_{max}$), we have the more simple expressions :*

$$(A(w)X)_{ij} \;:=\; \frac{\alpha_{ij,1}^2}{\Delta t} X_{ij} + \frac{\alpha_{ij,1}^2}{h}\mu(t,y_j)(X_{ij} - X_{i,j-1}) + \frac{1}{2}\sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}} (-X_{i-\xi_1,j-\xi_2} + 2X_{ij} - X_{i+\xi_1,j+\xi_2}),$$

and

$$b_{i,j}(w) := \frac{\alpha_{ij,1}^2}{\Delta t} v_h(t+\Delta t, x_i, y_j), \qquad \delta_{ij}(w) := \frac{\alpha_{ij,1}^2}{\Delta t}.$$

Remark IV.15. *Note that on the boundary $x=0$, if we assume that $v_h(t+\Delta t, 0, y) = g(0)$ then the scheme reads*

$$\min_{\alpha_1^2+\alpha_2^2=1} \left\{ \alpha_1^2 \frac{v_h(t,0,y) - g(0)}{\Delta t} + \alpha_1^2 \mu(t,y)\, \frac{v_h(t,0,y) - v_h(t,0,y-h)}{h} + \frac{1}{2}\alpha_2^2(-v_h(t,0,y-h) + 2v_h(t,0,y) - v_h(t,0,y+h)) \right\} = 0, \quad \forall y \in h\mathbb{N}^* \tag{4.6}$$

and with $v_h(t,0,0) = g(0)$. One can show that $v_h(t,0,y) = const := g(0)$ is the only bounded solution of (4.6) (using the results of Lemma 6.1, Proposition IV.21 and Proposition IV.23). Hence by recursion we see that $v_h(t,0,y) = g(0)$ for all t and $y \in h\mathbb{N}$. In order to simplify the presentation of $A(w)$ and $b(w)$ we have preferred not to add this knowledge in a boundary condition at $x = 0$.

Preliminary results. In order to find a solution of (4.4), we first consider the linear system

$$A(w)X = b(w),$$

for a given $w \in \mathcal{A}$. For expository reasons, some specific results for such systems have been postponed to the appendix. We can check that $(A(w), b(w))$ satisfy all the assumptions of Proposition IV.22. In particular, we can check that the matrix $A(w)$ is *monotone*, in the sense that if $X = (X_{i,j})_{i\geq 0, j\geq 1}$ is bounded (or bounded from below) and such that

$$\forall i \geq 0,\ \forall j \geq 1, \quad \delta_{ij}(w) = 0\ \Rightarrow (A(w)X)_{ij} = 0, \tag{4.7}$$

then

$$A(w)X \geq 0 \Rightarrow\ X \geq 0.$$

Here (4.7) is equivalent to

$$\forall i \geq 0,\ j \geq 1, \quad \alpha_{ij,1} = 0\ \Rightarrow -X_{i,j-1} + 2X_{ij} - X_{i,j+1} = 0.$$

Since $b(w)$ satisfies $\delta_{ij}(w) = 0 \Rightarrow b_{i,j}(w) = 0$, and that

$$\max_{i,j;\delta_{ij}(w)>0} \frac{|b_{ij}(w)|}{\delta_{ij}(w)} \leq \max(||v_h(t+\Delta t,.,.)||_\infty,\, ||g||_\infty),$$

we also obtain by Proposition IV.22 (ii) that there exists a unique bounded X such that $A(w)X = b(w)$, and satisfying furthermore

$$||X||_\infty := \max_{i\geq 0, j\geq 1} |X_{ij}| \leq \max(||v_h(t+\Delta t,.,.)||_\infty,\ ||g||_\infty).$$

Howard algorithm We can now consider the following Howard algorithm for solving (4.4).

Let $w^0 \in \mathcal{A}$ be a given initial control value

Iterate for $k \geq 0$,

- Find X^k bounded, such that $A(w^k)X^k = b(w^k)$.
- $w^{k+1} := \operatorname{argmin}_{w\in\mathcal{A}}(A(w)X^k - b(w))$.

In the second step note that the minimization is done component by component, since $(A(w)X^k - b(w))_{ij}$ depends only of the control α_{ij}; the minimum is also well defined since the control set $\bar{\mathcal{B}}$ for α_{ij} is compact.

Then we have the following result, whose proof is postponed to the appendix.

Proposition IV.16. *There exists a unique bounded solution X to the problem*

$$\min_{w\in\mathcal{A}}(A(w)X - b(w)) = 0,$$

and the sequence X^k converges pointwisely towards X, i.e., $\lim_{k\to\infty} X^k_{ij} = X_{ij}\ \forall i, j \geq 0$.

