



HAL
open science

I. Pénalisation des problèmes de Neumann non linéaires via les systèmes d'EDS-EDSR II. Inégalité de transport pour les EDS et EDSR

Soufiane Mouchtabih

► To cite this version:

Soufiane Mouchtabih. I. Pénalisation des problèmes de Neumann non linéaires via les systèmes d'EDS-EDSR II. Inégalité de transport pour les EDS et EDSR. Mathématiques générales [math.GM]. Université de Toulon; Université Cadi Ayyad (Marrakech, Maroc). Faculté des sciences Semlalia, 2021. Français. NNT : 2021TOUL0016 . tel-04695807

HAL Id: tel-04695807

<https://theses.hal.science/tel-04695807>

Submitted on 12 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale 548 - Mer et Sciences. UTLN
Centre des Études Doctorales Sciences et
Techniques. FSSM-UCA

THÈSE présentée par :

Soufiane MOUCHTABIH

Pour obtenir le grade de

DOCTEUR EN MATHÉMATIQUES

Spécialité : ANALYSE ET PROBABILITÉS

I. Pénalisation des problèmes de Neumann
non-linéaires via les systèmes d'EDS-EDSR
II. Inégalité de transport pour les EDS et
EDSR

Préparée dans le cadre d'une cotutelle entre

L'UNIVERSITÉ DE TOULON ET L'UNIVERSITÉ CADI AYYAD MARRAKECH

JURY

M. Bahlali Khaled	MCF-HDR, Université de Toulon	Directeur de thèse
M. BOUFOUSSI Brahim	Professeur, Université Cadi Ayyad Marrakech	Directeur de thèse
M. BUCKDAHN Rainer	Professeur, Université de Bretagne Occidentale, Brest	Rapporteur
M. DELARUE François	Professeur, Université de Nice Sophia-Antipolis	Examinateur
M. ERRAOUI Mohamed	Professeur, Université Chouaib Doukkali El Jadida	Examinateur
M. LEJAY Antoine	MCF-HDR, INRIA, Institut Elie Cartan de Lorraine	Rapporteur
M. TOUZI Nizar	Professeur, Ecole polytechnique Paris	Examinateur
M. OUKNINE Youssef	Professeur, Université Cadi Ayyad Marrakech	Président et Rapporteur

Remerciements

Je tiens à présenter ma gratitude à mes directeurs de thèse, M. Brahim Boufoussi et M. Khaled Bahlali. Je les remercie infiniment pour leur confiance, leur temps et leur soutien précieux durant ces années de thèse.

J'exprime mes plus profonds remerciements à M. Brahim BOUFOUSSI pour sa confiance qui m'a accordé depuis le PFE du master, sa disponibilité et son aide. Je le remercie pour ces qualités humaines et sa sympathie. Ses conseils pertinents m'ont été de grand intérêt.

J'adresse de chaleureux remerciements à M. Khaled BAHALALI pour sa disponibilité, pour ses conseils avisés et son écoute. Je le remercie également pour sa gentillesse et l'aide qu'il m'a apporté dans tous les aspects de mes séjours à Toulon.

Je remercie mes collaborateurs, M. Salah HAJJI et M. Ludovic TANGPI.

J'exprime ma gratitude à M. Rainer BUCKDAHN, M. Antoine LEJAY et M. Youssef OUKNINE pour l'honneur qu'ils m'ont fait en acceptant de rapporter cette thèse. La version finale de cette thèse a bénéficié de leur lecture minutieuse et de leurs commentaires précieux.

Je présente mes sincères remerciements à M. François DELARUE, M. Mohamed ERRAOUI et M. Nizar TOUZI qui ont bien voulu être examinateurs.

Je n'oublierais pas de remercier Adrian Zalinescu et Cédric GALUSINSKI pour les discussions sur les résultats du chapitre 2.

Un grand merci aux membres des laboratoires LIBMA et IMATH que j'ai eu à côtoyer ces quatre dernières années.

J'adresse mes sincères remerciements à mes amis pour leur soutien moral et leur disponibilité au moment difficile tout au long de la thèse.

Et enfin, une immense merci à ma famille pour leurs sacrifices si précieuses, leur soutien et leur encouragement.

Liste des travaux ayant contribué à la rédaction de la thèse

- K. Bahlali, B. Boufoussi and S. Mouchtabih, Transportation cost inequality for backward stochastic differential equations, *Statistics and Probability Letters*, Volume 155, December 2019, 108586. Chapitre [6](#)
- K. Bahlali, S. Mouchtabih and L. Tangpi, Quadratic transportation inequalities for SDEs with measurable drifts. Accepté pour publication dans *Proceedings of the AMS*. Chapitre [5](#)
- K. Bahlali, B. Boufoussi and S. Mouchtabih, Penalization for a PDE with a Nonlinear Neumann boundary condition and measurable coefficients. Accepté pour publication dans *Stochastics and Dynamics*. Chapitre [2](#)
- K. Bahlali, B. Boufoussi and S. Mouchtabih, Approximation of a degenerate semilinear PDE with a nonlinear Neumann boundary condition. Chapitres [3](#) et [4](#)

Résumé

Cette thèse est composée de deux parties. Dans la première, nous nous intéressons aux problèmes de Neumann non linéaires. En utilisant des outils probabilistes, nous établissons des méthodes de pénalisations pour des équations aux dérivées partielles (EDP) semi-linéaires avec conditions de Neumann non linéaires. Premièrement, nous construisons une solution de L^p -viscosité d'une EDP avec condition de Neumann et à coefficients mesurables comme étant limite d'une suite de solutions au sens de L^p -viscosité des EDP avec conditions terminales. Puis nous traitons le cas de la solution de viscosité d'une EDP avec une non-linéarité dépendant du gradient de la solution. Dans le même esprit, nous démontrons un résultat de stabilité pour la solution de viscosité d'un problème de Neumann non-linéaire.

La deuxième partie concerne l'inégalité de transport quadratique. Dans un premier temps, nous traitons cette inégalité pour les EDS avec des dérivées mesurables vérifiant une condition d'intégrabilité et des coefficients de diffusions appartenant à un espace de Sobolev approprié. Ensuite, nous étendons l'inégalité de transport au cas des équations différentielles stochastiques rétrogrades. Dans les deux cas, nous en déduisons des inégalités de concentration.

Mots clés : Diffusion réfléchie; Méthode de pénalisation; Solutions faibles des EDS; S -topologie; Convergence en loi; Équations différentielles stochastiques rétrogrades; Solutions de viscosité et L^p -viscosité pour EDP; Stabilité des EDP; Inégalités de concentration; Inégalités de transport; Transformation de Girsanov; Dérivées singulières; Régularité de Sobolev.

Abstract

This thesis consists of two parts. In the first one, we are interested in the non-linear Neumann problems. Using probabilistic techniques, we establish penalization methods for semi-linear partial differential equations (PDE) with non-linear Neumann conditions. First, we construct a L^p -viscosity solution of a PDE with Neumann condition and measurable coefficients as a limit of a sequence of L^p -viscosity solution of a sequence of PDE with terminal conditions. Then, we treat the case of an EDP with a non-linearity also depending on the gradient of the solution. In the same context, we demonstrate a stability result for the viscosity solution of a non-linear Neumann problem.

The second part concerns transportation inequalities. First, we establish the quadratic transportation inequality for stochastic differential equations with measurable drifts that verify an integrability condition and diffusion coefficients belonging to an appropriate Sobolev space. Then, we extend this inequality to the case of backward stochastic differential equations. In both cases we deduce concentration inequalities.

Keywords : Reflected diffusion; Penalization method; Weak solution of stochastic differential equation; S -topology; Convergence in law, Backward stochastic differential equations; Viscosity and L^p -viscosity solution for PDE; Stability for PDE; Concentration of measure; Transportation inequality; Girsanov transformation; Singular drifts, Sobolev regularity.

Contents

1	Introduction et principaux résultats	1
1.1	Problèmes de Neumann non linéaires	1
1.1.1	Pénalisation des problèmes de Neumann non-linéaires : Résultats connus	4
1.1.2	Pénalisation des problèmes de Neumann non-linéaires : Résultats nouveaux	6
1.1.2.1	Pénalisation des EDP avec conditions de Neu- mann non-linéaires et coefficients mesurables	6
1.1.2.2	Pénalisation des EDP dégénérées avec conditions de Neumann non-linéaires	10
1.1.3	Un résultat de stabilité des EDP dégénérées avec condi- tions de Neumann non-linéaires	14
1.2	Inégalités de coût de transport	18
1.2.1	Inégalité de transport pour les EDS avec dérivées mesurables	19
1.2.2	Inégalité de transport pour les EDSR	22
I	NONLINEAR NEUMANN BOUNDARY VALUE PROBLEMS	25
2	Penalization for a PDE with a nonlinear Neumann boundary condition and measurable coefficients	27
2.1	Introduction	27
2.2	Reflected stochastic differential equations	29
2.3	Backward stochastic differential equations	35
2.4	PDE with a nonlinear Neumann boundary condition	42
3	Approximation of a degenerate semilinear PDE with a nonlin- ear Neumann boundary condition	47
3.1	Introduction	48
3.2	Preliminaries and formulation of the problem	49
3.2.1	Penalization for reflected stochastic differential equation	50
3.2.2	Backward inequality	52
3.2.3	The BSDEs associated to the nonlinear Neumann problem	55
3.3	Penalization of the nonlinear Neumann PDE	57
3.3.1	Convergence of the penalized BSDE	57
3.3.2	Convergence of the penalized PDE	66

4	A Stability result for a nonlinear Neumann problem	69
4.1	Introduction	69
4.2	Preliminaries	70
4.2.1	Stability of Reflected SDEs	70
4.2.2	The BSDE associated to nonlinear Neumann PDE (4.1)	72
4.3	Main result	74
II	TRANSPORTATION COST INEQUALITIES	79
5	Quadratic transportation inequalities for SDEs with measurable drift	81
5.1	Introduction	81
5.2	Main results	83
5.3	Proofs of the main results	84
5.4	Examples	93
6	Transportation cost inequality for backward stochastic differential equations	95
6.1	Introduction	95
6.2	Transportation inequality for BSDEs	97
6.3	Concentration inequalities	101
6.4	Transportation cost inequality for FBSDE	102
A	<i>S</i>-topology	105
	Bibliographie	107

Chapitre 1

Introduction et principaux résultats

Cette thèse est composée de deux parties : la première est consacrée aux résultats de pénalisation et de stabilité des problèmes de Neumann non linéaires par les équations différentielles stochastiques rétrogrades (EDSR). La deuxième concerne les inégalités de coûts de transport pour les équations différentielles stochastiques progressives et rétrogrades.

1.1 Problèmes de Neumann non linéaires

L'objectif de cette première partie est d'utiliser des méthodes probabilistes basées sur la théorie des équations différentielles stochastiques rétrogrades (EDSR) pour des EDP avec conditions de Neumann au bord. Commençons par rappeler qu'une EDSR est une équation de la forme

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t \\ Y_T = \xi \end{cases} \quad (1.1)$$

où $(W_t)_{t \in [0, T]}$ est un mouvement brownien r -dimensionnel défini sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$, $(\mathcal{F}_t)_{t \in [0, T]}$ est la filtration générée par W augmentée par les \mathbb{P} -négligeables de \mathcal{F} , f est une fonction aléatoire progressivement mesurable définie sur $\Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^{d \times r}$, à valeurs dans \mathbb{R}^d , une telle fonction est appelée le générateur et ξ est une variable aléatoire \mathcal{F}_T -mesurable appelée condition terminale. La différence majeure avec une équation différentielle stochastique progressive est la donnée de la condition finale au lieu de la valeur initiale. Une solution à l'EDSR est une paire $(Y_t, Z_t)_{t \in [0, T]}$ de processus stochastiques adaptés. Le processus Y appelé variable d'état et Z est la variable de contrôle, ou processus de contrôle. L'EDSR décrit l'évolution du processus Y rétrograde dans le temps à partir de la condition terminale ξ , donc l'adaptabilité de Y n'est pas évidente et le rôle de Z est de corriger l'adaptabilité, ce qui permet l'existence des solutions adaptées. Le problème de l'existence et l'unicité d'une telle solution a été étudié en premier lieu par Pardoux et Peng dans [66], sous l'hypothèse de continuité de Lipschitz sur f par rapport aux variables

(y, z) , et une condition terminale ξ de carré intégrable. Depuis, de nombreux efforts ont été faits pour affaiblir les hypothèses sur les données de l'équation (ξ, f) . En outre, des formes plus générales de cette équation ont été envisagées.

Formule de Feynman-Kac

Considérons la diffusion suivante :

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) dr + \int_t^s \sigma(r, X_r^{t,x}) dW_r. \quad (1.2)$$

Il est connu que cette EDS fournit, par la formule de Feynmann-Kac, une solution probabiliste à l'EDP linéaire suivante :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \frac{1}{2} Tr(\sigma\sigma^* \Delta u)(t, x) + b \cdot \nabla u(t, x) - K(t, x)u(t, x) = 0, \\ \forall t \in [0, T], x \in \mathbb{R}^d, \\ u(T, x) = g(x). \end{cases} \quad (1.3)$$

Plus précisément, sous de bonnes conditions sur les coefficients b, σ, K, g , il est possible de montrer que la quantité :

$$u(t, x) = \mathbb{E} \left[\exp \left(- \int_t^T K(r, X_r^{t,x}) dr \right) g(X_T^{t,x}) \right]$$

est une solution classique de l'EDP (1.3). Dans le cas de coefficients moins réguliers u est une solution de viscosité de (1.3).

La théorie des EDSR permet de généraliser ce résultat au cas des EDP semi-linéaires. Dans [65], Pardoux et Peng ont établi une relation entre l'EDP suivante :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \frac{1}{2} Tr(\sigma\sigma^* \Delta u)(t, x) + b \cdot \nabla u(t, x) + f(t, x, u(t, x), (\nabla u \sigma)(t, x)) = 0 \\ u(T, x) = g(x) \end{cases} \quad (1.4)$$

et un système d'équations différentielles stochastiques progressives rétrogrades de la forme :

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) dr + \int_t^s \sigma(r, X_r^{t,x}) dW_r \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dW_r. \end{cases} \quad (1.5)$$

Ils ont établi que si u est une solution forte de l'EDP (1.4), alors $(Y_t, Z_t) := (u(s, X_s^{t,x}), \nabla u \sigma(s, X_s^{t,x}))$ est une solution de l'EDSR (1.5). Ils ont aussi démontré que la solution du système (1.5) induit une solution de viscosité à l'EDP (1.4) donnée par la formule de Feynman-Kac, $u(t, x) = Y_t^{t,x}$. En particulier, cette formule fournit une représentation probabiliste à la solution de l'EDP. La liaison entre les EDSR et les EDP s'utilise dans les deux sens : Les EDSR peuvent être utilisées pour obtenir des propriétés sur les EDP associées ou pour faire le contraire.

Il existe une littérature abondante sur les liens entre les EDSR et les EDP. Par exemple, El Karoui et al [27] et Cvitanic et Karatzas [19] ont connecté des

EDSR réfléchies avec des problèmes d'obstacles pour les EDP, Pardoux et Zhang [71] ont associé les EDSR généralisées aux EDP avec conditions de Neumann au bord, Kobylanski [39] et Briand et Hu [13] ont prouvé les liens entre les EDSR quadratiques et les EDP quadratiques. Le lien qui nous intéresse dans cette thèse est celui établi par Pardoux-Zhang entre les EDSR généralisées (EDSRG) et les EDP avec conditions de Neumann. Rappelons qu'une EDSRG c'est une EDSR plus une intégrale de Stieltjes contre un processus continu croissant. Plus précisément,

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt - h(t, Y_t)dk_t + Z_t dW_t \\ Y_T = \xi \end{cases} \quad (1.6)$$

où h est une fonction aléatoire progressivement mesurable et k est un processus réel continu et croissant. Dans [71], en considérant une condition terminale de carré intégrable, les auteurs ont démontré l'existence et l'unicité de la solution sous les hypothèses suivantes :

$$\begin{aligned} \langle y - y', f(t, y, z) - f(t, y', z) \rangle &\leq \alpha |y - y'|^2, & \alpha \in \mathbb{R}, \\ \langle y - y', h(t, y) - h(t, y') \rangle &\leq \beta |y - y'|^2, & \beta < 0, \\ |f(t, y, z) - f(t, y, z')| &\leq L |z - z'| & L > 0. \end{aligned}$$

Pardoux et Zhang ont couplé l'équation (1.6) avec une équation différentielle stochastique réfléchie, sur un domaine D de \mathbb{R}^d , comme suit :

$$\begin{cases} \forall s \in [t, T] \\ X_s^{t,x} = x + \int_t^s b(X_r^{t,x}) dr + \int_t^s \sigma(X_r^{t,x}) dW_r + \int_t^s \nabla l(X_r^{t,x}) dk_r^{t,x}, \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x} \\ \quad - \int_s^T Z_r^{t,x} dW_r, \end{cases} \quad (1.7)$$

et ils ont établi un lien avec un système des EDP avec condition de Neumann au bord de la forme :

$$\begin{cases} \frac{\partial u_i}{\partial t}(t, x) + \mathcal{L}u_i(t, x) + f_i(t, x, u(t, x), (\nabla u \sigma)(t, x)) = 0, \\ 1 \leq i \leq m, (t, x) \in [0, T) \times D, \\ u(T, x) = g(x), x \in D, \\ \frac{\partial u}{\partial n}(t, x) + h(t, x, u(t, x)) = 0, \forall (t, x) \in [0, T) \times \partial D, \end{cases} \quad (1.8)$$

où

$$\mathcal{L} = \frac{1}{2} \sum_{i,j} (\sigma \sigma^*(\cdot))_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(\cdot) \frac{\partial}{\partial x_i}. \quad (1.9)$$

En fait, ils ont prouvé que la quantité déterministe $u(t, x) = Y_t^{t,x}$ est une solution de viscosité du système (1.8).

1.1.1 Pénalisation des problèmes de Neumann non-linéaires : Résultats connus

Considérons la diffusion réfléchie suivante :

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(X_r^{t,x}) dr + \int_t^s \sigma(X_r^{t,x}) dW_r + K_s^{t,x}, \\ K_s^{t,x} = \int_t^s \nabla l(X_r^{t,x}) d|K^{t,x}|_{[t,r]}, \\ |K^{t,x}|_{[t,s]} = \int_t^s 1_{\{X_r^{t,x} \in \partial D\}} d|K^{t,x}|_{[t,r]}, \end{cases} \quad (1.10)$$

où D est un domaine de \mathbb{R}^d borné convexe et régulier. Nous supposons qu'il existe une fonction $l \in \mathcal{C}_b^2(\mathbb{R}^d)$ telle que

$$D = \{x \in \mathbb{R}^d : l(x) > 0\}, \quad \partial D = \{x \in \mathbb{R}^d : l(x) = 0\},$$

et pour tout $x \in \partial D$, $\nabla l(x)$ est le vecteur normal pointant vers l'intérieur de D . En vue de définir la procédure d'approximation, nous considérons l'application $x \mapsto \text{dist}^2(x, \bar{D})$, cette fonction est de classe \mathcal{C}^1 et convexe sur \mathbb{R}^d . En outre, nous pouvons choisir l de telle sorte que

$$\langle \nabla l(x), \delta(x) \rangle \leq 0, \quad \forall x \in \mathbb{R}^d,$$

où $\delta(x) := \nabla (\text{dist}^2(x, \bar{D}))$ est appelé le terme de pénalisation.

Il est connu que sous des conditions appropriées sur b et σ , la diffusion $X^{t,x}$ peut être approchée par une suite de solutions des EDS de la forme (voir [50, 56, 81]) :

$$X_s^{t,x,n} = x + \int_t^s [b(X_r^{t,x,n}) - n\delta(X_r^{t,x,n})] dr + \int_t^s \sigma(X_r^{t,x,n}) dW_r, \quad s \in [t, T], \quad (1.11)$$

Rappelons le résultat de [50, 56] à savoir :

$$\mathbb{E} \left[\sup_{t \leq s \leq T} |X_s^{t,x,n} - X_s^{t,x}|^q \right] \longrightarrow 0 \quad \text{quand } n \rightarrow \infty.$$

Cette convergence a une interprétation analytique (voir [50]). En particulier, si on considère le problème de Cauchy suivant :

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}u^n(t, x) - n\delta(x)\nabla u^n(t, x) = 0 \quad (t, x) \in (0, T] \times \mathbb{R}^d, \\ u^n(0, x) = u_0(x), \end{cases}$$

où \mathcal{L} , donné par (1.9), est le générateur infinitésimal qui correspond à la partie diffusion de $X^{t,x}$ et u_0 est une donnée initiale régulière. Alors pour tout $(t, x) \in [0, T] \times \bar{D}$, $u^n(t, x)$ converge vers $u(t, x)$, qui est la solution du problème de Neumann suivant :

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \mathcal{L}u(t, x) = 0, \quad (t, x) \in (0, T] \times D, \\ u(0, x) = u_0(x), \quad x \in D, \\ \frac{\partial u}{\partial n}(t, x) = 0, \quad \forall (t, x) \in (0, T] \times \partial D. \end{cases}$$

Dans [11], Boufoussi et J.V. Casteren ont généralisé ce résultat au cas d'une EDP semi-linéaire avec condition de Neumann non-linéaire de la forme :

$$\begin{cases} \frac{\partial u_i}{\partial t}(t, x) + \mathcal{L}u_i(t, x) + f_i(t, x, u(t, x)) = 0, & 1 \leq i \leq k, (t, x) \in [0, T) \times D, \\ u(T, x) = g(x), & x \in D, \\ \frac{\partial u_i}{\partial n}(t, x) + h_i(t, x, u(t, x)) = 0, & 1 \leq i \leq k, \forall (t, x) \in [0, T) \times \partial D. \end{cases} \quad (1.12)$$

Le résultat principal de [11] est l'approximation d'une solution de viscosité u du système (1.12) par une suite de solutions de viscosité satisfaisant des EDP de la forme :

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}u_i^n(t, x) + f_i(t, x, u^n(t, x)) \\ \quad - n \langle \delta(x), \nabla u_i^n(t, x) \rangle - n \langle \delta(x), \nabla l(x) \rangle + h_i(t, x, u^n(t, x)) = 0, \\ 1 \leq i \leq k, \forall (t, x) \in [0, T) \times \mathbb{R}^d, \\ u^n(T, x) = g(x) \quad \forall x \in \mathbb{R}^d. \end{cases} \quad (1.13)$$

Dans le même esprit, et en utilisant les techniques développées dans [11], Bahlali et al, [5] ont démontré ce résultat d'approximation sous des conditions plus faibles que celles considérées dans [11]. Particulièrement, les auteurs de [5] ont établi le résultat avec b et σ qui sont continus bornés au lieu de Lipschitziens bornés supposés dans [11]. Dans ce nouveau contexte, Bahlali et al ont utilisé la convergence faible de la procédure de pénalisation et l'unicité en loi de l'équation réfléchie (1.10) afin d'établir le résultat de convergence de u^n vers u .

La stratégie développée dans [11] et adoptée par [5] consiste à associer le système (1.12) à une EDSRG de la forme :

$$\begin{aligned} Y_s^{t,x} &= g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}) dr - \int_s^T U_r^{t,x} \sigma(X_r^{t,x}) dW_r \\ &\quad + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}, \end{aligned} \quad (1.14)$$

où

$$k_s^{t,x} := |K^{t,x}|_{[t,s]} = \int_t^s \langle \nabla l(X_r^{t,x}), dK_r^{t,x} \rangle \quad (1.15)$$

est la variation totale de $K^{t,x}$ sur $[t, s]$ et $(X_s^{t,x}; t \leq s \leq T)$ dénote la solution de l'équation (1.10). De plus, ils ont associé les EDP (1.13) aux EDSR suivantes :

$$\begin{aligned} Y_s^{t,x,n} &= g(X_T^{t,x,n}) + \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) dr - \int_s^T U_r^{t,x,n} \sigma(X_r^{t,x,n}) dW_r \\ &\quad + \int_s^T h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n}, \end{aligned} \quad (1.16)$$

avec

$$k_s^{t,x,n} := -n \int_t^s \langle \nabla l(X_r^{t,x,n}), \delta(X_r^{t,x,n}) \rangle dr \quad (1.17)$$

et $(X_r^{t,x,n}; t \leq s \leq T)$ dénote la solution de l'équation pénalisée (1.11).

En tenant compte de la convergence de la méthode de pénalisation pour le processus réfléchi $X^{t,x}$, i.e. la convergence de $X^{t,x,n}$ vers $X^{t,x}$. Les auteurs de [5] ont démontré que $Y^{t,x,n}$ (solution de (1.16)) converge faiblement sur l'espace des fonctions càdlàg muni de la topologie S (voir Appendice A) vers $Y^{t,x}$ solution de l'équation (1.14). En particulier, $u^n(t, x) = Y_t^{t,x,n}$ solution de viscosité de (1.13) va converger vers $u(t, x) = Y_t^{t,x}$ solution de viscosité de (1.12). Notons que cette approximation sort du cadre de stabilité de solutions des EDP, puisque le problème limite, qui est avec condition de Neumann au bord, est obtenu à partir d'une suite de problèmes avec conditions terminales. A cette étape, des questions naturelles se posent :

- Est-ce qu'on peut affaiblir les conditions sur (b, σ) encore une fois ?
- Est-ce qu'on peut autoriser à la non-linéarité f de dépendre du gradient de la solution ?

Notre objectif dans la première partie de la thèse est de donner des réponses affirmatives à ces deux questions.

1.1.2 Pénalisation des problèmes de Neumann non-linéaires : Résultats nouveaux

1.1.2.1 Pénalisation des EDP avec conditions de Neumann non-linéaires et coefficients mesurables

Dans le **chapitre 2**, nous allons démontrer une méthode de pénalisation pour le système (1.12) sous des hypothèses plus faibles que celles considérées dans [5]. Précisément, nous permettons aux coefficients b , σ et $x \mapsto f(., x, .)$ d'être discontinus. Dans ce nouveau contexte, nous perdrons la propriété de l'unicité en loi des équations (1.10) et (1.11), les convergences sont plus difficiles à établir et la notion de solution de viscosité classique n'est plus valable. Pour établir notre résultat principal, nous adoptons la méthode de [11, 5], c'est-à-dire nous utiliserons la propriété de tension de la suite $(Y^{t,x,n})$ par rapport à la topologie S et l'association des systèmes (1.12) et (1.13) aux EDSR (1.14) et (1.16) respectivement. Dans notre cas, l'inégalité de Krylov, qui est bien connue pour les EDS (voir [42, 87, 75]), joue un rôle crucial dans les preuves de nos résultats.

Nous supposons les hypothèses suivantes :

- (A.1) $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$ sont mesurables et bornés,
- (A.2) il existe $\alpha > 0$ tel que pour tout $x \in \mathbb{R}^d$ $\sigma \sigma^*(x) \geq \alpha I$,
- (A.3) l'équation (1.10) satisfait la propriété d'unicité faible,
- (A.4) l'équation (1.11) satisfait la propriété d'unicité faible,
- (A.5) ils existent des constantes positives C_1, C_2, l_h et $\mu_f \in \mathbb{R}$, $\beta < 0$ et $q \geq 1$ telles que $\forall t, s \in [0, T], \forall (x, x', y, y') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^k)^2$ on a

- (i) $\langle y' - y, f(t, x, y') - f(t, x, y) \rangle \leq \mu_f |y' - y|^2$,
- (ii) $|h(t, x', y') - h(s, x, y)| \leq l_h (|t - s| + |x' - x| + |y' - y|)$,
- (iii) $\langle y' - y, h(t, x, y') - h(t, x, y) \rangle \leq \beta |y' - y|^2$,
- (iv) $|f(t, x, y)| + |h(t, x, y)| \leq C_1 (1 + |y|)$,
- (v) $|g(x)| \leq C_2(1 + |x|^q)$.

g est continue et f est mesurable par rapport à x et continue par rapport à (t, y) .

Avant de présenter la convergence de la pénalisation de l'équation (1.10), rappelons d'abord que $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s\}, W, X, K)$ est une solution faible de (1.10) si $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s\})$ est une base stochastique, W est un mouvement Brownien défini sur cette base, X est un processus continu adapté et K est un processus continu à variation bornée tels que $X_s \in \bar{D}$ \mathbb{P} -p.s., $\forall s \in [t, T]$ et (X, K) satisfait le système (1.10).

D'autre part, sous les hypothèses (A.1)-(A.2), l'équation (1.11) admet une solution faible $X^{t,x,n}$ (voir [42, 74]) qui converge faiblement vers une solution de l'équation (1.10). En particulier, nous avons :

Proposition 1.1. Sous les hypothèses (A.1)-(A.3), l'équation (1.10) a une solution faible unique $(X^{t,x}, K^{t,x})$. De plus,

$$(X^{t,x,n}, K^{t,x,n}) \xrightarrow[U \times U]^* (X^{t,x}, K^{t,x}),$$

où $K_s^{t,x,n} := -n \int_t^s \delta(X_r^{t,x,n}) dr$ et $\xrightarrow[U \times U]^*$ dénote la convergence faible des processus par rapport à la topologie uniforme.

La preuve de cette convergence est basée sur la relative compacité de la suite $(X^{t,x,n}, K^{t,x,n})$ sur l'espace $\mathcal{C}([0, T], \mathbb{R}^d \times \mathbb{R}^d)$ muni de la topologie de convergence uniforme et sur l'unicité en loi de l'équation (1.10). Nous nous référons par exemple aux [75, 81].

Il est connu que lorsque la dérive et le coefficient de diffusion présentent des points de discontinuité, la propriété de l'unicité en loi généralement n'est pas vérifiée. Ce fait justifie les hypothèses (A.3) et (A.4). Présentons des cas où ces deux conditions sont vérifiées. Si l'une des hypothèses suivantes est satisfaite :

- D est un semi-compacte, la dimension $d \leq 2$ et les hypothèses (A.1)-(A.2) sont satisfaites.
- b est mesurable borné, σ est continu borné et $\sigma\sigma^*$ est uniformément elliptique.

Alors, l'équation (1.10) vérifie la propriété d'unicité faible. Voir [43, 82].

D'autre part, sous les hypothèses (A.1)-(A.2), si la dimension $d \leq 2$ alors la propriété (A.4) est vérifiée, c'est-à-dire que l'équation (1.11) admet une solution faible unique. voir [43, 40].

Passons au premier résultat principal, à savoir : la convergence faible de la solution de l'équation (1.16) vers celle de (1.14).

Proposition 1.2. Supposons que les hypothèses (A.1)-(A.5) sont satisfaites, alors il existe $(Y_s^{t,x,n}, U_s^{t,x,n})_{s \in [t, T]}$ et $(Y_s^{t,x}, U_s^{t,x})_{s \in [t, T]}$ des solutions uniques des équations (1.16) et (1.14) respectivement. De plus

$$(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n}) \xrightarrow[S \times S \times S]^* (Y^{t,x}, M^{t,x}, H^{t,x})$$

et

$$\lim_{n \rightarrow \infty} Y_t^{t,x,n} = Y_t^{t,x}$$

avec

$$M_s^{t,x,n} := \int_t^s U_r^{t,x,n} \sigma(X_r^{t,x,n}) dW_r, \quad H_s^{t,x,n} := \int_t^s h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n},$$

$$M_s^{t,x} := \int_t^s U_r^{t,x} \sigma(X_r^{t,x}) dW_r \quad \text{et} \quad H_s^{t,x} := \int_t^s h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}.$$

Ici $\xrightarrow[S]^*$ est la convergence faible par rapport à la topologie de Jakubowski, voir annexe A et [35].

Expliquons la stratégie de la preuve. L'idée consiste à démontrer que la suite $(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n})$ est tendue sur $(\mathcal{D}([0, T], \mathbb{R}^k))^3$ muni de la topologie de Jakubowski, par conséquent, le long d'une sous suite $(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n})$ va converger vers $(\bar{Y}, \bar{M}, \bar{H})$ un élément de $(\mathcal{D}([0, T], \mathbb{R}^k))^3$. Ensuite, nous passons à la limite dans l'EDSR (1.16). La dernière étape c'est l'identification de la limite c'est-à-dire $\bar{Y} = Y^{t,x}$ et $\bar{M} = M^{t,x}$.

Parmi les problèmes auxquels nous nous sommes confrontés est de définir une notion de solution pour le système (1.12). Comme les coefficients peuvent être discontinus, la notion classique de viscosité n'est pas définie dans notre contexte. Nous étendrons la notion de L^p -viscosité définie dans [15, 17] pour les EDP avec coefficients mesurables, au cas des EDP avec condition limite de Neumann. Introduisons la définition de la solution de L^p -viscosité adoptée pour le problème (1.12).

Définition 1.1. Soit p un entier, tel que $p > d + 2$.

- (i) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^k)$ est appelée sous-solution de L^p -viscosité du système (1.12) si $u_i(T, x) \leq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq k$, et si de plus pour tout $1 \leq i \leq k$ et toute fonction $\varphi \in W_{p,loc}^{1,2}([0, T] \times \bar{D})$, si $(\hat{t}, \hat{x}) \in [0, T] \times \bar{D}$ est un maximum local de $u_i - \varphi$, alors :

$$\text{ess} \liminf_{(t,x) \rightarrow (\hat{t}, \hat{x})} \left\{ -\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)) \right\} \leq 0, \quad \text{si } \hat{x} \in D,$$

$$\text{ess} \liminf_{(t,x) \rightarrow (\hat{t}, \hat{x})} \min \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)), \right. \\ \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \leq 0, \quad \text{si } \hat{x} \in \partial D.$$

(ii.) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^k)$ est appelée sur-solution de L^p -viscosité de (1.12) si $u_i(T, x) \geq g_i(x)$ pour tout $x \in \bar{D}$ et tout $1 \leq i \leq k$, et si de plus pour tout $1 \leq i \leq k$ et toute fonction $\varphi \in W_{p,loc}^{1,2}([0, T] \times \bar{D})$, si $(\hat{t}, \hat{x}) \in [0, T] \times \bar{D}$ est un minimum local de $u_i - \varphi$, alors :

$$\operatorname{ess\,lim\,sup}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \left\{ -\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)) \right\} \geq 0, \quad \text{si } \hat{x} \in D,$$

$$\operatorname{ess\,lim\,sup}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \max \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)), \right. \\ \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \geq 0, \quad \text{si } \hat{x} \in \partial D.$$

(iii.) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^k)$ est appelée solution de L^p -viscosité de (1.12) si c'est à la fois une sous et une sur-solution de L^p -viscosité.

Dans la suite, nous montrons l'existence d'une solution de L^p -viscosité pour le système (1.12), ensuite nous prouvons que cette solution peut être construite comme étant la limite d'une suite de solutions de L^p -viscosité de (1.13). Il faut noter que même les coefficients des EDP ne sont pas continus, les solutions de L^p -viscosité doivent être continues. Pour assurer cette régularité, nous montrons les résultats suivants :

Proposition 1.3. On suppose (A.1)-(A.2) et (A.4). Soit $X^{t,x,n}$ l'unique solution de l'équation (1.11), alors l'application $[0, T] \times \mathbb{R}^d \ni (t, x) \rightarrow X^{t,x,n}$ est continue en loi.

Proposition 1.4. Supposons que les hypothèses (A.1)-(A.3) sont satisfaites. Soit $(X^{t,x}, K^{t,x})$ la solution unique de l'équation (1.10), alors l'application $[0, T] \times \bar{D} \ni (t, x) \rightarrow (X^{t,x}, K^{t,x})$ est continue en loi.

On définit les quantités : $u^n(t, x) := Y_t^{t,x,n}$ où $Y^{t,x,n}$ est la solution de l'EDSR (1.16) et $u(t, x) := Y_t^{t,x}$ avec $Y^{t,x}$ est la solution de (1.14).

La proposition suivante nous donne la continuité nécessaire pour définir nos solutions de L^p -viscosité.

Proposition 1.5. Sous les hypothèses (A.1)-(A.5), les fonctions $u^n : (t, x) \in [0, T] \times \mathbb{R}^d \mapsto u^n(t, x)$ et $u : (t, x) \in [0, T] \times \bar{D} \mapsto u(t, x)$ sont continues.

Le deuxième résultat principal est l'application suivante aux EDP avec conditions de Neumann non-linéaires.

Théorème 1.1. *Supposons que les hypothèses (A.1)-(A.5) sont satisfaites. Pour tout $p > d + 2$, les fonctions $u^n(t, x) = Y_t^{t,x,n}$ et $u(t, x) = Y_t^{t,x}$ sont des solutions de L^p -viscosité des systèmes (1.13) et (1.12) respectivement. En outre,*

$$\lim_{n \rightarrow \infty} u^n(t, x) = u(t, x), \quad \forall (t, x) \in [0, T] \times \bar{D}.$$

1.1.2.2 Pénalisation des EDP dégénérées avec conditions de Neumann non-linéaires

Dans le **chapitre 3**, nous visons à généraliser les résultats précédents au cas où la non-linéarité f dépendrait également du gradient de la solution. De même, le générateur de l'EDSR correspondante dépendra également du processus de contrôle Z . Les techniques adoptées ci-dessus et dans [11, 5] reposent sur des propriétés de tension par rapport à la topologie S , le principal inconvénient de cette méthode est qu'elle ne permet pas de considérer une non-linéarité f dépendante de ∇u . Ici, la méthode utilisée est directe et ne nécessite aucune propriété de relative compacité, et la convergence de l'EDSR pénalisée est obtenue seulement grâce à la convergence de l'EDS pénalisée (1.11). Lorsque f dépend du gradient de la solution, des techniques des EDP sont généralement utilisées pour contrôler ∇u afin d'obtenir la convergence de l'EDSR associée. Tel n'est pas le cas ici. Le travail est entièrement basé sur les EDSR. Une autre contribution dans ce chapitre est la relaxation des conditions sur les coefficients b, σ et h imposées dans [11]. Précisément, nous supprimons les hypothèses que b, σ soient bornés, σ soit inversible et la condition de Lipschitz sur h . Nous considérons la version suivante du système (1.12)

$$\begin{cases} \frac{\partial u_i}{\partial t}(t, x) + \mathcal{L}u_i(t, x) + f_i(t, x, u(t, x), (\nabla u \sigma)(t, x)) = 0 \\ , 1 \leq i \leq k, (t, x) \in [0, T] \times D, \\ u(T, x) = g(x), x \in D, \\ \frac{\partial u}{\partial n}(t, x) + h(t, x, u(t, x)) = 0, \forall (t, x) \in [0, T] \times \partial D, \end{cases} \quad (1.18)$$

pour lequel, nous construisons une solution de viscosité comme limite d'une suite de solutions de viscosité d'un système de la forme ($1 \leq i \leq k, 0 \leq t \leq T, x \in \mathbb{R}^d, n \in \mathbb{N}^*$).

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}u_i^n(t, x) + f_i(t, x, u^n(t, x), (\nabla u^n \sigma)(t, x)) \\ - n \langle \delta(x), \nabla u_i^n(t, x) + \nabla l(x) \rangle h_i(t, x, u^n(t, x)) = 0; \\ u^n(T, x) = g(x). \end{cases} \quad (1.19)$$

À cette fin, nous introduisons les EDRS généralisées suivantes, pour tout $t \in [0, T]$ et $s \in [t, T]$

$$\begin{aligned} Y_s^{t,x,n} &= g(X_T^{t,x,n}) + \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) dr \\ &\quad + \int_s^T h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} - \int_s^T Z_r^{t,x,n} dW_r \end{aligned} \quad (1.20)$$

et

$$\begin{aligned} Y_s^{t,x} &= g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x} \\ &\quad - \int_s^T Z_r^{t,x} dW_r, \end{aligned} \quad (1.21)$$

où $X^{t,x,n}$ est la solution de l'équation (1.11), $k^{t,x,n}$ est donné par (1.17), $X^{t,x}$ est la solution de l'équation (1.10) et $k^{t,x}$ est défini par (1.15). Nous avons besoin des hypothèses suivantes :

(A.1) : Il existe des constantes positives C et μ telles que pour tout $(x, y) \in \mathbb{R}^d$:

- (i) $|b(x)| \vee \|\sigma(x)\| \leq C(1 + |x|)$,
- (ii) $|b(x) - b(y)| \vee \|\sigma(x) - \sigma(y)\| \leq \mu|x - y|$.

(A.2) : Il existe des constantes positives C , l_f et $\beta < 0$, $\mu_f \in \mathbb{R}$ telles que $\forall t \in [0, T]$, $\forall (x, x', y, y', z, z') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^k)^2 \times (\mathbb{R}^{k \times d'})^2$ nous avons :

- (i) $\langle y - y', f(t, x, y, z) - f(t, x, y', z) \rangle \leq \mu_f |y - y'|^2$,
- (ii) $|f(t, x, y, z) - f(t, x, y, z')| \leq l_f \|z - z'\|$,
- (iii) $|f(t, x, y, 0)| \leq C(1 + |y|)$,
- (iv) $\langle y - y', h(t, x, y) - h(t, x, y') \rangle \leq \beta |y - y'|^2$,
- (v) $|h(t, x, y)| \leq C(1 + |y|)$,
- (vi) $|g(x)| \leq C(1 + |x|)$.

