Theoretical and numerical study of problems nonlinear in the sense of McKean in finance

Alexandre Zhou

École doctorale : mathématiques et sciences
et technologies de l'information et de la communication

Thèse de doctorat

Spécialité : Mathématiques appliquées

présentée par

Alexandre ZHOU

Etude théorique et numérique de problèmes non linéaires au sens de McKean en finance

Thèse dirigée par Benjamin JOURDAIN
et préparée au CERMICS, Ecole des Ponts ParisTech

Soutenue le 17 Octobre 2018 devant le Jury composé de :

M. Frédéric ABERGEL	Centrale Supélec	Examinateur
Mme. Mireille BOSSY	INRIA Sophia Antipolis	Rapporteur
M. Jean-François CHASSAGNEUX	Université Paris Diderot	Rapporteur
M. Emmanuel GOBET	Ecole Polytechnique	Examinateur
M. Pierre HENRY-LABORDERE	Société Générale	Examinateur
M. Benjamin JOURDAIN	Ecole des Ponts ParisTech	Directeur de thèse

A Tong Kuan, Yue Min et Yu Ying.

Remerciements

Mes premiers remerciements s'adressent bien évidemment à mon directeur de thèse, Benjamin Jourdain. Son soutien permanent, tant sur le plan professionnel que personnel, a été essentiel pour le bon déroulement de ma thèse. Je lui suis extrêmement reconnaissant pour sa confiance, sa patience à mon égard et bien sûr, sa disponibilité. Il a su rendre ces trois années agréables et enrichissantes par son enthousiasme et l'étendue de ses connaissances. Si j'ai beaucoup appris à ses côtés, son efficacité et sa rigueur restent toujours pour moi un idéal à atteindre. J'espère pouvoir continuer à progresser ainsi.

Je tiens à remercier sincèrement Mireille Bossy et Jean-François Chassagneux pour l'intérêt qu'ils ont porté à mon travail en acceptant d'être les rapporteurs de ma thèse. Je suis également très honoré de la présence de Frédéric Abergel, d'Emmanuel Gobet et de Pierre Henry-Labordère dans mon jury de thèse. Leurs travaux ont beaucoup influencé mes recherches. Je suis très reconnaissant envers la chaire "Risques Financiers" de la Fondation du Risque pour avoir participé au financement de ma thèse.

Je remercie Stefano de Marco pour m'avoir initié aux mathématiques financières et de m'avoir montré les problématiques passionantes qu'offre cette discipline. Je salue également Florian Bourgey pour les calculs touffus que nous avons partagés ces derniers mois autour d'un café à l'IHP ou au CMAP. J'ai également une pensée pour Ankush Agarwal, Fanny Noubiagain, José Salas ainsi que tous les participants du CEMRACS 2017, pour nos six semaines riches en collaborations et en émotions, mais aussi pour les soirées incroyables que nous avons faites depuis.

Merci à Aymeric Bettinelli, Flavia Galante, Salmane Lahdachi et Alexandre Lethay de l'équipe "Hybrides" ainsi que tous mes collègues de la Société Générale pour m'avoir accompagné pour mes premiers pas dans l'industrie financière et pour tout ce que vous m'avez appris.

Enseigner à l'Ecole des Ponts et Chaussées a été une expérience plaisante et très formatrice. Je tiens en particulier à exprimer ma gratitude envers Aurélien Alfonsi et Vincent Lefieux qui m'ont confié des rôles d'enseignements pour leurs cours de probabilités et statistiques.

Ces trois années au CERMICS ont été une expérience aussi profitable qu'agréable, et je voudrais remercier tous mes collègues. Merci à Isabelle Simunic, la magicienne du laboratoire qui sait absolument gérer toutes les situations, et Fatna Baoudj pour son efficacité et sa sympathie. Je salue également tous les doctorants et stagiaires, que je remercie pour nos discussions sérieuses, nos fous rires et surtout le plaisir j'ai eu à les côtoyer au quotidien.

A l'heure de partir vers de nouveaux horizons, je me remémore aussi les leçons que j'ai tirées au conservatoire, en classe préparatoire ou autour d'un échiquier, auprès de mes anciens professeurs Agnès Guérémy, Séverin L'Hostis et Roger Mansuy. Merci encore à eux, pour tous leurs conseils qui me servent à mûrir et à grandir, aujourd'hui encore.

Un grand merci à tous mes amis, aujourd'hui dispersés aux quatre coins du monde. C'est toujours un plaisir de les retrouver autour d'un verre. Je pense à eux, qui m'ont tant apporté.

Enfin, ces quelques derniers mots ne suffiront pas à traduire la reconnaissance et l'affection que j'ai pour les membres de ma famille et nos proches. Malgré les épreuves que nous avons traversées, leur optimisme et leur bienveillance sont pour moi une source constante de motivation et d'inspiration. Cette thèse leur est dédiée, car je leur dois beaucoup. Merci pour la force qu'ils me donnent et pour tous les moments futurs que nous partagerons.

Titre : Etude théorique et numérique de problèmes non linéaires au sens de McKean en finance

Résumé : Cette thèse est consacrée à l'étude théorique et numérique de deux problèmes non linéaires au sens de McKean en finance. Nous abordons dans la première partie le problème de calibration d'un modèle à volatilité locale et stochastique pour tenir compte des prix d'options Européennes observés sur le marché. Ce problème se traduit par l'étude d'une équation différentielle stochastique (EDS) non linéaire au sens de McKean à cause de la présence de l'espérance conditionnelle du coefficient de volatilité stochastique par rapport à la solution de l'EDS dans le coefficient de diffusion. Nous obtenons l'existence du processus dans le cas particulier où la fonction de volatilité stochastique est un processus de sauts ayant un nombre fini d'états. Dans l'industrie, la calibration est effectuée efficacement à l'aide d'une régularisation de l'espérance conditionnelle par un estimateur à noyau de type Nadaraya-Watson, comme proposé par Guyon et Henry-Labordère dans [57]. Nous obtenons la convergence faible à l'ordre 1 de la discrétisation en temps de l'EDS non linéaire au sens de McKean. Nous proposons également un schéma numérique demi-pas de temps et étudions le système de particules associé que nous comparons à l'algorithme proposé par [57]. Dans la deuxième partie de la thèse, nous nous intéressons à un problème de valorisation de contrat avec appels de marge, une problématique apparue avec l'application de nouvelles régulations depuis la crise financière de 2008. Ce problème peut être modélisé par une équation différentielle stochastique rétrograde (EDSR) anticipative avec dépendance en la loi de la solution dans le générateur. Nous montrons que cette équation est bien posée et proposons une approximation de sa solution à l'aide d'EDSR standards linéaires lorsque la durée de liquidation de l'option en cas de défaut est petite. Enfin, comme le calcul des solutions d'EDSR approchant la solution de l'EDSR non linéaire au sens de McKean peut être amélioré à l'aide de la méthode de Monte Carlo multiniveaux introduite par Giles dans [46], nous effectuons l'analyse numérique de cette technique.

Mots-clés : Calibration, modèle à volatilité locale et stochastique, système de particules en interaction, EDS non linéaire au sens de McKean, EDSR anticipative de McKean, Monte-Carlo multiniveaux.

Title : Theoretical and numerical study of problems nonlinear in the sense of McKean in finance

Summary : This thesis is dedicated to the theoretical and numerical study of two problems which are nonlinear in the sense of McKean in finance. In the first part, we study the calibration of a local and stochastic volatility model taking into account the prices of European options observed in the market. This problem can be rewritten as a stochastic differential equation (SDE) nonlinear in the sense of McKean, due to the conditional expectation of the stochastic volatility factor computed w.r.t. the solution of the SDE in the diffusion coefficient. We obtain existence in the particular case where the stochastic volatility factor is a jump process with a finite number of states. In the industry, Guyon and Henry-Labordère proposed in [57] an efficient calibration procedure which consists in approximating the conditional expectation using a kernel estimator such as the Nadaraya-Watson one. We obtain weak convergence at order 1 for the Euler scheme discretizing in the time the SDE nonlinear in the sense of McKean. We also introduce a numerical half-step scheme and study the the associated particle system that we compare with the algorithm presented in [57]. In the second part of the thesis, we tackle a problem of derivative pricing with initial margin requirements, a recent problem that appeared along with new regulation since the 2008 financial crisis. This problem can be modelled by an anticipative backward stochastic differential equation (BSDE) with dependence in the law of the solution in the driver. We show that the equation is well posed and propose an approximation of its solution by standard linear BSDEs when the liquidation duration in case of default is small. Finally, as the computation of the solutions of the linear BSDEs approximating the solution of the BSDE nonlinear in the sense of McKean can be improved thanks to the multilevel Monte-Carlo method introduced by Giles in [46], we perform the numerical analysis of this technique.

Keywords : Calibration, local and stochastic volatility model, interacting particle system, SDE nonlinear in the sense of McKean, anticipative McKean BSDE, multilevel Monte-Carlo.

Contents

Chapitre 1

Introduction

1.1 Champ moyen et équations de McKean

La théorie de champ moyen est une approche efficace qui permet de relier deux échelles de modélisations : une échelle microscopique décrivant le comportement de nombreuses *particules* en interaction et une échelle macroscopique pour laquelle on dispose d'un modèle simplifié qui décrit la distribution statistique de ces particules. A l'origine, les systèmes de particules étaient étudiés en physique pour modéliser l'évolution temporelle de la distribution de particules chargées dans un plasma (modèle de Vlasov) ou décrire l'évolution d'un gaz hors équilibre (modèle de Boltzmann). Depuis, le mot particule est un terme générique qui peut désigner des particules physiques, un neurone, un agent financier, etc.

Lorsque le nombre de particules devient très grand, la quantité d'interactions à prendre en compte est telle que la plupart des modélisations fines sont impossibles à simuler. Dans le cas de champs moyens, il est possible d'obtenir une excellente approximation des interactions entre particules. Dans ce cadre, nous nous plaçons dans la limite $N \to \infty$ où la contribution de chaque particule au champ est infinitésimale. Plus précisément, la limite de champ moyen est une limite particulière où les interactions entre particules sont suffisamment faibles pour que les influences s'exerçant sur une particule donnée restent finies lorsque $N \to \infty$.

Nous introduisons le modèle financier de Carmona, Fouque et Sun [26]. C'est un modèle simple qui décrit un système de N banques qui empruntent ou prêtent de l'argent les unes aux autres. Pour $1 \leq i \leq N$, la position de la particule i représente les fonds log-monétaires que possède la banque i. Le système de particules en interaction ci-dessous modélise l'évolution des réserves, suivant la règle selon laquelle pour $1 \leq j \neq i \leq N$, la banque i emprunte à la banque j si la banque j est plus riche qu'elle, et prête à la banque j dans le cas contraire :

$$dX_t^{i,N} = \left(\frac{1}{N} \sum_{j=1}^{N} X_t^{j,N} - X_t^{i,N} \right) dt + dB_t^i, \ 1 \leq i \leq N, \tag{1.1.1}$$

Ici $\left(B^i\right)_{1 \leq i \leq N}$ est une famille de mouvements browniens unidimensionnels et indépendents. De plus, nous supposons que les valeurs initiales $\left(X_0^{1,N}, ..., X_0^{N,N}\right)$ sont indépendantes et identiquement distribuées (i.i.d.) de loi μ_0, où μ_0 est une mesure de probabilité sur $\mathbb{R}$. Avec une heuristique de type *loi des grand nombres*, on s'attend à ce que la dynamique d'une particule dans la limite $N \to \infty$ suive

$$d\overline{X}_t = \left(\mathbb{E}\left[\overline{X}_t\right] - \overline{X}_t\right) dt + dB_t, \tag{1.1.2}$$

où B est un mouvement Brownien unidimensionnel.

L'équation (1.1.2) est une équation différentielle stochastique (EDS) dite de McKean, à cause du coefficient de dérive qui dépend de la loi de la solution $\overline{X}$ au travers de l'espérance $\mathbb{E}\left[\overline{X}_t\right]$. En supposant que la condition initiale μ_0 est centrée dans cet exemple très simple, on vérifie facilement que l'EDS (1.1.2) est bien posée et que $\overline{X}$ est égal à la solution de l'EDS d'Ornstein-Uhlenbeck

$$d\overline{X}_t = -\overline{X}_t dt + dB_t.$$

Il est possible de prouver des vitesses de convergence en loi du système de particules vers le processus limite (1.1.2). Dans son cours fondateur pour le cas d'un coefficient de dérive plus général, Sznitman couple la particule

$X^{i,N}$ avec le processus $\overline{X}^i$ qui est solution de (1.1.2) en remplaçant le mouvement brownien B par B^i et $\overline{X}_0^i = X_0^{i,N}$. En exploitant ce couplage, nous obtenons :

$$X_t^{i,N} - \overline{X}_t^i = \int_0^t \left(-X_s^{i,N} + \overline{X}_s^i + \frac{1}{N}\sum_{j=1}^N X_s^{j,N} - \mathbb{E}\left[\overline{X}_s^1\right] \right) ds.$$

En introduisant la moyenne $\frac{1}{N}\sum_{j=1}^N \overline{X}_s^j$, il vient pour $s \geq 0$,

$$\left|X_s^{i,N} - \overline{X}_s^i\right| \leq \int_0^s \left|X_u^{i,N} - \overline{X}_u^i\right| + \left|\frac{1}{N}\sum_{j=1}^N \left(X_u^{j,N} - \overline{X}_u^j\right)\right| + \left|\frac{1}{N}\sum_{j=1}^N \overline{X}_u^j - \mathbb{E}\left[\overline{X}_u^1\right]\right| du.$$

Nous prenons ensuite le supremum en $s \in [0,t]$ et l'espérance de l'inégalité ci-dessus et en utilisant l'échangeabilité des particules, nous obtenons :

$$\mathbb{E}\left[\sup_{0\leq s\leq t}\left|X_s^{i,N} - \overline{X}_s^i\right|\right] \leq 2\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s}\left|X_u^{i,N} - \overline{X}_u^i\right|\right] ds + \int_0^t \mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^N \overline{X}_s^j - \mathbb{E}\left[\overline{X}_s^1\right]\right|\right] ds.$$

Le dernier terme est un terme d'erreur statistique. Comme les processus $\overline{X}^j$ sont i.i.d., par l'inégalité de Cauchy-Schwarz, il vient :

$$\mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^N \overline{X}_s^j - \mathbb{E}\left[\overline{X}_s^1\right]\right|\right] \leq \sqrt{\frac{\mathbf{Var}(\overline{X}_s^1)}{N}},$$

où la variance est finie si la mesure μ_0 possède un moment d'ordre 2 fini. Par application du Lemme de Gronwall, on conclut que :

$$\sup_N \sqrt{N}\mathbb{E}\left[\sup_{t\leq T}|X_t^{i,N} - \overline{X}_t^i|\right] < \infty. \tag{1.1.3}$$

Par ailleurs, remarquons qu'avec l'observation $\frac{1}{N}\sum_{i=1}^N X^{i,N} = \frac{1}{N}\sum_{i=1}^N W_t^i$, nous obtenons directement que sous l'hypothèse $X_0^{i,N} = 0$ pour $1 \leq i \leq N$,

$$X_t^i - \overline{X}_t^i = \frac{1}{N}\sum_{j=1}^N B_t^j - \frac{1}{N}\sum_{j=1}^N \left(e^{-t}\int_0^t e^s dB_s^j\right),$$

et permettrait d'obtenir plus rapidement (1.1.3), mais la technique de couplage de Sznitman se transpose au cas d'interactions non linéaires entre particules

$$dX_t^{i,N} = \frac{1}{N}\sum_{i=1}^N b\left(X_t^{i,N}, X_t^{j,N}\right) dt + \frac{1}{N}\sum_{i=1}^N \sigma\left(X_t^{i,N}, X_t^{j,N}\right) dB_t^i,\ 1 \leq i \leq N,$$

où les noyaux d'interaction $b : \mathbb{R}^2 \to \mathbb{R}$ et $\sigma : \mathbb{R}^2 \to \mathbb{R}$ sont Lipschitz et bornés. Le processus limite suit alors la dynamique :

$$d\overline{X}_t = \left(\int_{\mathbb{R}} b\left(\overline{X}_t, y\right) u_t(dy)\right) dt + \left(\int_{\mathbb{R}} \sigma\left(\overline{X}_t, y\right) u_t(dy)\right) dB_t, \tag{1.1.4}$$

où pour $t \geq 0$, u_t est la loi de $\overline{X}_t$. Notons que le couplage utilisé précédemment repose sur le caractère bien posé du processus limite (1.1.4).

Theorem (Sznitman, [93])**.** *Il y a existence et unicité, trajectorielle et en loi pour les solutions de l'EDS (1.1.4).*

La preuve repose sur un argument de linéarisation puis un argument de point fixe sur l'espace des trajectoires. En effet, une itérée de l'opérateur Φ qui à une mesure m sur $C([0,T] \times \mathbb{R})$ associe la loi de la solution de l'EDS

$$Z_t = \left(\int_{\mathbb{R}} b\left(Z_t, y\right) m_t(dy)\right) dt + \left(\int_{\mathbb{R}} \sigma\left(Z_t, y\right) m_t(dy)\right) dB_t,$$

est une contraction par rapport à la distance de Wasserstein sur l'ensemble des mesures de probabilité sur $C([0,T] \times \mathbb{R})$.

Theorem (Sznitman, [93]). *Pour tout $i \geq 1$ et $T > 0$,*

$$\sup_N \sqrt{N}\mathbb{E}\left[\sup_{t\leq T} |X_t^{i,N} - \overline{X}_t^i|\right] < \infty. \tag{1.1.5}$$

Ces résultats d'existence, unicité et approximation particulaire ont été étendus pour des processus à sauts par Jourdain, Méléard et Woyczynski dans [70], où le processus $t \to (t, B_t)$ est remplacé par un processus de Lévy.

Comme conséquence du résultat de convergence (1.1.5), pour tout système fixé $i_1 < ... < i_k$, la loi de $\left(X^{i_1,N}, ..., X^{i_k,N}\right)$ converge étroitement vers $u^{\otimes k}$ lorsque $N \to \infty$. L'indépendance à l'instant initial des variables $X_0^{i_1,N}, ..., X_0^{i_k,N}$ se transmet aux instants ultérieurs et les processus $X^{i_1,N}, ..., X^{i_k,N}$ convergent vers des processus indépendants. Ce phénomène, appelé *propagation du chaos*, est équivalent à la convergence en probabilité de la mesure empirique du système de N particules vers la loi du processus non linéaire. L'intérêt de la méthode particulaire réside dans cette convergence : il est possible d'approcher la solution d'équations de McKean à l'aide de la mesure empirique du système de N particules.

Problèmes singuliers

Lorsque les coefficients des modèles considérés ne satisfont plus les conditions classiques de type Lipschitz, on parle alors de problème de McKean singulier. Il s'agit d'exploiter la structure particulière de chaque type interaction pour montrer la propagation du chaos. Sans prétendre à l'exhaustivité, nous présentons quelques problèmes singuliers qui ont fait l'objet de travaux. Pour chaque exemple, les fonctions b, σ, lorsqu'elles sont présentes, sont a priori différentes.

Pour des EDS de la forme

$$X_t = X_0 + \int_0^t b(u, X_u, \mathbb{E}[\varphi_1(X_u)])du + \int_0^t \sigma(u, X_u, \mathbb{E}[\varphi_2(X_u)])dW_u, \tag{1.1.6}$$

où W est un mouvement Brownien d-dimensionnel, σ, φ_2 sont des fonctions régulières et bornées, b est Lipschitz en la troisième coordonnée mais seulement bornée en espace et φ_1 est Hölderien, l'existence et l'unicité trajectorielle d'une solution de (1.1.6) sont établis dans [28] lorsque X_0 est de carré intégrable et σ est uniformément elliptique. La singularité vient du fait que la fonction β définie pour $t \geq 0$, $x \in \mathbb{R}^d$ et une mesure ν de probabilité sur $\mathbb{R}^d$ par $\beta(t, x, \nu) = b\left(t, x, \int \varphi_1 d\nu\right)$ est bornée en espace et Hölder en la mesure ν au sens de la distance de Wasserstein.

Une fonction u est solution de l'équation de Burgers avec condition initiale u_0 si elle vérifie

$$\begin{aligned} \partial_t u &= \frac{1}{2}\partial_x^2 u - u\partial_x u, \\ u_{|t=0} &= u_0. \end{aligned}$$

Contrairement aux équations à dérivées partielles (EDP) de Fokker-Planck associées aux EDS non linéaires au sens de McKean avec des coefficients Lipschitz ou de la forme (1.1.6), l'intéraction décrite par l'équation de Burgers est locale dans le terme de dérive d'une interprétation probabiliste qui serait, de façon heuristique,

$$X_t = X_0 + B_t + \int_0^t u(s, X_s)ds,$$

où $u(s, \cdot)$ est la densité de X_s. Dans [93], l'auteur décrit u comme la loi du processus limite associé au système de particules

$$dX_t^i = dB_t^i + \frac{1}{2N}\sum_{j\neq i} dL_t^0(X^i - X^j),$$

où L_t^0 est le temps local symétrique en 0 à l'instant t. En intégrant l'équation de Burgers en espace et en étudiant la fonction de répartition de X_t plutôt que se densité, Bossy et Talay se ramènent dans [16] et [17] à une interprétation probabiliste avec interaction globale mais où le noyau est la fonction de Heaviside, qui est discontinue. Dans [69], les auteurs montrent l'existence, l'unicité et la propagation du chaos pour un problème de McKean où l'interaction est de plus locale en le coefficient de diffusion

$$dX_t = b(p(s, X_s))dt + \sigma(p(s, X_s))dW_t,$$

où $p(s,\cdot)$ est la densité de X_s.

Le modèle de Keller-Segel en dimension d, utilisé en biologie, est un système d'EDP qui décrit les évolutions temporelles de la densité ρ_t d'une population de cellules et de la concentration c_t d'une substance chimique attractive :

$$\begin{aligned}
\partial_t \rho(t,x) &= \frac{1}{2}\Delta\rho(t,x) - \nabla\cdot(\chi\rho\nabla c)(t,x),\ t>0, x\in\mathbb{R}^d\\
\alpha\partial_t c(t,x) &= \frac{1}{2}\Delta c(t,x) - \lambda c(t,x) + \rho(t,x),\ t>0, x\in\mathbb{R}^d,\\
\rho(0,x) = \rho_0(x),\ & c(0,x) = c_0(x).
\end{aligned}$$

Les interprétations probabilistes du modèle de Keller-Segel font intervenir, dans le coefficient de dérive, la convolution en espace de la densité de la solution d'une EDS avec un noyau singulier. L'évolution temporelle d'une cellule peut être représentée dans le cas $\alpha = 0$ et $d = 2$, étudié dans [62], [41] et [27], par

$$X_t = X_0 + B_t + \chi\int_0^t K * f_s(X_s)ds \tag{1.1.7}$$

où $K : x \in \mathbb{R}^2 \to \frac{-x}{2\pi|x|^2}$ et f_s est la densité de X_s pour $s \geq 0$. Dans le cas $\alpha > 0$, étudié dans [94] et [66] pour $d = 1$, apparaît une singularité dans le coefficient de dérive en $\left(\int_0^s \tilde{K}_{s-u} * f_u du\right)(X_s)$, où $\tilde{K}_t(x) = \chi e^{-\frac{\lambda}{\alpha}t}\nabla\left(\frac{e^{-\frac{|x|^2}{2\alpha t}}}{(2\pi\alpha t)^{d/2}}\right)$.

Une approche probabiliste de l'EDP de Navier-Stokes en dimension 2, proposée par [80] et [78], basée sur l'étude du rotationnel de la solution, est très proche du modèle de Keller-Segel ci-dessus pour $\alpha = 0$, puisque c'est l'EDS (1.1.7) où la fonction K est le noyau de Biot et Savart.

Par ailleurs d'importantes avancées ont été faites dans le cas où des espérances conditionnelles interviennent dans les coefficients. Parmi les modèles d'écoulements turbulents, les solutions de l'équation de Burgers avec viscosité et les solutions statistiques de l'EDP de Navier Stokes en dimension 2 avec condition initiale aléatoire admettent la représentation probabiliste suivante de type McKean :

$$\begin{aligned}
dX_t &= \mathbb{E}\left[b(x,X_t)|\theta\right]\big|_{x=X_t} dt + \mathbb{E}\left[\sigma(x,X_t)|\theta\right]\big|_{x=X_t} dW_t, \qquad (1.1.8)\\
(X_0,\theta) &\sim [\Phi(a)](x)dx\nu(da),
\end{aligned}$$

où θ est une variable aléatoire indépendante du mouvement brownien W. Dans [96], les auteurs montrent que le problème est bien posé et proposent un système de particules avec des poids aléatoires pour évaluer les moments des solutions statistiques de l'EDP de McKean-Vlasov-Fokker-Planck associée à (1.1.8). Ce résultat est étendu à des conditions initiales plus générales dans [98], où l'auteur propose de plus un système de particules à l'aide de régresseurs par ondelettes au lieu d'estimateurs classiques de type Nadaraya-Watson, accélérant alors le calcul des moments.

Une autre façon de modéliser les écoulements turbulents est donnée par les modèles Lagrangiens stochastiques qui décrivent l'évolution temporelle de la position X et de la vitesse $\mathcal{U}$ des particules d'un fluide. Ces modèles font aussi intervenir une espérance conditionnelle dans le coefficient de dérive puisqu'ils se mettent sous la forme

$$\begin{aligned}
X_t &= X_0 + \int_0^t \mathcal{U}_s ds,\\
\mathcal{U}_t &= \mathcal{U}_0 + \int_0^t \mathbb{E}\left[b(\mathcal{U}_s,u)|X_s\right]_{u=\mathcal{U}_s} ds + \int_0^t \sigma(s,X_s,\mathcal{U}_s)dW_s,
\end{aligned}$$

où b est une fonction d'intéraction bornée. Le caractère bien posé de cette solution et la propagation du chaos du système de particules associé sont établis dans [15]. Ces résultats sont étendus au cas où le processus évolue dans un compact et est soumis à des réflections spéculaires dans [14].

Pour simuler des dynamiques moléculaires, le processus Adaptive Biasing Force (ABF) permet d'empêcher le processus $X = (X^1, ..., X^d)$ de stationner dans les états métastables grâce à la présence d'une force qui repousse X_t de son état actuel s'il y reste trop longtemps. Sous la forme la plus simple, le processus X est solution de l'EDS

$$dX_t = \left(-\nabla V(X_t) + \mathbb{E}\left[\partial_1 V(X_t)|X_t^1\right] e_1\right) dt + \sqrt{2}dW_t, \tag{1.1.9}$$

où W est un brownien d-dimensionnel, V est une fonction de potentiel, $\partial_1 V$ sa dérivée par rapport à la première coordonnée et $e_1 = (1, 0, ..., 0)$ est le premier vecteur de la base canonique. Cette force de répulsion est représentée par une espérance conditionnelle dans le terme de dérive et induit une forte non linéarité en la loi de X. Dans [75], les auteurs prouvent la convergence en temps long des solutions de (1.1.9) vers la mesure stationnaire. Plus tard, l'existence et l'unicité de la solution à l'EDS (1.1.9) sont établies dans [68].

Dans cette thèse nous nous intéressons à deux problèmes de McKean singuliers où apparaissent des espérances conditionnelles. Nous considérons dans la première partie une EDS dont le coefficient de diffusion dépend de l'espérance conditionnelle d'un facteur stochastique par rapport à la solution de l'EDS. Dans la deuxième partie, nous étudions une EDSR dont le générateur dépend de la loi de l'évolution future de la solution conditionnelle à la connaissance du passé.

Jeux à champ moyen

Terminons cette section avec une source de regain d'intérêt pour les problèmes de champ moyen. Si l'on donne de plus à chaque banque la possibilité d'emprunter à une banque centrale, la dynamique de chaque particule devient :

$$dX_t^{i,N} = \left(\frac{1}{N} \sum_{i=1}^{N} X_t^{j,N} - X_t^{i,N} \right) dt + \alpha_t^i dt + dB_t^i, \ 1 \le i \le N.$$

La banque i peut alors décider du taux α^i pour minimiser la fonctionnelle

$$J^i(\alpha^1, ..., \alpha^N) = \mathbb{E} \left[\int_0^T f_i(\tilde{X}_t, \alpha_t^i) dt + g_i(X_T^{i,N}) \right],$$

où $\tilde{X}_t = \frac{1}{N} \sum_{j=1}^N X_t^{j,N}$ avec par exemple les coûts $f_i(\tilde{X}_t, \alpha_t^i) = \frac{1}{2} \left(\alpha^i\right)^2 - \alpha^i \left(\tilde{X}_t - X_t^{i,N} \right) + \frac{1}{2} \left(\tilde{X}_t - X_t^{i,N} \right)^2$ et $g_i(X_T^{i,N}) = \frac{1}{2} \left(\tilde{X}_T - X_T^{i,N} \right)^2$. Ce problème est un jeu à champ moyen. Ces jeux ont connu un grand succès ces dernières années grâce aux travaux de Lasry et Lions [74]. Elles constituent une adaptation de la théorie du champ moyen, où les particules sont remplacées par des agents qui interagissent entre eux dans des situations stratégiques. La difficulté principale est qu'un agent doit, pour élaborer sa stratégie, prendre en compte celle des autres agents. Dans ce cas, la nature du champ moyen est modifiée. L'information statistique recherchée n'est plus la position ou la vitesse de chaque particule, mais la stratégie de chaque agent. Une analyse probabiliste de ces jeux a été menée dans [25] entre autres, et conduit à l'étude de systèmes couplés d'EDS Forward-Backward.

1.2 Les modèles à volatilité locale et stochastique calibrés aux prix vanilles de marché

Dans la première partie de la thèse, nous abordons la modélisation du cours d'un actif S par le processus d'Ito suivant :

$$dS_t = rS_t dt + \sigma_t S_t dW_t, \ t \ge 0,$$

porté par le mouvement Brownien unidimensionnel W, en supposant un taux d'intérêt constant $r \in \mathbb{R}$. Jusqu'au début des années 2000, trois principales familles de modèles ont été introduites dans l'industrie pour le processus σ : le modèle de Black-Scholes où σ est une constante, la volatilité locale (LV) où σ_t est une fonction déterministe de S_t, et la volatilité stochastique (SV) où σ est un processus d'Ito . Les deux premières familles sont complètes. En effet comme le prix de l'actif est dirigé par le seul mouvement brownien W pour chaque payoff, il existe un unique portefeuille autofinançant de réplication et donc le prix est défini comme étant la valeur au temps $t = 0$ du portefeuille de réplication. Notons que dans la communauté académique, suite au travail de Gatheral, Jaisson et Rosenbaum [44], des modèles où la volatilité dite rugueuse est construite à partir d'un mouvement brownien fractionnaire ont été développés récemment.

La calibration des modèles pour tenir compte des prix du marché des options est pertinente lorsque les risques portés par l'option exotique sont bien contrôlés par l'échange d'options liquides disponibles sur le marché. En pratique, on observe sur le marché les prix de produits liquides pour un ensemble fini de paramètres. On se restreint ici aux vanilles et donc on observe en pratique le prix des calls $C(T, K)$ et puts $P(T, K)$ sur un nombre fini de maturités T et de strikes K. Si on fait l'hypothèse supplémentaire que l'on a accès aux prix des vanilles pour tous les strikes et toutes les maturités positifs, d'après le résultat de [18], cette information est équivalente à la connaissance de la loi du sous-jacent S_t pour tout $t \geq 0$. D'après Dupire, le processus

$$dS_t^{Dup} = rS_t^{Dup}dt + \sigma_{Dup}(t, S_t^{Dup})S_t^{Dup}dW_t, \ t \geq 0,$$

est calibré de façon exacte aux valeurs du marché pour le choix :

$$\sigma_{Dup}(T,K) = \sqrt{2\frac{\partial_T C(T,K) + rK\partial_K C(T,K)}{K^2\partial_K^2 C(T,K)}}, \ T, K \geq 0,$$

au sens où

$$\forall T, K \geq 0, \ \mathbb{E}\left[e^{-rT}(S_T - K)_+\right] = C(T,K).$$

Cependant, une fois calibrés, les modèles LV ne possèdent plus de flexibilité supplémentaire pour contrôler par exemple le risque de volatilité. En effet, le coefficient de diffusion du modèle de Dupire ne dépend que de la trajectoire du mouvement Brownien W qui est la seule source d'aléa.

D'un autre côté, les modèles à volatilité stochastique permettent de mieux gérer le risque de volatilité et de reproduire des faits stylisés sur le comportement d'options plus complexes, mais ces modèles sont incomplets. Dans ce cas, la volatilité est dirigée par un ou plusieurs mouvements browniens qui ne sont en général pas parfaitement corrélés à W. Nous perdons alors la réplication exacte et l'unicité du prix. Plusieurs concepts, comme la surréplication, ont alors été développés pour définir le prix d'options sous un modèle SV. Pour résumer, les dynamiques des modèles SV ont un comportement spot-vol beaucoup plus riche que celui des modèles LV. Pour calibrer les modèles SV de façon exacte au smile de marché et combiner les points forts des deux modèles, un modèle à volatilité locale et stochastique (LSV) a été introduit dans [76] et [89] :

$$dS_t = rS_tdt + f(Y_t)\sigma(t, S_t)S_tdW_t, \ t \geq 0. \tag{1.2.1}$$

Dans ce modèle, la fonction σ, déterministe en le temps et S_t est multipliée par un facteur de volatilité stochastique $f(Y_t)$ où Y est un processus stochastique et $f : \mathbb{R} \to \mathbb{R}$ est une fonction mesurable strictement positive.

Lorsque la fonction f et le processus Y sont connus, la procédure de calibration repose sur le résultat suivant de Gyöngy. Etant donné un processus d'Ito Z solution de

$$dZ_t = \beta_t dt + \alpha_t dB_t, \ t \geq 0,$$

où β et α sont des processus adaptés à la filtration d'un mouvement Brownien B en dimension $d \in \mathbb{N}^*$, il est possible de construire une diffusion

$$dZ_t^D = b(t, Z_t^D)dt + \sigma(t, Z_t^D)dB_t, \ t \geq 0,$$

de telle sorte que pour tout $t \geq 0$, Z_t et Z_t^D aient la même loi. Soit I_d la matrice identité en dimension d et A^* la transposée d'une matrice A.

Theorem (Gyöngy, [58])**.** *Si les coefficients β, α sont bornés et s'il existe une constante $\alpha_{\inf} > 0$ telle que presque sûrement, dt-presque partout $\alpha_t\alpha_t^* > \alpha_{\inf}^2 I_d$ au sens des matrices symétriques, alors il existe une solution Z^D à l'EDS*

$$dZ_t^D = b(t, Z_t^D)dt + \sigma(t, Z_t^D)dB_t, \ t \geq 0,$$

où les coefficients b, σ sont définis pour tout $t \geq 0, z \in \mathbb{R}^d$ par

$$\begin{aligned} b(t,z) &= \mathbb{E}\left[\beta_t | Z_t = z\right], \\ \sigma\sigma^*(t,z) &= \mathbb{E}\left[\alpha_t\alpha_t^* | Z_t = z\right], \end{aligned} \tag{1.2.2}$$

De plus, pour tout $t \geq 0$, Z_t a la même loi que Z_t^D.

En appliquant la Définition (1.2.2) avec le choix $Z = S$, défini par (1.2.1) et $Z^D = S^{Dup}$, on obtient formellement que pour tout $t \geq 0$, S_t et S_t^{Dup} ont la même loi si presque sûrement, dt-presque partout,

$$\mathbb{E}\left[\sigma^2(t, S_t)f^2(Y_t)|S_t\right] = \sigma_{Dup}^2(t, S_t).$$

La calibration du modèle LSV aux prix des options vanilles donnés par le marché impose donc le choix suivant pour la fonction σ :

$$\sigma(t,x) = \frac{\sigma_{Dup}(t,x)}{\sqrt{\mathbb{E}\left[f^2(Y_t)|S_t = x\right]}}, \ t \geq 0, x \in \mathbb{R},$$

et donc la dynamique suivante pour $X = \log(S)$:

$$\begin{aligned} dX_t &= \left(r - \frac{\sigma^2_{Dup}(t,e^{X_t})}{2}\frac{f^2(Y_t)}{\mathbb{E}\left[f^2(Y_t)|X_t\right]}\right)dt + \sigma_{Dup}(t,e^{X_t})\frac{f(Y_t)}{\sqrt{\mathbb{E}\left[f^2(Y_t)|X_t\right]}}dW_t, \\ \log(S_0) &= X_0, \end{aligned} \tag{1.2.3}$$

où X_0 est déterministe. L'équation (1.2.3) est nonlinéaire au sens de McKean à cause de la présence de l'espérance conditionnelle dans le dénominateur des coefficients de la dynamique et qui est hautement non linéaire en la loi de (X,Y). Comme les hypothèses Lipschitz standards ne sont pas vérifiées, les résultats classiques évoqués précédemment ne s'appliquent pas. La première question naturelle qui se pose concerne le caractère bien posé de l'équation (1.2.3).

Question : A-t-on existence et unicité d'une solution de l'EDS non linéaire au sens de McKean ?

Par ailleurs, la question de la simulation est également importante d'un point de vue pratique pour valoriser des contrats à l'aide du modèle LSV calibré. Ren Madan et Qian ont proposé une première approche dans [90]. Leur idée consiste dans un premier temps à résoudre numériquement l'équation de Fokker Planck associée au couple (X,Y) avant d'utiliser un schéma d'Euler pour diffuser le processus (X,Y) en approchant l'espérance conditionnelle à l'aide de l'approximation de la densité de (X_t,Y_t) obtenue aux instants de discrétisation temporelle. Un peu plus tard, en s'inspirant du point de vue champ moyen, Guyon et Henry-Labordère ont présenté dans [57] une méthode plus directe et intuitive en approchant l'espérance conditionnelle à l'aide d'un estimateur à noyau de type Nadaraya-Watson dans un système de N particules en interaction :

$$\mathbb{E}\left[f^2(Y_t)|X_t = x\right] \approx \frac{\frac{1}{N}\sum_{i=}^N f^2(Y_t^{i,N})K_\epsilon(X_t^{i,N} - x)}{\frac{1}{N}\sum_{i=}^N K_\epsilon(X_t^{i,N} - x)},$$

où K_ϵ est une fonction de régularisation de taille de fenêtre $\epsilon > 0$. Cette procédure efficace est désormais standard dans la plupart des banques, notamment sur le marché du change. La calibration s'opère avec les données de marché et la méthode particulaire converge sur une grande plage de paramètres, sauf quand le risque de volatilité est poussé à l'extrême. Cela est peut être dû à l'absence de solution pour le processus limite (1.2.3) ou bien à l'instabilité numérique sous ces jeux de paramètres.

Question : La méthode particulaire de calibration du modèle LSV converge-t-elle ? A quelle vitesse ?

Enfin, notons que la méthodologie de calibration introduite dans [57] se généralise dans [55] et [56] aux cas où la volatilité σ dépend de la trajectoire passée du cours de l'actif S ou lorsque l'on considère des options sur indices ou devises au lieu d'options vanilles dans le cadre de modèles multi-actifs.

Chapitre 2 : Existence d'un modèle à volatilité locale et changement de régime calibré et nouveaux faux mouvements browniens

Les premières réponses à la question de l'existence du processus limite solution de l'EDS (1.2.3) ont été apportées par Abergel et Tachet dans [1]. Les auteurs établissent l'existence locale en temps d'une solution à l'équation de Fokker-Planck restreinte à un compact en espace, en supposant que le facteur de volatilité stochastique est presque constant, ce qui fait du modèle LSV calibré obtenu une perturbation du modèle de Dupire.

Une des situations les plus simples pour aborder le problème d'existence globale pour l'EDS (1.2.3) est de se placer d'abord dans le cas où Y est un processus à valeurs dans l'ensemble fini de réels $\mathcal{Y} = \{y_1, ..., y_d\}$. Dans un premier temps, en choisissant $\sigma_{Dup} = 1$ et en négligeant le terme de dérive, nous nous ramenons à l'étude de l'EDS suivante :

$$dX_t = \frac{f(Y_t)}{\sqrt{\mathbb{E}\left[f^2(Y_t)|X_t\right]}}dW_t,$$

$$X_0 \sim \mu. \tag{1.2.4}$$

où μ est une mesure de probabilité sur $\mathbb{R}$. Pour obtenir l'existence du processus, nous considérons les densités $(p_i)_{1\leq i\leq d}$, où pour $1 \leq i \leq d$, $p_i(t,\cdot)$ est la densité de X_t conditionnellement à l'évènement $\{Y_t = y_i\}$ et multipliée par $\alpha_i := \mathbb{P}(Y_t = y_i)$. En d'autres termes, p_i est définie pour toute fonction mesurable bornée $\varphi : \mathbb{R} \to \mathbb{R}$ par

$$\mathbb{E}\left[\varphi(X_t)1_{\{Y_t=y_i\}}\right] = \int_{\mathbb{R}} \varphi(x)p_i(t,x)dx.$$

Dans le cas où Y est un processus constant en temps et la Condition (C) ci-dessous est satisfaite, nous établissons l'existence d'une solution au sens des distributions au système de Fokker-Planck vérifié par la famille $(p_i)_{1\leq i\leq d}$:

$$\begin{aligned} &\partial_t p_i = \frac{1}{2}\partial^2_{xx}\left(\frac{\sum_{k=1}^d p_k}{\sum_{k=1}^d f^2(y_k)p_k} f^2(y_i)p_i\right) \text{ dans } (0,T)\times\mathbb{R} \\ &p_i(0) = \alpha_i\mu \text{ dans } \mathbb{R}, \end{aligned} \tag{1.2.5}$$

Condition (C)**.** Il existe une matrice Γ symétrique définie positive de taille $d\times d$ telle que pour $1\leq k\leq d$, la matrice $\Gamma^{(k)}$ de taille $d\times d$ avec coefficients

$$\Gamma^{(k)}_{ij} = \frac{f^2(y_i)+f^2(y_j)}{2}(\Gamma_{ij}+\Gamma_{kk}-\Gamma_{ik}-\Gamma_{jk}),\ 1\leq i,j\leq k,$$

est définie positive sur l'espace vectoriel $\{y = (y_1,...,y_d) \in \mathbb{R}^d, y_k = 0\}$.

La Condition (C) assure la coercivité du système d'équations aux dérivées partielles (EDP) (1.2.5) et permet d'obtenir la stabilité des estimées d'énergie sur la famille $(p_i)_{1\leq i\leq d}$. Pour obtenir l'existence d'un processus solution du système (1.2.4), nous procédons en trois étapes.

1. Nous appliquons une méthode de Galerkine pour assurer l'existence d'une solution du système (1.2.5) au sens des distributions lorsque la condition initiale X_0 possède des moments d'ordre 2.
2. Pour une loi initiale quelconque, nous procédons par régularisation pour pouvoir appliquer le résultat de la première étape. En faisant tendre le paramètre de régularisation vers 0, nous extrayons une suite dont la limite est une solution du système (1.2.5) au sens des distributions.
3. D'après le résultat de Figalli [39, Theorem 2.6], nous faisons ensuite le lien entre l'existence d'une solution au système (1.2.5) et l'existence d'une solution faible à l'EDS (1.2.4).

Theorem. *Si la Condition (C) est vérifiée, alors il existe une solution faible à l'EDS (1.2.4). De plus, pour tout $t\geq 0$, X_t a la même loi que $M+W_t$, où M est une variable aléatoire de loi μ indépendante de W.*

Remarquons que si la mesure μ est le Dirac au point 0, alors le processus X est un *faux* mouvement Brownien, c'est-à-dire que pour tout $t\geq 0$, X_t suit la loi normale centrée de variance t mais X n'est pas un mouvement Brownien.

Une fois obtenue l'existence du processus solution de l'EDS (1.2.3) lorsque Y est constant en temps, il est possible de généraliser le résultat précédent en ajoutant une fonction de volatilité locale, un terme de dérive et en se plaçant dans le cas où Y est un processus de sauts prenant ses valeurs dans $\mathcal{Y}$. Le système de Fokker-Planck devient alors :

$$\partial_t p_i = \frac{1}{2}\partial^2_{xx}\left(\frac{\tilde\sigma^2_{Dup}\sum_{i=1}^d p_k}{\sum_{i=1}^d f^2(y_k)p_k} f^2(y_i)p_i\right) - \partial_x\left(\left[r - \frac{1}{2}\frac{f^2(y_i)\sum_{i=1}^d p_k}{\sum_{i=1}^d f^2(y_k)p_k}\tilde\sigma^2_{Dup}\right]p_i\right) + \sum_{j=1}^d q_{ji}p_j,$$

où $\tilde\sigma^2_{Dup}(t,x) = \sigma_{Dup}(t,e^x)$ pour $t\geq 0, x\in\mathbb{R}$. Du moment que les intensités de sauts sont bornées et que la fonction de volatilité locale de Dupire possède un peu de régularité et de l'ellipticité, le premier terme du membre de droite est similaire à celui que l'on retrouve dans (1.2.5) et les termes suppplémentaires qui apparaissent ne posent pas de difficulté pour la stabilité des estimées d'énergies. Sous la Condition (C), une méthodologie similaire et une généralisation du résultat de Figalli permettent alors d'obtenir l'existence du processus calibré.

Theorem. *Si la Condition (C) est vérifiée et sous les hypothèses du Théorème 2.2.2, il existe une solution faible à l'EDS (1.2.3). De plus, pour tout $t\geq 0$, X_t a la même loi que $X_t^{Dup} := \log\left(S_t^{Dup}\right)$.*

Remarquons que la méthodologie utilisée pour obtenir l'existence lorsque Y est un processus de sauts avec un nombre fini d'états ne nous permet pas de traiter le cas plus utilisé dans l'industrie où Y est un processus d'Ito.

Chapitre 3 : Discrétisation d'une classe de diffusions non linéaires au sens de McKean incluant le modèle à volatilité locale et stochastique calibré

Nous répondons à la question suivante.

Question : A quelle vitesse converge le schéma d'Euler discrétisant en temps l'EDS (1.2.3) dans le cadre général où Y est une diffusion d'Ito ?