Stability and monotonicity. First, the convergence proved in the previous proposition leads also to the bound $||v_h(t,.)||_\infty = ||X||_\infty \leq \max(||v_h(t+\Delta t,.)||_\infty, ||g||_\infty)$. Hence by recursion we obtain $||v_h(t,.)||_\infty \leq ||g||_\infty$, which shows the stability of the scheme.

Then, the monotonicity is also obtained directly from the definition of the scheme.

Remark IV.17. *Note that we have the following stronger monotonicity result : if $v_h^1(t+\Delta t)$ and $v_h^1(t+\Delta t)$ are two bounded vectors defined on the grid, and X^1 and X^2 denotes the two corresponding solutions of (4.4), then*

$$v_h^1(t+\Delta t,.) \leq v_h^2(t+\Delta t,.) \quad \Rightarrow \quad X^1 \leq X^2.$$

To see this, let us denote $b^q(w)$, for $q=1,2$, the vectors corresponding to $v_h^q(t+\Delta t)$ as defined in (4.6). We note that $b^1(w) \leq b^2(w)$, $\forall w \in \mathcal{A}$. Let w^1 be an optimal control for X^1. Then

$$\begin{aligned} A(w^1)X^1 - b^1(w^1) &= 0 = \min_{w\in\mathcal{A}}(A(w)X^2 - b^2(w)) \\ &\leq A(w^1)X^2 - b^2(w^1) \\ &\leq A(w^1)X^2 - b^1(w^1), \end{aligned}$$

and thus $A(w^1)(X^2 - X^1) \geq 0$. By the monotonicity property of $A(w^1)$ and the fact that if $\delta_{ij}(w^1) = 0$ then $b_{ij}^2(w^1) - b_{ij}^1(w^1) = 0$, we conclude to $X^1 \leq X^2$.

Remark IV.18. *Note that since we deal with an implicit scheme, the stability and monotonicity results are obtained inconditionnaly with respect to the mesh sizes h and $\Delta t > 0$, as expected.*

5 Convergence

Since the scheme is monotone, stable and consistent, we can use the same arguments as in [5, Theorem 2.1] to conclude the convergence of v_h toward ϑ, taking into account the comparison principle Theorem 2.5.

In order to prove this convergence, we first note that the following type of discrete comparison principle holds for the scheme.

Lemma 5.1. *Let $Y = Y_{h,\Delta t}(t,x,y)$ be defined on $(x,y) \in G_h$ and for $T - t \in \Delta t\mathbb{N}$. Suppose that Y is a super-solution of the scheme (resp sub-solution of the scheme), in the following sense :*
(i) $\forall\, t+\Delta t \leq T$, $(x,y) \in G_h$, $y > 0$, $S^\rho(t,x,y,Y(t,x,y),Y) \geq 0$
(resp. $S^\rho(t,x,y,Y(t,x,y)) \leq 0$),
(ii) $\forall\,(x,y) \in G_h$, $Y(T,x,y) \geq g(x)$ (resp $Y(T,x,y) \leq g(x)$),
(iii) $\forall\, t \leq T$, $(x,y) \in G_h$, $Y(t,x,0) \geq g(x)$ (resp $Y(t,x,0) \leq g(x)$),
(iv) $Y(t,x,y)$ is bounded from below (resp. from above).
Then $Y \geq v_h$ (resp $Y \leq v_h$), where $v_h = v_h(t,x,y)$ are the scheme values.

Proof. Indeed the proof can be obtained by recursion (using $Y(t+\Delta t,.) \geq v_h(t+\Delta t,.)$ to show that $Y(t,.) \geq v_h(t,.)$) following the same arguments as in Remark IV.17. In order to conclude from $A(w_1)(Y(t,.) - v_h(t,.)) \geq 0$ to $Y(t,.) - v_h(t,.) \geq 0$ (for a given control w_1), we use the fact that $Y(t,.) - v_h(t,.)$ is bounded from below and Prop IV.22 1). The proof for the sub-solution is similar. □

We can give now the convergence result.