L'élément clé pour prouver notre résultat principal est la convergence de la procédure de pénalisation pour la diffusion réfléchie (1.10). Plusieurs auteurs ont étudié le problème de l'existence de la solution de la diffusion réfléchie et de son approximation par des solutions d'équations avec des termes de pénalisation (1.11). Nous faisons référence, par exemple, aux [82, 84, 49, 45, 81]. Nous rappelons les résultats de [84, 81].

Théorème 1.2. *Sous l'hypothèse (A.1), nous avons*

(i) *l'équation (1.10) admet une solution forte unique,*

(ii) *pour tout $1 \leq q < \infty$, $0 < T < \infty$, la quantité $\mathbb{E} \left[\sup_{t \leq s \leq T} |X_s^{t,x,n} - X_s^{t,x}|^q \right]$*

converge vers 0 quand $n \rightarrow \infty$, pour $(t, x) \in [0, T] \times \bar{D}$.

Sous l'hypothèse (A.2), les EDSR (1.21) et (1.20) admettent des solutions uniques notées $(Y^{t,x}, Z^{t,x})$ et $(Y^{t,x,n}, Z^{t,x,n})$ respectivement. Voir [66, 71].

Nous rappelons également une inégalité rétrograde due à Pardoux et Rascanu [69, 70]. Cette inégalité constitue la clé de la preuve de notre résultat principal.

Lemme 1.1. *Soit $(Y, Z) \in \mathcal{S}_m^0 \times \mathcal{M}_{m \times d}^0$, la solution de l'équation suivante*

$$Y_t = Y_T + \int_t^T d\mathcal{K}_r - \int_t^T Z_r dW_r, \quad 0 \leq t \leq T, \quad \mathbb{P}\text{-p.s.}$$

où $\mathcal{K} \in \mathcal{S}_m^0$ et $\mathcal{K}(\omega) \in BV([0, T]; \mathbb{R}^m)$. Étant donné

- *un processus stochastique croissant L avec $L_0 = 0$,*
- *un processus stochastique $R(\omega) \in BV([0, T], \mathbb{R})$, avec $R_0 = 0$,*

- un processus stochastique continu $V_r(\omega) \in BV([0, T], \mathbb{R})$, tel que $V_0 = 0$ et

$$\mathbb{E} \left(\sup_{s \in [0, T]} \int_s^T e^{2V_r} dR_r \right) < +\infty.$$

On suppose

- (i) $\langle Y_r, d\mathcal{K}_r \rangle \leq \frac{\alpha}{2} \|Z_r\|^2 dr + |Y_r|^2 dV_r + |Y_r| dL_r + dR_r$ au sens des mesures sur $[0, T]$,
- (ii) $\mathbb{E} \sup_{r \in [0, T]} e^{2V_r} |Y_r|^2 < +\infty$.

Si $\alpha < 1$, alors il existe des constantes positives C_1, C_2 et C_3 dépendant uniquement de α telles que

$$\begin{aligned} & \mathbb{E} \left(\sup_{r \in [0, T]} |e^{V_r} Y_r|^2 \right) + \mathbb{E} \left(\int_0^T e^{2V_r} \|Z_r\|^2 dr \right) \\ & \leq C_1 \mathbb{E} |e^{V_T} Y_T|^2 + C_2 \mathbb{E} \left(\int_0^T e^{V_r} dL_r \right)^2 + C_3 \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2V_r} dR_r. \end{aligned}$$

Notre premier résultat principal est la convergence de la solution de l'équation (1.20), $(Y^{t,x,n}, Z^{t,x,n})$, vers celle de l'équation (1.21), $(Y^{t,x}, Z^{t,x})$, pour la norme classique des EDSR. Le théorème suivant généralise [11, théorème 3.1].

Théorème 1.3. *Supposons les conditions (A.1)-(A.2), alors*

$$\mathbb{E} \left(\sup_{r \in [0, T]} |Y_r^{t,x,n} - Y_r^{t,x}|^2 + \int_0^T \|Z_r^{t,x,n} - Z_r^{t,x}\|^2 dr \right) \rightarrow 0, \quad \text{quand } n \rightarrow +\infty.$$

L'idée de la preuve du théorème précédent est basée sur l'application du lemme 1.1 à l'EDSR suivante :

$$Y_s^{t,x,n} - Y_s^{t,x} = g(X_T^{t,x,n}) - g(X_T^{t,x}) + \int_s^T d\mathcal{K}_r^{t,x,n} - \int_s^T (Z_r^{t,x,n} - Z_r^{t,x}) dW_r$$

où

$$\begin{aligned} d\mathcal{K}_r^n & := [f(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) - f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})] dr \\ & \quad + h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} - h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}. \end{aligned}$$

Ainsi, par le lemme 1.1, il existe des constantes positives C_1, C_2 et C_3 telles que

$$\begin{aligned} & \mathbb{E} \left(\sup_{r \in [0, T]} |Y_r^{t,x,n} - Y_r^{t,x}|^2 \right) + \mathbb{E} \left(\int_0^T \|Z_r^{t,x,n} - Z_r^{t,x}\|^2 dr \right) \\ & \leq C_1 \mathbb{E} e^{2\lambda T} |g(X_T^{t,x,n}) - g(X_T^{t,x})|^2 + C_2 \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 \\ & \quad + C_3 \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \end{aligned} \tag{1.22}$$

où L^n et R^n sont définis par

$$\begin{aligned} dL_r^n &:= |f(r, X_r^{t,x,n}, Y_r^{t,x}, Z_r^{t,x}) - f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})|dr \\ &\quad + |h(r, X_r^{t,x,n}, Y_r^{t,x}) - h(r, X_r^{t,x}, Y_r^{t,x})|dk_r^{t,x,n}, \\ dR_r^n &:= \langle Y_r^{t,x,n} - Y_r^{t,x}, h(r, X_r^{t,x}, Y_r^{t,x}) \rangle (dk_r^{t,x,n} - dk_r^{t,x}). \end{aligned}$$

Ensuite, grâce à la convergence de $X^{t,x,n}$ vers $X^{t,x}$, nous démontrons que le terme de droite de l'inégalité (1.22) converge vers zéro quand n tend vers $+\infty$.

Après avoir établi la convergence de l'EDSR pénalisée, nous sommes maintenant prêts à présenter le résultat de la pénalisation pour le système (1.18). Comme on considère les solutions de viscosité, on introduit la condition

(A.3) f_i , la i -ème coordonnée de f , ne dépend que de la i -ème ligne de la matrice z .

Il convient de rappeler la définition de la solution de viscosité du système (1.18).

Définition 1.2. (i) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^m)$ est appelée une sous-solution de viscosité du système (1.18) si $u_i(T, x) \leq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq m$, et de plus pour tout $1 \leq i \leq m$, $\varphi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$, et $(t, x) \in (0, T] \times \bar{D}$ pour lequel $u_i - \varphi$ a un maximum local, on a

$$-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)) \leq 0, \quad \text{si } x \in D,$$

$$\begin{aligned} \min \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)), \right. \\ \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \leq 0, \quad \text{si } x \in \partial D. \end{aligned}$$

(ii) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^d)$ est appelée une sur-solution de viscosité de (1.18) si $u_i(T, x) \geq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq m$, et de plus pour tout $1 \leq i \leq m$, $\varphi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$, et $(t, x) \in (0, T] \times \bar{D}$ pour lequel $u_i - \varphi$ a un minimum local, on a

$$-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)) \geq 0, \quad \text{si } x \in D,$$

$$\begin{aligned} \max \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)), \right. \\ \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \geq 0, \quad \text{si } x \in \partial D. \end{aligned}$$

(iii) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^m)$ est appelée une solution de viscosité du système (1.18) si c'est à la fois une sous et sur-solution de viscosité.

Notre deuxième contribution principale est l'approximation donnée par le théorème suivante.

Théorème 1.4. *Supposons que les hypothèses (A.1), (A.2) et (A.3) sont vérifiées. Alors il existe une suite de fonctions continues $u^n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ et une fonction $u : [0, T] \times \bar{D} \rightarrow \mathbb{R}^m$ telles que : u^n est une solution de viscosité du système (1.19), u est une solution de viscosité du système (1.18) et de plus pour chaque $(t, x) \in [0, T] \times \bar{D}$ on a*

$$\lim_{n \rightarrow +\infty} u^n(t, x) = u(t, x).$$

L'idée principale de la preuve de ce résultat est l'association du système (1.19) avec l'EDSR (1.20) et l'EDP (1.18) avec (1.21). On pose,

$$u^n(t, x) := Y_t^{t,x,n} \quad \text{et} \quad u(t, x) := Y_t^{t,x}. \quad (1.23)$$

Il découle du théorème 3.2 de [68] que u^n est une solution de viscosité de l'EDP (1.19). Grâce à [70, 71], u est une solution de viscosité de l'EDP (1.21). D'autre part, nous avons pour chaque $(t, x) \in [0, T] \times \bar{D}$

$$|u^n(t, x) - u(t, x)|^2 = |Y_t^{t,x,n} - Y_t^{t,x}|^2 \leq \mathbb{E} \sup_{s \in [0, T]} |Y_s^{t,x,n} - Y_s^{t,x}|^2.$$

Ainsi,

$$\lim_{n \rightarrow +\infty} u^n(t, x) = u(t, x).$$

1.1.3 Un résultat de stabilité des EDP dégénérées avec conditions de Neumann non-linéaires

Le **chapitre 4** sera consacré à établir un résultat de stabilité pour le problème de Neumann non-linéaire (1.18). Précisément, nous considérons (u^n) une suite de solutions de viscosité des EDP avec conditions de Neumann non-linéaires suivantes :

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}^n u_i^n(t, x) + f_i^n(t, x, u^n(t, x), (\nabla u^n \sigma^n)(t, x)) = 0, \\ 1 \leq i \leq k, \quad (t, x) \in [0, T] \times D, \\ u^n(T, x) = g^n(x), \quad x \in D, \\ \frac{\partial u^n}{\partial n}(t, x) + h^n(t, x, u^n(t, x)) = 0, \quad \forall (t, x) \in [0, T] \times \partial D, \end{cases} \quad (1.24)$$

où $f^n : [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{k \times d} \rightarrow \mathbb{R}^k$, $h^n : [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ et $g^n : \mathbb{R}^d \rightarrow \mathbb{R}^k$ sont des fonctions données et

$$\mathcal{L}^n = \frac{1}{2} \sum_{i,j} (\sigma^n(\sigma^n)^*(\cdot))_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i^n(\cdot) \frac{\partial}{\partial x_i}.$$

Nous montrons que lorsque $(b^n, \sigma^n, f^n, g^n, h^n)$ convergent vers (b, σ, f, g, h) , alors u^n converge vers u solution de viscosité du système (1.18).

Pour montrer ce résultat de stabilité, il est nécessaire de disposer d'une propriété de stabilité pour la solution de l'équation (1.10). Pour cela, nous introduisons une suite de processus réfléchis donnée par :

$$\begin{cases} X_s^{t,x,n} = x + \int_t^s b^n(X_r^{t,x,n}) dr + \int_t^s \sigma^n(X_r^{t,x,n}) dW_r + K_s^{t,x,n}, \\ K_s^{t,x,n} = \int_t^s \nabla l(X_r^{t,x,n}) d|K^{t,x,n}|_{[t,r]}, \\ |K^{t,x,n}|_{[t,s]} = \int_t^s 1_{\{X_r^{t,x,n} \in \partial D\}} d|K^{t,x,n}|_{[t,r]}, \quad \forall s \in [t, T], \end{cases} \quad (1.25)$$

avec $b^n : \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma^n : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ sont des suites de fonctions mesurables satisfaisant les hypothèses suivantes :

(A.4) Il existe des constantes positives C, μ_n telles que pour tous $x, y \in \mathbb{R}^d$

- (i) $\sup_{n \geq 0} |b^n(x)| \vee \sup_{n \geq 0} \|\sigma^n(x)\| \leq C(1 + |x|)$,
- (ii) $|b^n(x) - b^n(y)| \vee \|\sigma^n(x) - \sigma^n(y)\| \leq \mu_n |x - y|$.

Afin d'établir notre principal résultat, nous devons montrer la convergence de la solution de (1.25) vers celle de l'équation (1.10), c'est-à-dire la convergence de $(X^{t,x,n}, K^{t,x,n})$ vers $(X^{t,x}, K^{t,x})$. Pour ce faire, nous considérons la condition suivante :

(A.5) Les suites de fonctions b^n et σ^n convergent localement uniformément vers b et σ respectivement.

Le lemme suivant est un résultat de stabilité pour l'équation (1.10).

Lemme 1.2. *Supposons que les hypothèses (A.1.)1.1.2.2, (A.4) et (A.5) soient satisfaites, alors*

(a) pour chaque $q \geq 1$ et $(t, x) \in [0, T] \times \bar{D}$, nous avons

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s^{t,x,n} - X_s^{t,x}|^{2q} + \sup_{0 \leq s \leq T} |K_s^{t,x,n} - K_s^{t,x}|^{2q} \right] = 0, \quad (1.26)$$

(b) pour chaque $q \geq 1$ et chaque $(t, x) \in [0, T] \times \bar{D}$

$$\lim_{n \rightarrow +\infty} \mathbb{E} \sup_{s \in [0, T]} |k_s^{t,x,n} - k_s^{t,x}|^q = 0$$

où

$$k_s^{t,x,n} = \int_t^s \langle \nabla l(X_r^{t,x,n}), dK_r^{t,x,n} \rangle$$

et $k^{t,x}$ est donné par (1.15).

Nous considérons une suite des EDSR liées aux problèmes de Neumann (1.24)

$$\begin{aligned} Y_s^{t,x,n} &= g^n(X_T^{t,x,n}) + \int_s^T f^n(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) dr \\ &\quad + \int_s^T h^n(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} - \int_s^T Z_r^{t,x,n} dW_r. \end{aligned} \quad (1.27)$$

Introduisons l'hypothèse :

(A.6) Les fonctions f^n , h^n et g^n sont continues et il existe des constantes positives C , l_{f^n} et $\beta_n < 0$, $\mu_{f^n} \in \mathbb{R}$ telles que pour chaque $t \in [0, T]$ et chaque $(x, x', y, y', z, z') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^k)^2 \times (\mathbb{R}^{k \times d'})^2$ nous avons

- (i) $\langle y - y', f^n(t, x, y, z) - f^n(t, x, y', z) \rangle \leq \mu_{f^n} |y - y'|^2$,
- (ii) $|f^n(t, x, y, z) - f^n(t, x, y, z')| \leq l_{f^n} \|z - z'\|$,
- (iii) $\sup_{n \geq 0} |f^n(t, x, y, z)| \leq C(1 + |y|)$,
- (iv) $\langle y - y', h^n(t, x, y) - h^n(t, x, y') \rangle \leq \beta_n |y - y'|^2$,
- (v) $\sup_{n \geq 0} |h^n(t, x, y)| \leq C(1 + |y|)$,
- (vi) $\sup_{n \geq 0} |g^n(x)| \leq C(1 + |x|)$.

Pour n fixé, sous (A.6), l'EDSR (1.27) a une solution unique $(Y_s^{t,x,n}, Z_s^{t,x,n})_{s \in [t, T]}$. Le prochain lemme donne des estimations a priori de $(Y_s^{t,x,n}, Z_s^{t,x,n})_{s \in [t, T]}$. Ces estimations seront utilisées dans la preuve de notre principal résultat.

Lemme 1.3. *Supposons que les hypothèses (A.4)(i) et (A.6) sont vérifiées. Soient $\lambda_q := 1 \wedge (q - 1)$, $V_s^n := \mu_{f^n}^+ s + \frac{l_{f^n}^2}{\lambda_q} s$ et supposons que pour tout $q > 1$ $\sup_{n \geq 0} \exp(qV_T^n) < +\infty$. Alors,*

$$\sup_{n \geq 0} \mathbb{E} \sup_{s \in [0, T]} |e^{V_s^n} Y_s^{t,x,n}|^q + \sup_{n \geq 0} \mathbb{E} \left(\int_0^T e^{2V_s^n} \|Z_s^{t,x,n}\|^2 ds \right)^{\frac{q}{2}} < +\infty.$$

Pour prouver la convergence de la solution du problème (1.24), on considère la condition

(A.7) : Pour tout compact (κ, κ') de $\mathbb{R}^d \times \mathbb{R}^k$ on a

$$\lim_{n \rightarrow +\infty} \sup_{r \in [0, T]} \sup_{x \in \kappa, y \in \kappa', z \in \mathbb{R}^{k \times d'}} |f^n(r, x, y, z) - f(r, x, y, z)| = 0$$

et

$$\lim_{n \rightarrow +\infty} \left[\sup_{x \in \kappa} |g^n(x) - g(x)| + \sup_{r \in [0, T]} \sup_{x \in \kappa, y \in \kappa'} |h^n(r, x, y) - h(r, x, y)| \right] = 0.$$

Pour donner un sens à la solution de viscosité des EDP (1.24), nous devons supposer

(A.8) : f_i^n , la i ème coordonnée de f^n , ne dépend que de la i ème ligne de la matrice z .

Le principal résultat de ce chapitre est le théorème suivant.

Théorème 1.5. *Supposons que les hypothèses (A.1).1.1.2.2, (A.2).1.1.2.2, (A.3).1.1.2.2, (A.4), (A.5), (A.6), (A.7) et (A.8) sont satisfaites. Supposons en outre que $\sup_{n \geq 0} \exp(qV_T^n) < +\infty$ pour tout $q > 1$. Alors, il existe des fonctions continues u^n , $u : [0, T] \times \bar{D} \rightarrow \mathbb{R}^m$ telles que*

- u^n est une solution de viscosité de (1.24),
- u est une solution de viscosité du problème de Neumann (1.18),
- $u^n(t, x)$ converge vers $u(t, x)$ pour tout $(t, x) \in [0, T] \times \bar{D}$.

L'idée de la preuve est la suivante : on considère $(Y^{t,x}, Z^{t,x})$ et $(Y^{t,x,n}, Z^{t,x,n})$ respectivement les solutions uniques des équations (1.14) et (1.27) puis on définit $u(t, x) := Y_t^{t,x}$ et $u^n(t, x) := Y_t^{t,x,n}$. Ainsi, [71, Théorème 4.3] garantit que u et u^n sont des solutions de viscosité des systèmes (1.18) et (1.24) respectivement. Pour prouver la convergence de u^n vers u , on montre que la suite de processus $Y^{t,x,n}$ converge vers $Y^{t,x}$. À cet égard, on applique le lemme 1.1 à l'EDSR

$$Y_s^{t,x,n} - Y_s^{t,x} = g^n(X_T^{t,x,n}) - g(X_T^{t,x}) + \int_s^T d\mathcal{K}_r^{t,x,n} - \int_s^T (Z_r^{t,x,n} - Z_r^{t,x}) dW_r$$

avec

$$d\mathcal{K}_r^{t,x,n} := [f^n(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) - f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})] dr + h^n(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} - h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}.$$

Il existe donc des constantes positives C_1, C_2 et C_3 telles que

$$\begin{aligned} & \mathbb{E} \left(\sup_{r \in [0, T]} |Y_r^{t,x,n} - Y_r^{t,x}|^2 \right) + \mathbb{E} \left(\int_0^T \|Z_r^{t,x,n} - Z_r^{t,x}\|^2 dr \right) \\ & \leq C_1 \mathbb{E} e^{2\lambda T} |g^n(X_T^{t,x,n}) - g(X_T^{t,x})|^2 + C_2 \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 \\ & \quad + C_3 \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \end{aligned} \quad (1.28)$$

où L^n et R^n sont donnés par

$$\begin{aligned} dL_r^n & := |f^n(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) - f(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n})| dr \\ & \quad + |f(r, X_r^{t,x,n}, Y_r^{t,x}, Z_r^{t,x}) - f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})| dr \\ & \quad + |h^n(r, X_r^{t,x,n}, Y_r^{t,x,n}) - h(r, X_r^{t,x,n}, Y_r^{t,x,n})| dk_r^{t,x,n} \\ & \quad + |h(r, X_r^{t,x,n}, Y_r^{t,x}) - h(r, X_r^{t,x}, Y_r^{t,x})| dk_r^{t,x,n} \end{aligned}$$

et

$$dR_r^n := \langle Y_r^{t,x,n} - Y_r^{t,x}, h(r, X_r^{t,x}, Y_r^{t,x}) \rangle (dk_r^{t,x,n} - dk_r^{t,x}).$$

En utilisant (1.26), nous prouvons que le côté droit de (1.28) converge vers zéro quand n tend vers l'infini.

1.2 Inégalités de coût de transport

Le problème de transport optimal proposé pour la première fois par le géomètre français G. Monge en 1781 [59] consiste à chercher la manière la plus économique pour transporter des tas de sable dans des trous. En terme de langage mathématique, le transport optimal de Monge peut être formulé comme suit

$$\inf_{T \in \mathcal{T}(\mu, \nu)} \int_{\mathcal{X}} c(x, T(x)) d\mu(x)$$

où μ et ν sont deux probabilités sur des espace polonais respectivement \mathcal{X} et \mathcal{Y} , $\mathcal{T}(\mu, \nu)$ est l'ensemble des applications mesurables $T : \mathcal{X} \rightarrow \mathcal{Y}$ vérifiant $\mu(T^{-1}(B)) = \nu(B)$ pour chaque Borélienne $B \subset \mathcal{Y}$ et $c : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}^+$ mesurable. En particulier, $c(x, y)$ représente la quantité de travail nécessaire pour déplacer une unité de masse de la position $x \in \mathcal{X}$ à $y \in \mathcal{Y}$.

En 1942, L. Kantorovich [38] a introduit une version détendue du problème de transport optimal de Monge. L'approche de Kantorovich consiste à rechercher une probabilité sur l'espace produit $\mathcal{X} \times \mathcal{Y}$ au lieu d'une application préservant la mesure comme dans le problème de Monge. Le problème de Kantorovich s'écrit

$$\inf_{\gamma \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\gamma(x, y),$$

où $\Pi(\mu, \nu)$ représente l'ensemble des probabilités sur l'espace $\mathcal{X} \times \mathcal{Y}$ avec les marginaux μ et ν . Pour plus de détails, nous renvoyons aux livres [89, 88].

Ce sujet a connu un formidable développement à partir des années 80, vu les nombreuses applications dans les équations aux dérivées partielles, la géométrie différentielle, la théorie des probabilités et d'autres domaines. Parmi les différentes généralisations du problème de transport optimal, nous nous intéressons à ce qu'on appelle les inégalités de coût de transport. Ces inégalités mettent en relation un coût $\mathcal{C}(\mu, \nu)$ de transport d'une mesure de probabilité générique $\mu \in P(\mathcal{X})$ sur une mesure de probabilité de référence $\nu \in P(\mathcal{X})$ avec une autre fonctionnelle $J(\mu, \nu)$, où $P(\mathcal{X})$ est l'espace de toutes les mesures de probabilité sur \mathcal{X} . Cette inégalité de transport s'écrit

$$\alpha(\mathcal{C}(\mu, \nu)) \leq J(\mu, \nu), \quad \text{pour tout } \mu \in P(\mathcal{X})$$

où $\alpha : [0, +\infty) \rightarrow [0, +\infty)$ est une fonction croissante avec $\alpha(0) = 0$. Les inégalités de transport classiques sont construites avec $\mathcal{C} = \mathcal{W}_p^d$, où \mathcal{W}_p^d représente la distance de Wasserstein d'ordre p , définie avec une certaine métrique d sur \mathcal{X} et $J(\cdot/\mu) = H(\cdot/\mu)$ est l'entropie relative donnée par

$$H(\nu/\mu) = \begin{cases} \int_{\mathcal{X}} \ln \frac{d\nu}{d\mu} d\mu, & \text{si } \nu \ll \mu \\ +\infty & \text{sinon.} \end{cases}$$

Une autre inégalité de transport classique est définie par l'information de Fisher $I(\cdot/\mu)$ au lieu de l'entropie relative. Voir [60].

Les recherches sur les inégalités de transport commencent en 1996 avec la parution de plusieurs articles sur ce sujet. La première inégalité de transport est

apparue sous le nom de l'inégalité Csiszar-Kullback-Pinsker, voir [23, 24, 51, 83]. Ces inégalités ont reçu plus d'attention avec Marton dans [52]. Dans cet article, un lien entre l'inégalité de transport et les inégalités de concentration a été établi. Les inégalités de transport ont été généralisées dans différentes directions, et de nouveaux liens avec les inégalités fonctionnelles classiques ont été explorés.

Parmi cette famille générale d'inégalités de transport, nous nous intéressons aux inégalités suivantes

$$\mathcal{W}_p^d(\nu, \mu) \leq \sqrt{2CH(\nu/\mu)}, \quad \forall \mu \in P(\mathcal{X}).$$

Dans une telle situation, on dit que la mesure ν satisfait une inégalité de coût de transport et on note $\nu \in T_p(C)$ (dite aussi inégalité de Talagrand). Les cas $p = 1$ et $p = 2$ sont particulièrement intéressants. En fait, les $T_1(C)$ sont liés aux concentrations des mesures telles qu'il a été formulé dans [51, 52]. Les $T_2(C)$ ont été établies initialement par Talagrand [83], pour la mesure gaussienne sur \mathbb{R}^d . Ces inégalités attirent beaucoup d'attention grâce à leurs liens avec d'autres propriétés telles que l'inégalité de Sobolev, l'inégalité de Poincaré et les équations de Hamilton-Jacobi [61, 91].

Les inégalités de transport et leur application aux équations différentielles stochastiques ont été largement étudiées. En outre, de nombreux arguments ont été développés pour établir ces inégalités de transport. La méthode la plus utilisée est la transformation de Girsanov introduite dans [83] et appliquée efficacement par nombreux auteurs. Par exemple, [92] pour les systèmes dynamiques en dimension infinie, [32, 31] pour les EDS singulières, [78] pour les EDS dirigées par un mouvement Brownien fractionnaire et [48] pour les équations d'évolution stochastique gouvernées par un mouvement Brownien fractionnaire. Récemment, [73] a étudié l'inégalité de transport pour la loi des ESD dirigées par des processus gaussiens en utilisant la théorie des trajectoires rugueuses de Lyons. Les auteurs de [12] ont étudié l'inégalité $T_2(C)$ pour l'équation de la chaleur stochastique avec un bruit blanc et un bruit fractionnaire.

Notre objectif dans cette partie est d'une part d'établir $T_2(C)$ pour les EDS sous des hypothèses plus faibles que celles généralement utilisées pour ce type d'inégalités (conditions de dissipativité ou continuité de Lipschitz). D'autre part, nous étendrons l'inégalité de Talagrand au cadre des EDSR.

1.2.1 Inégalité de transport pour les EDS avec dérivées mesurables

L'objectif du **chapitre 5** est d'étendre le résultat de [25] aux EDS avec coefficients non réguliers. La première tentative pour établir $T_2(C)$ en relâchant les conditions était avec Bartl et Tangpi dans [8], où ils ont considéré une EDS en dimension un avec une dérive dérivable par rapport au temps et mesurable par rapport à l'espace et un coefficient de diffusion Lipschitzien. Ici, nous nous intéressons au cas multidimensionnel et nous supposons des hypothèses plus faibles. Il faut souligner que les équations multidimensionnelles sont fondamentales. Par exemple, pour les applications aux résultats de la concentration et

asymptotiques pour les interactions des systèmes de particules, voir par exemple [22, Section 5]. Nous considérons une EDS de la forme :

$$X(t) = x + \int_0^t b(s, X(s)) ds + \int_0^t \sigma(s, X(s)) dW(s), \quad x \in \mathbb{R}^d, \quad (1.29)$$

où $x \in \mathbb{R}^d$, $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$. Pour $p \geq 1$, on dénote par $L_{\text{loc}}^p([0, T]) := L_{\text{loc}}^p([0, T] \times \mathbb{R}^d)$ l'espace de Lebesgue des classes des fonctions localement intégrable et pour tous $m_1, m_2 \in \mathbb{N}$, soit $W_{p, \text{loc}}^{m_1, m_2}([0, T]) := W_{p, \text{loc}}^{m_1, m_2}([0, T] \times \mathbb{R}^d)$ l'espace de Sobolev usuel des fonctions faiblement différentiables $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ telles que

$$\|f\|_{L_{\text{loc}}^p} + \sum_{|\alpha| \leq m_1} \|\partial_t^\alpha f\|_{L_{\text{loc}}^p} + \sum_{|\alpha| \leq m_2} \|\partial_x^\alpha f\|_{L_{\text{loc}}^p} < +\infty.$$

Soit $L_q^p([0, T]) := L^q([0, T], L^p(\mathbb{R}^d))$ l'espace des fonctions mesurables $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ telles que

$$\|f\|_{L_q^p} := \left(\int_0^T \left(\int_{\mathbb{R}^d} |f(s, x)|^p dx \right)^{q/p} ds \right)^{1/q} < \infty.$$

On suppose que le coefficient de diffusion σ satisfait la condition d'ellipticité uniforme

$$\xi^* \sigma(t, x) \xi \geq \Lambda_\sigma |\xi|^2 \quad \text{pour tout } (t, x, \xi) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \quad \text{pour un } \Lambda_\sigma > 0. \quad (1.30)$$

Les résultats principaux de ce chapitre sont les deux théorèmes suivants :

Théorème 1.6. *Soient $\sigma \in W_{2(d+1), \text{loc}}^{0,1}([0, T]) \cap L^\infty([0, T] \times \mathbb{R}^d)$ et $b \in W_{(d+1), \text{loc}}^{0,1}([0, T]) \cap L^\infty([0, T] \times \mathbb{R}^d)$. Supposons que σ vérifie (1.30). Alors, l'équation (1.29) admet une solution unique X , et*

$$\text{la loi } \mu_x \text{ de } X \text{ satisfait } T_2(C)$$

avec constante

$$C := \inf_{0 < \varepsilon < 1} 2 \|\sigma\|_\infty^2 \exp \left(7T \frac{32 + \varepsilon}{\varepsilon(1 - \varepsilon)} \right) \frac{1}{1 - \varepsilon}.$$

Ce résultat donne l'inégalité de transport pour les EDS avec des coefficients dans certains espaces de Sobolev. Il constitue en outre un argument essentiel dans la preuve du résultat suivant, où nous le combinons avec certaines estimations de gradient pour les solutions des EDP pour démontrer l'inégalité de transport pour les EDS avec des dérivées mesurables.

Théorème 1.7. *Supposons que σ satisfait (1.30) et que l'une des conditions suivantes soit vérifiée :*

$$(A) \quad \sigma, b \in L^\infty([0, T] \times \mathbb{R}^d), \text{ la fonction } \sigma \text{ est continue en } (t, x) \text{ et appartient à } W_{2(d+1), \text{loc}}^{0,1}([0, T]).$$

(B) $\sigma \in W_{2(d+1),\text{loc}}^{0,1}([0, T]) \cap L^\infty([0, T] \times \mathbb{R}^d)$, σ est uniformément continue en x . La fonction $b \in L_q^p([0, T])$ pour certains p, q tels que $d/p + 2/q < 1$, $2(d+1) \leq p$ et $q > 2$.

Alors, il existe $T > 0$ assez petit tel que l'équation (1.29) admet une unique solution forte X dont les trajectoires sont continues et

$$\text{la loi } \mu_x \text{ de } X \text{ satisfait } T_2(C)$$

avec une constante C dépendant des données $\|b\|_{L_q^p}, \|\sigma\|_\infty, T, x, d, p$ et q .

Nous expliquons notre stratégie pour montrer le théorème 1.7. En gros, dans les deux situations (A) et (B), la méthode reste la même et elle consiste à transformer l'équation (1.29) à une sans dérive. Lorsque l'hypothèse (A) est satisfaite, nous suivons la démarche développée par [3] dans la preuve d'unicité forte des EDS avec une dérive mesurable et un coefficient de diffusion localement Sobolev. Nous établissons d'abord l'inégalité de transport pour les EDS avec des coefficients appartenant à certains espaces Sobolev. Ensuite, nous utilisons la transformation de Zvonkin [93] pour traiter le cas où la dérive est seulement mesurable. Quand la condition (B) est vérifiée, c'est-à-dire le cas où la dérive appartient à un espace $L_q^p([0, T])$, nous enlevons également la dérive, mais la preuve est plus élaborée. Nous modifions légèrement la méthode de [41] afin de dériver des estimations de gradient appropriées pour les EDP paraboliques singulières. Ainsi, avec les estimations de gradient pour les solutions des EDP, le théorème 1.6 est un élément essentiel dans la preuve du résultat principal.

Présentons maintenant quelques exemples de modèles de diffusions multidimensionnels à coefficients non-Lipschitz qui satisferaient nos conditions.

Interaction des particules par le biais de leur rang

Soient W_1, \dots, W_n des mouvements Browniens indépendants. Les modèles d'interaction basés sur le rang sont donnés par

$$dX^{i,n}(t) = \sum_{j=1}^n \delta_j 1_{\{X^{i,n}(t) = X^{(j),n}(t)\}} dt + \sigma^i(t) dW_i(t), \quad X^{i,n}(0) = X_i$$

pour certains nombres réels δ_j , certaines fonctions mesurables et bornées σ^i , avec $X^{(1),n}(t) \leq X^{(2),n}(t) \leq \dots \leq X^{(n),n}(t)$ est le système dans l'ordre croissant. Plus généralement, ce modèle peut s'écrire comme suit

$$dX^{i,n}(t) = b \left(\frac{1}{n} \sum_{j=1}^n 1_{\{X^{n,j}(t) \leq X^{n,i}(t)\}} \right) dt + \sigma^i(t) dW_i(t), \quad X^{i,n}(0) = X_i$$

pour une fonction (déterministe) donnée b . Ce modèle a été introduit par [30] dans le contexte de la théorie du portefeuille stochastique. Les résultats de la concentration des mesures pour ces systèmes peuvent être trouvés dans [64]. Lorsque $0 < c \leq \inf_{i,t} |\sigma^i(t)| \leq \sup_{i,t} |\sigma^i(t)| \leq C$ pour certains c, C et $b \in L^\infty$ ou $b \in L^p(\mathbb{R}, dx)$, notre résultat principal montre que la loi de $(X^{1,n}, \dots, X^{n,n})$

satisfait $T_2(C)$, avec une constante $C > 0$. Ce résultat est également vrai pour le modèle Atlas (fini) de [7] donné par

$$dX^{i,n}(t) = \sum_{j=1}^n \delta 1_{\{X^{i,n}(t)=X^{p_i,n}(t)\}} dt + \sigma^i(t) dW_i(t) \quad X^{i,n}(0) = X_i,$$

pour une constante δ et une permutation (p_1, \dots, p_n) de $(1, \dots, n)$.

Les particules en interaction quantile

Les modèles d'interaction quantile sont donnés par

$$dX^{i,n}(t) = b(t, X^{n,i}(t), V^{\alpha,n}(t))dt + \sigma(t, X^{n,i}(t))dW_i(t), \quad X^{i,n}(0) = X_i,$$

où $V^{\alpha,n}(t)$ est le quantile au niveau $\alpha \in [0, 1]$ de la mesure empirique du système $(X^{1,n}(t), \dots, X^{n,n}(t))$. C'est-à-dire,

$$V^{\alpha,n}(t) := \inf \left\{ u \in \mathbb{R} : \frac{1}{n} \sum_{i=1}^n 1_{\{X^{i,n}(t) \leq u\}} \geq \alpha \right\}.$$

Ce modèle est considéré par exemple dans [18] en relation avec les systèmes de particules échangeables. Le théorème 1.7 peut être appliqué à ce cas sous des conditions d'intégrabilité sur b et des conditions de régularité faibles sur σ .

1.2.2 Inégalité de transport pour les EDSR

Dans [44, 85], un lien entre les inégalités de transport et les inégalités de concentration pour les mesures de risque a été envisagé. D'autre part, il existe une relation entre les EDSR et les mesures de risque dynamiques, à savoir : les EDSR permettent de définir des mesures de risque dynamiques. Réciproquement, les mesures de risque dynamiques peuvent être vues comme des solutions à des EDSR particulières. Motivé par ces faits, l'objectif du **chapitre 6** est de prouver une inégalité de transport pour les EDSR.

Considérons l'EDSR suivante :

$$Y_t = h(B_T) + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s. \quad (1.31)$$

où h et f sont des fonctions réelles déterministes. Nous supposons les hypothèses suivantes :

(H.1) h est bornée et il existe une constante positive l telle que pour tout $x, x' \in \mathbb{R}$

$$|h(x) - h(x')| \leq l|x - x'|$$

(H.2) $t \rightarrow f(t, 0, 0)$ est bornée et il existe des constantes positives $L_y > 0$ et $L_z > 0$ telles que pour tout (t, y, z, y', z') in $[0, T] \times \mathbb{R}^4$

$$|f(t, y, z) - f(t, y', z')| \leq L_y|y - y'| + L_z|z - z'|.$$

Nous pouvons maintenant énoncer notre principal résultat.

Théorème 1.8. *Sous les hypothèses (H.1) et (H.2), la loi \mathbb{P}_Y de Y satisfait une inégalité de transport, $T_2(C)$, sur l'espace $\mathcal{C}([0, T], \mathbb{R})$ muni de la distance uniforme où C est une constante positive dépendant de données de l'équation.*

Notons qu'après la publication de notre résultat [1], l'article [8] est paru traitant la même propriété $T_2(C)$ pour la loi d'une EDSR multidimensionnelle.

Comme conséquences du théorème précédent, nous obtenons les inégalités de concentration suivantes :

$$\mathbb{P}\left(\sup_{t \in [0, T]} |Y_t - h(B_T)| - \mathbb{E}\left[\sup_{t \in [0, T]} |Y_t - h(B_T)|\right] > r\right) \leq \exp\left(-\frac{r^2}{8C}\right),$$

et

$$\mathbb{E} \exp(F(Y) - \mathbb{E}F(Y)) \leq \exp\left(\frac{C}{2} \|F\|_{Lip}^2\right),$$

pour toute fonction Lipschitzienne F sur $\mathcal{C}([0, T], \mathbb{R})$.

Ensuite, nous étendons notre résultat au cas d'un système d'équations différentielles stochastiques progressives rétrogrades de la forme

$$\begin{cases} \forall s \in [t, T] \\ X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) ds + \int_t^s \sigma(r, X_r^{t,x}) dW_r, \\ Y_s^{t,x} = h(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dW_r. \end{cases} \quad (1.32)$$

Nous supposons que : il existe des constantes positives k , Λ et Δ telles que pour tous $t \in [0, T]$ et $(x, y, z, x', y', z') \in \mathbb{R}^6$:

$$(H.4) \quad (x - x')(b(t, x) - b(t, x')) \leq k|x - x'|^2 \quad \text{et} \quad |b(t, x)| \leq k(1 + |x|),$$

$$(H.5) \quad |\sigma(t, x) - \sigma(t, x')| \leq k|x - x'|, \quad |\sigma(t, x)| \leq k(1 + |x|) \quad \text{et} \quad \sigma^2(t, x) \geq \Lambda,$$

$$(H.6) \quad |h(x) - h(x')| \leq k|x - x'| \quad \text{et} \quad |h(x)| \leq \Delta,$$

$$(H.7) \quad |f(t, x, y, z) - f(t, x', y, z')| \leq k(|x - x'| + |z - z'|),$$

$$(H.8) \quad (y - y')(f(t, x, y, z) - f(t, x, y', z)) \leq k|y - y'|^2 \quad \text{et} \quad |f(t, x, y, z)| \leq k(1 + |y| + |z|),$$

$$(H.9) \quad \text{les fonctions } \sigma, x \mapsto b(t, x) \text{ et } y \mapsto f(t, x, y, z) \text{ sont continues.}$$

Notons par \mathbb{P}_x la loi du couple $(X^{t,x}, Y^{t,x})$ sur l'espace $\mathcal{B} = \mathcal{C}([t, T], \mathbb{R}) \times \mathcal{C}([t, T], \mathbb{R})$.