Nous introduisons un cadre général dans lequel s'inscrit la discrétisation en temps de l'EDS (1.2.3). Pour $k \geq 1$, nous notons $\mathcal{S}_k^+(\mathbb{R})$ l'espace des matrices symétriques positives de taille $k \times k$. Soit $d_1 \geq 1$ et Z une solution de l'EDS

$$\begin{aligned} dZ_t &= b(t, Z_t)\,dt + \sigma(t, Z_t)\,dB_t, \\ Z_0 &\sim \mu_{Z_0}. \end{aligned} \tag{1.2.6}$$

Ici B est un mouvement Brownien d_1-dimensionnel, μ_{Z_0} est une mesure de probabilité sur $\mathbb{R}^{d_1}$, $b : [0,\infty)\times\mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$ et $\sigma : [0,\infty)\times\mathbb{R}^{d_1} \to \mathcal{S}_{d_1}^+(\mathbb{R})$ sont des fonctions mesurables. Soient $d_2, q \geq 2$, et (X, Y) une solution de l'EDS

$$\begin{aligned} dX_t &= b_X(t, X_t, Y_t, \mathbb{E}[\phi(X_t, Y_t)|X_t])\,dt + \sigma_X(t, X_t, Y_t, \mathbb{E}[\phi(X_t, Y_t)|X_t])\,dW_t^1, \\ dY_t &= b_Y(t, X_t, Y_t)\,dt + \sigma_Y(t, X_t, Y_t)\,dW_t^2, \\ (X_0, Y_0) &\sim \mu_0, \end{aligned} \tag{1.2.7}$$

où $b_X : [0,\infty)\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\times\mathbb{R}^q \to \mathbb{R}^{d_1}$, $\sigma_X : [0,\infty)\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\times\mathbb{R}^q \to \mathcal{S}_{d_1}^+(\mathbb{R})$, $b_Y : [0,\infty)\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$, $\sigma_Y : [0,\infty)\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathcal{S}_{d_2}^+(\mathbb{R})$ et $\phi : \mathbb{R}^{d_1}\times\mathbb{R}^{d_2} \to \mathbb{R}^q$ sont des fonctions mesurables, W^1 (resp. W^2) est un mouvement Brownien d_1-dimensionnel (resp. d_2-dimensionnel), μ_0 est une mesure de probabilité sur $\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$ telle que son image par les d_1 première coordonnées est égale à μ_{Z_0}, et nous notons également $a_X := \sigma_X\sigma_X^*$, $a := \sigma\sigma^*$. De façon analogue à la projection markovienne (1.2.2), nous supposons que le couple (b_X, a_X) possède la propriété suivante : pour tout couple de variables aléatoires (A, B) à valeurs dans $\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$, vérifiant

$$\mathbb{E}[|\phi(A,B)|] < \infty, \text{ et } \forall t \geq 0, \quad \mathbb{E}[|b_X(t, A, B, \mathbb{E}[\phi(A,B)|A])|] < \infty, \quad \mathbb{E}[||a_X(t, A, B, \mathbb{E}[\phi(A,B)|A])||] < \infty,$$

nous avons presque sûrement

$$\forall t \geq 0, \; \mathbb{E}[b_X(t, A, B, \mathbb{E}[\phi(A,B)|A])|A] = b(t, A), \quad \mathbb{E}[a_X(t, A, B, \mathbb{E}[\phi(A,B)|A])|A] = a(t, A). \tag{1.2.8}$$

Il est facile dans ce cas de vérifier que le modèle LSV calibré correspond à la situation où pour $d \geq 2$ et $(t, x, y, z) \in [0,T]\times\mathbb{R}\times\mathbb{R}^{d-1}\times\mathbb{R}$,

$$b_X(t,x,y,z) = r - \frac{1}{2}\tilde{\sigma}_{Dup}^2(t,x)\frac{f^2(y)}{z}, \qquad \phi(x,y) = f^2(y), \qquad \sigma_X(t,x,y,z) = \tilde{\sigma}_{Dup}(t,x)\frac{f(y)}{\sqrt{z}}$$
$$b(t,x) = r - \frac{1}{2}\tilde{\sigma}_{Dup}^2(t,x), \quad \sigma(t,x) = \tilde{\sigma}_{Dup}(t,x),$$

et par exemple les fonctions b_Y, σ_Y sont globalement Lischitz et ne dépendent pas de X, de façon à ce que Y soit une diffusion bien posée. Pour un horizon fini $T > 0$, nous étudions dans un premier temps l'erreur faible entre Z_T and la composante X_T^n, $n \geq 1$, du schéma d'Euler explicite avec pas de temps constant $\Delta = \frac{T}{n}$ associé à (X, Y). Pour $n \in \mathbb{N}^*$, ce schéma d'Euler est donné par

$$\begin{aligned} dX_t^n &= b_X\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n, \mathbb{E}\left[\phi\left(X_{\tau_t}^n, Y_{\tau_t}^n\right)|X_{\tau_t}^n\right]\right)dt + \sigma_X\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n, \mathbb{E}\left[\phi\left(X_{\tau_t}^n, Y_{\tau_t}^n\right)|X_{\tau_t}^n\right]\right)dW_t^1, \\ dY_t^n &= b_Y\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n\right)dt + \sigma_Y\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n\right)dW_t^2, \\ (X_0^n, Y_0^n) &\sim \mu_0, \end{aligned} \tag{1.2.9}$$

où pour $t \in [0,T]$, $\tau_t := \lfloor \frac{nt}{T} \rfloor \frac{T}{n}$ est l'instant de discrétisation le plus proche de t par valeurs inférieures. Contrairement au problème de l'existence d'une solution des EDS (1.2.3) ou (1.2.7) qui est difficile, la discrétisation en temps à n pas égaux est bien définie dès que l'on parvient à contrôler les moments d'ordre 1 des coefficients de dérive et de diffusion

$$b_X\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n, \mathbb{E}\left[\phi\left(X_{\tau_t}^n, Y_{\tau_t}^n\right)|X_{\tau_t}^n\right]\right), \; \sigma_X\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n, \mathbb{E}\left[\phi\left(X_{\tau_t}^n, Y_{\tau_t}^n\right)|X_{\tau_t}^n\right]\right).$$

De plus, nous supposons que les coefficients b, σ intervenant dans la diffusion de Dupire sont suffisament réguliers pour que le processus Z soit bien défini. Dans ce cas, étant donnée une fonction test $\varphi : \mathbb{R}^{d_1} \to \mathbb{R}$, une technique classique pour estimer l'erreur faible

$$\mathbb{E}\left[\varphi(X_T^n)\right] - \mathbb{E}\left[\varphi(Z_T)\right],$$

est la méthode de Talay et Tubaro [95], où l'on considère la représentation de Feynman-Kac pour le processus Z, définie pour $t \in [0,T], x \in \mathbb{R}$ par

$$u(t,x) = \mathbb{E}\left[\varphi(Z_T)|Z_t = x\right].$$

L'erreur faible peut alors se réécrire comme une somme téléscopique

$$\mathbb{E}\left[\varphi(X_T^n)\right] - \mathbb{E}\left[\varphi(Z_T)\right] = \sum_{k=0}^{n-1} \mathbb{E}\left[u(t_{k+1}, X_{t_{k+1}}^n) - u(t_k, X_{t_k}^n)\right].$$

Il suffit alors de contrôler chaque terme de la somme. En appliquant le lemme d'Ito au processus $t \to u(t, X_t^n)$ sur chaque intervalle (t_k, t_{k+1}), nous exploitons la condition de structure (1.2.8) pour éliminer le terme d'ordre le plus bas. Il suffit ensuite de contrôler en $O(\frac{1}{n^2})$ les termes d'ordre supérieur. S'il ne semble pas raisonnable d'obtenir un développement de l'erreur faible, le résultat ci-dessous peut être néanmoins rapproché du résultat original de vitesse faible de convergence de Talay et Tubaro [95].

Theorem. *Sous les hypothèses du Théorème 3.2.1, il existe une constante C telle que*

$$\forall n \geq 1, \ \left|\mathbb{E}\left[\varphi(X_T^n)\right] - \mathbb{E}\left[\varphi(Z_T)\right]\right| \leq \frac{C}{n}.$$

Dans un contexte financier et un cadre unidimensionnel, on peut aussi s'intéresser à des payoffs moins réguliers comme le put et dans ce cas l'estimation de l'erreur faible est un peu dégradée, d'un facteur logarithmique.

Theorem. *Sous les hypothèses du Théorème 3.2.2, pour tout $K > 0$, il existe une constante C_P telle que*

$$\forall n \geq 1, \ \left|\mathbb{E}\left[(K - e^{X_T^n})_+\right] - \mathbb{E}\left[(K - e^{Z_T})_+\right]\right| \leq C_P \frac{\log(n)}{n}.$$

Pour estimer l'erreur faible dans le cas du put, l'astuce est la même que dans le cas régulier, mais les autres termes sont contrôlés différemment en utilisant des estimées de type d'Aronson pour la densité X_t^n ainsi que des estimées gaussiennes pour les dérivées spatiales de u.

En l'état, les schémas d'Euler discrétisant en temps ne sont pas directement simulables, à cause de l'espérance conditionnelle toujours présente. Sous l'hypothèse d'ellipticité du coefficient de diffusion, nous introduisons ensuite un schéma numérique demi-pas de temps qui permet d'obtenir une représentation de l'espérance conditionnelle comme un ratio de convolutions contre la densité gaussienne sans utiliser d'approximation de l'espérance conditionnelle par noyau. Nous introduisons ensuite une méthode de particules associée à ce schéma demi-pas de temps en remplaçant la loi du processus demi-pas de temps par la loi empirique du système de particules. Enfin nous donnons des estimées de convergence forte de la méthode de particules et nous comparons numériquement les résultats obtenus avec ceux de [57].

1.3 Correction de prix de contrats en présence d'appels de marge

Depuis la crise de 2008, l'approche de la valorisation des contrats financiers a peu à peu changé, pour prendre en compte les nouvelles régulations. De nos jours, les banques et les institutions financières doivent déposer une somme d'argent, appelée collatéral, chez une chambre de compensation (Central CounterParty, ou CCP) pour assurer leurs positions. Tous les jours, la CCP demande à chaque membre un dépôt dont la somme dépend de son exposition au risque de ses contrats de gré à gré. En particulier, la marge initiale correspond à un collatéral qui couvre en cas de défaut, une perte de valeur du contrat à sa liquidation. Nous étudions dans cette partie comment les appels de marge affectent la valorisation et la couverture de contrats dans le cas particulier rencontré en industrie financière où le collatéral est proportionnel à la Value-at-Risk conditionnelle (CVaR) du prix du contrat sur une courte période de temps, typiquement de l'ordre de 10 jours.

Nous introduisons un modèle simple où la présence d'appels de marge induit un terme non linéaire au sens de McKean dans l'EDS rétrograde (EDSR) de valorisation et de couverture. Supposons que le prix de l'actif risqué noté S évolue selon un mouvement Brownien géométrique

$$dS_t = \mu S_t dt + \sigma S_t dW_t, t \geq 0,$$

où $(\mu, \sigma) \in \mathbb{R}\times\mathbb{R}_*^+$ et W est un mouvement brownien unidimensionnel. Nous considérons la situation où un trader souhaite vendre une option européenne de maturité $T > 0$ et de payoff $\Phi(S_T)$ et se couvrir dynamiquement avec les actifs risqué S et sans risque S^0, où $S_t^0 = e^{rt}$ pour $t \in [0, T]$ et r est un taux d'intérêt sans risque. Dans le cadre financier classique (voir par exemple [83]), en notant V la valeur du portefeuille de converture autofinançant et π la somme d'argent investie dans l'actif risqué, le couple (V, π) satisfait l'équation stochastique suivante :

$$\begin{aligned} dV_t &= r(V_t - \pi_t)dt + \pi_t \frac{dS_t}{S_t} = rV_t dt + (\mu - r)\pi_t dt + \sigma\pi_t dW_t,\ t \in [0, T], \\ V_T &= \Phi(S_T). \end{aligned} \tag{1.3.1}$$

L'équation (1.3.1) est une EDSR à cause de la condition terminale imposée sur V, qui de plus est linéaire car ses coefficients sont linéaires en (V, π) (voir [72] pour un vaste aperçu sur les EDSR et leurs applications en finance).

Nous introduisons à présent un modèle simple prenant en compte les appels de marge. Cette contrainte induit un coût supplémentaire au portefeuille autofinançant de couverture. Dans ce modèle, nous supposons que le montant à déposer est proportionnel à la CVaR du portefeuille sur une durée Δ (typiquement 10 jours) au niveau de risque α (par exemple $\alpha = 99\%$). Le coût de ce dépôt est fixé par un taux d'intérêt R. Nous modélisons ainsi la contribution des appels de marge par un terme supplémentaire dans la dynamique du portefeuille autofinançant

$$dV_t = r(V_t - \pi_t)dt + \pi_t \frac{dS_t}{S_t} - R\mathbf{CVaR}^{\alpha}_{\mathcal{F}_t}(V_t - V_{(t+\Delta)\wedge T})dt,\ t \in [0, T],$$

où la CVaR d'une variable aléatoire L conditionnée à la filtration $\mathcal{F}$ du mouvement Brownien W est définie à l'instant t par

$$\mathbf{CVaR}^{\alpha}_{\mathcal{F}_t}(L) = \inf_{x\in\mathbb{R}} \mathbb{E}\left[\frac{(L-x)^+}{1-\alpha} + x \,\middle|\, \mathcal{F}_t \right].$$

En réécrivant l'EDSR sous forme intégrale, nous obtenons :

$$V_t = \Phi(S_T) + \int_t^T \left(-r(V_s - \pi_s) - \mu\pi_s + R\mathbf{CVaR}^{\alpha}_{\mathcal{F}_s}(V_s - V_{(s+\Delta)\wedge T})\right) ds - \int_t^T \pi_s \sigma dW_s,\ t \in [0, T]. \tag{1.3.2}$$

Le terme de CVaR conditionnelle est non linéaire au sens de McKean et porte sur la loi de l'évolution future du portefeuille conditionnelle à la connaissance du passé. Remarquons que dans ce modèle, (V, π) suit une règle endogène de valorisation au sens où la CVaR est calculée sur la solution V elle même.

Chapitre 4 : Approximation numérique d'EDSR anticipatives de McKean pour la valorisation d'options en présence d'appel de marge

Nous étudions d'abord l'EDSR anticipative nonlinéaire au sens de McKean. Nous introduisons un cadre général incluant l'EDSR (1.3.2) et nous montrons que le problème est bien posé. Nous introduisons les espaces classiques : pour $\beta \ge 0$ et $d \in \mathbb{N}^*$, $\mathbb{H}^2_{\beta,T}(\mathbb{R}^d)$ est l'espace des processus φ adaptés à valeur dans $\mathbb{R}^d$ et tels que $\mathbb{E}\left[\int_0^T e^{\beta t}|\varphi_t^2| dt\right] < \infty$ et $\mathbb{S}^2_{\beta,T}(\mathbb{R}^d)$ est l'espace des processus $\varphi \in \mathbb{H}^2_{\beta,T}(\mathbb{R}^d)$ tels que $\mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\varphi_t^2|\right] < \infty$. Soit W un mouvement Brownien d-dimensionnel, $\mathcal{F}$ la filtration de associée à W et ξ une variable aléatoire $\mathcal{F}_T$-mesurable. Il s'agit de trouver le couple $(Y, Z) \in \mathbb{S}^2_{0,T}(\mathbb{R}) \times \mathbb{H}^2_{0,T}(\mathbb{R}^d)$ tel que

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, \Lambda_s(Y_{s:T}))ds - \int_t^T Z_s dW_s,\ t \in [0, T], \tag{1.3.3}$$

où pour $(y, z, \lambda) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$, $f(\cdot, y, z, \lambda)$ est un processus adapté à valeurs dans $\mathbb{R}$ et pour $X \in \mathbb{S}^2_{0,T}(\mathbb{R})$, $(\Lambda_t(X_{t:T}))_{t\in[0,T]}$ est un processus adapté à valeurs dans $\mathbb{R}$.

Theorem. *Sous les hypothèses du Théorème 4.2.1, pour toute condition terminale de carré intégrable et $\mathcal{F}_T$-mesurable ξ, il existe une unique solution $(Y, Z) \in \mathbb{S}^2_{0,T}(\mathbb{R}) \times \mathbb{H}^2_{0,T}(\mathbb{R}^d)$ à l'EDSR (1.3.3).*

La preuve repose sur des arguments classiques. Dans le même esprit que [72] sur les solutions d'une EDSR, nous calculons des estimées a priori du couple (Y, Z). Ensuite, pour $\beta > 0$ bien choisi, nous utilisons la méthode de point fixe de Picard dans l'espace $\mathbb{S}^2_{\beta,T}(\mathbb{R}) \times \mathbb{H}^2_{\beta,T}(\mathbb{R}^d)$. Il suffit ensuite de vérifier que le terme non linéaire au sens de McKean avec la CVaR satisfait les hypothèses du théorème.

Corollary. *Pour toute condition terminale de carré intégrable ξ et $\mathcal{F}_T$-mesurable, le problème (1.3.2) d'appel de marge basé sur la CVaR est bien posé.*

A cause de la dépendance en loi, résoudre le problème (1.3.2) numériquement semble plus difficile que résoudre une EDSR standard à l'aide de techniques par exemple de Monte-Carlo avec régression comme dans [51]. Lorsque la durée de liquidation Δ est petite, en utilisant l'approximation à l'ordre le plus faible en $\sqrt{\Delta}$, nous avons que

$$V_s - V_{(s+\Delta)\wedge T} \approx -\int_s^{(s+\Delta)\wedge T} Z_s dW_u \overset{(d)}{=} -|Z_s|\sqrt{(s+\Delta)\wedge T - s} \times G,$$

où G est une gaussienne standard indépendente de $\mathcal{F}_s$. En notant la constante $C_\alpha = \mathbf{CVaR}^\alpha(G)$, nous obtenons une EDSR standard non linéaire

$$V_t^{NL} = \xi + \int_t^T \left(rV_s^{NL} + Z_s^{NL}\sigma^{-1}(r-\mu) + RC_\alpha\sqrt{(s+\Delta)\wedge T - s}|Z_s^{NL}|ds \right) - \int_t^T Z_s^{NL} dW_s, \ t \in [0,T],$$

du fait de la présence de la valeur absolue dans le générateur. En interprétant cette EDSR comme une équation paramétrisée par Δ et en suivant la méthodologie proposée dans [50], il est possible de linéariser l'EDSR standard non linéaire à l'aide d'un développement limité en Δ au voisinage de 0. Nous obtenons alors deux EDSR linéaires aux ordres 0 et 1 respectivement

$$\begin{aligned} V_t^{BS} =& \xi + \int_t^T \left(-rV_s^{BS} + Z_s^{BS}\sigma^{-1}(r-\mu) \right) ds - \int_t^T Z_s^{BS} dW_s, \ t \in [0,T]. \\ V_t^{L} =& \xi + \int_t^T \left(-rV_s^{L} + Z_s^{L}\sigma^{-1}(r-\mu) + RC_\alpha\sqrt{(s+\Delta)\wedge T - s}|Z_s^{BS}| \right) ds - \int_t^T Z_s^{NL} dW_s, \ t \in [0,T]. \end{aligned}$$

Theorem. *Sous les hypothèses du Théorème 4.3.1, nous obtenons que*

$$\begin{aligned} ||V^L - V^{BS}||^2_{\mathbb{S}^2_{0,T}} + ||Z^L - Z^{BS}||^2_{\mathbb{H}^2_{0,T}} &= \mathcal{O}(\Delta), \\ ||V^{NL} - V^{L}||^2_{\mathbb{S}^2_{0,T}} + ||Z^{NL} - Z^{L}||^2_{\mathbb{H}^2_{0,T}} &= \mathcal{O}(\Delta^2), \\ ||V - V^{NL}||^2_{\mathbb{S}^2_{0,T}} + ||Z - Z^{NL}||^2_{\mathbb{H}^2_{0,T}} &= \mathcal{O}(\Delta^2). \end{aligned}$$

Ces estimations illustrent le fait qu'il y a une différence d'ordre $\sqrt{\Delta}$ entre les EDSR satisfaites par les valeurs respectives V et V^{BS} des portefeuilles avec et sans appel de marge. En revanche, les différences entre V^L, V^{NL} et V sont d'ordre supérieur, en Δ. Ces estimées sont généralisées dans le Théorème 4.3.1, dans un cadre multidimensionnel où de plus les coefficients de dérive et de volatilité du processus S ne sont pas constants.

Chapitre 5 : Méthodes de Monte-Carlo imbriquées

Selon les estimées du chapitre précédent, les solutions V^L et V^{NL} sont de précision équivalente pour évaluer le prix des contrats avec appels de marge. Nous nous intéressons alors au calcul de V^L, solution d'une EDSR linéaire, qui semble plus simple que celui de V^{NL}. En faisant l'hypothèse simplificatrice que les opérations de couverture sont effectuées avant $T-\Delta$, la correction induite par les appels de marge est donnée par :

$$\sqrt{\Delta} RC_\alpha \mathbb{E}\left[\int_0^{T-\Delta} e^{-ru}|Z_u^{BS}|du \right], \tag{1.3.4}$$

où Z^{BS} admet la représentation

$$Z_s^{BS} = \sigma s \partial_s \mathbb{E}\left[e^{-r(T-t)}\Phi(S_T)|S_t = s \right] = \mathbb{E}\left[e^{-r(T-t)}\left(\Phi(S_T) - \Phi(S_t)\right)\frac{W_T - W_t}{T-t}|S_t = s \right].$$

Nous pouvons ainsi estimer le terme de droite de l'approximation (1.3.4) en traitant un problème de Monte-Carlo imbriqué qui se met sous la forme générique suivante

$$\mathbb{E}\left[g\left(\mathbb{E}\left[f(X,Y)|X \right] \right) \right] =: I, \tag{1.3.5}$$

où f, g sont des fonctions mesurables et X, Y deux variables aléatoires indépendantes. La correction réécrit en posant $X = (U, S_U)$ où U une variable aléatoire uniforme sur $[0, T-\Delta]$ et Y une gaussienne centrée réduite, g est la fonction valeur absolue et pour $(u, s) \in [0, T-\Delta] \times \mathbb{R}$ et $y \in \mathbb{R}$,

$$f(u,s,y) = \frac{ye^{-rT}}{\sqrt{T-u}} \left(\Phi\left(se^{\left(r-\frac{\sigma^2}{2}\right)(T-u)+\sigma y\sqrt{T-u}} \right) - \Phi(s) \right).$$

Nous présentons deux approches pour traiter le problème (1.3.5). Une façon naive d'estimer V_0^L serait d'utiliser l'estimateur

$$\hat{I}_{M,N} = \frac{1}{M}\sum_{i=1}^{M} g\left(\frac{1}{N}\sum_{j=1}^{N} f(X^i, Y_j^i) \right),$$

où X^i et Y_j^i sont des tirages indépendants et de même loi que X et Y respectivement. L'analyse de l'erreur quadratique (RMSE) s'effectue par une décomposition biais-variance

$$\mathbf{RMSE}^2 := \mathbb{E}\left[\left(\hat{I}_{M,N} - I\right)^2 \right] = \left(\mathbb{E}\left[\hat{I}_{M,N}\right] - I \right)^2 + \frac{1}{M}\mathbf{Var}\left(g\left(\frac{1}{N}\sum_{j=1}^{N} f(X^i, Y_j^i) \right) \right)$$

et le contrôle sur l'erreur **RMSE** $< \epsilon$ est garanti pour $\epsilon > 0$ avec une complexité $\mathcal{O}(\epsilon^{-3})$. Pour améliorer la complexité du Monte-Carlo imbriqué, nous utilisons la technique multiniveaux introduite par Giles dans [46] avec un estimateur antithétique. Pour un nombre de couches $L \in \mathbb{N}^*$, un nombre de simulations extérieures M_l et de simulations intérieures n_l à chaque niveau $0 \le l \le L$, nous introduisons l'estimateur antithétique multiniveaux

$$\hat{I}_{\mathbf{M},\mathbf{n}}^{ML} = \hat{I}_{M_0,n_0} + \sum_{l=1}^{L} \tilde{I}_{M_l,n_l},$$

où $\mathbf{M} = (M_0, ..., M_L)$, $\mathbf{n} = n_0(2^0, ..., 2^L)$,

$$\begin{aligned}\tilde{I}_{M_l,n_l} = \frac{1}{M_l}\sum_{m=1}^{M_l} \Bigg\{ & g\left(\frac{1}{n_l}\sum_{j=1}^{n_l} f\left(X_m^l, Y_j^{l,m}\right) \right) - \frac{1}{2} g\left(\frac{1}{n_{l-1}}\sum_{j=1}^{n_{l-1}} f\left(X_m^l, Y_j^{l,m}\right) \right) \\ & - \frac{1}{2} g\left(\frac{1}{n_{l-1}}\sum_{j=n_{l-1}+1}^{n_l} f\left(X_m^l, Y_j^{l,m}\right) \right) \Bigg\},\end{aligned}$$

et $(X_m^l, Y_j^{l,m})$ sont des copies i.i.d. de (X, Y) pour $m, l, j \ge 0$. Pour une tolérance d'ordre $\epsilon > 0$ sur la RMSE, en égalisant les contributions du biais et de la variance à l'erreur quadratique, le choix du nombre de niveaux L est défini par le biais $\mathbb{E}\left[\hat{I}_{M_L,n_L}\right] - I$ tandis que la complexité est donnée par la vitesse de décroissance de $V_l := M_l \mathbf{Var}\left(\tilde{I}_{M_l,n_l}\right)$, variance du résultat d'une simulation de la couche $0 \le l \le L$.

Theorem (Giles, [46])**.** *Supposons qu'il existe trois constantes $c, \beta > 0$ et $\alpha \ge \frac{1}{2}$ telles que $V^l \le c2^{-\beta l}$ et $\left|\mathbb{E}\left[\hat{I}_{1,n_l}\right] - I\right| \le c2^{-\alpha l}$ pour tout $l \ge 0$. Alors la tolérance $\epsilon > 0$ sur la RMSE est garantie avec une complexité*

$$\begin{cases} \mathcal{O}\left(\epsilon^{-2}\right) & si\ \beta > 1, \\ \mathcal{O}\left(\epsilon^{-2}(\log\epsilon)^2\right) & si\ \beta = 1, \\ \mathcal{O}\left(\epsilon^{-2-(1-\beta)/\alpha}\right) & si\ \beta < 1. \end{cases}$$

Dans le cas où g est de classe C^2 à dérivées bornées, il est possible d'obtenir (cf. [46]) l'erreur quadratique ϵ pour une complexité $\mathcal{O}(\epsilon^{-2})$. Comme dans notre cadre g est une fonction avec un nombre fini de points singuliers, ce résultat ne s'applique pas directement, mais nous démontrons le théorème suivant.

Theorem. *Sous les hypothèses du Théorème 5.2.3, l'erreur* **RMSE** *est plus petite qu'une tolérance $\epsilon > 0$ pour une complexité d'ordre $\mathcal{O}(\epsilon^{-2})$.*

Ce résultat est également une conséquence du théorème de Giles. Il suffit pour cela de vérifier la vitesse de décroissance des termes de biais et variance en fonction des niveaux. La difficulté principale est de traiter les points singuliers de la fonction g. L'idée est de comparer la position $\mathbb{E}[f(X,Y)|X]$ avec son estimateur empirique et de partitionner l'espace des évènements selon que $\mathbb{E}[f(X,Y)|X]$ et son estimateur se trouvent dans un même intervalle où g est régulière, que $\mathbb{E}[f(X,Y)|X]$ est loin de son estimateur empirique ou que $\mathbb{E}[f(X,Y)|X]$ est proche d'un point de singularité. Dans le premier cas, il suffit d'appliquer un développement de Taylor et exploiter la régularité de g. Dans le deuxième, il suffit d'utiliser l'inégalité de Markov. Enfin, les hypothèses assurent que la loi de la variable aléatoire $\mathbb{E}[f(X,Y)|X]$ est suffisamment régulière pour ne pas être trop concentrée au voisinage des points singuliers, ce qui permet de traiter le troisième cas. Nous vérifions ensuite que ces hypothèses sont valides dans le cadre de V^L pour pouvoir estimer V_0^L à l'aide de la technique multiniveaux. Dans notre analyse, similaire à [45], les hypothèses que nous utilisons sont moins restrictives et plus faciles à vérifier que celles de [53] et étendent les résultats de [47].

Une autre façon d'estimer V_0^L dans le cas où g est convexe est de l'encadrer par des bornes inférieure et supérieure dont le calcul ne nécessite pas de Monte-Carlo imbriqué. Nous commençons par le cas où g est de type partie positive.

Lemma. *Soit $K \in \mathbb{R}$ et $\mathcal{R}$ une variable aléatoire réelle. Pour toute fonction mesurable $\varphi : \mathbb{R} \to \mathbb{R}$ et variable aléatoire $\mathcal{O}$,*

$$\mathbb{E}\left[(\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)_+\right] \geq \mathbb{E}\left[(\mathcal{R} - K)1_{\varphi(\mathcal{O}) \geq K}\right] =: C_\varphi(K),$$
$$\mathbb{E}\left[(K - \mathbb{E}[\mathcal{R}|\mathcal{O}])_+\right] \geq \mathbb{E}\left[(K - \mathcal{R})1_{\varphi(\mathcal{O}) \leq K}\right] =: P_\varphi(K).$$

Ainsi que les minorants dans le lemme ne dépendent plus de façon explicite de la loi conditionnelle de $\mathcal{R}$ sachant $\mathcal{O}$. De plus, comme une fonction $g : \mathbb{R} \to \mathbb{R}$ convexe est différentiable en dehors d'un nombre dénombrable de points, et sur un point de différentiabilité z, nous avons la représentation suivante :

$$\forall x \in \mathbb{R},\ g(x) = g(z) + g'(z)(x - z) + \int_z^\infty (x-u)^+ \mu(du) + \int_{-\infty}^z (u-x)^+ \mu(du),$$

où la mesure μ est égale à la dérivée seconde de g au sens des distributions. Nous introduisons ensuite pour toute fonction mesurable $\varphi : \mathbb{R} \to \mathbb{R}$

$$J_\varphi = g(z) + g'(z)(\mathbb{E}[\mathcal{R}] - z) + \int_z^\infty C_\varphi(u)\mu(du) + \int_{-\infty}^z P_\varphi(u)\mu(du).$$

En s'inspirant des techniques de valorisation d'options américaines, nous obtenons la représentation suivante permettant d'obtenir des bornes inférieure et supérieure sans Monte Carlo imbriqué.

Proposition. *Soit $\mathcal{R}$ une variable aléatoire réelle et $g : \mathbb{R} \to \mathbb{R}$ convexe avec $g(\mathcal{R})$ intégrable, alors*

$$\sup_\varphi J_\varphi = \mathbb{E}\left[g(\mathbb{E}[\mathcal{R}|\mathcal{O}])\right] = \inf_\epsilon \mathbb{E}\left[g(\mathcal{R} - \epsilon)\right],$$

où $\epsilon \in \{\Theta$ intégrable et tel que $\mathbb{E}[\Theta|\mathcal{O}] = 0\}$. L'égalité est atteinte pour $\varphi(\mathcal{O}) = \mathbb{E}[\mathcal{R}|\mathcal{O}]$ et $\epsilon = \mathcal{R} - \mathbb{E}[\mathcal{R}|\mathcal{O}]$.

Enfin nous comparons numériquement ces deux méthodes pour l'estimation de V_0^L.

Part I

Calibrated Stochastic and Local Volatility Models

Chapter 2

Existence of a calibrated regime-switching local volatility model and new fake Brownian motions

Ce chapitre est un article écrit avec Benjamin Jourdain. Une version raccourcie a été soumise à Mathematical Finance*, sous le nom de* Existence of a calibrated regime-switching local volatility model*.*

Abstract

By Gyongy's theorem, a local and stochastic volatility (LSV) model is calibrated to the market prices of all European call options with positive maturities and strikes if its local volatility function is equal to the ratio of the Dupire local volatility function over the root conditional mean square of the stochastic volatility factor given the spot value. This leads to a SDE nonlinear in the sense of McKean. Particle methods based on a kernel approximation of the conditional expectation, as presented in [57], provide an efficient calibration procedure even if some calibration errors may appear when the range of the stochastic volatility factor is very large. But so far, no global existence result is available for the SDE nonlinear in the sense of McKean. In the particular case where the local volatility function is equal to the inverse of the root conditional mean square of the stochastic volatility factor multiplied by the spot value given this value and the interest rate is zero, the solution to the SDE is a fake Brownian motion. When the stochastic volatility factor is a constant (over time) random variable taking finitely many values and the range of its square is not too large, we prove existence to the associated Fokker-Planck equation. Thanks to [39], we then deduce existence of a new class of fake Brownian motions. We then extend these results to the special case of the LSV model called regime switching local volatility, where the stochastic volatility factor is a jump process taking finitely many values and with jump intensities depending on the spot level. Under the same condition on the range of its square, we prove existence to the associated Fokker-Planck PDE. Finally, we deduce existence of the calibrated model by extending the results in [39].

Keywords: local and stochastic volatility models, calibration, Dupire's local volatility, Fokker-Planck systems, diffusions nonlinear in the sense of McKean.

2.1 Introduction

The notion of *fake* Brownian motion was introduced by [84] to describe a martingale $(X_t)_{t\geq 0}$ such that for any $t \geq 0$, $X_t \sim \mathcal{N}(0,t)$, but the process $(X_t)_{t\geq 0}$ does not have the same distribution as the Brownian motion $(W_t)_{t\geq 0}$. [77], [61], and more recently [63] provided construction of discontinuous fake Brownian motions. The question of the existence of continuous fake Brownian motion was positively answered by [5], using products of Bessel processes, and [10] extended that result by exhibiting a sequence of continuous martingales with Brownian marginals and scaling property. Oleszkiewicz gave a simpler example of a continuous fake Brownian motion, based on the Box Muller transform. Then, [65] proved existence of a continous fake exponential Brownian motion by mixing diffusion processes. More generally, [38] and [64] provided construction of self-similar martingales with given marginal distributions.

One reason for raising interest in the search of processes that have marginal distributions matching given ones comes from mathematical finance, where calibration to the market prices of European call options is a major concern. According to [18], the knowledge of the market prices of those options for a continuum of strikes and maturities is equivalent to the knowledge of the marginal distributions of the underlying asset under the pricing measure. Thus, to be consistent with the market prices, a calibrated model must have marginal distributions that coincide with those given by the market. More specifically, in this paper, we address the question of existence of a special class of calibrated local and stochastic volatility (LSV) models.

LSV models, introduced by [76] and by [89], can be interpreted as an extension of Dupire's local volatility (LV) model, described in [36], to get more consistency with real markets. A typical LSV model has the dynamics

$$dS_t = rS_t + f(Y_t)\sigma(t,S_t)S_t dW_t$$

for the stock under the risk-neutral probability measure, where $(Y_t)_{t\geq 0}$ is a stochastic process which may be correlated with $(S_t)_{t\geq 0}$ and r is the risk free rate. Meanwhile, according to [36], given the European call option prices $C(t,K)$ for all positive maturities t and strikes K, the process $\left(S_t^D\right)_{t\geq 0}$, which follows the dynamics

$$dS_t^D = rS_t^D dt + \sigma_{Dup}\left(t, S_t^D\right) S_t^D dW_t,$$

where $\sigma_{Dup}(t,K) := \sqrt{2\frac{\partial_t C(t,K) + rK\partial_K C(t,K)}{K^2 \partial^2_{KK} C(t,K)}}$ is Dupire's local volatility function, is calibrated to the European option prices, that is,

$$\forall t, K > 0,\ C(t,K) = \mathbb{E}\left[e^{-rt}\left(S_t^D - K\right)^+\right].$$

Under mild assumptions, Gyongy's theorem in [58] gives that the choice $\sigma(t,x) = \frac{\sigma_{Dup}(t,x)}{\sqrt{\mathbb{E}[f^2(Y_t)|S_t = x]}}$ ensures that for any $t \geq 0$, S_t has the same law as S_t^D. This leads to the following SDE

$$dS_t = rS_t dt + \frac{f(Y_t)}{\sqrt{\mathbb{E}[f^2(Y_t)|S_t]}}\sigma_{Dup}(t,S_t)S_t dW_t,$$

which is nonlinear in the sense of McKean, because the diffusion term depends on the law of (S_t, Y_t) through the conditional expectation in the denominator. Getting existence and uniqueness to this kind of SDE is still an open problem, although some local in time existence results have been found by [1] using a perturbative approach. From a numerical point of view, calibration can be achieved through the resolution of the associated Fokker-Planck PDE by [90]. Moreover, particle methods based on a kernel approximation of the conditional expectation, as presented by [57], provide an efficient calibation procedure even if some calibration errors may appear when the range of the stochastic volatility process $(f(Y_t))_{t\geq 0}$ is very large.

Recently, advances have been made for analogous models. [6] proved existence and uniqueness of stochastic local intensity models calibrated to CDO tranche prices, where the discrete loss process makes the conditional expectation simpler to handle. Moreover, [54] showed that in a local volatility model enhanced with jumps, the particle method applied to a well defined nonlinear McKean SDE with a regularized volatility function gives call prices that converge to the market prices as the regularization parameter goes to 0, for all strikes and maturities.

In this paper, we first prove existence of a simplified version of the LSV model. More precisely, we set $r = \frac{1}{2}$, $\sigma_{Dup} \equiv 1$, and consider the dynamics of the asset's log-price, where for simplicity we also neglect the drift term as its conditional expectation given the spot is equal to 0. Moreover, let $d \geq 2$ and Y be a random variable (constant in time) which takes values in $\mathcal{Y} := \{1,...,d\}$. We assume that for $i \in \{1,...,d\}$, $\alpha_i := \mathbb{P}(Y = i) > 0$. Given a positive function f on $\mathcal{Y}$ and a probability measure μ on $\mathbb{R}$, the SDE that we study is thus:

$$\begin{aligned} dX_t &= \frac{f(Y)}{\sqrt{\mathbb{E}\left[f^2(Y)|X_t\right]}}dW_t, \\ X_0 &\sim \mu. \end{aligned} \tag{2.1.1}$$

We also suppose that X_0, Y and $(W_t)_{t\geq 0}$ are independent. Let us set $\lambda_i := f^2(i) > 0, i \in \{1, ..., d\}$ and denote by $p_i(t,x)$ the conditional density of X_t given $\{Y = i\}$ multiplied by α_i. It means that, for any measurable and nonnegative function ϕ,

$$\mathbb{E}\left[\phi\left(X_t\right) 1_{\{Y=i\}}\right] = \int_{\mathbb{R}} \phi(x)p_i(t,x)dx.$$

The Fokker-Planck equations derived from SDE (2.1.1) form a partial differential system (PDS) that writes for $1 \leq i \leq d$:

$$\partial_t p_i = \frac{1}{2}\partial^2_{xx}\left(\frac{\sum_{l=1}^d p_l}{\sum_{l=1}^d \lambda_l p_l}\lambda_i p_i\right) \text{ in } (0,T)\times\mathbb{R} \tag{2.1.2}$$

$$p_i(0) = \alpha_i \mu \text{ in } \mathbb{R}, \tag{2.1.3}$$

where the constant $T > 0$ is the finite time horizon. We call this PDS (FBM). Let us define $\lambda_{min} := \min\limits_{1\leq i\leq d}\lambda_i$ and $\lambda_{max} := \max\limits_{1\leq i\leq d}\lambda_i$. We notice that if $\lambda_{min} = \lambda_{max} = \lambda$, then each p_i is a solution to the heat equation with initial condition $\alpha_i\mu$. Results proving existence and uniqueness of solutions to variational formulations of the heat equation have already been widely developed, e.g. in [19], so we focus on the case where $\lambda_{min} < \lambda_{max}$, that is when the function f is not constant on $\mathcal{Y}$. We also notice that if we add the equalities (2.1.2) and (2.1.3) over the index $i \in \{1, ..., d\}$, we obtain that $\sum_{i=1}^d p_i$ satisfies the heat equation with initial condition μ, and it is well known that $\left(\sum_{i=1}^d p_i\right)(t,x) = \mu * h_t(x)$, with h the heat kernel defined as $h_t(x) := \frac{1}{\sqrt{2\pi t}}\exp\left(\frac{-x^2}{2t}\right)$ for $t > 0$ and $x \in \mathbb{R}$. This observation can also be made formally, using Gyongy's theorem. Indeed, as we see that

$$\mathbb{E}\left[\left(\frac{f(Y)}{\sqrt{\mathbb{E}\left[f^2(Y)|X_t\right]}}\right)^2|X_t\right] = 1, \tag{2.1.4}$$

then if Gyongy's theorem applies, X_t has the same law as $X_0 + W_t$, which has the density $\mu * h_t$.

We introduce the following condition on the family $(\lambda_i)_{1\leq i\leq d}$, under which we will obtain existence to SDE (2.1.1). We denote by $(e_1, ..., e_d)$ the canonical basis of $\mathbb{R}^d$ and for $x = (x_1, ..., x_d) \in \mathbb{R}^d$, we define $x^\perp = \{y = (y_1, ..., y_d) \in \mathbb{R}^d, \sum_{i=1}^d x_i y_i = 0\}$.

Condition (C)**.** There exists a symmetric positive definite $d \times d$ matrix Γ with real valued coefficients such that for $1 \leq k \leq d$, the $d \times d$ matrix $\Gamma^{(k)}$ with coefficients

$$\Gamma^{(k)}_{ij} = \frac{\lambda_i + \lambda_j}{2}\left(\Gamma_{ij} + \Gamma_{kk} - \Gamma_{ik} - \Gamma_{jk}\right), \ 1 \leq i,j \leq d, \tag{2.1.5}$$

is positive definite on the space $e_k^\perp$.

Under the Condition (C), which ensures a coercivity property that enables to establish energy estimates, existence to SDE (2.1.1) can be proved in three steps:

1. For μ having a square integrable density on $\mathbb{R}$, we define a variational formulation called $V_{L^2}(\mu)$ to (FBM) and we apply Galerkin's method to show that $V_{L^2}(\mu)$ has a solution.

2. For μ a probability measure on $\mathbb{R}$, we define a weaker variational formulation called $V(\mu)$. We take advantage of the fact that if $(p_1, ..., p_d)$ is a solution to $V(\mu)$ then $\sum_{i=1}^d p_i$ is solution to the heat equation with initial condition μ. This enables to get control of the explosion rate of the L^2 norm of $p_i, 1 \leq i \leq d$, as $t \to 0$, and we extend the results obtained in Step 1 to show existence to $V(\mu)$.

3. Thanks to the results in [39], which give equivalence between the existence of a solution to a Fokker-Planck equation and the existence of a solution to the associated martingale problem with time marginals given by the solution to the Fokker-Planck equation, we show that the existence result in Step 2 implies existence of a weak solution to SDE (2.1.1). Moreover, if f is non constant on $\mathcal{Y}$ and if $X_0 = 0$, that weak solution is a continuous fake Brownian motion.

To get a more realistic financial framework, we then adapt the previous strategy to obtain existence of a class of calibrated LSV models, where the process $(S_t)_{t\geq 0}$ describing the underlying asset follows the dynamics

$$dS_t = rS_t dt + \frac{f(Y_t)}{\sqrt{\mathbb{E}[f^2(Y_t)|S_t]}}\sigma_{Dup}(t,S_t)S_t dW_t, \tag{2.1.6}$$

the process $(Y_t)_{t\geq 0}$ taking values in $\mathcal{Y}$ and

$$\forall j \in \{1,...,d\}\backslash\{Y_t\}, \mathbb{P}\left(Y_{t+dt} = j | \sigma\left((S_s,Y_s), 0 \leq s \leq t\right)\right) = q_{Y_t j}\left(\log\left(S_t\right)\right)dt,$$

where the functions $(q_{ij})_{1\leq i\neq j\leq d}$ are non negative and bounded. The process $(S_t,Y_t)_{t\geq 0}$ is then a calibrated regime switching local volatility (RSLV) model.

The chapter is organized as follows. In Section 2.2, we state the main results that give existence to SDE (2.1.1). Section 2.3 is dedicated to the proofs of the results in Section 2.2. In Section 2.4, we state the existence of calibrated RSLV models and we prove that result in Section 2.5. Beforehand, we introduce some additional notation.

Notation

- For an interval $I \subset \mathbb{R}$, we denote by $L^2(I)$ the space of measurable real valued functions defined on I which are square integrable for the Lebesgue measure. For $k \geq 1$, and $u, v \in \left(L^2(I)\right)^k$, we use the notation

$$(u,v)_k = \int_I \sum_{i=1}^k u_i(x)v_i(x)dx, \ |u|_k = (u,u)_k^{\frac{1}{2}}.$$

  and we define $L(I) := \left(L^2(I)\right)^d$. We also define $L := L(\mathbb{R})$ and we denote by L' its dual space.

- For $m \geq 1$, we denote by $H^m(I)$ the Sobolev space of real valued functions on I that are square integrable together with all their distribution derivatives up to the order m. We define the space $H(I) := \left(H^1(I)\right)^d$, endowed with the usual scalar product and norm

$$\langle u,v\rangle_d = \int_I \sum_{i=1}^d \left(u_i(x)v_i(x) + \partial_x u_i(x)\partial_x v_i(x)\right)dx, \ ||u||_d = \langle u,u\rangle_d^{\frac{1}{2}},$$

  and we define $H := H(\mathbb{R})$. We denote by H' its dual space, and by $\langle \cdot,\cdot\rangle$ the duality product between H and H'.

- For $1 \leq p \leq \infty$, we denote by $W^{1,p}(\mathbb{R})$ the Sobolev space of functions belonging to $L^p(\mathbb{R})$, and that have a first order derivative in the sense of distributions that also belongs to $L^p(\mathbb{R})$.

- For $n \geq 1$, $\mathcal{O}$ an open subset of $\mathbb{R}^n$ and $0 \leq k \leq \infty$, we denote by $C^k(\mathcal{O})$ the set of functions $\mathcal{O} \to \mathbb{R}$ that are continuous and have continuous derivatives up to the order k, we denote by $C_c^k(\mathcal{O})$ the set of functions in $C^k(\mathcal{O})$ that have compact support on $\mathcal{O}$, and we denote by $C_b^k(\mathcal{O})$ the set of functions in $C^k(\mathcal{O})$ that are uniformly bounded on $\mathcal{O}$ together with their p-th order derivatives, for $p \leq k$.

- For $n \geq 1$, we denote by $\mathcal{M}_n(\mathbb{R})$ the set of $n \times n$ matrices with real-valued coefficients. We denote by $I_n \in \mathcal{M}_n(\mathbb{R})$ the identity matrix, and by $J_n \in \mathcal{M}_n(\mathbb{R})$ the matrix where all the coefficients are equal to 1.

- For $n \geq 1$, we denote by $\mathcal{S}_n^+(\mathbb{R})$ (resp. $\mathcal{S}_n^{++}(\mathbb{R})$) the set of $n \times n$ matrices in $\mathcal{M}_n(\mathbb{R})$ which are symmetric and positive semidefinite (resp. definite). For $S \in \mathcal{S}_n^+(\mathbb{R})$ (resp. $S \in \mathcal{S}_n^{++}(\mathbb{R})$) we denote by $\sqrt{S}$ the unique matrix in $\mathcal{S}_n^+(\mathbb{R})$ (resp. $\mathcal{S}_n^{++}(\mathbb{R})$) such that $\sqrt{S}\sqrt{S} = S$, and we denote by $l_{min}(S)$ and $l_{max}(S)$ respectively the smallest eigenvalue and the largest eigenvalue of S.

- Given $n, p \geq 1$ and $\mathcal{A}$ a $n \times p$ matrix, we denote its transpose by $\mathcal{A}^*$ and we define $||A||_\infty := \max\{|\mathcal{A}_{ij}|, 1 \leq i \leq n, 1 \leq j \leq p\}$.
- For $y \in \mathbb{R}$, we denote its positive part by $y^+ = \max(0, y)$ and its negative part by $y^- = \min(0, y)$. For $k \geq 2$ and $x = (x_1, ..., x_k) \in \mathbb{R}^k$, we denote its positive part by $x^+ := \left(x_1^+, ..., x_k^+\right)$, and its negative part by $x^- := \left(x_1^-, ..., x_k^-\right)$.
- We denote by $\mathcal{P}(\mathbb{R})$ (resp. $\mathcal{P}(\mathbb{R} \times \mathcal{Y})$) the set of probability measures on $\mathbb{R}$ (resp. $\mathbb{R} \times \mathcal{Y}$).
- We define $\mathcal{D} := (\mathbb{R}^+)^d \backslash (0, ..., 0)$.
- For notational simplicity, for a function g defined on $\mathbb{R}^2$ and $t \in \mathbb{R}$, we may sometimes use the notation $g(t) := g(t, \cdot)$.

2.2 A new class of fake Brownian motions

In this section, we give the main results concerning the SDE (1). We introduce a variational formulation to give sense to the PDS (FBM). Let us assume that $p := (p_1, ..., p_d)$ is a classical solution to the PDS (FBM) such that p takes values in $\mathcal{D}$ and for $t \in (0, T]$, $p(t, \cdot) \in H$. For $v := (v_1, ..., v_d) \in H$ and $i \in \{1, ..., d\}$ we multiply (2.1.2) by v_i and integrate over $\mathbb{R}$. Through integration by parts, we obtain that for $i \in \{1, ..., d\}$, the following equality holds in the classical sense and also in the sense of distributions on $(0, T)$:

$$\int_{\mathbb{R}} v_i(x)\partial_t p_i(t,x)dx = \frac{d}{dt}\int_{\mathbb{R}} v_i(x)p_i(t,x)dx = -\frac{1}{2}\int_{\mathbb{R}} \partial_x v_i \partial_x \left(\frac{\sum_{l=1}^d p_l(t,x)}{\sum_{l=1}^d \lambda_l p_l(t,x)}\lambda_i p_i(t,x)\right)dx. \tag{2.2.1}$$

The term $\partial_x\left(\frac{\sum_{l=1}^d p_l}{\sum_{l=1}^d \lambda_l p_l}\lambda_i p_i\right)$ rewrites

$$\partial_x\left(\frac{\sum_{l=1}^d p_l}{\sum_{l=1}^d \lambda_l p_l}\lambda_i p_i\right) = \left(1 + \frac{\left(\sum_{l\neq i}\lambda_l p_l\right)\left(\sum_{l\neq i}(\lambda_i - \lambda_l)p_l\right)}{\left(\sum_{l=1}^d \lambda_l p_l\right)^2}\right)\partial_x p_i + \sum_{j\neq i}\frac{\lambda_i p_i\left(\sum_{l\neq j}(\lambda_l - \lambda_j)p_l\right)}{\left(\sum_{l=1}^d \lambda_l p_l\right)^2}\partial_x p_j.$$

Let us introduce the function $M : \mathcal{D} \to \mathcal{M}_d(\mathbb{R})$ where for $\rho \in \mathcal{D}$, $M(\rho)$ is the matrix with coefficients:

$$\begin{aligned} M_{ii}(\rho) &= \frac{\left(\sum_{l\neq i}\lambda_l \rho_l\right)\left(\sum_{l\neq i}(\lambda_i - \lambda_l)\rho_l\right)}{\left(\sum_{l=1}^d \lambda_l \rho_l\right)^2}, \ 1 \leq i \leq d, \\ M_{ij}(\rho) &= \frac{\lambda_i \rho_i\left(\sum_{l\neq j}(\lambda_l - \lambda_j)\rho_l\right)}{\left(\sum_{l=1}^d \lambda_l \rho_l\right)^2}, \ 1 \leq i \neq j \leq d, \end{aligned}$$

We also define $A : \mathcal{D} \to \mathcal{M}_d(\mathbb{R})$, where for $\rho \in \mathcal{D}$,

$$A(\rho) = \frac{1}{2}\left(I_d + M(\rho)\right).$$

For $1 \leq i \leq d$, we have that $\frac{1}{2}\partial_x\left(\frac{\sum_{l=1}^d p_l}{\sum_{l=1}^d \lambda_l p_l}\lambda_i p_i\right) = (A(p)\partial_x p)_i$. Summing the equality (2.2.1) over the index $1 \leq i \leq d$, the following equality holds in the sense of distributions on $(0, T)$,

$$\forall v \in H, \ \frac{d}{dt}(v, p(t))_d + (\partial_x v, A(p(t))\partial_x p(t))_d = 0.$$

We denote by p_0 the measure μ multiplied by the vector $(\alpha_1, ..., \alpha_d)$ and introduce below a weak variational formulation for the PDS (FBM).

$$\text{Find } p = (p_1, ..., p_d) \text{ satisfying:} \tag{2.2.2}$$

$$p \in L^2_{loc}((0, T]; H) \cap L^\infty_{loc}((0, T]; L), \tag{2.2.3}$$

$$p \text{ takes values in } \mathcal{D}, \text{ a.e. on } (0,T)\times\mathbb{R}, \tag{2.2.4}$$

$$\forall v \in H,\ \frac{d}{dt}(v,p)_d + (\partial_x v, A(p)\partial_x p)_d = 0 \text{ in the sense of distributions on } (0,T), \tag{2.2.5}$$

$$p(t,\cdot) \underset{t\to 0}{\overset{\text{weakly-*}}{\rightarrow}} p_0, \tag{2.2.6}$$

where the last condition means that:

$$\forall v \in H, (v,p(t))_d \underset{t\to 0}{\rightarrow} \sum_{i=1}^{d} \alpha_i \int_{\mathbb{R}} v_i(x)\mu(dx).$$

We call $V(\mu)$ the problem defined by (2.2.2)-(2.2.6). If p only satisfies (2.2.3), then the initial condition (2.2.6) does not make sense. As we will see in the proof of Theorem 2.2.1, if p satisfies the conditions (2.2.3)-(2.2.5), then p is a.e. equal to a function continuous from $(0,T]$ into L. We will always consider its continuous representative and therefore (2.2.6) makes sense. The existence results to $V(\mu)$ and to the SDE (2.1.1) are stated in the following theorems.