Theorem 5.2. *We assume (A1)-(A3) and that g is C^2-regular and such that $-x^2 g''(x)$ be bounded from below. Suppose also that $(p_{max}h)$ satisfies the assumptions of Proposition IV.11. Then the scheme solution v_h converges locally uniformly to ϑ when $h, \Delta t \to 0$.*

Proof. In the following when we denote $h \to 0$ we also mean that $\Delta t \to 0$. Let $\bar{v}$ and $\underline{v}$ be defined by

$$\overline{v}(t,x,y) = \limsup_{h,\Delta t \to 0, (t',x',y') \to (t,x,y)} v_h(t',x',y'),$$
$$\underline{v}(t,x,y) = \liminf_{h,\Delta t \to 0, (t',x',y') \to (t,x,y)} v_h(t',x',y'),$$

(The function $v_h(t,x,y)$ defined for (x,y) in the grid G_h and for $T-t = n\Delta t$ can be extended to $[T,0] \times \mathbb{R}^+ \times \mathbb{R}^+$ by a P0 interpolation.) As in [5, Theorem 2.1], using properties (S1-S3) of the scheme, we can prove that $\bar{v}$ and $\underline{v}$ are respectively bounded viscosity sub- and super-solution of (3.1). If the following inequalities hold :

$$\overline{v}(T,x,y) \leq g(x) \leq \underline{v}(T,x,y) \tag{5.1}$$
$$\overline{v}(t,x,0) \leq g(x) \leq \underline{v}(t,x,0) \tag{5.2}$$

then, by the comparison principle (Theorem 2.5) we obtain $\overline{v} \leq \underline{v}$, hence $\overline{v} = \underline{v}$ and the convergence of v_h towards the unique viscosity solution of HJB, i.e. ϑ.

Now let us prove the claims (5.1) and (5.2).

Step 1 : $\underline{v}(T,x,y) \geq g(x)$, and $\underline{v}(t,x,0) \geq g(x)$.
Considering $Y(t,x,y) = g(x)$, we see that Y is a sub-solution of the scheme (3.10) (in the sense of Lemma 5.1). Hence $v_h \geq Y$ and we deduce the two inequalities $\underline{v}(T,x,y) \geq g(x)$ and $\underline{v}(t,x,0) \geq g(x)$.

Step 2 : $\bar{v}(T,x,y) \leq g(x)$, and $\bar{v}(t,x,0) \leq g(x)$.
Let $C \geq 0$ and $L \geq 0$ be some constants such that $-\frac{1}{2}x^2 g''(x) \geq -C$ and $\sigma^2(t,y) \leq Ly$. Let K be a constant such that $K \geq CL$, fix an arbitrary $\varepsilon > 0$ and, for $t \in [0,T]$, let

$$Y(t,x,y) := K\,(T-t)\,(y + \min(1;y)) + g(x).$$

Note that $-\frac{1}{2}x^2 \frac{1}{h^2}(-g(x-\xi_1 h) + 2g(x) - g(x+\xi_1 h)) = -\frac{1}{2}x^2 \xi_1^2 g"(\theta_{x,h}) \geq -C$ (for some $\theta_{x,h} \in [x - \xi_1 h, x + \xi_1 h]$). Then we can deduce that $S^\rho(t,x,y,Y(t,x,y),Y) \geq 0$, Y satisfies the assumptions (i)-(iv) of Lemma 5.1, and thus $Y \geq v_h$. In particular,

$$\overline{v}(T,x,y) = \limsup_{h \to 0, (t',x',y') \to (T,x,y)} v_h(t',x',y') \leq \limsup_{h \to 0, (t',x',y') \to (T,x,y)} Y(t,x,y) \leq g(x).$$

We prove that $\overline{v}(t,x,0) \leq g(x)$ in the same way. □

Appendix

6 Properties of some infinite linear system

In this section we give some basic results for solving some specific infinite matrix system that are involved in our scheme.

Notations. We say that $A = (a_{ij})_{1\le i,j}$, $i, j \in \mathbb{N}^*$, with $a_{ij} \in \mathbb{R}$ is an *infinite* matrix if $\{j \ge 1,\ a_{ij} \ne 0\}$ is finite $\forall i \ge 1$. If $X = (x_i)_{i\ge 1}$ then we denote $(AX)_i = \sum_{j\ge 1} a_{ij}x_j$. We also denote $X \ge 0$ if $x_i \ge 0$, $\forall i \ge 1$.

The following Lemma generalizes the monotony property of M-matrices.