Théorème 1.9. *Supposons que (H.4)-(H.9) et $\|\sigma\|_\infty := \sup\{|\sigma(t, x)|; x \in \mathbb{R}, t \in [0, T]\} < \infty$ soient vérifiées. Il existe une constante positive C telle que la probabilité \mathbb{P}_x satisfait $T_2(C)$ sur l'espace métrique \mathcal{B} muni la distance suivante*

$$d((x, y), (x', y')) = \left(\sup_{t \leq s \leq T} |x(s) - x'(s)|^2 + \sup_{t \leq s \leq T} |y(s) - y'(s)|^2 \right)^{\frac{1}{2}}.$$

Grâce au théorème précédent, on trouve l'inégalité de concentration suivante :

$$\mathbb{P}\left(\sup_{s \in [t, T]} |Y^{t,x}(s) - h(X_T^{t,x})| - \mathbb{E}\left[\sup_{s \in [t, T]} |Y^{t,x}(s) - h(X_T^{t,x})|\right] > r\right) \leq \exp\left(-\frac{r^2}{8C}\right).$$

Part I

NONLINEAR NEUMANN BOUNDARY VALUE PROBLEMS

Chapter 2

Penalization for a PDE with a nonlinear Neumann boundary condition and measurable coefficients

Contents

2.1 Introduction	27
2.2 Reflected stochastic differential equations	29
2.3 Backward stochastic differential equations	35
2.4 PDE with a nonlinear Neumann boundary condition	42

Abstract

We consider a system of semilinear partial differential equations with measurable coefficients and a nonlinear Neumann boundary condition. We then construct a sequence of penalized partial differential equations, which converges to our initial problem. Since the coefficients we consider may be discontinuous, we use the notion of solution in the L^p -viscosity sense. The method we use is based on backward stochastic differential equations and their S-tightness. The present work is motivated by the fact that many partial differential equations in physics have discontinuous coefficients. As a consequence, it follows that if the uniqueness holds, then the solution can be constructed by a penalization.

2.1 Introduction

Let D be a \mathcal{C}^2 convex open and bounded domain in \mathbb{R}^d . For $(t, x) \in [0, T] \times \bar{D}$, we consider the following reflecting stochastic differential equation

$$X_s = x + \int_t^s b(X_r)dr + \int_t^s \sigma(X_r)dW_r + K_s, \quad s \in [t, T], \quad (2.1)$$

where $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$ are given measurable functions and K is a bounded variation process satisfying some minimality conditions. Several authors have studied approximations of reflected diffusions in such domains. We refer, for instance, to [56, 84] in the case of a convex bounded domain D and uniformly Lipschitz coefficients. The non-convex case was treated in [49]. The latter has been extended to reflected diffusions on non necessary bounded domains in [67]. The case of non-Lipschitz coefficients and non-convex domain can be found in [75], where the authors studied the existence of a weak solution of reflected equations when the coefficients are merely measurable, and the diffusion coefficient may degenerate on some subset of the domain. Equation (2.1) is also used to handle linear partial differential equations (PDEs) with a Neumann boundary condition, see for instance [34, 77, 82].

Our aim in this paper is to show that there exists a L^p -viscosity solution to a system of PDEs with a nonlinear Neumann boundary condition and this solution can be constructed by a penalization. Our approach is based on backward stochastic differential equations (BSDEs) and their tightness properties. It is well known that the markovian BSDEs allows to provide probabilistic representations for solutions of different types of semilinear PDEs, see for instance [66, 65] for parabolic equations on \mathbb{R}^d , [20] for elliptic equations with Dirichlet boundary condition, and [71] for a nonlinear Neumann boundary condition. More references can be found in [69].

In order to describe our work, we shall give some notations which will be used in the sequel. We assume that there exists a function $l \in \mathcal{C}_b^2(\mathbb{R}^d)$ such that

$$D = \{x \in \mathbb{R}^d : l(x) > 0\}, \quad \partial D = \{x \in \mathbb{R}^d : l(x) = 0\}$$

and for all $x \in \partial D$, $\nabla l(x)$ is the unit normal pointing toward the interior of D . We explain how to define the approximation procedure. We consider the function $x \mapsto \text{dist}^2(x, \bar{D})$. This function is \mathcal{C}^1 and convex on \mathbb{R}^d . We choose l such that, for any $x \in \mathbb{R}^d$,

$$\langle \nabla l(x), \delta(x) \rangle \leq 0,$$

where $\delta(x) := \nabla (\text{dist}^2(x, \bar{D}))$ called the penalization term.

We also have, for every $x \in \mathbb{R}^d$,

$$\frac{1}{2}\delta(x) = \frac{1}{2}\nabla (\text{dist}^2(x, \bar{D})) = x - \pi_{\bar{D}}(x),$$

where $\pi_{\bar{D}}$ is the projection operator. Note also that, δ is a Lipschitz function and we have

$$\langle z - x, \delta(x) \rangle \leq 0, \quad \forall x \in \mathbb{R}^d, \quad \forall z \in \bar{D}. \tag{2.2}$$

We consider the following sequence of semilinear partial differential equations ($1 \leq i \leq k$, $0 \leq t \leq T$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$).

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L} u_i^n(t, x) + f_i(t, x, u^n(t, x)) \\ \quad - n \langle \delta(x), \nabla u_i^n(t, x) \rangle - n \langle \delta(x), \nabla l(x) \rangle h_i(t, x, u^n(t, x)) = 0, \\ u^n(T, x) = g(x), \end{cases} \tag{2.3}$$

where \mathcal{L} is the infinitesimal generator of the diffusion X and defined by

$$\mathcal{L} := \frac{1}{2} \sum_{i,j} (\sigma \sigma^*(\cdot))_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(\cdot) \frac{\partial}{\partial x_i}. \quad (2.4)$$

We first establish that for each n , u^n exists and defines a L^p -viscosity solution to the penalized PDE (2.3). Next, using the connection between BSDEs and semilinear PDEs, we prove that the sequence $u^n(t, x)$ converges to a function $u(t, x)$ which is the L^p -viscosity solution of the following system of PDEs with Neumann boundary condition, for $1 \leq i \leq k$,

$$\begin{cases} \frac{\partial u_i}{\partial t}(t, x) + \mathcal{L}u_i(t, x) + f_i(t, x, u(t, x)) = 0, & (t, x) \in [0, T) \times D, \\ u_i(T, x) = g_i(x), & x \in D \\ \frac{\partial u_i}{\partial n}(t, x) + h_i(t, x, u(t, x)) = 0, & (t, x) \in [0, T) \times \partial D, \end{cases} \quad (2.5)$$

where $\frac{\partial u_i}{\partial n}(t, x) = \langle \nabla l(x), \nabla u_i(t, x) \rangle$ for all $x \in \partial D$.

It turns out that the convergence of u^n to u follows from the uniqueness in law of the forward part, although the coefficients are merely measurable.

The penalization of nonlinear Neumann boundary problem (2.5) has been firstly considered in [11] when the coefficients b, σ are uniformly Lipschitz then extended to the case where (b, σ) are continuous in [5]. The result of the present paper can be seen as an extension of the works [5, 11] to the case where the coefficients b and σ and the nonlinear term f is measurable in the state variable x . Our method is inspired from those developed in [5, 11]. The difficulty in our framework is the discontinuity of the coefficients, which makes the convergence of the sequence of penalized equations more delicate. Note also that, due to the discontinuity of the coefficients, the classical viscosity solution, used in [5, 11, 71], can not be defined for our PDEs. Therefore, we use the notion of L^p -viscosity solution, for which we give here a probabilistic interpretation. The theory of L^p -viscosity solution has been introduced in [15] for elliptic PDEs and developed in [17] for parabolic PDEs.

The paper is outlined as follows. In Section 2, we prove the convergence of solutions of our penalized SDE. The continuity of the solution with respect to the initial data is also established for both penalized and reflected SDEs. In Section 3, the same properties are shown for the solutions of the BSDEs parts, which are our first main results. Section 4 is devoted to the approximation of the system of PDEs with nonlinear Neumann boundary conditions by a penalization, which is the second fundamental aim of this paper.

2.2 Reflected stochastic differential equations

Throughout the paper, T is a fixed strictly positive number and $d, d' \in \mathbb{N}^*$. Consider a stochastic differential equation with reflecting boundary condition

of the form

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(X_r^{t,x}) dr + \int_t^s \sigma(X_r^{t,x}) dW_r + K_s^{t,x}, \\ K_s^{t,x} = \int_t^s \nabla l(X_r^{t,x}) d|K^{t,x}|_{[t,r]}, \\ |K^{t,x}|_{[t,s]} = \int_t^s 1_{\{X_r^{t,x} \in \partial D\}} d|K^{t,x}|_{[t,r]}, \end{cases} \quad (2.6)$$

where $t \in [0, T]$, $s \in [t, T]$, and the notation $|K^{t,x}|_{[t,s]}$ stands for the total variation of $K^{t,x}$ on the interval $[t, s]$. We will denote this continuous increasing process by $k_s^{t,x}$. In particular, we have

$$k_s^{t,x} = \int_t^s \langle \nabla l(X_r^{t,x}), dK_r^{t,x} \rangle. \quad (2.7)$$

We say that $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s\}, W, X, K)$ is a weak solution of (3.3) if $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_s\})$ is a stochastic basis, W is a d' -dimensional Brownian motion defined on this basis, X is a continuous adapted process and K is a continuous bounded variation process such that $X_s \in \bar{D}$ \mathbb{P} -a.s., $\forall s \in [t, T]$, and (X, K) satisfies system (3.3).

We precise some notations. $\mathcal{C}([0, T], \mathbb{R}^d)$ denotes the space of \mathbb{R}^d -valued continuous functions. $\mathcal{D}([0, T], \mathbb{R}^d)$ is the space of \mathbb{R}^d -valued cadlag functions. $W_{p,loc}^{1,2}([0, T] \times \mathbb{R}^d)$ is the classical Sobolev space of functions φ with values in \mathbb{R} such that both φ and all the generalized derivatives $\partial_t \varphi$, $\partial_x \varphi$ and $\partial_{xx}^2 \varphi$ belong to $L_{loc}^p([0, T] \times \mathbb{R}^d)$. For a given sequence of processes $(Y^n)_n$ and a process Y , the convergence $Y^n \xrightarrow[U]{*} Y$ means the convergence in law with respect to the uniform topology, $Y^n \xrightarrow[S]{*} Y$ denotes the weak convergence with respect to the Jakubowski S -topology, see Appendix (A) for a brief presentation of this topology, and [35] for more details.

We consider the following assumptions

(A.1) $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$ are measurable and bounded functions,

(A.2) There exists $\alpha > 0$ such that for all $x \in \mathbb{R}^d$ $\sigma \sigma^*(x) \geq \alpha I$,

(A.3) The weak uniqueness holds for equation (3.3).

The reflected diffusions with measurable coefficients were considered in [75, 81], where the authors have proved some approximations, stability, and existence results. Note that when the coefficients are discontinuous, the uniqueness generally failed. The presence of assumption (A.3) is explained by the fact that the weak uniqueness is crucial in our approach.

Remark 2.1. We assume that one of the following sets of conditions is satisfied.

1. the dimension $d \leq 2$, assumptions (A.1)-(A.2) hold, and D verifies some assumptions given in [43],

2. b is measurable and bounded, σ is continuous and bounded and $\sigma\sigma^*$ is uniformly non degenerate (i.e. uniformly elliptic).

Then, the weak uniqueness holds for equation (3.3).

Indeed, let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}, W, X, K)$ be a weak solution of (3.3) and $f \in C^{1,2}([0, T] \times \bar{D})$. Itô's formula applied to $f(s, X_s)$ gives

$$\begin{aligned} f(s, X_s) &= f(t, x) + \int_t^s \left(\frac{\partial f}{\partial r} + \mathcal{L}f \right) (r, X_r) dr + \int_t^s \langle \nabla_x f(r, X_r), \nabla l(X_r) \rangle dk_r \\ &\quad + \int_t^s \langle \nabla_x f(r, X_r), \sigma(X_r) dW_r \rangle. \end{aligned} \quad (2.8)$$

Since $\sigma\sigma^*$ is non degenerate, we use Krylov's inequality for reflecting diffusions (see Theorem 5.1 in [75]) to get for every $s \in [t, T]$,

$$\begin{aligned} \mathbb{E} \int_t^s \left| \left(\frac{\partial f}{\partial r} + \mathcal{L}f \right) (r, X_r) \right| 1_{\{X_r \in \partial D\}} dr \\ \leq C \left(\int_t^s \int_D \det(\sigma\sigma^*)^{-1} \left(\frac{\partial f}{\partial r} + \mathcal{L}f \right)^{d+1} 1_{\{\partial D\}} ds dx \right)^{\frac{1}{d+1}} = 0. \end{aligned}$$

Thus, equality (2.8) becomes

$$\begin{aligned} f(s, X_s) &= f(t, x) + \int_t^s \left(\frac{\partial f}{\partial r} + \mathcal{L}f \right) (r, X_r) 1_{\{X_r \in D\}} dr + \int_t^s \langle \nabla_x f(r, X_r), \nabla l(X_r) \rangle dk_r \\ &\quad + \int_t^s \langle \nabla_x f(r, X_r), \sigma(X_r) dW_r \rangle, \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Therefore

$$f(s, X_s) - f(t, x) - \int_t^s \left(\frac{\partial f}{\partial r} + \mathcal{L}f \right) (r, X_r) 1_{\{X_r \in D\}} dr$$

is a \mathbb{P} -submartingale whenever $f \in C^{1,2}([0, T] \times \bar{D})$ satisfies

$$\langle \nabla_x f(s, x), \nabla l(x) \rangle \geq 0, \quad \forall x \in \partial D.$$

Assuming the first set of conditions, we deduce from Theorem 3 in [43] that the process $(X_s)_{s \in [t, T]}$ is unique in law. Alternatively, under the second set of conditions, we apply Theorem 5.7 in [82] with $\phi = l$, $\gamma := \nabla\phi$ and $\rho := 0$ to obtain that the solution to the submartingale problem is unique for each starting point (t, x) . Therefore, our solution process $(X_s)_{s \in [t, T]}$ is unique in law. The uniqueness in law of the couple (X, K) follows then from Theorem 6 in [28].

We now consider the penalized SDEs related to our reflected diffusion $X^{t,x}$.

$$X_s^{t,x,n} = x + \int_t^s [b(X_r^{t,x,n}) - n\delta(X_r^{t,x,n})] dr + \int_t^s \sigma(X_r^{t,x,n}) dW_r, \quad s \in [t, T]. \quad (2.9)$$

Let $n \in \mathbb{N}^*$ be fixed. Assuming (A.1)-(A.2), we deduce from Krylov [42] that there exists a weak solution to equation (3.5). Moreover, Krylov has also established that it is possible to select a strong Markov weak solution of equation (3.5). In the sequel, we shall need to show the continuity of the flow associated to equation (3.5). To this end, we assume

(A.4) The weak uniqueness (i.e. in law) holds for equation (3.5).

Remark 2.2. Note that assumption (A.4) is not needed when the dimension space is lower or equal to 2, since it holds when conditions (A.1)-(A.2) are satisfied, see [43, 40].

For $t \in [0, T]$ and $s \in [t, T]$, we put

$$K_s^{t,x,n} := \int_t^s -n\delta(X_r^{t,x,n})dr \quad \text{and} \quad k_s^{t,x,n} := \int_t^s \langle \nabla l(X_r^{t,x,n}), dK_r^{t,x,n} \rangle.$$

The following bound is classical, it can be found for instance in [5].

$$\sup_{n \geq 1} \mathbb{E} \sup_{s \in [t, T]} |X_s^{t,x,n}|^{2q} + \sup_{n \geq 1} \mathbb{E} |K^{t,x,n}|_{[t, T]}^q < +\infty, \quad \text{for each } q \geq 1. \quad (2.10)$$

Theorem 2.3. Under assumptions (A.1)-(A.3), the equation (3.3) has a unique weak solution $(X^{t,x}, K^{t,x})$. Moreover,

$$(X^{t,x,n}, K^{t,x,n}) \xrightarrow[U \times U]{*} (X^{t,x}, K^{t,x}).$$

Proof. The existence of weak solution follows from Theorem 4.1 in [75]. The convergence of the penalization results from Theorem 3.1 in [81]. \square

Remark 2.4. If assumptions (A.1)-(A.3) are satisfied, then according to Theorems 7 and 10 of [28], the solution $X^{t,x}$ of equation (3.3) is a strong Markov process.

We extend the processes $(X^{t,x}, K^{t,x})$ and $(X^{t,x,n}, K^{t,x,n})$ to $[0, t]$ by denoting

$$X_s^{t,x} = X_s^{t,x,n} := x, \quad K_s^{t,x} = K_s^{t,x,n} := 0, \quad s \in [0, t]. \quad (2.11)$$

The following proposition can be proved by using Itô's formula and the boundedness of the coefficients b and σ .

Proposition 2.5. Let $(X^{t,x}, K^{t,x})$ be a solution of equation (3.3). Assume (A.1) be satisfied. Then, for any convergent sequence (t_n, x_n) in $[0, T] \times \bar{D}$ and any $q \geq 1$, we have

$$\sup_{n \geq 0} \mathbb{E} \sup_{s \in [0, T]} |X_s^{t_n, x_n}|^{2q} < +\infty \quad \text{and} \quad \sup_{n \geq 0} \mathbb{E} |K^{t_n, x_n}|_T^q < +\infty. \quad (2.12)$$

Proposition 2.6. Assume that (A.1), (A.2) and (A.4) are satisfied. Let $X^{t,x,n}$ be the solution of the penalized equations (3.5). Then, for any $n \in \mathbb{N}^*$, the map $[0, T] \times \mathbb{R}^d \ni (t, x) \rightarrow X^{t,x,n}$ is continuous in law.

Proof. Let $(t_m, x_m) \rightarrow (t, x)$. Arguing as in Corollary 2 in [74] (see also [42]) and using the uniqueness in law, we get for each fixed $n \in \mathbb{N}^*$

$$X^{t_m, x_m, n} \xrightarrow[U]{*} X^{t, x, n}.$$

The proposition is proved. \square

The continuity in law of the solution, of equation (3.3), in its initial data will be established in the following proposition. This constitute a slight extension of Lemma 3.8 in [5] to the case where the coefficients are measurable.

Proposition 2.7. We suppose that (A.1)-(A.3) are in force. Let $(X^{t,x}, K^{t,x})$ be a solution of equation (3.3). Then, the map $[0, T] \times \bar{D} \ni (t, x) \rightarrow (X^{t,x}, K^{t,x})$ is continuous in law.

Proof. Let $(t, x) \in [0, T] \times \bar{D}$ be fixed and $(t_n, x_n) \rightarrow (t, x)$, as $n \rightarrow +\infty$. We set

$$(X_s^{t_n, x_n}, K_s^{t_n, x_n}) := (X_s^n, K_s^n). \quad (2.13)$$

We first prove that the sequence of processes (X^n, K^n) is tight on $\mathcal{C}([0, T], \mathbb{R}^d \times \mathbb{R}^d)$. By Itô's formula applied to $X_s^n - X_r^n$, where r is fixed and $s \geq r$ we get:

$$\begin{aligned} \mathbb{E}|X_s^n - X_r^n|^8 &\leq C|s - r|^4 + C\mathbb{E} \left(\sup_{v \in [r, s]} \left| \int_r^v \langle X_u^n - X_r^n, \sigma(X_u^n) dW_u^n \rangle \right|^4 \right) \\ &\leq C|s - r|^4 + C\mathbb{E} \left(\int_r^s |X_u^n - X_r^n|^2 |\sigma(X_u^n)|^2 du \right)^2 \\ &\leq C|s - r|^4 + C|s - r|^2 \leq C \max\{|s - r|^4, |s - r|^2\}. \end{aligned}$$

Concerning K^n , we have:

$$K_s^n - K_r^n = (X_s^n - X_r^n) - \int_r^s b(X_u^n) du - \int_r^s \sigma(X_u^n) dW_u^n.$$

Hence,

$$\begin{aligned} \mathbb{E}|K_s^n - K_r^n|^8 &\leq C\mathbb{E}|X_s^n - X_r^n|^8 + C\mathbb{E} \left(\int_r^s |b(X_u^n)| du \right)^8 \\ &\quad + C\mathbb{E} \left(\sup_{v \in [r, s]} \left| \int_r^v \sigma(X_u^n) dW_u^n \right|^8 \right) \\ &\leq C \max\{|s - r|^8, |s - r|^2\}. \end{aligned}$$

The above constant C may change from line to line and it does not depend on n . Therefore, (X^n, K^n, W^n) is tight on $\mathcal{C}([0, T], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$. By Prokhorov's theorem (see Chap I in [69]), there exist a process (X, K, W) and a subsequence that we still denote by (X^n, K^n, W^n) , such that

$$(X^n, K^n, W^n) \xrightarrow[U \times U \times U]{*} (X, K, W).$$

We shall show that $(X, K) \stackrel{law}{=} (X^{t,x}, K^{t,x})$. Thanks to Skorohod's representation theorem, there exist a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ and a sequence of processes $(\hat{X}^n, \hat{K}^n, \hat{W}^n)$ and a processes $(\hat{X}, \hat{K}, \hat{W})$ which are defined on this probability space and which satisfy

$$(\hat{X}^n, \hat{K}^n, \hat{W}^n) \stackrel{law}{=} (X^n, K^n, W^n), \quad (\hat{X}, \hat{K}, \hat{W}) \stackrel{law}{=} (X, K, W)$$

and $(\hat{X}^n, \hat{K}^n, \hat{W}^n) \rightarrow (\hat{X}, \hat{K}, \hat{W})$ $\hat{\mathbb{P}}$ -a.s., as $n \rightarrow \infty$. Moreover, for every n , $(\hat{W}^n, \mathcal{F}^{\hat{W}^n, \hat{X}^n})$ and $(\hat{W}, \mathcal{F}^{\hat{W}, \hat{X}})$ are Brownian motions.

We define

$$\begin{aligned}\hat{V}_s^n &:= x_n + \int_{t_n}^s b(\hat{X}_r^n) dr + \int_{t_n}^s \sigma(\hat{X}_r^n) d\hat{W}_r^n, \\ \hat{V}_s &:= x + \int_t^s b(\hat{X}_r) dr + \int_t^s \sigma(\hat{X}_r) d\hat{W}_r.\end{aligned}\tag{2.14}$$

Since the processes X^n and X have finite moments (uniformly in n) of any order, $\sigma\sigma^*$ is non degenerate and the coefficients b, σ are bounded, then using Skorokhod's theorem ([79] p. 32) and Krylov's estimate, one can show that:

$$\begin{aligned}\int_t^s b(\hat{X}_r^n) dr &\xrightarrow{\text{Proba}} \int_t^s b(\hat{X}_r) dr \quad \text{as } n \rightarrow +\infty, \\ \int_t^s \sigma(\hat{X}_r^n) d\hat{W}_r^n &\xrightarrow{\text{Proba}} \int_t^s \sigma(\hat{X}_r) d\hat{W}_r, \quad \text{as } n \rightarrow +\infty.\end{aligned}$$

Since b and σ are bounded, we deduce by the Lebesgue dominated theorem that the following limit holds in $L^q(\hat{\Omega})$ for each $q \geq 1$, and for each $s \in [t, T]$

$$\hat{\mathbb{E}} \left| \hat{V}_s^n - \hat{V}_s \right|^q \rightarrow 0, \quad \text{as } n \rightarrow +\infty.$$

Putting

$$V_s^n := x_n + \int_{t_n}^s b(X_r^n) dr + \int_{t_n}^s \sigma(X_r^n) dW_r^n,$$

it follows that $X_s^n = V_s^n + K_s^n$. Further, we remark that

$$(X^n, K^n, W^n, V^n) \stackrel{\text{law}}{=} (\hat{X}^n, \hat{K}^n, \hat{W}^n, \hat{V}^n) \quad \text{on } \mathcal{C}([0, T], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d'} \times \mathbb{R}^d)$$

then,

$$\hat{X}_s^n = \hat{V}_s^n + \hat{K}_s^n \quad \hat{\mathbb{P}}\text{-a.s.}$$

Passing to the limit on n , we get

$$\hat{X}_s = \hat{V}_s + \hat{K}_s \quad \hat{\mathbb{P}}\text{-a.s.}$$

Taking into account of (2.14), it follows that (\hat{X}, \hat{K}) is a solution of equation (3.3) with initial data (t, x) . By the uniqueness in law, we have $(\hat{X}, \hat{K}) = (X^{t,x}, K^{t,x})$. Therefore, (X^n, K^n) converges in law for the uniform topology. This completes the proof. \square

The next technical lemma is a stochastic version of Helly-Bray theorem, it can be found in Proposition 6 of [72].

Lemma 2.8. Let $(M^n, \eta^n) : (\Omega^n, \mathcal{F}^n, \mathbb{P}^n) \rightarrow \mathcal{C}([0, T], \mathbb{R}^d)^2$ be a sequence of random variables and $(M, \eta) : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathcal{C}([0, T], \mathbb{R}^d)^2$ such that

$$(M^n, \eta^n) \xrightarrow[U \times U]{*} (M, \eta).$$

If $(\eta^n)_n$ has an a.s. bounded variation and

$$\sup_{n \geq 1} \mathbb{P}^n (|\eta^n|_{[0,T]} > a) \rightarrow 0, \quad \text{as } a \rightarrow +\infty$$

then η has an a.s. bounded variation and

$$\int_0^\cdot \langle M_r^n, d\eta_r^n \rangle \xrightarrow{*}_U \int_0^\cdot \langle M_r, d\eta_r \rangle, \quad \text{as } n \rightarrow +\infty.$$

The following convergence are deduced from the previous lemma.

Lemma 2.9. Assume (A.1)-(A.4) be satisfied. Then,

$$k^{t,x,n} \xrightarrow{*}_U k^{t,x} \quad \text{and} \quad k^{t_n,x_n} \xrightarrow{*}_U k^{t,x}.$$

Proof. Since $(X^{t,x,n}, K^{t,x,n}) \xrightarrow{*}_{U \times U} (X^{t,x}, K^{t,x})$, then Lemma 2.8, applied to $(M^n, \eta^n) := (\nabla l(X^{t,x,n}), K^{t,x,n})$, allows us to get $k^{t,x,n} \xrightarrow{*}_U k^{t,x}$. We prove the second convergence. By the continuity in law with respect to the initial data, $(X^{t_n,x_n}, K^{t_n,x_n}) \xrightarrow{*}_{U \times U} (X^{t,x}, K^{t,x})$. Once again, Lemma 2.8, applied now with $(M^n, \eta^n) := (\nabla l(X^{t_n,x_n}), K^{t_n,x_n})$, allows to get $k^{t_n,x_n} \xrightarrow{*}_U k^{t,x}$. \square

2.3 Backward stochastic differential equations

Consider the functions $f, h : [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}^k$, satisfying the following assumptions:

(A.5) There exist positive constants C_1, C_2, l_h and $\mu_f \in \mathbb{R}$, $\beta < 0$ and $q \geq 1$ such that for every $s, t \in [0, T]$ and every $(x, x', y, y') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^k)^2$, we have

- (i) $\langle y' - y, f(t, x, y') - f(t, x, y) \rangle \leq \mu_f |y' - y|^2$,
- (ii) $|h(t, x', y') - h(s, x, y)| \leq l_h (|t - s| + |x' - x| + |y' - y|)$,
- (iii) $\langle y' - y, h(t, x, y') - h(t, x, y) \rangle \leq \beta |y' - y|^2$,
- (iv) $|f(t, x, y)| + |h(t, x, y)| \leq C_1 (1 + |y|)$,
- (v) $|g(x)| \leq C_2 (1 + |x|^q)$.

g is continuous and f is measurable with respect to x and continuous in (t, y) .

We assume without loss of generality that the processes $(X_s^{t,x,n}, K_s^{t,x,n})_{s \in [t, T]}$ and $(X_s^{t,x}, K_s^{t,x})_{s \in [t, T]}$ are defined on the canonical space. Consider the following generalized BSDEs on $[t, T]$

$$\begin{aligned} Y_s^{t,x,n} &= g(X_T^{t,x,n}) + \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) dr - \int_s^T U_r^{t,x,n} dM_r^{X^{t,x,n}} \\ &\quad + \int_s^T h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} \end{aligned} \quad (2.15)$$

and

$$\begin{aligned} Y_s^{t,x} &= g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}) dr - \int_s^T U_r^{t,x} dM_r^{X^{t,x}} \\ &\quad + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}, \end{aligned} \quad (2.16)$$

where

$$M_s^{X^{t,x,n}} := \int_t^s \sigma(X_r^{t,x,n}) dW_r \quad \text{and} \quad M_s^{X^{t,x}} := \int_t^s \sigma(X_r^{t,x}) dW_r. \quad (2.17)$$

Under assumption (A.5), there exist $(Y_s^{t,x,n}, U_s^{t,x,n})_{s \in [t, T]}$ and $(Y_s^{t,x}, U_s^{t,x})_{s \in [t, T]}$ which are respectively the unique solutions of equations (3.11) and (3.12) (see [68, 71]).

Remark 2.10. (i) The existence and uniqueness of the solutions to both BSDEs (3.11) and (3.12) are ensured when either (A.5)(ii) or (A.5)(iii) is satisfied.

(ii) Condition (A.5)(iii) will be used to establish some estimates which allow us to prove the tightness proprieties while (A.5)(ii) will be used to identify the limit.

The next proposition will be used to get the convergence of the solutions of the sequence of penalized PDEs.

Proposition 2.11. Assume that (A.1)-(A.5) are satisfied. Then, the following convergence holds

$$(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n}) \xrightarrow[S \times S \times S]{*} (Y^{t,x}, M^{t,x}, H^{t,x})$$

where

$$\begin{aligned} M_s^{t,x,n} &:= \int_t^s U_r^{t,x,n} dM_r^{X^{t,x,n}}, \quad H_s^{t,x,n} := \int_t^s h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n}, \\ M_s^{t,x} &:= \int_t^s U_r^{t,x} dM_r^{X^{t,x}} \quad \text{and} \quad H_s^{t,x} := \int_t^s h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}. \end{aligned} \quad (2.18)$$

Moreover, $\lim_{n \rightarrow \infty} Y_t^{t,x,n} = Y_t^{t,x}$.

Proof. Thanks to [11], the solution of equation (3.11) satisfies

$$\sup_{n \geq 0} \mathbb{E} \sup_{t \leq s \leq T} |Y_s^{t,x,n}|^2 + \sup_{n \geq 0} \mathbb{E} \int_t^T \|U_s^{t,x,n} \sigma(X_s^{t,x,n})\|^2 ds < +\infty,$$

To show the tightness property with respect to the S -topology, we have to compute the conditional variation CV_T defined by (A.3) in Appendix. Arguing as in [11], we show that $(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n})$ is tight with respect to the S -topology. Then, there is a subsequence still denoted by $(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n})$ and $(\bar{Y}, \bar{M}, \bar{H})$ in $(\mathcal{D}([0, T], \mathbb{R}^k))^3$, such that

$$(X^{t,x,n}, K^{t,x,n}, k^{t,x,n}, Y^{t,x,n}, M^{t,x,n}, H^{t,x,n}) \xrightarrow[U^3 \times S^3]{*} (X^{t,x}, K^{t,x}, k^{t,x}, \bar{Y}, \bar{M}, \bar{H}). \quad (2.19)$$

We shall prove the convergence of each term in BSDE (3.11). Thanks to the Skorohod representation theorem for the uniform topology and the S -topology, one can and assume that

(i) $(X^{t,x,n}, K^{t,x,n}, k^{t,x,n}, Y^{t,x,n}, M^{t,x,n}, H^{t,x,n}, X^{t,x}, K^{t,x}, k^{t,x}, \bar{Y}, \bar{M}, \bar{H})$ are defined in the same probability space and $(X^{t,x,n}, K^{t,x,n}, k^{t,x,n}, Y^{t,x,n}, M^{t,x,n}, H^{t,x,n})$ converges to $(X^{t,x}, K^{t,x}, k^{t,x}, \bar{Y}, \bar{M}, \bar{H})$ \mathbb{P} -a.s. on $\mathcal{C}([0, T], \mathbb{R}^d)^2 \times \mathcal{C}([0, T], \mathbb{R}) \times \mathcal{D}([0, T], \mathbb{R}^k)^3$.

(ii) $(X^{t,x,n}, K^{t,x,n}, k^{t,x,n})$ converges to $(X^{t,x}, K^{t,x}, k^{t,x})$ a.s. in ω uniformly in s ,

(iii) $(Y^{t,x,n}, M^{t,x,n}, H^{t,x,n})$ converges to $(\bar{Y}, \bar{M}, \bar{H})$ a.e. in (s, ω) and $(Y_T^{t,x,n}, M_T^{t,x,n}, H_T^{t,x,n})$ converges to $(\bar{Y}_T, \bar{M}_T, \bar{H}_T)$, \mathbb{P} -a.s.

In order to make the proof more easy to read, we will divide it into several steps.

Step 1. *The convergence of $\int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) dr$.* Since the function f may be discontinuous in x , So, to prove the convergence of this term, we proceed by mollification. For $\eta \in \mathbb{N}^*$, let $f_\eta(t, x, y)$ be continuous function in (t, x, y) such that $f_\eta(t, x, y)$ verifies (A.5)(iv) with a constant independent of η and $\sup_{\{|y| \leq M\}} |f(r, x, y) - f_\eta(r, x, y)|$ tends to 0 in $L^{d+1}([0, T] \times D_R)$ as η tends to $+\infty$.

We have

$$\mathbb{E} \left| \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f(r, X_r^{t,x}, \bar{Y}_r) dr \right| \leq J_1(n, \eta) + J_2(n, \eta) + J_3(\eta) \quad (2.20)$$

where

$$\begin{aligned} J_1(n, \eta) &:= \mathbb{E} \left| \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f_\eta(r, X_r^{t,x,n}, Y_r^{t,x,n}) dr \right|, \\ J_2(n, \eta) &:= \mathbb{E} \left| \int_s^T f_\eta(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f_\eta(r, X_r^{t,x}, \bar{Y}_r) dr \right|, \\ J_3(\eta) &:= \mathbb{E} \left| \int_s^T f_\eta(r, X_r^{t,x}, \bar{Y}_r) - f(r, X_r^{t,x}, \bar{Y}_r) dr \right|. \end{aligned}$$

Since the function f_η is continuous with respect to its three arguments, we use properties (ii)-(iii) and the dominated convergence theorem to obtain that $J_2(n, \eta)$ goes to 0 as n tends to $+\infty$.

We now prove the convergence of the term $J_1(n, \eta)$. Krylov's inequality will be an essential tool to do this. For $R > 0$ let $D_R := \{x \in \mathbb{R}^d, |x| \leq R\}$ and $\tau_R^n := \inf\{r > t, |X_r^{t,x,n}| > R\} \wedge T$, with convention $\inf\{\emptyset\} = \infty$. We set

$$J_1(n, \eta, R) := \mathbb{E} \left| \int_s^{T \wedge \tau_R^n} f(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f_\eta(r, X_r^{t,x,n}, Y_r^{t,x,n}) dr \right|.$$

Let $M > 0$. We have

$$\begin{aligned}
J_1(n, \eta, R) &\leq \mathbb{E} \int_s^{T \wedge \tau_R^n} |f(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f_\eta(r, X_r^{t,x,n}, Y_r^{t,x,n})| 1_{\{|Y_r^{t,x,n}| > M\}} dr \\
&\quad + \mathbb{E} \int_s^{T \wedge \tau_R^n} |f(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f_\eta(r, X_r^{t,x,n}, Y_r^{t,x,n})| 1_{\{|Y_r^{t,x,n}| \leq M\}} dr \\
&\leq C \mathbb{E} \int_s^{T \wedge \tau_R^n} (1 + |Y_r^{t,x,n}|) 1_{\{|Y_r^{t,x,n}| > M\}} dr \\
&\quad + \mathbb{E} \int_s^{T \wedge \tau_R^n} \sup_{\{|y| \leq M\}} |f(r, X_r^{t,x,n}, y) - f_\eta(r, X_r^{t,x,n}, y)| dr \\
&\leq \frac{C}{M^{1/2}} \left(\mathbb{E} \int_s^T (1 + |Y_r^{t,x,n}|)^2 dr \right)^{\frac{1}{2}} \left(\mathbb{E} \int_s^T |Y_r^{t,x,n}| dr \right)^{\frac{1}{2}} \\
&\quad + \mathbb{E} \int_s^{T \wedge \tau_R^n} |\zeta_\eta(r, X_r^{t,x,n})| dr
\end{aligned}$$

where $\zeta_\eta(r, x) := \sup_{\{|y| \leq M\}} |f(r, x, y) - f_\eta(r, x, y)|$. Thanks to Krylov's inequality, there exists a positive constant $N(T, R, d)$ such that

$$\begin{aligned}
J_1(n, \eta, R) &\leq \frac{C}{M^{1/2}} \left(\mathbb{E} \int_s^T (1 + |Y_r^{t,x,n}|)^2 dr \right)^{\frac{1}{2}} \left(\mathbb{E} \int_s^T |Y_r^{t,x,n}| dr \right)^{\frac{1}{2}} \\
&\quad + N(T, R, d) \|\zeta_\eta\|_{L^{d+1}([0, T] \times D_R)}.
\end{aligned}$$

Sending successively η and M to $+\infty$, we show that $J_1(n, \eta, R)$ tends to zero for all $n \in \mathbb{N}^*$. Since τ_R^n increases to infinity with R , then for R large enough one can replace $T \wedge \tau_R^n$ by T . It follows that $\lim_{\eta \rightarrow +\infty} \limsup_{n \rightarrow +\infty} J_1(n, \eta) = 0$.

Arguing as for $J_1(\eta)$, one can show that $J_3(\eta)$ tends to zero as η goes to $+\infty$. Note that in the proof of the convergence, some integrability condition should be required to the process \bar{Y} , but this integrability is ensured by Lemma A.2 in [47]. This shows that

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left| \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) - f(t, X_r^{t,x}, \bar{Y}_r) dr \right| = 0.$$

Step 2. $\bar{H} = H^{t,x}$ and the expression of \bar{Y} . We successively use properties (ii)-(iii), Lemma 3.3 in [11] and the fact that the processes \bar{Y} , \bar{M} and \bar{H} are càdlàg to show that $\bar{H} = H^{t,x}$ and for every $t \leq s \leq T$

$$\bar{Y}_s = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, \bar{Y}_r) dr - (\bar{M}_T - \bar{M}_s) + \int_s^T h(r, X_r^{t,x}, \bar{Y}_r) dk_r^{t,x}.$$

Step 3. $M^{t,x} = \int_t^\cdot U_r^{t,x} dM_r^{X^{t,x}}$ is an $\mathcal{F}^{X^{t,x}, \bar{Y}, \bar{M}}$ -martingale. Lemma A.1 in [47] ensures that the process \bar{M} is a $\mathcal{F}^{X^{t,x}, \bar{Y}, \bar{M}}$ -martingale. We shall show that $M^{X^{t,x}}$ is a $\mathcal{F}^{X^{t,x}, \bar{Y}, \bar{M}}$ -martingale. Let $\psi_s : \mathcal{C}([t, s], \mathbb{R}^d) \times \mathcal{D}([t, s], \mathbb{R}^k)^2 \rightarrow \mathbb{R}$ be a bounded continuous function and $\varphi \in \mathcal{C}_b^\infty(\mathbb{R}^d)$. Let \mathcal{L} be the infinitesimal generator of the diffusion $X^{t,x}$ defined by (2.4).

Itô's formula shows that the process $\varphi(X_s^{t,x,n}) - \varphi(x) - \int_t^s \mathcal{L}\varphi(X_r^{t,x,n})dr - \int_t^s \nabla\varphi(X_r^{t,x,n})dK_r^{t,x,n}$ is a $\mathcal{F}^{X^{t,x,n}}$ -martingale. For any $t \leq s_1 < s_2 \leq T$ and any $n \in \mathbb{N}^*$, we have

$$\mathbb{E} \left[\psi_{s_1}(X^{t,x,n}, Y^{t,x,n}, M^{t,x,n}) \left(\varphi(X_{s_2}^{t,x,n}) - \varphi(X_{s_1}^{t,x,n}) - \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr - \int_{s_1}^{s_2} \nabla\varphi(X_r^{t,x,n})dK_r^{t,x,n} \right) \right] = 0.$$

Moreover,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E} \left[\psi_{s_1}(X^{t,x,n}, Y^{t,x,n}, M^{t,x,n}) \left(\varphi(X_{s_2}^{t,x,n}) - \varphi(X_{s_1}^{t,x,n}) - \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr \right) \right] \\ &= \mathbb{E} \left[\psi_{s_1}(X^{t,x}, \bar{Y}, \bar{M}) \left(\varphi(X_{s_2}^{t,x}) - \varphi(X_{s_1}^{t,x}) - \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x})dr \right) \right]. \end{aligned}$$

To prove the previous limit, we restrict ourselves to the convergence of

$$\psi_{s_1}(X^{t,x,n}, Y^{t,x,n}, M^{t,x,n}) \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr,$$

the other terms are similarly handled. We have

$$\begin{aligned} & \mathbb{E} \left| \psi_{s_1}(X^{t,x,n}, Y^{t,x,n}, M^{t,x,n}) \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr - \psi_{s_1}(X^{t,x}, \bar{Y}, \bar{M}) \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x})dr \right| \\ & \leq \mathbb{E} \left| \left[\psi_{s_1}(X^{t,x,n}, Y^{t,x,n}, M^{t,x,n}) - \psi_{s_1}(X^{t,x}, \bar{Y}, \bar{M}) \right] \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr \right| \\ & \quad + \mathbb{E} \left| \psi_{s_1}(X^{t,x}, \bar{Y}, \bar{M}) \left[\int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr - \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x})dr \right] \right| \\ & := B_1(n) + B_2(n) \end{aligned}$$

Using properties (i) and (ii), the continuity of ψ and the boundedness of b , σ , φ , $\frac{\partial\varphi}{\partial x_i}$, and $\frac{\partial^2\varphi}{\partial x_i\partial x_j}$, we obtain by the dominated convergence theorem that $\lim_{n \rightarrow +\infty} B_1(n) = 0$.