Theorem 2.2.1. *Under Condition (C), $V(\mu)$ has a solution $p \in C((0,T],L)$ such that for almost every (t,x) in $(0,T]\times\mathbb{R}$, $\sum_{i=1}^{d} p_i(t,x) = (\mu * h_t)(x)$.*

Theorem 2.2.2. *Under Condition (C), SDE (2.1.1) has a weak solution, which has the same marginal law as $(Z+W_t)_{t\geq 0}$, where Z is independent from $(W_t)_{t\geq 0}$ and $Z \sim \mu$.*

The proofs of Theorems 2.2.1 and 2.2.2 are postponed to the following section. To end this section, let us remark that whenever the function f is not constant on $\mathcal{Y}$, the solution to the SDE (2.1.1) given by Theorem 2.2.2 is a continuous fake Brownian motion.

Proposition 2.2.3. *Under Condition (C), if f is not constant on $\mathcal{Y}$ then the solutions to SDE (2.1.1) with initial condition $\mu = \delta_0$ are continuous fake Brownian motions.*

Proof. Let X be a solution to SDE (2.1.1). The term $\frac{f(Y)}{\sqrt{\mathbb{E}[f^2(Y)|X_t]}}$ is bounded, so X is a continuous martingale and by [58, Theorem 4.6], for $t \in [0,T]$, X_t has the law $\mathcal{N}(0,t)$. Let us remark that a solution to SDE (2.1.1) with the properties stated by Theorem 2.2.2 satisfies the marginal constraints without using Gyongy's theorem. We consider the quadratic variation $\langle X\rangle$ of X, which satisfies $d\langle X\rangle_t = \frac{f^2(Y)}{\mathbb{E}[f^2(Y)|X_t]}dt$. If almost surely, $\forall t \geq 0$, $\langle X\rangle_t = t$, then a.s., dt-a.e., $\frac{f^2(Y)}{\mathbb{E}[f^2(Y)|X_t]} = 1$ and $f^2(Y)$ is $\sigma(X_t)$ measurable. As moreover f is positive, $X_t = W_t$, and there exists $t \geq 0$ such that $f^2(Y)$ is a measurable function of W_t. As Y is independent from $(W_t)_{t\geq 0}$, we have that $f^2(Y)$ is constant. Then by contraposition, if f is not constant on $\mathcal{Y}$, we do not have that for $t > 0$, $\langle X\rangle_t = t$, so $(X_t)_{t\geq 0}$ is a continuous fake Brownian motion. □

2.3 Proofs of Section 2

Condition (C), developed in Subsection 2.3.1, enables to establish a priori energy estimates of solutions to $V(\mu)$, computed in Subsection 2.3.2, where we give a stronger variational formulation to the PDS (FBM) under the assumption that the initial distribution μ has a square integrable density. Under Condition (C), we prove existence to that formulation. That result is extended in Subsection 2.3.3 to prove Theorem 2.2.1. In Subsection 2.3.4, we establish a link between $V(\mu)$ and the variational formulation in the sense of distributions defined in [39], and we prove Theorem 2.2.2, thanks to [39, Theorem 2.6].

2.3.1 Condition (C)

Let us introduce the notion of uniform coercivity.

Definition 2.3.1. Given a domain $D \subset \mathbb{R}^d$, a function $G : D \to \mathcal{M}_d(\mathbb{R})$ is uniformly coercive on D with a coefficient $c > 0$ if:

$$\forall \rho \in D, \forall \xi \in \mathbb{R}^d,\ \xi^* G(\rho)\xi \geq c\xi^*\xi.$$

First, let us notice that the function A is bounded.

Lemma 2.3.2. *Any matrix B in the image $A(\mathcal{D})$ of $\mathcal{D}$ by A satisfies $||B||_\infty \leq \frac{1}{2}\left(1+\frac{\lambda_{max}}{\lambda_{min}}\right)$.*

Proof. For $\rho \in \mathcal{D}, A(\rho) = \frac{1}{2}(I_d + M(\rho))$. It is sufficient to check that $||M(\rho)||_\infty \leq \frac{\lambda_{max}}{\lambda_{min}}$. Since for $1 \leq l \leq d$, $\lambda_l > 0$, $\rho_l \geq 0$ and $\sum_{l=1}^d \lambda_l \rho_l > 0$, we have for $1 \leq i \neq j \leq d$,

$$\begin{aligned}
-\frac{\lambda_{max}}{\lambda_{min}} \leq -\lambda_j \frac{\lambda_i\rho_i}{\left(\sum_l \lambda_l\rho_l\right)} \frac{\sum_{l\neq j}\rho_l}{\left(\sum_l \lambda_l\rho_l\right)} &\leq M_{ij}(\rho) \leq \frac{\lambda_i\rho_i}{\left(\sum_l \lambda_l\rho_l\right)} \frac{\sum_{l\neq j}\lambda_l\rho_l}{\left(\sum_l \lambda_l\rho_l\right)} \leq 1 \leq \frac{\lambda_{max}}{\lambda_{min}},\\
-\frac{\lambda_{max}}{\lambda_{min}} \leq -1 \leq \frac{\sum_{l\neq i}\lambda_l\rho_l \sum_{l\neq i}(-\lambda_l)\rho_l}{\left(\sum_l \lambda_l\rho_l\right)^2} &\leq M_{ii}(\rho) \leq \lambda_i \frac{\sum_{l\neq i}\lambda_l\rho_l}{\sum_l \lambda_l\rho_l} \frac{\sum_{l\neq i}\rho_l}{\left(\sum_l \lambda_l\rho_l\right)} \leq \frac{\lambda_{max}}{\lambda_{min}}.
\end{aligned}$$

□

The role of Condition (C) is to ensure the existence of a matrix $\Pi \in \mathcal{S}_d^{++}(\mathbb{R})$ such that ΠA is uniformly coercive on $\mathcal{D}$. By a slight abuse of notation, we will say that a matrix Γ satisfies (C) if $\Gamma \in \mathcal{S}_d^{++}(\mathbb{R})$, and for $1 \leq k \leq d$, the matrix $\Gamma^{(k)}$ defined by (2.1.5) is positive definite on $e_k^\perp$. We consider matrices with the form $J_d + \epsilon\Gamma$, where $\Gamma \in S_d^{++}(\mathbb{R})$ and $\epsilon > 0$, and we show that if Γ satisfies (C), then for ϵ small enough, $(J_d + \epsilon\Gamma)A$ achieves the coercivity property. The intuition behind the choice of J_d is the observation that if $(p_1, ..., p_d)$ is a solution to $V(\mu)$ then $\sum_{i=1}^d p_i$ is a solution to the heat equation. This also translates into the algebraic property that $J_d M = 0$. We define $\mathbf{1} := (1, ..., 1) \in \mathbb{R}^d$.

Proposition 2.3.3. *If $\Gamma \in \mathcal{S}_d^{++}(\mathbb{R})$ satisfies Condition (C), then there exists $z > 0$ such that*

$$\forall x \in \mathbf{1}^\perp, \forall \rho \in \mathcal{D}, x^*\Gamma A(\rho)x \geq z x^* x. \tag{2.3.1}$$

Moreover, for $\epsilon > 0$ small enough, the function $(J_d + \epsilon\Gamma)A$ is uniformly coercive on $\mathcal{D}$.

Proof. For $\rho \in \mathcal{D}$, we introduce the notation $\overline{\lambda}(\rho) := \sum_k \lambda_k \frac{\rho_k}{\sum_l \rho_l}$ and remark that $\overline{\lambda}(\mathcal{D}) = [\lambda_{min}, \lambda_{max}]$. The matrix $A(\rho)$ rewrites as the convex combination $A(\rho) = \sum_{k=1}^d w_k(\rho) A_k(\rho)$ with the weight $w_k(\rho) = \frac{\lambda_k\rho_k}{\sum_l \lambda_l\rho_l}$ of the matrix $A_k(\rho)$ which has non zero coefficients only on the diagonal and the k-th row, and is defined by:

$$(A_k(\rho))_{ij} = \frac{1}{2}\left(1_{\{i=j\}}\frac{\lambda_i}{\overline{\lambda}(\rho)} + 1_{\{i=k\}}\left(1 - \frac{\lambda_j}{\overline{\lambda}(\rho)}\right)\right), \ 1 \leq i, j \leq d.$$

We prove that for $1 \leq k \leq d$, there exists $z_k > 0$ such that

$$\forall x \in \mathbf{1}^\perp, \forall \rho \in \mathcal{D}, 2x^*\overline{\lambda}(\rho)\Gamma A_k(\rho)x \geq z_k x^* x, \tag{2.3.2}$$

and we can set $z = \min\limits_{1\leq k\leq d} \frac{z_k}{2\lambda_{max}} > 0$ to obtain (2.3.1). We have that

$$2\overline{\lambda}(\rho)\Gamma A_k(\rho) = \left(\lambda_j\left(\Gamma_{ij} - \Gamma_{ik}\right)\right)_{1\leq i,j\leq d} + \left(\overline{\lambda}(\rho)\left(\Gamma_{ik}\right)\right)_{1\leq i,j\leq d}$$

For $x \in \mathbf{1}^\perp$, we have

$$\begin{aligned}
2x^*\overline{\lambda}(\rho)\Gamma A_k(\rho)x &= \sum_{i,j=1}^d \lambda_j\left(\Gamma_{ij} - \Gamma_{ik}\right)x_i x_j + \sum_{i,j=1}^d \overline{\lambda}(\rho)\left(\Gamma_{ik}\right)x_i x_j && (2.3.3)\\
&= \sum_{i\neq k, j\neq k} \lambda_j\left(\Gamma_{ij} - \Gamma_{ik} - \Gamma_{kj} + \Gamma_{kk}\right)x_i x_j && (2.3.4)\\
&= \sum_{i\neq k, j\neq k} \frac{\lambda_i + \lambda_j}{2}\left(\Gamma_{ij} - \Gamma_{ik} - \Gamma_{kj} + \Gamma_{kk}\right)x_i x_j, && (2.3.5)
\end{aligned}$$

where between the first and the second equality, we used the fact that $\sum_{i=1}^d x_i = 0$ and then replaced x_k by $-\sum_{i\neq k} x_i$. As $\Gamma^{(k)}$ is positive definite on $e_k^\perp$, there exists $\tilde{z}_k > 0$ such that

$$2x^*\overline{\lambda}(\rho)\Gamma A_k(\rho)x \geq \tilde{z}_k \sum_{i\neq k} x_i^2.$$

By Cauchy-Schwarz inequality, $x_k^2 = \left(\sum_{i\neq k} x_i\right)^2 \leq (d-1)\sum_{i\neq k} x_i^2$, so $\sum_{i\neq k} x_i^2 \geq \frac{1}{d}x^*x$ and we can set $z_k = \frac{\tilde{z}_k}{d}$ to satisfy (2.3.2).

Now, we show that for $\epsilon > 0$ small enough, the function $(J_d + \epsilon\Gamma)\,A$ is uniformly coercive on $\mathcal{D}$. For $\rho \in \mathcal{D}$, as $A(\rho) = \frac{1}{2}\left(I_d + M(\rho)\right)$ and for $1 \leq j \leq d$, as it is easy to check that $\sum_{i=1}^d M_{ij}(\rho) = 0$, we have that $J_d M = 0$ and

$$(J_d + \epsilon\Gamma)\,A(\rho) = \frac{1}{2}J_d + \epsilon\Gamma A(\rho).$$

For $x \in \mathbb{R}^d$, we decompose $x = u + v$ where $u \in \mathbb{R}\mathbf{1}$ and $v \in \mathbf{1}^\perp$. As $J_d v = 0$, we have

$$x^*\left(\frac{1}{2}J_d + \epsilon\Gamma A(\rho)\right)x \;=\; u^*\left(\frac{1}{2}J_d + \epsilon\Gamma A(\rho)\right)u + \epsilon v^*\Gamma A(\rho)v + \epsilon u^*\Gamma A(\rho)v + \epsilon v^*\Gamma A(\rho)u.$$

We will use Young's inequality:

$$\forall\eta > 0, \forall a,b \in \mathbb{R},\; ab \leq \eta a^2 + \frac{1}{4\eta}b^2.$$

For $1 \leq i,j \leq d$, by Young's inequality and Lemma 2.3.2, $(\Gamma A(\rho))_{ij} = \sum_{k=1}^d \Gamma_{ik}A(\rho)_{kj} \leq \frac{d}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)$, so for $a,b \in \mathbb{R}$, and $\eta > 0$,

$$(\Gamma A(\rho))_{ij}\,ab \geq -\frac{d}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left(\eta a^2 + \frac{1}{4\eta}b^2\right).$$

We obtain, for $\tilde{x},\tilde{y} \in \mathbb{R}^d$, and $\eta > 0$,

$$\tilde{x}^*\Gamma A(\rho)\tilde{y} = \sum_{i,j=1}^d (\Gamma A(\rho))_{ij}\,\tilde{x}_i\tilde{y}_j \geq -\frac{d^2}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left(\eta\tilde{x}^*\tilde{x} + \frac{1}{4\eta}\tilde{y}^*\tilde{y}\right).$$

Then for $\eta > 0$,

$$\begin{aligned} v^*\Gamma A(\rho)u &\geq -\frac{d^2}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left(\eta v^*v + \frac{1}{4\eta}u^*u\right), \\ u^*\Gamma A(\rho)v &\geq -\frac{d^2}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left(\eta v^*v + \frac{1}{4\eta}u^*u\right), \end{aligned}$$

and moreover,

$$u^*\Gamma A(\rho)u \geq -\frac{d^2}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)u^*u.$$

Using (2.3.1) and $J_d u = du$, we get that for $\eta > 0$,

$$x^*\left(\frac{1}{2}J_d + \epsilon\Gamma A(\rho)\right)x \geq \left(\frac{d}{2} - \epsilon\frac{d^2}{2}||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left(1 + \frac{1}{2\eta}\right)\right)u^*u + \epsilon\left(z - d^2||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\eta\right)v^*v.$$

For $0 < \epsilon < \left(d||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left(1 + \frac{d^2||\Gamma||_\infty\left(1+\frac{\lambda_{max}}{\lambda_{min}}\right)}{2z}\right)\right)^{-1}$, we check that

$$\eta_1 := \left(\frac{2}{\epsilon d||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)} - 2\right)^{-1} < \frac{z}{d^2||\Gamma||_\infty\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)} =: \eta_2,$$

and with the choice $\eta \in (\eta_1,\eta_2)$, we see that the function $(J_d + \epsilon\Gamma)\,A$ is uniformly coercive on $\mathcal{D}$. ☐

Corollary 2.3.4. *Condition (C) is equivalent to the existence of a matrix $\Pi \in \mathcal{S}_d^{++}(\mathbb{R})$ such that the function ΠA is uniformly coercive on $\mathcal{D}$ with a coefficient $\kappa \in \left(0, \frac{l_{min}(\Pi)}{2}\right)$.*

Proof. If a matrix $\Gamma \in \mathcal{S}_d^{++}(\mathbb{R})$ satisfies (C), then by Proposition 2.3.3, for $\epsilon > 0$ small enough, the function $(J_d + \epsilon\Gamma)A$ is uniformly coercive on $\mathcal{D}$. Moreover, $(J_d + \epsilon\Gamma) \in \mathcal{S}_d^{++}(\mathbb{R})$ as $J_d \in \mathcal{S}_d^{+}(\mathbb{R})$.

Conversely, if $\Pi \in \mathcal{S}_d^{++}(\mathbb{R})$ is such that the function ΠA is uniformly coercive on $\mathcal{D}$ with a coefficient $c > 0$, with the same computation as (2.3.3)-(2.3.5), we have that for $x \in \mathbf{1}^\perp$ and $1 \le k \le d$,

$$\sum_{i\neq k, j\neq k} \frac{\lambda_i + \lambda_j}{2} \left(\Pi_{ij} - \Pi_{ik} - \Pi_{kj} + \Pi_{kk}\right) x_i x_j = 2x^* \overline{\lambda}(\rho) \Pi A_k(\rho) x \ge 2\lambda_{min} c \sum_{i \neq k} x_i^2,$$

so Π satisfies (C). To conclude, it is obvious that if ΠA is uniformly coercive, then ΠA is uniformly coercive with a coefficient $\kappa \le \frac{l_{min}(\Pi)}{2}$. □

Making Condition (C) explicit does not seem to be an easy task in general, but we give here simpler criteriae which are all proved in Appendix 2.B, for particular situations.

- For $d = 2$, Condition (C) is satisfied, for the choice $\Gamma = I_2$.
- For $d = 3$, we define

  $$r_1 = \frac{\lambda_3}{\lambda_2} + \frac{\lambda_2}{\lambda_3} \ge 2, \; r_2 = \frac{\lambda_3}{\lambda_1} + \frac{\lambda_1}{\lambda_3} \ge 2, \; r_3 = \frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} \ge 2.$$

  Condition (C) is satisfied if and only if

  $$\frac{1}{\sqrt{(r_1-2)(r_2-2)}} + \frac{1}{\sqrt{(r_2-2)(r_3-2)}} + \frac{1}{\sqrt{(r_1-2)(r_3-2)}} > \frac{1}{4},$$

  with the convention that $\frac{1}{0} = +\infty$.
- For $d \ge 4$, we give in Appendix 2.B a numerical procedure to check if there exists a diagonal matrix that satisfies Condition (C).
- For $d \ge 4$, if

  $$\max_{1\le k\le d} \sqrt{\sum_{i\neq k} \lambda_i \sum_{i \neq k} \frac{1}{\lambda_i}} < d+1,$$

  then Condition (C) is satisfied, for the choice $\Gamma = I_d$.

2.3.2 μ has a square integrable density

We first suppose that the measure μ has a square integrable density with respect to the Lebesgue measure. In this case, we denote by p_0, the element of L obtained by multiplication of that density by the vector $(\alpha_1, ..., \alpha_d)$, and we define the stronger variational formulation:

$$\text{Find} \qquad p = (p_1, ..., p_d) \text{ satisfying :} \tag{2.3.6}$$

$$p \in L^2([0,T]; H) \cap L^\infty([0,T]; L), \tag{2.3.7}$$

$$p \text{ takes values in } \mathcal{D}, \text{ a.e. on } [0,T]\times\mathbb{R}, \tag{2.3.8}$$

$$\forall v \in H, \; \frac{d}{dt}(v,p)_d + (\partial_x v, A(p)\partial_x p)_d = 0 \text{ in the sense of distributions on } (0,T), \tag{2.3.9}$$

$$p_i(0,\cdot) = p_{0,i} \;, \forall i \in \{1, ..., d\}. \tag{2.3.10}$$

We call $V_{L^2}(\mu)$ the problem defined by (2.3.6)-(2.3.10). If p only satisfies (2.3.7), then the initial condition (2.3.10) does not make sense. We will show that if p satisfies the conditions (2.3.7)-(2.3.9), then p is a.e. equal to a function continuous from $[0,T]$ into L. We consider in what follows this continuous representative and therefore the initial condition (2.3.10) makes sense. We now present an existence result to $V_{L^2}(\mu)$.

Theorem 2.3.5. *Under Condition (C), $V_{L^2}(\mu)$ has a solution $p \in C([0,T],L)$ such that for almost every (t,x) in $[0,T]\times\mathbb{R}$, $\sum_{i=1}^d p_i(t,x) = (\mu * h_t)(x) = \int_{\mathbb{R}} h_t(x-y)\mu(y)dy$.*

To prove Theorem 2.3.5, we apply Galerkin's procedure, as in [97, III. 1.3]. Since H is a separable Hilbert space, there exists a sequence $(w_k)_{k\geq 1} = ((w_{k1},...,w_{kd}))_{k\geq 1}$ of linearly independent elements which is total in H. It is not at all obvious to preserve the condition that p takes values in $\mathcal{D}$ a.e. on $[0,T]\times\mathbb{R}$ at the discrete level. That is why, for $\epsilon > 0$, we introduce for $\rho \in (\mathbb{R}_+)^d$ the approximation M_ϵ of M defined on $(\mathbb{R}_+)^d$ by

$$\begin{aligned} M_{\epsilon,ii}(\rho) &= \frac{\sum_{l\neq i}\lambda_l\rho_l\sum_{l\neq i}(\lambda_i-\lambda_l)\rho_l}{\epsilon^2\vee\left(\sum_l\lambda_l\rho_l\right)^2},\ 1\leq i\leq d,\\ M_{\epsilon,ij}(\rho) &= \frac{\lambda_i\rho_i\sum_{l\neq j}(\lambda_l-\lambda_j)\rho_l}{\epsilon^2\vee\left(\sum_l\lambda_l\rho_l\right)^2},\ 1\leq i\neq j\leq d. \end{aligned}$$

We introduce the approximation A_ϵ of A defined on $(\mathbb{R}_+)^d$ by $A_\epsilon = \frac{1}{2}(I_d+M_\epsilon)$, and the approximate variational formulation $V_\epsilon(\mu)$ defined by

$$\text{Find} \qquad p_\epsilon = (p_{\epsilon,1},...,p_{\epsilon,d}) \text{ satisfying :} \tag{2.3.11}$$

$$p_\epsilon \in L^2([0,T];H)\cap L^\infty([0,T];L) \tag{2.3.12}$$

$$\forall v\in H,\ \frac{d}{dt}(v,p_\epsilon)_d + (\partial_x v, A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon)_d = 0 \text{ in the sense of distributions on } (0,T), \tag{2.3.13}$$

$$p_{\epsilon,i}(0,\cdot) = p_{0,i},\ \forall i\in\{1,...,d\} \tag{2.3.14}$$

For the same reasons as for the variational formulation $V_{L^2}(\mu)$, any solution of $V_\epsilon(\mu)$ has a continuous representative in $C([0,T],L)$, and therefore (2.3.14) makes sense.

We first prove existence to $V_\epsilon(\mu)$ by Galerkin's procedure. Then we will check that a.e. on $[0,T]\times\mathbb{R}$, p_ϵ takes values in $(\mathbb{R}_+)^d$. We will then obtain existence to $V_{L^2}(\mu)$ by taking the limit $\epsilon\to 0$. In what follows, whenever Condition (C) is satisfied, we denote by Π an element of $\mathcal{S}_d^{++}(\mathbb{R})$ such that ΠA is uniformly coercive on $\mathcal{D}$ with a coefficient $\kappa\in\left(0,\frac{l_{min}(\Pi)}{2}\right)$, and both exist by Corollary 2.3.4. Let us remark that as the function $v\to\Pi v$ is a bijection from H to H and Π is symmetric, Equality (2.3.13) is equivalent to:

$$\forall v\in H,\ \frac{d}{dt}(v,\Pi p_\epsilon)_d + (\partial_x v,\Pi A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon)_d = 0,\ \text{ in the sense of distributions on } (0,T). \tag{2.3.15}$$

This formulation will help us take advantage of the coercivity property of the function ΠA. For $m\in\mathbb{N}^*$, we denote by p_0^m the orthogonal projection of p_0 onto the subspace of L spanned by $(w_1,...,w_m)$. We first solve an approximate formulation named $V_\epsilon^m(\mu)$:

$$\text{Find } g_{\epsilon,1}^m,...,g_{\epsilon,m}^m\in C^1([0,T],\mathbb{R}),\ \text{such that the function } t\in[0,T]\to p_\epsilon^m(t) := \textstyle\sum_{j=1}^m g_{\epsilon,j}^m(t)w_j \text{ satisfies:} \tag{2.3.16}$$

$$\forall i\in\{1,...,m\},\ \frac{d}{dt}\left(w_i,\Pi p_\epsilon^m(t)\right)_d + \left(\partial_x w_i,\Pi A_\epsilon\left(\left(p_\epsilon^m\right)^+(t)\right)\partial_x p_\epsilon^m(t)\right)_d = 0, \tag{2.3.17}$$

$$p_\epsilon^m(0) = p_0^m. \tag{2.3.18}$$

We denote by $W^{(m)}\in\mathcal{S}_m^{++}(\mathbb{R})$ the non singular Gram matrix of the linearly independent family $\left(\sqrt{\Pi}w_i\right)_{1\leq i\leq m}$, with coefficients $W_{ij}^{(m)} = (w_i,\Pi w_j)_d$ for $1\leq i,j\leq m$. We introduce $g_{\epsilon,0}^m := (g_{01}^m,...,g_{0m}^m)\in\mathbb{R}^m$, which is p_0^m expressed on the basis $(w_1,...,w_m)$, and $g_\epsilon^m := \left(g_{\epsilon,1}^m,...,g_{\epsilon,m}^m\right)$. We define the function $K_\epsilon^m:\mathbb{R}^m\to\mathcal{M}_m(\mathbb{R})$ such that for $z\in\mathbb{R}^m$, $K_\epsilon^m(z)$ is the matrix with coefficients

$$K_\epsilon^m(z)_{ij} = \left(\partial_x w_i,\Pi A_\epsilon\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right)\partial_x w_j\right)_d,$$

for $1 \leq i, j \leq m$ and we define the function F_ϵ^m on $\mathbb{R}^m$ by:

$$F_\epsilon^m(z) = -\left(W^{(m)}\right)^{-1} K_\epsilon^m(z)z.$$

Solving $V_\epsilon^m(\mu)$ becomes equivalent to solving the following ODE for g_ϵ^m:

$$\left(g_\epsilon^m\right)'(t) = F_\epsilon^m(g_\epsilon^m) \tag{2.3.19}$$

$$g_\epsilon^m(0) = g_{\epsilon,0}^m. \tag{2.3.20}$$

To show existence of a unique solution to $(V_\epsilon^m(\mu))$, for $m \geq 1$, we check that the function F_ϵ^m is locally Lipschitz and that the function A_ϵ is bounded. The proof of Lemma 2.3.6 below is postponed to Appendix 2.A.

Lemma 2.3.6. *For $m \geq 1$, the function $z \in \mathbb{R}^m \to F_\epsilon^m(z)$ is locally Lipschitz.*

Lemma 2.3.7. *For $\epsilon > 0$ and $B \in A_\epsilon\left((\mathbb{R}^+)^d\right)$, $||B||_\infty \leq \frac{1}{2}\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)$.*

Proof. For $\epsilon > 0$, $A_\epsilon(0) = \frac{1}{2}I_d$, so $||A_\epsilon(0)||_\infty = \frac{1}{2}$. For $\rho \in \mathcal{D}$ and $\epsilon > 0$, it is clear that $||M_\epsilon(\rho)||_\infty \leq ||M(\rho)||_\infty \leq \frac{\lambda_{max}}{\lambda_{min}}$, where the inequality on the r.h.s comes from the proof of Lemma 2.3.2, so $||A_\epsilon(\rho)||_\infty \leq \frac{1}{2}\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)$. □

Lemma 2.3.8. *For every $m \geq 1$, there exists a unique solution to $(V_\epsilon^m(\mu))$.*

Proof. For $m \geq 1$, we use Lemma 2.3.6 and the Cauchy-Lipschitz theorem to get existence and uniqueness of a maximal solution g_ϵ^m on the interval $[0, T^*)$ for a certain $T^* > 0$. It is sufficient to show that $T^* > T$, to ensure that g_ϵ^m is defined on $[0, T]$. As A_ϵ is uniformly bounded by $\frac{1}{2}\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)$ by Lemma 2.3.7, we have that all the coefficients of $\left(W^{(m)}\right)^{-1} K_\epsilon^m$ are uniformly bounded by

$$\gamma := \frac{d}{2}\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)\left|\left|\left(W^{(m)}\right)^{-1}\right|\right|_\infty ||\Pi||_\infty \max_{1 \leq j \leq m}\left(\sum_{i=1}^m \sum_{a,b=1}^d \int_{\mathbb{R}} |\partial_x w_{ia}||\partial_x w_{jb}| dx\right).$$

For $1 \leq i \leq m$,

$$\left|\left(g_{\epsilon,i}^m\right)'(t)\right| = \left|\sum_{j=1}^m \left(\left(W^{(m)}\right)^{-1} K_\epsilon^m(g_\epsilon^m(t))\right)_{ij} g_{\epsilon,j}^m(t)\right| \leq \gamma \sum_{j=1}^m |g_{\epsilon,j}^m(t)|.$$

Summing over the index $i \in \{1, ..., d\}$, we have for $t \in [0, T^*)$:

$$\sum_{i=1}^m |g_{\epsilon,i}^m(t)| \leq \sum_{i=1}^m |g_{\epsilon,i}^m(0)| + \sum_{i=1}^m \int_0^t |\left(g_{\epsilon,i}^m\right)'(t)| dt \leq \sum_{i=1}^m |g_{\epsilon,i}^m(0)| + m\gamma \int_0^t \left(\sum_{i=1}^d |g_{\epsilon,i}^m(t)|\right) dt,$$

and by Gronwall's lemma, for $0 \leq t < T^*$, $\left(\sum_{i=1}^m |g_{\epsilon,i}^m(t)|\right) \leq \left(\sum_{i=1}^m |g_{\epsilon,i}^m(0)|\right) \exp(m\gamma t)$. If $T^* < \infty$, the function $t \to \left(\sum_{i=1}^m |g_{\epsilon,i}^m(t)|\right)$ would explode as $t \to T^*$. We conclude that $T^* = \infty$, g_ϵ^m is defined on $[0, T]$ and $p_\epsilon^m = \sum_{i=1}^m g_{\epsilon,i}^m(t) w_i$ is the solution to $V_\epsilon^m(\mu)$. □

Before showing the existence of a converging subsequence of $(p_\epsilon^m)_{m \geq 1}$ whose limit is a solution of $V_\epsilon(\mu)$, we check that ΠA_ϵ is uniformly coercive on $\mathcal{D}$, uniformly in $\epsilon > 0$.

Lemma 2.3.9. *If ΠA is uniformly coercive on $\mathcal{D}$ with coefficient $\kappa \in \left(0, \frac{l_{min}(\Pi)}{2}\right)$, then for $\epsilon > 0$, ΠA_ϵ is uniformly coercive on $(\mathbb{R}^+)^d$ with coefficient κ.*

Proof. For $\epsilon > 0$, $A_\epsilon(0) = \frac{1}{2}I_d$ so $\forall \xi \in \mathbb{R}^d, \xi^* \Pi A_\epsilon(0)\xi = \frac{1}{2}\xi^*\Pi\xi \geq \frac{l_{min}(\Pi)}{2}\xi^*\xi \geq \kappa\xi^*\xi$. For $\rho \in \mathcal{D}$, if $\epsilon \leq \sum_l \lambda_l \rho_l$, then $A_\epsilon(\rho) = A(\rho)$ and by hypothesis $\forall \xi \in \mathbb{R}^d, \xi^*\Pi A_\epsilon(\rho)\xi \geq \kappa\xi^*\xi$. If $\epsilon > \sum_l \lambda_l p_l$, then for $1 \leq i, j \leq d$, $M_{\epsilon,ij}(\rho) = \left(\frac{1}{\epsilon}\sum_l \lambda_l \rho_l\right)^2 M_{ij}(\rho)$, with $\left(\frac{1}{\epsilon}\sum_l \lambda_l \rho_l\right)^2 < 1$. If $\xi^*\Pi M(\rho)\xi \leq 0$, then $\xi^*\Pi M_\epsilon(\rho)\xi \geq \xi^*\Pi M(\rho)\xi$ and $\xi^*\Pi A_\epsilon(\rho)\xi \geq \xi^*\Pi A(\rho)\xi \geq \kappa\xi^*\xi$. If $\xi^*\Pi M(\rho)\xi > 0$, then $\xi^*\Pi M_\epsilon(\rho)\xi \geq 0$ and $\xi^*\Pi A_\epsilon(\rho)\xi \geq \frac{1}{2}\xi^*\Pi\xi \geq \kappa\xi^*\xi$, so ΠA_ϵ is uniformly coercive on $(\mathbb{R}^+)^d$ with coefficient κ. □

We now state an existence result for $(V_\epsilon(\mu))$.

Proposition 2.3.10. *Under Condition (C), for $\epsilon > 0$, there exists a solution $p_\epsilon \in C([0,T], L)$ to $(V_\epsilon(\mu))$.*

Proof. We compute standard energy estimates of p_ϵ^m for $m \geq 1$. We multiply (2.3.17) by $g_{\epsilon,i}^m(t)$ and add these equations for $i = 1, ..., m$. This gives

$$\frac{1}{2}\frac{d}{dt}\left|\sqrt{\Pi}p_\epsilon^m\right|_d^2 + \left(\partial_x p_\epsilon^m(t), \Pi A_\epsilon\left(\left(p_\epsilon^m\right)^+(t)\right)\partial_x p_\epsilon^m(t)\right)_d = 0. \tag{2.3.21}$$

By the uniform coercivity of the function ΠA_ϵ, we obtain that

$$-\frac{1}{2}\frac{d}{dt}\left|\sqrt{\Pi}p_\epsilon^m\right|_d^2 = \left(\partial_x p_\epsilon^m(t), \Pi A_\epsilon\left(\left(p_\epsilon^m\right)^+(t)\right)\partial_x p_\epsilon^m(t)\right)_d \geq \kappa\left|\partial_x p_\epsilon^m(t)\right|_d^2 \geq 0.$$

Therefore, we get the following inequalities:

$$l_{min}(\Pi)\sup_{t\in[0,T]}|p_\epsilon^m(t)|_d^2 \leq \sup_{t\in[0,T]}|\sqrt{\Pi}p_\epsilon^m(t)|_d^2 \leq |\sqrt{\Pi}p_\epsilon^m(0)|_d^2 \leq l_{max}(\Pi)|p_\epsilon^m(0)|_d^2 \leq l_{max}(\Pi)|p_0|_d^2, \tag{2.3.22}$$

$$\int_0^T \kappa|\partial_x p_\epsilon^m(t)|_d^2 dt \leq \frac{1}{2}\left(|\sqrt{\Pi}p_\epsilon^m(0)|_d^2 - |\sqrt{\Pi}p_\epsilon^m(T)|_d^2\right) \leq \frac{l_{max}(\Pi)}{2}|p_0|_d^2, \tag{2.3.23}$$

$$\int_0^T ||p_\epsilon^m(t)||_d^2 dt = \int_0^T |p_\epsilon^m(t)|_d^2 dt + \int_0^T |\partial_x p_\epsilon^m(t)|_d^2 dt \leq \left(T\frac{l_{max}(\Pi)}{l_{min}(\Pi)} + \frac{l_{max}(\Pi)}{2\kappa}\right)|p_0|_d^2. \tag{2.3.24}$$

We see that the sequence $(p_\epsilon^m)_{m\geq1}$ remains in a bounded set of $L^2([0,T];H)\cap L^\infty([0,T];L)$, so there exists an element $p_\epsilon \in L^2([0,T];H)\cap L^\infty([0,T];L)$ and a subsequence, for notational simplicity also called $(p_\epsilon^m)_{m\geq1}$, that has the following convergence:

$$\begin{aligned} p_\epsilon^m &\underset{m\to\infty}{\rightarrow} p_\epsilon \text{ in } L^2([0,T];H) \text{ weakly} \\ p_\epsilon^m &\underset{m\to\infty}{\rightarrow} p_\epsilon \text{ in } L^\infty([0,T];L) \text{ weakly-*}. \end{aligned}$$

We now show that there exists a subsequence of $(p_\epsilon^m)_{m\geq1}$ that converges a.e. on $(0,T)\times\mathbb{R}$ to p_ϵ. For $q \in H$, let us define the function $G_\epsilon q \in H'$, by

$$\langle G_\epsilon q, v\rangle = \left(\partial_x v, \Pi A_\epsilon\left(q^+\right)\partial_x q\right)_d.$$

for $v \in H$. Then the equality (2.3.17) rewrites:

$$\forall i \in \{1, ..., m\}, \frac{d}{dt}\left(w_i, \Pi p_\epsilon^m\right)_d + \langle G_\epsilon p_\epsilon^m, w_i\rangle = 0. \tag{2.3.25}$$

As by Lemma 2.3.7, the matrices in $A_\epsilon\left(\left(\mathbb{R}^+\right)^d\right)$ are uniformly bounded by $\frac{1}{2}\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)$, we have for $q \in H$:

$$||G_\epsilon q||_{H'} = \sup_{v\in H, ||v||_H\leq1}\left(\partial_x v, \Pi A_\epsilon\left(q^+\right)\partial_x q\right)_d \leq \frac{d^2}{2}\left(1 + \frac{\lambda_{max}}{\lambda_{min}}\right)||\Pi||_\infty||q||_d. \tag{2.3.26}$$

Since the family $(p_\epsilon^m)_{m\geq1}$ is bounded in $L^2([0,T],H)$, the family $(G_\epsilon p_\epsilon^m)_{m\geq1}$ is bounded in $L^2([0,T],H')$ and the computations in [97, (iii) p. 285] give that for any bounded open subset $\mathcal{O} \subset \mathbb{R}$, modulo the extraction of a subsequence,

$$p_{\epsilon|\mathcal{O}}^m \rightarrow p_{\epsilon|\mathcal{O}} \text{ in } L^2([0,T], L(\mathcal{O})) \text{ strongly and a.e. on } [0,T]\times\mathcal{O}.$$

We define for $n \geq 1$, $\mathcal{O}_n = (-n,n)$, so $\mathcal{O}_n \subset \mathcal{O}_{n+1}$, and $\bigcup_{n\geq1}\mathcal{O}_n = \mathbb{R}$. By diagonal extraction, we get from $(p_\epsilon^m)_{m\geq1}$ a subsequence, called $\left(p_\epsilon^{\phi(m)}\right)_{m\geq1}$ such that

$$\forall n \geq 1, p_{\epsilon|\mathcal{O}_n}^{\phi(m)} \underset{m\to\infty}{\rightarrow} p_{\epsilon|\mathcal{O}_n} \text{ strongly in } L^2([0,T]; L(\mathcal{O}_n)), \tag{2.3.27}$$

$$p_\epsilon^{\phi(m)} \rightarrow p_\epsilon \text{ a.e. on } [0,T]\times\mathbb{R}. \tag{2.3.28}$$

We show that p_ϵ is a solution to the variational formulation $V_\epsilon(\mu)$. For $1 \le j \le m$ and $\psi \in C^1([0,T],\mathbb{R})$ with $\psi(T) = 0$, we have, through integration by parts, the equality:

$$-\int_0^T \left(\psi'(t)w_j, \Pi p_\epsilon^{\phi(m)}(t)\right)_d dt + \int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right)\partial_x p_\epsilon^{\phi(m)}(t)\right)_d dt = \left(w_j, \Pi p_0^{\phi(m)}\right)_d \psi(0).$$

The sequence $\left(p_0^{\phi(m)}\right)_{m\ge0}$ converges strongly to p_0 in L so

$$\left(w_j, \Pi p_0^{\phi(m)}\right)_d \psi(0) \to (w_j, \Pi p_0)_d \psi(0).$$

The sequence $\left(p_\epsilon^{\phi(m)}\right)_{m\ge0}$ converges weakly to p_ϵ in $L^2([0,T],L)$ so

$$-\int_0^T \left(\psi'(t)w_j, \Pi p_\epsilon^{\phi(m)}(t)\right)_d dt \to -\int_0^T \left(\psi'(t)w_j, \Pi p_\epsilon(t)\right)_d dt.$$

For the remaining term, we have

$$\begin{aligned}
&\left|\int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon(p_\epsilon^+(t))\partial_x p_\epsilon(t)\right)_d dt - \int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right)\partial_x p_\epsilon^{\phi(m)}(t)\right)_d dt\right| \\
&\quad\le \left|\int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon(p_\epsilon^+(t))\left(\partial_x p_\epsilon(t) - \partial_x p_\epsilon^{\phi(m)}(t)\right)\right)_d dt\right| \qquad (2.3.29)\\
&\quad+ \left|\int_0^T \left(\psi(t)\partial_x w_j, \Pi\left(A_\epsilon\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right) - A_\epsilon(p_\epsilon^+(t))\right)\partial_x p_\epsilon^{\phi(m)}(t)\right)_d dt\right|. \qquad (2.3.30)
\end{aligned}$$

For the term in (2.3.29), the sequence $\left(\partial_x p_\epsilon^{\phi(k)}\right)_{k\ge0}$ converges weakly to $\partial_x p_\epsilon$ in $L^2([0,T];L)$. In addition, by Lemma 2.3.7, A_ϵ is bounded, so the function $t \to \psi(t)A_\epsilon^*(p_\epsilon^+(t))\Pi\partial_x w_j$ belongs to $L^2([0,T];L)$ and we have the convergence

$$\left|\int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon(p_\epsilon^+(t))\left(\partial_x p_\epsilon(t) - \partial_x p_\epsilon^{\phi(m)}(t)\right)\right)_d dt\right| \underset{m\to\infty}{\to} 0.$$

For the term in (2.3.30), using the Cauchy-Schwarz inequality we have that

$$\begin{aligned}
&\left|\int_0^T \left(\psi(t)\left(A_\epsilon^*\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right) - A_\epsilon^*\left(p_\epsilon^+(t)\right)\right)\Pi\partial_x w_j, \partial_x p_\epsilon^{\phi(m)}(t)\right)_d dt\right|^2 \\
&\quad\le \left(\int_0^T \left|\psi(t)\left(A_\epsilon^*\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right) - A_\epsilon^*(p_\epsilon^+(t))\right)\Pi\partial_x w_j\right|_d^2 dt\right)\left(\int_0^T |\partial_x p_\epsilon^{\phi(k)}(t)|_d^2 dt\right) \\
&\quad\le \frac{l_{max}(\Pi)}{2\kappa}|p_0|_d^2 \int_0^T \left|\psi(t)\left(A_\epsilon^*\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right) - A_\epsilon^*(p_\epsilon^+(t))\right)\Pi\partial_x w_j\right|_d^2 dt,
\end{aligned}$$

where we used the energy estimate (2.3.23) in the last inequality. The function A_ϵ is continuous on $(\mathbb{R}^+)^d$ as shown in the proof of Lemma 2.3.6, and $p_\epsilon^{\phi(m)} \to p_\epsilon$ a.e. on $[0,T]\times\mathbb{R}$, so $A_\epsilon\left(\left(p_\epsilon^{\phi(m)}\right)^+\right) \to A_\epsilon\left(p_\epsilon^+\right)$ a.e. on $[0,T]\times\mathbb{R}$. By Lemma 2.3.7, A_ϵ is uniformly bounded so we have through dominated convergence,

$$\int_0^T \left|\psi(t)\left(A_\epsilon^*\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right) - A_\epsilon^*(p_\epsilon^+(t))\right)\Pi\partial_x w_j\right|_d^2 dt \underset{m\to\infty}{\to} 0.$$

Thus we have shown that

$$\int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon\left(\left(p_\epsilon^{\phi(m)}\right)^+(t)\right)\partial_x p_\epsilon^{\phi(m)}(t)\right)_d dt \underset{m\to\infty}{\to} \int_0^T \left(\psi(t)\partial_x w_j, \Pi A_\epsilon\left(p_\epsilon^+(t)\right)\partial_x p_\epsilon(t)\right)_d dt. \quad (2.3.31)$$

When we gather the convergence results and use the fact that the sequence $(w_j)_{j\geq 1}$ is total in H, we obtain that

$$\forall v \in H, -\int_0^T (\psi'(t)v, \Pi p_\epsilon(t))_d dt + \int_0^T (\psi(t)\partial_x v, \Pi A_\epsilon(p_\epsilon^+(t))\partial_x p_\epsilon(t))_d dt = (v, \Pi p_0)_d \psi(0),$$

which rewrites:

$$\forall v \in H, -\int_0^T (\psi'(t)v, p_\epsilon(t))_d dt + \int_0^T (\psi(t)\partial_x v, A_\epsilon(p_\epsilon^+(t))\partial_x p_\epsilon(t))_d dt = (v, p_0)_d \psi(0), \tag{2.3.32}$$

If moreover ψ belongs to $C_c^\infty((0,T))$, we obtain that p_ϵ satisfies (2.3.13) in the distributional sense on $(0,T)$. For $v \in H$, the function $t \to (v, p_\epsilon(t))_d$ belongs to $H^1((0,T))$, as the functions $t \in (0,T) \to (v, p_\epsilon(t))_d$ and $t \in (0,T) \to (\partial_x v, A_\epsilon(p_\epsilon^+(t))\partial_x p_\epsilon(t))_d$ both belong to $L^2((0,T))$. Thanks to [19, Corollary 8.10], the following integration by parts formula also holds:

$$\forall v \in H, -\int_0^T \psi'(t)(v, p_\epsilon(t))_d dt + \int_0^T \psi(t)(\partial_x v, A_\epsilon(p_\epsilon^+(t))\partial_x p_\epsilon(t))_d dt = (v, p_\epsilon(0))_d \psi(0). \tag{2.3.33}$$

If we choose $\psi(0) \neq 0$, then by comparing (2.3.33) with (2.3.32), we have that:

$$\forall v \in H, (v, p_\epsilon(0) - p_0)_d = 0, \tag{2.3.34}$$

and this concludes the proof for the existence of a solution to $(V_\epsilon(\mu))$. We now show that $p_\epsilon \in C([0,T], L)$. The function p_ϵ satisfies:

$$\forall v \in H, \frac{d}{dt}(v, p_\epsilon)_d + \left(v, A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon\right)_d = 0,$$

in the distributional sense on $(0,T)$, with

$$p_\epsilon \in L^2([0,T], H), \tag{2.3.35}$$
$$A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon \in L^2([0,T], L). \tag{2.3.36}$$

Then by [97, III. Lemma 1.2], p_ϵ is a.e. equal to a function belonging to $C([0,T], L)$. □

Proposition 2.3.11. *Under Condition (C), for $\epsilon > 0$, the solutions to $V_\epsilon(\mu)$ are non negative.*

Proof. Let $p_\epsilon = (p_{\epsilon,1}, ..., p_{\epsilon,d})$ be a solution to the variational problem $V_\epsilon(\mu)$. For $x \in \mathbb{R}$, let $x^- = \min(x, 0)$. We take $p_\epsilon^- = (p_{\epsilon,1}^-, ..., p_{\epsilon,d}^-)$ as a test function in (2.3.13). Thanks to (2.3.35)-(2.3.36), we obtain by [97, Lemma 1.2 p. 261]:

$$\frac{1}{2}\frac{d}{dt}(p_\epsilon^-, p_\epsilon)_d + (\partial_x p_\epsilon^-, A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon)_d = 0 \tag{2.3.37}$$

in the sense of distributions. For $f \in H^1(\mathbb{R}), \partial_x f^- = 1_{\{f<0\}}\partial_x f$, so we have $\forall i \in \{1, ..., d\}$, $p_{\epsilon,i}^+ \partial_x p_{\epsilon,i}^- = 0$, $\partial_x p_{\epsilon,i}^- \partial_x p_{\epsilon,i} = (\partial_x p_{\epsilon,i}^-)^2$. As a consequence,

$$\partial_x p_{\epsilon,i}^- \sum_{j\neq i} M_{\epsilon,ij}(p_\epsilon^+)\partial_x p_{\epsilon,j} = \partial_x p_{\epsilon,i}^- \frac{\lambda_i p_{\epsilon,i}^+}{\epsilon \vee \left(\sum_l \lambda_l p_{\epsilon,l}^+\right)} \left(\sum_{j\neq i} \frac{\sum_{l\neq j}(\lambda_l - \lambda_j)p_{\epsilon,l}^+}{\epsilon \vee \left(\sum_l \lambda_l p_{\epsilon,l}^+\right)} \partial_x p_{\epsilon,j}\right) = 0,$$

$$(\partial_x p_\epsilon^-, A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon)_d = \int_{\mathbb{R}} \sum_{i,j=1}^d \partial_x p_{\epsilon,i}^- A_{\epsilon,ij}(p_\epsilon^+)\partial_x p_{\epsilon,j} dx = \int_{\mathbb{R}} \sum_{i=1}^d \partial_x p_{\epsilon,i}^- A_{\epsilon,ii}(p_\epsilon^+)\partial_x p_{\epsilon,i} dx.$$

Equality (2.3.37) simplifies into

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}} \sum_{i=1}^d \left(p_{\epsilon,i}^-\right)^2 dx + \int_{\mathbb{R}} \sum_{i=1}^d A_{\epsilon,ii}(p_\epsilon^+) \left(\partial_x p_{\epsilon,i}^-\right)^2 dx = 0.$$

Let us show that for $1 \leq i \leq d$,

$$A_{\epsilon,ii}(p_\epsilon^+) \left(\partial_x p_{\epsilon,i}^-\right)^2 \geq \frac{\lambda_{min}}{2\lambda_{max}} \left(\partial_x p_{\epsilon,i}^-\right)^2 \geq 0. \tag{2.3.38}$$

For $1 \leq i \leq d$,

$$2A_{\epsilon,ii}(p_\epsilon^+) = 1 + \frac{\left(\sum_{l\neq i}\lambda_l p_{\epsilon,l}^+\right)\left(\lambda_i \sum_{l\neq i} p_{\epsilon,l}^+\right)}{\epsilon^2 \vee \left(\sum_{l=1}^d \lambda_l p_{\epsilon,l}^+\right)^2} - \frac{\left(\sum_{l\neq i}\lambda_l p_{\epsilon,l}^+\right)^2}{\epsilon^2 \vee \left(\sum_{l=1}^d \lambda_l p_{\epsilon,l}^+\right)^2}.$$

We distinguish the cases $p_{\epsilon,i}^+ > 0$ and $p_{\epsilon,i}^+ = 0$. In the first case, $\partial_x p_{\epsilon,i}^- = 1_{\{p_{\epsilon,i}<0\}}\partial_x p_{\epsilon,i} = 0$, so (2.3.38) is true. In the second case, let us define $z := \left(\sum_{l\neq i}\lambda_l p_{\epsilon,l}^+\right) = \left(\sum_{l=1}^d \lambda_l p_{\epsilon,l}^+\right)$, and notice that $\left(\lambda_i \sum_{l\neq i} p_{\epsilon,l}^+\right) \geq \frac{\lambda_{min}}{\lambda_{max}} z$. Thus we obtain:

$$2A_{\epsilon,ii}(p_\epsilon^+) \geq 1 + \left(\frac{\lambda_{min}}{\lambda_{max}} - 1\right)\frac{z^2}{\epsilon^2 \vee z^2} \geq \frac{\lambda_{min}}{\lambda_{max}},$$

as $0 \leq \frac{z^2}{\epsilon^2 \vee z^2} \leq 1$, and (2.3.38) is also true.