Lemma 6.1 (monotony)**.** *Let $A = (a_{ij})_{1\le i,j}$ be a real infinite matrix such that*
(i) For all $i \ge 1$, $\exists \delta_i \ge 0$, $a_{ii} = \delta_i + \sum_{j\ne i} |a_{ij}|$,
(ii) $a_{ij} \le 0$ $\forall i \ne j$,
(iii) $\delta_1 > 0$,
(iv) $\forall i \ge 1$, $\sum_j j\ a_{ij} \ge 0$.
(v) $\forall i \ge 2$, if $\delta_i = 0$ then $\exists q_i > 0$ such that $(AX)_i = q_i(-x_{i-1} + 2x_i - x_{i+1})$.
Then A is monotone in the following sense : if $X = (x_i)_{i\ge 1}$ is bounded from below and such that $\forall i \ge 1$, $\delta_i = 0 \Rightarrow (AX)_i = 0$, then

$$AX \ge 0 \Rightarrow X \ge 0.$$

Remark IV.19. *Note that from Lemma 6.1 we deduce the uniqueness of bounded solutions of $AX = b$ for any b such that $\delta_i = 0 \Rightarrow b_i = 0$.*

Proof of Lemma 6.1. Let $m = \min_{i\ge 1} x_i$.

Step 1. We first assume that there exists $i \ge 1$ such that $m = x_i$. Then

$$0 \le a_{ii}x_i + \sum_{j\ne i} a_{ij}x_j = \delta_i x_i + \sum_{j\ne i} |a_{ij}|(x_i - x_j) \le \delta_i x_i$$

If $\delta_i > 0$, then $x_i \ge 0$. In the case $\delta_i = 0$, by assumption (v) we obtain that $m = x_i = x_{i-1} = x_{i+1}$. In particular the minimum m is also reached by x_{i-1}. Since $\delta_1 > 0$, by a recursion argument we will arrive at a point j such that $\delta_j > 0$ and thus $x_j \ge 0$.

Step 2. In the general case we consider $Y = (y_i)$ with $y_i := x_i + \varepsilon\ i$ for some $\varepsilon > 0$. We note that $y_i \to +\infty$, hence $i \to y_i$ has a minimum. Also, $(AY)_i = (AX)_i + \varepsilon \sum_j j\ a_{ij} \ge 0$. Hence $AY \ge 0$ and $Y \ge 0$ by Step 1. Since this is true for any $\varepsilon > 0$, we conclude that $X \ge 0$. □

Remark IV.20. *Note that in Lemma 6.1 we can relax the assumption (x_i) bounded from below by $\liminf_{i\to\infty} \frac{x_i}{i} \ge 0$.*

Proposition IV.21 (Existence of solutions for linear systems)**.** *We consider A, an infinite matrix, such that*

(i) $\forall i \geq 1$, $\exists \delta_i \geq 0$, $a_{ii} = \delta_i + \sum_{j \neq i} |a_{ij}|$,
(ii) $\delta_1 > 0$.
(iii) $\forall i \geq 2$, *if* $\delta_i = 0$ *then* $\exists q_i > 0$, *such that* $(AX)_i = q_i(-x_{i-1} + 2x_i - x_{i+1})$.
Let also $b = (b_i)_{i \geq 1}$ *be such that*

$$\forall i, \delta_i = 0 \Rightarrow b_i = 0, \qquad \textit{and} \qquad \max_{k \geq 1,\ \delta_k \neq 0} \frac{|b_k|}{\delta_k} < \infty.$$

Then there exists a unique X, *in the space of bounded sequences, such that* $AX = b$, *and furthermore we have*

$$\max_{k \geq 1} |x_k| \leq \max_{k \geq 1,\ \delta_k \neq 0} \frac{|b_k|}{\delta_k}.$$

Proof. We look for solutions $x^{(n)} = (x_1^{(n)}, \ldots, x_n^{(n)})^T \in \mathbb{R}^n$ of the first n linear equations of $AX = b$, and set also $x_k^{(n)} := 0$, $\forall k > n$. (Dirichlet type boundary conditions on the right border). This leads to solve the finite dimensional system

$$A^{(n)} x^{(n)} = b^{(n)} \tag{6.1}$$

where $A^{(n)} := (a_{ij})_{1 \leq i,j \leq n}$ and $b^{(n)} := (b_1, \ldots, b_n)^T$,

Lemma 6.2. *There exists a unique* $x^{(n)}$ *solution of* (6.1) *and furthermore it satisfies the inequality*

$$\max_{1 \leq k \leq n} |x_k^{(n)}| \leq \max_{1 \leq k \leq n,\ \delta_k \neq 0} \frac{|b_k|}{\delta_k}. \tag{6.2}$$