We now show that $\lim_{n \rightarrow +\infty} B_2(n) = 0$. We have

$$\begin{aligned} B_2(n) & := \mathbb{E} \left| \psi_{s_1}(X^{t,x}, \bar{Y}, \bar{M}) \left[\int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x,n})dr - \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x})dr \right] \right| \\ & \leq C \mathbb{E} \int_{s_1}^{s_2} |\mathcal{L}\varphi(X_r^{t,x,n}) - \mathcal{L}\varphi(X_r^{t,x})| dr \\ & \leq C \sum_i^d \mathbb{E} \int_{s_1}^{s_2} \left| b_i(X_r^{t,x,n}) \frac{\partial\varphi}{\partial x_i}(X_r^{t,x,n}) - b_i(X_r^{t,x}) \frac{\partial\varphi}{\partial x_i}(X_r^{t,x}) \right| dr \\ & \quad + C \sum_{i,j}^d \mathbb{E} \int_{s_1}^{s_2} \left| (\sigma\sigma^*(X_r^{t,x,n}))_{ij} \frac{\partial^2\varphi}{\partial x_i\partial x_j}(X_r^{t,x,n}) - (\sigma\sigma^*(X_r^{t,x}))_{ij} \frac{\partial^2\varphi}{\partial x_i\partial x_j}(X_r^{t,x}) \right| dr. \end{aligned}$$

We use the boundedness of b , σ , $\frac{\partial \varphi}{\partial x_i}$ and $\frac{\partial^2 \varphi}{\partial x_i \partial x_j}$ combined with Krylov's estimate and we proceed as in the proof of (2.20) to conclude that $\lim_{n \rightarrow +\infty} B_2(n) = 0$.

On the other hand, we use (i) and (ii), the boundedness and the continuity of ψ , the regularity of $\nabla \varphi$, estimate (3.7) and Lemma 2.2 in [11] to get

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E} \left[\psi_{s_1} (X^{t,x,n}, Y^{t,x,n}, M^{t,x,n}) \int_{s_1}^{s_2} \nabla \varphi(X_r^{t,x,n}) dK_r^{t,x,n} \right] \\ &= \mathbb{E} \left[\psi_{s_1} (X^{t,x}, \bar{Y}, \bar{M}) \int_{s_1}^{s_2} \nabla \varphi(X_r^{t,x}) dK_r^{t,x} \right]. \end{aligned}$$

Hence,

$$\mathbb{E} \left[\psi_{s_1} (X^{t,x}, \bar{Y}, \bar{M}) \left(\varphi(X_{s_2}^{t,x}) - \varphi(X_{s_1}^{t,x}) - \int_{s_1}^{s_2} \mathcal{L}\varphi(X_r^{t,x}) dr - \int_{s_1}^{s_2} \nabla \varphi(X_r^{t,x}) dK_r^{t,x} \right) \right] = 0.$$

Itô's formula applied to $\varphi(X_s^{t,x})$ between $s_1 < s_2$ allows to show that

$$\mathbb{E} \left[\psi_{s_1} (X^{t,x}, \bar{Y}, \bar{M}) \int_{s_1}^{s_2} \nabla \varphi(X_r^{t,x}) dM_r^{X^{t,x}} \right] = 0.$$

Hence, $M^{X^{t,x}}$ is a $\mathcal{F}^{X^{t,x}, \bar{Y}, \bar{M}}$ -martingale. Since $Y^{t,x}$ and $U^{t,x}$ are $\mathcal{F}^{X^{t,x}}$ -adapted, $M^{t,x} = \int_t^\cdot U_r^{t,x} dM_r^{X^{t,x}}$ is also a $\mathcal{F}^{X^{t,x}, \bar{Y}, \bar{M}}$ -martingale.

Thanks to Lemma A.2 in [47], \bar{Y} and \bar{M} are square-integrable processes. Using the Burkholder-Davis-Gundy inequality, we show that $\int_t^\cdot \bar{Y}_r d\bar{M}_r$ is an $\mathcal{F}^{X^{t,x}, \bar{Y}, \bar{M}}$ -martingale.

Step 4. We show that $\bar{Y} = Y^{t,x}$ and $\bar{M} = M^{t,x}$. We use Itô's formula to get, for all $s \in [t, T]$,

$$\begin{aligned} & |Y_s^{t,x} - \bar{Y}_s|^2 + ([M^{t,x} - \bar{M}]_T - [M^{t,x} - \bar{M}]_s) = \\ & 2 \int_s^T \langle Y_r^{t,x} - \bar{Y}_r, f(r, X_r^{t,x}, Y_r^{t,x}) - f(r, X_r^{t,x}, \bar{Y}_r) \rangle dr \\ & + 2 \int_s^T \langle Y_r^{t,x} - \bar{Y}_r, h(r, X_r^{t,x}, Y_r^{t,x}) - h(r, X_r^{t,x}, \bar{Y}_r) \rangle dk_r^{t,x} \\ & - 2 \int_s^T \langle Y_r^{t,x} - \bar{Y}_r, d(M_r^{t,x} - \bar{M}_r) \rangle. \end{aligned}$$

We take expectation in the previous equality then we use assumptions (A.5)(i) and (A.5)(iii) to obtain

$$\mathbb{E} |Y_s^{t,x} - \bar{Y}_s|^2 + \mathbb{E} ([M^{t,x} - \bar{M}]_T - [M^{t,x} - \bar{M}]_s) \leq 2\mu_f^+ \mathbb{E} \int_s^T |Y_r^{t,x} - \bar{Y}_r|^2 dr.$$

Hence, by Gronwall's inequality, we infer that $Y^{t,x} = \bar{Y}$ and $M^{t,x} = \bar{M}$.

In order to complete the proof, we shall show that $Y_t^{t,x}$ satisfies the initial BSDE. Given that $M_T^{t,x,n} \rightarrow M_T^{t,x}$ almost surely and since

$$\begin{aligned} Y_t^{t,x,n} &= g(X_T^{t,x,n}) + \int_t^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}) dr - M_T^{t,x,n} \\ &\quad + \int_t^T h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n}, \end{aligned}$$

we then pass to the limit to get

$$Y_t^{t,x} = g(X_T^{t,x}) + \int_t^T f(r, X_r^{t,x}, Y_r^{t,x}) dr - M_T^{t,x} + \int_t^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}.$$

The proof is finished. \square

We shall establish the continuity of the map $(t, x) \rightarrow Y_t^{t,x}$. To this end, we extend $(Y^{t,x}, U^{t,x})$ and $M^{X^{t,x}}$ to $[0, t]$ as follows

$$Y_s^{t,x} := Y_t^{t,x}, \quad U_s^{t,x} := 0 \quad \text{and} \quad M_s^{X^{t,x}} := 0, \quad \text{for all } s \in [0, t]. \quad (2.21)$$

Proposition 2.12. Under assumptions (A.1)-(A.5), the functions $u^n : (t, x) \in [0, T] \times \mathbb{R}^d \mapsto u^n(t, x) := Y_t^{t,x,n}$ and $u : (t, x) \in [0, T] \times \bar{D} \mapsto u(t, x) := Y_t^{t,x}$ are continuous.

Proof. We will show only that the function u is continuous. Let $(t_n, x_n) \rightarrow (t, x)$, as $n \rightarrow +\infty$. We have

$$\begin{aligned} \mathbb{E} \sup_{0 \leq s \leq T} |Y_s^{t_n, x_n}|^2 + \mathbb{E} \int_0^T \|U_s^{t_n, x_n} \sigma(X_s^{t_n, x_n})\|^2 ds &\leq C \quad \text{for all } n \in \mathbb{N}, \\ \mathbb{E} \sup_{0 \leq s \leq T} |Y_s^{t, x}|^2 + \mathbb{E} \int_0^T \|U_s^{t, x} \sigma(X_s^{t, x})\|^2 ds &\leq C, \end{aligned}$$

where C is a positive constant independent of n , see [71]. We recall that the quantities $M^{t,x}$ and $H^{t,x}$ are given by (2.18). Since

$$\begin{aligned} \sup_{n \geq 0} \left(\text{CV}_T(Y^{t_n, x_n}) + \mathbb{E} \sup_{0 \leq s \leq T} |Y_s^{t_n, x_n}| + \mathbb{E} \sup_{0 \leq s \leq T} |M_s^{t_n, x_n}| \right. \\ \left. + \text{CV}_T(H^{t_n, x_n}) + \mathbb{E} \sup_{0 \leq s \leq T} |H_s^{t_n, x_n}| \right) < \infty, \end{aligned}$$

it follows that $(Y^{t_n, x_n}, M^{t_n, x_n}, H^{t_n, x_n})$ is tight with respect to the S -topology. So there exists a subsequence still denoted by $(Y^{t_n, x_n}, M^{t_n, x_n}, H^{t_n, x_n})$ and $(\bar{Y}, \bar{M}, \bar{H})$ in $(\mathcal{D}([0, T], \mathbb{R}^k))^3$, such that

$$(X^{t_n, x_n}, K^{t_n, x_n}, k^{t_n, x_n} Y^{t_n, x_n}, M^{t_n, x_n}, H^{t_n, x_n}) \xrightarrow[U^3 \times S^3]^* (X^{t, x}, K^{t, x}, k^{t, x} \bar{Y}, \bar{M}, \bar{H}).$$

The same arguments used in the proof of Proposition 2.11 ensure that for every $s \in [t, T]$

$$\bar{Y}_s = g(X_T^{t,x}) + \int_s^T 1_{[t, T]} f(r, X_r^{t,x}, \bar{Y}_r) dr - (\bar{M}_T - \bar{M}_s) + \int_s^T h(r, X_r^{t,x}, \bar{Y}_r) dk_r^{t,x}$$

and

$$Y^{t,x} = \bar{Y} \quad \text{and} \quad M^{t,x} = \bar{M},$$

On the other hand, Since

$$\begin{aligned} u(t_n, x_n) = Y_{t_n}^{t_n, x_n} &= Y_0^{t_n, x_n} = g(X_T^{t_n, x_n}) + \int_0^T 1_{[t_n, T]} f(r, X_r^{t_n, x_n}, Y_r^{t_n, x_n}) dr - M_T^{t_n, x_n} \\ &\quad + \int_0^T h(r, X_r^{t_n, x_n}, Y_r^{t_n, x_n}) dk_r^{t_n, x_n}, \end{aligned}$$

then arguing as in the proof of Proposition 2.11, we deduce that $\lim_{n \rightarrow \infty} u(t_n, x_n) = u(t, x)$ with

$$\begin{aligned} u(t, x) = Y_t^{t,x} &= Y_0^{t,x} = g(X_T^{t,x}) + \int_0^T 1_{[t,T]} f(r, X_r^{t,x}, Y_r^{t,x}) dr - M_T^{t,x} \\ &\quad + \int_0^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x}. \end{aligned}$$

Proposition 2.12 is proved. \square

2.4 PDE with a nonlinear Neumann boundary condition

Since the coefficients of our PDEs are not continuous, we cannot define the solutions in the classical viscosity sense. We then use the L^p -viscosity solution. This notion of solutions has been introduced by [15] in order to study nonlinear PDEs with measurable coefficients. The definition of the L^p -viscosity solution for the PDE (2.3) is given below after the following notations:

$$\mathcal{L}_n \varphi := \mathcal{L} \varphi - n \langle \delta(\cdot), \nabla \varphi \rangle$$

and

$$\bar{f}^n(r, x, y) := f(r, x, y) - n \langle \nabla l(x), \delta(x) \rangle h(r, x, y).$$

Definition 2.13. Let p be an integer such that $p > d + 1$.

1. A function $u \in \mathcal{C}([0, T] \times \mathbb{R}^d, \mathbb{R}^k)$ is a L^p -viscosity sub-solution of the system of PDEs (2.3) if for every $x \in \mathbb{R}^d$, $u_i(T, x) \leq g_i(x)$, $1 \leq i \leq k$ and, for every $\varphi \in W_{p,loc}^{1,2}([0, T] \times \mathbb{R}^d)$ and $(\hat{t}, \hat{x}) \in [0, T] \times \mathbb{R}^d$ at which $u_i - \varphi$ has a local maximum, one has

$$\operatorname{ess\,lim\,inf}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \left\{ -\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}_n \varphi(t, x) - \bar{f}_i^n(t, x, u(t, x)) \right\} \leq 0.$$

2. A function $u \in \mathcal{C}([0, T] \times \mathbb{R}^d, \mathbb{R}^k)$ is a L^p -viscosity super-solution of the PDEs (2.3) if for every $x \in \mathbb{R}^d$, $u_i(T, x) \geq g_i(x)$, $1 \leq i \leq k$ and, for every $\varphi \in W_{p,loc}^{1,2}([0, T] \times \mathbb{R}^d)$ and $(\hat{t}, \hat{x}) \in [0, T] \times \mathbb{R}^d$ at which $u_i - \varphi$ has a local minimum, one has

$$\operatorname{ess\,lim\,sup}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \left\{ -\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}_n \varphi(t, x) - \bar{f}_i^n(t, x, u(t, x)) \right\} \geq 0.$$

3. A function $u \in \mathcal{C}([0, T] \times \mathbb{R}^d, \mathbb{R}^k)$ is a L^p -viscosity solution if it is both a L^p -viscosity sub-solution and super-solution.

Remark 2.14. Assertion (1) of the previous definition means that: for every $\varepsilon > 0$, $r > 0$, there exists a set $A \subset B_r(\hat{t}, \hat{x})$ of positive measure such that

$$-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}_n \varphi(t, x) - \bar{f}_i^n(t, x, u(t, x)) \leq \varepsilon, \quad \forall (t, x) \in A.$$

We now define the L^p -viscosity solution for system (2.5), which can be seen as a natural extension of the notion of viscosity solution of PDEs with nonlinear Neumann boundary condition, to the case of PDEs with measurable coefficients.

Definition 2.15. Let p be an integer satisfying $p > d + 1$.

- (i) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^k)$ is called a L^p -viscosity subsolution of system (2.5) if $u_i(T, x) \leq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq k$ and, for any $1 \leq i \leq k$ any $\varphi \in W_{p,loc}^{1,2}([0, T] \times \bar{D})$ and any $(\hat{t}, \hat{x}) \in [0, T] \times \bar{D}$ at which $u_i - \varphi$ has a local maximum, one has

$$\operatorname{ess\,lim\,inf}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \left\{ -\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)) \right\} \leq 0, \quad \text{if } \hat{x} \in D,$$

$$\operatorname{ess\,lim\,inf}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \min \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)), \right. \\ \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \leq 0, \quad \text{if } \hat{x} \in \partial D.$$

- (ii) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^k)$ is called a L^p -viscosity super-solution of (2.5) if $u_i(T, x) \geq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq k$ and, for any $1 \leq i \leq k$ any $\varphi \in W_{p,loc}^{1,2}([0, T] \times \bar{D})$ and any $(\hat{t}, \hat{x}) \in [0, T] \times \bar{D}$ at which $u_i - \varphi$ has a local minimum, one has

$$\operatorname{ess\,lim\,sup}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \left\{ -\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)) \right\} \geq 0, \quad \text{if } \hat{x} \in D,$$

$$\operatorname{ess\,lim\,sup}_{(t,x) \rightarrow (\hat{t}, \hat{x})} \max \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x)), \right. \\ \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \geq 0, \quad \text{if } \hat{x} \in \partial D.$$

- (iii) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^k)$ is called a L^p -viscosity solution of system (2.5) if it is both a L^p -viscosity sub- and super-solution.

Remark 2.16. Note that if the coefficients in the definition above are continuous and the test function φ is in $\mathcal{C}^{1,2}$, we then recover the classical viscosity solution of PDEs with Neumann boundary condition defined in [71].

The main result of this part is.

Theorem 2.17. Let assumptions (A.1)-(A.5) be satisfied and $p > d + 1$. Let $u^n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^k$ and $u : [0, T] \times \bar{D} \rightarrow \mathbb{R}^k$ be the functions defined in Proposition 2.12. Then, for every n , u^n and u are respectively L^p -viscosity solutions to systems (2.3) and (2.5) and we have

$$\lim_{n \rightarrow \infty} u^n(t, x) = u(t, x), \quad \text{for every } (t, x) \in [0, T] \times \bar{D},$$

We divide the proof of Theorem 2.17 into two lemmas. We also use Proposition 2.11 in the proof.

Lemma 2.18. The function u^n is a L^p -viscosity solution to system (2.3).

Proof. We follow the idea used in Proposition 5.1 of [2]. We will show only that u^n is a L^p -viscosity subsolution, the property of the super-solution property can be handled similarly. Let $\varphi \in W_{p,loc}^{1,2}([0, T] \times \mathbb{R}^d)$. Let $(\hat{t}, \hat{x}) \in [0, T] \times \mathbb{R}^d$ be a point which is a local maximum of $u_i^n - \varphi$. Since $p > d + 1$, then φ admits a continuous version which we consider from now on. We assume without loss of generality that

$$u_i^n(\hat{t}, \hat{x}) = \varphi(\hat{t}, \hat{x}). \quad (2.22)$$

We will argue by contradiction. Assume that there exists $\varepsilon, \alpha > 0$ such that

$$\frac{\partial \varphi}{\partial t}(t, x) + \mathcal{L}_n \varphi(t, x) + \bar{f}_i^n(t, x, u^n(t, x)) < -\varepsilon, \quad \lambda\text{-a.e. in } B_\alpha(\hat{t}, \hat{x}) \quad (2.23)$$

where λ denotes the Lebesgue measure and $B_\alpha(\hat{t}, \hat{x})$ is the ball of center (\hat{t}, \hat{x}) and radius α . Since (\hat{t}, \hat{x}) is a local maximum of $u_i^n - \varphi$, then there exists a positive number α' (which we can suppose equal to α) such that

$$u_i^n(t, x) \leq \varphi(t, x) \quad \text{for all } (t, x) \text{ in } B_\alpha(\hat{t}, \hat{x}). \quad (2.24)$$

Define the stopping time

$$\tau = \inf\{s \geq \hat{t}; |X_s^{\hat{t}, \hat{x}, n} - \hat{x}| > \alpha\} \wedge (\hat{t} + \alpha)$$

Since $X^{t,x,n}$ is a Markov diffusion, one can show that for every $r \in [\hat{t}, \hat{t} + \alpha]$, $Y_r^{\hat{t}, \hat{x}, n} = u^n(r, X_r^{\hat{t}, \hat{x}, n})$. Hence, the process $(\bar{Y}_s, \bar{U}_s) := \left(Y_s^{\hat{t}, \hat{x}, n, i}, 1_{[0, \tau]} U_s^{\hat{t}, \hat{x}, n, i} \right)_{s \in [\hat{t}, \hat{t} + \alpha]}$ solves the following BSDE in the interval $[\hat{t}, \hat{t} + \alpha]$

$$\begin{aligned} \bar{Y}_s &= u_i^n(\tau, X_\tau^{\hat{t}, \hat{x}, n}) + \int_s^{\hat{t} + \alpha} 1_{[0, \tau]} \bar{f}_i^n(r, X_r^{\hat{t}, \hat{x}, n}, u^n(r, X_r^{\hat{t}, \hat{x}, n})) dr \\ &\quad - \int_s^{\hat{t} + \alpha} \bar{U}_r dM_r^{X^{\hat{t}, \hat{x}, n}}. \end{aligned} \quad (2.25)$$

On the other hand, by Itô-Krylov's formula (see Chap. 2 Sec. 2 and 3 in [42]), the process $(\hat{Y}_s, \hat{U}_s)_{s \in [\hat{t}, \hat{t} + \alpha]}$ defined by

$$(\hat{Y}_s, \hat{U}_s) := \left(\varphi(s \wedge \tau, X_{s \wedge \tau}^{\hat{t}, \hat{x}, n}), 1_{[0, \tau]} \nabla \varphi(s, X_s^{\hat{t}, \hat{x}, n}) \right)$$

satisfies

$$\hat{Y}_s = \varphi(\tau, X_\tau^{\hat{t}, \hat{x}, n}) - \int_s^{\hat{t} + \alpha} 1_{[0, \tau]} \left(\frac{\partial \varphi}{\partial r} + \mathcal{L}_n \varphi \right) (r, X_r^{\hat{t}, \hat{x}, n}) dr - \int_s^{\hat{t} + \alpha} \hat{U}_r dM_r^{X^{\hat{t}, \hat{x}, n}}.$$

By the choice of τ , we have $(\tau, X_\tau^{\hat{t}, \hat{x}, n}) \in B_\alpha(\hat{t}, \hat{x})$ and hence $u_i^n(\tau, X_\tau^{\hat{t}, \hat{x}, n}) \leq \varphi(\tau, X_\tau^{\hat{t}, \hat{x}, n})$.

Consider the set

$$A := \left\{ (t, x) \in B_\alpha(\hat{t}, \hat{x}), \left(\frac{\partial \varphi}{\partial t} + \mathcal{L}_n \varphi + \bar{f}_i^n(\cdot, \cdot, u^n(\cdot, \cdot)) \right) (t, x) < -\varepsilon \right\}$$

and $A^c := B_\alpha(\hat{t}, \hat{x}) \setminus A$ is the complement of A . By assumption (2.23) we get $\lambda(A^c) = 0$. Since the process $X_r^{\hat{t}, \hat{x}, n}$ is non degenerate, Krylov's inequality implies that $1_{A^c}(r, X_r^{\hat{t}, \hat{x}, n}) = 0$ $dr \times \mathbb{P}$ -a.e. It follows that

$$0 < \mathbb{E}(\tau - \hat{t})\varepsilon \leq \mathbb{E} \int_{\hat{t}}^{\hat{t}+\alpha} -1_{[0, \tau]} \left[\left(\frac{\partial \varphi}{\partial r} + \mathcal{L}_n \varphi \right) (r, X_r^{\hat{t}, \hat{x}, n}) + \bar{f}_i^n(r, X_r^{\hat{t}, \hat{x}, n}, u^n(r, X_r^{\hat{t}, \hat{x}, n})) \right] dr.$$

This implies that

$$-1_{[0, \tau]} \left[\left(\frac{\partial \varphi}{\partial r} + \mathcal{L}_n \varphi \right) (r, X_r^{\hat{t}, \hat{x}, n}) + \bar{f}_i^n(r, X_r^{\hat{t}, \hat{x}, n}, u^n(r, X_r^{\hat{t}, \hat{x}, n})) \right] > 0$$

on a set of $dt \times d\mathbb{P}$ -positive measure. Therefore, Theorem 2.4 in [68] shows that $\hat{Y}_{\hat{t}} > \bar{Y}_{\hat{t}}$, that is $\varphi(\hat{t}, \hat{x}) > u_i^n(\hat{t}, \hat{x})$, which contradicts assumption (2.22). \square

Lemma 2.19. The function u is a L^p -viscosity solution of system (2.5) in the sense of Definition 2.15.

Proof. We shall prove that u is a L^p -viscosity subsolution. Let $\varphi \in W_{p, loc}^{1,2}([0, T] \times \bar{D})$ and let $(\hat{t}, \hat{x}) \in [0, T] \times \bar{D}$ be a point which is a local maximum of $u_i - \varphi$. We consider a continuous version of φ and we assume without loss of generality that

$$u_i(\hat{t}, \hat{x}) = \varphi(\hat{t}, \hat{x}). \quad (2.26)$$

We skip the proof in the case $\hat{x} \in D$ because of it is similar with that of u^n in the previous lemma. Let us consider the case $\hat{x} \in \partial D$, we suppose that there exist $\varepsilon, \alpha > 0$ such that

$$\begin{aligned} \frac{\partial \varphi}{\partial t}(t, x) + \mathcal{L}\varphi(t, x) + f_i(t, x, u(t, x)) &< -\varepsilon \quad \text{and} \\ \frac{\partial \varphi}{\partial n}(t, x) + h_i(t, x, u(t, x)) &< -\varepsilon \quad \lambda\text{-a.e. in } B_\alpha(\hat{t}, \hat{x}). \end{aligned} \quad (2.27)$$

Since (\hat{t}, \hat{x}) is a local maximum of $u_i - \varphi$, we have

$$u_i(t, x) \leq \varphi(t, x) \quad \text{in } B_\alpha(\hat{t}, \hat{x})$$

Define

$$\tau := \inf\{s \geq \hat{t} : |X_s^{\hat{t}, \hat{x}} - \hat{x}| > \alpha\} \wedge (\hat{t} + \alpha).$$

Since $X^{t,x}$ is a Markov process, we have for every $r \in [\hat{t}, \hat{t} + \alpha]$, $Y_r^{\hat{t}, \hat{x}} = u(r, X_r^{\hat{t}, \hat{x}})$. Moreover, the process $(\bar{Y}_s, \bar{U}_s) := (Y_{s \wedge \tau}^{\hat{t}, \hat{x}, i}, 1_{[0, \tau]}(s)U_s^{\hat{t}, \hat{x}, i})$ solves the following equation in the interval $[\hat{t}, \hat{t} + \alpha]$,

$$\begin{aligned} \bar{Y}_s &= u_i(\tau, X_\tau^{\hat{t}, \hat{x}}) + \int_s^{\hat{t}+\alpha} 1_{[0, \tau]} f_i(r, X_r^{\hat{t}, \hat{x}}, u(r, X_r^{\hat{t}, \hat{x}})) dr - \int_s^{\hat{t}+\alpha} \bar{U}_r dM_r^{X^{\hat{t}, \hat{x}}} \\ &\quad + \int_s^{\hat{t}+\alpha} h_i(r, X_r^{\hat{t}, \hat{x}}, u(r, X_r^{\hat{t}, \hat{x}})) dk_r^{\hat{t}, \hat{x}}. \end{aligned}$$

On the other hand, by Itô-Krylov's formula, the process $(\hat{Y}_s, \hat{U}_s) := (\varphi(s \wedge \tau, X_{s \wedge \tau}^{\hat{t}, \hat{x}}), 1_{[0, \tau]} \nabla \varphi(s, X_s^{\hat{t}, \hat{x}}))$ solves the following BSDE

$$\begin{aligned} \hat{Y}_s &= \varphi(\tau, X_\tau^{\hat{t}, \hat{x}}) - \int_s^{\hat{t}+\alpha} 1_{[0, \tau]}(r) \left(\frac{\partial \varphi}{\partial t} + \mathcal{L}\varphi \right) (r, X_r^{\hat{t}, \hat{x}}) dr - \int_s^{\hat{t}+\alpha} \hat{U}_r dM_r^{X_r^{\hat{t}, \hat{x}}} \\ &\quad - \int_s^{\hat{t}+\alpha} 1_{[0, \tau]}(r) \frac{\partial \varphi}{\partial n} (r, X_r^{\hat{t}, \hat{x}}) dk_r^{\hat{t}, \hat{x}}. \end{aligned}$$

We consider the set

$$A = \{(t, x) \in B_\alpha(\hat{t}, \hat{x}) : \frac{\partial \varphi}{\partial t}(t, x) + \mathcal{L}\varphi(t, x) + f_i(t, x, u(t, x)) < -\varepsilon\}$$

then $\lambda(A^c) = 0$, where A^c is the complement set of A . By Krylov's inequality for semimartingales (see [6, 55, 75]) we get $1_{A^c}(r, X_r^{\hat{t}, \hat{x}}) = 0$ $dr \times d\mathbb{P}$ -a.e., it follows that

$$\mathbb{E} \int_{\hat{t}}^{\hat{t}+\alpha} -1_{[0, \tau]}(r) \left[\left(\frac{\partial \varphi}{\partial t} + \mathcal{L}\varphi \right) (r, X_r^{\hat{t}, \hat{x}}) + f_i(r, X_r^{\hat{t}, \hat{x}}, u(r, X_r^{\hat{t}, \hat{x}})) \right] dr \geq \mathbb{E}(\tau - \hat{t})\varepsilon > 0.$$

Therefore,

$$-1_{[0, \tau]}(r) \left(\frac{\partial \varphi}{\partial t} + \mathcal{L}\varphi \right) (r, X_r^{\hat{t}, \hat{x}}) > 1_{[0, \tau]}(r) f_i(r, X_r^{\hat{t}, \hat{x}}, u(r, X_r^{\hat{t}, \hat{x}}))$$

on a set of $dr \times d\mathbb{P}$ positive measure. Furthermore, by Theorem 1.4 in [71] we get $\hat{Y}_{\hat{t}} > \bar{Y}_{\hat{t}}$, which contradicts our assumption (2.26). \square

Chapter 3

Approximation of a degenerate semilinear PDE with a nonlinear Neumann boundary condition

Contents

3.1	Introduction	48
3.2	Preliminaries and formulation of the problem	49
3.2.1	Penalization for reflected stochastic differential equation	50
3.2.2	Backward inequality	52
3.2.3	The BSDEs associated to the nonlinear Neumann problem	55
3.3	Penalization of the nonlinear Neumann PDE	57
3.3.1	Convergence of the penalized BSDE	57
3.3.2	Convergence of the penalized PDE	66

Abstract

We consider a system of semilinear partial differential equations (PDEs) where the nonlinearity depends on both the solution and its gradient. The Neumann boundary condition depends on the solution in a nonlinear manner. We show that this problem admits a viscosity solution which can be approximated by a penalization. The uniform ellipticity is not required to the diffusion coefficient. Our approach is probabilistic. We extend, in particular, the result of [11] and, in some sense, those obtained in [5, 4]. In contrast to the previous works, we do not use the weak compactness of the laws of the stochastic system associated to our problem. We construct a sequence of penalized stochastic system of Markovian backward stochastic differential equations (BSDEs) then we directly show its strong convergence. This allows us to deal with the case where the nonlinearity depends on both the solution and its gradient.

3.1 Introduction

Let D be a regular convex, open and bounded subset of \mathbb{R}^d . We introduce the function $\rho \in \mathcal{C}_b^1(\mathbb{R}^d)$ such that $\rho = 0$ in \bar{D} , $\rho > 0$ in $\mathbb{R}^d \setminus \bar{D}$ and $\rho(x) = (d(x, \bar{D}))^2$ in a neighborhood of \bar{D} . On the other hand, since the domain D is smooth (say \mathcal{C}^3), it is possible to consider an extension $l \in \mathcal{C}_b^2(\mathbb{R}^d)$ of the function $d(\cdot, \partial D)$ such that D and ∂D are characterized by

$$D = \{x \in \mathbb{R}^d : l(x) > 0\} \quad \text{and} \quad \partial D = \{x \in \mathbb{R}^d : l(x) = 0\},$$

and for every $x \in \partial D$, $\nabla l(x)$ coincides with the unit normal pointing toward the interior of D (see for example [49, Remark 3.1]). In particular we may and do choose ρ and l such that

$$\langle \nabla l(x), \delta(x) \rangle \leq 0, \quad \text{for all } x \in \mathbb{R}^d, \quad (3.1)$$

where $\delta(x) := \nabla \rho(x)$ and is called the penalization term. We have

$$\frac{1}{2}\delta(x) = x - \pi_{\bar{D}}(x), \quad \forall x \in \mathbb{R}^d$$

where $\pi_{\bar{D}}$ is the projection operator on \bar{D} . Consider the second-order differential operator

$$\mathcal{L} = \frac{1}{2} \sum_{i,j} (\sigma \sigma^*(\cdot))_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(\cdot) \frac{\partial}{\partial x_i}$$

where $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$ are given functions satisfying suitable assumptions.

Our first aim is to establish the existence of a viscosity solution via penalization to the following system of partial differential equations with nonlinear Neumann boundary condition, defined for $1 \leq i \leq m$, $0 \leq t \leq T$, $x \in D$.

$$\begin{cases} \frac{\partial u_i}{\partial t}(t, x) + \mathcal{L}u_i(t, x) + f_i(t, x, u(t, x), (\nabla u \sigma)(t, x)) = 0 \\ u(T, x) = g(x), \quad x \in D \\ \frac{\partial u}{\partial n}(t, x) + h(t, x, u(t, x)) = 0, \quad (t, x) \in [0, T) \times \partial D. \end{cases} \quad (3.2)$$

To this end, consider a sequence (u^n) of viscosity solutions of the following semi-linear partial differential equations ($1 \leq i \leq m$, $0 \leq t \leq T$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$).

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}u_i^n(t, x) + f_i(t, x, u^n(t, x), (\nabla u^n \sigma)(t, x)) \\ \quad - n < \delta(x), \nabla u_i^n(t, x) + \nabla l(x) > h_i(t, x, u^n(t, x)) = 0; \\ u^n(T, x) = g(x). \end{cases}$$

The authors of [5, 4, 11] considered the case where f depends only on u but not on ∇u . Using the connection between backward stochastic differential equations

(BSDEs) and partial differential equations (PDEs), the convergence of u^n to u has been established in [11] for bounded and uniformly Lipschitz coefficients b and σ . The authors of [5] extend the result of [11] to the case where b and σ are bounded continuous. The case where b , σ and f are bounded measurable is considered in [4] in the framework of L^p -viscosity solution. The techniques developed in the previous works rely on tightness properties of the associated sequence of BSDEs in the Jakubowski S -topology. The main drawback of this method is that it does not allow to deal with the nonlinearity f depending on ∇u . Here, our method is direct and does not require weak compactness properties. The convergence of the penalized BSDE is provided only by the convergence of the penalized forward SDE. Usually, when the nonlinearity f depends in the gradient of the solution, PDEs techniques are used to control ∇u in order to get the convergence of the associated BSDE. Here, our approach is completely different. Namely, we use a purely probabilistic method, which allows us to deal with (possibly) degenerate PDEs. Our proof of convergence of the penalized BSDEs adopts the idea used by the authors of [70] in order to prove the continuity of the solution of a system of SDE-BSDE in its initial variables (t, x) . By bringing some modifications in the computations of [70], we prove the convergence of our sequence of penalized BSDEs. In the the same spirit, we establish the stability of the system (3.2) with respect to its coefficients. Precisely, we consider a sequence (u^n) of viscosity solutions of the following nonlinear Neumann PDEs

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}^n u_i^n(t, x) + f_i^n(t, x, u^n(t, x), (\nabla u^n \sigma^n)(t, x)) = 0, \\ 1 \leq i \leq m, (t, x) \in [0, T] \times D, \\ u^n(T, x) = g^n(x), x \in D \\ \frac{\partial u^n}{\partial n}(t, x) + h^n(t, x, u^n(t, x)) = 0, \forall (t, x) \in [0, T] \times \partial D, \end{cases}$$

where

$$\mathcal{L}^n = \frac{1}{2} \sum_{i,j} (\sigma^n (\sigma^n)^*(\cdot))_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i^n(\cdot) \frac{\partial}{\partial x_i}.$$

We show that when $(b^n, \sigma^n, f^n, g^n, h^n)$ converge to (b, σ, f, g, h) , then u^n converges to u which is a viscosity solution of the system (3.2).

The paper is organized as follows: Section 2 contains some facts about reflected stochastic differential equations (SDEs) and generalized BSDEs. This mainly consists in approximation, stability, existence, uniqueness results and a priori estimates of the solutions. Section 3 is devoted to the penalization of the nonlinear Neumann PDE. In section 4, we investigate the stability of solutions of a nonlinear Neumann boundary value problem.

3.2 Preliminaries and formulation of the problem

Throughout the paper, for a fixed $T > 0$, $(W_t; t \in [0, T])$ is a d -dimensional Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and for

every $t \in [0, T]$, \mathcal{F}_s^t is the σ -algebra $\sigma(W_r; t \leq r \leq s) \vee \mathcal{N}$ if $s \geq t$ and $\mathcal{F}_s^t = \mathcal{N}$ if $s \leq t$, where \mathcal{N} is the \mathbb{P} -zero sets of \mathcal{F} . For $q \geq 0$, we denote by $\mathcal{S}_d^q[0, T]$ the space of continuous progressively measurable stochastic processes $X : \Omega \times [0, T] \rightarrow \mathbb{R}^d$, such that for $q > 0$ we have

$$\mathbb{E} \sup_{t \in [0, T]} |X_t|^q < +\infty.$$

For $q \geq 0$, we denote by $\mathcal{M}_d^q(0, T)$ the space of progressively measurable stochastic processes $X : \Omega \times [0, T] \rightarrow \mathbb{R}^d$ such that:

$$\mathbb{E} \left[\left(\int_0^T |X_t|^2 dt \right)^{\frac{q}{2}} \right] < +\infty \quad \text{if } q > 0; \quad \text{and} \quad \int_0^T |X_t|^2 dt < +\infty \quad \mathbb{P}\text{-a.s.} \quad \text{if } q = 0.$$

3.2.1 Penalization for reflected stochastic differential equation

Let $(t, x) \in [0, T] \times \bar{D}$. The reflected SDE under consideration is

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(X_r^{t,x}) dr + \int_t^s \sigma(X_r^{t,x}) dW_r + K_s^{t,x}, \\ K_s^{t,x} = \int_t^s \nabla l(X_r^{t,x}) d|K^{t,x}|_{[t,r]}, \\ |K^{t,x}|_{[t,s]} = \int_t^s 1_{\{X_r^{t,x} \in \partial D\}} d|K^{t,x}|_{[t,r]}, \quad s \in [t, T], \end{cases} \quad (3.3)$$

where the notation $|K^{t,x}|_{[t,s]}$ stands for the total variation of $K^{t,x}$ on the interval $[t, s]$, we denote this continuous increasing process by $k_s^{t,x}$. In particular we have

$$k_s^{t,x} = \int_t^s \langle \nabla l(X_r^{t,x}), dK_r^{t,x} \rangle. \quad (3.4)$$

Several authors have studied the problem of the existence of solutions of the reflected diffusion and its approximation by solutions of equations with penalization terms, we refer for example to [49, 45, 81, 82, 84]. We consider the following sequence of penalized SDEs associated with our reflected diffusion $X^{t,x}$

$$X_s^{t,x,n} = x + \int_t^s [b(X_r^{t,x,n}) - n\delta(X_r^{t,x,n})] dr + \int_t^s \sigma(X_r^{t,x,n}) dW_r, \quad s \in [t, T]. \quad (3.5)$$

For $s \in [t, T]$, we put

$$K_s^{t,x,n} := \int_t^s -n\delta(X_r^{t,x,n}) dr \quad \text{and} \quad k_s^{t,x,n} := \int_t^s \langle \nabla l(X_r^{t,x,n}), dK_r^{t,x,n} \rangle. \quad (3.6)$$

We introduce the following assumption

(A.1): There exist positive constants C and μ such that for every $(x, y) \in \mathbb{R}^d$:

- (i) $|b(x)| + |\sigma(x)| \leq C(1 + |x|)$,
- (ii) $|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq \mu|x - y|$.

It is known that under assumption (A.1) equation (3.5) admits, for any fixed $n \in \mathbb{N}$, a unique strong solution, and we have for every $q \geq 1$:

$$\sup_{n \geq 0} \mathbb{E} \sup_{s \in [t, T]} |X_s^{t,x,n}|^{2q} + \sup_{n \geq 0} \mathbb{E} \sup_{s \in [t, T]} |K_s^{t,x,n}|^{2q} + \sup_{n \geq 0} \mathbb{E} |K_{[t, T]}^{t,x,n}|^q < +\infty. \quad (3.7)$$

The proof of the previous estimates can be found e.g. in [5, Lemma 3.1].

The first assertion of the following theorem is proved in [84], while the second one follows from [81].

Theorem 3.1. Under assumption (A.1), we have

- (i) the system (3.3) admits a unique solution,
- (ii) for every $1 \leq q < \infty$ and $0 < T < \infty$,

$$\mathbb{E} \left[\sup_{t \leq s \leq T} |X_s^{t,x,n} - X_s^{t,x}|^q \right] \longrightarrow 0, \quad \text{as } n \rightarrow \infty,$$

the limit is uniform in $(t, x) \in [0, T] \times \bar{D}$.

We extend the processes $(X^{t,x}, K^{t,x})$ and $(X^{t,x,n}, K^{t,x,n})$ to $[0, t)$ by putting

$$X_s^{t,x} = X_s^{t,x,n} := x, \quad K_s^{t,x} = K_s^{t,x,n} := 0, \quad \text{for } s \in [0, t).$$

As a consequence of Theorem 3.1, we have the following convergence, which is established in [11, Lemma 2.2].