Consequently, $\int_{\mathbb{R}} \sum_{i=1}^d \left(p_{\epsilon,i}^-\right)^2(t)$ is non increasing in time. As $\int_{\mathbb{R}} \sum_{i=1}^d \left(p_{\epsilon,i}^-\right)^2(0)dx = \int_{\mathbb{R}} \sum_{i=1}^d \left(p_{0,i}^-\right)^2 dx = 0$, we conclude that p_ϵ has every component non-negative. Let us remark that to obtain the non negativity of the solutions to $V_\epsilon(\mu)$, it is sufficient and easier to prove directly that $A_{\epsilon,ii} \geq 0$ for $1 \leq i \leq d$, but Inequality (2.3.38) will be reused for the proof of Proposition 2.5.7 below. □

Before showing the existence of a solution to $V_{L^2}(\mathbb{R})$, we compute the energy estimates on p_ϵ.

Lemma 2.3.12. *Let $\epsilon > 0$, and let p_ϵ be a solution to $V_\epsilon(\mu)$. The following energy estimates hold.*

$$\sup_{t\in[0,T]} |p_\epsilon(t)|_d^2 \leq \frac{l_{max}(\Pi)}{l_{min}(\Pi)}|p_0|_d^2, \tag{2.3.39}$$

$$\int_0^T |\partial_x p_\epsilon(t)|_d^2 dt \leq \frac{l_{max}(\Pi)}{2\kappa}|p_0|_d^2, \tag{2.3.40}$$

$$\int_0^T ||p_\epsilon||_d^2 dt = \int_0^T |p_\epsilon|_d^2 dt + \int_0^T |\partial_x p_\epsilon|_d^2 dt \leq \left(T\frac{l_{max}(\Pi)}{l_{min}(\Pi)} + \frac{l_{max}(\Pi)}{2\kappa}\right)|p_0|_d^2. \tag{2.3.41}$$

Proof. We obtain the energy estimates by taking p_ϵ as the test function in (2.3.13), and the computations are the same as for (2.3.22)-(2.3.24). □

Proposition 2.3.13. *Under Condition (C), $V_{L^2}(\mu)$ has a solution p, which satisfies the same energy estimates as the solutions to $V_\epsilon(\mu)$, and that are given by (2.3.39)-(2.3.41). Moreover, a.e. on $(0,T)\times\mathbb{R}$, $\sum_{i=1}^d p_i(t,x) = \mu * h_t(x)$.*

Proof. For $\epsilon > 0$, let us now denote by p_ϵ a solution to $V_\epsilon(\mu)$, which exists under Condition (C) according to Proposition 2.3.10. The family $(p_\epsilon)_{\epsilon>0}$ is bounded in $L^2([0,T];H) \cap L^\infty([0,T];L)$, as the energy estimates (2.3.39)-(2.3.41) have bounds that do not depend on ϵ. There exists a limit function $p \in L^2([0,T];H) \cap L^\infty([0,T];L)$ and a subsequence $(\epsilon_k)_{k\geq 1}$ decreasing to 0, such that

$$\begin{aligned} p_{\epsilon_k} &\rightarrow p \text{ in } L^2([0,T];H) \text{ weakly,} \\ p_{\epsilon_k} &\rightarrow p \text{ in } L^\infty([0,T];L) \text{ weakly-*,} \end{aligned}$$

and p also satisfies Inequalities (2.3.39)-(2.3.41), where we replace p_ϵ by p. Moreover, in the sense of [97, III, Lemma 1.1], we check with arguments similar to (2.3.26) that the family $\left(\frac{dp_{\epsilon_k}}{dt}\right)_{k\geq 1}$ is bounded in $L^2([0,T],H')$ and by [97, III, Theorem 2.3], for any bounded open subset $\mathcal{O} \subset \mathbb{R}$, there is a subsequence of $(p_{\epsilon_k})_{k\geq 1}$ converging to p in $L^2([0,T],L(\mathcal{O}))$ strongly and a.e. on $[0,T]\times\mathcal{O}$. Then by diagonal extraction, similarly to the proof of Proposition 2.3.10, we can also assume that:

$$p_{\epsilon_k} \rightarrow p \text{ a.e. on } [0,T]\times\mathbb{R},$$

and that p is non negative as p_ϵ is non negative for $\epsilon > 0$. For any function $\psi \in C^1([0,T],\mathbb{R})$, with $\psi(T) = 0$, and $k \geq 1$, we have

$$\forall v \in H, -\int_0^T (\psi'(t)v, p_{\epsilon_{\phi(k)}}(t))_d dt + \int_0^T (\psi(t)\partial_x v, A_{\epsilon_{\phi(k)}}(p_{\epsilon_{\phi(k)}}(t))\partial_x p_{\epsilon_{\phi(k)}}(t))_d dt = (v, p_0)_d \psi(0).$$

The sequence $\left(p_{\epsilon_{\phi(k)}}\right)_{k\geq 0}$ converges weakly to p in $L^2([0,T];H)$ so

$$-\int_0^T (\psi'(t)v, p_{\epsilon_{\phi(k)}}(t))_d dt \underset{k\to\infty}{\to} -\int_0^T (\psi'(t)v, p(t))_d dt.$$

We show that p takes values in $\mathcal{D}$ a.e. on $[0,T]\times\mathbb{R}$. As for $\rho\in(\mathbb{R}_+)^d$, and $1\leq j\leq d$, $\sum_{i=1}^d M_{\epsilon,ij}(\rho)=0$, we have

$$\sum_{i=1}^d \left(A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon\right)_i = \frac{1}{2}\partial_x\left(\sum_{i=1}^d p_{\epsilon,i}\right) + \frac{1}{2}\sum_{j=1}^d\left(\sum_{i=1}^d M_{\epsilon,ij}(p_\epsilon^+)\right)\partial_x p_{\epsilon,j} = \frac{1}{2}\partial_x\left(\sum_{i=1}^d p_{\epsilon,i}\right).$$

Thus, in the sense of distributions on $(0,T)$, for $\tilde{v}\in H^1(\mathbb{R})$, and $v:=(\tilde{v},...,\tilde{v})\in H$,

$$0=\frac{d}{dt}(v,p_\epsilon)_d + (\partial_x v, A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon)_d \quad = \quad \frac{d}{dt}\int_{\mathbb{R}}\tilde{v}\left(\sum_{i=1}^d p_{\epsilon,i}\right)dx + \frac{1}{2}\int_{\mathbb{R}}\partial_x\tilde{v}\partial_x\left(\sum_{i=1}^d p_{\epsilon,i}\right)dx.$$

We also have for $\epsilon>0$, $\sum_{i=1}^d p_{\epsilon,i}(0,x)dx=\mu(dx)$. Then the function $\sum_{i=1}^d p_{\epsilon,i}$ is a solution to the formulation $H_{L^2}(\mu)$ defined by :

$$z\in L^2([0,T];H^1(\mathbb{R}))\cap L^\infty([0,T];L^2(\mathbb{R})) \tag{2.3.42}$$

$$\forall w\in H^1(\mathbb{R}),\ \frac{d}{dt}\int_{\mathbb{R}} wz dx + \frac{1}{2}\int_{\mathbb{R}}(\partial_x w)(\partial_x z)\,dx = 0 \text{ in the sense of distributions on } (0,T), \tag{2.3.43}$$

$$z(0)=\mu. \tag{2.3.44}$$

The problem $H_{L^2}(\mu)$ is a variational formulation to the heat PDE

$$\begin{aligned}\partial_t z &= \frac{1}{2}\partial_{xx} z\\ z(0) &= \mu,\end{aligned}$$

and it is well known that the solution of $H_{L^2}(\mu)$ is unique and expressed as the convolution of μ with the heat kernel $h_t(x)=\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{x^2}{2t}\right)$. Consequently, for a subsequence $(p_{\epsilon_k})_{k\geq 1}$ such that $p_{\epsilon_k}\underset{k\to\infty}{\to} p$ a.e. on $(0,T]\times\mathbb{R}$, the sequence $(z_{\epsilon_k})_{k\geq 1}$, defined by $z_{\epsilon_k}=\sum_{i=1}^d p_{\epsilon_k,i}$ satisfies

$$\text{for almost every } (t,x)\in(0,T]\times\mathbb{R}, \forall k\geq 1,\ z_{\epsilon_k}(t,x)=\mu * h_t(x)>0,$$

and the value of $z_{\epsilon_k}(t,x)$ is independent from ϵ_k. If we define $z_0=\sum_{i=1}^d p_i$, we have that $z_0(t,x)=\mu*h_t(x)>0$ a.e. on $(0,T]\times\mathbb{R}$. As the limit p has every component non-negative, we can thus conclude that p takes values in $\mathcal{D}$, a.e. on $(0,T]\times\mathbb{R}$. Then for almost every $(t,x)\in[0,T]\times\mathbb{R}$, there exists $k_0(t,x)$ such that $\forall k\geq k_0(t,x)$, $\sum\lambda_i p_{\epsilon_{\phi(k)},i}(t,x)\geq\frac{1}{2}\sum_i\lambda_i p_i(t,x)>\epsilon_{\phi(k)}$. For $k\geq k_0(t,x)$, $A_{\epsilon_{\phi(k)}}(p_{\epsilon_{\phi(k)}}(t,x))=A(p_{\epsilon_{\phi(k)}}(t,x))$, and as A is continuous on $\mathcal{D}$, we have

$$A_{\epsilon_{\phi(k)}}(p_{\epsilon_{\phi(k)}}(t,x))\underset{k\to\infty}{\to} A(p(t,x)) \text{ a.e.}. \tag{2.3.45}$$

By Lemma 2.3.7, the family $\left(A_{\epsilon_{\phi(k)}}\right)_{k\geq 1}$ has bounds uniform in k. Hence

$$\int_0^T(\psi(t)\partial_x v, A_{\epsilon_{\phi(k)}}(p_{\epsilon_{\phi(k)}}(t))\partial_x p_{\epsilon_{\phi(k)}}(t))_d dt \underset{k\to\infty}{\to}\int_0^T(\psi(t)\partial_x v, A(p(t))\partial_x p(t))_d dt,$$

in the same manner as in (2.3.30). Consequently, we have that

$$\forall v\in H, -\int_0^T(\psi'(t)v,p(t))_d dt + \int_0^T(\psi(t)\partial_x v, A(p(t))\partial_x p(t))_d dt = (v,p_0)_d\psi(0). \tag{2.3.46}$$

Similarly to the proof of Proposition 2.3.10, we prove (2.3.9) and obtain $p\in C([0,T],L)$. Finally we repeat the same arguments as in (2.3.34) to conclude that p is a solution to $V_{L^2}(\mu)$. $\square$

2.3.3 Proof of Theorem 2.2.1

We first introduce a lemma that will be used later in the proof.

Lemma 2.3.14. *Let $\gamma \geq 0$ and let ϕ be a non-negative function, s.t. $\forall t \in (0,T]$, $\int_t^T \phi^2(s)ds \leq \frac{\gamma}{\sqrt{t}}$.*

Then for $t \in (0,T]$ $\int_0^t \phi(s)ds \leq \frac{\sqrt{\gamma}t^{\frac{1}{4}}}{2^{\frac{1}{4}}-1}$.

Proof. With monotone convergence, then using the Cauchy-Schwarz inequality, we obtain, for $t \in (0,T]$:

$$\begin{aligned}\int_0^t \phi(s)ds &= \sum_{k=0}^{\infty}\int_{t2^{-(k+1)}}^{t2^{-k}} \phi(s)ds \\ &\leq \sum_{k=0}^{\infty}\sqrt{t}2^{-(\frac{k+1}{2})}\sqrt{\int_{t2^{-(k+1)}}^{t2^{-k}} \phi^2(s)ds} \leq \sum_{k=0}^{\infty}\sqrt{t}2^{-(\frac{k+1}{2})}\sqrt{\int_{t2^{-(k+1)}}^{T} \phi^2(s)ds} \\ &\leq \sqrt{t}\sum_{k=0}^{\infty}2^{-(\frac{k+1}{2})}\sqrt{\frac{\gamma}{2^{-(\frac{k+1}{2})}\sqrt{t}}} = \frac{\sqrt{\gamma}t^{\frac{1}{4}}}{2^{\frac{1}{4}}-1}.\end{aligned}$$

□

Now let $(\sigma_k)_{k\geq 0}$ be a sequence decreasing to 0 as $k \to \infty$. In order to rely on Theorem 2.3.5, we approximate μ by a sequence of measures μ_{σ_k}, weakly converging to μ and that have densities in $L^2(\mathbb{R})$. To do so, we apply a convolution product on μ by setting $\mu_{\sigma_k} := \mu * h_{\sigma_k^2}$. The measure μ_{σ_k} is absolutely continuous with respect to the Lebesgue measure, with the density

$$\mu_{\sigma_k}(x) := \frac{1}{\sqrt{2\pi\sigma_k^2}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2\sigma_k^2}}d\mu(y).$$

With Jensen's inequality and Fubini's theorem, we check that $\mu_{\sigma_k} \in L^2(\mathbb{R})$. Indeed,

$$\int_{\mathbb{R}} \mu_{\sigma_k}^2(x)dx = \int_{\mathbb{R}}\left(\frac{1}{\sqrt{2\pi\sigma_k^2}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2\sigma_k^2}}d\mu(y)\right)^2 dx \leq \frac{1}{2\pi\sigma_k^2}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{\sigma_k^2}}dx\right)d\mu(y) = \frac{1}{2\sqrt{\pi}\sigma_k}.$$

Let us define $p_{0,\sigma_k} := (\alpha_1\mu_{\sigma_k}, ..., \alpha_d\mu_{\sigma_k}) \in L$ and let us denote by $p_{\sigma_k} = (p_{\sigma_k,1}, ..., p_{\sigma_k,d})$ a solution to the variational formulation $V_{L^2}(\mu_{\sigma_k})$, which exists as an application of Theorem 2.3.5, and that satisfies equality (2.3.46) with initial condition p_{0,σ_k} as in the proof of Proposition 2.3.13. We compute the same energy estimates as previously. As $p_{\sigma_k,i}$ is non-negative for $1 \leq i \leq d$ and $\sum_{i=1}^d p_{\sigma_k,i} = \mu_{\sigma_k} * h_t$ a.e. on $(0,T) \times \mathbb{R}$,

$$\text{for a.e. } t \text{ on } (0,T),\ |p_{\sigma_k}(t)|_d^2 \leq \left|\left|\sum_{i=1}^d p_{\sigma_k,i}(t)\right|\right|_{L^2}^2 = \left|\left|\mu * h_{\sigma_k^2} * h_t\right|\right|_{L^2}^2 \leq \frac{1}{2\sqrt{\pi t}},$$

$$\int_0^T |p_{\sigma_k}(t)|_d^2 dt \leq \int_0^T \frac{1}{2\sqrt{\pi t}}dt = \sqrt{\frac{T}{\pi}},$$

Using Πp_{σ_k} as a test function in (2.3.9) and the fact that ΠA is uniformly coercive on $\mathcal{D}$ with coefficient κ,

$$\forall t \in (0,T),\ \int_t^T |\partial_x p_{\sigma_k}(s)|_d^2 ds \leq \frac{1}{2\kappa}|\sqrt{\Pi}p_{\sigma_k}(t)|_d^2 \leq \frac{l_{max}(\Pi)}{4\kappa\sqrt{\pi t}}. \tag{2.3.47}$$

Let us remark that the estimates $\int_0^T |p_{\sigma_k}(t)|_d^2 dt$ and $\int_t^T |\partial_x p_{\sigma_k}(t)|_d^2 ds$, for $t \in (0,T)$ have bounds that are independent from the choice of σ_k. Then for a sequence $(s_n)_{n\geq 1}$ with values in $(0,T)$ and decreasing to 0 as $n \to \infty$, there exists a function p defined a.e. on $(0,T] \times \mathbb{R}$ such that for each $n \geq 1$, there exists a converging subsequence called again p_{σ_k} with the following convergences:

$$p_{\sigma_k} \to p \text{ in } L^2((0,T];L) \text{ weakly},$$

$$
\begin{aligned}
p_{\sigma_k} &\rightarrow p \text{ in } L^2([s_n,T];H) \text{ weakly},\\
p_{\sigma_k} &\rightarrow p \text{ in } L^\infty([s_n,T];L) \text{ weakly-*}.
\end{aligned}
$$

Similarly to the proof of Proposition 2.3.13, we can also suppose, modulo the extraction of a subsection that

$$
p_{\sigma_k} \rightarrow p \text{ a.e. on } [s_n,T]\times\mathbb{R}.
$$

By diagonal extraction, we obtain a subsequence, called again p_{σ_k} such that

$$
\begin{aligned}
p_{\sigma_k} &\rightarrow p \text{ in } L^2((0,T];L) \text{ weakly},\\
p_{\sigma_k} &\rightarrow p \text{ in } L^2_{loc}((0,T];H) \text{ weakly},\\
p_{\sigma_k} &\rightarrow p \text{ in } L^\infty_{loc}((0,T];L) \text{ weakly-*},\\
p_{\sigma_k} &\rightarrow p \text{ in } (0,T]\times\mathbb{R}.
\end{aligned}
$$

Let us remark that $p \geq 0$ as $p_{\sigma_k} \geq 0$ for $k \geq 1$. Moreover, for a.e. $(t,x) \in (0,T]\times\mathbb{R}$, $\mu_{\sigma_k} * h_t(x) = \sum_{i=1}^d p_{\sigma_k,i}(t,x) \underset{k\rightarrow\infty}{\rightarrow} \sum_{i=1}^d p(t,x) = \mu_{\sigma_k} * h_t(x)$, so p takes values in $\mathcal{D}$ a.e. on $(0,T]\times\mathbb{R}$. Let us prove that p is a solution to $V(\mu)$. For $k \geq 0$, as p_{σ_k} satisfies (2.3.46) with initial condition p_{0,σ_k}, for $\psi \in \mathcal{C}^1([0,T],\mathbb{R})$ such that $\psi(T) = 0, v \in H^1(\mathbb{R})$, we have that

$$
-\int_0^T (\psi'(s)v, p_{\sigma_k}(s))_d ds + \int_0^T (\psi(s)\partial_x v, A(p_{\sigma_k})\partial_x p_{\sigma_k})_d ds \quad = \quad (v, p_{0,\sigma_k})_d \psi(0).
$$

We study the limit as $k \rightarrow \infty$. First,

$$
\int_0^T (\psi'(s)v, p_{\sigma_k}(s))_d \, ds \quad \rightarrow \quad \int_0^T (\psi'(s)v, p(s))_d \, ds,
$$

as $p_{\sigma_k} \rightarrow p$ in $L^2((0,T];L)$ weakly. Since $p_{0,\sigma_k} \rightarrow p_0$ weakly and $v \in H$ is continuous and bounded, we have the following convergence

$$
(v, p_{0,\sigma_k})_d \psi(0) \rightarrow \psi(0) \sum_{i=1}^d \alpha_i \int v_i(x)\mu(dx).
$$

To show that

$$
\int_0^T (\psi(s)\partial_x v, A(p_{\sigma_k})\partial_x p_{\sigma_k})_d \, ds \rightarrow \int_0^T (\psi(s)\partial_x v, A(p)\partial_x p)_d \, ds, \tag{2.3.48}
$$

we check that the family $(\partial_x p_{\sigma_k})_{k\geq 1}$ and $\partial_x p$ belong to a bounded subset of $L^1((0,T],L)$. More precisely, for $q \in (\partial_x p_{\sigma_k})_{k\geq 1} \cup \{\partial_x p\}$, using (2.3.47) we have that

$$
\forall t \in (0,T), \quad \int_t^T |\partial_x q(s)|_d^2 ds \leq \frac{l_{max}(\Pi)}{4\kappa\sqrt{\pi t}},
$$

and by Lemma 2.3.14, we have that

$$
\forall t \in (0,T), \quad \int_0^t |\partial_x q(s)|_d ds \leq \sqrt{\frac{l_{max}(\Pi)}{4\kappa\sqrt{\pi}}} \frac{t^{\frac{1}{4}}}{2^{\frac{1}{4}}-1}. \tag{2.3.49}
$$

We then write

$$
\begin{aligned}
\left|\int_0^T (\psi(s)\partial_x v, A(p_{\sigma_k})\partial_x p_{\sigma_k} - A(p)\partial_x p)_d ds\right| \quad &\leq \quad \left|\int_t^T (\psi(s)\partial_x v, A(p_{\sigma_k})\partial_x p_{\sigma_k} - A(p)\partial_x p)_d ds\right| \\
&\quad + \quad \left|\int_0^t (\psi(s)\partial_x v, A(p_{\sigma_k})\partial_x p_{\sigma_k})_d ds\right| + \left|\int_0^t (\psi(s)\partial_x v, A(p)\partial_x p)_d ds\right|.
\end{aligned}
$$

For fixed $t > 0$, the first term of the r.h.s. goes to 0 as $k \to \infty$ with the same reasoning used to obtain (2.3.31). Moreover,

$$\left|\int_0^t (\psi(s)\partial_x v, A(p_{\sigma_k})\partial_x p_{\sigma_k})_d ds\right| \leq \frac{d}{2}\left(1+\frac{\lambda_{max}}{\lambda_{min}}\right)||\psi||_\infty |\partial_x v|_d \int_0^t |\partial_x p_{\sigma_k}|_d ds, \quad (2.3.50)$$

$$\left|\int_0^t (\psi(s)\partial_x v, A(p)\partial_x p)_d ds\right| \leq \frac{d}{2}\left(1+\frac{\lambda_{max}}{\lambda_{min}}\right)||\psi||_\infty |\partial_x v|_d \int_0^t |\partial_x p|_d ds, \quad (2.3.51)$$

by the Cauchy-Schwarz inequality and Lemma 2.3.7. Using (2.3.49), the l.h.s. of Inequalities (2.3.50) and (2.3.51) can be made arbitrarily small, uniformly in k, choosing t small enough, so we obtain (2.3.48). Gathering the convergence results we get

$$-\int_0^T (\psi'(s)v, p(s))_d ds + \int_0^T (\psi(s)\partial_x v, A(p)\partial_x p)_d ds = \psi(0)\sum_{i=1}^d \alpha_i \int v_i(x)\mu(dx). \quad (2.3.52)$$

and this is enough to obtain (2.2.5). As it is easy to check that the function $s \to (v, p(s))$ belongs to $H^1((t,T))$ for $t \in (0,T)$, similarly to (2.3.33), we also have the following integration by parts formula:

$$\forall t \in (0,T), -\int_t^T (\psi'(s)v, p(s))_d ds + \int_t^T (\psi(s)\partial_x v, A(p)\partial_x p)_d ds = (v, p(t))_d \psi(t). \quad (2.3.53)$$

The integrals on the l.h.s converge to $-\int_0^T (\psi'(s)v, p(s))_d ds + \int_0^T (\psi(s)\partial_x v, A(p)\partial_x p)_d ds$ as $t \to 0$, as it is easy to check by Cauchy-Schwarz inequality that the functions $s \to (\psi'(s)v, p(s))_d$ and $s \to (\psi(s)\partial_x v, A(p)\partial_x p)_d$ belong to $L^1((0,T),\mathbb{R})$. If we choose $\psi(0) \neq 0$, we obtain that

$$\lim_{t\to 0}(p(t), v)_d = \sum_{i=1}^d \alpha_i \int v_i(x)\mu(dx), \quad (2.3.54)$$

by comparing (2.3.52) with (2.3.53) and this gives existence to $V(\mu)$. Moreover, we obtain, with the same arguments as the end of Proposition 2.3.10 $p \in C((0,T], L)$, and this concludes the proof.

2.3.4 Proof of Theorem 2.2.2

The existence of a solution to the variational formulation $V(\mu)$ will give the existence of a weak solution to the original SDE (2.1.1). This essentially comes from the equivalence between existence to a Fokker-Planck equation and existence to the corresponding martingale problem, established by [39]. To use that result, we first give a lemma that makes $V(\mu)$ compatible with the variational formulation described in [39].

Lemma 2.3.15. *Let p be a solution to $V(\mu)$. Then, for $1 \leq i \leq d$ and a.e. $s \in (0,T)$, $\left(\frac{\sum_k p_k}{\sum_k \lambda_k p_k}\lambda_i p_i\right)(s,\cdot) \in H^1(\mathbb{R})$ and $\frac{1}{2}\partial_x\left(\frac{\sum_k p_k}{\sum_k \lambda_k p_k}\lambda_i p_i\right)(s,\cdot) = (A(p)\partial_x p)_i(s,\cdot)$.*

Proof. It is sufficient to show that for $1 \leq i,j \leq d$ and a.e. $s \in (0,T)$, $\frac{p_i p_j}{\sum_k \lambda_k p_k}(s,\cdot) \in H^1(\mathbb{R})$ and

$$\partial_x\left(\frac{p_i p_j}{\sum_k \lambda_k p_k}\right)(s,\cdot) = \left(\frac{p_i\partial_x p_j + p_j \partial_x p_i}{\sum_k \lambda_k p_k} - \frac{p_i p_j}{\left(\sum_k \lambda_k p_k\right)^2}\sum_k \lambda_k \partial_x p_k\right)(s,\cdot),$$

as we conclude by linearity. For a.e. $s \in (0,T)$, $\frac{p_i p_j}{\sum_k \lambda_k p_k}(s,\cdot)$ and $\left(\frac{p_i\partial_x p_j + p_j \partial_x p_i}{\sum_k \lambda_k p_k} - \frac{p_i p_j}{\left(\sum_k \lambda_k p_k\right)^2}\sum_k \lambda_k \partial_x p_k\right)(s,\cdot)$ belong to $L^2(\mathbb{R})$, as $\forall i \in \{1,...,d\}$, $\frac{p_i}{\sum_k \lambda_k p_k} \in [0; \frac{1}{\lambda_{min}}]$, a.e. on $(0,T)\times\mathbb{R}$. It is sufficient to show that for $K \subset \mathbb{R}$ compact, $\phi \in C_c^\infty(\mathbb{R})$ with support included in K, and a.e. $s \in (0,T)$,

$$\int_K \partial_x\phi(x)\left(\frac{p_i p_j}{\sum_k \lambda_k p_k}\right)(s,x)dx = -\int_K \phi(x)\left(\frac{p_i\partial_x p_j + p_j \partial_x p_i}{\sum_k \lambda_k p_k} - \frac{p_i p_j}{\left(\sum_k \lambda_k p_k\right)^2}\sum_k \lambda_k \partial_x p_k\right)(s,x)dx.$$

Let $(\rho_n)_{n\geq 1}$ be a regularizing sequence, where $\rho_n \in C_c^\infty(\mathbb{R})$, with a support included in $(-\frac{1}{n},\frac{1}{n})$, $\int_\mathbb{R} \rho_n = 1$, and $\rho_n \geq 0$ for $n \geq 1$. We define $p_{n,i}(s,\cdot) := \rho_n * p_i(s,\cdot)$. Then for a.e. $s \in (0,T)$, $(p_{n,i}(s,\cdot))_{n\geq 0}, i \in \{1,...,d\}$ are sequences of non negative functions in $C^\infty(\mathbb{R})\cap H^1(\mathbb{R})$ such that for $1 \le i \le d$, the following strong convergences hold, by [19, Theorem 4.22]:

$$\begin{array}{rcl} p_{n,i}(s,\cdot) & \underset{n\to\infty}{\to} & p_i(s,\cdot) \text{ in } L^2(\mathbb{R}), \\ \partial_x p_{n,i}(s,\cdot) & \underset{n\to\infty}{\to} & \partial_x p_i(s,\cdot) \text{ in } L^2(\mathbb{R}). \end{array}$$

We first check that for $n \ge 1$, $x \in \mathbb{R}$ and a.e. $s \in (0,T)$, $\sum_{k=1}^d \lambda_k p_{n,k}(s,x) > 0$. Indeed, as $\sum_i p_i$ is a solution to the heat equation, $\sum_{i=1}^d p_i(s,x) = \mu * h_s(x)$ a.e. on $(0,T)\times\mathbb{R}$, we have that for $x \in \mathbb{R}$ and a.e. $s \in (0,T)$, $\sum_{k=1}^d \lambda_k p_{n,k}(s,x) \ge \lambda_{min} \sum_{k=1}^d p_{n,k}(s,x) = \lambda_{min}\rho_n * \mu * h_s(x) > 0$. Then for a.e. $s \in (0,T)$, we have the equality:

$$\int_K \partial_x\phi(x)\left(\frac{p_{n,i}p_{n,j}}{\sum_k \lambda_k p_{n,k}}\right)(s,x)dx = -\int_K \phi(x)\left(\frac{p_{n,i}\partial_x p_{n,j} + p_{n,j}\partial_x p_{n,i}}{\sum_k \lambda_k p_{n,k}} - \frac{p_{n,i}p_{n,j}}{\left(\sum_k \lambda_k p_{n,k}\right)^2}\sum_k \lambda_k \partial_x p_{n,k}\right)(s,x)dx. \tag{2.3.55}$$

For a.e. $s \in (0,T)$, modulo the extraction of a subsequence (which can depend on s), we can assume that for $1 \le l \le d$ and a.e. $x \in K$, $p_{n,l}(s,x) \underset{n\to\infty}{\to} p_l(s,x)$. As $\sum_{l=1}^d p_l(s,x) = \mu * h_s(x) > 0$ for a.e. $x \in \mathbb{R}$, we have $\frac{p_{n,l}}{\sum_k \lambda_k p_{n,k}}(s,x) \to \frac{p_l}{\sum_k \lambda_k p_k}(s,x)$ for a.e. $x \in K$. Then,

$$\begin{aligned} \left|\left|\left(\frac{p_{n,i}p_{n,j}}{\sum_k \lambda_k p_{n,k}} - \frac{p_i p_j}{\sum_k \lambda_k p_k}\right)(s,\cdot)\right|\right|_{L^2(K)} &\le \left|\left|\left(\frac{p_{n,i}p_{n,j}}{\sum_k \lambda_k p_{n,k}} - \frac{p_{n,i}p_j}{\sum_k \lambda_k p_{n,k}}\right)(s,\cdot)\right|\right|_{L^2(K)} \\ &+ \left|\left|\left(\frac{p_{n,i}p_j}{\sum_k \lambda_k p_{n,k}} - \frac{p_i p_j}{\sum_k \lambda_k p_k}\right)(s,\cdot)\right|\right|_{L^2(K)} \\ &\le \frac{1}{\lambda_{min}}||\left(p_{n,j}-p_j\right)(s,\cdot)||_{L^2(K)} \\ &+ \left|\left|\left(p_j\left(\frac{p_{n,i}}{\sum_k \lambda_k p_{n,k}} - \frac{p_i}{\sum_k \lambda_k p_k}\right)\right)(s,\cdot)\right|\right|_{L^2(K)}. \end{aligned}$$

The first term of the r.h.s converges to 0 as $n \to \infty$, as $p_{n,i}(s,\cdot) \underset{n\to\infty}{\to} p_i(s,\cdot)$ in $L^2(\mathbb{R})$, and the second term also converges to 0 by dominated convergence as $\frac{p_{n,i}}{\sum_k \lambda_k p_{n,k}}(s,\cdot) \to \frac{p_i}{\sum_k \lambda_k p_k}(s,\cdot)$ a.e. on K, and $\forall i \in \{1,...,d\}, \forall n \ge 1$, $\frac{p_{n,i}}{\sum_k \lambda_k p_{n,k}}(s,x) \in [0; \frac{1}{\lambda_{min}}]$. This ensures that

$$\int_K \partial_x\phi(x)\left(\frac{p_{n,i}p_{n,j}}{\sum_k \lambda_k p_{n,k}}\right)(s,x)dx \underset{n\to\infty}{\to} \int_K \partial_x\phi(x)\left(\frac{p_i p_j}{\sum_k \lambda_k p_k}\right)(s,x)dx,$$

for a.e. $s \in (0,T)$. With similar arguments, we let $n \to \infty$ in the r.h.s. of (2.3.55), we have the convergence of the r.h.s. term to

$$-\int_K \phi(x)\left(\frac{p_i\partial_x p_j + p_j\partial_x p_i}{\sum_k \lambda_k p_k} - \frac{p_i p_j}{\left(\sum_k \lambda_k p_k\right)^2}\sum_k \lambda_k \partial_x p_k\right)(s,x)dx,$$

for a.e. $s \in (0,T)$ and this concludes the proof. □

We can now prove Theorem 2.2.2. By Lemma 2.3.15 the solution of the variational formulation $V(\mu)$ satisfies :

$$\forall i \in \{1,...,d\}, \forall \phi \in H^1(\mathbb{R}), \frac{d}{dt}\int_x \phi p_i dx + \frac{1}{2}\int_x \left(\partial_x\phi\right)\partial_x\left(\frac{\sum_k p_k}{\sum_k \lambda_k p_k}\lambda_i p_i\right)dx = 0,$$

in the sense of distributions. Then through an integration by parts,

$$\forall i \in \{1,...,d\}, \forall \phi \in C_c^\infty(\mathbb{R}), \frac{d}{dt}\int_x \phi p_i dx - \frac{1}{2}\int_x \left(\partial_{xx}^2\phi\right)\left(\frac{\sum_k p_k}{\sum_k \lambda_k p_k}\lambda_i p_i\right)dx = 0.$$

For $y_i \in \mathcal{Y}$, [39, Theorem 2.6] gives the existence of a probability measure $\mathbb{P}^{y_i}$ on the space $C([0,T],\mathbb{R})$ with canonical process $(X_t)_{0\le t\le T}$ satisfying :

$$\begin{aligned}
& X_0 \quad \sim \quad \mu \text{ under } \mathbb{P}^{y_i}, \\
& \forall \phi \in C_c^\infty(\mathbb{R}), \qquad M_t^{\phi,y_i} := \phi(X_t) - \phi(X_0) - \frac{1}{2}\int_0^t \partial_{xx}^2 \phi(X_s) f^2(y_i) \frac{\sum_k p_k(s,X_s)}{\sum_k f^2(y_k)p_k(s,X_s)} ds \text{ is a martingale under } \mathbb{P}^{y_i},
\end{aligned}$$

and for $t > 0$, X_t has the density $\frac{p_i(t,\cdot)}{\alpha_i}$ under $\mathbb{P}^{y_i}$. We then form the measure $\mathbb{Q} = \sum_{i=1}^d \alpha_i \mathbb{P}^{y_i}(dX) \otimes \delta_{y_i}(dY)$ and show that it solves the following martingale problem:

$$\begin{aligned}
& X_0 \quad \sim \quad \mu \text{ under } \mathbb{Q}, \\
& \forall \phi \in C_c^\infty(\mathbb{R}), \qquad M_t^{\phi,Y} := \phi(X_t) - \phi(X_0) - \frac{1}{2}\int_0^t \partial_{xx}^2 \phi(X_s) f^2(Y) \frac{\sum_{i=1}^d p_i(s,X_s)}{\sum_{i=1}^d f^2(y_i)p_i(s,X_s)} ds \text{ is a martingale under } \mathbb{Q}.
\end{aligned}$$

For $\phi \in C_c^\infty(\mathbb{R})$ and $t \geq 0$, $M_t^{\phi,Y}$ is bounded as $\left|\frac{f^2(Y)\sum_{i=1}^d p_i(s,X_s)}{\sum_{i=1}^d f^2(y_i)p_i(s,X_s)}\right| \leq \frac{\lambda_{max}}{\lambda_{min}}$. For $s \geq 0$, we define $\mathcal{F}_s = \sigma(\{X_u, u \leq s\})$. To obtain our result it is enough to check that for $0 \leq s \leq t$,

$$\mathbb{E}^{\mathbb{Q}}\left[M_t^{\phi,Y} - M_s^{\phi,Y} | \mathcal{F}_s, Y\right] = 0.$$

For $\phi \in C_c^\infty(\mathbb{R})$ and g measurable and bounded on $\mathbb{R}^p \times \mathcal{Y}$, $p \geq 1$, and $0 \leq s_1 \leq ... \leq s_p \leq s$,

$$\mathbb{E}^{\mathbb{Q}}\left[(M_t^{\phi,Y} - M_s^{\phi,Y}) g(X_{s_1}, ..., X_{s_p}, Y)\right] = \sum_{i=1}^d \alpha_i \mathbb{E}^{\mathbb{P}^{y_i}}\left[(M_t^{\phi,y_i} - M_s^{\phi,y_i}) g(X_{s_1}, ..., X_{s_p}, y_i)\right] = 0,$$

as M_t^{ϕ,y_i} is a $\mathbb{P}^{y_i}$-martingale for $1 \leq i \leq d$. So $\mathbb{E}^{\mathbb{Q}}\left[M_t^{\phi,Y} | \mathcal{F}_s, Y\right] = M_s^{\phi,Y}$, and $M_t^{\phi,Y}$ is a $\mathbb{Q}$-martingale. For $0 \leq s \leq T$, we compute the conditional expectation $\mathbb{E}^{\mathbb{Q}}\left[f^2(Y)|X_s\right]$. Given a bounded measurable function g on $\mathbb{R}$,

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[f^2(Y)g(X_s)\right] &= \sum_{i=1}^d \alpha_i \mathbb{E}^{\mathbb{P}^{y_i}}\left[f^2(y_i)g(X_s)\right] = \sum_{i=1}^d \alpha_i \int_{\mathbb{R}} f^2(y_i) g(x) \frac{p_i(s,x)}{\alpha_i} dx \\
&= \int_{\mathbb{R}} \sum_{i=1}^d f^2(y_i) g(x) p_i(s,x) dx = \int_{\mathbb{R}} \left(\frac{\sum_{i=1}^d f^2(y_i)p_i(s,x)}{\sum_{i=1}^d p_i(s,x)}\right) g(x) \sum_{i=1}^d p_i(s,x) dx, \\
\mathbb{E}^{\mathbb{Q}}\left[g(X_s)\right] &= \sum_{i=1}^d \alpha_i \mathbb{E}^{\mathbb{P}^{y_i}}\left[g(X_s)\right] = \int_{\mathbb{R}} g(x) \sum_{i=1}^d p_i(s,x) dx.
\end{aligned}$$

Thus under $\mathbb{Q}$, X_s has the density $\sum_{i=1}^d p_i(s,\cdot)$, which is equal to $\mu * h_s(\cdot)$ by Theorem 2.2.1, so X_s has the same law as $Z + W_s$, where $Z \sim \mu$ and Z is independent from $(W_t)_{t \geq 0}$. Moreover, we have the equality:

$$\mathbb{E}^{\mathbb{Q}}\left[f^2(Y)|X_s\right] = \frac{\sum_{i=1}^d f^2(y_i)p_i(s,X_s)}{\sum_{i=1}^d p_i(s,X_s)} \text{ a.s..}$$

Therefore $\mathbb{Q}$ is a solution to the martingale problem:

$$\begin{aligned}
& X_0 \quad \sim \quad \mu \text{ under } \mathbb{Q}, \\
& \forall \phi \in C_c^\infty(\mathbb{R}), \qquad \phi(X_t) - \phi(X_0) - \frac{1}{2}\int_0^t \partial_{xx}^2 \phi(X_s) \frac{f^2(Y)}{\mathbb{E}^{\mathbb{Q}}\left[f^2(Y)|X_s\right]} ds \text{ is a martingale under } \mathbb{Q}.
\end{aligned}$$

and this ensures the existence of a weak solution to the SDE (2.1.1).

2.4 Calibrated RSLV models

We extend the results obtained in the previous sections to the case when the asset price S follows the dynamics:

$$dS_t = rS_t dt + \frac{f(Y_t)}{\sqrt{\mathbb{E}\left[f^2(Y_t)|S_t\right]}} \sigma_{Dup}(t,S_t) S_t dW_t,$$

$$(\log (S_0), Y_0) \quad \sim \quad \mu,$$

where Y_0 is a random variables with values in $\mathcal{Y}$, μ is a probability measure on $\mathbb{R}\times\mathcal{Y}$, Y is a process evolving in $\mathcal{Y}$, with

$$\mathbb{P}\left(Y_{t+dt}=j|\sigma\left(\left(S_s,Y_s\right),0\le s\le t\right)\right)=q_{Y_tj}(\log(S_t))dt$$

for $j\neq Y_t$, and for $1\le i\neq j\le d$, the function $q_{ij}:\mathbb{R}\to\mathbb{R}$ is non negative. Moreover, for $1\le i\le d$, we define $q_{ii}:=-\sum_{j\neq i}q_{ij}$. We assume that $(\log(S_0),Y_0)$ and $(W_t)_{t\ge0}$ are independent. In addition, we assume that there exists $\overline{q}>0$ such that $||q_{ij}||_\infty\le\overline{q}$ for $1\le i,j\le d$, and that the risk free rate r is constant. We define $\tilde{\sigma}_{Dup}(t,x):=\sigma_{Dup}(t,e^x)$. The asset log-price $X_t=\log S_t$ follows the dynamics:

$$dX_t \quad = \quad \left(r-\frac{1}{2}\frac{f^2(Y_t)}{\mathbb{E}\left[f^2(Y_t)|X_t\right]}\tilde{\sigma}^2_{Dup}(t,X_t)\right)dt+\frac{f(Y_t)}{\sqrt{\mathbb{E}\left[f^2(Y_t)|X_t\right]}}\tilde{\sigma}_{Dup}(t,X_t)dW_t, \tag{2.4.1}$$

$$(X_0,Y_0) \quad \sim \quad \mu. \tag{2.4.2}$$

Formally, if we apply Gyongy's theorem, any solution to the SDE (2.4.1) should have the same time marginals as the solution to the Dupire SDE for the asset's log-price:

$$dX^D_t \quad = \quad \left(r-\frac{1}{2}\tilde{\sigma}^2_{Dup}(t,X^D_t)\right)dt+\tilde{\sigma}_{Dup}(t,X^D_t)dW_t, \tag{2.4.3}$$

$$X^D_0 \quad \sim \quad \mu_{X_0} \tag{2.4.4}$$

where μ_{X_0} is the law of X_0. The Fokker-Planck PDS associated with the SDE (2.4.1)-(2.4.2) writes, for $1\le i\le d$:

$$\partial_t p_i \quad = \quad -\partial_x\left(\left[r-\frac{1}{2}\frac{\lambda_i\sum_{i=1}^d p_k}{\sum_{i=1}^d\lambda_k p_k}\tilde{\sigma}^2_{Dup}\right]p_i\right)+\frac{1}{2}\partial^2_{xx}\left(\frac{\lambda_i\sum_{i=1}^d p_k}{\sum_{i=1}^d\lambda_k p_k}\tilde{\sigma}^2_{Dup}p_i\right)+\sum_{j=1}^d q_{ji}p_j \tag{2.4.5}$$

$$p_i(0,\cdot) \quad = \quad \alpha_i\mu_i, \tag{2.4.6}$$

where for $1\le i\le d$, μ_i is the conditional law of X_0 given $\{Y_0=i\}$, and as before, $\alpha_i=\mathbb{P}\left(Y_0=i\right)\ge0$. We denote by Λ the diagonal $d\times d$ matrix with coefficients $(\lambda_i)_{1\le i\le d}$. We define, for $\rho\in\mathcal{D}$, $R(\rho):=\frac{\sum_{i=1}^d\rho_i}{\sum_{i=1}^d\lambda_i\rho_i}$, and for $x\in\mathbb{R}$, $Q(x)=(q_{ij}(x))_{1\le i,j\le d}$. Moreover, we assume that the European call prices given by the market have sufficient regularity so that the following assumption holds.

Assumption (B). *The function $\tilde{\sigma}_{Dup}$ belongs to the space $L^\infty([0,T],W^{1,\infty}(\mathbb{R}))$, and there exists a constant $\underline{\sigma}>0$ such that a.e. on $[0,T]\times\mathbb{R}$, $\underline{\sigma}\le\tilde{\sigma}_{Dup}$.*

We will denote by $\overline{\sigma}>0$ some constant such that a.e. on $[0,T]\times\mathbb{R}$, $\tilde{\sigma}_{Dup}\le\overline{\sigma}$. For the PDS (2.4.5)-(2.4.6), we introduce an associated variational formulation, called $V_{Fin}(\mu)$:

$$\text{Find}\qquad p=(p_1,...,p_d)\ \text{satisfying:}$$

$$p\in L^2_{loc}((0,T];H)\cap L^\infty_{loc}((0,T];L),$$

$$p\ \text{takes values in}\ \mathcal{D},\ \text{a.e. on}\ (0,T)\times\mathbb{R},$$

$$\forall v\in H,\ \frac{d}{dt}(v,p)_d-r\left(\partial_x v,p\right)_d+\left(\partial_x v,\frac{1}{2}R(p)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+2\partial_x\tilde{\sigma}_{Dup}\right)\Lambda p\right)_d+(\partial_x v,\tilde{\sigma}^2_{Dup}A(p)\partial_x p)_d=(Qv,p)_d$$

$$\text{in the sense of distributions on}\ (0,T),\ \text{and}$$

$$p(t,\cdot)\underset{t\to0^+}{\overset{\text{weakly-*}}{\rightarrow}}p_0:=(\alpha_1\mu_1,...,\alpha_d\mu_d).$$

Let us remark that if we sum the PDS (2.4.5) over the index i, as $\sum_{i=1}^d q_{ji}=0$ for $1\le j\le d$, then $\sum_{i=1}^d p_i$ satisfies the Fokker-Planck equation associated to the SDE (2.4.4):

$$\partial_t\sum_{i=1}^d p_i \quad = \quad -\partial_x\left(\left[r-\frac{1}{2}\tilde{\sigma}^2_{Dup}\right]\sum_{i=1}^d p_i\right)+\frac{1}{2}\partial^2_{xx}\left(\tilde{\sigma}^2_{Dup}\sum_{i=1}^d p_i\right) \tag{2.4.7}$$

$$\sum_{i=1}^{d} p_i(0,\cdot) \quad = \quad \mu_{X_0}. \tag{2.4.8}$$

In the same way, if p is a solution to $V_{Fin}(\mu)$, then $u := \sum_{i=1}^{d} p_i$ solves $LV(\mu_{X_0})$, where for $\nu \in \mathcal{P}(\mathbb{R})$, $LV(\nu)$ is defined by:

$$u \in L^2_{loc}((0,T]; H^1(\mathbb{R})) \cap L^\infty_{loc}((0,T]; L^2(\mathbb{R})),\ u \geq 0,$$

$$\forall v \in H^1(\mathbb{R}),\ \frac{d}{dt}(v,u)_1 - r\,(\partial_x v, u)_1 + \left(\partial_x v, \frac{1}{2}\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup} + 2\partial_x \tilde{\sigma}_{Dup}\right) u\right)_1 + \frac{1}{2}\left(\partial_x v, \tilde{\sigma}^2_{Dup}\partial_x u\right)_1 = 0,$$

in the sense of distributions on $(0,T)$, and

$$u(t,\cdot) \underset{t\to 0^+}{\overset{\text{weakly-}*}{\rightarrow}} \nu.$$

Lemma 2.4.1. *Under Assumption (B), for $\nu \in \mathcal{P}(\mathbb{R})$, the solutions of $LV(\nu)$ are continuous and positive on $(0,T] \times \mathbb{R}$.*

The proof of Lemma 2.4.1 is postponed to Appendix 2.C. We make here an additional assumption on the regularity of the function $\tilde{\sigma}_{Dup}$.