Proof of Lemma 6.2. Suppose that $x^{(n)}$ exists, and let i be such that $|x_i^{(n)}| = \max_{1 \leq j \leq n} |x_j^{(n)}|$. Note that we still have $\forall 1 \leq i \leq n$, $a_{ii}^{(n)} = \delta_i + \sum_{j \neq i} |a_{ij}^{(n)}|$. If $\delta_i > 0$,

$$|b_i| \geq |a_{ii}^{(n)} x_i^{(n)}| - \sum_{j \neq i} |a_{ij}^{(n)}||x_j^{(n)}| \geq \delta_i |x_i^{(n)}|$$

thus $|x_i^{(n)}| \leq \frac{|b_i|}{\delta_i}$. If $\delta_i = 0$, we consider

$$i_0 := \sup\{k < i,\ \delta_k > 0\}.$$

(i_0 exists since $\delta_1 > 0$). Then $-x_{k-1}^{(n)} + 2x_k^{(n)} - x_{k+1}^{(n)} = b_k/q_k = 0$ for $k = i_0 + 1, \ldots, i$, and $x_{k+1}^{(n)} - x_k^{(n)} = \textit{const} = c_0$ for $k = i_0, \ldots, i$. But $x_i^{(n)}$ is an extremum of $x_{i-1}^{(n)}$, $x_i^{(n)}$ and $x_{i+1}^{(n)}$. This implies that $x_{i-1}^{(n)} = x_i^{(n)} = x_{i+1}^{(n)}$, and thus $c_0 = 0$ and $x_{i_0}^{(n)} = x_i^{(n)}$ is also an extremum. Since $\delta_{i_0} > 0$, we can estimate $|x_{i_0}^{(n)}|$ as before. This implies the invertibility of $A^{(n)}$, and thus the uniqueness of $x^{(n)}$. □

Now we shall prove that the sequence $X^{(n)} = (x^{(n)}, 0, 0, \ldots)^T$, which satisfies already $||X^{(n)}||_\infty \leq C := \max_{\delta_k \neq 0} \frac{|b_k|}{\delta_k}$, converges pointwisely towards a solution X of the problem.

We first suppose that $b \geq 0$. We can see that $A^{(n)}$ is still a monotone matrix (following the proof of Lemma 6.1). Hence $x^{(n)} \geq 0$. Now we consider $x^{(n+1)}$ and for $i \leq n$ we see that

$$\left(A^{(n)}x^{(n+1)}\right)_i = b_i - a_{i,n+1}x^{(n+1)}_{n+1} \geq b_i = \left(A^{(n)}x^{(n)}\right)_i.$$

Hence we obtain that

$$(x_1^{(n+1)}, \dots, x_n^{(n+1)})^T \geq (x_1^{(n)}, \dots, x_n^{(n)})^T,$$

and in particular $X^{(n)} \leq X^{(n+1)}$. Since $||X||_\infty \leq C$, we obtain the (pointwise) convergence of $X^{(n)}$ towards some vector X such that $||X||_\infty \leq C$. In the general case, we can decompose $b = b^+ - b^-$ with $b^+ = \max(b, 0)$, $b^- = \max(-b, 0)$, and proceed in the same way. We obtain the pointwise convergence of $X^{(n)} = X^{(n),+} - X^{(n),-}$ towards some X, with $X^{(n),\pm} \geq 0$ and $||X^{(n),\pm}||_\infty \leq C$, hence also $||X||_\infty \leq C$.

Since $\{j,\ a_{ij}^{(n)} \neq 0\}$ is finite, for any given i we can pass to the limit $n \to \infty$ in $\sum_{j\geq 1} a_{ij}^{(n)} x_j^{(n)} = b_i$, and obtain $(AX)_i = b_i$. □

Case of infinite 2d matrices. We say that the set of real numbers $A = (A_{(i,j),(k,\ell)})_{1\leq i,j,k,\ell}$ is an *infinite 2d* matrix if $\{(k,\ell),\ A_{(i,j),(k,l)} \neq 0\}$ is finite $\forall i, j \geq 1$ (A is also an "infinite" tensor). If $X = (X_{i,j})_{i,j\geq 1}$ then we denote $(AX)_{i,j} = \sum_{k,\ell\geq 1} A_{(i,j),(k,\ell)}X_{k,\ell}$. We also denote $X \geq 0$ if $X_{i,j} \geq 0$, $\forall i, j$.

The previous results can be easily generalized to infinite 2d matrices. We state here the results without proof.