Lemma 3.2. Under assumptions of Theorem 3.1, we have, for any $q \geq 1$:

- (i) $\lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq s \leq T} |K_s^{t,x,n} - K_s^{t,x}|^q \right] = 0$;
- (ii) $\lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq s \leq T} \left| \int_0^s \varphi(X_r^n) dK_r^{t,x,n} - \int_0^s \varphi(X_r) dK_r^{t,x} \right|^q \right] = 0$ for every $\varphi \in \mathcal{C}_b^1(\mathbb{R}^d)$.

Remark 3.3. (i) Using Lemma 3.2 and the representations (3.4), (3.6), it holds that for any $q \geq 1$ and any $(t, x) \in [0, T] \times \bar{D}$

$$\lim_{n \rightarrow +\infty} \mathbb{E} \sup_{s \in [0, T]} |k_s^{t,x,n} - k_s^{t,x}|^q = 0. \quad (3.8)$$

- (ii) From [80, Corollary 2.5], it follows that for each $q \geq 1$ and each $(t, x) \in [0, T] \times \bar{D}$

$$\mathbb{E} \sup_{s \in [0, T]} |X_s^{t,x}|^{2q} + \mathbb{E} \sup_{s \in [0, T]} |K_s^{t,x}|^{2q} + \mathbb{E} |K_{[0, T]}^{t,x}|^q < +\infty. \quad (3.9)$$

3.2.2 Backward inequality

We state a lemma, which is a version of [69, Proposition 6.80, Annex C]. We also give the proof for the convenience.

Lemma 3.4. Let $(Y, Z) \in \mathcal{S}_m^0 \times \mathcal{M}_{m \times d}^0$, satisfying

$$Y_t = Y_T + \int_t^T d\mathcal{K}_r - \int_t^T Z_r dW_r, \quad 0 \leq t \leq T, \quad \mathbb{P} - a.s.,$$

where $\mathcal{K} \in \mathcal{S}_m^0$ and $\mathcal{K}(\omega) \in BV([0, T]; \mathbb{R}^m)$ (the space of bounded variation processes).

Assume be given

- a non-decreasing stochastic process L with $L_0 = 0$,
- a stochastic process $R(\omega) \in BV([0, T], \mathbb{R})$, with $R_0 = 0$,
- a continuous stochastic process $V(\omega) \in BV([0, T], \mathbb{R})$, such that $V_0 = 0$,

$$\mathbb{E} \left(\sup_{s \in [0, T]} \int_s^T e^{2V_r} dR_r \right) < \infty,$$

and

- (i) $\langle Y_r, d\mathcal{K}_r \rangle \geq \frac{\alpha}{2} \|Z_r\|^2 dr + |Y_r|^2 dV_r + |Y_r| dL_r + dR_r$ as measures on $[0, T]$,
- (ii) $\mathbb{E} \sup_{r \in [0, T]} e^{2V_r} |Y_r|^2 < +\infty$.

We then have the following conclusion : if $\alpha < 1$, then there exist positive constants C_1 , C_2 and C_3 , depending only on α , such that

$$\begin{aligned} & \mathbb{E} \left(\sup_{r \in [0, T]} |e^{V_r} Y_r|^2 \right) + \mathbb{E} \left(\int_0^T e^{2V_r} \|Z_r\|^2 dr \right) \\ & \leq C_1 \mathbb{E} |e^{V_T} Y_T|^2 + C_2 \mathbb{E} \left(\int_0^T e^{V_r} dL_r \right)^2 + C_3 \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2V_r} dR_r. \end{aligned}$$

Proof. By Itô's formula, we have

$$\begin{aligned} |e^{V_t} Y_t|^2 &= |e^{V_T} Y_T|^2 - 2 \int_t^T e^{2V_s} |Y_s|^2 dV_s + 2 \int_t^T e^{2V_s} \langle Y_s, d\mathcal{K}_s \rangle \\ &\quad - \int_t^T e^{2V_s} \|Z_s\|^2 ds - 2 \int_t^T \langle e^{V_s} Y_s, e^{V_s} Z_s dW_s \rangle \\ &= |e^{V_T} Y_T|^2 + 2 \int_t^T e^{2V_s} (\langle Y_s, d\mathcal{K}_s \rangle - |Y_s|^2 dV_s) - \int_t^T e^{2V_s} \|Z_s\|^2 ds \\ &\quad - 2 \int_t^T \langle e^{V_s} Y_s, e^{V_s} Z_s dW_s \rangle. \end{aligned}$$

Using (i) of Lemma 3.4, we get

$$\begin{aligned} |e^{V_t}Y_t|^2 + (1 - \alpha) \int_t^T e^{2V_s} \|Z_s\|^2 ds \leq \\ |e^{V_T}Y_T|^2 + 2 \int_t^T e^{2V_s} (dR_s + |Y_s|dL_s) - 2 \int_t^T \langle e^{V_s}Y_s, e^{V_s}Z_s dW_s \rangle. \end{aligned}$$

We consider the following sequence of stopping time

$$T_n := T \wedge \inf\{s \geq t : \sup_{r \in [t, s]} |e^{V_r}Y_r - e^{V_t}Y_t| + \int_t^s e^{2V_r} \|Z_r\|^2 dr + \int_t^s e^{V_r} dL_r \geq n\}.$$

For $s \in [t, T]$, we put $N_s := 2 \int_0^s 1_{[t, T_n]}(r) \langle e^{V_r}Y_r, e^{V_r}Z_r dW_r \rangle$.

Since

$$\begin{aligned} \mathbb{E}(\langle N \rangle_T)^{\frac{1}{2}} &\leq C \mathbb{E} \left(\int_t^{T_n} e^{4V_r} |Y_r|^2 \|Z_r\|^2 dr \right)^{\frac{1}{2}} \\ &\leq \sqrt{2} C \mathbb{E} \left(\left[\sup_{r \in [t, T_n]} |e^{V_r}Y_r - e^{V_t}Y_t|^2 + |e^{V_t}Y_t|^2 \right] \int_t^{T_n} e^{2V_r} |Z_r|^2 dr \right)^{\frac{1}{2}} \\ &\leq \sqrt{2} C \mathbb{E} (|e^{V_t}Y_t| + n) \sqrt{n} < +\infty. \end{aligned}$$

Then, the process $\{N_s : s \in [0, T]\}$ is a martingale.

By the forgoing, we have

$$\begin{aligned} |e^{V_t}Y_t|^2 + (1 - \alpha) \int_t^{T_n} e^{2V_s} \|Z_s\|^2 ds \leq |e^{V_{T_n}}Y_{T_n}|^2 + 2 \int_t^{T_n} e^{2V_s} (dR_s + |Y_s|dL_s) \\ -(N_{T_n} - N_t). \end{aligned}$$

Taking expectation, it follows that

$$\begin{aligned} \mathbb{E} \left(|e^{V_t}Y_t|^2 + (1 - \alpha) \int_t^{T_n} e^{2V_s} \|Z_s\|^2 ds \right) \leq \\ \mathbb{E} \left(|e^{V_{T_n}}Y_{T_n}|^2 + 2 \int_t^{T_n} e^{2V_s} (dR_s + |Y_s|dL_s) \right). \quad (3.10) \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \mathbb{E} \sup_{r \in [0, T_n]} |e^{V_r}Y_r|^2 \leq \\ \mathbb{E} \left(|e^{V_{T_n}}Y_{T_n}|^2 + 2 \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r + 2 \int_0^{T_n} e^{2V_s} |Y_s| dL_s + 2 \sup_{r \in [0, T_n]} |N_r| \right). \end{aligned}$$

The Burkholder-Davis-Gundy inequality shows that

$$\begin{aligned}
\mathbb{E} \sup_{r \in [0, T_n]} |e^{V_r} Y_r|^2 &\leq \mathbb{E} \left(|e^{V_{T_n}} Y_{T_n}|^2 + 2 \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r + 2 \int_0^{T_n} e^{2V_s} |Y_s| dL_s \right) \\
&\quad + 2C_{BDG} \mathbb{E} \left(\int_0^{T_n} e^{4V_r} |Y_r|^2 \|Z_r\|^2 dr \right)^{\frac{1}{2}} \\
&\leq \mathbb{E} \left(|e^{V_{T_n}} Y_{T_n}|^2 + 2 \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r + 2 \int_0^{T_n} e^{2V_s} |Y_s| dL_s \right) \\
&\quad + 2C_{BDG} \mathbb{E} \left(\sup_{r \in [0, T_n]} e^{2V_r} |Y_r|^2 \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr \right)^{\frac{1}{2}} \\
&\leq \mathbb{E} \left(|e^{V_{T_n}} Y_{T_n}|^2 + 2 \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r + 2 \int_0^{T_n} e^{2V_s} |Y_s| dL_s \right) \\
&\quad + \mathbb{E} \left(\frac{1}{2} \sup_{r \in [0, T_n]} e^{2V_r} |Y_r|^2 + 2C_{BDG}^2 \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr \right) \\
&\leq 2\mathbb{E} \left(|e^{V_{T_n}} Y_{T_n}|^2 + 2 \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r + 2 \int_0^{T_n} e^{2V_s} |Y_s| dL_s \right) \\
&\quad + 4C_{BDG}^2 \mathbb{E} \left(\int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr \right),
\end{aligned}$$

where C_{BDG} denotes the Burkholder-Davis-Gundy constant. From inequality (3.10), we get

$$\begin{aligned}
\mathbb{E} \sup_{r \in [0, T_n]} |e^{V_r} Y_r|^2 + \frac{1}{2} \mathbb{E} \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr &\leq \mu \mathbb{E} (|e^{V_{T_n}} Y_{T_n}|^2) \\
+ 2\mu \mathbb{E} \left(\sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r \right) + 2\mu \mathbb{E} \left(\int_0^{T_n} e^{2V_s} |Y_s| dL_s \right),
\end{aligned}$$

where

$$\mu = 2 + \frac{2C_{BDG}^2}{1 - \alpha} + \frac{1}{2(1 - \alpha)}.$$

It follows that

$$\begin{aligned}
\mathbb{E} \sup_{r \in [0, T_n]} |e^{V_r} Y_r|^2 + \frac{1}{2} \mathbb{E} \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr &\leq \mathbb{E} \left(\mu |e^{V_{T_n}} Y_{T_n}|^2 + 2\mu \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r \right) \\
&\quad + 2\mu \mathbb{E} \left(\sup_{r \in [0, T_n]} e^{V_r} |Y_r| \int_0^{T_n} e^{V_r} dL_r \right) \\
\mathbb{E} \sup_{r \in [0, T_n]} |e^{V_r} Y_r|^2 + \frac{1}{2} \mathbb{E} \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr &\leq \mathbb{E} \left(\mu |e^{V_{T_n}} Y_{T_n}|^2 + 2\mu \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r \right) \\
&\quad + \frac{1}{2} \mathbb{E} \sup_{r \in [0, T_n]} e^{2V_r} |Y_r|^2 + 2\mu^2 \mathbb{E} \left(\int_0^{T_n} e^{V_r} dL_r \right)^2 \\
\mathbb{E} \sup_{r \in [0, T_n]} |e^{V_r} Y_r|^2 + \mathbb{E} \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr &\leq \mathbb{E} \left(2\mu |e^{V_{T_n}} Y_{T_n}|^2 + 4\mu \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r \right) \\
&\quad + 4\mu^2 \mathbb{E} \left(\int_0^{T_n} e^{V_r} dL_r \right)^2.
\end{aligned}$$

Put

$$C_1 := 2\mu, \quad C_2 := 4\mu \quad \text{and} \quad C_3 := 4\mu^2.$$

We then have,

$$\begin{aligned}
\mathbb{E} \left(\sup_{r \in [0, T_n]} |e^{V_r} Y_r|^2 + \int_0^{T_n} e^{2V_r} \|Z_r\|^2 dr \right) &\leq C_1 \mathbb{E} |e^{V_{T_n}} Y_{T_n}|^2 + C_2 \mathbb{E} \sup_{s \in [0, T_n]} \int_s^{T_n} e^{2V_r} dR_r \\
&\quad + C_3 \mathbb{E} \left(\int_0^T e^{V_r} dL_r \right)^2.
\end{aligned}$$

Letting n tends to $+\infty$, we conclude by using the Beppo-Levi theorem for the left-hand side term of the previous inequality, and the Lebesgue dominated convergence theorem for the right-hand side term. \square

3.2.3 The BSDEs associated to the nonlinear Neumann problem

We introduce the generalized BSDEs which we have to use. Let $f : [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m \times d'} \rightarrow \mathbb{R}^m$, $h : [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}^m$ be continuous functions, satisfying the following assumptions:

(A.2) There exist C, l_f positive constants and $\beta < 0, \mu_f \in \mathbb{R}$ such that for every $t \in [0, T]$ and every $(x, x', y, y', z, z') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^m)^2 \times (\mathbb{R}^{m \times d'})^2$ we have:

- (i) $\langle y - y', f(t, x, y, z) - f(t, x, y', z) \rangle \leq \mu_f |y - y'|^2$,
- (ii) $|f(t, x, y, z) - f(t, x, y, z')| \leq l_f \|z - z'\|$,
- (iii) $|f(t, x, y, 0)| \leq C(1 + |y|)$,
- (iv) $\langle y - y', h(t, x, y) - h(t, x, y') \rangle \geq \beta |y - y'|^2$,

- (v) $|h(t, x, y)| \leq C(1 + |y|)$,
 (vi) $|g(x)| \leq C(1 + |x|)$.

For every $t \in [0, T]$ and $s \in [t, T]$, consider the following generalized BSDEs

$$\begin{aligned} Y_s^{t,x,n} &= g(X_T^{t,x,n}) + \int_s^T f(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) dr + \int_s^T h(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} \\ &\quad - \int_s^T Z_r^{t,x,n} dW_r, \end{aligned} \quad (3.11)$$

and

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}) dk_r^{t,x} - \int_s^T Z_r^{t,x} dW_r. \quad (3.12)$$

According to [68, 71], assumption (A.2) ensures the existence of unique solutions to equations (3.11) and (3.12). The solutions of equations (3.11) and (3.12) will be respectively denoted by $(Y_s^{t,x,n}, Z_s^{t,x,n})_{s \in [t, T]}$ and $(Y_s^{t,x}, Z_s^{t,x})_{s \in [t, T]}$.

Lemma 3.5. Under assumptions (A.1)(i) and (A.2), it holds that for any $t \in [0, T]$,

$$\sup_{n \geq 1} \mathbb{E} \left(\sup_{r \in [t, T]} |Y_r^{t,x,n}|^2 + \int_t^T |Y_r^{t,x,n}|^2 dk_r^{t,x,n} + \int_t^T \|Z_r^{t,x,n}\|^2 dr \right) < +\infty \quad (3.13)$$

and

$$\mathbb{E} \left(\sup_{r \in [t, T]} |Y_r^{t,x}|^q + \mathbb{E} \left(\int_t^T \|Z_r^{t,x}\|^2 dr \right)^{\frac{q}{2}} \right) < +\infty, \quad \forall q > 1. \quad (3.14)$$

Proof. We prove the first assertion. Itô's formula gives

$$\begin{aligned} |Y_s^{t,x,n}|^2 + \int_s^T \|Z_r^{t,x,n}\|^2 dr &= |g(X_T^{t,x,n})|^2 \\ &+ 2 \int_t^T \langle Y_r^{t,x,n}, f(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) \rangle dr \\ &+ 2 \int_t^T \langle Y_r^{t,x,n}, h(r, X_r^{t,x,n}, Y_r^{t,x,n}) \rangle dk_r^{t,x,n} - 2 \int_t^T \langle Y_r^{t,x,n}, Z_r^{t,x,n} dW_r \rangle. \end{aligned}$$

Using assumption (A.2), we find

$$\begin{aligned} |Y_s^{t,x,n}|^2 + \int_s^T \|Z_r^{t,x,n}\|^2 dr &\leq C^2 T + 2C^2 + 2C^2 |X_T^{t,x,n}|^2 \\ &+ (1 + 2\mu_f + 2l_f^2) \int_s^T |Y_r^{t,x,n}|^2 dr + 2 \int_s^T (\beta |Y_r^{t,x,n}|^2 + C |Y_r^{t,x,n}|) dk_r^{t,x,n} \\ &\quad - 2 \int_s^T \langle Y_r^{t,x,n}, Z_r^{t,x,n} dW_r \rangle. \end{aligned}$$

Since $-\beta > 0$, we use the inequality $2ab \leq (-\beta)a^2 + \frac{b^2}{(-\beta)}$, to get

$$\begin{aligned} |Y_s^{t,x,n}|^2 + \int_s^T \|Z_r^{t,x,n}\|^2 dr &\leq C^2 T + C^2 + C^2 |X_T^{t,x,n}|^2 \\ &\quad + (1 + 2\mu_f + 2l_f^2) \int_s^T |Y_r^{t,x,n}|^2 dr - \frac{C^2}{\beta} k_T^{t,x,n} + \beta \int_s^T |Y_r^{t,x,n}|^2 dk_r^{t,x,n} \\ &\quad - 2 \int_s^T \langle Y_r^{t,x,n}, Z_r^{t,x,n} dW_r \rangle. \end{aligned}$$

Since the process $(\int_0^\cdot \langle Y_r^{t,x,n}, Z_r^{t,x,n} \rangle dW_r)$ is a uniformly integrable martingale, we take expectation in the previous inequality to show that

$$\begin{aligned} \mathbb{E} \left(|Y_s^{t,x,n}|^2 + |\beta| \int_s^T |Y_r^{t,x,n}|^2 dk_r^{t,x,n} + \int_s^T \|Z_r^{t,x,n}\|^2 dr \right) &\leq C^2 T + C^2 + C^2 \mathbb{E} |X_T^{t,x,n}|^2 \\ &\quad + \frac{C^2}{|\beta|} \mathbb{E} k_T^{t,x,n} + (1 + 2\mu_f + 2l_f^2) \mathbb{E} \int_s^T |Y_r^{t,x,n}|^2 dr. \end{aligned}$$

Using estimate (3.7) and Gronwall's inequality, we obtain

$$\sup_{n \geq 1} \sup_{s \in [t, T]} \mathbb{E} \left(|Y_s^{t,x,n}|^2 + \int_s^T |Y_r^{t,x,n}|^2 dk_r^{t,x,n} + \int_s^T \|Z_r^{t,x,n}\|^2 dr \right) < +\infty.$$

Burkholder-Davis-Gundy inequality shows that

$$\sup_{n \geq 1} \mathbb{E} \left(\sup_{s \in [t, T]} |Y_s^{t,x,n}|^2 + \int_t^T |Y_r^{t,x,n}|^2 dk_r^{t,x,n} + \int_t^T \|Z_r^{t,x,n}\|^2 dr \right) < +\infty.$$

Inequality (3.13) is proved. Using [54, Proposition A.2], we prove inequality (3.14). \square

We extend the processes $(Y^{t,x,n}, Z^{t,x,n})$ and $(Y^{t,x}, Z^{t,x})$ to $[0, t)$ as follows

$$Y_s^{t,x,n} := Y_t^{t,x,n}, \quad Y_s^{t,x} := Y_t^{t,x} \quad \text{and} \quad Z_s^{t,x,n} = Z_s^{t,x} := 0, \quad s \in [0, t). \quad (3.15)$$

3.3 Penalization of the nonlinear Neumann PDE

We divide this section into two parts. The first one concerns the convergence of the solution of the BSDE (3.11). The second one is an application of our convergence to the nonlinear Neumann boundary problem.

3.3.1 Convergence of the penalized BSDE

For $(t, x) \in [0, T] \times \bar{D}$, let $(Y_s^{t,x,n}, Z_s^{t,x,n})_{s \in [0, T]}$ and $(Y_s^{t,x}, Z_s^{t,x})_{s \in [0, T]}$ be respectively, the solutions of BSDEs (3.11) and (3.12). Our first main result is

Theorem 3.6. Let assumptions (A.1) and (A.2) hold. Then, we have the following convergence

$$\mathbb{E} \left(\sup_{r \in [0, T]} |Y_r^{t,x,n} - Y_r^{t,x}|^2 + \int_0^T \|Z_r^{t,x,n} - Z_r^{t,x}\|^2 dr \right) \rightarrow 0, \quad \text{as } n \rightarrow +\infty.$$

Proof. We adapt the proof of [70, Theorem 3.1] to our situation by bringing some modifications. From now on, we suppress the superscripts (t, x) , and C will denote a nonnegative constant, which may vary from one line to another, but does not depend on n . We shall apply Lemma 3.4 to the following BSDE

$$Y_s^n - Y_s = g(X_T^n) - g(X_T) + \int_s^T d\mathcal{K}_r^n - \int_s^T (Z_r^n - Z_r) dW_r$$

where

$$d\mathcal{K}_r^n := [f(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r, Y_r, Z_r)] dr + h(r, X_r^n, Y_r^n) dk_r^n - h(r, X_r, Y_r) dk_r.$$

Using (A.2)(i)-(A.2)(ii), we get for every $0 \leq t \leq s \leq T$,

$$\begin{aligned} & \int_s^T \langle Y_r^n - Y_r, f(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r, Y_r, Z_r) \rangle dr \\ &= \int_s^T \langle Y_r^n - Y_r, f(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r, Z_r^n) \rangle dr \\ & \quad + \int_s^T \langle Y_r^n - Y_r, f(r, X_r^n, Y_r, Z_r^n) - f(r, X_r^n, Y_r, Z_r) \rangle dr \\ & \quad + \int_s^T \langle Y_r^n - Y_r, f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r) \rangle dr \\ & \leq \int_s^T \mu_f |Y_r^n - Y_r|^2 dr + l_f |Y_r^n - Y_r| \|Z_r^n - Z_r\| dr \\ & \quad + \int_s^T |Y_r^n - Y_r| |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr \\ & \leq \int_s^T (l_f^2 + \mu_f) |Y_r^n - Y_r|^2 dr + \frac{1}{4} \|Z_r^n - Z_r\|^2 dr \\ & \quad + \int_s^T |Y_r^n - Y_r| |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr. \end{aligned}$$

On the other hand, thanks to assumption (A.2)(iv), we obtain

$$\begin{aligned}
& \int_s^T \langle Y_r^n - Y_r, h(r, X_r^n, Y_r^n) dk_r^n - h(r, X_r, Y_r) dk_r \rangle \\
&= \int_s^T \langle Y_r^n - Y_r, h(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r) \rangle dk_r^n \\
&\quad + \int_s^T \langle Y_r^n - Y_r, h(r, X_r^n, Y_r) - h(r, X_r, Y_r) \rangle dk_r^n \\
&\quad + \int_s^T \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r) \\
&\leq \beta \int_s^T |Y_r^n - Y_r|^2 dk_r^n + \int_s^T |Y_r^n - Y_r| |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \\
&\quad + \int_s^T \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r) \\
&\leq \int_s^T |Y_r^n - Y_r| |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \\
&\quad + \int_s^T \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r).
\end{aligned}$$

By the foregoing, it holds that, for $\lambda = (l_f^2 + \mu_f) \vee l_f^2$,

$$\begin{aligned}
\langle Y_r^n - Y_r, d\mathcal{K}_r^n \rangle &\leq \frac{1}{4} \|Z_r^n - Z_r\|^2 dr + \lambda |Y_r^n - Y_r|^2 dr \\
&\quad + |Y_r^n - Y_r| |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \\
&\quad + |Y_r^n - Y_r| |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr \\
&\quad + \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r) \\
&= \frac{1}{4} \|Z_r^n - Z_r\|^2 dr + \lambda |Y_r^n - Y_r|^2 dr + |Y_r^n - Y_r| dL_r^n + dR_r^n,
\end{aligned}$$

where L^n and R^n are defined by

$$dL_r^n := |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr + |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \quad (3.16)$$

$$dR_r^n := \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r). \quad (3.17)$$

Therefore, by Lemma 3.4, there exist positive constants C_1, C_2 and C_3 such that

$$\begin{aligned}
& \mathbb{E} \left(\sup_{r \in [0, T]} e^{2\lambda r} |Y_r^n - Y_r|^2 \right) + \mathbb{E} \left(\int_0^T e^{2\lambda r} \|Z_r^n - Z_r\|^2 dr \right) \\
&\leq C_1 \mathbb{E} e^{2\lambda T} |g(X_T^n) - g(X_T)|^2 + C_2 \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 + C_3 \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n.
\end{aligned} \quad (3.18)$$

We shall give several auxiliary assertions ensuring that the right-hand side term of the previous inequality converges to zero as n goes to $+\infty$.

Lemma 3.7. Under assumptions (A.1) and (A.2)(vi), the following convergence holds

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(e^{2\lambda T} |g(X_T^n) - g(X_T)|^2 \right) = 0.$$

Proof. Taking into account the convergence of X_T^n to X_T , the continuity of g , assumption (A.4)(vi) and estimates (3.7) and (3.9), the result follows by using the Lebesgue dominated convergence theorem. \square

Lemma 3.8. Let L^n be the processes given by equation (3.16). Assume that (A.1) and (A.4) are satisfied. Then,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 = 0.$$

Proof. Using the triangular inequality, we obtain

$$\begin{aligned} \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 &\leq 2\mathbb{E} \left(\int_0^T e^{\lambda r} |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr \right)^2 \\ &\quad + 2\mathbb{E} \left(\int_0^T e^{\lambda r} |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \right)^2 \\ &:= I_1^n + I_2^n. \end{aligned}$$

We shall show that I_1^n and I_2^n tend to zero as n tends to ∞ . Hölder's inequality leads to

$$\begin{aligned} I_1^n &= \mathbb{E} \left(\int_0^T e^{\lambda r} |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr \right)^2 \\ &\leq 2T e^{2\lambda T} \mathbb{E} \left(\int_0^T |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)|^2 dr \right). \end{aligned} \quad (3.19)$$

Again by the convergence of X^n to X in each L^q with respect to the uniform norm and the continuity of f we deduce that the sequence $|f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)|^2$ converges to zero in probability, *a.e.* $r \in [0, T]$. Since by assumptions (A.2)(ii) and (A.2)(iii) on f we have

$$|f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)|^2 \leq C(1 + |Y_r|^2 + \|Z_r\|^2), \quad \textit{a.e. } r \in [0, T],$$

it follows that

$$\mathbb{E}|f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)|^2 \rightarrow 0, \quad \textit{as } n \rightarrow +\infty, \quad \textit{a.e. } r \in [0, T].$$

Using the Lebesgue dominated convergence theorem, we get $\lim_{n \rightarrow +\infty} I_1^n = 0$. Concerning I_2^n , Hölder's inequality yields

$$\begin{aligned} I_2^n &= 2\mathbb{E} \left(\int_0^T e^{\lambda r} |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \right)^2 \\ &\leq 2e^{2\lambda T} \left(\mathbb{E} \sup_{r \in [0, T]} |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)|^4 \right)^{\frac{1}{2}} \left(\sup_{n \geq 1} \mathbb{E} (k_T^n)^4 \right)^{\frac{1}{2}} \end{aligned} \quad (3.20)$$

On the other hand, by the linear growth assumption on h , we have for each $q > 1$

$$\mathbb{E} \sup_{r \in [0, T]} |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)|^{4q} \leq C(1 + \mathbb{E} \sup_{r \in [0, T]} |Y_r|^{4q}).$$

It follows from estimates (3.14) that the sequence of random variables $\sup_{r \in [0, T]} |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)|^4$ is uniformly integrable. Since X^n converges to X in each L^q for the uniform norm, we deduce that the sequence $\sup_{r \in [0, T]} |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)|^4$ converges to zero in probability as n goes to $+\infty$. This combined with estimate (3.7) ensure that $\lim_{n \rightarrow +\infty} I_2^n = 0$. Lemma 3.8 is proved. \square

We will show an estimate for the solution Y^n that will be used to control the term $\mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n$. To this end, let $N \in \mathbb{N}$, $N > T$ and the partition of $[0, T]$, $r_i = \frac{iT}{N}$ $i = 0, \dots, N$. We put $r/N := \max\{r_i; r_i \leq r\}$. Given a continuous stochastic process $(H_r)_{r \in [0, T]}$, we define

$$H_r^N := \sum_{i=0}^{N-1} H_{r_i} 1_{[r_i, r_{i+1})}(r) + H_T 1_{\{T\}}(r) = H_{r/N}.$$

Lemma 3.9. Assume (A.1) and (A.2) hold. Then, for any $q \in]1, 2[$, there exists a positive constant C depending on T , q and independent on N , such that:

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left(\int_0^T |Y_r^n - Y_r^{n, N}|^q (dk_r^n + dk_r) \right) \leq \frac{C}{N^{q/2}} + C \left[\mathbb{E} \max_{i=1, \dots, N} (k_{r_i} - k_{r_{i-1}})^{\frac{2q}{2-q}} \right]^{\frac{2-q}{4}}.$$

Proof. We write BSDE (3.11) between s/N and s

$$\begin{aligned} Y_s^{n, N} &= Y_s^n + \int_{s/N}^s f(r, X_r^n, Y_r^n, Z_r^n) dr \\ &\quad + \int_{s/N}^s h(r, X_r^n, Y_r^n) dk_r^n - \int_{s/N}^s Z_r^n dW_r. \end{aligned}$$

Hölder's inequality gives

$$\begin{aligned} |Y_s^{n, N} - Y_s^n|^q &\leq \frac{C}{N^{q/2}} \left[\int_{s/N}^s |f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right]^{q/2} \\ &\quad + C (k_s^n - k_{s/N}^n)^{q/2} \left[\int_{s/N}^s |h(r, X_r^n, Y_r^n)|^2 dk_r^n \right]^{q/2} + C \left| \int_{s/N}^s Z_r^n dW_r \right|^q. \end{aligned}$$

It follows that

$$\mathbb{E} \left(\int_0^T |Y_r^n - Y_r^{n, N}|^q (dk_r^n + dk_r) \right) \leq J_1^{n, N} + J_2^{n, N} + J_3^{n, N}$$

where

$$J_1^{n, N} := \frac{C}{N^{q/2}} \mathbb{E} \int_0^T \left[\int_{s/N}^s |f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right]^{q/2} (dk_s^n + dk_s),$$

$$J_2^{n,N} := C \mathbb{E} \int_0^T (k_s^n - k_{s/N}^n)^{q/2} \left(\int_{s/N}^s |h(r, X_r^n, Y_r^n)|^2 dk_r^n \right)^{q/2} (dk_s^n + dk_s),$$

$$J_3^{n,N} := C \mathbb{E} \int_0^T \left| \int_{s/N}^s Z_r^n dW_r \right|^q (dk_s^n + dk_s).$$

We shall estimate $J_1^{n,N}$, $J_2^{n,N}$ and $J_3^{n,N}$. We use Hölder's inequality to obtain

$$\begin{aligned} J_1^{n,N} &= \frac{C}{N^{q/2}} \mathbb{E} \int_0^T \left[\int_{s/N}^s |f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right]^{q/2} (dk_s^n + dk_s) \\ &\leq \frac{C}{N^{q/2}} \mathbb{E} \left((k_T^n + k_T) \left[\int_0^T |f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right]^{q/2} \right) \\ &\leq \frac{C}{N^{q/2}} \left(\mathbb{E}(k_T^n + k_T)^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}} \left(\mathbb{E} \int_0^T |f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right)^{q/2}. \end{aligned}$$

By the linear growth of f in its third variable and Lipschitz continuity with respect to the fourth argument, we get

$$\begin{aligned} J_1^{n,N} &\leq \frac{C}{N^{q/2}} \left(\mathbb{E}(k_T^n + k_T)^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}} \left(1 + \mathbb{E} \sup_{r \in [0, T]} |Y_r^n|^2 + \mathbb{E} \int_0^T \|Z_r^n\|^2 dr \right)^{q/2} \\ &\leq \frac{C}{N^{q/2}} \end{aligned}$$

where in the last line we have used inequalities (3.7), (3.9) and (3.13).

Concerning $J_2^{n,N}$, we use Hölder's inequality, the monotony of k^n and the

linear growth condition on h , to obtain

$$\begin{aligned}
J_2^{n,N} &= C \mathbb{E} \int_0^T (k_s^n - k_{s/N}^n)^{q/2} \left(\int_{s/N}^s |h(r, X_r^n, Y_r^n)|^2 dk_r^n \right)^{q/2} (dk_s^n + dk_s) \\
&\leq C \mathbb{E} \left(\int_0^T |h(r, X_r^n, Y_r^n)|^2 dk_r^n \right)^{q/2} \sum_{i=1}^N \int_{r_{i-1}}^{r_i} (k_s^n - k_{s/N}^n)^{q/2} (dk_s^n + dk_s) \\
&\leq C \left(\mathbb{E} \int_0^T |h(r, X_r^n, Y_r^n)|^2 dk_r^n \right)^{q/2} \\
&\quad \times \left[\mathbb{E} \left(\sum_{i=1}^N \int_{r_{i-1}}^{r_i} (k_s^n - k_{s/N}^n)^{q/2} (dk_s^n + dk_s) \right)^{\frac{2-q}{2}} \right]^{\frac{2-q}{2}} \\
&\leq C \left(\mathbb{E} \int_0^T |h(r, X_r^n, Y_r^n)|^2 dk_r^n \right)^{q/2} \\
&\quad \times \left[\mathbb{E} \left(\sum_{i=1}^N (k_{r_i}^n - k_{r_{i-1}}^n)^{q/2} (k_{r_i}^n + k_{r_i} - k_{r_{i-1}}^n - k_{r_{i-1}}) \right)^{\frac{2-q}{2}} \right]^{\frac{2-q}{2}} \\
&\leq C \left(\mathbb{E} \int_0^T 1 + |Y_r^n|^2 dk_r^n \right)^{q/2} \\
&\quad \times \left[\mathbb{E} \left(\sum_{i=1}^N (k_{r_i}^n - k_{r_{i-1}}^n)^{q/2} (k_{r_i}^n + k_{r_i} - k_{r_{i-1}}^n - k_{r_{i-1}}) \right)^{\frac{2-q}{2}} \right]^{\frac{2-q}{2}} \\
&\leq C \left(\mathbb{E} k_T^n + \mathbb{E} \int_0^T |Y_r^n|^2 dk_r^n \right)^{q/2} \\
&\quad \times \left[\mathbb{E} \left(\sum_{i=1}^N (k_{r_i}^n - k_{r_{i-1}}^n)^{q/2} (k_{r_i}^n + k_{r_i} - k_{r_{i-1}}^n - k_{r_{i-1}}) \right)^{\frac{2-q}{2}} \right]^{\frac{2-q}{2}}.
\end{aligned}$$

Hence, by inequalities (3.7) and (3.13), we see that

$$J_2^{n,N} \leq C \left[\mathbb{E} \left(\sum_{i=1}^N (k_{r_i}^n - k_{r_{i-1}}^n)^{q/2} (k_{r_i}^n + k_{r_i} - k_{r_{i-1}}^n - k_{r_{i-1}}) \right)^{\frac{2-q}{2}} \right]^{\frac{2-q}{2}}.$$

Taking into account the convergence of k^n to k (Remark 3.3 (i)), estimates (3.7) and (3.9), we pass to the limit as n goes $+\infty$ then we use the Lebesgue

dominated convergence theorem to get

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} J_2^{n,N} &\leq C \left[\mathbb{E} \left(\sum_{i=1}^N (k_{r_i} - k_{r_{i-1}})^{q/2} (k_{r_i} - k_{r_{i-1}}) \right)^{\frac{2}{2-q}} \right]^{\frac{2-q}{2}} \\
 &\leq C \left[\mathbb{E} \left(\max_{i=1, \dots, N} (k_{r_i} - k_{r_{i-1}})^{q/2} k_T \right)^{\frac{2}{2-q}} \right]^{\frac{2-q}{2}} \\
 &\leq C \left[\mathbb{E} \max_{i=1, \dots, N} (k_{r_i} - k_{r_{i-1}})^{\frac{q}{2-q}} k_T^{\frac{2}{2-q}} \right]^{\frac{2-q}{2}} \\
 &\leq C \left[\mathbb{E} \max_{i=1, \dots, N} (k_{r_i} - k_{r_{i-1}})^{\frac{2q}{2-q}} \right]^{\frac{2-q}{4}} \left[\mathbb{E} k_T^{\frac{4}{2-q}} \right]^{\frac{2-q}{4}}.
 \end{aligned}$$

For $J_3^{n,N}$, we have

$$\begin{aligned}
 J_3^{n,N} &= C \mathbb{E} \int_0^T \left| \int_{s/N}^s Z_r^n dW_r \right|^q (dk_s^n + dk_s) \\
 &= C \mathbb{E} \sum_{i=1}^N \int_{r_{i-1}}^{r_i} \left| \int_{s/N}^s Z_r^n dW_r \right|^q (dk_s^n + dk_s) \\
 &\leq C \sum_{i=1}^N \mathbb{E} \sup_{s \in [r_{i-1}, r_i]} \left| \int_{s/N}^s Z_r^n dW_r \right|^q (k_{r_i}^n - k_{r_{i-1}}^n + k_{r_i} - k_{r_{i-1}}) \\
 &\leq C \sum_{i=1}^N \left(\mathbb{E} \sup_{s \in [r_{i-1}, r_i]} \left| \int_{r_{i-1}}^s Z_r^n dW_r \right|^2 \right)^{\frac{q}{2}} \left(\mathbb{E} (k_{r_i}^n - k_{r_{i-1}}^n + k_{r_i} - k_{r_{i-1}})^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}}.
 \end{aligned}$$

Using the Burkholder-Davis-Gundy inequality, we obtain

$$J_3^{n,N} \leq C \sum_{i=1}^N \left(\mathbb{E} \int_{r_{i-1}}^{r_i} \|Z_r^n\|^2 dr \right)^{\frac{q}{2}} \left(\mathbb{E} (k_{r_i}^n - k_{r_{i-1}}^n + k_{r_i} - k_{r_{i-1}})^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}}.$$

Again by Hölder's inequality, we find

$$\begin{aligned}
 J_3^{n,N} &\leq C \left(\sum_{i=1}^N \mathbb{E} \int_{r_{i-1}}^{r_i} \|Z_r^n\|^2 dr \right)^{\frac{q}{2}} \left(\sum_{i=1}^N \mathbb{E} (k_{r_i}^n - k_{r_{i-1}}^n + k_{r_i} - k_{r_{i-1}})^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}} \\
 &\leq C \left(\mathbb{E} \int_0^T \|Z_r^n\|^2 dr \right)^{\frac{q}{2}} \left(\sum_{i=1}^N \mathbb{E} (k_{r_i}^n - k_{r_{i-1}}^n + k_{r_i} - k_{r_{i-1}})^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}}.
 \end{aligned}$$

Keeping in mind inequality (3.13) and the convergence of k^n to k , we pass to

the limit as $n \rightarrow +\infty$, to obtain

$$\begin{aligned}
 \lim_{n \rightarrow \infty} J_3^{n,N} &\leq C \left(\sum_{i=1}^N \mathbb{E} (k_{r_i} - k_{r_{i-1}})^{\frac{2}{2-q}} \right)^{\frac{2-q}{2}} \\
 &\leq C \left(\mathbb{E} \max_{i=1,N} (k_{r_i} - k_{r_{i-1}})^{\frac{q}{2-q}} \sum_{i=1}^N (k_{r_i} - k_{r_{i-1}}) \right)^{\frac{2-q}{2}} \\
 &\leq C \left(\mathbb{E} \max_{i=1,N} (k_{r_i} - k_{r_{i-1}})^{\frac{q}{2-q}} k_T \right)^{\frac{2-q}{2}} \\
 &\leq C (\mathbb{E} k_T^2)^{\frac{2-q}{4}} \left(\mathbb{E} \max_{i=1,N} (k_{r_i} - k_{r_{i-1}})^{\frac{2q}{2-q}} \right)^{\frac{2-q}{4}} \\
 &\leq C \left(\mathbb{E} \max_{i=1,N} (k_{r_i} - k_{r_{i-1}})^{\frac{2q}{2-q}} \right)^{\frac{2-q}{4}}.
 \end{aligned}$$

This completes the proof of Lemma 3.9. \square

Lemma 3.10. Let R^n be the process defined by (3.17). Under assumptions (A.1) and (A.2), the following inequality holds

$$\limsup_{n \rightarrow \infty} \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \leq 0.$$

Proof. Set $h_r = h(r, X_r, Y_r)$ and $|h|_\infty = \sup_{r \in [0, T]} |h_r|$, we have

$$\begin{aligned}
 \langle Y_r^n - Y_r, h_r \rangle &= \langle Y_r^{n,N} - Y_r^N, h_r - h_r^N \rangle + \langle Y_r^N - Y_r, h_r \rangle \\
 &\quad + \langle Y_r^{n,N} - Y_r^N, h_r^N \rangle + \langle Y_r^n - Y_r^{n,N}, h_r \rangle.
 \end{aligned}$$

$$\begin{aligned}
 &\mathbb{E} \left(\sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \right) = \\
 &\mathbb{E} \left(\sup_{s \in [0, T]} \int_s^T e^{2\lambda r} \langle Y_r^n - Y_r, h(r, X_r, Y_r) (dk_r^n - dk_r) \rangle \right) \\
 &\leq \mathbb{E} \left[(|Y^n|_\infty + |Y|_\infty) |h - h^N|_\infty + |Y^N - Y|_\infty |h|_\infty e^{2\lambda T} (k_T^n + k_T) \right] \\
 &\quad + \mathbb{E} \left(\sup_{s \in [0, T]} \sum_{i=1}^N \langle Y_{r_{i-1}}^n - Y_{r_{i-1}}, h_{r_{i-1}} \rangle \int_{s \wedge r_{i-1}}^{r_i} e^{2\lambda r} d(k_r^n - k_r) \right) \\
 &\quad + \mathbb{E} \left(e^{2\lambda T} |h|_\infty \int_0^T |Y_r^{n,N} - Y_r^n| (dk_r^n + dk_r) \right).
 \end{aligned}$$

Let $1 < q < 2$. Using Hölder's inequality repeatedly to obtain

$$\begin{aligned}
 & \mathbb{E} \left(\sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \right) \leq \\
 & \left[\mathbb{E} \left[e^{2\lambda T} (|Y^n|_\infty + |Y|_\infty)^2 \right] \right]^{\frac{1}{2}} \left[\mathbb{E} (k_T^n + k_T)^4 \right]^{\frac{1}{4}} \left[\mathbb{E} |h - h^N|_\infty^4 \right]^{\frac{1}{4}} \\
 & + \left[\mathbb{E} \left[e^{2\lambda T} (k_T^n + k_T) |h|_\infty \right] \right]^{\frac{1}{2}} \left[\mathbb{E} |Y^N - Y|_\infty^2 \right]^{\frac{1}{2}} \\
 & + 2N e^{2\lambda T} (1 + \lambda T) \left[\mathbb{E} \left[e^{2\lambda T} (|Y^n|_\infty + |Y|_\infty)^2 \right] \right]^{\frac{1}{2}} \left[\mathbb{E} |h|_\infty^4 \right]^{\frac{1}{4}} \left(\mathbb{E} \left[\sup_{s \in [0, T]} |k_s^n - k_s|^4 \right] \right)^{\frac{1}{4}} \\
 & + e^{2\lambda T} \left[\mathbb{E} |h|_\infty^{\frac{2q}{q-1}} \right]^{\frac{q-1}{2q}} \left[\mathbb{E} (k_T^n + k_T)^2 \right]^{\frac{q-1}{2q}} \left[\mathbb{E} \int_0^T |Y_r^{n, N} - Y_r^n|^q (dk_r^n + dk_r) \right]^{\frac{1}{q}}. \quad (3.21)
 \end{aligned}$$

The linear growth hypothesis on h combined with estimate (3.14) show that for every $p \geq 1$,

$$\mathbb{E} |h|_\infty^p = \mathbb{E} \sup_{r \in [0, T]} |h(r, X_r, Y_r)|^p \leq C(1 + \mathbb{E} \sup_{r \in [0, T]} |Y_r|^p) < +\infty.$$

On the other hand, we use inequalities (3.7), (3.9), (3.13) and (3.14) along with Lemma 3.9 then we pass to the limit as n goes to $+\infty$ in inequality (3.21) to get, for all $N \in \mathbb{N}^*$,

$$\begin{aligned}
 \limsup_{n \rightarrow +\infty} \mathbb{E} \left(\sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \right) & \leq C \left(\mathbb{E} |h - h^N|_\infty^4 \right)^{1/4} + C \left(\mathbb{E} |Y^N - Y|_\infty^2 \right)^{1/2} \\
 & + \left[\frac{C}{N^{q/2}} + C \left[\mathbb{E} \max_{i=1, N} (k_{r_i} - k_{r_{i-1}})^{\frac{2q}{2-q}} \right]^{\frac{2-q}{4}} \right]^{1/q}.
 \end{aligned}$$

Since the integrands are uniformly integrable, then passing to the limit as $N \rightarrow +\infty$, we get the result. Lemma 3.10 is proved. \square

Now, combining inequality (3.18) with Lemmas 3.7, 3.8 and 3.10, we complete the proof of Theorem 3.6. \square

3.3.2 Convergence of the penalized PDE

This subsection is devoted to an application of our convergence of the BSDE. Namely, we will establish the convergence of a viscosity solution of the following systems

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L} u_i^n(t, x) + f_i(t, x, u^n(t, x), (\nabla u^n \sigma)(t, x)) \\ \quad - n \langle \delta(x), \nabla u_i^n(t, x) + \nabla l(x) \rangle h_i(t, x, u^n(t, x)) = 0, \\ u^n(T, x) = g(x), 1 \leq i \leq m, 0 \leq t \leq T, x \in \mathbb{R}^d, n \in \mathbb{N} \end{cases} \quad (3.22)$$

to a viscosity solution of a system of the form

$$\begin{cases} \frac{\partial u_i}{\partial t}(t, x) + \mathcal{L}u_i(t, x) + f_i(t, x, u(t, x), (\nabla u\sigma)(t, x)) = 0, \\ 1 \leq i \leq m, (t, x) \in [0, T] \times D, \\ u(T, x) = g(x), x \in D, \\ \frac{\partial u}{\partial n}(t, x) + h(t, x, u(t, x)) = 0, \forall (t, x) \in [0, T] \times \partial D. \end{cases} \quad (3.23)$$

Since, we consider viscosity solutions, we introduce the following condition

(A.3) f_i , the i -th coordinate of f , depends only on the i -th row of the matrix z .