Assumption (H)**.** *The function $\tilde{\sigma}_{Dup}$ is continuous and there exist two constants $H_0 > 0$ and $\chi \in (0,1]$ such that*

$$\forall s,t \in [0,T], \forall x \in \mathbb{R},\ |\tilde{\sigma}_{Dup}(s,x) - \tilde{\sigma}_{Dup}(t,x)| \leq H_0|t-s|^\chi.$$

If $\tilde{\sigma}_{Dup} \in L^\infty([0,T], W^{1,\infty}(\mathbb{R}))$ and $\tilde{\sigma}_{Dup}$ is continuous, then $\tilde{\sigma}_{Dup}$ and $\tilde{\sigma}^2_{Dup}$ have the Lipschitz property in the space variable, uniformly in time so that existence and trajectorial uniqueness hold for (2.4.3)-(2.4.4). Moreover, Assumptions (H) and (B) imply that there exists $\tilde{H}_0 > 0$ such that:

$$\forall s,t \in [0,T], \forall x,y \in \mathbb{R},\ |\tilde{\sigma}_{Dup}(s,x) - \tilde{\sigma}_{Dup}(t,y)| \leq \tilde{H}_0\left(|t-s|^\chi + |x-y|\right).$$

Moreover, for $\nu \in \mathcal{P}(\mathbb{R})$, they are sufficient to obtain uniqueness to $LV(\nu)$ and Aronson-like upper-bounds on the solution of $LV(\nu)$. The proof of Proposition 2.4.2 that follows is also postponed to Appendix 2.C.

Proposition 2.4.2. *Under Assumptions (B) and (H), there exists a unique solution u to $LV(\mu_{X_0})$ and the time marginals of the solution $\left(X^D_t\right)_{t\in(0,T]}$ to the SDE (2.4.3)-(2.4.4) are given by $(u(t,\cdot))_{t\in(0,T]}$. Moreover, there exists a finite constant ζ, independent from μ_{X_0} and such that u satisfies $||u(t)||^2_{L^2} \leq \frac{\zeta}{\sqrt{t}}$ for a.e. $t \in (0,T]$.*

We give here the main results on the calibrated RSLV model, that we prove in Section 2.5.

Theorem 2.4.3. *Under Condition (C), Assumptions (B) and (H), there exists a solution $p \in C((0,T], L)$ to $V_{Fin}(\mu)$ such that $\sum_{i=1}^{d} p_i$ is the unique solution to $LV(\mu_{X_0})$.*

Theorem 2.4.4. *Under Condition (C), Assumptions (B) and (H), there exists a weak solution to the SDE (2.4.1)-(2.4.2). which has the same time marginals as those of the solution to the SDE (2.4.3)-(2.4.4).*

2.5 Proofs of Section 4

2.5.1 Proof of Theorem 2.4.3

Similarly to the proof of Theorem 2.2.1, in Subsection 5.1.1, under the hypothesis that for $1 \leq i \leq d$, μ_i has a square integrable density with respect to the Lebesgue measure, we prove existence to a variational formulation slightly stronger than $V_{Fin}(\mu)$. Then, in Subsection 5.1.2, when μ is a general probability measure on $\mathbb{R} \times \mathcal{Y}$, we mollify μ_i for $1 \leq i \leq d$, in order to use the results of Subsection 5.1.1 and obtain a solution to $V_{Fin}(\mu)$.

Case when μ has square integrable densities

In this section, the measures $(\mu_i)_{1\le i\le d}$ are assumed to have square integrable densities which are also denoted by μ_i for notational simplicity. We define $p_0 := (\alpha_1\mu_1, ..., \alpha_d\mu_d) \in L$. We define the variational formulation $V_{Fin,L^2}(\mu)$:

$$\text{Find} \qquad p = (p_1, ..., p_d) \text{ satisfying :} \tag{2.5.1}$$

$$p \in L^2([0,T];H) \cap L^\infty([0,T];L) \text{ and takes values in } \mathcal{D} \text{ a.e. on } (0,T)\times\mathbb{R}, \tag{2.5.2}$$

$$\forall v \in H,\ \frac{d}{dt}(v,p)_d - r\,(\partial_x v, p)_d + \left(\partial_x v, \frac{1}{2}R(p)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup} + 2\partial_x\tilde\sigma_{Dup}\right)\Lambda p\right)_d + (\partial_x v, \tilde\sigma^2_{Dup}A(p)\partial_x p)_d = (Qv,p)_d \tag{2.5.3}$$

in the sense of distributions on $(0,T)$,

$$p(0,\cdot) = p_0. \tag{2.5.4}$$

To show existence to $V_{Fin,L^2}(\mu)$, we use Galerkin's procedure, as in the proof of Theorem 2.2.1. It is not obvious that p takes values in $\mathcal{D}$ a.e. on $(0,T)\times\mathbb{R}$ at the discrete level, that is why for $\epsilon > 0$, we define, for $\rho \in (\mathbb{R}_+)^d$, $R_\epsilon(\rho) = \frac{\sum_{i=1}^d \rho_i}{\epsilon \vee \left(\sum_{i=1}^d \lambda_i\rho_i\right)}$ and we introduce the variational formulation $V_{Fin,\epsilon}(\mu)$:

$$\text{Find} \qquad p_\epsilon = (p_{\epsilon,1}, ..., p_{\epsilon,d}) \text{ satisfying :} \tag{2.5.5}$$

$$p_\epsilon \in L^2([0,T];H) \cap L^\infty([0,T];L) \tag{2.5.6}$$

$$\forall v \in H,\ \frac{d}{dt}(v,p_\epsilon)_d - r\,(\partial_x v, p_\epsilon)_d + \left(\partial_x v, \frac{1}{2}R_\epsilon(p_\epsilon^+)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup} + \partial_x\tilde\sigma_{Dup}\right)\Lambda p_\epsilon\right)_d + \left(\partial_x v, \tilde\sigma^2_{Dup}A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon\right)_d = (Qv,p_\epsilon)_d \tag{2.5.7}$$

in the sense of distributions on $(0,T)$,

$$p_\epsilon(0,\cdot) = p_0. \tag{2.5.8}$$

Let us remark that if p (resp. p_ϵ) satisfies (2.5.1)-(2.5.3) (resp. (2.5.5)-(2.5.7)), those conditions imply that $\frac{dp}{dt}$ (resp. $\frac{dp_\epsilon}{dt}$) belongs to $L^2([0,T],H')$ in the sense of [97, III, Lemma 1.1], so by [97, III, Lemma 1.2], p (resp. p_ϵ) is equal a.e. on $[0,T]$ to a function of $C([0,T],L)$, so that the initial condition (2.5.4) (resp. 2.5.8) makes sense.

To take advantage of the fact that under Condition (C), there exists $\Pi \in S_d^{++}(\mathbb{R})$ and $\kappa > 0$ such that ΠA and ΠA_ϵ, for $\epsilon > 0$, are uniformly coercive on $\mathcal{D}$ with the coefficient κ, we introduce, for $\epsilon > 0$ and $m \ge 1$, the approximate variational formulation $V^m_{Fin,\epsilon}(\mu)$:

$$\text{Find } g^m_{\epsilon,1}, ..., g^m_{\epsilon,m} \in C^0([0,T],\mathbb{R}),\ \text{such that:}$$

$$\text{the function } t \in [0,T] \to p^m_\epsilon(t) = \sum_{j=1}^m g^m_{\epsilon,j}(t)w_j \text{ satisfies, for } 1 \le i \le m,$$

$$\begin{aligned}(Q\Pi w_i + r\Pi\partial_x w_i, p^m_\epsilon(t))_d &= \frac{d}{dt}\left(w_i, \Pi p^m_\epsilon(t)\right)_d + \left(\partial_x w_i, \tilde\sigma^2_{Dup}\Pi A_\epsilon\left(\left(p^m_\epsilon\right)^+(t)\right)\partial_x p^m_\epsilon(t)\right)_d \\ &+ \frac{1}{2}\left(\partial_x w_i, R_\epsilon\left(\left(p^m_\epsilon\right)^+(t)\right)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup} + 2\partial_x\tilde\sigma_{Dup}\right)\Pi\Lambda p^m_\epsilon(t)\right)_d \end{aligned} \tag{2.5.9}$$

$$p^m_\epsilon(0) = p^m_0,$$

where p^m_0 is the orthogonal projection of p_0 in L on the space spanned by $(w_j)_{1\le j\le m}$. For $z \in \mathbb{R}^m$ and $t \ge 0$, we define $K^m_{\epsilon,1}(t,z)$ the matrix, where for $1 \le i,j \le m$:

$$K^m_{\epsilon,1,ij}(t,z) = \frac{1}{2}\left(\partial_x w_i, R_\epsilon\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right)\tilde\sigma_{Dup}(t)\left(\tilde\sigma_{Dup}(t) + 2\partial_x\tilde\sigma_{Dup}(t)\right)\Pi\Lambda w_j\right)_d,$$

$K^m_{\epsilon,2}(t,z)$ the matrix where for $1\leq i,j\leq m$,

$$K^m_{\epsilon,2,ij}(t,z) = \left(\partial_x w_i, \tilde{\sigma}^2_{Dup}(t)\Pi A_\epsilon\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right)\partial_x w_j\right)_d,$$

and $K^m_{\epsilon,3}$ the constant matrix where for $1\leq i,j\leq m$:

$$K^m_{\epsilon,3,ij} = (Q\Pi w_i + r\Pi\partial_x w_i, w_j)_d.$$

We then define $F^m_{\epsilon,k}(t,z) := \left(W^{(m)}\right)^{-1} K^m_{\epsilon,k}(t,z)z$, for $k=1,2$, and $F^m_{\epsilon,3}(z) := \left(W^{(m)}\right)^{-1} K^m_{\epsilon,3}z$, Finally we define for $z\in\mathbb{R}^m$, $G^m_\epsilon(t,z) := -F^m_{\epsilon,1}(t,z) - F^m_{\epsilon,2}(t,z) + F^m_{\epsilon,3}(z)$. The ODE for g^m_ϵ rewrites:

$$\begin{array}{rcl}(g^m_\epsilon)'(t) & = & G^m_\epsilon(t, g^m_\epsilon(t))\\ g^m_\epsilon(0) & = & g^m_{\epsilon,0},\end{array}$$

where the vector $g^m_{\epsilon,0}$ is the expression of p^m_0 on the basis $(w_i)_{1\leq i\leq m}$. We prove existence and uniqueness to $V^m_{Fin,\epsilon}(\mu)$. We clearly have the following lemma.

Lemma 2.5.1. *The functions R and R_ϵ, for $\epsilon>0$, are uniformly bounded. More precisely, $||R_\epsilon||_{L^\infty((\mathbb{R}_+)^d)} \leq \frac{1}{\lambda_{min}}$, $||R||_{L^\infty(\mathcal{D})} \leq \frac{1}{\lambda_{min}}$.*

Lemma 2.5.2. *Under Assumption (B), for $m\geq 1$, the functions $K^m_{\epsilon,1}$ and $K^m_{\epsilon,2}$ are uniformly bounded.*

Proof. Using Assumption (B), Lemma 2.5.1 and Lemma 2.3.7, for $t\in[0,T]$, $x\in\mathbb{R}$, $\rho\in\mathcal{D}$, we have

$$\begin{array}{rcl}\left|\frac{1}{2}R_\epsilon(\rho)\tilde{\sigma}_{Dup}(t,x)\left(\tilde{\sigma}_{Dup}(t,x)+2\partial_x\tilde{\sigma}_{Dup}(t,x)\right)\right| & \leq & \frac{1}{2\lambda_{min}}\overline{\sigma}(\overline{\sigma}+2||\partial_x\tilde{\sigma}_{Dup}||_\infty),\\ ||A_\epsilon(\rho)\tilde{\sigma}^2_{Dup}(t,x)||_\infty & \leq & \frac{1}{2}\left(1+\frac{\lambda_{max}}{\lambda_{min}}\right)\overline{\sigma}^2.\end{array}$$

This is sufficient to show that $K^m_{\epsilon,1}$ and $K^m_{\epsilon,2}$ are uniformly bounded, as the functions $(w_{ik}\partial_x w_{jl})_{1\leq i,j\leq m, 1\leq k,l\leq d}$ and $(\partial_x w_{ik}\partial_x w_{jl})_{1\leq i,j\leq m,1\leq k,l\leq d}$ belong to $L^1(\mathbb{R})$. ☐

Lemma 2.5.3. *Under Assumption (B), for $m\geq 1$, the functions $F^m_{\epsilon,1}$ and $F^m_{\epsilon,2}$ are locally Lipschitz in z, uniformly in $t\in[0,T]$.*

The proof of Lemma 2.5.3 is similar to the proof of Lemma 2.3.6 and is postponed to Appendix A.

Lemma 2.5.4. *Under Assumption (B), $V^m_{Fin,\epsilon}(\mu)$ has a unique solution.*

Proof. In addition to Lemma 2.5.3, $F_{\epsilon,3}$ is clearly a Lipschitz function. Therefore the function G_ϵ is locally Lipschitz in z uniformly in t. Caratheodory's theorem (see e.g. [59, Theorems 5.2, 5.3]) gives the existence of a unique maximal absolutely continuous solution g^m_ϵ on an interval $[0,T^*)$, with $T^*>0$. In addition, as the function $(t,z)\to -W^{-1}K_{\epsilon,1}(t,z) - W^{-1}K_{\epsilon,2}(t,z) + W^{-1}K_{\epsilon,3}$ is uniformly bounded on $\mathbb{R}_+\times\mathbb{R}^m$, we conclude as in the proof of Lemma 2.3.8, using Gronwall's lemma, that $T^*=\infty$. Consequently, g^m_ϵ is defined on $[0,T]$, and there exists a unique solution to $V^m_{Fin,\epsilon}(\mu)$. ☐

We now compute energy estimates on the solution p^m_ϵ to $V^m_{Fin,\epsilon}(\mu)$, for $m\geq 1$ and $\epsilon>0$. Taking p^m_ϵ as a test function in (2.5.9), we have

$$\begin{array}{rcl}\frac{1}{2}\frac{d}{dt}|\sqrt{\Pi}p^m_\epsilon|^2_d - (Q\Pi p^m_\epsilon, p^m_\epsilon)_d & = & -\left(\partial_x p^m_\epsilon, \tilde{\sigma}^2_{Dup}\Pi A_\epsilon\left((p^m_\epsilon)^+\right)\partial_x p^m_\epsilon\right)_d\\ & - & \left(\partial_x p^m_\epsilon, \left(\frac{1}{2}R_\epsilon((p^m_\epsilon)^+(t))\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+\partial_x\tilde{\sigma}_{Dup}\right)\Pi\Lambda - r\Pi\right)p^m_\epsilon(t)\right)_d.\end{array}$$

For $\eta>0$, by Young's inequality we have that

$$\left|\left(\partial_x p^m_\epsilon, \left(\frac{1}{2}R_\epsilon((p^m_\epsilon)^+)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+2\partial_x\tilde{\sigma}_{Dup}\right)\Pi\Lambda - r\Pi\right)p^m_\epsilon\right)_d\right| \leq C\left(\eta|\partial_x p^m_\epsilon|^2_d + \frac{1}{4\eta}|p^m_\epsilon|^2_d\right), \tag{2.5.10}$$

where $C := d||\Pi||_\infty \left(\frac{\lambda_{max}}{2\lambda_{min}}\overline{\sigma}(\overline{\sigma}+2||\partial_x\tilde{\sigma}||_\infty)+r\right)$. As for $\epsilon>0$, ΠA_ϵ is uniformly coercive with coefficient κ, and $\tilde{\sigma}_{Dup}$ is bounded from below by $\underline{\sigma}>0$, so

$$\left(\partial_x p_\epsilon^m, \tilde{\sigma}_{Dup}^2 \Pi A_\epsilon\left((p_\epsilon^m)^+\right)\partial_x p_\epsilon^m\right)_d \quad \geq \quad \kappa\underline{\sigma}^2|\partial_x p_\epsilon^m|_d^2.$$

We then choose $\eta = \frac{\kappa}{2C}\underline{\sigma}^2$, so that $C\eta = \frac{\kappa}{2}\underline{\sigma}^2$. Moreover, for $b := d^2\overline{q}||\Pi||_\infty$ we have that

$$\forall \xi\in\mathbb{R}^d, |\xi^* Q\Pi\xi| \leq b\xi^*\xi.$$

Gathering the previous inequalities we have that

$$\frac{1}{2}\frac{d}{dt}|\sqrt{\Pi}p_\epsilon^m|_d^2 - (Q\Pi p_\epsilon^m, p_\epsilon^m)_d \quad \leq \quad -\kappa\underline{\sigma}^2|\partial_x p_\epsilon^m|_d^2 + \frac{\kappa}{2}\underline{\sigma}^2|\partial_x p_\epsilon^m|_d^2 + \frac{C^2}{2\kappa\underline{\sigma}^2}|p_\epsilon^m|_d^2. \tag{2.5.11}$$

Consequently, we have

$$\frac{1}{2}\frac{d}{dt}|\sqrt{\Pi}p_\epsilon^m|_d^2 - \left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)|p_\epsilon^m|_d^2 \leq -\frac{\kappa}{2}\underline{\sigma}^2|\partial_x p_\epsilon^m|_d^2 \leq 0.$$

As $|p_\epsilon^m|_d^2 \leq \frac{1}{l_{min}(\Pi)}|\sqrt{\Pi}p_\epsilon^m|_d^2$, we also have

$$\frac{1}{2}\frac{d}{dt}|\sqrt{\Pi}p_\epsilon^m|_d^2 - \frac{1}{l_{min}(\Pi)}\left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)|\sqrt{\Pi}p_\epsilon^m|_d^2 \leq -\frac{\kappa}{2}\underline{\sigma}^2|\partial_x p_\epsilon^m|_d^2 \leq 0. \tag{2.5.12}$$

Integrating the inequality (2.5.12), and using the fact that

$$l_{min}(\Pi)|p_\epsilon^m|_d^2 \leq |\sqrt{\Pi}p_\epsilon^m|_d^2 \leq l_{max}(\Pi)|p_\epsilon^m|_d^2,$$

we obtain the following lemma.

Lemma 2.5.5. *The following energy estimates hold.*

$$\sup_{t\in[0,T]}|p_\epsilon^m(t)|_d^2 \quad \leq \quad \frac{l_{max}(\Pi)}{l_{min}(\Pi)}e^{\frac{2}{l_{min}(\Pi)}\left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)T}|p_0|_d^2, \tag{2.5.13}$$

$$\int_0^T|\partial_x p_\epsilon^m(t)|_d^2 dt \quad \leq \quad \frac{l_{max}(\Pi)}{\kappa\underline{\sigma}^2}e^{\frac{2}{l_{min}(\Pi)}\left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)T}|p_0|_d^2, \tag{2.5.14}$$

$$\int_0^T||p_\epsilon^m||_d^2 dt \quad \leq \quad \left(T\frac{l_{max}(\Pi)}{l_{min}(\Pi)}+\frac{l_{max}(\Pi)}{\kappa\underline{\sigma}^2}\right)e^{\frac{2}{l_{min}(\Pi)}\left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)T}|p_0|_d^2. \tag{2.5.15}$$

Now we can prove existence to $V_{Fin,\epsilon}(\mu)$.

Proposition 2.5.6. *There exists a solution p_ϵ to $(V_{Fin,\epsilon}(\mu))$.*

Proof. Given $\epsilon>0$, the family $(p_\epsilon^m)_{m\geq 0}$ has standard energy estimates with bounds independent from m. We also check that (2.5.9) rewrites for $1\leq j\leq m$

$$\frac{d}{dt}(w_j,\Pi p_\epsilon^m)_d + \langle w_j, Gp_\epsilon^m\rangle = 0$$

where for $q\in H$, $G_\epsilon q\in H'$, and

$$\begin{aligned}\forall v\in H, \langle v, G_\epsilon q\rangle \quad &= \quad -r(\partial_x v,\Pi q)_d + \left(\partial_x v, \left(\frac{1}{2}R_\epsilon\left(q^+\right)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+2\partial_x\tilde{\sigma}_{Dup}\right)\right)\Pi\Lambda p\right)_d \\ &+ \quad \left(\partial_x v, \tilde{\sigma}_{Dup}^2\Pi A_\epsilon\left(q^+\right)\partial_x q\right)_d - (Q\Pi v, q)_d.\end{aligned}$$

We see that for almost every $t\in[0,T]$,

$$||G_\epsilon p(t)||_{H'} \leq d||\Pi||_\infty\left(r+\frac{\lambda_{max}}{2\lambda_{min}}\overline{\sigma}(\overline{\sigma}+2||\partial_x\tilde{\sigma}_{Dup}||_\infty)+\frac{d}{2}\left(1+\frac{\lambda_{max}}{\lambda_{min}}\right)\overline{\sigma}^2+d\overline{q}\right)||p(t)||_H. \tag{2.5.16}$$

By Lemma 2.5.5 the family $(p_\epsilon^m)_{m\geq 1}$ is bounded in $L^2([0,T],H)$ and through the equality (2.5.16) the family $(G_\epsilon p_\epsilon^m)_{m\geq 1}$ is bounded in $L^2([0,T],H')$. What follows is a repetition of arguments used in the proof of Proposition 2.3.10. We extract a subsequence of $(p_\epsilon^m)_{m\geq 1}$, also denoted by $(p_\epsilon^m)_{m\geq 1}$, and a function $p_\epsilon \geq 0$ such that as $m \to \infty$,

$$\begin{array}{lll} p_\epsilon^m & \to & p_\epsilon \text{ in } L^2([0,T];H) \text{ weakly},\\ p_\epsilon^m & \to & p_\epsilon \text{ in } L^\infty([0,T];L) \text{ weakly-*},\\ p_\epsilon^m & \to & p_\epsilon \text{ a.e. on } (0,T]\times\mathbb{R}. \end{array}$$

For $j\geq 1$, $\psi \in C^1(\mathbb{R})$, with $\psi(T)=0$, and $m\geq j$, we have:

$$\begin{aligned} \int_0^T \psi(t)(Q\Pi w_j + r\Pi\partial_x w_j, p_\epsilon^m(t))_d dt &= -\int_0^T \psi'(t)(w_j,\Pi p_\epsilon^m(t))_d dt - \psi(0)(w_j,\Pi p_\epsilon^m(0))_d \\ &+ \int_0^T \psi(t)\left(\partial_x w_j, \tilde\sigma_{Dup}^2(t)\Pi A_\epsilon\left((p_\epsilon^m)^+(t)\right)\partial_x p_\epsilon^m(t)\right)_d dt \\ &+ \frac{1}{2}\int_0^T \psi(t)\left(\partial_x w_j, R_\epsilon\left((p_\epsilon^m)^+(t)\right)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup}+\partial_x\tilde\sigma_{Dup}\right)\Pi\Lambda p_\epsilon^m(t)\right)_d dt. \end{aligned}$$

The following convergences hold, using the same arguments as in the proof of Proposition 2.3.10.

$$\begin{aligned} \psi(0)(w_j,\Pi p_\epsilon^m(0))_d &\underset{m\to\infty}{\to} \psi(0)(w_j,\Pi p_0)_d, \\ -\int_0^T \psi'(t)(w_j,\Pi p_\epsilon^m(t))_d dt &\underset{m\to\infty}{\to} -\int_0^T \psi'(t)(w_j,\Pi p_\epsilon(t))_d dt, \\ \int_0^T \psi(t)(\partial_x w_j, \tilde\sigma_{Dup}^2(t)\Pi A_\epsilon\left((p_\epsilon^m)^+(t)\right)\partial_x p_\epsilon^m(t))_d dt &\underset{m\to\infty}{\to} \int_0^T \psi(t)(\partial_x w_j, \tilde\sigma_{Dup}^2(t)\Pi A_\epsilon\left((p_\epsilon)^+(t)\right)\partial_x p_\epsilon(t))_d dt, \end{aligned}$$

We only explain how to deal with the additional terms. As $p_\epsilon^m \to p_\epsilon$ weakly in $L^2([0,T];L)$, we have that:

$$\int_0^T \psi(t)(Q\Pi w_j + r\Pi\partial_x w_j, p_\epsilon^m(t))_d dt \underset{m\to\infty}{\to} \int_0^T \psi(t)(Q\Pi w_j + r\Pi\partial_x w_j, p_\epsilon(t))_d dt,$$

As the function R_ϵ is continuous and bounded on $(\mathbb{R}_+)^d$, and as $p_\epsilon^m \underset{m\to\infty}{\to} p_\epsilon$ a.e. on $(0,T]\times\mathbb{R}$, we have that $R_\epsilon\left((p_\epsilon^m)^+\right) \underset{m\to\infty}{\to} R_\epsilon\left((p_\epsilon)^+\right)$ a.e. on $[0,T]\times\mathbb{R}$. As the bounds (2.5.13)-(2.5.15) are independent from m, we show, with the same arguments as for (2.3.31), that:

$$\begin{aligned} &\int_0^T \psi(t)\left(\partial_x w_j, R_\epsilon\left((p_\epsilon^m)^+(t)\right)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup}+2\partial_x\tilde\sigma_{Dup}\right)\Pi\Lambda p_\epsilon^m(t)\right)_d dt \\ &\underset{m\to\infty}{\to} \int_0^T \psi(t)\left(\partial_x w_j, R_\epsilon\left((p_\epsilon)^+(t)\right)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup}+2\partial_x\tilde\sigma_{Dup}\right)\Pi\Lambda p_\epsilon(t)\right)_d dt. \end{aligned}$$

As the family $(w_j)_{j\geq 1}$ is total in H, and the function $v\to \Pi v$ is a bijection from H to H, we conclude that:

$$\begin{aligned} \forall v\in H, \int_0^T \psi(t)(Qv + r\partial_x v, p_\epsilon(t))_d dt &= -\int_0^T \psi'(t)(v,p_\epsilon(t))_d dt - \psi(0)(v,p_0) \\ &+ \int_0^T \psi(t)(\partial_x v, A_\epsilon\left((p_\epsilon)^+(t)\right)\tilde\sigma_{Dup}^2(t)\partial_x p_\epsilon(t))_d dt \\ &+ \frac{1}{2}\int_0^T \psi(t)\left(\partial_x v, R_\epsilon\left((p_\epsilon)^+(t)\right)\tilde\sigma_{Dup}\left(\tilde\sigma_{Dup}+2\partial_x\tilde\sigma_{Dup}\right)\Lambda p_\epsilon(t)\right)_d dt. \end{aligned}$$

Thus we have checked that p_ϵ satisfies (2.5.7). With the same arguments as the end of the proof of Proposition 2.3.10, we can show that for $v\in H$, the function $t\to(v,p_\epsilon(t))_d$ belongs to $H^1(0,T)$, and that p_ϵ satisfies the initial condition (2.5.8), so that p_ϵ is a solution to $V_{Fin,\epsilon}(\mu)$. ☐

We now show that the solutions of $V_{Fin,\epsilon}(\mu)$ are non negative.

Proposition 2.5.7. *The solutions of $V_{Fin,\epsilon}(\mu)$ are non negative.*

Proof. Let us take p_ϵ^- as a test function in (2.5.7). We obtain

$$\frac{1}{2}\frac{d}{dt}|p_\epsilon^-|_d^2 + \left(\partial_x p_\epsilon^-, \left(\frac{1}{2}R_\epsilon(p_\epsilon^+)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup} + 2\partial_x\tilde{\sigma}_{Dup}\right)\Lambda - rI_d\right)p_\epsilon\right)_d + \left(\partial_x p_\epsilon^-, \tilde{\sigma}_{Dup}^2 A_\epsilon(p_\epsilon^+)\partial_x p_\epsilon\right)_d = (Qp_\epsilon^-, p_\epsilon)_d$$

In the proof of Proposition 2.3.11, we have shown that $A_{\epsilon,ii}(p_\epsilon^+)\left(\partial_x p_{\epsilon,i}^-\right)^2 \geq \frac{\lambda_{min}}{2\lambda_{max}}\left(\partial_x p_{\epsilon,i}^-\right)^2$, so

$$\left(\partial_x p_\epsilon^-, \tilde{\sigma}_{Dup}^2 A_\epsilon\left(p_\epsilon^+\right)\partial_x p_\epsilon\right)_d \geq \underline{\sigma}^2\frac{\lambda_{min}}{2\lambda_{max}}|\partial_x p_\epsilon^-|_d^2.$$

By the Young inequality, for $\eta > 0$,

$$\left(\partial_x p_\epsilon^-, \left(\frac{1}{2}R_\epsilon(p_\epsilon^+)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup} + 2\partial_x\tilde{\sigma}_{Dup}\right)\Lambda - rI_d\right)p_\epsilon\right)_d \geq -K\left(\eta|\partial_x p_\epsilon^-|_d^2 + \frac{1}{4\eta}|p_\epsilon^-|_d^2\right),$$

where $K = \frac{1}{2}\frac{\lambda_{max}}{\lambda_{min}}\overline{\sigma}\left(\overline{\sigma} + 2||\partial_x\tilde{\sigma}_{Dup}||_\infty\right) + r$. We set $\eta = \underline{\sigma}^2\frac{\lambda_{min}}{4K\lambda_{max}}$, so that

$$\frac{1}{2}\frac{d}{dt}|p_\epsilon^-|_d^2 - (Qp_\epsilon^-, p_\epsilon^-)_d - \frac{K^2}{\underline{\sigma}^2}\frac{\lambda_{max}}{\lambda_{min}}|p_\epsilon^-|_d^2 \leq (Qp_\epsilon^-, p_\epsilon^+)_d - \underline{\sigma}^2\frac{\lambda_{min}}{4\lambda_{max}}|\partial_x p_\epsilon^-|_d^2. \tag{2.5.17}$$

We also check that $(Qp_\epsilon^-, p_\epsilon^+)_d \leq 0$. Indeed, as for $i \in \{1, ..., d\}$, $p_{\epsilon,i}^- p_{\epsilon,i}^+ = 0$ and $q_{ij} \geq 0$ for $j \neq i$,

$$(Qp_\epsilon^-, p_\epsilon^+)_d = \int_{\mathbb{R}}\sum_{i=1}^d p_{\epsilon,i}^+\left(\sum_{j=1}^d q_{ij}p_{\epsilon,j}^-\right)dx = \int_{\mathbb{R}}\sum_{i\neq j} q_{ij}p_{\epsilon,j}^- p_{\epsilon,i}^+ dx \leq 0,$$

so the r.h.s of (2.5.17) is nonpositive. Moreover, $(Qp_\epsilon^-, p_\epsilon^-)_d \leq d\overline{q}|p_\epsilon^-|_d^2$, so we have the inequality

$$\frac{1}{2}\frac{d}{dt}|p_\epsilon^-|_d^2 - \left(d\overline{q} + \frac{K^2}{\underline{\sigma}^2}\frac{\lambda_{max}}{\lambda_{min}}\right)|p_\epsilon^-|_d^2 \leq 0$$

We thus obtain that the function $t \to \exp\left(-2\left(d\overline{q} + \frac{K^2}{\underline{\sigma}^2}\frac{\lambda_{max}}{\lambda_{min}}\right)t\right)|p_\epsilon^-(t)|_d^2$ is non increasing. As $p_\epsilon^-(0) = 0$, we can conclude that $p_\epsilon^- \equiv 0$. □

We check that if p is the limit of a sequence $(p_{\epsilon_k})_{k\geq 1}$, with p_{ϵ_k} being a solution of $V_{Fin,\epsilon_k}(\mu)$ for $k \geq 1$, then p takes values in $\mathcal{D}$.

Lemma 2.5.8. *Let $(p_{\epsilon_k})_{k\geq 1}$ be a sequence such that for $k \geq 1$, p_{ϵ_k} is a solution to $V_{Fin,\epsilon_k}(\mu)$, and $\epsilon_k \underset{k\to\infty}{\rightarrow} 0$. If the sequence $(p_{\epsilon_k})_{k\geq 1}$ has the limit p, in the sense:*

$$\begin{aligned}
p_{\epsilon_k} &\rightarrow p \text{ in } L^2([0,T];H) \text{ weakly,}\\
p_{\epsilon_k} &\rightarrow p \text{ in } L^\infty([0,T];L) \text{ weakly-*},\\
p_{\epsilon_k} &\rightarrow p \text{ a.e. on } (0,T] \times \mathbb{R},
\end{aligned}$$

then under Assumptions (B) and (H), $\sum_{i=1}^d p_i$ is the unique solution to $LV(\mu_{X_0})$ and p takes values in $\mathcal{D}$.

Proof. For $k \geq 1$, we define $u_{\epsilon_k} := \sum_{i=1}^d p_{\epsilon_k,i}$, and $u := \sum_{i=1}^d p_i$. Then u_{ϵ_k} satisfies, for $\psi \in C_c^\infty((0,T),\mathbb{R})$ and $v \in H^1(\mathbb{R})$,

$$\begin{aligned}
0 &= -\int_0^T \psi'(t)(v, u_{\epsilon_k})_1 dt - \int_0^T \psi(t)r(\partial_x v, u_{\epsilon_k})_1 dt + \int_0^T \psi(t)\frac{1}{2}\left(\partial_x v, \tilde{\sigma}_{Dup}^2\partial_x u_{\epsilon_k}\right)_1 dt\\
&+ \int_0^T \psi(t)\left(\partial_x v, \frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x\tilde{\sigma}_{Dup}\right))\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)}u_{\epsilon_k}\right)_1 dt,
\end{aligned}$$

as for $i \in \{1,...,d\}$, $\sum_{j=1}^d q_{ij} = 0$ and $\sum_{j=1}^d A_{\epsilon,ij} = \frac{1}{2}$. As $p_{\epsilon_k} \to p$ in $L^2([0,T];H)$ weakly, the terms on the r.h.s of the first line converge to

$$-\int_0^T \psi'(t)(v,u)_1 dt - \int_0^T \psi(t) r(\partial_x v, u)_1 dt + \int_0^T \psi(t)\frac{1}{2}\left(\partial_x v, \tilde{\sigma}^2_{Dup}\partial_x u\right)_1 dt$$

It is sufficient to show that the term on the r.h.s of the second line converges to

$$\int_0^T \psi(t)\left(\partial_x v, \frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))u\right)_1 dt.$$

For a.e. $(t,x) \in [0,T]\times\mathbb{R}$, $u_{\epsilon_k}(t,x) \to u(t,x)$. Let us remark that by Proposition 2.5.7, for $k \geq 1$, $p_{\epsilon_k} \geq 0$, so p and u are nonnegative. If $u(t,x) > 0$, then $\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)}(t,x) \to 1$. If $u(t,x) = 0$, $\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)} u_{\epsilon_k}(t,x) \to 0$ as $\left|\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)}\right| \leq 1$. Let us use the following decomposition : $u_{\epsilon_k} = 1_{\{u>0\}}u_{\epsilon_k} + 1_{\{u=0\}}u_{\epsilon_k}$. We first study the limit, as $k \to \infty$, of

$$I_1(k) := \int_0^T \psi(t)\left(\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)} 1_{\{u>0\}}\partial_x v, u_{\epsilon_k}\right)_1 dt$$

The function u_{ϵ_k} converges weakly to u in $L^2([0,T], L^2(\mathbb{R}))$. Moreover the function

$$\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)} 1_{\{u>0\}}\partial_x v$$

converges strongly to $\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))1_{\{u>0\}}\partial_x v$, in $L^2([0,T], L^2(\mathbb{R}))$. Indeed, the convergence is a.e. and

$$\left|\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)} 1_{\{u>0\}}\partial_x v\right| \leq \frac{1}{2}\overline{\sigma}(\overline{\sigma} + 2||\partial_x \tilde{\sigma}||_\infty)|\partial_x v| \in L^2([0,T], L^2(\mathbb{R}))$$

and we conclude by dominated convergence. Therefore, $I_1(k)$ converges, as $k \to \infty$, to

$$\int_0^T \psi(t)\left(\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))\partial_x v, 1_{\{u>0\}}u\right)_1 dt = \int_0^T \psi(t)\left(\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))\partial_x v, u\right)_1 dt,$$

because $u \geq 0$. We then study the term

$$I_2(k) := \int_0^T \psi(t)\left(\frac{1}{2}\tilde{\sigma}_{Dup}(\tilde{\sigma}_{Dup} + 2\left(\partial_x \tilde{\sigma}_{Dup}\right))\frac{\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}}{\epsilon_k \vee \left(\sum_{i=1}^d \lambda_i p_{\epsilon_k,i}\right)} 1_{\{u=0\}}\partial_x v, u_{\epsilon_k}\right)_1 dt.$$

We notice that $I_2(k) \leq \frac{1}{2}\overline{\sigma}(\overline{\sigma} + 2||\partial_x \tilde{\sigma}||_\infty)\int_0^T |\psi(t)|\left(|\partial_x v|1_{\{u=0\}}, u_{\epsilon_k}\right)_1$, and the r.h.s. converges to 0 as $k \to \infty$, as $u_{\epsilon_k} \to u$ weakly in $L^2([0,T], L)$. We prove that the initial condition is satisfied with the arguments at the end of the proof of Proposition 2.3.10. So we conclude that u is, by Proposition 2.4.2, the unique solution to $LV(\mu_{X_0})$ and by Lemma 2.4.1, $u > 0$ a.e. on $(0,T)\times\mathbb{R}$. Finally, as p is nonnegative, p takes values in $\mathcal{D}$ a.e. on $[0,T]\times\mathbb{R}$. □

Proposition 2.5.9. *Under Assumption (B), (H) and Condition (C), there exists a solution $p \in C([0,T], L)$ to $V_{Fin,L^2}(\mu)$ such that $\sum_{i=1}^d p_i$ is the unique solution to $LV(\mu_{X_0})$.*

Proof. It is easy to check that the family $(p_\epsilon)_{\epsilon>0}$ satisfies uniform in ϵ energy estimates. Using (2.5.16), we also obtain that the family $\left(\frac{dp_\epsilon}{dt}\right)_{\epsilon>0}$ is bounded in $L^2\left([0,T], H'\right)$. Similarly to the proof of Proposition 2.3.13, there exists a subsequence $(p_{\epsilon_k})_{k\geq 1}$ converging to a function $p \geq 0$, with the convergences as $\epsilon_k \to 0$.

$$p_{\epsilon_k} \quad \to \quad p \text{ in } L^2([0,T];H) \text{ weakly},$$

$$\begin{aligned} p_{\epsilon_k} &\to p \text{ in } L^\infty([0,T];L) \text{ weakly-*}, \\ p_{\epsilon_k} &\to p \text{ a.e. on } [0,T]\times\mathbb{R}, \end{aligned}$$

We check that for $v \in H$ and $\psi \in C^1([0,T],\mathbb{R})$ with $\psi(T)=0$,

$$\int_0^T \left(\psi(t)\partial_x v, \frac{1}{2}R_{\epsilon_k}(p_{\epsilon_k})\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+\partial_x\tilde{\sigma}_{Dup}\right)\Lambda p_{\epsilon_k}\right)_d dt \to \int_0^T \left(\psi(t)\partial_x v, \frac{1}{2}R(p)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+\partial_x\tilde{\sigma}_{Dup}\right)\Lambda p\right)_d dt.$$

The convergence is justified as p takes values in $\mathcal{D}$ a.e. by Lemma 2.5.8, $R_{\epsilon_k}(p_{\epsilon_k}) \underset{k\to\infty}{\to} R(p)$ a.e. on $[0,T]\times\mathbb{R}$ (with the same argument that justifies $A_\epsilon(p_\epsilon)\to A(p)$ a.e. in (2.3.45)), the functions R_ϵ and R have uniform bounds by Lemma 2.5.1, and $p_{\epsilon_k}\to p$ weakly in $L^2([0,T],H)$. Similarly we also obtain the convergence:

$$\int_0^T (\psi(t)\partial_x v, A_{\epsilon_k}(p_{\epsilon_k})\tilde{\sigma}^2_{Dup}\partial_x p_{\epsilon_k})_d dt \to \int_0^T (\psi(t)\partial_x v, A(p)\tilde{\sigma}^2_{Dup}\partial_x p)_d dt.$$

As $p_{\epsilon_k}\to p$ weakly in $L^2([0,T],L)$, we finally obtain the equality:

$$\begin{aligned} \int_0^T \psi(t)(Qv + r\partial_x v, p(t))_d dt &= -\int_0^T (\psi'(t)v,p)_d dt - (v,p_0)_d\psi(0) \\ &+ \int_0^T (\psi(t)\partial_x v, A(p)\tilde{\sigma}^2_{Dup}\partial_x p)_d dt \\ &+ \int_0^T \left(\psi(t)\partial_x v, \frac{1}{2}R(p)\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+\partial_x\tilde{\sigma}_{Dup}\right)\Lambda p\right)_d dt, \end{aligned}$$

so that (2.5.3) is verified and we conclude the proof with the same arguments as in the end of the proof of Proposition 2.3.10 to obtain existence to $V_{Fin,L^2}(\mu)$. By Lemma 2.5.8, $\sum_{i=1}^d p_i$ solves $LV(\mu_{X_0})$, which has a unique solution by Proposition 2.4.2. Finally, if p is a solution to $V_{Fin,L^2}(\mu)$, then p is a.e. equal to a function that belongs to $C([0,T],L)$ and this concludes the proof. □

Case when μ is a general probability measure on $\mathbb{R}\times\mathcal{Y}$

We now prove Theorem 2.4.3. When μ is a general probability measure on $\mathbb{R}\times\mathcal{Y}$, for $\sigma>0$, we mollify each $\mu_i, 1\le i\le d$ into measures $\mu_{\sigma,i} := \mu_i * h_\sigma$ with square integrable densities, also denoted by $\mu_{\sigma,i}$, and we define μ_σ as the measure under which the conditional law of X_0 given $\{Y_0=i\}$ is given by $\mu_{\sigma,i}$ and $\mathbb{P}(Y_0=i)=\alpha_i$. Let $(\sigma_k)_{k\ge 0}$ be a decreasing sequence converging to 0 as $k\to\infty$. Let us denote by p_{σ_k} a solution to $V_{Fin,L^2}(\mu_{\sigma_k})$, which exists by Proposition 2.5.9. Let us remark that for $k\ge 0$, $\sum_{i=1}^d p_{\sigma_k,i}$ satisfies $LV\left(\sum_{i=1}^d \alpha_i\mu_{\sigma_k,i}\right)$. We compute energy estimates, using Proposition 2.4.2:

$$\begin{aligned} |p_{\sigma_k}(t)|_d^2 &\le \left|\left|\sum_{i=1}^d p_{\sigma_k,i}(t)\right|\right|_{L^2}^2 \le \frac{\zeta}{\sqrt{t}}, \\ \int_0^T |p_{\sigma_k}(t)|_d^2 dt &\le 2\zeta\sqrt{T}, \\ \int_t^T |\partial_x p_{\sigma_k}(s)|_d^2 ds &\le \frac{l_{max}(\Pi)}{\kappa\underline{\sigma}}\frac{\zeta}{\sqrt{t}}e^{\frac{2}{l_{min}(\Pi)}\left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)T}, \end{aligned} \tag{2.5.18}$$

where we used (2.5.14) for the last inequality. Repeating the arguments of the proof in Subsection 2.3.3, we obtain a subsequence again called p_{σ_k}, such that:

$$\begin{aligned} p_{\sigma_k} &\to p \text{ in } L^2((0,T];L) \text{ weakly}, \\ p_{\sigma_k} &\to p \text{ in } L^2_{loc}((0,T];H) \text{ weakly}, \\ p_{\sigma_k} &\to p \text{ in } L^\infty_{loc}((0,T];L) \text{ weakly-*} \\ p_{\sigma_k} &\to p \text{ a.e. on } (0,T]\times\mathbb{R}. \end{aligned}$$

We show that p takes values in $\mathcal{D}$ a.e. on $(0,T]\times\mathbb{R}$. To do so, we show that $u := \sum_{i=1}^d p_i$ is a solution to (LV) with initial condition μ_{X_0}. For $\psi \in C^1([0,T],\mathbb{R})$, and $v \in H^1(\mathbb{R})$, as $u_{\sigma_k} := \sum_{i=1}^d p_{\sigma_k,i} \to u$ in $L^2((0,T],L^2(\mathbb{R}))$ weakly,

$$-\int_0^T \psi'(t)(v,u_{\sigma_k})_1 dt \;\;\to\;\; -\int_0^T \psi'(t)(v,u)_1 dt,$$

$$\int_0^T \psi(t)\left(\partial_x v, \left(-r+\frac{1}{2}\tilde{\sigma}^2_{Dup}+\tilde{\sigma}_{Dup}\left(\partial_x\tilde{\sigma}_{Dup}\right)\right)u_{\sigma_k}\right)_1 dt \;\;\to\;\; \int_0^T \psi(t)\left(\partial_x v, \left(-r+\frac{1}{2}\tilde{\sigma}^2_{Dup}+\tilde{\sigma}_{Dup}\left(\partial_x\tilde{\sigma}_{Dup}\right)\right)u\right)_1 dt.$$

Let us define $\varUpsilon := \frac{1}{2^{1/4}-1}\left(\frac{l_{max}(\Pi)\zeta}{\kappa\underline{\sigma}} e^{\frac{2}{l_{min}(\Pi)}\left(b+\frac{C^2}{2\kappa\underline{\sigma}^2}\right)T}\right)^{1/2}$. By Lemma 2.3.14, we have that for $\forall t\in(0,T)$ and $k\geq 1$, $\int_0^t |\partial_x p_{\sigma_k}(s)|_d ds \leq \varUpsilon t^{1/4}$ and $\int_0^t |\partial_x p(s)|_d ds \leq \varUpsilon t^{1/4}$. Then, $\forall t \in (0,T)$,

$$\int_0^t \left|\partial_x u_{\sigma_k}\right|_d dt = \int_0^t \left|\sum_{i=1}^d \partial_x p_{\sigma_k,i}(s)\right|_1 ds \leq \sqrt{d}\int_0^t |\partial_x p_{\sigma_k}(s)|_d ds \leq \sqrt{d}\varUpsilon t^{1/4},$$

and $\int_0^t \left|\partial_x u\right|_d dt \leq \sqrt{d}\varUpsilon t^{1/4}$. The same arguments as those leading to (2.3.48) enable to prove that:

$$\int_0^T \psi(t)\left(\partial_x v, \tilde{\sigma}^2_{Dup}\partial_x u_{\sigma_k}\right) dt \to \int_0^T \psi(t)\left(\partial_x v, \tilde{\sigma}^2_{Dup}\partial_x u\right) dt.$$

The initial condition is treated as before. This is sufficient to prove that u solves $LV(\mu_{X_0})$ and assert that p takes values in $\mathcal{D}$ by Lemma 2.4.1. For $\psi \in C^1([0,T],\mathbb{R})$, such that $\psi(T)=0$ and $v\in H$, as $p_{\sigma_k}\to p$ weakly in $L^2((0,T],L)$,

$$\int_0^T (\psi'(s)v, p_{\sigma_k}(s))_d ds \underset{k\to\infty}{\to} \int_0^T (\psi'(s)v, p(s))_d ds$$

$$\int_0^T \psi(t)(Qv, p_{\sigma_k})_d dt \underset{k\to\infty}{\to} \int_0^T \psi(t)(Qv, p)_d dt$$

We check that:

$$\int_0^T \psi(t)\left(\partial_x v, R(p_{\sigma_k}(t))\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+\partial_x\tilde{\sigma}_{Dup}\right)\Lambda p_{\sigma_k}(t)\right)_d dt \underset{k\to\infty}{\to} \frac{1}{2}\int_0^T \psi(t)\left(\partial_x v, R(p(t))\tilde{\sigma}_{Dup}\left(\tilde{\sigma}_{Dup}+\partial_x\tilde{\sigma}_{Dup}\right)\Lambda p(t)\right)_d,$$

which is the case as the sequence $(p_{\sigma_k})_{k\geq 0}$ converges to p weakly in $L^2((0,T],L)$, R is bounded and continuous on $\mathcal{D}$, and $R(p_{\sigma_k})\to R(p)$ a.e. on $(0,T]\times\mathbb{R}$. We also check that:

$$\int_0^T (\psi(s)\partial_x v, \tilde{\sigma}^2_{Dup}A(p_{\sigma_k})\partial_x p_{\sigma_k})_d ds \underset{k\to\infty}{\to} \int_0^T (\psi(s)\partial_x v, \tilde{\sigma}^2_{Dup}A(p)\partial_x p)_d ds,$$

using the same argument as for (2.3.48). Arguments similar to (2.3.54) enable to check the initial condition and thus assert existence to $V_{Fin}(\mu)$. As any solution to $V_{Fin}(\mu)$ has a representative in $C((0,T],L)$, this concludes the proof of Theorem 2.4.3.