Proposition IV.22. *Let $A = (A_{(i,j),(k,\ell)})_{1\leq i,j,k,\ell}$ be an infinite 2d matrix such that*
(i) For all $i, j \geq 1$, $A_{(i,j),(i,j)} = \delta_{ij} + \sum_{(k,\ell)\neq(i,j)} |A_{(i,j),(k,\ell)}|$ with $\delta_{ij} \geq 0$,
(ii) $A_{(i,j),(k,\ell)} \leq 0\ \forall (i,j) \neq (k,\ell)$,
(iii) $\delta_{i1} > 0$, $\forall i \geq 1$,
(iv) $\forall i, j \geq 1$, $\sum_{(k,\ell)}(k+\ell)A_{(i,j),(k,\ell)} \geq 0$,
(v) $\forall i \geq 1$, $\forall j \geq 2$, if $\delta_{ij} = 0$ then $\exists q_{ij} > 0$ such that

$$(AX)_{ij} = q_{ij}(-X_{i,j-1} + 2X_{i,j} - X_{i,j+1}).$$

1) Then A is monotone in the following sense : if $X = (X_{i,j})_{i,j\geq 1}$ is bounded from below and such that $\forall i, j \geq 1$, $\delta_{i,j} = 0 \Rightarrow (AX)_{i,j} = 0$, then

$$AX \geq 0 \Rightarrow X \geq 0.$$

2) If $b = (b_{ij})_{i,j\geq 1}$ is such that $\delta_{ij} = 0 \Rightarrow b_{i,j} = 0$, and $\max_{i,j\geq 1,\ \delta_{ij}>0} \frac{|b_{ij}|}{\delta_{ij}} < \infty$, then there is a unique bounded X such that $AX = b$, and furthermore

$$\max_{i,j\geq 1} |X_{ij}| \leq \max_{i,j\geq 1,\ \delta_{ij}>0} \frac{|b_{ij}|}{\delta_{ij}}.$$

7 Convergence of the Howard algorithm

In this section we prove the following result.

Proposition IV.23. *Let S be a compact set, and $\mathcal{A} := S^{\mathbb{N}}$, the set of infinite sequences of S. For all $w \in \mathcal{A}$, let $A(w) := (a_{ij}(w))_{i,j\geq 1}$ be an infinite matrix, and $b(w) := (b_i(w))_{i\geq 1}$. We assume furthermore that*
(i) If $w = (w_i)_{i\geq 1}$, $a_{ij}(w)$ depends only of w_i, and also $b_i(w)$ depends only of w_i, and this dependence is continuous.
(ii) $\forall i$, $\sup_{w\in\mathcal{A}} \big(Card\{j,\ a_{ij}(w) \neq 0\}\big) < \infty$.
(iii) (monotony) For all $w \in \mathcal{A}$ and X bounded,

$$A(w)X \geq 0 \quad \Rightarrow \quad X \geq 0.$$

(iv) $\exists C \geq 0$, $\forall w \in \mathcal{A}$, $\exists X$ solution of $A(w)X = b(w)$ and such that

$$||X||_\infty \leq C.$$

Then
(i) there exists a unique bounded solution X to the problem

$$\min_{w\in\mathcal{A}}(A(w)X - b(w)) = 0. \tag{7.1}$$

(ii) the Howard algorithm as defined in section 4 converges pointwisely towards X.

Remark IV.24. *Proposition IV.23 can then be adapted in order to prove Proposition IV.16. The proof is left to the reader.*

Proof. Let us first check the uniqueness. Let X and Y be two solutions, and let $\bar{w}$ be an optimal control associated to Y. Then

$$\begin{aligned} A(\bar{w})Y - b(\bar{w}) &= 0 \\ &= \min_{w\in\mathcal{A}}(A(w)X - b(w)) \\ &\leq A(\bar{w})X - b(\bar{w}). \end{aligned}$$

Hence $A(\bar{w})(Y - X) \leq 0$ and thus $Y \leq X$ using the monotony property. We can prove $Y \geq X$ in the same way, hence $X = Y$ which proves uniqueness.

The existence now is obtained by considering the sequence X^k and controls w^k as in the Howard algorithm of section 4.

We first remark that for all $k \geq 0$, $X^k \leq X^{k+1}$, because

$$\begin{aligned} A(w^{k+1})X^{k+1} - b(w^{k+1}) &= 0 \\ &= A(w^k)X^k - b(w^k) \\ &\geq \min_w(A(w)X^k - b(w)) \\ &\geq A(w^{k+1})X^k - b(w^{k+1}) \end{aligned}$$

and using the monotony of $A(w^{k+1})$. Also, X^k is bounded. Hence X^k converges pointwisely towards some bounded X. It remains to show that X satisfies (7.1).