For the self-contained, we recall the definition of the viscosity solution of system (3.23).

Definition 3.11. (i) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^m)$ is called a viscosity sub solution of system (3.23) if $u_i(T, x) \leq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq m$ and, for any $1 \leq i \leq m$, $\varphi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$, and $(t, x) \in (0, T] \times \bar{D}$ at which $u_i - \varphi$ has a local maximum, one has

$$-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)) \leq 0, \quad \text{if } x \in D,$$

$$\begin{aligned} & \min \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)), \right. \\ & \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \leq 0, \quad \text{if } x \in \partial D. \end{aligned}$$

(ii) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^d)$ is called a viscosity super-solution of (3.23) if $u_i(T, x) \geq g_i(x)$, $x \in \bar{D}$, $1 \leq i \leq m$ and, for any $1 \leq i \leq m$, $\varphi \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^d)$, and $(t, x) \in (0, T] \times \bar{D}$ at which $u_i - \varphi$ has a local minimum, one has

$$-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)) \geq 0, \quad \text{if } x \in D,$$

$$\begin{aligned} & \max \left(-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}\varphi(t, x) - f_i(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)), \right. \\ & \left. -\frac{\partial \varphi}{\partial n}(t, x) - h_i(t, x, u(t, x)) \right) \geq 0, \quad \text{if } x \in \partial D. \end{aligned}$$

(iii) $u \in \mathcal{C}([0, T] \times \bar{D}, \mathbb{R}^m)$ is called a viscosity solution of system (3.23) if it is both a viscosity sub and super-solution.

We recall a continuity of the map $(t, x) \mapsto Y_t^{t,x}$ where $Y_t^{t,x}$ is the solution of BSDE (3.12). This continuity has been proved in [70].

Proposition 3.12. ([70]) Under assumptions (A.1) and (A.2), the mapping $(t, x) \rightarrow Y_t^{t,x}$ is continuous.

Our second main result is

Theorem 3.13. Assume (A.1), (A.2) and (A.3). There exist a sequence of continuous functions $u^n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ and a function $u : [0, T] \times \bar{D} \rightarrow \mathbb{R}^m$ such that: u^n is a viscosity solution to system (3.22), u is a viscosity solution to system (3.23) and the following convergence holds for every $(t, x) \in [0, T] \times \bar{D}$

$$\lim_{n \rightarrow +\infty} u^n(t, x) = u(t, x).$$

Proof. We set,

$$u^n(t, x) := Y_t^{t,x,n} \quad \text{and} \quad u(t, x) := Y_t^{t,x}. \quad (3.24)$$

It follows from Theorem 3.2 of [68] that u^n is a viscosity solution of PDEs (3.22). Thanks to [70, 71], u is a viscosity solution of PDEs (3.23). Further, we have for each $(t, x) \in [0, T] \times \bar{D}$

$$|u^n(t, x) - u(t, x)|^2 = |Y_t^{t,x,n} - Y_t^{t,x}|^2 \leq \mathbb{E} \sup_{s \in [0, T]} |Y_s^{t,x,n} - Y_s^{t,x}|^2.$$

Thanks to Theorem 3.6, we have $\lim_{n \rightarrow \infty} \mathbb{E} \sup_{s \in [0, T]} |Y_s^{t,x,n} - Y_s^{t,x}|^2 = 0$. It follows that,

$$\lim_{n \rightarrow +\infty} u^n(t, x) = u(t, x).$$

The theorem is proved. □

Remark 3.14. When u is the unique viscosity solution of system (3.23), then it is constructible by penalization. This is the case when $d = d'$, since by [69, Theorem 5.43, p 423] or [70, Theorem 5.1] the viscosity solution of system (3.23) is unique.

Chapter 4

A Stability result for a nonlinear Neumann problem

Contents

4.1	Introduction	69
4.2	Preliminaries	70
4.2.1	Stability of Reflected SDEs	70
4.2.2	The BSDE associated to nonlinear Neumann PDE (4.1)	72
4.3	Main result	74

Abstract

We prove a stability result for the viscosity solution of the non-linear Neumann problem (3.23).

4.1 Introduction

In the the same spirit, we establish the stability of the system (3.2) with respect to its coefficients. Precisely, we consider a sequence (u^n) of viscosity solutions of the following nonlinear Neumann PDEs

$$\begin{cases} \frac{\partial u_i^n}{\partial t}(t, x) + \mathcal{L}^n u_i^n(t, x) + f_i^n(t, x, u^n(t, x), (\nabla u^n \sigma^n)(t, x)) = 0, \\ 1 \leq i \leq m, (t, x) \in [0, T) \times D, \\ u^n(T, x) = g^n(x), x \in D \\ \frac{\partial u^n}{\partial n}(t, x) + h^n(t, x, u^n(t, x)) = 0, \forall (t, x) \in [0, T) \times \partial D, \end{cases} \quad (4.1)$$

where

$$\mathcal{L}^n = \frac{1}{2} \sum_{i,j} (\sigma^n(\sigma^n)^*(\cdot))_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i^n(\cdot) \frac{\partial}{\partial x_i}.$$

⁰This chapter uses notations and assumptions from chapter 3

We show that when $(b^n, \sigma^n, f^n, g^n, h^n)$ converge to (b, σ, f, g, h) , then u^n converges to u which is a viscosity solution of the system (3.2).

4.2 Preliminaries

4.2.1 Stability of Reflected SDEs

In order to show the stability result for the nonlinear Neumann PDEs, we need a stability of solutions for equation (3.3). To this end, we consider the following sequence of reflected SDEs

$$\begin{cases} X_s^{t,x,n} = x + \int_t^s b^n(X_r^{t,x,n}) dr + \int_t^s \sigma^n(X_r^{t,x,n}) dW_r + K_s^{t,x,n}, \\ K_s^{t,x,n} = \int_t^s \nabla l(X_r^{t,x,n}) d|K^{t,x,n}|_{[t,r]}, \\ |K^{t,x,n}|_{[t,s]} = \int_t^s 1_{\{X_r^{t,x,n} \in \partial D\}} d|K^{t,x,n}|_{[t,r]}, \quad \forall s \in [t, T], \end{cases} \quad (4.2)$$

where $b^n : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma^n : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d'}$ are sequences of measurable functions satisfying the following assumptions

(A.4) There exist positive constants C and μ_n , $n \in \mathbb{N}$, such that for all $x, y \in \mathbb{R}^d$

- (i) $\sup_{n \geq 0} |b^n(x)| + \sup_{n \geq 0} \|\sigma^n(x)\| \leq C(1 + |x|)$,
- (ii) $|b^n(x) - b^n(y)| + \|\sigma^n(x) - \sigma^n(y)\| \leq \mu_n |x - y|$.

Remark that

$$k_s^{t,x,n} = \int_t^s \langle \nabla l(X_r^{t,x,n}), dK_r^{t,x,n} \rangle. \quad (4.3)$$

We extend the processes to $[0, T]$ by setting

$$X_s^{t,x,n} := x \quad \text{and} \quad K_s^{t,x,n} := 0 \quad \text{for } s \in [0, t]. \quad (4.4)$$

The following lemma is useful to prove the stability result for equation (3.3).

Lemma 4.1. Assume that (A.4) is satisfied. Then, for each $n \in \mathbb{N}$, equation (4.2) has a unique solution $(X^{t,x,n}, K^{t,x,n})$. Moreover this solution satisfies for each $q \geq 1$

$$\sup_{n \geq 0} \mathbb{E} \sup_{s \in [0, T]} |X_s^{t,x,n}|^{2q} + \sup_{n \geq 0} \mathbb{E} \sup_{s \in [0, T]} |K_s^{t,x,n}|^{2q} + \sup_{n \geq 0} \mathbb{E} |K^{t,x,n}|_{[0, T]}^q < +\infty. \quad (4.5)$$

Proof. The existence and uniqueness of solution for equation (4.2) follows from [84]. Arguing as in the proof of [5, Lemma 3.1], one can establish the estimates (4.5). \square

We introduce now the condition

(A.5) For any compact subset κ of \mathbb{R}^d it holds

$$\lim_{n \rightarrow +\infty} \left(\sup_{x \in \kappa} |b^n(x) - b(x)| + \sup_{x \in \kappa} \|\sigma^n(x) - \sigma(x)\| \right) = 0.$$

The next lemma is a stability result for equation (3.3), it is proved in [37] by means of the tightness property. For the reader's convenience, we recall the main lines of the proof under our assumptions.

Lemma 4.2. Let assumptions (A.1).3.2.1, (A.4) et (A.5) be satisfied, then

(a) For any $q \geq 1$ and $(t, x) \in [0, T] \times \bar{D}$, we have

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s^{t,x,n} - X_s^{t,x}|^{2q} + \sup_{0 \leq s \leq T} |K_s^{t,x,n} - K_s^{t,x}|^{2q} \right] = 0 \quad (4.6)$$

where $(X^{t,x}, K^{t,x})$ is the solution of equation (3.3).

(b) For every $q \geq 1$ and every $(t, x) \in [0, T] \times \bar{D}$

$$\lim_{n \rightarrow +\infty} \mathbb{E} \sup_{s \in [0, T]} |k_s^{t,x,n} - k_s^{t,x}|^q = 0. \quad (4.7)$$

where $k^{t,x}$ is given by (3.4).

Proof. We suppress the superscript (t, x) from the notation of the processes $(X^{t,x,n}, K^{t,x,n}, X^{t,x}, K^{t,x})$. By Itô's formula, we have

$$\begin{aligned} & |X_s^n - X_s|^2 - 2 \int_t^s \langle X_r^n - X_r, d(K_r^n - K_r) \rangle \leq \\ & \int_t^s |X_r^n - X_r|^2 dr + 2 \int_t^s |b^n(X_r^n) - b(X_r^n)|^2 dr \\ & + 2 \int_t^s |b(X_r^n) - b(X_r)|^2 dr + 2 \int_t^s \|\sigma^n(X_r^n) - \sigma(X_r^n)\|^2 dr + 2 \int_t^s \|\sigma(X_r^n) - \sigma(X_r)\|^2 dr \\ & + 2 \int_t^s \langle X_r^n - X_r, (\sigma^n(X_r^n) - \sigma(X_r)) dW_r \rangle. \end{aligned}$$

Since

$$- \int_t^s \langle X_r^n - X_r, d(K_r^n - K_r) \rangle \geq 0$$

we use assumptions (A.1).3.2.1 and (A.4)(i) together with the Burkholder-Davis-Gundy inequality to obtain for all $q \geq 1$ and $N \in \mathbb{N}^*$

$$\begin{aligned} \mathbb{E} \sup_{t \leq s \leq T} |X_s^n - X_s|^{2q} & \leq C \mathbb{E} \int_t^T |X_r^n - X_r|^{2q} dr + C \sup_{|x| \leq N} |b^n(x) - b(x)|^{2q} \\ & + C \sup_{|x| \leq N} \|\sigma^n(x) - \sigma(x)\|^{2q} + C \mathbb{E} \int_t^T (1 + |X_r^n|^{2q}) 1_{\{|X_r^n| > N\}} dr, \end{aligned}$$

where C is a constant varying from line to another and independent of n and N . By Hölder's inequality, we infer

$$\begin{aligned} \mathbb{E} \sup_{t \leq s \leq T} |X_s^n - X_s|^{2q} &\leq C \mathbb{E} \int_t^T |X_r^n - X_r|^{2q} dr + C \sup_{|x| \leq N} |b^n(x) - b(x)|^{2q} \\ &\quad + C \sup_{|x| \leq N} \|\sigma^n(x) - \sigma(x)\|^{2q} + \frac{C}{N^{\frac{1}{2}}} \left(\mathbb{E} \int_t^T (1 + |X_r^n|^{4q}) dr \right)^{\frac{1}{2}} \left(\mathbb{E} \int_t^T |X_r^n| dr \right)^{\frac{1}{2}}. \end{aligned}$$

By estimates (4.5) and Gronwall's lemma, we get

$$\mathbb{E} \sup_{t \leq s \leq T} |X_s^n - X_s|^{2q} \leq C \sup_{|x| \leq N} |b^n(x) - b(x)|^{2q} + C \sup_{|x| \leq N} \|\sigma^n(x) - \sigma(x)\|^{2q} + \frac{C}{N^{\frac{1}{2}}}.$$

Given assumption (A.5), we successively send n and N to infinity, to obtain

$$\lim_{n \rightarrow +\infty} \mathbb{E} \sup_{t \leq s \leq T} |X_s^n - X_s|^{2q} = 0.$$

On the other hand, we use assumptions (A.1).3.2.1 and (A.4)(i) then we proceed as before to show that for all $q \geq 1$ and $N \in \mathbb{N}^*$ we have

$$\begin{aligned} \mathbb{E} \sup_{t \leq s \leq T} |K_s^n - K_s|^{2q} &\leq C \mathbb{E} \sup_{t \leq s \leq T} |X_s^n - X_s|^{2q} + C \sup_{|x| \leq N} |b^n(x) - b(x)|^{2q} \\ &\quad + C \sup_{|x| \leq N} \|\sigma^n(x) - \sigma(x)\|^{2q} + \frac{C}{N^{\frac{1}{2}}}. \end{aligned}$$

Letting n and N go to infinity, we conclude that

$$\lim_{n \rightarrow +\infty} \mathbb{E} \sup_{t \leq s \leq T} |K_s^n - K_s|^{2q} = 0.$$

This proves assertion (a). Arguing as in the proof of [11, Lemma 2.2] and using assertion (a) along with the representations of k^n and k in (4.3) and (3.4) respectively, we get the claim (b). \square

4.2.2 The BSDE associated to nonlinear Neumann PDE (4.1)

We consider a sequence of BSDEs associated with the Neumann boundary value (4.1)

$$\begin{aligned} Y_s^{t,x,n} &= g^n(X_T^{t,x,n}) + \int_s^T f^n(r, X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}) dr \\ &\quad + \int_s^T h^n(r, X_r^{t,x,n}, Y_r^{t,x,n}) dk_r^{t,x,n} - \int_s^T Z_r^{t,x,n} dW_r, \quad s \in [t, T]. \end{aligned} \quad (4.8)$$

Let us introduce the following assumptions:

(A.6) For any $n \in \mathbb{N}$, the functions f^n , h^n and g^n are continuous and there exist positive constants C , l_{f^n} and $\beta_n < 0$, $\mu_{f^n} \in \mathbb{R}$ such that for every $t \in [0, T]$ and every $(x, x', y, y', z, z') \in (\mathbb{R}^d)^2 \times (\mathbb{R}^m)^2 \times (\mathbb{R}^{m \times d'})^2$, we have

- (i) $\langle y - y', f^n(t, x, y, z) - f^n(t, x, y', z) \rangle \leq \mu_{f^n} |y - y'|^2$,
- (ii) $|f^n(t, x, y, z) - f^n(t, x, y, z')| \leq l_{f^n} \|z - z'\|$,
- (iii) $\sup_{n \geq 0} |f^n(t, x, y, z)| \leq C(1 + |y|)$,
- (iv) $\langle y - y', h^n(t, x, y) - h^n(t, x, y') \rangle \leq \beta_n |y - y'|^2$,
- (v) $\sup_{n \geq 0} |h^n(t, x, y)| \leq C(1 + |y|)$,
- (vi) $\sup_{n \geq 0} |g^n(x)| \leq C(1 + |x|)$.

For fixed n , under (A.6) the BSDE (4.8) has a unique solution $(Y_s^{t,x,n}, Z_s^{t,x,n})_{s \in [t, T]}$. As before, we extend the solution to $[0, T]$ by putting

$$Y_s^{t,x,n} := Y_t^{t,x,n} \quad \text{and} \quad Z_s^{t,x,n} = 0 \quad \text{for } s \in [0, t). \quad (4.9)$$

The next Lemma gives a priori estimates which will be employed in the sequel.

Lemma 4.3. Let assumptions (A.4)(i) and (A.6) hold. Let $\lambda_q := 1 \wedge (q - 1)$, $V_s^n := \mu_{f^n}^+ s + \frac{l_{f^n}^2}{\lambda_q} s$ and assume that $\sup_{n \geq 0} \exp(qV_T^n) < +\infty$, for every $q > 1$. Then,

$$\sup_{n \geq 0} \mathbb{E} \sup_{s \in [0, T]} |e^{V_s^n} Y_s^{t,x,n}|^q + \sup_{n \geq 0} \mathbb{E} \left(\int_0^T e^{2V_s^n} \|Z_s^{t,x,n}\|^2 ds \right)^{\frac{q}{2}} < +\infty. \quad (4.10)$$

Remark 4.4. The assumption $\sup_{n \geq 0} \exp(qV_T^n) < +\infty$ is obviously equivalent to the boundedness of $(\mu_{f^n}^+)_{n \geq 1}$ and $(l_{f^n})_{n \geq 1}$.

Proof. Under condition (A.6), it follows from [54, Proposition A.2] that for any $n \in \mathbb{N}$ and any $q > 1$, $\mathbb{E} \sup_{s \in [0, T]} |e^{V_s^n} Y_s^{t,x,n}|^q < +\infty$. In the other hand, we use [69, Proposition 6.80, page 603], to get the following estimates for $(Y^{t,x,n}, Z^{t,x,n})$: for each $q > 1$ there exists a constant C_q , depending only on q , such that

$$\begin{aligned} \mathbb{E} \sup_{s \in [0, T]} |e^{V_s^n} Y_s^{t,x,n}|^q + \mathbb{E} \left(\int_0^T e^{2V_s^n} \|Z_s^{t,x,n}\|^2 ds \right)^{\frac{q}{2}} &\leq C_q \mathbb{E} |e^{V_T^n} g^n(X_T^{t,x,n})|^q \\ &+ \mathbb{E} \left(\int_0^T e^{V_s^n} [|f^n(s, X_s^{t,x,n}, 0, 0)| + |h^n(s, X_s^{t,x,n}, 0)|] d(s + k_s^{t,x,n}) \right)^q. \end{aligned}$$

Using assumptions (A.6)(iii), (v) and (vi) we get

$$\begin{aligned} \mathbb{E} \sup_{s \in [0, T]} |e^{V_s^n} Y_s^{t,x,n}|^q + \mathbb{E} \left(\int_0^T e^{2V_s^n} \|Z_s^{t,x,n}\|^2 ds \right)^{\frac{q}{2}} &\leq C_q e^{qV_T^n} \mathbb{E} (1 + |X_T^{t,x,n}|^q) \\ &+ C_q e^{qV_T^n} \left(T^q + T^{2q} + \mathbb{E} \left((k_T^{t,x,n})^q + (k_T^{t,x,n})^{2q} + \sup_{s \in [0, T]} |X_s^{t,x,n}|^{2q} \right) \right). \end{aligned}$$

Taking into account the estimates (4.5), we get the result. \square

The following proposition guarantees the continuity of the viscosity solution to the system (4.1). Its proof can be found in [70, Theorem 3.1].

Proposition 4.5. Assume that (A.4) and (A.6) are satisfied. Then, for every $n \in \mathbb{N}$, the map $(t, x) \in [0, T] \times \bar{D} \mapsto Y_t^{t,x,n}$ is continuous.

4.3 Main result

To prove the convergence of solutions of the nonlinear Neumann problem (4.1), we consider the following conditions:

(A.7): For any compact (κ, κ') of $\mathbb{R}^d \times \mathbb{R}^m$ it holds that

$$\lim_{n \rightarrow +\infty} \sup_{r \in [0, T]} \sup_{x \in \kappa, y \in \kappa', z \in \mathbb{R}^m \times d'} |f^n(r, x, y, z) - f(r, x, y, z)| = 0$$

and

$$\lim_{n \rightarrow +\infty} \left[\sup_{x \in \kappa} |g^n(x) - g(x)| + \sup_{r \in [0, T]} \sup_{x \in \kappa, y \in \kappa'} |h^n(r, x, y) - h(r, x, y)| \right] = 0.$$

To make sense to the viscosity solution of Neumann problem (4.1), we need to assume

(A.8): f_i^n , the i -th coordinate of f^n , depends only on the i -th row of the matrix z .

The main result of this section is

Theorem 4.6. Let assumptions (A.1).3.2.1, (A.2).3.2.3, (A.3).3.3.2, (A.4), (A.5), (A.6), (A.7) et (A.8) be satisfied. Assume moreover that the sequence (V^n) satisfies $\sup_{n \geq 0} \exp(qV_T^n) < +\infty$ for each $q > 1$. Then, there exist a sequence of continuous functions (u^n) and a continuous function u such that for any n , u^n and u are defined on $[0, T] \times \bar{D}$ with values in \mathbb{R}^m and satisfy

- (i) u^n is a viscosity solution of (4.1),
- (ii) u is a viscosity solution of the nonlinear Neumann problem (3.23),
- (iii) for all $(t, x) \in [0, T] \times \bar{D}$, $u^n(t, x)$ converges to $u(t, x)$.

Proof. Let $(Y^{t,x}, Z^{t,x})$ and $(Y^{t,x,n}, Z^{t,x,n})$ be respectively the unique solutions of equations (3.12) and (4.8). We define $u(t, x) := Y_t^{t,x}$ and $u^n(t, x) := Y_t^{t,x,n}$. From proposition 3.12, u is continuous and by Proposition 4.5, u^n is continuous in (t, x) for each n . Thanks to [71, Theorem 4.3], u and u^n are respectively viscosity solutions of the systems (3.23) and (4.1). In order to prove the convergence of u^n to u , we shall show that the sequence of processes $Y^{t,x,n}$ converges to $Y^{t,x}$. Throughout the proof, we suppress the superscript (t, x) and C will denote a positive constant changing from one line to another but does not depend on n . We put

$$\begin{aligned} d\mathcal{K}_r^n &:= [f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r, Y_r, Z_r)] dr \\ &\quad + h^n(r, X_r^n, Y_r^n) dk_r^n - h(r, X_r, Y_r) dk_r. \end{aligned}$$

The idea consists in applying Lemma 3.4 to the BSDE

$$Y_s^n - Y_s = g^n(X_T^n) - g(X_T) + \int_s^T d\mathcal{K}_r^n - \int_s^T (Z_r^n - Z_r) dW_r.$$

Using (A.2).3.2.3(i)-(ii), we have, for every $0 \leq t \leq s \leq T$

$$\begin{aligned} & \int_s^T \langle Y_r^n - Y_r, f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r, Y_r, Z_r) \rangle dr \\ & \leq \int_s^T (l_f^2 + \mu_f) |Y_r^n - Y_r|^2 dr + \frac{1}{4} \|Z_r^n - Z_r\|^2 dr \\ & \quad + \int_s^T |Y_r^n - Y_r| |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)| dr \\ & \quad + \int_s^T |Y_r^n - Y_r| |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr. \end{aligned}$$

On the other hand, thanks to assumption (A.2).3.2.3(iv) we obtain

$$\begin{aligned} & \int_s^T \langle Y_r^n - Y_r, h^n(r, X_r^n, Y_r^n) dk_r^n - h(r, X_r, Y_r) dk_r \rangle \\ & \leq \int_s^T |Y_r^n - Y_r| |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)| dk_r^n \\ & \quad + \int_s^T |Y_r^n - Y_r| |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \\ & \quad + \int_s^T \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r). \end{aligned}$$

By the foregoing, it holds that for $\lambda = (l_f^2 + \mu_f) \vee l_f^2$

$$\langle Y_r^n - Y_r, d\mathcal{K}_r^n \rangle \leq \frac{1}{4} \|Z_r^n - Z_r\|^2 dr + \lambda |Y_r^n - Y_r|^2 dr + |Y_r^n - Y_r| dL_r^n + dR_r^n,$$

where L^n and R^n are given by

$$\begin{aligned} dL_r^n & := |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)| dr \\ & \quad + |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)| dr + |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)| dk_r^n \\ & \quad + |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)| dk_r^n \end{aligned} \quad (4.11)$$

and

$$dR_r^n := \langle Y_r^n - Y_r, h(r, X_r, Y_r) \rangle (dk_r^n - dk_r). \quad (4.12)$$

Using Lemma 3.4, it follows that there exist positive constants C_1, C_2 and C_3 such that

$$\begin{aligned} & \mathbb{E} \left(\sup_{r \in [0, T]} e^{2\lambda r} |Y_r^n - Y_r|^2 \right) + \mathbb{E} \left(\int_0^T e^{2\lambda r} \|Z_r^n - Z_r\|^2 dr \right) \\ & \leq C_1 \mathbb{E} e^{2\lambda T} |g^n(X_T^n) - g(X_T)|^2 + C_2 \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 + C_3 \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n. \end{aligned} \quad (4.13)$$

It remains to show that the right-hand side of (4.13) converges to zero as n goes to infinity. For the first term, we use the continuity of g , the convergence of g^n to g and X^n to X along with estimate (4.5) to get

$$\lim_{n \rightarrow +\infty} \mathbb{E} |g^n(X_T^n) - g(X_T)|^2 = 0.$$

For the third term of the right side of (4.13), arguing as in the proof of Lemma 3.10, one can show that

$$\limsup_{n \rightarrow \infty} \mathbb{E} \sup_{s \in [0, T]} \int_s^T e^{2\lambda r} dR_r^n \leq 0.$$

Concerning the second term, we have

$$\begin{aligned} \mathbb{E} \left(\int_0^T e^{\lambda r} dL_r^n \right)^2 &\leq 4T e^{2\lambda T} \mathbb{E} \left(\int_0^T |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right) \\ &\quad + 4T e^{2\lambda T} \mathbb{E} \left(\int_0^T |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)|^2 dr \right) \\ &\quad + 4e^{2\lambda T} \mathbb{E} k_T^n \mathbb{E} \left(\int_0^T |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)|^2 dk_r^n \right) \\ &\quad + 4e^{2\lambda T} \mathbb{E} k_T^n \mathbb{E} \left(\int_0^T |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)|^2 dk_r^n \right). \end{aligned}$$

We denote

$$\begin{aligned} J_1^n &:= \mathbb{E} \left(\int_0^T |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right) \\ J_2^n &:= \mathbb{E} \left(\int_0^T |f(r, X_r^n, Y_r, Z_r) - f(r, X_r, Y_r, Z_r)|^2 dr \right) \\ J_3^n &:= \mathbb{E} \left(\int_0^T |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)|^2 dk_r^n \right) \\ J_4^n &:= \mathbb{E} \left(\int_0^T |h(r, X_r^n, Y_r) - h(r, X_r, Y_r)|^2 dk_r^n \right). \end{aligned}$$

The convergence of J_2^n and J_4^n to zero can be handled as the terms I_1^n and I_2^n in (3.19) and (3.20) respectively. We shall show that J_1^n tends to 0 as n tends to $+\infty$. Let $N \in \mathbb{N}^*$, we use assumptions (A.2).3.2.3(iii) and (A.6)(iii) together with estimates (4.5) and (4.10) to obtain

$$\begin{aligned} J_1^n &= \mathbb{E} \left(\int_0^T |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)|^2 dr \right) \\ &= \mathbb{E} \left(\int_0^T |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)|^2 1_{\{|X_r^n| + |Y_r^n| \leq N\}} dr \right) \\ &\quad + \mathbb{E} \left(\int_0^T |f^n(r, X_r^n, Y_r^n, Z_r^n) - f(r, X_r^n, Y_r^n, Z_r^n)|^2 1_{\{|X_r^n| + |Y_r^n| > N\}} dr \right) \\ &\leq \int_0^T \sup_{|x|, |y| \leq N, z \in \mathbb{R}^m \times d'} |f^n(r, x, y, z) - f(r, x, y, z)|^2 dr \\ &\quad + \frac{C}{N^{\frac{1}{2}}} \left(\mathbb{E} \int_0^T (1 + |Y_r^n|^2)^2 dr \right)^{\frac{1}{2}} \left(\mathbb{E} \int_0^T |X_r^n| + |Y_r^n| dr \right)^{\frac{1}{2}} \\ &\leq T \sup_{r \in [0, T]} \sup_{|x|, |y| \leq N, z \in \mathbb{R}^m \times d'} |f^n(r, x, y, z) - f(r, x, y, z)|^2 + \frac{C}{N^{\frac{1}{2}}}. \end{aligned}$$

Passing successively to the limit in n and N and using assumption (A.7), we obtain $\lim_{n \rightarrow +\infty} J_1^n = 0$. We now show that J_3^n tends to 0 as n tends to ∞ . Let $N \in \mathbb{N}^*$, we use assumptions (A.2).3.2.3(v) and (A.6)(v), estimates (4.5) and (4.10) to find

$$\begin{aligned}
J_3^n &= \mathbb{E} \left(\int_0^T |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)|^2 dk_r^n \right) \\
&= \mathbb{E} \left(\int_0^T |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)|^2 1_{\{|X_r^n| + |Y_r^n| \leq N\}} dk_r^n \right) \\
&\quad + \mathbb{E} \left(\int_0^T |h^n(r, X_r^n, Y_r^n) - h(r, X_r^n, Y_r^n)|^2 1_{\{|X_r^n| + |Y_r^n| > N\}} dk_r^n \right) \\
&\leq \mathbb{E} \left(\int_0^T \sup_{|x|, |y| \leq N} |h^n(r, x, y) - h(r, x, y)|^2 dk_r^n \right) \\
&\quad + \frac{C}{N^{\frac{1}{2}}} \left[\mathbb{E} \left(\int_0^T (1 + |Y_r^n|^4) dk_r^n \right) \right]^{\frac{1}{2}} \left[\mathbb{E} \left(\int_0^T |X_r^n| + |Y_r^n| dk_r^n \right) \right]^{\frac{1}{2}} \\
&\leq \mathbb{E} (k_T^n) \sup_{r \in [0, T]} \sup_{|x|, |y| \leq N} |h^n(r, x, y) - h(r, x, y)|^2 + \frac{C}{N^{\frac{1}{2}}}.
\end{aligned}$$

Passing successively to the limits in n and N , we get that J_3^n tends to 0 as n tends to $+\infty$. Theorem 4.6 is proved. \square

Part II

TRANSPORTATION COST INEQUALITIES

Chapter 5

Quadratic transportation inequalities for SDEs with measurable drift

Contents

5.1	Introduction	81
5.2	Main results	83
5.3	Proofs of the main results	84
5.4	Examples	93

Abstract

Let X be the solution of a stochastic differential equation in Euclidean space driven by standard Brownian motion, with measurable drift and Sobolev diffusion coefficient. In our main result we show that when the drift is measurable and the diffusion coefficient belongs to an appropriate Sobolev space, the law of X satisfies Talagrand's inequality with respect to the uniform distance.

5.1 Introduction

The concentration of measure phenomenon, initiated by [58] quantifies the deviation of a (Lipschitz continuous) function f of a random vector; $f(\xi^1, \dots, \xi^n)$ to its mean. It can be seen as a vast improvement of the classical Chebyshev inequality in large deviation theory, see e.g. the texts by [46] and [76] for modern presentations. [51] showed that the transportation inequality first established by [83] can be efficiently used to explain the concentration of measure phenomenon. The quadratic transportation inequality (or Talagrand inequality) compares the Wasserstein distance and the Kullback-Leibler divergence: Given a constant C , a probability measure μ is said to satisfy Talagrand's $T_2(C)$ in-

equality (or quadratic transportation inequality) if

$$\mathcal{W}_2(\mu, \nu) \leq \sqrt{CH(\nu|\mu)} \text{ for all probability measures } \nu.$$

We define the (second order) Wasserstein distance and the Kullback-Leibler divergence respectively by

$$\mathcal{W}_2(\mu, \nu) := \left(\inf_{\pi} \int_{\Omega \times \Omega} \|\omega - \eta\|_{\infty}^2 \pi(d\omega, d\eta) \right)^{1/2} \quad \text{and} \quad H(\nu|\mu) := \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu$$

where the infimum is taken over all, couplings π of μ and ν , that is, probability measures on $\Omega \times \Omega$ such that if $(\zeta, \eta) \sim \pi$ then $\zeta \sim \mu$ and $\eta \sim \nu$, and we used the convention $d\nu/d\mu = +\infty$ if ν is not absolutely continuous w.r.t. μ . The transportation inequality has since found numerous applications, for instance to isoperimetric problems, to randomized algorithms [26], or to quantitative finance [44, 85] and to various problems of probability in high dimensions [22, 8, 53]. Transportation inequalities are also related to other functional inequalities as Poincaré inequality, log-Sobolev inequality, inf-convolution and hypercontractivity, see [10, 61]. [83] proved a quadratic transportation inequality for the multidimensional Gaussian distribution with optimal constant $C = 2$. Using stochastic analysis techniques, notably Girsanov's theorem, Talagrand's work was then extended to Wiener measure on the path space by [31]. The case of SDEs was first analyzed by [25] using a technique based on Girsanov's transform that we also employ here. Their results gave rise to an interesting literature, including the papers [16, 62, 63, 86] on SDEs driven by Brownian motion and [73, 78] on SDEs driven by abstract Gaussian noise. Note that almost all the aforementioned works on SDEs assume that the coefficients are Lipschitz-continuous or satisfy a dissipative condition, with the exception of [16, 63] who assume a contraction condition of the form $(b(t, x) - b(t, y)) \cdot (x - y) \leq K \|x - y\|^2$. It is worth noting that [63] additionally deal with equations with reflections. The extension of [25] to diffusions with non-smooth coefficients was started by [8] where it is proved that $T_2(C)$ holds for *one-dimensional* equations, if b is measurable in space and *differentiable* in time and σ *Lipschitz* continuous. The idea of [8] is based on a transformation that is tailor-made for the one-dimensional case. The present paper deals with the multidimensional case and further weakens the regularity requirements imposed in [8]. In this case we use Zvonkin's transformation which is well-known in SDE theory, see for instance [3, 87, 93]. Note that considering multidimensional equations is, for instance, fundamental for applications to concentration and asymptotic results on *interacting particle systems*, see e.g. [22, Section 5] and the various examples we give in the final section. The proof of Theorem 5.2 is given in the next section, and section 5.4 presents some examples.