2.5.2 Existence of a weak solution to the SDE (2.4.1)

To prove Theorem 2.4.4, we use Theorem 2.5.11 below, which is a generalization of [39, Theorem 2.6] to make a link between the existence to a Fokker-Planck system and the existence to the corresponding martingale problem. Theorem 2.5.11 is proved in Appendix 2.D. There exist several generalizations of [39, Theorem 2.6], among whom we can mention [40], where the coefficients of the generator are no longer bounded but have linear growth, and [103], where the author deals with a partial integro differential equation with a Lévy generator.

Generalization of [39, Theorem 2.6]

Given functions $(b_i)_{1\leq i\leq d}, (a_i)_{1\leq i\leq d}, (q_{ij})_{1\leq i,j\leq d}$ defined on $[0,T]\times\mathbb{R}$ and a finite measure μ_0 on $\mathbb{R}\times\mathcal{Y}$, we study the following PDS, where for $1\leq i\leq d$:

$$\partial_t\mu_i + \partial_x(b_i\mu_i) - \frac{1}{2}\partial^2_{xx}(a_i\mu_i) - \sum_{j=1}^d q_{ji}\mu_j = 0 \text{ in } (0,T)\times\mathbb{R} \tag{2.5.19}$$

$$\mu_i(0) = \mu_0(\cdot,\{i\}). \tag{2.5.20}$$

Definition 2.5.10. A family of vectors of Borel measures $(\mu_1(t,\cdot),...,\mu_d(t,\cdot))_{t\in(0,T]}$ is a solution to the PDS (2.5.19)-(2.5.20) if for any function ϕ defined on $\mathcal{S}$ such that $\forall i\in\{1,...,d\}, \phi(\cdot,i)\in C_c^\infty(\mathbb{R})$,

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}}\sum_{i=1}^d\phi(x,i)\mu_i(t,dx) &= \sum_{i=1}^d\int_{\mathbb{R}}\left(b_i(t,x)\partial_x\phi(x,i)+\frac{1}{2}a_i(t,x)\partial^2_{xx}\phi(x,i)\right)\mu_i(t,dx)\\ &+ \sum_{i=1}^d\sum_{j=1}^d\left(\int_{\mathbb{R}}q_{ji}(t,x)\phi(x,i)\mu_j(t,dx)\right),\end{aligned} \tag{2.5.21}$$

in the distributional sense on $(0,T)$, and for $1\leq i\leq d$ and $\psi\in C_b^2(\mathbb{R})$, the function $t\to\int_{\mathbb{R}}\psi(x)\mu_i(t,dx)$ is continuous on $(0,T]$ and converges to $\int_{\mathbb{R}}\psi(x)\mu_0(dx,\{i\})$ as $t\to 0$.

We suppose that all the coefficients $b_i, a_i, q_{ij}, 1\leq i,j\leq d$, are uniformly bounded on $[0,T]\times\mathbb{R}$, that the coefficients $(a_i)_{1\leq i\leq d}$ are non negative, that the coefficients $(q_{ij})_{1\leq i,j\leq d}$ are non negative functions for $i\neq j$, and $q_{ii} = -\sum_{j\neq i}q_{ij}$. We introduce the SDE

$$dX_t = b_{Y_t}(t,X_t)dt + \sqrt{a_{Y_t}(t,X_t)}dW_t, \tag{2.5.22}$$

where Y_t is a stochastic process with values in $\mathcal{Y}$, and that satisfies, for $j\neq Y_t$,

$$\mathbb{P}\left(Y_{t+dt}=j|\,(X_s,Y_s)_{0\leq s\leq t}\right) = q_{Y_tj}(t,X_t)dt.$$

We define $E=\{(X,Y), X\in C([0,T],\mathbb{R}), Y \text{ càdlàg with values in } \mathcal{Y}\}$ endowed with the Skorokhod topology. For a probability measure m on $\mathbb{R}\times\mathcal{Y}$, a probability measure ν on E is a martingale solution to the SDE (2.5.22) with initial condition m if under the probability ν, the canonical process (X,Y) on E satisfies $(X_0,Y_0)\sim m$ and for any function ϕ defined on $\mathbb{R}\times\mathcal{Y}$ s.t. $\forall i\in\{1,...,d\}, \phi(\cdot,i)\in C_b^2(\mathbb{R})$, the process

$$\phi(X_t,Y_t)-\phi(X_0,Y_0)-\int_0^t\left(\frac{1}{2}a_{Y_s}(s,X_s)\partial^2_{xx}\phi(X_s,Y_s)+b_{Y_s}(s,X_s)\partial_x\phi(X_s,Y_s)+\sum_{l=1}^d q_{Y_sl}(s,X_s)\phi(X_s,l)\right)ds$$

is a ν-martingale. Now we can state the generalization of [39, Theorem 2.6], that will be used in the following section.

Theorem 2.5.11. *Let $(\mu_1(t,\cdot),...,\mu_d(t,\cdot))_{t\in(0,T]}$ be a solution to the PDS (2.5.19)-(2.5.20) where the initial condition μ_0 is a probability measure on $\mathbb{R}\times\mathcal{Y}$. We moreover suppose that there exists $B>0$, s.t. for $1\leq i\leq d$, and $t\in(0,T]$, $\mu_i(t,\mathbb{R})\leq B$. Then the SDE (2.5.22) with initial distribution μ_0 has a martingale solution ν which satisfies the following representation formula: for $1\leq i\leq d$ and $\psi\in C_c^\infty(\mathbb{R})$,*

$$\int_{\mathbb{R}}\psi(x)\mu_i(t,dx) = \int_E\psi(X_t)1_{\{Y_t=i\}}d\nu(X,Y).$$

Proof of Theorem 2.4.4

Let p be a solution to $V_{Fin}(\mu)$ with the properties stated in Theorem 2.4.3. To show that $(p_1(t,x)dx,...,p_d(t,x)dx)_{t\in(0,T]}$ satisfies the variational formulation in the sense of distributions (2.5.21), we check that for $1\leq i\leq d$, $\tilde{\sigma}^2_{Dup}\frac{\sum_{k=1}^d p_k}{\sum_{k=1}^d\lambda_kp_k}\lambda_ip_i\in H^1(\mathbb{R})$, and $\partial_x\left(\tilde{\sigma}^2_{Dup}\frac{\sum_{k=1}^d p_k}{\sum_{k=1}^d\lambda_kp_k}\lambda_ip_i\right) = \tilde{\sigma}^2_{Dup}\left(A(p)p\right)_i + \frac{\sum_{k=1}^d p_k}{\sum_{k=1}^d\lambda_kp_k}\lambda_ip_i\partial_x\tilde{\sigma}^2_{Dup}$. As $\tilde{\sigma}_{Dup}\in$

$L^{\infty}([0,T],W^{1,\infty}(\mathbb{R}))$, it is sufficient to check that $\frac{\sum_{k=1}^d p_k}{\sum_{k=1}^d \lambda_k p_k}\lambda_i p_i \in H^1(\mathbb{R})$ and that $\partial_x\left(\frac{\sum_{k=1}^d p_k}{\sum_{k=1}^d \lambda_k p_k}\lambda_i p_i\right) = (A(p)p)_i$, and the proof is exactly the same as the one of Lemma 2.3.15, thanks to Lemma 2.4.1 which ensures the positivity of $\sum_{i=1}^d p_i$ a.e. on $(0,T]\times\mathbb{R}$. Then by Theorem 2.5.11, there exists a measure ν under which $(X_0,Y_0)\sim\mu$ and for any function ϕ defined on $\mathcal{S}$ s.t. $\forall i\in\{1,...,d\}, \phi(\cdot,i)\in C_b^2(\mathbb{R})$, the function

$$\phi(X_t,Y_t)-\phi(X_0,Y_0)-\int_0^t \frac{1}{2}\tilde{\sigma}^2_{Dup}f^2(Y_s)\frac{\sum_{i=1}^d p_k}{\sum_{i=1}^d \lambda_k p_k}(s,X_s)\partial^2_{xx}\phi(X_s,Y_s)ds$$
$$-\int_0^t\left(r-\frac{1}{2}\tilde{\sigma}^2_{Dup}f^2(Y_s)\frac{\sum_{i=1}^d p_k}{\sum_{i=1}^d \lambda_k p_k}(s,X_s)\right)\partial_x\phi(X_s,Y_s)-\sum_{l=1}^d q_{Y_s l}(X_s)\phi(X_s,l)ds,$$

is a ν-martingale. Moreover, for $h:\mathcal{Y}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ continuous and bounded functions, we have that:

$$\mathbb{E}\left(h^2(Y_s)g(X_s)\right)=\sum_{i=1}^d\mathbb{E}\left(h^2(Y_s)1_{\{Y_s=y_i\}}g(X_s)\right)=\int_{\mathbb{R}}g(x)\sum_{i=1}^d h^2(y_i)p_i(t,x)dx.$$

Taking $h\equiv 1$ and $h\equiv f$, we check that the time marginals of X are given by $\sum_{i=1}^d p_i$ and that $\mathbb{E}\left(f^2(Y_s)|X_s\right) = \frac{\sum_{k=1}^d \lambda_k p_k}{\sum_{k=1}^d p_k}(s,X_s)$. Therefore ν is a solution to the martingale problem associated to the SDE (2.4.1), and we obtain existence of a weak solution to the SDE (2.4.1) by [73, Theorem 2.3], and which has the same time marginals as the solution of the SDE (2.4.4).

2.5.3 A more general fake Brownian motion

We consider the SDE:

$$dX_t=\frac{f(Y_t)}{\sqrt{\mathbb{E}[f^2(Y_t)|X_t]}}dW_t, \tag{2.5.23}$$

where the process $(Y_t)_{t\geq 0}$ takes values in $\mathcal{Y}$ and $\mathbb{P}\left(Y_{t+dt}=j|\left(X_s,Y_s\right),0\leq s\leq t\right)=q_{Y_t j}(X_t)dt$, for $1\leq j\leq d$ and $j\neq Y_t$, with the functions $(q_{ij})_{1\leq i\neq j\leq d}$ non negative and bounded, and $q_{ii}=-\sum_{j\neq i}q_{ij}$ for $1\leq i\leq d$. The vector (X_0,Y_0) has the probability distribution μ on $\mathbb{R}\times\mathcal{S}$ and is independent from $(W_t)_{t\geq 0}$. The associated Fokker-Planck PDS writes:

$$\begin{aligned}\forall i\in\{1,...,d\},\ \partial_t p_i &= \frac{1}{2}\partial^2_{xx}\left(\frac{\sum_{i=1}^d p_k}{\sum_{i=1}^d \lambda_k p_k}\lambda_i p_i\right)+\sum_{j=1}^d q_{ji}p_j\\ p_i(0,\cdot) &= \alpha_i\mu_i,\end{aligned}$$

with $(\alpha_i)_{1\leq i\leq d}$ and $(\mu_i)_{1\leq i\leq d}$ defined as in Section 2.4. We introduce an associated variational formulation $V_{Jump}(\mu)$:

$$\begin{gathered}\text{Find}\quad p=(p_1,...,p_d)\text{ satisfying:}\\ p\in L^2_{loc}((0,T];H)\cap L^\infty_{loc}((0,T];L),\\ p\text{ takes values in }\mathcal{D},\text{ a.e. on }(0,T)\times\mathbb{R},\\ \forall v\in H,\ \frac{d}{dt}(v,p)_d+(\partial_x v,A(p)\partial_x p)_d = (Qv,p)_d,\\ \text{in the sense of distributions on }(0,T),\text{ and}\\ p(t,\cdot)\underset{t\to 0^+}{\overset{\text{weakly-*}}{\rightarrow}}p_0:=(\alpha_1\mu_1,...,\alpha_d\mu_d)\end{gathered}$$

With the same arguments used to prove Theorems 2.4.3 and 2.4.4, we can prove the following results. The main difference is that $\sum_{i=1}^d p$ is solution to the heat equation and not the Dupire PDE, but existence and uniqueness, positivity and Aronson-like estimates of the solution hold in both cases, so it does not affect the proof.

Theorem 2.5.12. *Under Condition (C), $V_{Jump}(\mu)$ has a solution $p \in C((0,T], L)$. Moreover, $\sum_{i=1}^{d} p_i(t,x) = \mu_{X_0} * h_t(x)$ a.e. on $(0,T] \times \mathbb{R}$.*

Theorem 2.5.13. *Under Condition (C), SDE (2.5.23) has a weak solution, and its time marginals are those of $(Z + W_t)_{t\geq 0}$, where Z has the law μ_{X_0} and is independent from $(W_t)_{t\geq 0}$.*

Moreover, the solutions to SDE (2.5.23) are also continuous fake Brownian motions provided that $f(Y_0)$ can take at least two distinct values with positive probability. Let us define $\tilde{\mathcal{Y}} := \{i \in \mathcal{Y},\ \mathbb{P}(Y_0 = i) > 0\}$.

Proposition 2.5.14. *Under Condition (C), if f is non constant on $\tilde{\mathcal{Y}}$, the solutions to SDE (2.5.23) with initial condition $\mu = \delta_0$ are continuous fake Brownian motions.*

Proof. Let $(X_t)_{t\geq 0}$ be a solution to SDE (2.5.23), with initial condition $X_0 = 0$. The process X is a continuous martingale, and by [58, Theorem 4.6], for $t \geq 0$, $X_t \sim \mathcal{N}(0,t)$. We consider its quadratic variation $d\langle X\rangle_t = \frac{f^2(Y_t)}{\mathbb{E}[f^2(Y_t)|X_t]}dt$. We reason by contraposition and first suppose that a.s., for a.e. $t > 0$, the equality $f^2(Y_t) = \mathbb{E}\left[f^2(Y_t)\,|X_t\right]$. Then a.s., for $t > 0$, $X_t = W_t$ and there exists a measurable function ψ_t, such that $\psi_t(X_t) = f^2(Y_t)$. For $i \in \tilde{\mathcal{Y}}$ and $g : \mathbb{R} \to \mathbb{R}$ measurable and non negative, let us fix $t > 0$ such that those properties hold and consider the term:

$$
\begin{aligned}
\mathbb{E}\left[g(X_t)\,1_{\{\forall s\in[0,t],Y_s=i\}}\right] &= \mathbb{E}\left[g(X_t)\,\mathbb{E}\left[1_{\{\forall s\in[0,t],Y_s=i\}}|\,(X_s)_{0\leq s\leq t}\,,Y_0\right]\right] \\
&= \mathbb{E}\left[g(X_t)\exp\left(\int_0^t q_{ii}(s,X_s)\,ds\right)1_{\{Y_0=i\}}\right] \\
&\geq \exp(-\overline{q}t)\,\mathbb{E}\left[1_{\{Y_0=i\}}g(W_t)\right] = \alpha_i\frac{\exp(-\overline{q}t)}{\sqrt{2\pi t}}\int_{\mathbb{R}^2} g(y)\exp\left(-\frac{y^2}{2t}\right)dy,
\end{aligned}
$$

by independence between Y_0 and $(W_t)_{t\geq 0}$. Therefore on the event $\{\forall s \in [0,t], Y_s = y_i\}$, the random variable X_t belongs with positive probability to any Borel set with positive Lebesgue measure. By the equality $\psi_t(X_t) = f^2(Y_t)$, the previous observation implies that f is constant on $\tilde{\mathcal{Y}}$. To conclude the proof, if f is non constant on $\tilde{\mathcal{Y}}$, by contraposition, the function $t \to \langle X\rangle_t$ is not equal to the identity function, and $(X_t)_{t\geq 0}$ is a fake Brownian motion. ☐

2.A Local Lipschitz property

2.A.1 Proof of Lemma 2.3.6

Proof. Let $m \geq 1$. It is sufficient to show that the function $z \in \mathbb{R}^m \to K_\epsilon^m(z)$ is locally Lipschitz to have $z \in \mathbb{R}^m \to F_\epsilon^m(z)$ locally Lipschitz. Let us remark that for $a, b \in \{1, ..., m\}$, $(K_\epsilon(z) - K_\epsilon(\tilde{z}))_{ab}$ only contains integrals of type:

$$\int_{\mathbb{R}} \Theta\left(M_{\epsilon,ij}\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right) - M_{\epsilon,ij}\left(\left(\sum_{k=1}^m \tilde{z}_k w_k\right)^+\right)\right)dx,\ \ 1 \leq i,j \leq d,$$

where $\Theta \in L^1(\mathbb{R})$, as for $c, d \in \{1, ..., d\}$, $\partial_x w_{ac}\partial_x w_{bd} \in L^1(\mathbb{R})$. For $\Theta \in L^1(\mathbb{R})$, $1 \leq i,j \leq d$ and $z \in \mathbb{R}^m$, let us define

$$g_{ij}(z) := \int_{\mathbb{R}} \Theta M_{\epsilon,ij}\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right)dx.$$

Let C be a compact subset of $\mathbb{R}^m$. Let us show that the function $z \to g_{ij}(z)$ is Lipschitz on C and conclude by linearity. For $i \neq j$, $\rho, \tilde{\rho} \in \mathbb{R}^d$,

$$
\begin{aligned}
|M_{\epsilon,ij}(\rho^+) - M_{\epsilon,ij}(\tilde{\rho}^+)| &= \left|\lambda_i\rho_i^+\frac{\sum_{l\neq j}(\lambda_l-\lambda_j)\rho_l^+}{\left(\epsilon\vee\left(\sum_l\lambda_l\rho_l^+\right)\right)^2} - \lambda_i\tilde{\rho}_i^+\frac{\sum_{l\neq j}(\lambda_l-\lambda_j)\tilde{\rho}_l^+}{\left(\epsilon\vee\left(\sum_l\lambda_l\tilde{\rho}_l^+\right)\right)^2}\right| \\
&\leq |\Delta_1(\rho,\tilde{\rho})| + |\Delta_2(\rho,\tilde{\rho})| + |\Delta_3(\rho,\tilde{\rho})|
\end{aligned}
$$

with

$$
\begin{aligned}
|\Delta_1(\rho,\tilde\rho)| := \left|\lambda_i\left(\rho_i^+ - \tilde\rho_i{}^+\right)\frac{\sum_{l\neq j}(\lambda_l-\lambda_j)\rho_l^+}{\left(\epsilon\vee\left(\sum_l\lambda_l\rho_l^+\right)\right)^2}\right| &\leq \left|\frac{\sum_{l\neq j}(\lambda_l-\lambda_j)\rho_l^+}{\epsilon^2}\right|\lambda_i|\rho_i-\tilde\rho_i| \\
&\leq \frac{d-1}{\epsilon^2}|\lambda_{max}-\lambda_{min}|\lambda_{max}||\rho||_\infty||\rho-\tilde\rho||_\infty \\
|\Delta_2(\rho,\tilde\rho)| := \left|\lambda_i\tilde\rho_i^+\frac{\sum_{l\neq j}(\lambda_l-\lambda_j)\left(\rho_l^+-\tilde\rho_l{}^+\right)}{\left(\epsilon\vee\left(\sum_l\lambda_l\rho_l^+\right)\right)^2}\right| &\leq \frac{d-1}{\epsilon^2}|\lambda_{max}-\lambda_{min}|\lambda_{max}||\tilde\rho||_\infty||\rho-\tilde\rho||_\infty
\end{aligned}
$$

$$
\begin{aligned}
|\Delta_3(\rho,\tilde\rho)| &:= \left|\lambda_i\tilde\rho_i^+\left(\sum_{l\neq j}(\lambda_l-\lambda_j)\tilde\rho_l{}^+\right)\left(\frac{1}{\left(\epsilon\vee\left(\sum_l\lambda_l\rho_l^+\right)\right)^2}-\frac{1}{\left(\epsilon\vee\left(\sum_l\lambda_l\tilde\rho_l^+\right)\right)^2}\right)\right| \\
&\leq \frac{1}{\epsilon^4}\lambda_i\tilde\rho_i^+\left|\sum_{l\neq j}(\lambda_l-\lambda_j)\left(\tilde\rho_l{}^+\right)\right|\left|2\epsilon+\sum_l\lambda_l\left(\rho_l^++\tilde\rho_l^+\right)\right|\left|\sum_l\lambda_l\left(\rho_l-\tilde\rho_l\right)\right| \\
&\leq \frac{\lambda_{max}^2}{\epsilon^4}d(d-1)|\lambda_{max}-\lambda_{min}|\left(2\epsilon+d\lambda_{max}\left(||\rho||_\infty+||\tilde\rho||_\infty\right)\right)||\tilde\rho||_\infty^2||\rho-\tilde\rho||_\infty.
\end{aligned}
$$

where we used the fact that $\forall a,b\in\mathbb{R}, |a^+-b^+|\leq|a-b|$ and that in the last inequality, $\forall\epsilon,a,b\geq 0, |\left(\epsilon\vee a\right)^2-\left(\epsilon\vee b\right)^2| = |\epsilon\vee a+\epsilon\vee b||\epsilon\vee a-\epsilon\vee b|\leq|2\epsilon+a+b||a-b|$. We now replace ρ by $\sum_{k=1}^m z_k w_k$, and $\tilde\rho$ by $\sum_{k=1}^m\tilde z_k w_k$. As the sequence $(w_k)_{k\geq1}$ belongs to H, by [19, Corollary VIII.8], for $k\geq 1$ and $1\leq i\leq d$, $w_{ki}\in L^\infty(\mathbb{R})$ and the function $(x,z)\in\mathbb{R}\times C\to\sum_{k=1}^m z_k w_k(x)\in\mathbb{R}^d$ takes values for a.e. $x\in\mathbb{R}$ in a bounded subset of $\mathbb{R}^d$. Then there exists an uniform bound $B<\infty$ s.t. $\forall z,\tilde z\in C$,

$$
\left|\left|\Delta_1\left(\sum_{k=1}^m z_k w_k,\sum_{k=1}^m\tilde z_k w_k\right)\right|\right|_\infty+\left|\left|\Delta_2\left(\sum_{k=1}^m z_k w_k,\sum_{k=1}^m\tilde z_k w_k\right)\right|\right|_\infty+\left|\left|\Delta_3\left(\sum_{k=1}^m z_k w_k,\sum_{k=1}^m\tilde z_k w_k\right)\right|\right|_\infty\leq B||z-\tilde z||_\infty,
$$

so after integration against $\Theta\in L^1(\mathbb{R})$,

$$
|g_{ij}(z)-g_{ij}(\tilde z)|\leq||\Theta||_{L^1}B||z-\tilde z||_\infty,
$$

and g_{ij} is Lipschitz on C. Exactly in the same way, we also obtain that the functions $(g_{ii})_{i\in\{1,\ldots,d\}}$, are Lipschitz on C, so K_ϵ^m is locally Lipschitz and this concludes the proof. ☐

2.A.2 Proof of Lemma 2.5.3

Proof. Let $m\geq 0$. It is sufficient to prove that $K_{\epsilon,1}^m$ and $K_{\epsilon,2}^m$ are locally Lipschitz in z uniformly in t. In the proof of Lemma 2.3.6, we have shown that the function

$$
z\in\mathbb{R}^m\to\int_{\mathbb{R}}\Theta A_{\epsilon,ij}\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right)dx,
$$

is locally Lipschitz, uniformly for Θ in a bounded subset of $L^1(\mathbb{R})$. As the family

$$
\left(\tilde\sigma_{Dup}^2(t,\cdot)\partial_x w_{ik}\partial_x w_{jl}\right)_{1\leq i,j\leq m,1\leq k,l\leq d,t\in[0,T]},
$$

belongs to a bounded subset of $L^1(\mathbb{R})$, we then have that $F_{\epsilon,2}$ is locally Lipschitz in z uniformly in t . To show the local Lipschitz property of $F_{\epsilon,1}$, it is sufficient to prove that for any function $\Theta\in L^1(\mathbb{R})$, the function

$$
z\in\mathbb{R}^m\to\int_{\mathbb{R}}\Theta R_\epsilon\left(\left(\sum_{k=1}^m z_k w_k\right)^+\right)dx,
$$

is locally Lipschitz in z uniformly in t, as the functions $(w_{jk}\partial_x w_{il})_{1\leq i,j\leq m,1\leq k,l\leq d}$ belong to $L^1(\mathbb{R})$, and $\tilde\sigma_{Dup}$ and $\partial_x\tilde\sigma_{Dup}$ are uniformly bounded. The result is obtained since the function $\rho\to R_\epsilon\left(\rho^+\right)$ is locally Lipschitz and the functions $(w_{ki})_{1\leq k\leq m,1\leq i\leq d}$ belong to $L^\infty(\mathbb{R})$, as in the proof of Lemma 2.3.6. ☐

2.B About Condition (C)

In Subsection 2.B.1, we give a necessary and sufficient condition for a diagonal matrix to satisfy Condition (C). We also give a numerical procedure to check if there exists a diagonal matrix that satisfies (C). Then, in Subsection 2.B.2, we focus on the case $d = 3$, and give a simple necessary and sufficient condition for (C) to be satisfied. When $d = 3$, (C) is satisfied if and only if it is satisfied by a diagonal matrix. When $d \geq 4$, we do not know if this property still holds.

2.B.1 The diagonal case

For $k \geq 1$, $\delta := (\delta_1, ..., \delta_k) \in \mathbb{R}^k$, let us denote by $Diag(\delta) \in \mathcal{M}_k(\mathbb{R})$ the diagonal matrix with coefficients $\delta_1, ..., \delta_k$.

Proposition 2.B.1. *For $d \geq 2$ and $\alpha := (\alpha_1, ..., \alpha_d) \in \left(\mathbb{R}_+^*\right)^d$, $Diag(\alpha)$ satisfies Condition (C) if and only if*

$$\frac{2}{\alpha_k} + \sum_{i \neq k} \frac{1}{\alpha_i} > \sqrt{\sum_{i \neq k} \frac{\lambda_i}{\alpha_i} \sum_{i \neq k} \frac{1}{\lambda_i \alpha_i}}, \; 1 \leq k \leq d. \tag{2.B.1}$$

Proof. For $1 \leq k \leq d$, the symmetric matrix $D^{(k)}$ with coefficients

$$D^{(k)}_{ij} = \frac{\lambda_i + \lambda_j}{2} \left(\alpha_i 1_{\{i=j\}} + \alpha_k - \alpha_k 1_{\{i=k\}} - \alpha_k 1_{\{j=k\}}\right)$$

for $1 \leq i, j \leq d$, is positive definite on $e_k^\perp$ if and only if the matrix $\tilde{D}^{(k)}$ defined as $D^{(k)}$ with its k-th row and k-th column removed, is positive definite on $\mathbb{R}^{d-1}$. Here we only show how to deal with the case $k = d$, but the same arguments can be used for the indices $1 \leq k \leq d-1$. The matrix $\tilde{D}^{(d)}$ has coefficients

$$\tilde{D}^{(d)}_{ij} = \frac{\lambda_i + \lambda_j}{2} \left(\alpha_i 1_{\{i=j\}} + \alpha_d\right).$$

for $1 \leq i, j \leq d-1$. We define $\Delta := Diag\left(\left(\sqrt{\lambda_i \alpha_i}\right)_{1 \leq i \leq d-1}\right)$. The matrix $\tilde{D}^{(d)}$ rewrites

$$\begin{aligned}
\tilde{D}^{(d)} &= \Delta\Delta + \frac{\alpha_d}{2}\left((\lambda_i + \lambda_j)\right)_{1 \leq i,j \leq d-1} \\
&= \Delta\left(I_{d-1} + \frac{\alpha_d}{2}\left(\frac{\lambda_i + \lambda_j}{\sqrt{\lambda_i \alpha_i}\sqrt{\lambda_j \alpha_j}}\right)_{1 \leq i,j \leq d-1}\right)\Delta \\
&= \Delta\left(I_{d-1} + \frac{\alpha_d}{2}(ab^* + ba^*)\right)\Delta,
\end{aligned}$$

where $a = \left(\sqrt{\frac{\lambda_i}{\alpha_i}}\right)_{1 \leq i \leq d-1}$ and $b = \left(\frac{1}{\sqrt{\lambda_i \alpha_i}}\right)_{1 \leq i \leq d-1}$. The matrix $\tilde{D}^{(d)}$ is positive definite if and only if the matrix

$$\left(I_{d-1} + \frac{\alpha_d}{2}(ab^* + ba^*)\right),$$

has positive eigenvalues. The columns of the matrix $ab^* + ba^*$ are linear combinations of a and b. If a and b are not colinear (resp. colinear), then the matrix $ab^* + ba^*$ has eigenvalues 0 with multiplicity $d-2$ and $\sum_{i=1}^{d-1} a_i b_i - \sqrt{\sum_{i=1}^{d-1} a_i^2}\sqrt{\sum_{i=1}^{d-1} b^2} < 0$ for the eigenvector $a - \frac{\sqrt{\sum_{i=1}^{d-1} a_i^2}}{\sqrt{\sum_{i=1}^{d-1} b^2}} b$ (resp. 0 with multiplicity $d-1$), and $\sum_{i=1}^{d-1} a_i b_i + \sqrt{\sum_{i=1}^{d-1} a_i^2}\sqrt{\sum_{i=1}^{d-1} b^2} > 0$ for the eigenvector $a + \frac{\sqrt{\sum_{i=1}^{d-1} a_i^2}}{\sqrt{\sum_{i=1}^{d-1} b^2}} b$. Thus, $\tilde{D}^{(d)}$ is definite positive if and only if

$$1 + \frac{\alpha_d}{2}\left(\sum_{i \neq d} \frac{1}{\alpha_i} - \sqrt{\sum_{i \neq d} \frac{\lambda_i}{\alpha_i} \sum_{i \neq d} \frac{1}{\lambda_i \alpha_i}}\right) > 0.$$

which is equivalent to (2.B.1) for $k = d$. Using the same arguments on $\tilde{D}^{(k)}$ for $1 \leq k \leq d-1$, we obtain (2.B.1). $\square$

The choice $\alpha = (1, ..., 1) \in \mathbb{R}^d$ in Inequality (2.B.1) gives a sufficient condition for the identity matrix I_d to satisfy Condition (C).

Corollary 2.B.2. *If the condition*

$$\max_{1\leq k\leq d} \sqrt{\sum_{i\neq k} \lambda_i \sum_{i\neq k} \frac{1}{\lambda_i}} < d+1, \tag{2.B.2}$$

is satisfied then Condition (C) is satisfied for the choice $\Gamma = I_d$. In particular, if $\lambda_1 = ... = \lambda_d$, then (C) is satisfied.

Moreover, for $d = 2$, Inequality (2.B.2) is always satisfied, as for $k = 1, 2$, $\sqrt{\lambda_k \frac{1}{\lambda_k}} = 1 < 3$.

Corollary 2.B.3. *If $d = 2$, then Condition (C) is satisfied for the choice $\Gamma = I_2$.*

From Inequality (2.B.1), we deduce a method to check numerically whether there exists a diagonal matrix that satisfies Condition (C). We suppose that the values of $\lambda_1, ..., \lambda_d$ are not equal, otherwise by Corollary 2.B.2, I_d satisfies Condition (C). For $z = (z_1, ..., z_d) \in \left(\mathbb{R}^2\right)^d$, let us denote by $\mathcal{C}(z) = \{x \in \mathbb{R}^2 | \exists \mu_1, ..., \mu_d > 0, \sum_{i=1}^d \mu_i = 1, x = \sum_{i=1}^d \mu_i z_i\}$, the strict convex envelope of z. Let $\alpha := (\alpha_1, ..., \alpha_d) \in \left(\mathbb{R}_+^*\right)^d$, such that $Diag(\alpha)$ satisfies Condition (C). By Proposition 2.B.1, (2.B.1) holds and rewrites

$$\left(\frac{1}{\alpha_k} + \sum_{i=1}^d \frac{1}{\alpha_i}\right)^2 > \left(\sum_{i=1}^d \frac{\lambda_i}{\alpha_i} - \frac{\lambda_k}{\alpha_k}\right)\left(\sum_{i=1}^d \frac{1}{\lambda_i\alpha_i} - \frac{1}{\lambda_k\alpha_k}\right),\ 1 \leq k \leq d. \tag{2.B.3}$$

If we define $\tilde{\lambda} = \sum_{i=1}^d \lambda_i \frac{\frac{1}{\alpha_i}}{\sum_{k=1}^d \frac{1}{\alpha_k}} \in \mathcal{C}\left((\lambda_i)_{1\leq i\leq d}\right)$, $\widetilde{\lambda^{-1}} = \sum_{i=1}^d \frac{1}{\lambda_i} \frac{\frac{1}{\alpha_i}}{\sum_{k=1}^d \frac{1}{\alpha_k}} \in \mathcal{C}\left(\left(\frac{1}{\lambda_i}\right)_{1\leq i\leq d}\right)$, then Inequality (2.B.3) writes

$$1 + \frac{\frac{1}{\alpha_k}}{\sum_{i=1}^d \frac{1}{\alpha_i}}\left(2 + \frac{1}{\lambda_k}\tilde{\lambda} + \lambda_k\widetilde{\lambda^{-1}}\right) > \tilde{\lambda}\widetilde{\lambda^{-1}},\ 1 \leq k \leq d. \tag{2.B.4}$$

We deduce that there exists a diagonal matrix that satisfies Condition (C) if and only if there exists $(x, y) \in \mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_d, \frac{1}{\lambda_d}\right)\right)$ and a probability distribution $(p_1, ..., p_d) \in \left(\mathbb{R}_*^+\right)^d$ such that

$$\sum_{i=1}^d \lambda_i p_i = x,\ \sum_{i=1}^d \frac{1}{\lambda_i} p_i = y,\ \text{and}\ \ \forall k \in \{1, ..., d\},\ p_k > (xy - 1)\left(2 + \frac{x}{\lambda_k} + \lambda_k y\right)^{-1}. \tag{2.B.5}$$

For $(x, y) \in \mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_d, \frac{1}{\lambda_d}\right)\right)$, let us define

$$\begin{aligned}
M_0(x, y) &= (xy - 1)\sum_{i=1}^d \left(2 + \frac{x}{\lambda_i} + \lambda_i y\right)^{-1}, \\
M_{-1}(x, y) &= (xy - 1)\sum_{i=1}^d \frac{1}{\lambda_i}\left(2 + \frac{x}{\lambda_i} + \lambda_i y\right)^{-1}, \\
M_1(x, y) &= (xy - 1)\sum_{i=1}^d \lambda_i\left(2 + \frac{x}{\lambda_i} + \lambda_i y\right)^{-1},
\end{aligned}$$

and if $M_0(x, y) < 1$,

$$X(x, y) = \frac{x - M_1(x, y)}{1 - M_0(x, y)},\ Y(x, y) = \frac{y - M_{-1}(x, y)}{1 - M_0(x, y)}.$$

Proposition 2.B.4. *If $\lambda_1, ..., \lambda_d$ are not all equal, there exists a diagonal matrix that satisfies Condition (C) if and only if there exists $(x, y) \in \mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_d, \frac{1}{\lambda_d}\right)\right)$ such that*

$$M_0(x,y) < 1 \text{ and } (X(x,y), Y(x,y)) \in \mathcal{C}\left(\left(\lambda_i, \frac{1}{\lambda_i}\right)_{1\le i\le d}\right). \tag{2.B.6}$$

Proof. If a diagonal matrix $Diag(\alpha)$, where $\alpha := (\alpha_1, ..., \alpha_d) \in \left(\mathbb{R}_+^*\right)^d$, satisfies Condition (C), then (2.B.4) holds and it is easy to check that if we set $p_i = \frac{\frac{1}{\alpha_i}}{\sum_{k=1}^d \frac{1}{\alpha_k}}$ for $1 \le i \le d$, $x = \sum_{i=1}^d \lambda_i p_i$ and $y = \sum_{i=1}^d \frac{p_i}{\lambda_i}$, the conditions in (2.B.5) hold so the conditions in (2.B.6) are satisfied, as

$$\begin{pmatrix} X(x,y) \\ Y(x,y) \end{pmatrix} = \frac{1}{1 - M_0(x,y)} \sum_{k=1}^d \left(p_k - (xy-1)\left(2 + \frac{x}{\lambda_k} + \lambda_k y\right)^{-1}\right)\begin{pmatrix} \lambda_k \\ \frac{1}{\lambda_k} \end{pmatrix}.$$

Conversely, if $(x,y) \in \mathcal{C}\left(\left(\lambda_i, \frac{1}{\lambda_i}\right)_{1\le i\le d}\right)$ and satisfies (2.B.6), then there exists a probability distribution $(q_1, ..., q_d) \in (\mathbb{R}_*^+)^d$ such that $X(x,y) = \sum_{i=1}^d \lambda_i q_i$ and $Y(x,y) = \sum_{i=1}^d \frac{1}{\lambda_i} q_i$ and it is easy to check that if we set $p_i = (1 - M_0(x,y))\, q_i + (xy-1)\left(2 + \frac{x}{\lambda_i} + \lambda_i y\right)^{-1}$ for $1 \le i \le d$ then the conditions in (2.B.5) are satisfied. □

Given a discretization parameter n and the vector $(l_1, ..., l_d)$, which is the nondecreasing reordering of $(\lambda_1, ..., \lambda_d)$, the numerical procedure that follows builds a grid $\mathcal{G}$ that consists in $(d-1)(n-1)^2$ points of $\mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_d, \frac{1}{\lambda_d}\right)\right)$ and returns the list V of the points $(x,y) \in \mathcal{G}$ satisfying (2.B.6). The border of the convex $\mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_d, \frac{1}{\lambda_d}\right)\right)$ has a simple shape. Indeed, it is a polygon with vertices $\left(l_i, \frac{1}{l_i}\right)_{1\le i\le d}$ and edges $\left\{\left(l_i, \frac{1}{l_i}\right), \left(l_{i+1}, \frac{1}{l_{i+1}}\right)\right\}_{1\le i\le d}$, where we define $\left(l_{d+1}, \frac{1}{l_{d+1}}\right) = \left(l_1, \frac{1}{l_1}\right)$.

In Figure 1 below, we give an idea of the shape of this convex and we illustrate the output of the numerical procedure for $n = 200$, $d = 5$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $\lambda_4 = 5$ and $\lambda_5 = 10$. The border of the convex envelope $\mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_5, \frac{1}{\lambda_5}\right)\right)$ is colored in red, and the points (x,y) in $\mathcal{C}\left(\left(\lambda_1, \frac{1}{\lambda_1}\right), ..., \left(\lambda_5, \frac{1}{\lambda_5}\right)\right)$ satisfying (2.B.6) are colored in black. Condition (C) is thus satisfied in this situation.

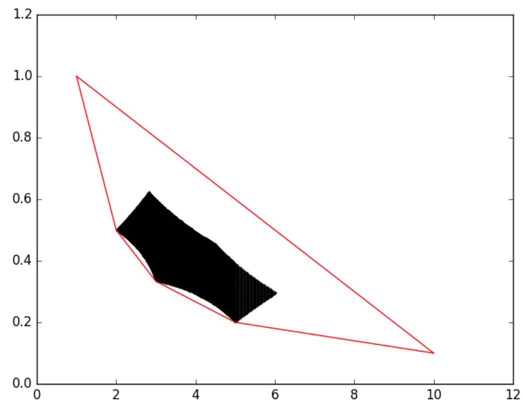


Figure 2.1: Condition (C) is satisfied for $d = 5$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $\lambda_4 = 5$ and $\lambda_5 = 10$, where $n = 200$.

The advantage of such a numerical procedure is that it operates in a bounded convex of $\mathbb{R}^2$ with a simple shape, instead of a bounded convex of $\mathbb{R}^d$, as Inequality (2.B.1) would suggest.

2.B.2 The case $d = 3$

In the following, we study the case $d = 3$. We recall that

$$r_1 = \frac{\lambda_3}{\lambda_2} + \frac{\lambda_2}{\lambda_3} \ge 2,\ r_2 = \frac{\lambda_3}{\lambda_1} + \frac{\lambda_1}{\lambda_3} \ge 2,\ r_3 = \frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} \ge 2.$$

```
V = []
for i = 1, ..., d − 1 do
    for k_1 = 1, ..., n − 1 do
        x = l_i + (l_{i+1} − l_i) k_1/n
        y_min = (1/l_i) (n−k_1)/n + (1/l_{i+1}) k_1/n
        y_max = 1/l_1 − (x−l_1)/(l_1 l_d)
        for k_2 = 1, ..., n − 1 do
            y = y_min + (y_max − y_min) k_2/n
            M_0 = (xy − 1) ∑_{k=1}^{d} (2 + x/l_i + l_i y)^{−1}
            if M_0 < 1 then
                M_1 = (xy − 1) ∑_{k=1}^{d} l_i (2 + x/l_i + l_i y)^{−1}
                X = (x−M_1)/(1−M_0)
                if l_1 < X < l_d then
                    j = Sum(X > l)
                    z_min = 1/l_j − (X−l_j)/(l_j l_{j+1})
                    z_max = 1/l_1 − (X−l_1)/(l_1 l_d)
                    M_{−1} = (xy − 1) ∑_{k=1}^{d} (1/l_i) (2 + x/l_i + l_i y)^{−1}
                    Y = (y−M_{−1})/(1−M_0)
                    if z_min < Y < z_max then
                        V = (x, y) :: V
                    end if
                end if
            end if
        end for
    end for
end for
return V
```

Let us first explicit the link between the values of r_1, r_2, r_3.

Lemma 2.B.5. *The values of r_1, r_2, r_3 are linked by*

$$r_3 \in \left\{ \frac{1}{2}\left(r_1 r_2 - \sqrt{(r_1^2-4)(r_2^2-4)}\right), \frac{1}{2}\left(r_1 r_2 + \sqrt{(r_1^2-4)(r_2^2-4)}\right) \right\}.$$

Proof. As $r_1 = \frac{\lambda_3}{\lambda_2} + \frac{\lambda_2}{\lambda_3} \geq 2$ and $r_2 = \frac{\lambda_1}{\lambda_3} + \frac{\lambda_3}{\lambda_1} \geq 2$, we have that $\frac{\lambda_2}{\lambda_3}, \frac{\lambda_3}{\lambda_2} \in \left\{ \frac{1}{2}\left(r_1 - \sqrt{r_1^2-4}\right), \frac{1}{2}\left(r_1 + \sqrt{r_1^2-4}\right) \right\}$ and $\frac{\lambda_1}{\lambda_3}, \frac{\lambda_3}{\lambda_1} \in \left\{ \frac{1}{2}\left(r_2 - \sqrt{r_2^2-4}\right), \frac{1}{2}\left(r_2 + \sqrt{r_2^2-4}\right) \right\}$. As $\frac{\lambda_1}{\lambda_2} = \frac{\lambda_1}{\lambda_3}\frac{\lambda_3}{\lambda_2}$, we deduce the two values that r_3 can take. $\square$

We now give the main result concerning the case $d = 3$.

Proposition 2.B.6. *There is equivalence between:*

(i) The inequality

$$\frac{1}{\sqrt{(r_1-2)(r_2-2)}} + \frac{1}{\sqrt{(r_2-2)(r_3-2)}} + \frac{1}{\sqrt{(r_1-2)(r_3-2)}} > \frac{1}{4}, \tag{2.B.7}$$

holds, with the convention $\frac{1}{0} = +\infty$.

(ii) Condition (C) is satisfied by a diagonal matrix.

(iii) Condition (C) is satisfied.

We now prove Proposition 2.B.6. To show that $(i) \Rightarrow (ii)$, we first study the case where there cardinality of $\{\lambda_1, \lambda_2, \lambda_3\}$ is smaller than 3, which is equivalent to $\min(r_1, r_2, r_3) = 2$ and implies (2.B.7).

Lemma 2.B.7. *If* $\min(r_1, r_2, r_3) = 2$*, then there exists a diagonal matrix that satisfies Condition* (C)*, i.e.* (ii) *holds.*

Proof. Let $\alpha_1, \alpha_2, \alpha_3 > 0$ and $p_i = \frac{\frac{1}{\alpha_i}}{\frac{1}{\alpha_1}+\frac{1}{\alpha_2}+\frac{1}{\alpha_3}}$, $i = 1, 2, 3$. Inequality (2.B.3) rewrites, for $k = 1$,

$$1 + 2p_1 + p_1^2 > \left(\frac{\lambda_2}{\lambda_3} + \frac{\lambda_3}{\lambda_2}\right) p_2 p_3 + p_2^2 + p_3^2.$$

Writing $p_1^2 = (1 - p_2 - p_3)^2$, and then using the fact that $p_1 + p_2 + p_3 = 1$, we obtain that

$$4 > p_2 p_3 \frac{(r_1 - 2)}{p_1}. \tag{2.B.8}$$

With similar computations for $k = 2, 3$ we moreover obtain

$$4 > p_1 p_3 \frac{(r_2 - 2)}{p_2} \text{ and } 4 > p_1 p_2 \frac{(r_3 - 2)}{p_3}. \tag{2.B.9}$$

To prove Lemma 2.B.7, it is sufficient to exhibit a probability distribution $(p_1, p_2, p_3) \in (\mathbb{R}_*^+)^3$ such that Inequalities (2.B.8)- (2.B.9) are satisfied. In the case where $r_1 = r_2 = 2$, we have that $\lambda_1 = \lambda_2 = \lambda_3$ so $r_3 = 2$, and the choice $p_1 = p_2 = p_3 = \frac{1}{3}$ satisfies (2.B.8)-(2.B.9). In the case where $r_1 = 2$, $r_2 > 2$ and $r_3 > 2$, we have that it is sufficient to choose $p_1 \in \left(0, \min\left(1, \frac{4}{r_2-2}, \frac{4}{r_3-2}\right)\right)$ and $p_2 = p_3 = \frac{1-p_1}{2}$ to satisfy (2.B.8)-(2.B.9). ☐

To complete the proof of $(i) \Rightarrow (ii)$, we show that in the case $\min(r_1, r_2, r_3) > 2$, if (2.B.7) holds, then Condition (C) is satisfied by a diagonal matrix.

Lemma 2.B.8. *Let us assume that* $\min(r_1, r_2, r_3) > 2$*. If Inequality (2.B.7) holds, then* (ii) *holds.*

Proof. Using the proof of Lemma 2.B.7, it is easy to check that that if Inequality (2.B.7) is satisfied then (2.B.8)-(2.B.9) are satisfied for the choice $p_i = \frac{\sqrt{r_i-2}}{\sqrt{r_1-2}+\sqrt{r_2-2}+\sqrt{r_3-2}} > 0$, for $i = 1, 2, 3$, so that the matrix $Diag\left(\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{p_3}\right)$ satisfies Condition (C). ☐

As the relation $(ii) \Rightarrow (iii)$ is trivial, we have obtained $(i) \Rightarrow (ii) \Rightarrow (iii)$. To prove $(iii) \Rightarrow (i)$, by Lemma 2.B.7, it is sufficient to show that in the case $\min(r_1, r_2, r_3) > 2$, if Condition (C) is satisfied then (2.B.7) holds. Let us remark that we can assume without loss of generality that $r_1 \leq r_2 \leq r_3$, so in what follows we suppose that

$$2 < r_1 \leq r_2 \leq r_3.$$

The next lemma deals with the case $\frac{r_3-2}{r_3+2} \geq \frac{r_1-2}{r_1+2} + \frac{r_2-2}{r_2+2}$.