Let $F_i(X)$ be the i-th component of $\min_{w\in\mathcal{A}}(A(w)X - b(w))$, i.e.

$$F_i(X) = \min_{w\in\mathcal{A}}(A(w)X - b(w))_i$$

For a given i, since $(A(w)X)_i$ involves only a finite number of matrix continuous coefficients $(a_{ij}(w))_{j\leq j_{\max}}$, we obtain that $\lim_{k\to\infty} F_i(X^k) = F_i(X)$. Also by compactness of S, by a diagonal extraction argument, there exists a subsequence of $(w^k)_{k\geq 0}$, denoted w^{ϕ_k}, that converges pointwisely towards some $w \in \mathcal{A}$.

Passing to the limit in $(A(w^{\phi_k})X^{\phi_k} - b(w^{\phi_k}))_i = 0$, we obtain $(A(w)X - b(w))_i = 0$. On the other hand,

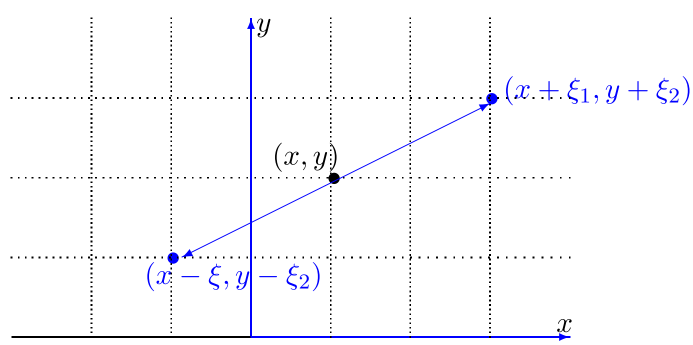
$$\begin{aligned} F_i(X) &= \lim_{k\to\infty} F_i(X^{\phi_k - 1}) \\ &= \lim_{k\to\infty} \Big(A(w^{\phi_k})X^{\phi_k} - b(w^{\phi_k})\Big)_i \\ &= \big(A(w)X - b(w)\big)_i \end{aligned}$$

Hence $F_i(X) = 0$, $\forall i$, which concludes the proof. □

8 Points on the boundary

We present in this section another way to consider the point near to the boundary in the discretization of the second order term. Consider the grid points that are close to the boundaries $x = 0$, $y = 0$. Fixes an order p_{max}, for theses points, the discretization of the second order term could involve some author points which are out of the grid.

Then we modify the expression of the elementary diffusion. Let us explain this modification on a simple example drawn in the following picture

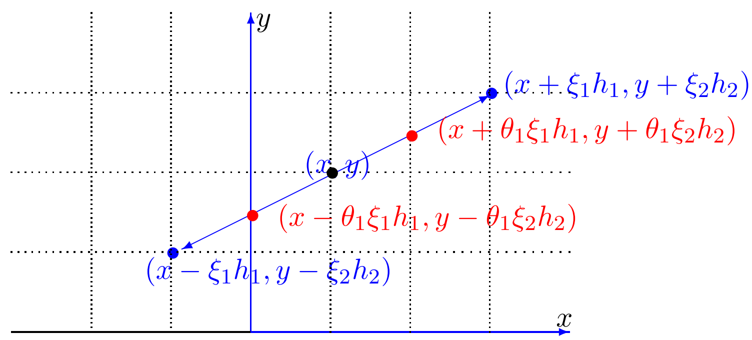


The direction of the diffusion (the vector →) points toward a grid points $(x\pm\xi_1 h_1, y\pm\xi_2 h_2)$ in the neighborhood of order 2. However, $(x - \xi_1 h_1, y - \xi_2 h_2)$ is out of the grid, which is delimited by the positive part of the x-axis and the positive part of the y-axis.

Introduce a parameter $\theta_1 \in [0,1]$ and the associated point $(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2)$. The real θ_1 is chosen in such a way that the point $(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2)$ is on the diffusion direction and belonging to the y-axis (the intersection between the y-axis and the vector formed by (x,y) and $(x-\xi_1 h_1, y-\xi_2 h_2)$). Although $(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2)$ is not a grid point, we will use it in the scheme, because the function v_h is known on the axis $x=0$. The elementary diffusion becomes

$$\Delta_{\xi,\theta}\phi(x,y) = \frac{\phi(x+\theta_1\xi_1 h_1, y+\theta_1\xi_2 h_2) + \phi(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2) - \phi(x,y)}{\theta_1^2}, \tag{8.1}$$

where $\theta_1 \in [0,1]$ is chosen such that $(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2)$ and $(x+\theta_1\xi_1 h_1, y+\theta_1\xi_2 h_2)$ are in the domain $[0,+\infty)^2$.