5.2 Main results

Our objective is to investigate transportation inequalities for stochastic differential equations of the form

$$X(t) = x + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dB(s) \quad \text{for } t \in [0, T], x \in \mathbb{R}^d \quad (5.1)$$

under minimal regularity assumptions on the coefficients $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$. Note that this equation is understood in the \mathbb{P} -almost sure sense on the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ of the d -dimensional Brownian motion denoted by B and equipped with the \mathbb{P} -completion of the raw filtration $\sigma(B(s), s \leq t)$ generated by B . That is $\Omega = C([0, T], \mathbb{R}^d)$ endowed with the supremum norm, and $B(t, \omega) = \omega(t)$. Further denote by $\mathcal{P}(\Omega)$ the set of all Borel probability measures on Ω . In all of the article we assume that the diffusion coefficient σ satisfies the uniform ellipticity condition

$$\xi^\top \sigma(t, x) \xi \geq \Lambda_\sigma |\xi|^2 \quad \text{for all } (t, x, \xi) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \text{ for some } \Lambda_\sigma > 0. \quad (5.2)$$

To state our main result, let us define the following functional spaces. For $p \geq 1$ denote by $L_{\text{loc}}^p([0, T]) := L_{\text{loc}}^p([0, T] \times \mathbb{R}^d)$ the (Lebesgue) space of classes of locally integrable functions and for every $m_1, m_2 \in \mathbb{N}$, let $W_p^{m_1, m_2}([0, T]) := W_p^{m_1, m_2}([0, T] \times \mathbb{R}^d)$ be the usual Sobolev space of weakly differentiable functions $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$\|f\|_{W_p^{m_1, m_2}} := \sum_{i=0}^{m_1} \|\partial_t^i f\|_{L^p} + \sum_{|\alpha| \leq m_2} \|\partial_x^\alpha f\|_{L^p} < \infty$$

where α is a d -dimensional multiindex, i.e. a d -tuple $\alpha = (\alpha_1, \dots, \alpha_d)$ and $|\alpha| := \alpha_1 + \dots + \alpha_d$. Denote by $W_{p, \text{loc}}^{m_1, m_2}([0, T])$ the space of weakly differentiable functions $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ such that

$$\|f\|_{L_{\text{loc}}^p} + \sum_{|\alpha| \leq m_1} \|\partial_t^\alpha f\|_{L_{\text{loc}}^p} + \sum_{|\alpha| \leq m_2} \|\partial_x^\alpha f\|_{L_{\text{loc}}^p} < \infty.$$

Further let $L_q^p([0, T]) := L^q([0, T], L^p(\mathbb{R}^d))$ be the space of (classes of) measurable functions $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$\|f\|_{L_q^p} := \left(\int_0^T \left(\int_{\mathbb{R}^d} |f(s, x)|^p dx \right)^{q/p} ds \right)^{1/q} < \infty.$$

The aim of this note is to prove the following two results:

Theorem 5.1. Let $\sigma \in W_{2(d+1), \text{loc}}^{0,1}([0, T]) \cap L^\infty([0, T] \times \mathbb{R}^d)$ and $b \in W_{(d+1), \text{loc}}^{0,1}([0, T]) \cap L^\infty([0, T] \times \mathbb{R}^d)$. Assume that σ satisfies (5.2). Then, equation (5.1) admits a unique strong solution X , and

$$\text{the law } \mu_x \text{ of } X \text{ satisfies } T_2(C)$$

with constant

$$C := \inf_{0 < \varepsilon < 1} 2 \|\sigma\|_\infty^2 \exp \left(7T \frac{32 + \varepsilon}{\varepsilon(1 - \varepsilon)} \right) \frac{1}{1 - \varepsilon}.$$

This result gives the transportation inequality for SDEs with coefficients in some Sobolev spaces. It moreover constitute one essential argument in the proof of the next result where we combine it with some gradient estimates for solutions of PDEs established by [41] to establish transportation inequality for SDEs with measurable drifts.

Theorem 5.2. Assume that σ satisfies (5.2) and that one of the following conditions is satisfied:

- (A) $\sigma, b \in L^\infty([0, T] \times \mathbb{R}^d)$, the function σ is continuous in (t, x) and belongs to $W_{2(d+1), \text{loc}}^{0,1}([0, T])$.
- (B) $\sigma \in W_{2(d+1), \text{loc}}^{0,1}([0, T]) \cap L^\infty([0, T] \times \mathbb{R}^d)$, σ is uniformly continuous in x . The function b satisfies $b \in L_q^p([0, T])$ for some p, q such that $d/p + 2/q < 1$, $2(d+1) \leq p$ and $q > 2$.

Then there exists $T > 0$ small enough so that equation (5.1) admits a unique strong solution X with continuous paths and

$$\text{the law } \mu_x \text{ of } X \text{ satisfies } T_2(C)$$

for some constant C depending on the data, namely $\|b\|_{L_q^p}, \|\sigma\|_\infty, T, x, d, p$ and q .

Since b is merely assumed measurable, Theorem 5.2 gives transportation inequality for *singular* SDEs as $dX(t) = \text{sgn}(X(t))dt + dB(t)$, or for "regime switching" models as

$$dX(t) = \{b_1(t, X(t))1_A(X(t)) + b_2(t, X(t))1_{A^c}(X(t))\}dt + \sigma(t, X(t))dB(t)$$

with A a measurable subset of \mathbb{R}^d . Other examples are discussed at the end of the article.

5.3 Proofs of the main results

We explain our strategy to prove Theorem 5.2. Roughly speaking, in both situations (A) and (B), the path remains the same and it consists in reducing equation (5.1) to one without drift. When assumption (A) is satisfied, we follow the path developed by [3] in the proof of strong uniqueness for SDEs with measurable drift and a locally Sobolev diffusion coefficient: We first derive the transportation inequality for SDEs with coefficients belonging to some Sobolev spaces, this is the goal of the next subsection. After that, we use Zvonkin's transformation [93] to deal with the case where the drift is only measurable. When assumption (B) is satisfied, i.e. the case where the drift belongs to some $L_q^p([0, T])$ -space, we also remove the drift but the proof is more elaborated: We slightly modify the method of [41] in order to derive suitable gradient estimates for singular second order parabolic PDEs. Thus, along with gradient estimates for solutions of singular PDEs, Theorem 5.1 is an essential building block for the proof of the main result.

Technical lemmas

The aim of this section is to present two lemmas that will be used in the proof of Theorem 5.1. The proof of the first lemma can be found in Step 1 of the proof of [25, Theorem 5.6].

Lemma 5.3. Let $\nu \in \mathcal{P}(\Omega)$ be such that $\nu \ll \mu_x$ and $H(\nu|\mu_x) < \infty$, and let X be the solution of (5.1). Then, the probability measure \mathbb{Q} given by

$$d\mathbb{Q} := \frac{d\nu}{d\mu_x}(X)d\mathbb{P}$$

satisfies

$$H(\nu|\mu_x) = \mathbb{E}_{\mathbb{Q}} \left[\frac{1}{2} \int_0^T |q(s)|^2 ds \right] \quad (5.3)$$

for some progressively measurable, square integrable process q such that $\tilde{B} := B - \int_0^\cdot q(s)ds$ is a \mathbb{Q} -Brownian motion.

The next lemma is a crucial point in the proof of Theorem 5.1, it provides the argument allowing to go around the uniform Lipschitz (or dissipativity) condition usually imposed on the coefficients. The main idea probably originated from [87] and further developed in [3] in order to establish the pathwise uniqueness of multidimensional SDEs when the diffusion coefficient is locally in some Sobolev space, that is when the usual Gronwall's lemma can not be directly used.

Lemma 5.4. Let X_1, X_2 be two square integrable semi-martingales of the form

$$X_i(t) = x_i + \int_0^t \beta_i(u)du + \int_0^t \alpha_i(u)dB(u), \quad i = 1, 2$$

where β_i, α_i are bounded adapted processes with $\xi^\top \alpha_i(t)\xi \geq \Lambda|\xi|^2$ for all $(t, \xi) \in [0, T] \times \mathbb{R}^d$. For $N > 0$, we put

$$\tau^N := \inf\{t > 0 : |X_1(t)| > N \text{ or } |X_2(t)| > N\} \wedge T, \quad N \in \mathbb{N}.$$

(i) Let $f \in W_{2(d+1), \text{loc}}^{0,1}$ and $A_f^N(s)$ be the increasing process defined by

$$A_f^N(t) := \int_0^{t \wedge \tau^N} \int_0^1 |\partial_x f(s, \lambda X_1(s) + (1 - \lambda)X_2(s))|^2 d\lambda ds.$$

Then, for every stopping time τ with values in $[0, T]$ it holds

$$\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |f(s, X_1(s)) - f(s, X_2(s))|^2 ds \right] \leq 6\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 dA_f^N(s) \right]$$

(ii) Let $g \in W_{(d+1), \text{loc}}^{0,1}$ and $A_g^N(s)$ be defined as follows

$$A_g^N(t) := \int_0^{t \wedge \tau^N} \int_0^1 |\partial_x g(s, \lambda X_1(s) + (1 - \lambda)X_2(s))| d\lambda ds.$$

Then, for every stopping time τ with values in $[0, T]$ it holds

$$\begin{aligned} \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)| |g(s, X_1(s)) - g(s, X_2(s))| ds \right] \\ \leq \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 dA_g^N(s) \right]. \end{aligned}$$

Proof. It must be noted that, in contrast to the Lipschitz case, the difference ratios of the functions f and g are not bounded in our situation, not even locally. We will see that, since the coefficients α_i of X_i are non degenerate, Krylov's estimate ([42, Theorem 2.2.4]) allows us to avoid this difficulty. Using Krylov's estimate, one can show that the processes $A_f^N(s)$ and $A_g^N(s)$ are well defined. We denote by K_N the ball $\{x \in \mathbb{R}^d, |x| \leq N\}$ and put $K := [0, T] \times K_N$. Let $f_n \in C^\infty(K)$ be such that,

$$\|f_n - f\|_{W_{2(d+1)}^{0,1}(K)} \rightarrow 0, \quad (5.4)$$

see e.g. [29, Theorem 5.3.2] for the existence of such a sequence. Using subsequently the classical inequality $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ and Krylov's estimate, we have that

$$\begin{aligned} I &:= \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |f(s, X_1(s)) - f(s, X_2(s))|^2 ds \right] \\ &\leq 3\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |f(s, X_1(s)) - f_n(s, X_1(s))|^2 ds \right] \\ &\quad + 3\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |f(s, X_2(s)) - f_n(s, X_2(s))|^2 ds \right] \\ &\quad + 3\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |f_n(s, X_1(s)) - f_n(s, X_2(s))|^2 ds \right] \\ &\leq C_{T,N,d} \|f - f_n\|_{L^2(d+1)(K)} \\ &\quad + 3\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 \int_0^1 |\partial_x f_n(s, Z(s, \lambda))|^2 d\lambda ds \right] \end{aligned} \quad (5.5)$$

with $Z(s, \lambda) := \lambda X_1(s) + (1 - \lambda)X_2(s)$, and where $C_{T,N,d}$ is a positive constant which depends on T, N and d . Using Fubini's theorem and Krylov's estimate

once again, we get

$$\begin{aligned}
& \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 \int_0^1 |\partial_x f_n(s, Z(s, \lambda))|^2 d\lambda ds \right] \\
& \leq \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 \right. \\
& \quad \left. \times \int_0^1 2|\partial_x f_n(s, Z(s, \lambda)) - \partial_x f(s, Z(s, \lambda))|^2 + 2|\partial_x f(s, Z(s, \lambda))|^2 d\lambda ds \right] \\
& \leq 8N^2 \int_0^1 \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |\partial_x f_n(s, Z(s, \lambda)) - \partial_x f(s, Z(s, \lambda))|^2 ds \right] d\lambda \\
& \quad + 2\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 \int_0^1 |\partial_x f(s, Z(s, \lambda))|^2 d\lambda ds \right] \\
& \leq 8N^2 C_{T,N,d} \|\partial_x f_n - \partial_x f\|_{L^2(d+1)(K)} + 2\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 dA_f^N(s) \right].
\end{aligned}$$

Putting this together with the estimate (5.5) for I leads to

$$\begin{aligned}
I & \leq C_{T,N,d} \|f_n - f\|_{L^2(d+1)(K)} + 24N^2 C_{T,N,d} \|\partial_x f_n - \partial_x f\|_{L^2(d+1)(K)} \\
& \quad + 6\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 dA_f^N(s) \right] \\
& \leq C'_{T,N,d} \|f_n - f\|_{W_{2(d+1)}^{0,1}(K)} + 6\mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 dA_f^N(s) \right]
\end{aligned}$$

for some constant $C'_{T,N,d}$. Hold N fixed and take the limit as n goes to infinity in the last inequality then use (5.4) to obtain the first claim.

The second inequality is obtained similarly. Indeed, let (g_n) be a sequence of functions such that $g_n \in C^\infty(K)$ and

$$\|g_n - g\|_{W_{(d+1)}^{0,1}(K)} \rightarrow 0.$$

We have

$$\begin{aligned}
& \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)| |g(s, X_1(s)) - g(s, X_2(s))| ds \right] \\
& \leq \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)| \{ |g(s, X_1(s)) - g_n(s, X_1(s))| \right. \\
& \quad \left. + |g_n(s, X_1(s)) - g_n(s, X_2(s))| + |g(s, X_2(s)) - g_n(s, X_2(s))| \} ds \right] \\
& \leq 2N^2 C_{T,N,d} \|g - g_n\|_{L^{(d+1)}(K)} \\
& \quad + \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 \int_0^1 |\partial_x g_n(s, Z(s, \lambda))| d\lambda ds \right] \\
& \leq 2N^2 C_{T,N,d} \|g - g_n\|_{W_{(d+1)}^{0,1}(K)} \\
& \quad + \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 \int_0^1 |\partial_x g(s, Z(s, \lambda))| d\lambda ds \right] \\
& = 2N^2 C_{T,N,d} \|g - g_n\|_{W_{(d+1)}^{0,1}(K)} + \mathbb{E} \left[\int_0^{\tau \wedge \tau^N} |X_1(s) - X_2(s)|^2 dA_g^N(s) \right].
\end{aligned}$$

Again, hold N fixed and let n go to infinity to obtain the second claim. \square

Proof of Theorem 5.1

That (5.1) admits a unique strong solution follows from e.g. [3, Theorem 2.1]. Let $\nu \in \mathcal{P}(\Omega)$ be absolutely continuous with respect to μ_x . We can assume without loss of generality that $H(\nu|\mu_x) < \infty$. Let \mathbb{Q} and q be as in Lemma 5.3. Under the probability measure \mathbb{Q} , the SDE (5.1) takes the form

$$dX(t) = \sigma(t, X(t)) d\tilde{B}(t) + \{ \sigma(t, X(t)) q(t) + b(t, X(t)) \} dt, \quad \text{with } X(0) = x \quad (5.6)$$

and the law of X under \mathbb{Q} is ν . Furthermore, the SDE

$$dY(t) = \sigma(t, Y(t)) d\tilde{B}(t) + b(t, Y(t)) dt, \quad \text{with } Y(0) = x$$

admits a unique solution (see [3]) and the law of Y under \mathbb{Q} is μ_x . That is, (X, Y) under \mathbb{Q} is a coupling of (ν, μ_x) . Thus,

$$\mathcal{W}_2^2(\nu, \mu_x) \leq \mathbb{E}_{\mathbb{Q}} \left[\sup_{0 \leq t \leq T} |X(t) - Y(t)|^2 \right]. \quad (5.7)$$

We now estimate the right hand side above. By Itô's formula, we have

$$\begin{aligned}
|X(t) - Y(t)|^2 &= \int_0^t 2(X(s) - Y(s)) \sigma(s, X(s)) q(s) + |\sigma(s, X(s)) - \sigma(s, Y(s))|^2 ds \\
& \quad + \int_0^t 2(X(s) - Y(s)) (b(s, X(s)) - b(s, Y(s))) ds \\
& \quad + \int_0^t 2(X(s) - Y(s)) (\sigma(s, X(s)) - \sigma(s, Y(s))) d\tilde{B}(s), \quad (5.8)
\end{aligned}$$

where we simply denote by ab the inner product between two vectors a and b . The difficulty is to deal with the terms $\sigma(s, X(s)) - \sigma(s, Y(s))$ and $b(s, X(s)) - b(s, Y(s))$. This was done in Lemma 5.4. As in that lemma, we introduce the following random times: First consider the sequence of stopping times

$$\tau^N := \inf\{t > 0 : |X(t)| > N \text{ or } |Y(t)| > N\} \wedge T.$$

It is clear that $\tau^N \uparrow T$. For each λ in $[0, 1]$ and t in $[0, T]$, we put $Z(t, \lambda) := \lambda X(t) + (1 - \lambda)Y(t)$. For every $N \geq 0$ and $t \in [0, \infty)$, define

$$A^N(t) := \int_0^{t \wedge \tau^N} \int_0^1 \left(|\partial_x \sigma(s, Z(s, \lambda))|^2 + |\partial_x b(s, Z(s, \lambda))| \right) d\lambda ds \quad (5.9)$$

$$:= A_\sigma^N(t) + A_b^N(t), \quad (5.10)$$

where the (weak) derivative operator acts on the spacial variable. Notice that the function $t \mapsto A^N(t)$ is not necessarily strictly increasing. Thus, consider the process $k^N(t) := t + A^N(t)$. It is continuous, strictly increasing and satisfies $k^N(0) = 0$. Moreover, k^N maps $[0, \infty)$ onto itself. We denote by γ^N the unique inverse map of k^N . Observe that for each N , we have $\gamma_t^N \uparrow \infty$ as $t \uparrow \infty$. Using (5.8), Cauchy-Schwarz and Burkholder-Davis-Gundy inequalities, one can show that for each $t \in [0, \infty)$ it holds that

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}} \left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2 \right] \\ & \leq \mathbb{E}_{\mathbb{Q}} \left[\int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 ds + \|\sigma\|_\infty^2 \int_0^{\gamma_t^N \wedge \tau^N} |q(s)|^2 ds \right] \\ & \quad + \mathbb{E}_{\mathbb{Q}} \left[\int_0^{\gamma_t^N \wedge \tau^N} |\sigma(s, X(s)) - \sigma(s, Y(s))|^2 \right. \\ & \quad \left. + 2|X(s) - Y(s)| |b(s, X(s)) - b(s, Y(s))| ds \right] \\ & \quad + 2C_{BDG} \mathbb{E}_{\mathbb{Q}} \left[\left(\int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 |\sigma(s, X(s)) - \sigma(s, Y(s))|^2 ds \right)^{1/2} \right] \\ & := I_1 + I_2 + I_3, \end{aligned} \quad (5.11)$$

for a (universal) constant $C_{BDG} > 0$.

We shall estimate each of the terms I_2 and I_3 separately, since they contain the differences $\sigma(t, X(t)) - \sigma(t, Y(t))$ and $b(t, X(t)) - b(t, Y(t))$. To start with, notice that I_3 can be estimated in terms of I_2 . In fact, letting $\varepsilon > 0$, by the Young's inequality $2ab \leq \varepsilon a^2 + \frac{1}{\varepsilon} b^2$, we have

$$\begin{aligned} I_3 & \leq \varepsilon \mathbb{E}_{\mathbb{Q}} \left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2 \right] \\ & \quad + \frac{C_{BDG}^2}{\varepsilon} \mathbb{E}_{\mathbb{Q}} \left[\int_0^{\gamma_t^N \wedge \tau^N} |\sigma(s, X(s)) - \sigma(s, Y(s))|^2 ds \right] \\ & \leq \varepsilon \mathbb{E}_{\mathbb{Q}} \left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2 \right] + \frac{C_{BDG}^2}{\varepsilon} I_2. \end{aligned}$$

We subsequently use Lemma 5.4 (with $f := \sigma$ and $g := b$) and identity (5.9) to get

$$\begin{aligned} I_2 &\leq 6\mathbb{E}\left[\int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 dA_\sigma^N(s) + \int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 dA_b^N(s)\right] \\ &= 6\mathbb{E}\left[\int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 dA^N(s)\right]. \end{aligned}$$

Coming back to (5.11), since $k^N(t) := t + A^N(t)$, we have

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}}\left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2\right] \leq \\ &\mathbb{E}_{\mathbb{Q}}\left[\int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 dk^N(s) + \|\sigma\|_\infty^2 \int_0^{\gamma_t^N \wedge \tau^N} |q(s)|^2 ds\right] \\ &\quad + \varepsilon \mathbb{E}_{\mathbb{Q}}\left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2\right] \\ &\quad + 6 \frac{C_{BDG}^2 + \varepsilon}{\varepsilon} \mathbb{E}_{\mathbb{Q}}\left[\int_0^{\gamma_t^N \wedge \tau^N} |X(s) - Y(s)|^2 dk^N(s)\right]. \end{aligned}$$

The time change $s \equiv \gamma_r^N$ (which is equivalent to $r = k^N(s)$) gives

$$\begin{aligned} (1 - \varepsilon) \mathbb{E}_{\mathbb{Q}}\left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2\right] &\leq \\ \|\sigma\|_\infty^2 \mathbb{E}_{\mathbb{Q}}\left[\int_0^{\gamma_t^N \wedge \tau^N} |q(s)|^2 ds\right] + 7 \frac{C_{BDG}^2 + \varepsilon}{\varepsilon} \mathbb{E}_{\mathbb{Q}}\left[\int_0^{t \wedge \tau^N} |X(\gamma_r^N) - Y(\gamma_r^N)|^2 dr\right] \end{aligned}$$

and using the fact that the function γ^N is increasing, we then have

$$\begin{aligned} (1 - \varepsilon) \mathbb{E}_{\mathbb{Q}}\left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2\right] &\leq \\ \|\sigma\|_\infty^2 \mathbb{E}_{\mathbb{Q}}\left[\int_0^{\gamma_t^N \wedge \tau^N} |q(s)|^2 ds\right] + 7 \frac{C_{BDG}^2 + \varepsilon}{\varepsilon} \mathbb{E}_{\mathbb{Q}}\left[\int_0^t \sup_{s \in [0, \gamma_r^N \wedge \tau^N]} |X(s) - Y(s)|^2 dr\right]. \end{aligned}$$

Choosing $\varepsilon < 1$ and using Gronwall's lemma, we get

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}}\left[\sup_{s \in [0, \gamma_t^N \wedge \tau^N]} |X(s) - Y(s)|^2\right] \leq \\ &\quad \frac{1}{1 - \varepsilon} \|\sigma\|_\infty^2 \mathbb{E}_{\mathbb{Q}}\left[\int_0^T |q(s)|^2 ds\right] \exp\left(7 \frac{C_{BDG}^2 + \varepsilon}{\varepsilon(1 - \varepsilon)} T\right) \end{aligned}$$

where we also used the fact that $\tau^N \wedge \gamma_t^N \leq T$. Letting successively t then N go to infinity, it follows by Fatou's lemma, $\gamma_t^N \uparrow \infty$, $\tau^N \uparrow T$ and the continuity of X and Y that

$$\mathbb{E}_{\mathbb{Q}}\left[\sup_{s \in [0, T]} |X(s) - Y(s)|^2\right] \leq \exp\left(7 \frac{C_{BDG}^2 + \varepsilon}{\varepsilon(1 - \varepsilon)} T\right) \frac{1}{1 - \varepsilon} \|\sigma\|_\infty^2 \mathbb{E}_{\mathbb{Q}}\left[\int_0^T |q(s)|^2 ds\right].$$

Hence, we conclude from (5.3) and (5.7) that

$$\mathcal{W}_2^2(\mu_x, \nu) \leq 2 \exp\left(7 \frac{C_{BDG}^2 + \varepsilon}{\varepsilon(1-\varepsilon)} T\right) \frac{1}{1-\varepsilon} \|\sigma\|_\infty^2 H(\nu|\mu_x).$$

This concludes the proof. \square

Proof of Theorem 5.2

(A): We start by the case when assumption (A) is satisfied. By [3, Theorem 3.1] equation (5.1) admits a unique strong solution. As in [3], the idea consists in using Zvonkin's transform in order to transform equation (5.1) into an SDE without drift then using Theorem 5.1 to conclude. In the rest of the paper, we denote by \mathcal{L} the differential operator defined by

$$\begin{aligned} \mathcal{L}\phi &:= b\partial_x\phi + \frac{1}{2}\text{tr}(\sigma^\top\sigma\partial_{xx}\phi) \\ &= \sum_{j=1}^d b_j\partial_{x_j}\phi + \frac{1}{2}\sum_{ij=1}^d a_{ij}\partial_{x_i x_j}\phi, \end{aligned}$$

where $(a_{ij})_{i,j=1,\dots,d}$ is the matrix $\sigma^\top\sigma$. According to [93, Theorem 2], there exists a $T > 0$ small enough such that the PDE

$$\begin{cases} \partial_t\varphi + \mathcal{L}\varphi = 0 \\ \varphi(T, x) = x \end{cases}$$

admits a unique solution $\varphi \in W_{p,\text{loc}}^{1,2}([0, T])$ such that: for every t , the function $x \mapsto \varphi(t, x)$ is one-to-one from \mathbb{R}^d onto \mathbb{R}^d , both φ and its inverse ψ belong to $W_{p,\text{loc}}^{1,2}([0, T])$ for each $p > 1$, both $\varphi(t, \cdot)$ as well as its inverse $\psi(t, \cdot)$ are Lipschitz continuous, with Lipschitz constants depending on $d, T, \|b\|_\infty$ and $\|\sigma\|_\infty$.

Applying Itô-Krylov's formula, see [42, Theorem 2.10.1] to $\varphi(t, X(t)) := Y(t)$, it follows that Y satisfies the drift-less SDE

$$Y(t) = Y(0) + \int_0^t \tilde{\sigma}(s, Y(s))dB(s)$$

with $\tilde{\sigma}(t, y) := (\sigma^\top\partial_x\varphi)(t, \psi(t, y))$. Since σ belongs to $W_{2(d+1),\text{loc}}^{0,1}([0, T])$, and φ belongs to $W_{p,\text{loc}}^{1,2}([0, T])$ for each $p > 1$ and both φ and ψ are Lipschitz, it follows that $\tilde{\sigma} \in W_{2(d+1),\text{loc}}^{0,1}([0, T])$, more details can be found in, [3, page 14, proof of Theorem 3.1]. Hence, by Lemma 5.1, the law μ_y of Y satisfies $T_2(C)$, where C is the constant in Theorem 5.1. But $X(t) = \psi(t, Y(t))$ and ψ is Lipschitz continuous. Thus, the result follows from [25, Lemma 2.1].

(B): We now assume that condition (B) is fulfilled. We need to introduce the following Banach spaces: For every $k \geq 0$ and $m \geq 1$, let $H_m^k := (I - \Delta)^{-k/2}L^m$ be the usual space of Bessel potentials on \mathbb{R}^d and denote

$$\mathbb{H}_p^{2,q}([0, T]) := L^q([0, T], H_p^2),$$

and

$$H_p^{2,q}([0, T]) := \{u : [0, T] \rightarrow H_p^2 \text{ and } \partial_t u \in L_p^q([0, T])\}.$$

The space H_p^2 is equipped with the norm

$$\|u\|_{H_p^2} := \|(I - \Delta)u\|_{L^p}$$

making it isomorphic to the Sobolev space $W_p^2(\mathbb{R}^d)$.

Under assumption (B), the existence and uniqueness of X follow e.g. from [, Theorem 1.1]. We now show that the law μ_x of X satisfies $T_2(C)$ for some $C > 0$. By [41, Theorem 10.3 and Remark 10.4], the PDE

$$\begin{cases} \partial_t u_i + \mathcal{L}u_i + b_i = 0 \\ u_i(T, x) = 0 \end{cases}$$

admits a unique solution $u_i \in H_p^{2,q}([0, T])$ and this solution satisfies

$$\|\partial_t u_i\|_{L_p^q} + \|u_i\|_{\mathbb{H}_p^{2,q}([0, T])} \leq C_1 \|b_i\|_{L_p^q}$$

for some constant C_1 depending on d, p, q, T and $\|b\|_{L_p^q}$. Furthermore, since $d/p + 2/q < 1$, it follows by [41, Lemma 10.2] that

$$|\partial_x u_i| \leq C_2 T^{\varepsilon/2} \|u_i\|_{\mathbb{H}_p^{2,q}([0, T])}^{1-1/q-\varepsilon/2} \|\partial_t u_i\|_{L_p^q}^{1/q+\varepsilon/2}$$

with for $\varepsilon \in (0, 1)$ such that $\varepsilon + d/p + 2/q < 1$ and C_2 a constant depending on p, q and ε . Therefore, it follows that

$$|\partial_x u_i| \leq C_1 C_2 T^{\varepsilon/2} \|b_i\|_{L_p^q}, \quad (5.12)$$

so that choosing T small enough, we have $|\partial_x u| < \frac{1}{2d}$. Now consider the function $\phi_i(t, x) := X_i + u_i(t, x)$, $i = 1, \dots, d$. It is easily checked that the function ϕ_i solves the PDE

$$\begin{cases} \partial_t \phi_i + \mathcal{L}\phi_i = 0 \\ \phi_i(T, x) = X_i. \end{cases} \quad (5.13)$$

Put $\phi(t, x) = (\phi_1(t, x), \dots, \phi_d(t, x))$. Due to (5.12), it holds that

$$(1 - 2d|\partial_x u|^2)|x - y|^2 \leq 2|\phi(t, x) - \phi(t, y)|^2 \leq 2(2 + |\partial_x u|^2)|x - y|^2 \quad \text{for all } x, y \in \mathbb{R}^d.$$

As a consequence, ϕ is one-to-one, (see e.g. the corollary on page 87 of [36]), and its inverse $\psi := \phi^{-1}$ is Lipschitz continuous.

Since for every t , $u(t, \cdot)$ belongs to H_p^2 , then it can be seen as an element of $W_p^2(\mathbb{R}^d)$. Moreover, the derivative of u with respect to t belongs to L^p , it thus follows that u belongs to $W_p^{1,2}([0, T])$. Hence, the function $\phi(t, x) := x + u(t, x)$ belongs to $W_{p,loc}^{1,2}([0, T])$. Itô-Krylov's formula applied to ϕ gives

$$\begin{aligned} Y(t) := \phi(t, X(t)) &= \phi(0, x) + \int_0^t (\partial_t \phi + \mathcal{L}\phi)(s, X(s)) ds + \int_0^t \partial_x \phi(s, X(s)) \sigma dB(s) \\ &= \phi(0, x) + \int_0^t \tilde{\sigma}(s, Y(s)) dB(s) \end{aligned}$$

with $\tilde{\sigma}(t, y) := (\sigma^\top \partial_x \phi)(t, \psi(t, y))$, and where the second equation follows by (5.13). The rest of the proof follows as in the case of assumption (A). \square

Remark 5.5. The estimate (5.12) in the proof of Theorem 5.2 allows to see that when T is arbitrary, the conclusion of Theorem 5.2 still holds provided that $\|b\|_{L^q_p}$ is small enough.

5.4 Examples

Let us now present a few examples of multidimensional diffusion models with non-Lipschitz coefficients which fit to our framework.

Particles interacting through their rank

Let B_1, \dots, B_n be n independent Brownian motions. Rank-based interaction models are given by

$$dX^{i,n}(t) = \sum_{j=1}^n \delta_j 1_{\{X^{i,n}(t) = X^{(j),n}(t)\}} dt + \sigma^i(t) dB_i(t) \quad X^{i,n}(0) = X_i$$

for some real numbers δ_j , some measurable, bounded functions σ^i , with $X^{(1),n}(t) \leq X^{(2),n}(t) \leq \dots \leq X^{(n),n}(t)$ is the system in increasing order. More generally, this model can be written as

$$dX^{i,n}(t) = b\left(\frac{1}{n} \sum_{j=1}^n 1_{\{X^{n,j}(t) \leq X^{n,i}(t)\}}\right) dt + \sigma^i(t) dB_i(t) \quad X^{i,n}(0) = X_i$$

for a given (deterministic) functions b . This model was introduced by [30] in the context of stochastic portfolio theory. Concentration of measures results for such systems can be found in [64]. When $0 < c \leq \inf_{i,t} |\sigma^i(t)| \leq \sup_{i,t} |\sigma^i(t)| \leq C$ for some c, C and $b \in L^\infty$ or $b \in L^p(\mathbb{R}, dx)$ (with appropriate p, d), our main result shows that the law of $(X^{1,n}, \dots, X^{n,n})$ satisfies $T_2(C)$ for some $C > 0$. This result is also valid for the so-called (finite) Atlas model of [7] given by

$$dX^{i,n}(t) = \sum_{j=1}^n \delta 1_{\{X^{i,n}(t) = X^{p_i,n}(t)\}} dt + \sigma^i(t) dB_i(t) \quad X^{i,n}(0) = X_i,$$

for some constant δ and a permutation (p_1, \dots, p_n) of $(1, \dots, n)$.

Particles in quantile interaction

Quantile interaction models are given by

$$dX^{i,n}(t) = b(t, X^{n,i}(t), V^{\alpha,n}(t)) dt + \sigma(t, X^{n,i}(t)) dB_i(t) \quad X^{i,n}(0) = X_i,$$

where $V^{\alpha,n}(t)$ is the quantile at level $\alpha \in [0, 1]$ of the empirical measure of the system

$(X^{1,n}(t), \dots, X^{n,n}(t))$. That is,

$$V^{\alpha,n}(t) := \inf \left\{ u \in \mathbb{R} : \frac{1}{n} \sum_{i=1}^n 1_{\{X^{i,n}(t) \leq u\}} \geq \alpha \right\}.$$

This model is considered for instance in [18] in connection to exchangeable particle systems. Theorem 5.2 can be applied to this case under integrability conditions on b and mild regularity conditions σ .

Brownian motion with random drift

In addition to particle systems, our main result can also allow to derive transportation inequalities for semimartingales. We illustrate this in the next corollary. Let g be a progressive stochastic process. We call Brownian motion with drift the process

$$X(t) = x + \int_0^t g(s)ds + \sigma dB(t). \quad (5.14)$$

We have the following corollary of Theorem 5.2:

Corollary 5.6. Assume that the constant matrix σ satisfies (5.2). If the drift g is bounded and T small enough, then the law μ_t^x of $X(t)$ given by (5.14) satisfies $T_2(C)$ for some $C > 0$ depending on T, σ, d and $\|g\|_\infty$.

Proof. Consider the Borel measurable function

$$b(t, x) := \mathbb{E}[g(t)|X(t) = x].$$

By [14, Corollary 3.7], we have $\mu_t^x = \tilde{\mu}_t$, where $\tilde{\mu}_t$ is the law of the weak solution $\tilde{X}(t)$ of the SDE

$$\tilde{X}(t) = x + \int_0^t b(s, \tilde{X}(s))ds + \sigma dB(s). \quad (5.15)$$

Since g is bounded so is the function b . Thus, the SDE (5.15) admits a unique strong solution, see e.g. [3, 87]. Thus, \tilde{X} is necessarily a strong solution and by Theorem 5.2 $\tilde{\mu}$ satisfies $T_2(C)$, which concludes the argument. \square

Chapter 6

Transportation cost inequality for backward stochastic differential equations

Contents

6.1	Introduction	95
6.2	Transportation inequality for BSDEs	97
6.3	Concentration inequalities	101
6.4	Transportation cost inequality for FBSDE	102

Abstract

We prove that probability laws of a backward stochastic differential equation, satisfy a quadratic transportation cost inequality under the uniform metric. That is, a comparison of the Wasserstein distance from the law of the solution of the equation to any other absolutely continuous measure with finite relative entropy. From this we derive concentration properties of Lipschitz functions of the solution.

6.1 Introduction

Let (E, d) be a metric space equipped with a σ -field \mathcal{F} such that $d(\cdot, \cdot)$ is $\mathcal{F} \times \mathcal{F}$ measurable. Given $p \geq 1$ and two probability measures μ and ν on E , we define the Wasserstein distance of order p between μ and ν by

$$W_p^d(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{E \times E} d(x, y)^p d\pi(x, y) \right)^{1/p},$$

where $\Pi(\mu, \nu)$ is the set of all probability measures on the product space $E \times E$ with marginals μ and ν . The relative entropy of ν with respect to μ is

defined as

$$\mathbb{H}(\nu/\mu) = \begin{cases} \int_E \ln \frac{d\nu}{d\mu} d\nu, & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise.} \end{cases}$$

The probability measure μ satisfies the L^p -transportation inequality (also called the Talagrand inequality) on (E, d) if there exists a constant $C \geq 0$ such that for any probability measure ν , $W_p^d(\mu, \nu) \leq \sqrt{2C\mathbb{H}(\nu/\mu)}$. In this case we say that $\mu \in T_p(C)$.

The transportation inequalities of order 1, $T_1(C)$, are used to drive concentration inequalities, which provide upper bounds on the probability that a random variable, distributed according to a given measure, deviates from its mean by a given amount. Concentration inequalities are useful in fields as diverse as probability, statistics, convex geometry, functional analysis, statistical physics, stochastic finance and information theory. The property $T_2(C)$ was first established in [83] for the Gaussian measure and generalized in [32, 31] to an abstract Wiener space. Several other recent articles have showed intimate links between transportation inequalities and various other functional inequalities.

Moreover, the transportation inequalities are related to the concentration inequalities for risk measure as was formulated in [44, 85]. In the same vein there is a strong relation between Backward stochastic differential equations (BSDEs) and time-consistent dynamic risk measures, that is: BSDEs allow to define dynamic risk measures. Reciprocally time-consistent dynamic risk measures can be seen as solutions of BSDEs.

Motivated by the above quoted facts, the goal of this note is to prove a transportation inequality for the law of the solution Y of a BSDE and derive some concentration inequalities, like

$$\mathbb{P}\left(\sup_{t \in [0, T]} |Y_t - h(B_T)| - \mathbb{E}\left[\sup_{t \in [0, T]} |Y_t - h(B_T)|\right] > r\right) \leq \exp\left(-\frac{r^2}{8C}\right),$$

and

$$\mathbb{E} \exp(F(Y) - \mathbb{E}F(Y)) \leq \exp\left(\frac{C}{2}\|F\|_{Lip}^2\right).$$

This kind of inequalities are well known in the forward stochastic differential equations. Unlike for backward framework, to the best of our knowledge, there is no work reported on such as concentration inequalities.

Let $T > 0$ be a given finite horizon. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be probability space endowed with a real Brownian motion $B = (B_t)_{t \geq 0}$ and $(\mathcal{F}_t)_{t \in [0, T]}$ the natural filtration generated by B which satisfies the usual conditions. We denote

- \mathcal{S}^2 the set of real valued \mathcal{F} -adapted continuous processes Y on $[0, T]$ such that $\mathbb{E}\left[\sup_{0 \leq r \leq T} |Y_r|^2\right] < \infty$,
- \mathcal{S}_∞ the set of real valued \mathcal{F} -adapted continuous processes Y on $[0, T]$ such that $\|Y\|_{\mathcal{S}_\infty} = \sup_{\omega \in \Omega} \sup_{0 \leq r \leq T} |Y_r(\omega)| < \infty$,
- \mathcal{H}^2 the set of predictable real processes Z such that $\mathbb{E}\left[\int_0^T Z_r^2 dr\right] < \infty$,

- \mathcal{H}_P^2 the set of progressively measurable real processes Z such that $\mathbb{E} \left[\int_0^T Z_r^2 dr \right] < \infty$.

The paper is organized as follows. In section 2, we establish the property $T_2(C)$ for the law of Y solution of BSDEs. In section 3, we derive some concentration inequalities. In the last section, we explain how these results can be derived for a system of forward backward stochastic differential equations (FBSDEs).

6.2 Transportation inequality for BSDEs

We establish the property $T_2(C)$ for backward stochastic differential equations of the form

$$Y_t = h(B_T) + \int_t^T g(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s. \quad (6.1)$$

where $h : \mathbb{R} \rightarrow \mathbb{R}$ and $g : [0, T] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ are deterministic functions. We suppose the following assumptions

- ($\mathcal{H}.1$) h is bounded and there exists a positive constant l such that for all $x, x' \in \mathbb{R}$

$$|h(x) - h(x')| \leq l|x - x'|$$

- ($\mathcal{H}.2$) $\|g(\cdot, 0, 0)\|_\infty$ is bounded and there exist two constants $L_y > 0$ and $L_z > 0$ such that for all (t, y, z, y', z') in $[0, T] \times \mathbb{R}^4$

$$|g(t, y, z) - g(t, y', z')| \leq L_y|y - y'| + L_z|z - z'|.$$

Remark 6.1. It is known, since [66], that under assumptions ($\mathcal{H}.1$)-($\mathcal{H}.2$), the BSDE (6.1) has a unique solution (Y, Z) in $\mathcal{S}^2 \times \mathcal{H}^2$. Arguing as in [33], one can show that the solution (Y, Z) is bounded. In particular, there exist a positive constant c such that

$$\|Z\|_{\mathcal{S}_\infty} \leq c.$$

The main result is given in the next theorem.

Theorem 6.2. Under assumptions ($\mathcal{H}.1$) and ($\mathcal{H}.2$). There exists a positive constant C such that the law \mathbb{P}_Y of Y satisfies a transportation cost inequality, $T_2(C)$, on the space $\mathcal{C}([0, T], \mathbb{R})$ endowed with the uniform metric.

Proof. Let \mathbb{Q} be a probability measure on $\mathcal{C}([0, T], \mathbb{R})$ absolutely continuous with respect to \mathbb{P}_Y . We consider:

$$\tilde{\mathbb{Q}} = \frac{d\mathbb{Q}}{d\mathbb{P}_Y}(Y) \mathbb{P}, \quad M_s = \mathbb{E} \left(\frac{d\mathbb{Q}}{d\mathbb{P}_Y}(Y) / \mathcal{F}_s \right), \quad s \in [0, T]. \quad (6.2)$$

Since \mathbb{Q} is a probability measure we have:

$$\int_\Omega \frac{d\mathbb{Q}}{d\mathbb{P}_Y}(Y) d\mathbb{P} = \int_{\mathcal{C}([0, T], \mathbb{R})} \frac{d\mathbb{Q}}{d\mathbb{P}_Y} d\mathbb{P}_Y = \mathbb{Q}(\mathcal{C}([0, T], \mathbb{R})) = 1.$$

Then (M_s) is a uniformly integrable martingale which will be chosen continuous. Let $\tau = \inf\{s \in [0, T]; M_s = 0\} \wedge T$, with the convention that $\inf(\emptyset) = +\infty$. Then τ is an \mathcal{F}_s -stopping time such that $\tilde{\mathbb{Q}}(\tau = T) = 1$. Indeed, we can easily show by the martingale property that $\tilde{\mathbb{Q}}[M_s > 0] = 1$, for any $0 \leq s \leq T$.