Lemma 2.B.9. *Let us assume that* $\min(r_1, r_2, r_3) > 2$*. Then Inequality (2.B.7) is equivalent to*

$$\left\{\sqrt{(r_1-2)(r_2-2)} \leq 4\right\} \text{ or } \left\{\sqrt{(r_1-2)(r_2-2)} > 4 \text{ and } r_3 < 16\left(\frac{\sqrt{r_1-2}+\sqrt{r_2-2}}{\sqrt{(r_1-2)(r_2-2)}-4}\right)^2 + 2\right\}. \tag{2.B.10}$$

In particular, if moreover $\frac{r_3-2}{r_3+2} \geq \frac{r_1-2}{r_1+2} + \frac{r_2-2}{r_2+2}$*, then Inequality (2.B.7) holds.*

Proof. Under the assumption that $\min(r_1, r_2, r_3) > 2$, Inequality (2.B.7) rewrites $\frac{1}{\sqrt{r_3-2}}\left(\frac{1}{\sqrt{r_1-2}} + \frac{1}{\sqrt{r_2-2}}\right) > \frac{1}{4} - \frac{1}{\sqrt{(r_1-2)(r_2-2)}}$, so it is equivalent to (2.B.10). The term $\frac{r_1-2}{r_1+2} + \frac{r_2-2}{r_2+2}$ rewrites

$$\frac{r_1-2}{r_1+2} + \frac{r_2-2}{r_2+2} = \frac{2\sqrt{(r_1-2)(r_2-2)}\left(\sqrt{(r_1-2)(r_2-2)}-4\right) + 4\left(\sqrt{r_1-2}+\sqrt{r_2-2}\right)^2}{\left(\sqrt{(r_1-2)(r_2-2)}-4\right)^2 + 4\left(\sqrt{r_1-2}+\sqrt{r_2-2}\right)^2}.$$

If $\sqrt{(r_1-2)(r_2-2)} > 4$, then $\frac{r_1-2}{r_1+2} + \frac{r_2-2}{r_2+2} > 1$. Thus, since $1 > \frac{r_3-2}{r_3+2}$, if

$$\frac{r_3-2}{r_3+2} \geq \frac{r_1-2}{r_1+2} + \frac{r_2-2}{r_2+2},$$

then $\sqrt{(r_1-2)(r_2-2)} \leq 4$ and Inequality (2.B.7) holds. ☐

Let us now suppose that Condition (C) is satisfied by a matrix $\Gamma \in \mathcal{S}_3^{++}(\mathbb{R})$. We consider, for $k = 1, 2, 3$, the matrices $\Gamma^{(k)}$ with coefficients

$$\Gamma_{ij}^{(k)} = \frac{\lambda_i + \lambda_j}{2}\left(\Gamma_{ij} + \Gamma_{kk} - \Gamma_{ik} - \Gamma_{jk}\right),\ 1 \leq i, j \leq 3.$$

We define

$$v_1 := \Gamma_{22} + \Gamma_{33} - 2\Gamma_{23},\ v_2 := \Gamma_{11} + \Gamma_{33} - 2\Gamma_{13},\ v_3 := \Gamma_{11} + \Gamma_{22} - 2\Gamma_{12}.$$

The matrix $\Gamma^{(3)}$ rewrites

$$\Gamma^{(3)} = \begin{pmatrix} \lambda_1 v_2 & \frac{\lambda_1+\lambda_2}{2} \times \frac{v_1+v_2-v_3}{2} & 0 \\ \frac{\lambda_1+\lambda_2}{2} \times \frac{v_1+v_2-v_3}{2} & \lambda_2 v_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We deduce that $\Gamma^{(3)}$ is positive definite on $e_3^{\perp}$ if and only if

$$v_1 > 0, v_2 > 0 \tag{2.B.11}$$

and the determinant of the matrix $\left(\Gamma^{(3)}\right)_{1\leq i,j\leq 2}$ is positive, which rewrites $16v_1v_2 > \left(\sqrt{\frac{\lambda_2}{\lambda_1}} + \sqrt{\frac{\lambda_1}{\lambda_2}}\right)^2 (v_1 + v_2 - v_3)^2 = (r_3 + 2)(v_1 + v_2 - v_3)^2$ and therefore

$$-4\frac{r_3 - 2}{r_3 + 2} v_1 v_2 > v_1^2 + v_2^2 + v_3^2 - 2(v_1v_2 + v_2v_3 + v_1v_3), \tag{2.B.12}$$

With similar computations for $\Gamma^{(1)}$ and $\Gamma^{(2)}$, we obtain the additional inequalities

$$\begin{aligned} v_3 &> 0, && (2.B.13) \\ -4\frac{r_2-2}{r_2+2}v_1v_3 &> v_1^2 + v_2^2 + v_3^2 - 2(v_1v_2 + v_2v_3 + v_1v_3), && (2.B.14) \\ -4\frac{r_1-2}{r_1+2}v_2v_3 &> v_1^2 + v_2^2 + v_3^2 - 2(v_1v_2 + v_2v_3 + v_1v_3). && (2.B.15) \end{aligned}$$

If Condition (C) is satisfied, then there exists v_1, v_2, v_3 satisfying (2.B.11)-(2.B.15). Let us define for $i = 1, 2, 3$, $\gamma_i = 4\frac{r_i-2}{r_i+2} \in (0, 4)$. If we assume that moreover, $\gamma_3 v_1 v_2 = \gamma_1 v_2 v_3 = \gamma_2 v_1 v_3$, that is $\frac{\gamma_3}{v_3} = \frac{\gamma_1}{v_1} = \frac{\gamma_2}{v_2}$, then we have that

$$0 > \gamma_1^2 + \gamma_2^2 + \gamma_3^2 - 2(\gamma_1\gamma_2 + \gamma_2\gamma_3 + \gamma_1\gamma_3) + \gamma_1\gamma_2\gamma_3. \tag{2.B.16}$$

In the case $\frac{r_3-2}{r_3+2} \geq \frac{r_2-2}{r_2+2} + \frac{r_1-2}{r_1+2}$, by Lemma 2.B.9, Inequality (2.B.7) holds. We show in Lemma 2.B.11 below that under the assumption $\frac{r_3-2}{r_3+2} < \frac{r_2-2}{r_2+2} + \frac{r_1-2}{r_1+2}$, Condition (C) implies Inequality (2.B.16). To conclude the proof of $(iii) \Rightarrow (i)$ and therefore the proof of Proposition 2.B.6, we show in Lemma 2.B.10 below that (2.B.16) implies (2.B.7).

Lemma 2.B.10. *Let us assume that $2 < r_1 \leq r_2 \leq r_3$, then Inequality (2.B.16) implies Inequality (2.B.7).*

Proof. We see the term on the r.h.s. of Inequality (2.B.16) as a second degree polynomial in the variable γ_3, which has two distinct roots $z_- < z_+$. Indeed, $\gamma_1, \gamma_2 \in (0, 4)$, and the discriminant of the polynomial is $\gamma_1\gamma_2(4 - \gamma_1)(4 - \gamma_2) > 0$. As $\gamma_3 = 4\frac{r_3-2}{r_3+2}$, Inequality (2.B.16) is equivalent to $r_3 \in \left(\frac{8+2z_-}{4-z_-}, \frac{8+2z_+}{4-z_+}\right)$, where

$$\frac{8 + 2z_\pm}{4 - z_\pm} = 16\left(\frac{\sqrt{r_1 - 2} \pm \sqrt{r_2 - 2}}{\sqrt{(r_1 - 2)(r_2 - 2)} \mp 4}\right)^2 + 2.$$

By Lemma 2.B.9, we conclude that Inequality (2.B.16) implies Inequality (2.B.7). □

Lemma 2.B.11. *Let us assume that $2 < r_1 \leq r_2 \leq r_3$, and that $\frac{r_3-2}{r_3+2} < \frac{r_2-2}{r_2+2} + \frac{r_1-2}{r_1+2}$. If Condition (C) is satisfied then Inequality (2.B.16) holds.*

Proof. Let us define the function $f : \left(\mathbb{R}_+^*\right)^3 \to \mathbb{R}$ by $f(a_1, a_2, a_3) = 2(a_1 + a_2 + a_3) - \frac{a_1 a_2}{a_3} - \frac{a_2 a_3}{a_1} - \frac{a_1 a_3}{a_2}$. Reformulating Inequalities (2.B.12)-(2.B.15) with the change of variables $a_i = \frac{1}{v_i}, i = 1, 2, 3$, we obtain that

$$f(a_1, a_2, a_3) > \max\{\gamma_1 a_1, \gamma_2 a_2, \gamma_3 a_3\}.$$

Under the assumption that $2 < r_1 \leq r_2 \leq r_3$, we have that $0 < \gamma_1 \leq \gamma_2 \leq \gamma_3$, and let us remark that

$$\max\{\gamma_1 a_1, \gamma_2 a_2, \gamma_3 a_3\} \geq \max\{\gamma_3 a_{(1)}, \gamma_2 a_{(2)}, \gamma_1 a_{(3)}\},$$

where $\left(a_{(1)}, a_{(2)}, a_{(3)}\right)$ is the nondecreasing reordering of (a_1, a_2, a_3). Let $\mathcal{R}$ be the set of elements $(a_1, a_2, a_3) \in \left(\mathbb{R}_+^*\right)^3$ such that $a_1 \leq a_2 \leq a_3$ and

$$f(a_1, a_2, a_3) > \max\{\gamma_1 a_3, \gamma_2 a_2, \gamma_3 a_1\}. \tag{2.B.17}$$

If Condition (C) is satisfied then $\mathcal{R}$ is not empty. Let us remark first that as both sides of (2.B.17) are homogeneous of order 1, $\mathcal{R}$ is stable by scaling: if $(a_1, a_2, a_3) \in \mathcal{R}$ then for $\zeta > 0$, $(\zeta a_1, \zeta a_2, \zeta a_3) \in \mathcal{R}$. Moreover, if we assume that there exists $(a_1, a_2, a_3) \in \mathcal{R}$ such that $\gamma_1 a_3 = \gamma_2 a_2 = \gamma_3 a_1 =: \Delta$, then we can check that Inequality (2.B.16) is satisfied. Indeed, we have that $a_3 = \frac{\Delta}{\gamma_1}, a_2 = \frac{\Delta}{\gamma_2}, a_3 = \frac{\Delta}{\gamma_1}$, and

$$f\left(\frac{\Delta}{\gamma_3}, \frac{\Delta}{\gamma_2}, \frac{\Delta}{\gamma_1}\right) > \Delta. \tag{2.B.18}$$

Multiplying both sides of (2.B.18) by $\frac{\gamma_1\gamma_2\gamma_3}{\Delta}$, we obtain (2.B.16).

Let $(q_1, q_2, q_3) \in \mathcal{R}$. To prove that there exists $(u_1, u_2, u_3) \in \mathcal{R}$, such that $\gamma_1 u_3 = \gamma_2 u_2 = \gamma_3 u_1$, and conclude with the previous argument, we construct a path included into $\mathcal{R}$ that goes from (q_1, q_2, q_3) to (u_1, u_2, u_3). We now distinguish the three cases.

Case 1: $\gamma_3 q_1 \leq \max(\gamma_2 q_2, \gamma_1 q_3) = \gamma_1 q_3$. Let us remark that if $a_1 \leq a_2 \leq a_3$, the partial derivative $\partial_{a_1} f$ satisfies

$$\partial_{a_1} f(a_1, a_2, a_3) = \frac{a_2 a_3}{a_1^2} - \left(\sqrt{\frac{a_2}{a_3}} - \sqrt{\frac{a_3}{a_2}}\right)^2 \geq \frac{a_3}{a_2}\left(\left(\frac{a_2}{a_1}\right)^2 - 1\right) \geq 0. \tag{2.B.19}$$

For $x \in \left[q_1, \min\left(q_2, \frac{\gamma_1}{\gamma_3} q_3\right)\right]$, $f(x, q_2, q_3) \geq f(q_1, q_2, q_3)$. If $\min\left(q_2, \frac{\gamma_1}{\gamma_3} q_3\right) = \frac{\gamma_1}{\gamma_3} q_3$, then for $\tilde{q}_1 = \frac{\gamma_1}{\gamma_3} q_3$, we have that

$$f(\tilde{q}_1, q_2, q_3) \geq f(q_1, q_2, q_3) > \gamma_1 q_3 = \gamma_3 \tilde{q}_1 \geq \gamma_2 q_2. \tag{2.B.20}$$

We scale (2.B.20) and define $\zeta = \frac{\gamma_1}{\tilde{q}_1} = \frac{\gamma_3}{q_3} > 0$ so that $\zeta \tilde{q}_1 = \gamma_1$ and $\zeta q_3 = \gamma_3$. We have that

$$f(\zeta \tilde{q}_1, \zeta q_2, \zeta q_3) \geq f(\zeta q_1, \zeta q_2, \zeta q_3) > \gamma_3 \gamma_1 \geq \gamma_2 \zeta q_2, \tag{2.B.21}$$

We now increase q_2 in (2.B.21). For $a_1, a_2, a_3 > 0$, such that $a_1 \leq a_2 \leq a_3$ and $a_2 \leq \frac{a_3 a_1}{a_3 - a_1}$ we have that

$$\partial_{a_2} f(a_1, a_2, a_3) = a_1 a_3 \left(\frac{1}{a_2^2} - \left(\frac{1}{a_1} - \frac{1}{a_3}\right)^2\right) \geq 0.$$

By hypothesis,

$$q_2 \leq \frac{\gamma_1 \gamma_3}{\zeta \gamma_2} \leq \frac{1}{\zeta} \frac{\gamma_1 \gamma_3}{\gamma_3 - \gamma_1} = \frac{q_3 \tilde{q}_1}{q_3 - \tilde{q}_1},$$

so for $z \in [q_2, \frac{\gamma_1 \gamma_3}{\zeta \gamma_2}]$, $f(\zeta \tilde{q}_1, \zeta z, \zeta q_3) \geq f(\zeta \tilde{q}_1, \zeta q_2, \zeta q_3)$ and in particular for $\hat{q}_2 = \frac{\gamma_1 \gamma_3}{\zeta \gamma_2}$, we have that

$$f(\zeta \tilde{q}_1, \zeta \hat{q}_2, \zeta q_3) > \gamma_1 \zeta q_3 = \gamma_3 \zeta \tilde{q}_1 = \gamma_2 \zeta \hat{q}_2,$$

so we obtain (2.B.16). If $\min\left(q_2, \frac{\gamma_1}{\gamma_3} q_3\right) = q_2$, as the function $x \to f(x, q_2, q_3)$ is nondecreasing for $x \in [q_1, q_2]$, we have that

$$f(q_2, q_2, q_3) \geq f(q_1, q_2, q_3) > \gamma_1 q_3 \geq \gamma_3 q_2 \geq \gamma_2 q_2.$$

As the function $g : \mathbb{R}^2 \to \mathbb{R}$ defined by $g(a_2, a_3) = f(a_2, a_2, a_3)$ satisfies $\partial_{a_2} g(a_2, a_3) = 4 - 2\frac{a_2}{a_3} \geq 0$ if $a_2 \leq a_3$, we have that for $y \in \left[q_2, \frac{\gamma_1 q_3}{\gamma_3}\right]$, $f(y, y, q_3) \geq f(q_1, q_2, q_3)$. In particular for $\tilde{q}_1 = \tilde{q}_2 = \frac{\gamma_1 q_3}{\gamma_3}$, we have

$$f(\tilde{q}_1, \tilde{q}_2, q_3) \geq f(q_1, q_2, q_3) > \gamma_1 q_3 = \gamma_3 \tilde{q}_1 \geq \gamma_2 \tilde{q}_2. \tag{2.B.22}$$

and (2.B.22) is treated in the same way as (2.B.20).

Case 2: $\gamma_3 q_1 \leq \max(\gamma_2 q_2, \gamma_1 q_3) = \gamma_2 q_2$. Using (2.B.19), as $\partial_{a_1} f(x, q_2, q_3) \geq 0$ for $x \in [q_1, q_2]$, we have that for $\tilde{q}_1 = \frac{\gamma_2}{\gamma_3} q_2 \in [q_1, q_2]$,

$$f(\tilde{q}_1, q_2, q_3) \geq f(q_1, q_2, q_3) > \gamma_2 q_2 = \gamma_3 \tilde{q}_1 \geq \gamma_1 q_3.$$

For $\theta = \frac{\gamma_3}{q_2} = \frac{\gamma_2}{\tilde{q}_1}$, we have that

$$f(\theta\tilde{q}_1, \theta q_2, \theta q_3) > \gamma_3\gamma_2 \geq \gamma_1 \theta q_3.$$

As $\gamma_3 - \gamma_2 < \gamma_1$, we have that

$$q_3 \leq \frac{\gamma_2\gamma_3}{\theta\gamma_1} < \frac{1}{\theta}\frac{\gamma_2\gamma_3}{\gamma_3 - \gamma_2} = \frac{\tilde{q}_1 q_2}{q_2 - \tilde{q}_1},$$

Moreover, for $0 < a_1 \leq a_2 \leq a_3$ such that $a_3 \leq \frac{a_2 a_1}{a_2 - a_1}$,

$$\partial_{a_3} f(a_1, a_2, a_3) = a_1 a_2 \left(\frac{1}{a_3^2} - \left(\frac{1}{a_1} - \frac{1}{a_2} \right)^2 \right) \geq 0, \tag{2.B.23}$$

so for $\hat{q}_3 = \frac{\gamma_2\gamma_3}{\theta\gamma_1}$, we obtain

$$f(\theta\tilde{q}_1, \theta q_2, \theta\hat{q}_3) \geq f(\theta\tilde{q}_1, \theta q_2, \theta q_3) > \gamma_2\theta q_2 = \gamma_3\theta\tilde{q}_1 = \gamma_1\theta\hat{q}_3,$$

and we deduce (2.B.16).

Case 3: $\gamma_3 q_1 > \max(\gamma_2 q_2, \gamma_1 q_3)$. We show that there exists $\tilde{q}_3 \geq q_3$ such that

$$f(q_1, q_2, \tilde{q}_3) \geq f(q_1, q_2, q_3) > \gamma_3 q_1 = \gamma_1 \tilde{q}_3 \geq \gamma_2 q_2. \tag{2.B.24}$$

Let $\eta = \frac{\gamma_1}{q_1}$ so that $\eta q_1 = \gamma_1$. We have that

$$f(\eta q_1, \eta q_2, \eta q_3) > \gamma_3\gamma_1 > \max(\gamma_1\eta q_3, \gamma_2\eta q_2).$$

As $q_2 < \frac{\gamma_1\gamma_3}{\eta\gamma_2}$ and $\gamma_3 - \gamma_2 < \gamma_1$, we have that $\frac{\gamma_1}{\gamma_3 - \gamma_2} > 1$ and

$$q_3 \leq \frac{\gamma_3}{\eta} < \frac{1}{\eta}\frac{\gamma_1\gamma_3}{\gamma_3 - \gamma_2} = \frac{1}{\eta}\frac{\gamma_1\frac{\gamma_3\gamma_1}{\gamma_2}}{\frac{\gamma_3\gamma_1}{\gamma_2} - \gamma_1} \leq \frac{q_1 q_2}{q_2 - q_1}.$$

Using (2.B.23), we deduce that for $y \in \left[q_3, \frac{\gamma_3}{\eta}\right]$, the function $y \to f(\eta q_1, \eta q_2, \eta y)$ is non decreasing and for $\tilde{q}_3 = \frac{\gamma_3}{\eta}$, we obtain

$$f(\eta q_1, \eta q_2, \eta\tilde{q}_3) \geq f(\eta q_1, \eta q_2, \eta q_3) > \gamma_3\eta q_1 = \gamma_1\eta\tilde{q}_3 \geq \gamma_2\eta q_2,$$

which is equivalent to (2.B.24) and we conclude in the same manner as for (2.B.20) in Case 1. $\square$

2.C Additional proofs of Section 4

2.C.1 Proof of Lemma 2.4.1

Proof. Let $\nu \in \mathcal{P}(\mathbb{R})$ and let u be a solution to $LV(\nu)$. As $\tilde{\sigma}_{Dup} \in L^\infty([0,T], W^{1,\infty}(\mathbb{R}))$ and $u \in L^2_{loc}((0,T]; H^1(\mathbb{R}))$, we have that dt-a.e. on $(0,T]$, $\left(\tilde{\sigma}^2_{Dup} u\right)(t,\cdot) \in H^1(\mathbb{R})$ and $\tilde{\sigma}^2_{Dup}(t,\cdot)\partial_x u(t,\cdot) = \partial_x\left(\tilde{\sigma}^2_{Dup} u\right)(t,\cdot) - \left(u\partial_x\tilde{\sigma}^2_{Dup}\right)(t,\cdot)$ in the sense of distributions on $\mathbb{R}$. Then for any function ϕ defined for $(t,x) \in [0,T]\times\mathbb{R}$ by $\phi(t,x) := g_1(t)g_2(x)$, with $g_1 \in C_c^\infty((0,T))$ and $g_2 \in C_c^\infty(\mathbb{R})$, the Borel measure $dm := u dx dt$ satisfies the equality:

$$\int_{(0,T)\times\mathbb{R}} \left[\partial_t\phi + \frac{1}{2}\tilde{\sigma}^2_{Dup}\partial^2_{xx}\phi + \partial_x\left(\frac{1}{2}\tilde{\sigma}^2_{Dup}\right)\partial_x\phi + \left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup} - \partial_x\left(\frac{1}{2}\tilde{\sigma}^2_{Dup}\right)\right)\partial_x\phi \right] dm = 0.$$

By density of the space spanned by the functions of type $g_1(t)g_2(x)$ in $C_c^\infty((0,T)\times\mathbb{R})$ for the norm $\phi \in C_c^\infty((0,T)\times\mathbb{R}) \to ||\phi||_\infty + ||\partial_x\phi||_\infty + ||\partial^2_{xx}\phi||_\infty + ||\partial_t\phi||_\infty$, the previous equality is also satisfied for any function $\phi \in C_c^\infty((0,T)\times\mathbb{R})$. The variational formulation of the PDE (2.4.7) as defined in [13, equality 1.5] is then satisfied.

We recall that $h_1(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ and it is easy to check that $h_1 \in H^1(\mathbb{R})$. As $(u(t), h_1)_1 \underset{t\to 0}{\to} \int h_1 d\nu > 0$ and as u is non negative, we obtain that for any $\tau \in (0,T)$, $\int_0^\tau (u(t), h_1)_1 dt > 0$, so $\underset{(0,\tau)\times\mathbb{R}}{ess\sup}\, u(t,x) > 0$. In addition, with Assumption (B), the functions $\frac{1}{2}\tilde{\sigma}^2_{Dup}$, $\partial_x\left(\frac{1}{2}\tilde{\sigma}^2_{Dup}\right)$ and $\left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup} - \partial_x\left(\frac{1}{2}\tilde{\sigma}^2_{Dup}\right)\right)$ are uniformly bounded, and $\underline{\sigma}^2 \leq \tilde{\sigma}^2_{Dup}$ a.e. on $[0,T]\times\mathbb{R}$, so by [13, Corollary 3.1] we obtain that u is continuous and positive on $(0,T]\times\mathbb{R}$. $\square$

2.C.2 Proof of Proposition 2.4.2

Let us first remark that if the initial condition ν has a density $u_0 \in L^2(\mathbb{R})$, then by energy estimates, uniqueness holds without Assumption (H) for a slightly stronger variational formulation called $LV^*(u_0)$ which is

$$u \in L^2([0,T]; H^1(\mathbb{R})) \cap L^\infty([0,T]; L^2(\mathbb{R})), u \geq 0$$

$$\forall v \in H^1(\mathbb{R}),\ 0 \quad = \quad \frac{d}{dt}(v,u)_1 - \left(\partial_x v, \left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup} - \partial_x\left(\frac{1}{2}\tilde{\sigma}^2_{Dup}\right)\right) u\right)_1 + \frac{1}{2}\left(\partial_x v, \tilde{\sigma}^2_{Dup}\partial_x u\right)_1,$$

in the sense of distributions on $(0,T)$, and $u(0,\cdot) = u_0$.

We now prove Proposition 2.4.2.

Proof. The main ingredient to obtain uniqueness to $LV(\nu)$ is [39, Proposition 4.2]. In the proof of Theorem 2.4.3, we obtained existence of a solution to $LV(\nu)$. Moreover, if u is a solution to $LV(\nu)$, then with the same arguments as in the proof of Lemma 2.4.1, we show that the Borel measure $udtdx$ solves the PDE (2.4.7) with initial condition ν in the sense of distributions, which means that for any $\phi \in C_c^\infty(\mathbb{R})$, the equality

$$\frac{d}{dt}\int_{\mathbb{R}} \phi(x)u(t,x)dx \quad = \quad \int_{\mathbb{R}} \left[\left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup}(t,x)\right)\phi'(x) + \frac{1}{2}\tilde{\sigma}^2_{Dup}(t,x)\phi''(x)\right] u(t,x)dx, \qquad (2.C.1)$$

holds in the sense of distributions on $(0,T)$, and u converges to ν in duality with $C_c^\infty(\mathbb{R})$ as $t \to 0$. Under Assumptions (B) and (H), by [39, Proposition 4.2], the measure $udtdx$ is the unique solution to the PDE (2.4.7) with initial condition ν in the sense of distributions and therefore u is the unique solution to $LV(\nu)$.

It is then sufficient to exhibit a solution $udtdx$ to the PDE (2.4.7) in the distributional sense with initial condition ν and such that $\int u^2(t,x)dx \leq \frac{\zeta}{\sqrt{t}}$ for a.e. $t \in (0,T)$, where $\zeta > 0$ is a constant that does not depend on ν. Under Assumptions (B) and (H), the martingale problem associated to the SDE (2.4.4) is well posed by [92, Theorems 6.1.7 and 7.2.1], and as mentioned in [82, Paragraph 4.1], there exists two constants $c := c(\underline{\sigma},\chi)$ and $C := C(T,\underline{\sigma},H_0,\chi)$ such that for $y \in \mathbb{R}$, the solution $(X_t^y)_{t\geq 0}$ to the SDE (2.4.4) with initial distribution δ_y has the density $p^y(t,x)$ that satisfies $\forall (t,x) \in (0,T] \times \mathbb{R}$, $p^y(t,x) \leq Cu_c(t,x,y)$, with $u_c(t,x,y) := \sqrt{\frac{c}{2\pi t}}\exp\left(-c\frac{(x-y)^2}{2t}\right)$ for $t \in (0,T], x,y \in \mathbb{R}$, and the function $y \to p^y(t,x)dx$ is measurable. For $\phi \in C_c^\infty(\mathbb{R})$, by Itô's Lemma,

$$\mathbb{E}\left[\phi(X_t^y)\right] \quad = \quad \phi(y) + \int_0^t \mathbb{E}\left[\phi'(X_s^y)\left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup}(s,X_s^y)\right) + \frac{1}{2}\phi''(X_s^y)\tilde{\sigma}^2_{Dup}(s,X_s^y)\right] ds.$$

In the sense of distributions, we have

$$\frac{d}{dt}\int_{\mathbb{R}} \phi(x)p^y(t,x)dx = \int_{\mathbb{R}} \left[\left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup}(t,x)\right)\phi'(x) + \frac{1}{2}\tilde{\sigma}^2_{Dup}(t,x)\phi''(x)\right] p^y(t,x)dx,$$

and $\lim_{t\to 0}\int_{\mathbb{R}} \phi(x)p^y(t,x)dx = \phi(y)$. We set $u := \int_{\mathbb{R}} p^y\nu(dy)$ and we check that $udtdx$ solves the PDE (2.4.7) with initial condition ν in the distributional sense. Indeed, for $\psi \in C_c^\infty((0,T),\mathbb{R})$, we check, using Fubini's theorem, that

$$-\int_0^T \psi'(t)\int_{\mathbb{R}} \phi(x)u(t,x)dxdt \quad = \quad \int_0^T \psi(t)\int_{\mathbb{R}} \left[\left(r - \frac{1}{2}\tilde{\sigma}^2_{Dup}(t,x)\right)\phi'(x) + \frac{1}{2}\tilde{\sigma}^2_{Dup}(t,x)\phi''(x)\right] u(t,x)dxdt,$$

and using Lebesgue's theorem, that $\lim_{t\to 0}\int_{\mathbb{R}^2} \phi(x)p^y(t,x)dx\nu(dy) = \int_{\mathbb{R}}\left(\lim_{t\to 0}\int_{\mathbb{R}} \phi(x)p^y(t,x)dx\right)\nu(dy) = \int_{\mathbb{R}} \phi(y)\nu(dy)$, and we conclude that u is the unique solution to $LV(\nu)$ and that moreover u coincides with the time marginals of the solution to the SDE (2.4.4) with initial distribution ν. Finally we define $\zeta := \frac{C^2\sqrt{c}}{2\sqrt{\pi}}$ and for a.e. $t \in (0,T]$, by Jensen's inequality, $\int_{\mathbb{R}} u^2(t,x)dx \leq \int_{\mathbb{R}^2}\left(p^y(t,x)\right)^2\nu(dy)dx \leq \frac{cC^2}{2\pi t}\int_{\mathbb{R}}\exp\left(-c\frac{x^2}{t}\right)dx \leq \frac{\zeta}{\sqrt{t}}$. ☐

2.D Proof of Theorem 2.5.11

Let us define $S := \mathbb{R} \times \mathcal{Y}$, $\bar{b} := \max_{i\in\{1,...,d\}} ||b_i||_\infty$, $\bar{a} := \max_{i\in\{1,...,d\}} ||a_i||_\infty$ and $\bar{q} := \max_{i,j\in\{1,...,d\}} ||q_{ij}||_\infty$. For $(x, y) \in \mathcal{S}$ we denote by $\delta_{x,y}$ the Dirac distribution on $\{(x, y)\}$. Lemma 2.D.1 below is a consequence of the constant expectation of a martingale combined with Fubini's theorem.

Lemma 2.D.1. *Let $\{\nu_{x,y}\}_{(x,y)\in\mathcal{S}}$ be a measurable family of probability measures on E such that for $\mu_0 - a.e.$ (x, y), $\nu_{x,y}$ is a martingale solution to the SDE (2.5.22) with initial distribution $\delta_{x,y}$. Define the measure $\mu_i(t, \cdot)$, for $1 \le i \le d$ and $t \in (0, T]$ by*

$$\int_{\mathbb{R}} \psi(x)\mu_i(t, dx) := \int_{E\times\mathcal{S}} \psi(X_t)1_{\{Y_t=i\}} d\nu_{x,y}(X, Y)\mu_0(dx, dy)$$

for any function $\psi \in C_b^2(\mathbb{R})$. Then $(\mu_1(t, \cdot), ..., \mu_{d,t}(t, \cdot))_{t\in(0,T]}$ is a solution to the PDS (2.5.19)-(2.5.20).

We now prove Theorem 2.5.11.

Step 1: We first establish the result for a regularized version of the PDS (2.5.19)-(2.5.20). Let ρ_X and ρ_T be convolution kernels defined by $\rho_X(x) = \rho_T(x) = C_0 e^{-\sqrt{1+x^2}}$, for $x \in \mathbb{R}$, and where $C_0 = \left(\int_{\mathbb{R}^2} e^{-\sqrt{1+x^2}} dx\right)^{-1}$. For $\epsilon > 0$ and $(t, x) \in \mathbb{R}^2$, we define the functions $\rho_X^\epsilon(x) := \frac{1}{\epsilon}\rho_X(\frac{x}{\epsilon})$, $\rho_T^\epsilon := \frac{1}{\epsilon}\rho_T(\frac{t}{\epsilon})$ and $\rho^\epsilon(t, x) := \rho_T^\epsilon(t)\rho_X^\epsilon(x)$. For $1 \le i \le d$, we extend the definition of $t \to \mu_i(t, \cdot)$ to $\mathbb{R}$, by setting for $t \le 0$, $\mu_i(t, \cdot) = \mu_0(\cdot, \{i\})$ and for $t > T$, $\mu_i(t, \cdot) = \mu_i(T, \cdot)$. As a consequence, given $\psi \in C_b^2(\mathbb{R})$ and $1 \le i \le d$, the extended function $t \to \int_{\mathbb{R}} \psi(x)\mu_i(t, dx)$ is now continuous and bounded on $\mathbb{R}$. For $1 \le i \le d$, let us also extend the definitions of the functions $(a_i)_{1\le i\le d}$ on $\mathbb{R}^2$ by setting $a_i(t, \cdot) = 0$ if $t \notin [0, T]$. In the same way, we extend the functions $(b_i)_{1\le i\le d}$ and $(q_{ij})_{1\le i,j\le d}$. Then the family $(\mu_1(t, \cdot), ..., \mu_{d,t}(t, \cdot))_{t\in\mathbb{R}}$ satisfies the equality (2.5.21) in the distributional sense on $\mathbb{R}$. We define for $1 \le i \le d$ and $(t, x) \in \mathbb{R}^2$,

$$\mu_i^\epsilon(t, x) := \mu_i * \rho^\epsilon(t, x) = \int_{\mathbb{R}^2} \rho^\epsilon(t - s, x - y)\mu_i(s, dy)ds.$$

We define $a_i^\epsilon = \frac{(a_i\mu_i)*\rho^\epsilon}{\mu_i^\epsilon}$, $b_i^\epsilon = \frac{(b_i\mu_i)*\rho^\epsilon}{\mu_i^\epsilon}$, $q_{ij}^\epsilon = \frac{(q_{ij}\mu_i)*\rho^\epsilon}{\mu_i^\epsilon}$ for $1 \le i, j \le d$. The family $(\mu_1^\epsilon(t, \cdot), ..., \mu_d^\epsilon(t, \cdot))_{t\in\mathbb{R}}$ is a smooth solution of the following PDS, denoted by $(PDS)_\epsilon$, where for $1 \le i \le d$,

$$\begin{aligned} \partial_t\mu_i^\epsilon + \partial_x(b_i^\epsilon\mu_i^\epsilon) - \frac{1}{2}\partial_{xx}^2(a_i^\epsilon\mu_i^\epsilon) - \sum_{j=1}^{d} q_{ji}^\epsilon\mu_i^\epsilon &= 0 \\ \mu_i^\epsilon(0) &= \mu_i * \rho^\epsilon(0, \cdot). \end{aligned} \tag{2.D.1}$$

Without loss of generality, we suppose that for $1 \le i \le d$, μ_i^ϵ has a positive density on $\mathbb{R}^2$. If not, then $\mu_i(t, \mathbb{R})$ is equal to zero for all $t \in (0, T]$. In that case, it is sufficient to consider the PDS (2.5.19)-(2.5.20) without the state i. Under this assumption, the functions $a_i^\epsilon, b_i^\epsilon, q_{ij}^\epsilon$ for $1 \le i, j \le d$ are well defined and for $i, j \in \{1, ..., d\}$, $||a_i^\epsilon||_\infty \le ||a_i||_\infty$, $||b_i^\epsilon||_\infty \le ||b_i||_\infty$, $||q_{ji}^\epsilon||_\infty \le ||q_{ji}||_\infty$. It is easy to check that for $x \in \mathbb{R}$, and $k \ge 1$, there exists constants $C_{X,k} > 0$ s.t. $\left|\frac{\partial^k \rho_X}{\partial x^k}(x)\right| \le C_{X,k}|\rho_X(x)|$, so $a_i^\epsilon, b_i^\epsilon, q_{ij}^\epsilon$ are continuous and bounded on $\mathbb{R}^2$, as well as their derivatives. Let us denote by $(SDE)_\epsilon$, the SDE

$$dX_t^\epsilon = b_{Y_t^\epsilon}^\epsilon(t, X_t^\epsilon)dt + \sqrt{a_{Y_t^\epsilon}^\epsilon(t, X_t^\epsilon)}dW_t,$$

where Y_t^ϵ is a stochastic process with values in $\mathcal{Y}$, and that satisfies

$$\mathbb{P}\left(Y_{t+dt}^\epsilon = j | (X_s^\epsilon, Y_s^\epsilon)_{0\le s\le t}\right) = q_{Y_t^\epsilon j}^\epsilon(t, X_t^\epsilon)dt,$$

for $j \ne Y_t^\epsilon$. As the functions $(a_i^\epsilon)_{1\le i\le d}, (b_i^\epsilon)_{1\le i\le d}, (q_{ij}^\epsilon)_{1\le i,j\le d}$ are continuous, Lipschitz and bounded, by [81, Theorem 5.3] and the Kunita-Watanabe theorem, for any $(x, y) \in \mathcal{S}$ there exists a unique martingale solution $\nu_{x,y}$ to $(SDE)_\epsilon$ with initial distribution $\delta_{(x,y)}$ and by [81, Proposition 5.52], the function $(x, y) \to \nu_{x,y}^\epsilon$ is measurable. We define $\nu^\epsilon := \sum_{i=1}^d \int_{\mathbb{R}} \nu_{x,i}^\epsilon (\mu_i * \rho^\epsilon)(0, x)dx$. For $t \in (0, T]$, $1 \le i \le d$, we define the measure $\tilde{\mu}_i^\epsilon(t, \cdot)$ by

$$\int_{\mathbb{R}} \psi(x)\tilde{\mu}_i^\epsilon(t, dx) = \int_E \psi(X_t)1_{\{Y_t=i\}} d\nu^\epsilon(X, Y),$$

for $\psi \in C_b^2(\mathbb{R})$. By Lemma 2.D.1, $(\tilde{\mu}_i^\epsilon)_{1\le i\le d}$ solves $(PDS)_\epsilon$ with initial condition $(\mu_i * \rho^\epsilon(0,\cdot))_{1\le i\le d}$. Since $(PDS)_\epsilon$ has a unique solution by Proposition 2.D.3 below, we obtain that for $t \in (0,T]$, $1 \le i \le d$, $\tilde{\mu}_i^\epsilon(t,\cdot) = \mu_i^\epsilon(t,\cdot)$, and for $\psi \in C_b^2(\mathbb{R})$,

$$\int_{\mathbb{R}} \psi(x)\mu_i^\epsilon(t,dx) = \int_E \psi(X_t)1_{\{Y_t=i\}} d\nu^\epsilon(X,Y). \tag{2.D.2}$$

Step 2: Let $(\epsilon_n)_{n\ge 0}$ be a positive sequence decreasing to 0, we check the family of measures $(\nu^{\epsilon_n})_{n\ge0}$ has a converging subsequence. For $n \ge 0$, as we might not have $\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^\epsilon(0,x)dx = 1$, we define $\overline{\nu}^{\epsilon_n} := \frac{\nu^{\epsilon_n}}{\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,x)dx}$, so that $\overline{\nu}^{\epsilon_n}$ is a probability measure. We use Aldous' criterion to show the tightness of the family $(\overline{\nu}^{\epsilon_n})_{n\ge0}$. Let us denote by $(X^{\epsilon_n}, Y^{\epsilon_n})$ the solution to $(SDE)_{\epsilon_n}$ where the initial law satisfies, for a non negative and measurable function ψ on $\mathbb{R}$:

$$\mathbb{E}\left[\psi(X_0^{\epsilon_n})1_{\{Y_0^{\epsilon_n}=i\}}\right] = \int_{\mathbb{R}} \psi(x)\frac{\mu_i * \rho^{\epsilon_n}(0,x)}{\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,y)dy}dx,$$

for $1 \le i \le d$. The process $(X^{\epsilon_n}, Y^{\epsilon_n})$ has the law $\overline{\nu}^{\epsilon_n}$. First we check that for any $\eta > 0$, there exists a constant $K_\eta > 0$ such that $\forall n \ge 0, \mathbb{P}(\sup_{t\in[0,T]} |X_t^{\epsilon_n}| + |Y_t^{\epsilon_n}| > K_\eta) \le \eta$. For $K > 0$, $n \ge 0$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,T]}(|X_t^{\epsilon_n}| + |Y_t^{\epsilon_n}|) > K\right) &\le \mathbb{P}\left(|X_0^{\epsilon_n}| + \sup_{t\in[0,T]}|X_t^{\epsilon_n} - X_0^{\epsilon_n}| + \sup_{t\in[0,T]}|Y_t^{\epsilon_n}| > K\right)\\
&\le \mathbb{P}\left(\sup_{t\in[0,T]}|Y_t^{\epsilon_n}| > K/3\right) + \mathbb{P}\left(\sup_{t\in[0,T]}|X_t^{\epsilon_n} - X_0^{\epsilon_n}| > K/3\right) + \mathbb{P}\left(|X_0^{\epsilon_n}| > K/3\right).
\end{aligned}$$

In the second line, for the first term of the r.h.s., since Y^{ϵ_n} only takes values in $\{1,...,d\}$, we have that $\forall n \ge 0, \forall K > 3d$, $\mathbb{P}(\sup_{t\in[0,T]}|Y_t^{\epsilon_n}| \ge K/3) = 0$. For the second term, by Markov's inequality and Lemma 2.D.2 below, we have that for $n \ge 0$, $\mathbb{P}\left(\sup_{t\in[0,T]}|X_t^{\epsilon_n} - X_0^{\epsilon_n}| \ge K/3\right) \le \frac{3}{K}\mathbb{E}\left(\sup_{t\in[0,T]}|X_t^{\epsilon_n} - X_0^{\epsilon_n}|\right) \le \frac{3}{K}\left(T\overline{b} + 2\sqrt{T}\overline{a}^{1/2}\right)$. For $K \ge K_1 := \frac{6}{\eta}\left(T\overline{b} + 2\sqrt{T}\overline{a}^{1/2}\right)$, $\mathbb{P}\left(\sup_{t\in[0,T]}|X_t^{\epsilon_n} - X_0^{\epsilon_n}| \ge K/3\right) \le \frac{\eta}{2}$. To study the last term, let us prove that for $t \in \mathbb{R}$, $1 \le i \le d$ and $\phi \in C_b^2(\mathbb{R})$,

$$\int_{\mathbb{R}} \phi(x)\mu_i^{\epsilon_n}(t,dx) \underset{n\to\infty}{\to} \int_{\mathbb{R}} \phi(x)\mu_i(t,dx).$$

We study the difference:

$$\begin{aligned}
\left|\int_{\mathbb{R}} \phi(x)\mu_i^{\epsilon_n}(t,dx) - \int_{\mathbb{R}} \phi(x)\mu_i(t,dx)\right| &= \left|\int_{\mathbb{R}^2} \rho_T^{\epsilon_n}(t-s)\left((\phi * \rho_X^{\epsilon_n}(x))\,\mu_i(s,dx) - \phi(x)\mu_i(t,dx)\right)ds\right|\\
&\le \left|\int_{\mathbb{R}^2} \rho_T^{\epsilon_n}(t-s)\phi(x)\left(\mu_i(s,dx) - \mu_i(t,dx)\right)ds\right|\\
&+ \left|\int_{\mathbb{R}^2} \rho_T^{\epsilon_n}(t-s)\left(\phi * \rho_X^{\epsilon_n}(x) - \phi(x)\right)\mu_i(s,dx)ds\right|.
\end{aligned}$$

The first term of the r.h.s converges to 0 as $n \to \infty$ as the function $t \to \int \phi(x)\mu_i(t,dx)$ is continuous and bounded. For the last term, as $\phi \in C_b^2(\mathbb{R})$, ϕ is globally lipschitz and we observe that for $x \in \mathbb{R}$,

$$\begin{aligned}
|\phi * \rho_X^{\epsilon_n}(x) - \phi(x)| &\le \int_{\mathbb{R}} |\phi(y) - \phi(x)|\rho_X^{\epsilon_n}(x-y)dy \le ||\phi'||_\infty \int_{\mathbb{R}} |y-x|\rho_X^{\epsilon_n}(x-y)dy\\
&\le \epsilon_n C_0 ||\phi'||_\infty \left(\int_{\mathbb{R}} |y| \exp\left(-\sqrt{1+y^2}\right)dy\right),
\end{aligned}$$

so $\phi * \rho_X^{\epsilon_n}$ converges uniformly to ϕ and as $\mu(t,\mathbb{R}) \le B$ for $t \in \mathbb{R}$, we conclude that for $t \in [0,T]$ and $i \in \{1,...,d\}$,

$$\left|\int_{\mathbb{R}} \phi(x)\mu_i^{\epsilon_n}(t,dx) - \int_{\mathbb{R}} \phi(x)\mu_i(t,dx)\right| \underset{n\to\infty}{\to} 0. \tag{2.D.3}$$

In particular, as the random variable $X_0^{\epsilon_n}$ has the density $\frac{\sum_{i=1}^d \mu_i * \rho^{\epsilon_n}(0,x)}{\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,x)dx} dx$, the convergence (2.D.3) implies that for $\psi \in C_b^2(\mathbb{R})$,

$$\mathbb{E}\left[\psi\left(X_0^{\epsilon_n}\right)\right] \to \sum_{i=1}^d \int_{\mathbb{R}} \psi(x)\mu_0(dx,\{i\}),$$

so there exists $K_2 > 0$ such that $\mathbb{P}\left(|X_0^{\epsilon_n}| \geq K_2/3\right) \leq \frac{\eta}{2}$. As a consequence, for $K_\eta := \max\left(3d, K_1, K_2\right)$ and $\forall n \geq 0$,

$$\mathbb{P}\left(\sup_{t\in[0,T]} \left(|X_t^{\epsilon_n}| + |Y_t^{\epsilon_n}|\right) > K_\eta\right) \leq \eta.$$

We now check that for $\zeta > 0, K > 0$, there exists $\delta_{\zeta,K} > 0$ such that $\forall n \geq 0$,

$$\sup_{\substack{\tau\in\mathcal{T}^f_{[0,T]} \\ 0\leq\delta\leq\delta_{\zeta,K}}} \mathbb{P}\left(\max\left(|X^{\epsilon_n}_{\tau+\delta} - X^{\epsilon_n}_\tau|, |Y^{\epsilon_n}_{\tau+\delta} - Y^{\epsilon_n}_\tau|\right) > K\right) < \zeta,$$

where $\mathcal{T}^f_{[0,T]}$ denotes the set of stopping times taking values in a finite subset of $[0,T]$. For $K > 0$, $\delta \geq 0$, $n \geq 0$ and $\tau \in \mathcal{T}^f_{[0,T]}$, we have the following inequality,

$$\mathbb{P}\left(\max\left(|X^{\epsilon_n}_{\tau+\delta} - X^{\epsilon_n}_\tau|, |Y^{\epsilon_n}_{\tau+\delta} - Y^{\epsilon_n}_\tau|\right) > K\right) \leq \mathbb{P}\left(|X^{\epsilon_n}_{\tau+\delta} - X^{\epsilon_n}_\tau| > K\right) + \mathbb{P}\left(|Y^{\epsilon_n}_{\tau+\delta} - Y^{\epsilon_n}_\tau| > K\right) =: P_1 + P_2.$$

Using Markov's inequality on P_1, we have

$$\begin{aligned} P_1 &\leq \frac{1}{K^2}\mathbb{E}\left[\left|\int_\tau^{\tau+\delta} b^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})ds + \int_\tau^{\tau+\delta} \sqrt{a^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})}dW_s\right|^2\right] \\ &\leq \frac{2}{K^2}\mathbb{E}\left[\left|\int_\tau^{\tau+\delta} b^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})ds\right|^2 + \left|\int_\tau^{\tau+\delta} \sqrt{a^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})}dW_s\right|^2\right]. \end{aligned}$$

On the one hand $\mathbb{E}\left[\left|\int_\tau^{\tau+\delta} b^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})ds\right|^2\right] \leq \delta\mathbb{E}\left[\int_\tau^{\tau+\delta}\left(b^{\epsilon_n}_{Y_s^{\epsilon_n}}\right)^2(s, X_s^{\epsilon_n})ds\right] \leq \delta^2\overline{b}^2$ by the Cauchy-Schwarz inequality, and on the other hand $\mathbb{E}\left[\left|\int_\tau^{\tau+\delta} \sqrt{a^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})}dW_s\right|^2\right] = \mathbb{E}\left[\left|\int_0^T 1_{\{s\in[\tau,(\tau+\delta)\wedge T]\}} \sqrt{a^{\epsilon_n}_{Y_s^{\epsilon_n}}(s, X_s^{\epsilon_n})}dW_s\right|^2\right] \leq \delta\overline{a}$ by Ito's isometry. For P_2 we notice that

$$P_2 \leq 1 - \mathbb{P}\left(\forall s \in [\tau, \tau+\delta], Y^{\epsilon_n}_\tau = Y^{\epsilon_n}_{\tau+s}\right) \leq 1 - e^{-\overline{q}\delta}.$$

We gather the upper bounds on P_1 and P_2, and to satisfy Aldous' criterion, it is sufficient to choose $\delta_{\zeta,K}$ small enough so that

$$\frac{2}{K^2}\delta_{\zeta,K}\left(\delta_{\zeta,K}\overline{b}^2 + \overline{a}\right) + 1 - e^{-\overline{q}\delta_{\zeta,K}} \leq \epsilon.$$

We have then shown that the family of processes $(X^{\epsilon_n}, Y^{\epsilon_n})_{n\geq0}$ is tight so the family of measures $\left(\overline{\nu}^{\epsilon_n}\right)_{n\geq0}$ is tight. As E is Polish, by Prohorov's theorem there exists a measure ν which is a limit point of $\left(\overline{\nu}^{\epsilon_n}\right)_{n\geq0}$. Let us remark that by (2.D.3), $\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,x)dx \to 1$, so ν is also a limit point of $\left(\nu^{\epsilon_n}\right)_{n\geq0}$. For notational simplicity, we suppose in what follows that $\left(\nu^{\epsilon_n}\right)_{n\geq0}$ and $\left(\overline{\nu}^{\epsilon_n}\right)_{n\geq0}$ converge to ν as $n\to\infty$. By [37, Lemma 7.7 and Theorem 7.8], $D(\nu) := \{t \in [0,T], \nu\left((X_{t^-}, Y_{t^-}) = (X_t, Y_t)\right) = 1\}$ has a complement in $[0,T]$ which is at most countable. Moreover, for any $k \geq 1$, and $t_1, t_2, ..., t_k \in D(\nu)$, the sequence $\left(\left(X^{\epsilon_n}_{t_1}, Y^{\epsilon_n}_{t_1}\right), ..., \left(X^{\epsilon_n}_{t_k}, Y^{\epsilon_n}_{t_k}\right)\right)_{n\geq0}$ converges to $\left((X_{t_1}, Y_{t_1}), ..., (X_{t_k}, Y_{t_k})\right)$ in distribution as $n\to\infty$. Consequently, for $t \in D(\nu)$, $1 \leq i \leq d$, and $\phi \in C_c^\infty(\mathbb{R})$,

$$\int_E \phi(X_t)1_{\{Y_t=i\}}d\nu^{\epsilon_n}(X,Y) \underset{n\to\infty}{\to} \int_E \phi(X_t)1_{\{Y_t=i\}}d\nu(X,Y).$$

Moreover, by (2.D.2) and (2.D.3) we obtain the following equality, for $t \in D(\nu)$:

$$\int_{\mathbb{R}} \phi(x)\mu_i(t,dx) = \int_E \phi(X_t)1_{\{Y_t=i\}}d\nu(X,Y),$$

As the function $t \to \int_E \phi(X_t)1_{\{Y_t=i\}}d\nu(X,Y)$ is right-continuous and the function $t \to \int_{\mathbb{R}} \phi(x)\mu_i(t,dx)$ is continuous, the previous equality holds for $t \in (0,T]$.