In general case, for ξ is in the stencil $\mathcal{S}_{p_{\max}}$ and (x,y) in the grid G_h, if the points $(x\pm\xi_1 h_1, y\pm\xi_2 h_2)$ should be used in the approximation of the covariance matrix and if thery are out of the domain $[0,+\infty)^2$, then we modify the elementary diffusion Δ_ξ by :

$$\Delta_{\xi,\theta}\phi(x,y) = \frac{\theta_2\phi(x+\theta_1\xi_1, y+\theta_1\xi_2) + \theta_1\phi(x-\theta_2\xi_1, y-\theta_2\xi_2) - 2(\theta_1+\theta_2)\phi(x,y)}{\theta_1\theta_2(\theta_1+\theta_2)},$$

where $\theta_1, \theta_2 \in [0,1]$ are such that $(x+\theta_1\xi_1 h_1, y+\theta_1\xi_2 h_2)$ and $(x-\theta_2\xi_1 h_1, y-\theta_2\xi_2 h_2)$ are in the domain.

Therefore, the scheme (3.4) should be written :

$$v_h(T,x,y) = g(x) = v_h(t,x,0), \quad v_h(t,0,y) = g(0), \tag{8.2}$$

$$\min_{\alpha_1^2+\alpha_2^2=1}\{-\alpha_1^2\delta_t v_h(t,x,y) - \alpha_1^2\mu\delta_y v_h(t,x,y) - \frac{1}{2}\sum_{\xi\in\mathcal{S}}\gamma_\xi^{\alpha_1,\alpha_2}\Delta_{\xi,\theta}v_h(t,x,y)\} = 0, \tag{8.3}$$

for $t < T-\Delta t$.

This scheme satisfies consistency property, in the sense of Proposition IV.11.

Bibliographie

[1] M. Akian, J.L. Menaldi, and A. Sulem. On an investement-consumption model with transaction costs. *SIAM J. Control Optim.*, 34 :329–364, 1996.

[2] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems and Control : Foundations and Applications. Birkhäuser, Boston, 1997.

[3] G. Barles. An approach of deterministic control problems with unbounded data. *Ann. Inst. H. Poincaré, Anal. Non Linéaire*, 7(4) :235–258, 1990.

[4] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume 17 of *Mathématiques et Applications*. Springer, Paris, 1994.

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

[6] J.F. Bonnans, E. Ottenwaelter, and H. Zidani. Numerical schemes for the two dimensional second-order HJB equation. *ESAIM : M2AN*, 38 :723–735, 2004.

[7] J.F. Bonnans and H. Zidani. Characterization of consistency of some numerical schemes for the stochastic HJB equation. In J.L. Menaldi, E. Rofman, and A. Sulem, editors, *Optimal Control and PDE - Innovations et Applications*. IOS Press, Amsterdam, 2000.

[8] B. Bruder. Superreplication of european options with a derivative asset under constrained finite variation strategies. preprint, 2005.

[9] J.-Ph. Chancelier, B. Øksendal, and A. Sulem. Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. *Tr. Mat. Inst. Steklova*, 237(Stokhast. Finans. Mat.) :149–172, 2002.

[10] P. Cheridito, H. Soner, and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Annales de l'Institut Henri Poincaré : Analyse Non Linéaire*. To appear.

[11] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. American Mathematical Society (New Series)*, 27 :1–67, 1992.

[12] J. Cvitanic, H. Pham, and N. Touzi. Super-replication in stochastic volatility models with portfolio constraints. *Journal of Applied Probability*, 36 :523–545, 1999.

[13] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics, A Foundation For Computer Science*. Addison-Wesley, Paris, 1994.

[14] N. El Karoui and M.C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.*, 33(1) :29–66, 1995.

[15] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 2 : viscosity solutions and uniqueness. *Communications in partial differential equations*, 8 :1229–1276, 1983.

[16] H. Pham. On some recent aspects of stochastic control and their applications. *Probability Surveys*, 2 :1–549, 2005.

[17] M. Soner and N. Touzi. Super replication under gamma constraints. *SIAM J. on Control and Optimization*, 39 :73–96, 2000.

[18] M. Soner and N. Touzi. Stochastic target problems, dynamic programming and viscosity solutions. *SIAM J. on Control and Optimization*, 41 :404–424, 2002.