We can write

$$M_s = \exp\left(L_s - \frac{1}{2} \langle L \rangle_s\right) 1_{s < \tau},$$

where $L_s := \int_0^s \frac{dM_r}{M_r}$, $\forall s < \tau$.

We have

$$\mathbb{H}(\mathbb{Q}/\mathbb{P}_Y) = \mathbb{H}(\tilde{\mathbb{Q}}/\mathbb{P}) = \mathbb{E}(M_T \ln M_T) = \mathbb{E}_{\tilde{\mathbb{Q}}}(\ln M_T). \quad (6.3)$$

Consider the sequence of stopping times

$$\tau_n = \inf\{s \in [0, \tau); \langle L \rangle_s = n\} \wedge \tau, \quad n \geq 1,$$

with the same convention that $\inf \emptyset = +\infty$. M is \mathbb{P} -a.s continuous, and since $\tau_n \rightarrow \tau$, we have $\lim_n M_{\tau_n \wedge T} \log M_{\tau_n \wedge T} = M_{\tau \wedge T} \log M_{\tau \wedge T}$ \mathbb{P} -a.s.. Note that the entropy $\mathbb{H}(\mathbb{Q}/\mathbb{P}_Y)$ is assumed to be finite, otherwise the result of Theorem 6.2 is obvious. So, since $\mathbb{E}(M_T \ln M_T) < \infty$, and by using the uniform integrability of $(M_{\tau_n \wedge T})_{n \geq 1}$ together with the help of some technical arguments of martingale theory, we can easily show that the sequence of variables $(M_{\tau_n \wedge T} \log M_{\tau_n \wedge T})_{n \geq 1}$ is uniformly integrable. Consequently, we get

$$\lim_n \mathbb{E}(M_{\tau_n \wedge T} \log M_{\tau_n \wedge T}) = \mathbb{E}(M_{\tau \wedge T} \log M_{\tau \wedge T}).$$

Martingale optional sampling theorem gives

$$\lim_n \mathbb{E}_{\tilde{\mathbb{Q}}}(\log M_{\tau_n \wedge T}) = \mathbb{E}_{\tilde{\mathbb{Q}}}(\log M_{\tau \wedge T}). \quad (6.4)$$

We have by equality (6.3) and (6.4)

$$\mathbb{H}(\mathbb{Q}/\mathbb{P}_Y) = \mathbb{H}(\tilde{\mathbb{Q}}/\mathbb{P}) = \lim_n \mathbb{E}_{\tilde{\mathbb{Q}}}\left(L_{T \wedge \tau_n} - \frac{1}{2} \langle L \rangle_{T \wedge \tau_n}\right). \quad (6.5)$$

Using Girsanov's theorem, we show that $(L_{t \wedge \tau_n} - \langle L \rangle_{t \wedge \tau_n})$ is a $\tilde{\mathbb{Q}}$ -martingale which is null in 0. We then get from (6.5)

$$\mathbb{H}(\mathbb{Q}/\mathbb{P}_Y) = \mathbb{H}(\tilde{\mathbb{Q}}/\mathbb{P}) = \frac{1}{2} \lim_n \mathbb{E}_{\tilde{\mathbb{Q}}}(\langle L \rangle_{T \wedge \tau_n}) = \frac{1}{2} \mathbb{E}_{\tilde{\mathbb{Q}}}(\langle L \rangle_T).$$

Therefore, the process $(L_{s \wedge \tau_n})$ is \mathbb{P} -square integrable martingale and $\tilde{\mathbb{Q}}[\lim_n \tau_n = \tau = T] = 1$. Hence, by representation theorem there is a predictable process $(\rho(s)_{0 \leq s < \tau})$ such that $\int_0^s \rho(r)^2 dr < +\infty$, \mathbb{P} -a.s and

$$L_s = \int_0^s \rho(r) dB_r.$$

It follows that

$$\mathbb{H}(\mathbb{Q}/\mathbb{P}_Y) = \mathbb{H}(\tilde{\mathbb{Q}}/\mathbb{P}) = \frac{1}{2} \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds. \quad (6.6)$$

Thanks to (6.2), we have $\frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}_Y}(Y) = M_T$, since is \mathcal{F}_T -measurable. Then $\tilde{\mathbb{Q}} = M_T\mathbb{P}$ and

$$M_T = \exp\left(\int_0^T \rho(r)dB_r - \frac{1}{2}\int_0^T \rho(r)^2 dr\right), \quad \tilde{\mathbb{Q}}\text{-a.s.}$$

By Girsanov's theorem

$$\bar{B}_s := B_s - \int_0^s \rho(r)dr,$$

is $\tilde{\mathbb{Q}}$ -Brownian motion. Consequently, the processes (Y, Z) satisfy

$$Y_t = h(B_T) + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s \rho(s)ds - \int_t^T Z_s d\bar{B}_s, \quad \tilde{\mathbb{Q}}\text{-a.s.} \quad (6.7)$$

Let (\bar{Y}, \bar{Z}) be the solution of the following backward equations

$$\bar{Y}_t = h(\bar{B}_T) + \int_t^T g(s, \bar{Y}_s, \bar{Z}_s)ds - \int_t^T \bar{Z}_s d\bar{B}_s, \quad \tilde{\mathbb{Q}}\text{-a.s.} \quad (6.8)$$

Under $\tilde{\mathbb{Q}}$, the law of \bar{Y} is \mathbb{P}_Y . Thus, (Y, \bar{Y}) is under $\tilde{\mathbb{Q}}$ a coupling of $(\mathbb{Q}, \mathbb{P}_Y)$ and it follows that

$$\left(W_2^{d_\infty}(\mathbb{Q}, \mathbb{P}_Y)\right)^2 \leq \mathbb{E}_{\tilde{\mathbb{Q}}}\left(\sup_{0 \leq s \leq T} |Y_s - \bar{Y}_s|^2\right).$$

Setting $\hat{Y}_t := Y_t - \bar{Y}_t$ and $\hat{Z}_t := Z_t - \bar{Z}_t$. We shall estimate \hat{Y}_t . Using Itô formula and Lipschitz assumption we get:

$$\begin{aligned} \hat{Y}_t^2 + \int_t^T \hat{Z}_s^2 ds &\leq l^2 |B_T - \bar{B}_T|^2 + 2L_y \int_t^T \hat{Y}_s^2 ds + 2L_z \int_t^T |\hat{Y}_s| |\hat{Z}_s| ds \\ &\quad - 2 \int_t^T \hat{Y}_s Z_s \rho(s) ds - 2 \int_t^T \hat{Y}_s \hat{Z}_s d\bar{B}_s, \end{aligned}$$

it follows

$$\begin{aligned} \hat{Y}_t^2 + \frac{1}{2} \int_t^T \hat{Z}_s^2 ds &\leq Tl^2 \int_0^T \rho(s)^2 ds + 2L_y \int_t^T \hat{Y}_s^2 ds + 2L_z^2 \int_t^T \hat{Y}_s^2 ds + \int_t^T \hat{Y}_s^2 ds \\ &\quad + \int_t^T Z_s^2 \rho(s)^2 ds - 2 \int_t^T \hat{Y}_s \hat{Z}_s d\bar{B}_s, \end{aligned}$$

take into account of the boundedness of Z we get

$$\begin{aligned} \hat{Y}_t^2 + \frac{1}{2} \int_t^T \hat{Z}_s^2 ds &\leq Tl^2 \int_0^T \rho(s)^2 ds + (1 + 2L_y + 2L_z^2) \int_t^T \hat{Y}_s^2 ds + c \int_t^T \rho(s)^2 ds \\ &\quad - 2 \int_t^T \hat{Y}_s \hat{Z}_s d\bar{B}_s, \end{aligned} \quad (6.9)$$

taking expectation under $\tilde{\mathbb{Q}}$, and using Gronwall's inequality we get

$$\mathbb{E}_{\tilde{\mathbb{Q}}}\hat{Y}_t^2 \leq C \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + C \int_t^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds$$

where C is a positive constant varying from line to another. We also have

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{Q}}}\int_t^T \hat{Z}_s^2 ds &\leq C \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + C \mathbb{E}_{\tilde{\mathbb{Q}}}\int_t^T \hat{Y}_s^2 ds + C \int_t^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds \\
&\leq C \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + C \int_t^T \int_s^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(u)^2 duds \\
&\leq C \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + C \int_t^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds.
\end{aligned} \tag{6.10}$$

On the other hand, using inequality (6.9), we obtain

$$\begin{aligned}
\sup_{0 \leq s \leq t} \hat{Y}_s^2 &\leq Tl^2 \int_0^T \rho(s)^2 ds + (1 + 2L_y + 2L_z^2) \int_0^T \hat{Y}_s^2 ds + c \int_0^T \rho(s)^2 ds \\
&\quad + 2 \sup_{0 \leq s \leq t} \left| \int_s^T \hat{Y}_s \hat{Z}_s d\bar{B}_s \right|.
\end{aligned}$$

Using the Burkholder-Davis -Gundy inequality we find

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq t} \hat{Y}_s^2 &\leq Tl^2 \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds + (1 + 2L_y + 2L_z^2) \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\hat{Y}_s^2 ds \\
&\quad + c \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + 2\mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq t} \left| \int_s^T \hat{Y}_s \hat{Z}_s d\bar{B}_s \right| \\
\mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq t} \hat{Y}_s^2 &\leq Tl^2 \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds + (1 + 2L_y + 2L_z^2) \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\hat{Y}_s^2 ds \\
&\quad + c \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + 2C_B \mathbb{E}_{\tilde{\mathbb{Q}}}\left(\sup_{0 \leq s \leq T} \hat{Y}_s^2 \int_0^T \hat{Z}_s^2 ds \right)^{\frac{1}{2}} \\
&\leq Tl^2 \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds + (1 + 2L_y + 2L_z^2) \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\hat{Y}_s^2 ds \\
&\quad + c \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + \frac{1}{2} \mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq T} \hat{Y}_s^2 + C_B^2 \mathbb{E}_{\tilde{\mathbb{Q}}}\left(\int_0^T \hat{Z}_s^2 ds \right),
\end{aligned}$$

then,

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq T} \hat{Y}_s^2 &\leq 2Tl^2 \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds + 2(1 + 2L_y + 2L_z^2) \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\hat{Y}_s^2 ds \\
&\quad + 2c \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\rho(s)^2 ds + 2C_B^2 \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \hat{Z}_s^2 ds.
\end{aligned}$$

Using inequality (6.10), it holds

$$\mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq T} \hat{Y}_s^2 \leq C \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds + 2(1 + 2L_y + 2L_z^2) \int_0^T \mathbb{E}_{\tilde{\mathbb{Q}}}\hat{Y}_s^2 ds$$

Gronwall's inequality implies that

$$\mathbb{E}_{\tilde{\mathbb{Q}}}\sup_{0 \leq s \leq T} \hat{Y}_s^2 \leq C \mathbb{E}_{\tilde{\mathbb{Q}}}\int_0^T \rho(s)^2 ds.$$

Thus,

$$W_2^{d_\infty}(\mathbb{Q}, \mathbb{P}_Y)^2 \leq 2 C \text{H}(\mathbb{Q}/\mathbb{P}_Y).$$

This completes the proof. \square

Remark 6.3. The entropy representation formula (6.6) is the key of the proof of Theorem 6.2. For its proof, we have followed here the ideas used in [25]. However, another proof with a different point of view can be found in the paper [86].

6.3 Concentration inequalities

The link between the concentration inequalities and the L^1 -transportation inequality is proved in [9]. We recall that a measure μ on the metric space (E, d) satisfies the property $T_1(C)$ if and only if for any μ -integrable and Lipschitz function $F : (E, d) \rightarrow \mathbb{R}$ and for every $\lambda \in \mathbb{R}$ we have the following Gaussian concentration inequality

$$\int_E \exp\left(\lambda\left(F - \int_E F d\mu\right)\right) d\mu \leq \exp\left(C \frac{\lambda^2}{2} \|F\|_{Lip}^2\right),$$

where

$$\|F\|_{Lip} = \sup_{x \neq y} \frac{|F(x) - F(y)|}{d(x, y)}.$$

By Chebychev's inequality and an optimization argument, we obtain the following inequality of Hoeffding's type:

$$\mu\left(F - \int_E F d\mu \geq r\right) \leq \exp\left(-\frac{r^2}{2C \|F\|_{Lip}^2}\right), \forall r > 0. \quad (6.11)$$

1. The property $T_1(C)$ holds for the probability measure \mathbb{P}_Y , then we have for any Lipschitz function $F : \mathcal{C}([0, T], \mathbb{R}) \rightarrow \mathbb{R}$

$$\mathbb{P}_Y(F - E_{\mathbb{P}_Y} F > r) \leq \exp\left(-\frac{r^2}{2C \|F\|_{Lip}^2}\right), \quad \forall r > 0. \quad (6.12)$$

Let $V : \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz function, such that $\|V\|_{Lip} \leq \alpha$. We define F_V and F_∞ on $\mathcal{C}([0, T], \mathbb{R})$ by

$$F_V(\gamma) = \frac{1}{T} \int_0^T V(\gamma(t)) dt, \quad (6.13)$$

$$F_\infty(\gamma) = \sup_{t \in [0, T]} |\gamma(t) - \gamma(T)|. \quad (6.14)$$

The function F_V is α -Lipschitzian with respect to d_∞ . Similarly F_∞ , is 2-Lipschitzian map with respect to d_∞ . Using (6.12), we have the following

Hoeffding-type inequalities for Y on the metric space $\mathcal{C}([0, T], \mathbb{R})$, endowed with the metric d_∞ . For each $r > 0$ we have

$$\mathbb{P}\left(\frac{1}{T} \int_0^T V(Y_t) - \mathbb{E}F_V(Y) dt > r\right) \leq \exp\left(-\frac{r^2}{2\alpha^2 C}\right),$$

and using the functional F_∞ we get

$$\mathbb{P}\left(\sup_{t \in [0, T]} |Y_t - h(B_T)| - \mathbb{E}\left[\sup_{t \in [0, T]} |Y_t - h(B_T)|\right] > r\right) \leq \exp\left(-\frac{r^2}{8C}\right).$$

2. The following result was established in [10] on \mathbb{R}^d and extended after to $\mathcal{C}([0, T], \mathbb{R}^d)$ in [90]. Let F be a lower bounded measurable function on $\mathcal{C}([0, T], \mathbb{R})$, and consider

$$Q_c F(\gamma) := \inf_{\phi \in \mathcal{C}([0, T], \mathbb{R})} \left(F(\gamma + \phi) + \frac{1}{2c} \|\phi\|_\infty^2 \right)$$

the inf-convolution on $\mathcal{C}([0, T], \mathbb{R})$ with respect to metric d_∞ . Then $T_2(C)$ in Theorem 6.2 implies that

$$\mathbb{E} \exp\left(Q_c F(Y)\right) \leq \exp\left(\mathbb{E}F(Y)\right).$$

If in addition F is Lipschitz, since $Q_c F \geq F - \frac{C}{2} \|F\|_{Lip}^2$, we have the following concentration inequality

$$\mathbb{E} \exp\left(F(Y) - \mathbb{E}F(Y)\right) \leq \exp\left(\frac{C}{2} \|F\|_{Lip}^2\right).$$

6.4 Transportation cost inequality for FBSDE

Let us now consider the following forward backward stochastic differential equations

$$\begin{cases} \forall s \in [t, T] \\ X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) ds + \int_t^s \sigma(r, X_r^{t,x}) dB_r, \\ Y_s^{t,x} = h(X_T^{t,x}) + \int_s^T g(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dB_r. \end{cases} \quad (6.15)$$

where $h : \mathbb{R} \rightarrow \mathbb{R}$, $b, \sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ and $g : [0, T] \times \mathbb{R}^3 \rightarrow \mathbb{R}$ are given maps. We make the following assumptions: there exist positive constants k, Λ and Δ such that for all $t \in [0, T]$ and $(x, y, z, x', y', z') \in \mathbb{R}^6$:

$$(\mathcal{H}.1) \quad (x - x')(b(t, x) - b(t, x')) \leq k|x - x'|^2 \quad \text{and} \quad |b(t, x)| \leq k(1 + |x|),$$

$$(\mathcal{H}.2) \quad |\sigma(t, x) - \sigma(t, x')| \leq k|x - x'|, \quad |\sigma(t, x)| \leq k(1 + |x|) \quad \text{and} \quad \sigma^2(t, x) \geq \Lambda,$$

$$(\mathcal{H}.3) \quad |h(x) - h(x')| \leq k|x - x'| \quad \text{and} \quad |h(x)| \leq \Delta,$$

$$(\mathcal{H}.4) \quad |g(t, x, y, z) - g(t, x', y, z')| \leq k(|x - x'| + |z - z'|),$$

$$(\mathcal{H}.5) \quad (y - y')(g(t, x, y, z) - g(t, x, y', z)) \leq k|y - y'|^2 \quad \text{and} \quad |g(t, x, y, z)| \leq k(1 + |y| + |z|).$$

(\mathcal{H}.6) The functions σ , $x \mapsto b(t, x)$ and $y \mapsto g(t, x, y, z)$ are continuous.

The next theorem is proven in [21]. (Theorem 2.6)

Theorem 6.4. ([21]) Suppose that (\mathcal{H}.1)-(\mathcal{H}.6) hold. Then the system (6.15) admits a unique solution $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})_{t \leq s \leq T}$ in $\mathcal{S}^2 \times \mathcal{S}^2 \times \mathcal{H}_P^2$. Therefore, there exist versions of the processes $(Y_s^{t,x})$ and $(Z_s^{t,x})$ whose trajectories are uniformly bounded.

In the sequel we consider the bounded version of $(Y^{t,x}, Z^{t,x})$. We denote the law of $(X^{t,x}, Y^{t,x})$ on the metric space $\mathcal{B} = \mathcal{C}([t, T], \mathbb{R}) \times \mathcal{C}([t, T], \mathbb{R})$ by \mathbb{P}_x . We have the next result that can be proven by a similar arguments as those used in Section 6.2.

Theorem 6.5. Under assumptions (\mathcal{H}.1)-(\mathcal{H}.6) and $\|\sigma\|_\infty := \sup\{|\sigma(t, x)|; x \in \mathbb{R}, t \in [0, T]\} < \infty$. There exists a positive constant C such that the probability measure \mathbb{P}_x satisfies $T_2(C)$ on the metric space \mathcal{B} with respect to the metric

$$d((x, y), (x', y')) = \left(\sup_{t \leq s \leq T} |x(s) - x'(s)|^2 + \sup_{t \leq s \leq T} |y(s) - y'(s)|^2 \right)^{\frac{1}{2}}.$$

The next lemma permits to obtain the property $T_2(C)$ for the process Y on the space $\mathcal{C}([t, T], \mathbb{R})$ endowed with the uniform metric.

Lemma 6.6. ([25]) Let (E, d_E) and (F, d_F) be two metric spaces and $\psi : (E, d_E) \rightarrow (F, d_F)$ is a Lipschitz application, such that

$$d_F(\psi(x), \psi(y)) \leq \alpha d_E(x, y) \quad \forall x, y \in E,$$

If $\mu \in T_2(C)$ on (E, d_E) , then $\tilde{\mu} := \mu \circ \psi^{-1} \in T_2(\alpha^2 C)$ on (F, d_F) .

Remark 6.7. In view of the lemma above, we obtain that the probability measure \mathbb{P}_Y satisfies $T_2(C)$, consequently a transportation inequality of order 1, $T_1(C)$. As before, we consider:

$$F_\infty(\gamma) = \sup_{s \in [t, T]} |\gamma(s) - \gamma(T)|.$$

Then we have the following concentration inequality: for every $r > 0$

$$\mathbb{P} \left(\sup_{s \in [t, T]} |Y^{t,x}(s) - h(X_T^{t,x})| - \mathbb{E} \left[\sup_{s \in [t, T]} |Y^{t,x}(s) - h(X_T^{t,x})| \right] > r \right) \leq \exp \left(-\frac{r^2}{8C} \right).$$

Remark 6.8. We have proved our results under assumptions ensuring the boundedness of the term Z . Which is in some sense a natural condition, because this process plays the role of the diffusion term in the forward stochastic equations context. It's known that for SDEs the boundedness of the diffusion coefficient is required to obtain T_2 - inequalities.

Appendix A

S -topology

The S -topology on the space $\mathcal{D}([0, T], \mathbb{R}^d)$ was introduced by Jakubowski [35]. It is weaker than the Skorokhod topology. Its compactness criterion is simple to obtain as for the Meyer-Zheng one [57]. We recall below here some relevant results about the S -topology in the case of real space \mathbb{R} , they can be easily extended to finite dimensional space \mathbb{R}^d . The proof of the following properties can be found in Jakubowski [35] and [47].

Proposition A.1. (i) $K \subset \mathcal{D}([0, T], \mathbb{R})$ is relatively S -compact if and only if

$$\sup_{x \in K} \sup_{t \in [0, T]} |x_t| < +\infty \quad (\text{A.1})$$

and for all $a, b \in \mathbb{R}$ such that $a < b$

$$\sup_{x \in K} N^{a,b}(x) < +\infty \quad (\text{A.2})$$

where $N^{a,b}$ is the usual number of up-crossings given levels $a < b$, that is, $N^{a,b}(x) \geq k$ if one can find numbers $0 \leq t_1 < t_2 < \dots < t_{2k-1} < t_{2k} \leq T$ such that $x_{t_{2i-1}} < a$ and $x_{t_{2i}} > b$, $i = 1, 2, \dots, k$.

(ii) x^n converges to x in the S -topology if and only if (x^n) satisfies (A.1), (A.2) and for every subsequence (n_k) , one can find a further subsequence (n_{k_l}) and a countable subset $Q \subset [0, T]$ such that $x_t^{n_{k_l}} \rightarrow x_t$, $t \in [0, T] \setminus Q$.

Corollary A.2. If (x^n) is relatively S -compact and there exists a countable subset Q such that for every $t \in [0, T] \setminus Q$, $x_t^n \rightarrow x_t$, then (x^n) converges to x .

We now recall that a sequence of processes $(X^n)_n$ converges weakly to X in the S -topology, $X^n \xrightarrow[S]{*} X$, if for every subsequence (X^{n_k}) , we can find a further subsequence $(X^{n_{k_l}})$ and a stochastic processes (Y_l) and Y defined on $([0, 1], \mathcal{B}_{[0,1]}, \lambda)$, such that the laws of Y_l and $X^{n_{k_l}}$ are the same, $l \in \mathbb{N}$, for each $\omega \in [0, 1]$ $Y_l(\omega)$ converges to $Y(\omega)$ in the S -topology, and for each $\varepsilon > 0$, there exists an S -compact subset $K_\varepsilon \subset \mathcal{D}([0, T], \mathbb{R})$ such that

$$\lambda(\{\omega \in [0, 1] : Y_l(\omega) \in K_\varepsilon, l = 1, 2, \dots\}) > 1 - \varepsilon.$$

Proposition A.3. The following two properties are equivalent

- (i.) (X^n) is S -tight.
- (ii.) (X^n) is relatively compact with respect to the convergence " $\xrightarrow[S]{*}$ "

Proposition A.4. If (X^n) is S -tight and there exists a countable subset $Q \subset [0, T]$ such that for every $j \in \mathbb{N}$ and every $t_1, t_2, \dots, t_j \in [0, T] \setminus Q$

$$(X_{t_1}^n, X_{t_2}^n, \dots, X_{t_j}^n) \xrightarrow{*} (X_{t_1}, X_{t_2}, \dots, X_{t_j})$$

where X is a process with trajectories in $\mathcal{D}([0, T], \mathbb{R})$. Then $X^n \xrightarrow[S]{*} X$.

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration \mathcal{F}_t , let X be an adapted process with paths a.s in $\mathcal{D}([0, T], \mathbb{R})$. If X_t is integrable for all $t \in [0, T]$, we define the conditional variation of X by

$$\text{CV}_T(X) = \sup_{\pi} \sum_{i=1}^n \mathbb{E} [|\mathbb{E}[X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}]|], \quad (\text{A.3})$$

where the supremum is taken over all subdivisions π of the interval $[0, T]$. If $\text{CV}_T(X) < \infty$ then the process X is called a quasi-martingale. Notice that for martingales X the quantity $\text{CV}_T(X) = 0$.

We have the following criterion, for the proof we refer for example to [47] and the references therein.

Theorem A.5. Let $(X^n)_{n \geq 1}$ be a family of stochastic process in $\mathcal{D}([0, T], \mathbb{R})$. If

$$\sup_{n \geq 1} \left(\text{CV}_T(X^n) + \mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s^n| \right] \right) < \infty, \quad (\text{A.4})$$

then the sequence $(X^n)_{n \geq 1}$ is S -tight and there exists a subsequence $(X^{n_k})_{k \geq 1}$ of $(X^n)_{n \geq 1}$, a process X belonging to $\mathcal{D}([0, T], \mathbb{R})$, and a countable subset $Q \subset [0, T)$ such that for every $j \geq 1$ and for any finite subset $\{t_1, \dots, t_j\}$ of $[0, T] \setminus Q$ the following convergence is true:

$$(X_{t_1}^{n_k}, \dots, X_{t_j}^{n_k}) \xrightarrow{*} (X_{t_1}, \dots, X_{t_j}) \quad \text{as } k \rightarrow \infty.$$

Remark A.6. Note that T is not in the countable subset Q . More precisely the projection $\pi_T : \mathcal{D}([0, T], \mathbb{R}) \rightarrow \mathbb{R}$, which assigns to x the value $x(T)$, is continuous with respect to the S -topology (cfr Remark 2.4. p.8 in [35]).

Bibliography

- [1] K. Bahlali, B. Boufoussi, and S. Mouchtabih. Transportation cost inequality for backward stochastic differential equations. *Stat. Probab. Lett.*, 155, 2019.
- [2] K. Bahlali, A. Elouaffin, and E. Pardoux. Homogenization of semilinear PDEs with discontinuous averaged coefficients. *Electron. J. Probab.*, 14:477–499, 2009.
- [3] Khaled Bahlali. Flows of homeomorphisms of stochastic differential equations with measurable drift. *Stochastics Stoch. Reports*, 67(1-2):53–82, jul 1999.
- [4] Khaled Bahlali, Brahim Boufoussi, and Soufiane Mouchtabih. Penalization for a PDE with a Nonlinear Neumann boundary condition and measurable coefficients, <https://hal.archives-ouvertes.fr/hal-02507957>. nov 2019.
- [5] Khaled Bahlali, Lucian Maticiuc, and Adrian Zălinescu. Penalization method for a nonlinear Neumann PDE via weak solutions of reflected SDEs. *Electron. J. Probab.*, 18:1–20, 2013.
- [6] Khaled Bahlall and Brahim Mezerdi. Some properties of Solutions of stochastic differential equations driven by semi-martingales. *Random Oper. Stoch. Equations*, 9(4):308–318, jan 2001.
- [7] Adrian D. Banner, Robert Fernholz, and Ioannis Karatzas. Atlas models of equity markets. *Ann. Appl. Probab.*, 15(4):2296–2330, nov 2005.
- [8] Daniel Bartl and Ludovic Tangpi. Functional inequalities for forward and backward diffusions. *Electron. J. Probab.*, 25:1–22, 2020.
- [9] S. G. Bobkov and F. Götze. Exponential Integrability and Transportation Cost Related to Logarithmic Sobolev Inequalities. *J. Funct. Anal.*, 163(1):1–28, apr 1999.
- [10] Sergey G. Bobkov, Ivan Gentil, and Michel Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. des Math. Pures Appl.*, 80(7):669–696, sep 2001.
- [11] B. Boufoussi and J. van Casteren. An approximation result for a nonlinear Neumann boundary value problem via BSDEs. *Stoch. Process. their Appl.*, 114(2):331–350, dec 2004.

-
- [12] Brahim Boufoussi and Salah Hajji. Transportation inequalities for stochastic heat equations. *Stat. Probab. Lett.*, 139:75–83, aug 2018.
- [13] Philippe Briand and Ying Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Relat. Fields*, 141(3-4):543–567, 2008.
- [14] Gerard Brunick and Steven Shreve. Mimicking an itô process by a solution of a stochastic differential equation. *Ann. Appl. Probab.*, 23(4):1584–1628, aug 2013.
- [15] L. Caffarelli, M. G. Crandall, M. Kocan, and A. Świąch. On viscosity solutions of fully nonlinear equations with measurable ingredients. *Commun. Pure Appl. Math.*, 49(4):365–398, 1996.
- [16] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. *Lect. Notes Math.*, 2123:231–292, 2014.
- [17] M G Crandall, M Kocan, P L Lions, and A ´ Swie. EXISTENCE RESULTS FOR BOUNDARY PROBLEMS FOR UNIFORMLY ELLIPTIC AND PARABOLIC FULLY NONLINEAR EQUATIONS. *Electron. J. Differ. Equations*, 1999(24):1–20, 1999.
- [18] Dan Crisan, Thomas G. Kurtz, and Yoonjung Lee. Conditional distributions, exchangeable particle systems, and stochastic partial differential equations. *Ann. l’institut Henri Poincare Probab. Stat.*, 50(3):946–974, 2014.
- [19] Jakša Cvitanić and Ioannis Karatzas. Backward stochastic differential equations with reflection and Dynkin games. *Ann. Probab.*, 24(4):2024–2056, 1996.
- [20] R. W.R. Darling and Etienne Pardoux. Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *Ann. Probab.*, 25(3):1135–1159, 1997.
- [21] François Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stoch. Process. their Appl.*, 99(2):209–286, jun 2002.
- [22] François Delarue, Daniel Lacker, and Kavita Ramanan. From the master equation to mean field game limit theory: A central limit theorem. *Electron. J. Probab.*, 24, 2019.
- [23] A Dembo. Information inequalities and concentration of measure. *Ann. Prob*, 25(2), 1997.
- [24] Amir Dembo and Ofer Zeitouni. TRANSPORTATION APPROACH TO SOME CONCEN-TRATION INEQUALITIES IN PRODUCT SPACES. *Elect. Comm. in Probab*, 1:83–90, 1996.

- [25] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32(3 B):2702–2732, 2004.
- [26] Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, jan 2009.
- [27] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M. C. Quenez. Reflected solutions of backward SDE's, and related obstacle problems for PDE's. *Ann. Probab.*, 25(2):702–737, 1997.
- [28] Nicole El Karoui. Processus de réflexion dans \mathbb{R}^n . In *Séminaire Probab. IX, Lect. Notes Math.*, pages 534–554. Springer, Berlin, 1975.
- [29] Lawrence Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, mar 1997.
- [30] R Fernholz and I Karatzas. Stochastic portfolio theory: A survey. In *Math. Model. Numer. Methods Financ.* Elsevier, Amsterdam, 2009.
- [31] D. Feyel and A. S. Üstünel. Monge-Kantorovitch Measure Transportation and Monge-Ampère Equation on Wiener Space. *Probab. Theory Relat. Fields*, 128(3):347–385, mar 2004.
- [32] Denis Feyel and Ali Süleyman Üstünel. Transport de mesure sur l'espace de Wiener et théorème de Girsanov. *Comptes Rendus Math.*, 334(11):1025–1028, jun 2002.
- [33] S. Hamadene. EDSR : Le cas localement Lipschitzien. *Ann. IHP, Probab. Stat.*, 32(5):645–659, 1996.
- [34] Pei Hsu. Probabilistic approach to the neumann problem. *Commun. Pure Appl. Math.*, 38(4):445–472, jul 1985.
- [35] Adam Jakubowski. A non-skorohod topology on the skorohod space. *Electron. J. Probab.*, 2(2):1–21, 1997.
- [36] Fritz John. On quasi-isometric mappings, I. *Commun. Pure Appl. Math.*, 21(1):77–110, jan 1968.
- [37] H Kaneko and S Nakao. A NOTE ON APPROXIMATION FOR STOCHASTIC DIFFERENTIAL EQUATIONS. *Séminaire Probab. Strasbg.*, 22:155–162, 1988.
- [38] L V Kantorovich. On the translocation of masses. *Dokl. Akad. Nauk. USSR 37 (1942), 199–201. English Transl. J. Math. Sci.*, 133(20):1381–1382, 2006.

-
- [39] Magdalena Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.*, 28(2):558–602, 2000.
- [40] N. V. Krylov. On weak uniqueness for some diffusions with discontinuous coefficients. *Stoch. Process. their Appl.*, 113(1):37–64, sep 2004.
- [41] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Relat. Fields*, 131(2):154–196, feb 2005.
- [42] Krylov N. V. *Controlled diffusion processes*. Springer, New York, 1982.
- [43] Krylov N.V. 6) 11(x). *Sib. Mat. J. 10 (2), 1969, 343-354 (in Russ. (English Transl. Sib. Math. J.(2), 1969, 343-354 (in Russ. (English Transl. Sib. Math. J., 10(2), 1969.*
- [44] Daniel Lacker. Liquidity, risk measures, and concentration of measure. *Math. Oper. Res.*, 43(3):813–837, aug 2018.
- [45] Weronika Łaukajtys and Leszek Słomiński. Penalization methods for reflecting stochastic differential equations with jumps. *Stochastics Stoch. Reports*, 75(5):275–293, oct 2003.
- [46] M. Ledoux. *The Concentration of Measure Phenomenon*, volume 89. 2001.
- [47] Antoine Lejay. BSDE driven by Dirichlet process and semi-linear parabolic PDE. Application to homogenization. *Stoch. Process. their Appl.*, 97(1):1–39, jan 2002.
- [48] Z Li and J Luo. Transportation inequalities for stochastic delay evolution equations driven by fractional Brownian motion. *Front. Math. China*, 10(2):303–321, 2015.
- [49] P. L. Lions and A. S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Commun. Pure Appl. Math.*, 37(4):511–537, 1984.
- [50] P.L. Lions, J.L. Menaldi, and A.S. Sznitman. Construction de processus de diffusion réfléchis par pénalisation du domaine. *C. R. Acad. Sci. Paris, Ser. I*, 292:559–562, 1981.
- [51] K. Marton. Bounding \bar{d} -distance by informational divergence: A method to prove measure concentration. *Ann. Probab.*, 24(2):857–866, 1996.
- [52] Katalin Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.
- [53] Pascal Massart. Concentration inequalities and model selection. *Lect. Notes Math.*, 1896:1–324, 2007.

-
- [54] Lucian Maticiuc and Aurel Rascanu. Backward stochastic variational inequalities on random interval. *Bernoulli*, 21(2):1166–1199, may 2015.
- [55] A. V. Mel'nikov. Stochastic Equations and Krylov'S Estimates for Semimartingales. *Stochastics*, 10(2):81–102, 1983.
- [56] Jose Menaldi. Stochastic variational inequality for reflected diffusion, 1983.
- [57] P. MEYER and W. ZHENG. Tightness criteria for laws of semimartingales. *Ann. l'I.H.P. Probab. Stat.*, 20(4):353–372, 1984.
- [58] V.D. Milman. Asymptotic properties of functions of several variables that are defined on homogeneous spaces. *Dokl. Akad. Nauk SSSR*, 199:1247–1250, 1971.
- [59] G Monge. Mémoire sur la théorie des déblais et des remblais. *l'Imprimerie R.*, 1781.
- [60] N. Gozlan; C. Leonard. Transport Inequalities. A Survey. *Markov Process. Relat. Fields*, 16(4):635–736, 2010.
- [61] F Otto and Villani C. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173:361–400, 2000.
- [62] Soumik Pal. Concentration for multidimensional diffusions and their boundary local times. *Probab. Theory Relat. Fields*, 154(1-2):225–254, jun 2012.
- [63] Soumik Pal and Andrey Sarantsev. A note on transportation cost inequalities for diffusions with reflections. *Electron. Commun. Probab.*, 24, 2019.
- [64] Soumik Pal and Mykhaylo Shkolnikov. Concentration of measure for Brownian particle systems interacting through their ranks. *Ann. Appl. Probab.*, 24(4):1482–1508, 2014.
- [65] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In LNCIS 176 B.L. Rozovskii , R. Sowers eds, editor, *Stoch. Partial Differ. Equations Their Appl.*, pages 200–217. Springer-Verlag, oct 1992.
- [66] E. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Syst. Control Lett.*, 14(1):55–61, 1990.
- [67] E Pardoux and R J Williams. Symmetric reflected diffusions. *Ann. l'I.H.P. Probab. Stat.*, 30(1):13–62, 1994.
- [68] Étienne Pardoux. BSDEs, weak convergence and homogenization of semilinear PDEs. In *Nonlinear Anal. Differ. Equations Control*, pages 503–549. Springer Netherlands, 1999.

-
- [69] Etienne Pardoux and Aurel Rascanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, volume 69. 2014.
- [70] Etienne Pardoux and Aurel Răşcanu. Continuity of the Feynman–Kac formula for a generalized parabolic equation. *Stochastics*, 89(5):726–752, 2017.
- [71] Etienne Pardoux and Shuguang Zhang. Generalized BSDEs and nonlinear Neumann boundary value problems. *Probab. Theory Relat. Fields*, 110(4):535–558, 1998.
- [72] Luca Di Persio, Lucian Maticiuc, and Adrian Z[˘] Alinescu. Continuity with respect to parameters of the solutions of time-delayed BSDEs with Stieltjes integral.
- [73] Sebastian Riedel. Transportation–cost inequalities for diffusions driven by Gaussian processes. *Electron. J. Probab.*, 22, 2017.
- [74] Andrzej Rozkosz and Leszek Słomiński. On existence and stability of weak solutions of multidimensional stochastic differential equations with measurable coefficients. *Stoch. Process. their Appl.*, 37(2):187–197, apr 1991.
- [75] Andrzej Rozkosz and Leszek Słomiński. On stability and existence of solutions of SDEs with reflection at the boundary. *Stoch. Process. their Appl.*, 68(2):285–302, 1997.
- [76] G. Lugosi S. Boucheron and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.
- [77] K Sato and T Ueno. Multi-dimensional diffusion and the Markov. 3:529–605, 1965.
- [78] Bruno Saussereau. Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion. *Bernoulli*, 18(1):1–23, feb 2012.
- [79] A.V. Skorohod. *Studies in the theory of random processes*. Addison Wesley, Massachussetts, 1965.
- [80] Leszek Slomiński. Euler’s approximations of solutions of SDEs with reflecting boundary. *Stoch. Process. their Appl.*, 94(2):317–337, aug 2001.
- [81] Leszek Słomiński. Weak and strong approximations of reflected diffusions via penalization methods. *Stoch. Process. their Appl.*, 123(3):752–763, 2013.
- [82] Daniel W. Stroock and S. R.S. Varadhan. Diffusion processes with boundary conditions. *Commun. Pure Appl. Math.*, 24(2):147–225, 1971.
- [83] Michel Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

-
- [84] Hiroshi Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 9(1):163–177, 1979.
- [85] Ludovic Tangpi. Concentration of dynamic risk measures in a Brownian filtration. *Stoch. Process. their Appl.*, 129(5):1477–1491, may 2019.
- [86] Ali Suleyman Üstünel. Transportation Cost Inequalities for Diffusions Under Uniform Distance. In *Springer Proc. Math. Stat.*, volume 22, pages 203–214. Springer New York LLC, 2012.
- [87] A. Y. Veretennikov. On the strong solutions and explicit formulas for solutions of stochastic differential equations. *Math. USSR Sb.*, 39(3):387–403, 1981.
- [88] C Villani. *Optimal Transport, Old and New*.
- [89] C Villani. *Topics in Optimal Transportation*.
- [90] C. Villani. *Topics in optimal Transportation*. Amer. Math. Soc., Providence, RI., 2003.
- [91] F.-Y Wang. Probability distance inequalities on Riemannian manifolds and path spaces. *J. Funct. Anal.*, 206:167–190, 2004.
- [92] Wu. L Zhang. Z. Talagrand's T2-transportation inequality and log-Sobolev inequality for dissipative SPDEs and applications to reaction-diffusion equations. *Chin Ann Math Ser B*, 27:243–262, 2006.
- [93] A K Zvonkin. A TRANSFORMATION OF THE PHASE SPACE OF A DIFFUSION PROCESS THAT REMOVES THE DRIFT. *Math. USSR-Sbornik*, 22(1):129–149, feb 1974.