Step 3: we check that ν is a martingale solution of the SDE (2.5.22) with initial distribution μ_0. We define

$$L_t\phi(x,i) := b_i(t,x)\partial_x\phi(x,i) + \frac{1}{2}a_i(t,x)\partial^2_{xx}\phi(x,i) + \sum_{j=1}^d q_{ij}(t,x)\phi(x,j),$$

and for $n \geq 0$,

$$L_t^{\epsilon_n}\phi(x,i) := b_i^{\epsilon_n}(t,x)\partial_x\phi + \frac{1}{2}a_i^{\epsilon_n}(t,x)\partial^2_{xx}\phi + \sum_{j=1}^d q_{ij}^{\epsilon_n}(t,x)\phi(x,j).$$

Let $s \in [0,T]$, $p \in \mathbb{N}^*$, $0 \leq s_1 \leq ... \leq s_p \leq s$ and let $\psi_1, ..., \psi_p$ be bounded and continuous functions on $\mathcal{S}$, with $||\psi_i||_\infty \leq 1$ for $1 \leq i \leq d$. Since for $n \geq 0$, $\overline{\nu}^{\epsilon_n}$ is a martingale solution to $(SDE)_{\epsilon_n}$, and $\nu^{\epsilon_n} = \left(\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,x)dx\right)\overline{\nu}^{\epsilon_n}$, we have that for $t \geq s$,

$$\int_E \left[\phi(X_t,Y_t) - \phi(X_s,Y_s) - \int_s^t L_u^{\epsilon_n}\phi(X_u,Y_u)du\right]\prod_{i=1}^p \psi_i(X_{s_i},Y_{s_i})d\nu^{\epsilon_n}(X,Y) = 0.$$

Let $\tilde{b}_i : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, $\tilde{a}_i : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, $\tilde{q}_{ij} : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be bounded and continuous functions for $1 \leq i,j \leq d$. Let us also define

$$\tilde{L}_t\phi(x,i) := \tilde{b}_i(t,x)\partial_x\phi + \frac{1}{2}\tilde{a}_i(t,x)\partial^2_{xx}\phi + \sum_{j=1}^d \tilde{q}_{ij}(t,x)\phi(x,j),$$

and for $n \geq 0$,

$$\tilde{L}_t^{\epsilon_n}\phi(x,i) := \tilde{b}_i^{\epsilon_n}(t,x)\partial_x\phi + \frac{1}{2}\tilde{a}_i^{\epsilon_n}(t,x)\partial^2_{xx}\phi + \sum_{j=1}^d \tilde{q}_{ij}^{\epsilon_n}(t,x)\phi(x,j),$$

where $\tilde{b}_i^{\epsilon_n}, \tilde{a}_i^{\epsilon_n}, \tilde{q}_{ij}^{\epsilon_n}$ are built analogously to $b_i^{\epsilon_n}, a_i^{\epsilon_n}, q_{ij}^{\epsilon_n}$. Then, recalling that $||\psi_i||_\infty \leq 1$ for $1 \leq i \leq d$, we get

$$\left|\int_E \left[\phi(X_t,Y_t) - \phi(X_s,Y_s) - \int_s^t \tilde{L}_u^{\epsilon_n}\phi(X_u,Y_u)du\right]\prod_{i=1}^p \psi_i(X_{s_i},Y_{s_i})d\nu^{\epsilon_n}(X,Y)\right| \tag{2.D.4}$$

$$\begin{aligned}
&\leq \int_E \left[\int_s^t \left|\left(L_u^{\epsilon_n} - \tilde{L}_u^{\epsilon_n}\right)\phi(X_u,Y_u)\right| du\right] d\nu^{\epsilon_n}(X,Y) \\
&\leq \frac{1}{2}\sum_{i=1}^d \int_s^t \int_{\mathbb{R}} \left|\left(\frac{(a_i\mu_i) * \rho^{\epsilon_n}}{\mu_i^{\epsilon_n}} - \frac{(\tilde{a}_i\mu_i) * \rho^{\epsilon_n}}{\mu_i^{\epsilon_n}}\right)\partial^2_{xx}\phi(x,i)\right| \mu_i^{\epsilon_n}(u,dx)du \\
&+ \sum_{i=1}^d \int_s^t \int_{\mathbb{R}} \left|\left(\frac{(b_i\mu_i) * \rho^{\epsilon_n}}{\mu_i^{\epsilon_n}} - \frac{(\tilde{b}_i\mu_i) * \rho^{\epsilon_n}}{\mu_i^{\epsilon_n}}\right)\partial_x\phi(x,i)\right| \mu_i^{\epsilon_n}(u,dx)du \\
&+ \sum_{i,j=1}^d \int_s^t \int_{\mathbb{R}} \left|\left(\frac{(q_{ij}\mu_i) * \rho^{\epsilon_n}}{\mu_i^{\epsilon_n}} - \frac{(\tilde{q}_{ij}\mu_i) * \rho^{\epsilon_n}}{\mu_i^{\epsilon_n}}\right)\phi(x,j)\right| \mu_i^{\epsilon_n}(u,dx)du
\end{aligned}$$

$$\leq \frac{1}{2}\sum_{i=1}^d \int_s^t \left(\int_{\mathbb{R}^2} \rho_T^{\epsilon_n}(u-v)|a_i(v,x) - \tilde{a}_i(v,x)| \left(\rho_X^{\epsilon_n} * |\partial^2_{xx}\phi|(x,i)\right)\mu_i(v,dx)dv\right)du \tag{2.D.5}$$

$$+ \sum_{i=1}^d \int_s^t \left(\int_{\mathbb{R}^2} \rho_T^{\epsilon_n}(u-v)|b_i(v,x) - \tilde{b}_i(v,x)| \left(\rho_X^{\epsilon_n} * |\partial_x\phi|(x,i)\right)\mu_i(v,dx)dv\right)du \tag{2.D.6}$$

$$+ \sum_{i,j=1}^d \int_s^t \left(\int_{\mathbb{R}^2} \rho_T^{\epsilon_n}(u-v)|q_{ij}(v,x) - \tilde{q}_{ij}(v,x)| \left(\rho_X^{\epsilon_n} * |\phi|(x,i)\right)\mu_i(v,dx)dv\right)du \tag{2.D.7}$$

Our goal now is to let $n \to \infty$ in the terms (2.D.4)-(2.D.7), and obtain that for any $s_1, ..., s_p, s, t \in [0,T]$, such that $0 \le s_1 \le ... \le s_p \le s \le t$,

$$\left| \int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) - \int_s^t \tilde{L}_u \phi(X_u, Y_u) du \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu(X,Y) \right| \tag{2.D.8}$$

$$\le \frac{1}{2} \sum_{i=1}^d \int_s^t \int_{\mathbb{R}} |a_i(u,x) - \tilde{a}_i(u,x)| |\partial^2_{xx} \phi(x,i)| \mu_i(u,dx) du \tag{2.D.9}$$

$$+ \sum_{i=1}^d \int_s^t \int_{\mathbb{R}} |b_i(u,x) - \tilde{b}_i(u,x)| |\partial_x \phi(x,i)| \mu_i(u,dx) du \tag{2.D.10}$$

$$+ \sum_{i,j=1}^d \int_s^t \int_{\mathbb{R}} |q_{ij}(u,x) - \tilde{q}_{ij}(u,x)| |\partial_x \phi(x,i)| \mu_i(u,dx) du. \tag{2.D.11}$$

Here we explain how to conclude the proof of Theorem 2.5.11, if we suppose that (2.D.8)-(2.D.11) hold. Let us notice that the term

$$\left| \int_E \left[\int_s^t \left(\tilde{L}_u - L_u \right) \phi(X_u, Y_u) du \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu(X,Y) \right|,$$

is also dominated by the sum of the terms in the lines (2.D.9)-(2.D.11). By [24, Theorem 3.45], for $1 \le i \le d$, we can choose sequences of continuous functions $\left(\tilde{a}_i^k\right)_{k\ge 0}, \left(\tilde{b}_i^k\right)_{k \ge 0}$, and $\left(\tilde{q}_{ij}^k\right)_{k\ge 0}$ for $1 \le j \le d$ converging respectively to a_i, b_i, q_{ij} in $L^1([0,T] \times \mathbb{R}, \eta_i)$, with $\eta_i := \mu_i(t, \cdot)dt$, to finally obtain that

$$\int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) - \int_s^t L_u \phi(X_u, Y_u) du \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu(X,Y) = 0.$$

This is enough to conclude that ν is a martingale solution to the SDE (2.5.22) with initial condition μ_0 and end the proof of Theorem 2.5.11.

In what follows, we show how to obtain (2.D.8)-(2.D.11). As the function

$$(s_1, ..., s_k, s, t) \to \int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) - \int_s^t \tilde{L}_u \phi(X_u, Y_u) du \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu(X,Y),$$

is right-continuous, and as the terms in (2.D.9)-(2.D.11) are continuous in the variables (s,t), it is sufficient to show that the inequality in the lines (2.D.8)-(2.D.11) holds if we moreover assume that $s_1, ..., s_p, s, t \in D(\nu)$. Let us show that for $s_1, ..., s_p, s, t \in D(\nu)$,

$$\begin{aligned} &\int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) - \int_s^t \tilde{L}_u^{\epsilon_n} \phi(X_u, Y_u) du \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu^{\epsilon_n}(X,Y) \\ &\to \int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) - \int_s^t \tilde{L}_u \phi(X_u, Y_u) du \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu(X,Y). \end{aligned} \tag{2.D.12}$$

By convergence of the finite dimensional distributions of $(\overline{\nu}^{\epsilon_n})_{n \ge 0}$ and the fact that $\nu^{\epsilon_n} = \left(\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,x) dx \right) \overline{\nu}^{\epsilon_n}$ with $\sum_{i=1}^d \int_{\mathbb{R}} \mu_i * \rho^{\epsilon_n}(0,x) dx \underset{n \to \infty}{\to} 1$, the following convergence holds:

$$\begin{aligned} &\int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu^{\epsilon_n}(X,Y) \\ &\to \int_E \left[\phi(X_t, Y_t) - \phi(X_s, Y_s) \right] \prod_{i=1}^p \psi_i \left(X_{s_i}, Y_{s_i} \right) d\nu(X,Y). \end{aligned}$$

With the same argument, we have that for $u \in D(\nu) \cap (s,t)$,

$$\int_E \tilde{L}_u\phi(X_u, Y_u) \prod_{i=1}^p \psi_i\left(X_{s_i}, Y_{s_i}\right) d\nu^{\epsilon_n}(X,Y) \to \int_E \tilde{L}_u\phi(X_u, Y_u) \prod_{i=1}^p \psi_i\left(X_{s_i}, Y_{s_i}\right) d\nu(X,Y). \tag{2.D.13}$$

Then by Fubini's theorem,

$$\begin{aligned}
&\int_s^t \left| \int_E \tilde{L}_u^{\epsilon_n}\phi(X_u, Y_u) \prod_{i=1}^p \psi_i\left(X_{s_i}, Y_{s_i}\right) d\nu^{\epsilon_n}(X,Y) - \int_E \tilde{L}_u\phi(X_u, Y_u) \prod_{i=1}^p \psi_i\left(X_{s_i}, Y_{s_i}\right) d\nu(X,Y) \right| du \\
&\leq \int_s^t \left| \int_E \tilde{L}_u\phi(X_u, Y_u) \prod_{i=1}^p \psi_i\left(X_{s_i}, Y_{s_i}\right) d\nu^{\epsilon_n}(X,Y) - \int_E \tilde{L}_u\phi(X_u, Y_u) \prod_{i=1}^p \psi_i\left(X_{s_i}, Y_{s_i}\right) d\nu(X,Y) \right| du \qquad (2.D.14) \\
&+ \int_s^t \int_E |\tilde{L}_u^{\epsilon_n}\phi(X_u, Y_u) - \tilde{L}_u\phi(X_u, Y_u)| \prod_{i=1}^p \left|\psi_i\left(X_{s_i}, Y_{s_i}\right)\right| d\nu^{\epsilon_n}(X,Y) du \qquad (2.D.15)
\end{aligned}$$

The term on line (2.D.14) converges to zero as $n \to \infty$, by (2.D.13) and Lebesgue's theorem. For the term (2.D.15), we write

$$\int_s^t \int_E |\tilde{L}_u^{\epsilon_n}\phi(X_u, Y_u) - \tilde{L}_u\phi(X_u, Y_u)| \prod_{i=1}^p \left|\psi_i\left(X_{s_i}, Y_{s_i}\right)\right| d\nu^{\epsilon_n}(X,Y) du \leq \sum_{k=1}^d \int_s^t \int_{\mathbb{R}} |\tilde{L}_u^{\epsilon_n}\phi(x,k) - \tilde{L}_u\phi(x,k)| \mu_k^{\epsilon_n}(u,x) dx du.$$

As the terms in $\left(\tilde{L}_u^{\epsilon_n}\phi(\cdot,k)\right)_{1\leq k\leq d}$ and $\left(\tilde{L}_u\phi(\cdot,k)\right)_{1\leq k\leq d}$ are continuous and bounded, it is sufficient to show that for $1 \leq k \leq d$, $z_k : \mathbb{R}^2 \to \mathbb{R}$ a continuous and bounded function and $z_k^\epsilon := \frac{(z_k\mu_k)*\rho^\epsilon}{\mu_k^\epsilon}$, for $\epsilon > 0$, the following convergence holds:

$$\int_s^t \int_{\mathbb{R}} |z_k^\epsilon(u,x) - z_k(u,x)| \mu_k^\epsilon(u,x) dx du \underset{\epsilon\to 0}{\to} 0.$$

The previous term rewrites:

$$\begin{aligned}
\int_s^t \int_{\mathbb{R}} |z_k^\epsilon(u,x) - z_k(u,x)| \mu_k^\epsilon(u,x) dx du &= \int_s^t \int_{\mathbb{R}} |\left(z_k\mu_k\right) * \rho^\epsilon(u,x) - z_k(u,x)\mu_k^\epsilon(u,x)| dx du \\
\leq \int_s^t \int_{\mathbb{R}} \rho_T^\epsilon(u-v) \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \rho_X^\epsilon(x-y) |z_k(v,y) - z_k(u,x)| dx \right) \mu_k(v,dy) \right) dv du. && (2.D.16)
\end{aligned}$$

Let us notice that $\lim\limits_{M\to\infty} \sup\limits_{t\in\mathbb{R}} \mu_k(t, \mathbb{R}\backslash(-M,M)) = \lim\limits_{M\to\infty} \sup\limits_{t\in[0,T]} \mu_k(t, \mathbb{R}\backslash[-M,M]) = 0$. Indeed, if there exists $\delta > 0$, and a sequence $(t_n)_{n\geq 1}$ in the compact $[0,T]$ converging to $t \in [0,T]$ such that $\forall n \geq 1, \mu_k(t_n, \mathbb{R}\backslash(-n,n)) > \delta$, then by continuity, $\forall n \geq 1, \mu_k(t, \mathbb{R}\backslash(-n,n)) > \delta$, which is impossible.

For $\zeta > 0$, let $M > 0$ be such that $\sup\limits_{t\in\mathbb{R}} \mu_k(t, \mathbb{R}\backslash[-M,M]) \leq \zeta$, and let $\eta \in (0,1)$ such that $|z(v,y) - z(u,x)| \leq \zeta$ if $v, u \in [s-1, t+1]$, $x, y \in [-M-1, M+1]$, and $\max(|x-y|, |u-v|) \leq \eta$. Then, as

$$1 \leq 1_{\{|x-y|>\eta\}} + 1_{\{|u-v|>\eta\}} + 1_{\{|x-y|\leq\eta, |u-v|\leq\eta, (x,y)\in[-M-1,M+1]^2\}} + 1_{\{|x-y|\leq\eta, |u-v|\leq\eta, (x,y)\notin[-M-1,M+1]^2\}},$$

we have that $\forall u \in [s,t]$,

$$\begin{aligned}
&\int_{\mathbb{R}} \rho_T^\epsilon(u-v) \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \rho_X^\epsilon(x-y) |z(v,y) - z(u,x)| dx \right) \mu_i(v,dy) \right) dv \\
&\leq 2||z||_\infty B \left(\int_{|u-v|>\eta} \rho_T^\epsilon(u-v) dv + \int_{|x-y|>\eta} \rho_T^\epsilon(x-y) dx \right) + \zeta B + 2||z||_\infty \zeta \leq \left(4||z||_\infty + B\right)\zeta,
\end{aligned}$$

for ϵ small enough. Then we obtain that the term in line (2.D.16) converges to 0 as $\epsilon \to 0$. We now analyze the term on the line (2.D.5). We define the function $g : \mathbb{R} \to \mathbb{R}$ by

$$g(v) = \int_{\mathbb{R}} |a_i(v,x) - \tilde{a}_i(v,x)| |\partial_{xx}^2\phi(x,i)| \mu_i(v,dx),$$

and for $\epsilon > 0$, the function $g^\epsilon : \mathbb{R} \to \mathbb{R}$ by

$$g^\epsilon(v) = \int_{\mathbb{R}} |a_i(v,x) - \tilde{a}_i(v,x)| \left(\rho_X^\epsilon * |\partial^2_{xx}\phi(x,i)|\right) \mu_i(v,dx).$$

We study, for $c > 0$,

$$\begin{aligned}\int_s^t \left(\int_{\mathbb{R}} \rho_T^\epsilon(u-v) g^\epsilon(v)dv\right) du \quad &= \quad \int_s^t \left(\int_{\mathbb{R}} 1_{\{|u-v|>c\epsilon\}} \rho_T^\epsilon(u-v) g^\epsilon(v)dv\right) du \\ &+ \quad \int_s^t \left(\int_{\mathbb{R}} 1_{\{|u-v|\leq c\epsilon\}} \rho_T^\epsilon(u-v) g^\epsilon(v)dv\right) du.\end{aligned}$$

For the first term of the r.h.s., as there exists $\gamma > 0$ such that $\sup_{\epsilon>0} ||g^\epsilon||_\infty \leq \gamma$, the change of variables $w := \frac{u-v}{\epsilon}$ gives

$$\int_s^t \left(\int_{\mathbb{R}} 1_{\{|u-v|>c\epsilon\}} \rho_T^\epsilon(u-v) g^\epsilon(v)dv\right) du \leq \gamma(t-s) \int_{\mathbb{R}} 1_{\{|w|>c\}} \rho_T(w)dw.$$

For the second term of the r.h.s,

$$\int_s^t \left(\int_{\mathbb{R}} 1_{\{|u-v|\leq c\epsilon\}} \rho_T^\epsilon(u-v) g^\epsilon(v)dv\right) du \leq \int_{s-c\epsilon}^{t+c\epsilon} g^\epsilon(v) \left(\int_s^t \rho_T^\epsilon(u-v)du\right) dv \leq \int_{s-c\epsilon}^{t+c\epsilon} g^\epsilon(v)dv.$$

We then set $c = \frac{1}{\sqrt{\epsilon}}$, and obtain:

$$\int_s^t \left(\int_{\mathbb{R}} \rho_T^\epsilon(u-v) g^\epsilon(v)dv\right) du \leq \gamma(t-s) \int_{\mathbb{R}} 1_{\{|w|>\frac{1}{\sqrt{\epsilon}}\}} \rho_T(w) + \int_{s-\sqrt{\epsilon}}^{t+\sqrt{\epsilon}} g^\epsilon(u)du,$$

where the r.h.s of the inequality converges to $\int_s^t g(u)du$ as $\epsilon \to 0$, since $\rho_X^{\epsilon_n} * |\partial^2_{xx}\phi| \to |\partial^2_{xx}\phi|$ uniformly. A similar argument applies to the two last terms of Inequality (2.D.6)-(2.D.7), so that (2.D.8)-(2.D.11) holds.

Lemma 2.D.2. *Let $\epsilon > 0$, and let (X^ϵ, Y^ϵ) be the solution to $(SDE)_\epsilon$. Then,*

$$\mathbb{E}[\sup_{t\in[0,T]} |X_t^\epsilon - X_0^\epsilon|] \leq T\bar{b} + 2\sqrt{T}\bar{a}^{1/2}.$$

Proof. For $t \in [0,T]$,

$$\begin{aligned}|X_t^\epsilon - X_0^\epsilon| \quad &\leq \quad \left|\int_0^t b^\epsilon_{Y_s^\epsilon}(X_s)ds\right| + \left|\int_0^t \sqrt{a^\epsilon_{Y_s^\epsilon}(X_s^\epsilon)}dW_s\right| \\ &\leq \quad \int_0^t |b^\epsilon_{Y_s}(X_s^\epsilon)|ds + \sup_{u\leq t}\left|\int_0^u \sqrt{a^\epsilon_{Y_s}(X_s^\epsilon)}dW_s\right| \\ &\leq \quad T\bar{b} + \sup_{u\leq T}\left|\int_0^u \sqrt{a_{Y_s^\epsilon}(X_s^\epsilon)}dW_s\right|.\end{aligned}$$

Taking the supremum on $t \in [0,T]$ and then the expectation, we obtain:

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t^\epsilon - X_0^\epsilon|\right] \leq \left(T\bar{b} + \mathbb{E}\left[\sup_{u\leq T}\left|\int_0^u \sqrt{a^\epsilon_{Y_s}(X_s)}dW_s\right|\right]\right)$$

Finally, according to Doob's inequality,

$$\mathbb{E}\left[\sup_{u\leq T}\left|\int_0^u \sqrt{a^\epsilon_{Y_s}(X_s^\epsilon)}dW_s\right|\right] \leq 2\sqrt{T}\bar{a}^{1/2},$$

and this concludes the proof. □

Proposition 2.D.3. *For $\epsilon > 0$, $(PDS)_\epsilon$ has a unique solution.*

Proof. For $0 < s \leq t \leq T$, let us define $(X_t(s,x,y), Y_t(s,x,y))$ the value at t of the solution to the SDE (2.5.22) starting at s from $(x,y) \in \mathcal{S}$. For $\phi : \mathcal{S} \to \mathbb{R}$ such that for $i \in \{1,...,d\}$, $\phi(\cdot,i) \in C_c^\infty(\mathbb{R})$, we define $\mathcal{T}_t\phi(s,x,y) = \mathbb{E}\left[\phi\left(X_t(s,x,y), Y_t(s,x,y)\right)\right]$, and for $f : [0,t] \times \mathcal{S} \to \mathbb{R}$, such that for $1 \leq i \leq d$, $f(\cdot,\cdot,i) \in C_b^{1,2}([0,t] \times \mathbb{R})$, we define

$$L^\epsilon f(s,x,i) := b_i^\epsilon(s,x)\partial_x f(s,x,i) + \frac{1}{2}a_i^\epsilon(s,x)\partial_{xx}^2 f(s,x,i) + \sum_{j=1}^d q_{ij}^\epsilon(s,x) f(s,x,j).$$

By [100, Theorem 5.2], for $1 \leq y \leq d$, the function $(s,x) \to (\mathcal{T}_t\phi)(s,x,y)$ belongs to $C_b^{1,2}([0,t]\times\mathbb{R})$, and the function $s,x,y \to \mathcal{T}_t\phi(s,x,y)$ satisfies the backward Kolmogorov equation $\partial_s(\mathcal{T}_t\phi) + L^\epsilon(\mathcal{T}_t\phi) = 0$. Now, let $(\hat{\mu}_1^\epsilon(t,\cdot),...,\hat{\mu}_d^\epsilon(t,\cdot))_{t\in(0,T]}$ and $(\check{\mu}_1^\epsilon(t,\cdot),...,\check{\mu}_d^\epsilon(t,\cdot))_{t\in(0,T]}$ be two solutions to $(PDS)_\epsilon$ with the same initial condition, and let us define for $1 \leq i \leq d$ and $t \in (0,T]$, $z_i(t,\cdot) := \hat{\mu}_i^\epsilon(t,\cdot) - \check{\mu}_i^\epsilon(t,\cdot)$. It is sufficient to prove

$$\begin{aligned}\frac{d}{ds}\sum_{i=1}^d \int (\mathcal{T}_t\phi)(s,x,i) z_i(s,dx) &= \sum_{i=1}^d \int \left[\partial_s(\mathcal{T}_t\phi)(s,x,i) + L^\epsilon(\mathcal{T}_t\phi)(s,x,i)\right] z_i(s,dx) &\quad (2.D.17)\\ &= 0. &\quad (2.D.18)\end{aligned}$$

to obtain uniqueness. Indeed, $s \to \sum_{i=1}^d \int (\mathcal{T}_t\phi)(s,x,i) z_i(s,dx)$ would be constant on $(0,t)$, and moreover, for $s \in (0,t)$,

$$\begin{aligned}\sum_{i=1}^d \int (\mathcal{T}_t\phi)(s,x,i) z_i(s,dx) &= \sum_{i=1}^d \int \left|(\mathcal{T}_t\phi)(s,x,i) - (\mathcal{T}_t\phi)(0,x,i)\right| z_i(s,dx) + \sum_{i=1}^d \int (\mathcal{T}_t\phi)(0,x,i) z_i(s,dx)\\ &\leq 2Bs\sum_{i=1}^d ||\partial_s\mathcal{T}_t(\cdot,\cdot,i)||_\infty + \sum_{i=1}^d \int (\mathcal{T}_t\phi)(0,x,i) z_i(s,dx),\end{aligned}$$

so that

$$\sum_{i=1}^d \int (\mathcal{T}_t\phi)(s,x,i) z_i(s,dx) \underset{s\to 0}{\to} 0,$$

as $\hat{\mu}_i^\epsilon$ and $\check{\mu}_i^\epsilon$ satisfy the same initial conditions. Moreover, at the limit $s \to t$, we obtain that

$$\sum_{i=1}^d \int \phi(x,i) z_i(t,dx) = 0,$$

for any function $\phi : \mathcal{S} \to \mathbb{R}$ such that for $i \in \{1,...,d\}$, $\phi(\cdot,i) \in C_c^\infty(\mathbb{R})$. Thus, we can conclude that $z_i(t,\cdot) = 0$ for $1 \leq i \leq d$ and $t \in (0,T]$, hence the uniqueness.

We now prove equality (2.D.17) by a density argument. Let $\kappa \in C_c^\infty((0,t))$, and let us show that:

$$\int_0^t -\kappa'(s)\sum_{i=1}^d \int (\mathcal{T}_t\phi)(s,x,i) z_i(s,dx)ds = \int_0^t \kappa(s)\sum_{i=1}^d \int \left[\partial_s(\mathcal{T}_t\phi)(s,x,i) + L^\epsilon(\mathcal{T}_t\phi)(s,x,i)\right] z_i(s,dx)ds.$$

For ψ, φ two real valued functions defined on $\mathcal{S}$ such that for $i \in \{1,...,d\}$, $\varphi(\cdot,i) \in C_c^\infty((0,t))$ and $\psi(\cdot,i) \in C_c^\infty(\mathbb{R})$, we have in the sense of distributions,

$$\frac{d}{ds}\sum_{i=1}^d \int \varphi(s,i)\psi(x,i) z_i(s,dx) = \sum_{i=1}^d \int \left[\varphi'(s,i)\psi(x,i) + L^\epsilon(\varphi\psi)(s,x,i)\right] z_i(s,dx).$$

Let $h \in C_c^\infty(\mathbb{R})$ be such that $0 \leq h \leq 1$, $h(0) = 1$, and $h(x) = 0$ if $x \notin (-1,1)$. For $M > 0$, we define $h_M \in C_c^\infty(\mathbb{R})$ such that $h_M(x) = 1$ if $x \in [-M,M]$, $h_M(x) = 0$ if $x \notin [-M-1, M+1]$, $h_M(x) = h(x+M)$ if $x \in [-M-1,-M]$, and $h_M(x) = h(x-M)$ if $x \in [M, M+1]$. Let us remark that the family $(h_M T_t\phi)_{M>0}$ has uniform in M bounds. Let $\eta_1 > 0$, and choose M such that $\sum_{i=1}^d \int_0^t \int_{\mathbb{R}\setminus[-M,M]} \hat{\mu}_1^\epsilon(s,dx) < \eta_1$, and $\sum_{i=1}^d \int_0^t \int_{\mathbb{R}\setminus[-M,M]} \check{\mu}_1^\epsilon(s,dx) < \eta_1$. For $\eta_2 > 0$, there exists $p \in \mathbb{N}^*$, a family $(\psi_k)_{1\leq k\leq p}$ of functions defined on $(0,t)\times\{1,...,d\}$, such that for $1 \leq i \leq d$, and $1 \leq k \leq p$, $\psi_k(\cdot,i) \in C_c^\infty((0,t))$ and a family $(\varphi_k)_{1\leq k\leq p}$ of functions defined on $\mathcal{S}$ such that for $1 \leq i \leq d$, and $1 \leq k \leq p$, $\varphi_k(\cdot,i) \in C_c^\infty(\mathbb{R})$, and moreover, $\sum_{k=0}^p \psi_k(\cdot,i)\varphi_k(\cdot,i)$ has support in $(0,t)\times[-M-2;M+2]$, and such that the function $\sum_{k=0}^{p}\psi_k\varphi_k - h_M\mathcal{T}_t\phi$ and its derivatives are smaller than η_2. Using the decomposition $\mathcal{T}_t\phi = (h_M+1-h_M)\,\mathcal{T}_t\phi$, it is easy to check that

$$\left|\sum_{i=1}^{d}\int_0^t -\kappa'(s)\int_{\mathbb{R}}\left(\sum_{k=0}^{p}\psi_k\varphi_k - \mathcal{T}_t\phi\right)(s,x,i)z_i(s,dx)ds\right| \leq K_1\left(\eta_1+\eta_2\right),$$

$$\left|\sum_{i=1}^{d}\int_0^t \kappa(s)\int_{\mathbb{R}}\left(\sum_{k=0}^{p}\left(\psi_k^{'}\varphi_k + L^{\epsilon}\left(\psi_k\varphi_k\right)\right) - \left(\partial_s\left(\mathcal{T}_t\phi\right) + L^{\epsilon}\left(\mathcal{T}_t\phi\right)\right)\right)(s,x,i)z_i(s,dx)ds\right| \quad \leq \quad K_2\left(\eta_1+\eta_2\right),$$

where $K_j, j=1,2$, are positive values that do not depend on M or p, and this concludes the proof. $\square$

Chapter 3

Discretisation in time of a class of mean-field diffusions including the calibrated local and stochastic volatility model

Ce chapitre est un travail réalisé avec Benjamin Jourdain.

Abstract

Using the technique developed by Talay and Tubaro, we show the weak convergence at order 1 for the explicit Euler scheme with constant time step discretizing diffusions nonlinear in the sense of McKean, with coefficients containing conditional expectations computed with respect to the coordinates of the solution and satisfying a structure condition. An example is given by the Local and Stochastic Volatility model calibrated to the market prices of vanilla options. When the diffusion coefficient is uniformly elliptic, we propose a half-step scheme that allows a representation of the conditional expectation as a ratio of convolutions against heat kernels. We then study an interacting particles system based on the half-step scheme and under a slight Lipschitz modification of the heat kernel, we estimate its rate of weak convergence.

Keywords: Euler schemes, diffusions nonlinear in the sense of McKean, interacting stochastic particles

3.1 Introduction

In the field of mathematical finance, calibration of models to market prices of vanilla options is a major concern for pricing and hedging purposes. Under the assumption that we have access to the European call prices $C(t, K)$ for a continuum of strikes $K > 0$ and maturities $t > 0$, the Local and Stochastic Volatility model (LSV) calibrated to the market price of European call options, as introduced by Lipton [76] and Piterbarg [89], gives the dynamics for the log-price X of the underlying asset

$$dX_t = \left(r - \frac{1}{2}\sigma_{Dup}^2(t, X_t) \frac{f^2(Y_t)}{\mathbb{E}[f^2(Y_t)|X_t]} \right) dt + \sigma_{Dup}(t, X_t) \frac{f(Y_t)}{\sqrt{\mathbb{E}[f^2(Y_t)|X_t]}} dW_t. \quad (3.1.1)$$

Here, r is the interest rate, Y is a stochastic process, $f : \mathbb{R} \to \mathbb{R}$ is a positive function, the stochastic volatility factor at time t is $f(Y_t)$ and W is an unidimensional Brownian motion. Moreover, for $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$, $\sigma_{Dup}(t, x) = \hat{\sigma}_{Dup}(t, e^x)$, where for $t, K \geq 0$, $\hat{\sigma}_{Dup}(t, K) = \sqrt{2\frac{\partial_t C(t,K) + rK\partial_{KK}C(t,K)}{K^2\partial_{KK}^2 C(t,K)}}$ is the Dupire local volatility function, introduced in [36]. Due to the presence of the conditional expectation in the denominator in the drift and the diffusion coefficients, that SDE is nonlinear in the sense of McKean. Getting existence and uniqueness to this SDE is a challenging problem, as the conditional expectation does not satisfy Lipschitz property w.r.t. the Wasserstein distance. In [1], the authors prove local in time existence where f is a small perturbation of a constant thus studying a perturbation of the Dupire local volatility model, and Y an Ito process. In the last chapter, global existence is established in the particular case where Y is a jumping process taking a finite number of values. In the other cases, when for instance when Y is an autonomous Ito diffusion, global existence and uniqueness remain open problems.

As we formally observe that

$$\mathbb{E}\left[\sigma_{Dup}^2(t, X_t) \frac{f^2(Y_t)}{\mathbb{E}[f^2(Y_t)|X_t]} \middle| X_t \right] = \sigma_{Dup}^2(t, X_t), \quad (3.1.2)$$

we have under mild assumptions, by Gyongy's theorem [58, Thm. 4.6], that for $t \geq 0$, X_t has the same distribution as Z_t, where Z is a solution to the Dupire local volatility SDE

$$dZ_t = \left(r - \frac{1}{2}\sigma_{Dup}^2(t, Z_t) \right) dt + \sigma_{Dup}(t, Z_t) dW_t.$$

Motivated by that example, we first consider in this work the weak error between the time-discretized Euler scheme associated with a diffusion process X with a nonlinearity in the sense of McKean given by conditional expectations computed w.r.t. X, and a simpler diffusion process Z. We will work under the following general framework. Let $d_1 \geq 1$ and Z be a solution to the SDE

$$\begin{aligned} dZ_t &= b(t, Z_t) dt + \sigma(t, Z_t) dB_t, \\ Z_0 &\sim \mu_{Z_0}. \end{aligned} \quad (3.1.3)$$

Here B is a d_1−dimensional Brownian motion, μ_{Z_0} is a probability measure on $\mathbb{R}^{d_1}$, $b : [0, \infty) \times \mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$ and $\sigma : [0, \infty) \times \mathbb{R}^{d_1} \to \mathcal{S}_{d_1}^+(\mathbb{R})$ are measurable functions. Now let $d_2, q \geq 2$, and let (X, Y) be a solution to the SDE

$$\begin{aligned} dX_t &= b_X(t, X_t, Y_t, \mathbb{E}[\phi(X_t, Y_t)|X_t]) dt + \sigma_X(t, X_t, Y_t, \mathbb{E}[\phi(X_t, Y_t)|X_t]) dW_t^1, \\ dY_t &= b_Y(t, X_t, Y_t) dt + \sigma_Y(t, X_t, Y_t) dW_t^2, \\ (X_0, Y_0) &\sim \mu_0, \end{aligned} \quad (3.1.4)$$

where $b_X : [0, \infty) \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q \to \mathbb{R}^{d_1}$, $\sigma_X : [0, \infty) \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q \to \mathcal{S}_{d_1}^+(\mathbb{R})$, $b_Y : [0, \infty) \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$, $\sigma_Y : [0, \infty) \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathcal{S}_{d_2}^+(\mathbb{R})$ and $\phi : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}^q$ are measurable functions, W^1 (resp. W^2) is a d_1-dimensional (resp. d_2-dimensional) Brownian motion, μ_0 is a probability measure on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ such that its image by the d_1 first coordinates is equal to μ_{Z_0}. The brownian motions W^1, W^2 may be correlated. We will work in a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and we denote the augmented filtration of (W^1, W^2) by $\mathcal{F}$. Let us introduce the notation $a_X := \sigma_X \sigma_X^*$ and $a := \sigma\sigma^*$. Throughout this paper, and we assume that (b_X, a_X) satisfies the following property: for any random variable (A, B) with values in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, such that

$$\mathbb{E}[|\phi(A, B)|] < \infty, \text{ and } \forall t \geq 0, \quad \mathbb{E}[|b_X(t, A, B, \mathbb{E}[\phi(A, B)|A])|] < \infty, \quad \mathbb{E}[||a_X(t, A, B, \mathbb{E}[\phi(A, B)|A])||] < \infty,$$

we have that almost surely,

$$\forall t \geq 0,\ \mathbb{E}\left[b_X\left(t,A,B,\mathbb{E}\left[\phi\left(A,B\right)|A\right]\right)|A\right] = b\left(t,A\right), \qquad \mathbb{E}\left[a_X\left(t,A,B,\mathbb{E}\left[\phi\left(A,B\right)|A\right]\right)|A\right] = a\left(t,A\right) \tag{3.1.5}$$

Due to Property (3.1.2), it is easy to check that the LSV model is a particular case of Property (3.1.5). Indeed, the LSV model corresponds to the situation where for $d \geq 2$ and $(t,x,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^{d-1} \times \mathbb{R}$,

$$b_X\left(t,x,y,z\right) = r - \frac{1}{2}\sigma_{Dup}^2\left(t,x\right)\frac{f^2\left(y\right)}{z}, \qquad \phi\left(x,y\right) = f^2\left(y\right), \qquad \sigma_X\left(t,x,y,z\right) = \sigma_{Dup}\left(t,x\right)\frac{f\left(y\right)}{\sqrt{z}}$$
$$b\left(t,x\right) = r - \frac{1}{2}\sigma_{Dup}^2\left(t,x\right), \quad \sigma\left(t,x\right) = \sigma_{Dup}\left(t,x\right),$$

and for instance the functions b_Y, σ_Y satisfy the Lipschitz property and do not depend on the component X, so that Y is a well-posed autonomous Ito diffusion. Given a finite time horizon $T > 0$, we will first study the weak error between the law of Z_T and the one of the component X_T^n, $n \geq 1$, in the explicit Euler scheme with constant time step $\Delta = \frac{T}{n}$ associated with (X,Y). For $n \in \mathbb{N}^*$, this Euler scheme is given by

$$\begin{aligned} dX_t^n &= b_X\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n, \mathbb{E}\left[\phi\left(X_{\tau_t}^n, Y_{\tau_t}^n\right)|X_{\tau_t}^n\right]\right)dt + \sigma_X\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n, \mathbb{E}\left[\phi\left(X_{\tau_t}^n, Y_{\tau_t}^n\right)|X_{\tau_t}^n\right]\right)dW_t^1, \\ dY_t^n &= b_Y\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n\right)dt + \sigma_Y\left(\tau_t, X_{\tau_t}^n, Y_{\tau_t}^n\right)dW_t^2, \\ \left(X_0^n, Y_0^n\right) &\sim \mu_0, \end{aligned} \tag{3.1.6}$$

where for $t \in [0,T]$, $\tau_t := \lfloor \frac{nt}{T} \rfloor \frac{T}{n}$ is the last discretization time before t. While getting existence and uniqueness results for the SDE (3.1.4) is a challenging problem in general, existence and uniqueness to (3.1.6) holds under mild assumptions, as we will see in Section 3.2.2. Using the technique introduced by Talay-Tubaro [95], we obtain that under regularity assumptions on the coefficients of the process Z and the test function, the weak error is bounded by a term of order 1.

In Section 3.2, we show the bound of order 1 for the weak error in a regular setting and we also give the weak error of the Euler scheme discretizing in time the LSV model when the test function is the payoff of a put. Then under an ellipticity condition, we show in Section 3.3 that the conditional expectation can be expressed as a ratio of convolutions w.r.t. the heat kernel. After regularization in order to obtain a Lipschitz version of that ratio, we propose a half step scheme discretizing in time the SDE nonlinear in the sense of McKean and we also study its weak error. Finally in order to be implemented as a simulation, we introduce an associated particles system. We study its rate of convergence and show some numerical results.

Notation

- For $d \geq 1$ and $x = (x_1, ..., x_d) \in \mathbb{R}^d$, we define $|x| = \left(\sum_{j=1}^d x_j^2\right)^{1/2}$.
- Given a matrix A, we denote by A^* its transpose and by $||A||$ its operator norm. Moreover, if A is square, $\mathbf{Tr}\left(A\right)$ denotes its trace.
- $\mathcal{S}_d^+\left(\mathbb{R}\right)$ is the set of symmetric positive semidefinite real $d \times d$ matrices.
- For $m \geq 0$, $C^m\left(\mathbb{R}^d\right)$ is the space of functions $\psi : \mathbb{R}^d \to \mathbb{R}$ that have continuous derivatives of order up to m.
- For $m, n \geq 0$, $C^{m,n}\left([0,\infty) \times \mathbb{R}^d\right)$ is the space of functions $[0,\infty) \times \mathbb{R}^d \ni (t,x) \to \psi\left(t,x\right) \in \mathbb{R}$ such that ψ has m continuous derivatives w.r.t. t and n continuous derivatives w.r.t. x.
- For $d, q \geq 1$, a function $\psi : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^q$ has polynomial growth in the space variable if there exist constants $K, p > 0$ that do not depend on the time variable t and such that
$$\forall\left(t,x\right) \in [0,\infty) \times \mathbb{R}^d,\ |\psi(t,x)| \leq K\left(1 + |x|^p\right).$$
- If a function $u : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ has continuous derivatives up to the order 4 in the spatial variable, we denote by $\nabla_x u = (\partial_1 u, ..., \partial_d u)$ its gradient vector and $\nabla_x^2 u = \left(\partial_{ij}^2 u\right)_{1 \leq i,j \leq d}$ its Hessian matrix. We also use the notation $\nabla_x^3 u = \left(\partial_{ijk}^3 u\right)_{1 \leq i,j,k \leq d}$, $\nabla_x^4 u = \left(\partial_{ijkl}^4 u\right)_{1 \leq i,j,k,l \leq d}$. Moreover, for $x, y \in \mathbb{R}^d$, $x^* \nabla_x^3 u$ is the matrix $\left(\sum_{k=1}^d x_k \partial_{ijk}^3 u\right)_{1 \leq i,j \leq d}$ and $x^*\left(\nabla_x^4 u\right) y$ is the matrix $\left(\sum_{k,l=1}^d x_k y_l \partial_{ijkl}^4 u\right)_{1 \leq i,j \leq d}$.

3.2 Euler discretization of the SDE

3.2.1 Weak error estimates

Let $\varphi : \mathbb{R}^{d_1} \to \mathbb{R}$ be a test function. We will study the weak error $|\mathbb{E}[\varphi(X_T^n)] - \mathbb{E}[\varphi(Z_T)]|$ under the following assumptions.

- (A) The functions b_i, σ_{ij} belong to $C^{1,4}\left([0,T] \times \mathbb{R}^{d_1}\right)$ for $1 \leq i,j \leq d_1$, and their derivatives of positive order are bounded. The function φ belongs to $C^4\left(\mathbb{R}^{d_1}\right)$. Moreover, φ and its derivatives have polynomial growth.

- (SL) There exist $K_X, K_Y, K_\phi > 0$ such that for any random variable (A,B) with values in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and such that $\mathbb{E}[|\phi(A,B)|] < \infty$, we have almost surely that for $t \geq 0$ and $(x,y,z) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q$,

$$\begin{aligned} |b_X| \vee ||\sigma_X||\,(t,A,B,\mathbb{E}[\phi(A,B)|A]) &\leq K_X\left(1+|A|+|B|+|\mathbb{E}[\phi(A,B)|A]|\right), \\ |b_Y| \vee ||\sigma_Y||\,(t,A,B) &\leq K_Y\left(1+|A|+|B|\right), \\ |\phi(A,B)| &\leq K_\phi\left(1+|A|+|B|\right). \end{aligned} \tag{3.2.1}$$

- (IC) For any $p > 0$, $\int_{\mathbb{R}} \left(|x|^p + |y|^p\right) \mu_0(dx,dy) < \infty$.

In a financial models, we typically have that $d_1 = 1$ and μ_0 is a Dirac measure, so (IC) is satisfied. (A) is satisfied if σ_{Dup} belongs to $C^{1,4}([0,T] \times \mathbb{R})$ and is bounded as well as its derivatives and φ is a regular payoff function that belongs to $C^4(\mathbb{R})$. Moreover, if the function f^2 has sublinear growth and f has a positive lower bound f_{min}, then as $\mathbb{E}\left[f^2(A)\,|B\right] \geq f_{min}^2$ almost surely, there exists $C > 0$ such that

$$|b_X| \vee ||\sigma_X||\,(t,A,B,\mathbb{E}[\phi(A,B)|A]) \leq C\left|r + \frac{1}{2}\sigma^2(t,A)\frac{1+|B|}{f_{min}^2}\right| \vee |\sigma(t,A)|\frac{\sqrt{1+|B|}}{f_{min}},$$

and thus (SL) is satisfied as σ is bounded. We now give the main result of this section, that we prove in Section 3.2.2.

Theorem 3.2.1. *Under Assumptions (A), (SL) and (IC), there exists a constant $\zeta > 0$ such that*

$$\forall n \geq 1,\ |\mathbb{E}[\varphi(X_T^n)] - \mathbb{E}[\varphi(Z_T)]| \leq \frac{\zeta}{n}.$$

The case of put options In a financial framework, we are also interested in the weak convergence of the Euler scheme associated with the LSV model in the case where the test function is the payoff of a vanilla option. In this paragraph, we check that calibration is achieved for put options. We assume here that $d_1 = d_2 = q = 1$. Let us introduce slightly stronger assumptions.

- (A') The functions b, σ belong to the space $C^{1,6}([0,T] \times \mathbb{R})$ and their derivatives are bounded.

- (B) The functions b_X, σ_X are bounded. Moreover, the functions σ and σ_X are uniformly elliptic, that is there exists a positive constant $\underline{\sigma}$ such that $\underline{\sigma} \leq \sigma_X \wedge \sigma$.

Theorem 3.2.2. *Under Assumptions (A'), (B) and (IC), there exists a constant $\zeta > 0$ such that,*

$$\forall n \geq 2, \left|\mathbb{E}\left[\left(K - e^{X_T^n}\right)_+\right] - \mathbb{E}\left[\left(K - e^{Z_T}\right)_+\right]\right| \leq \zeta \frac{\log(n)}{n}. \tag{3.2.2}$$

Theorem 3.2.2 is proved in Section 3.2.3. Let us remark that in [11], the authors obtain a development of the weak error $|\mathbb{E}[\varphi(Z_T^n)] - \mathbb{E}[\varphi(Z_T)]|$, where Z^n is the explicit Euler scheme with constant time step Δ associated with the diffusion Z even if the terminal condition φ is irregular.

3.2.2 Proof of Theorem 3.2.1

Let us first establish existence and uniqueness to Equation (3.1.6). Let $0 \leq k \leq n-1$. If we have that $\mathbb{E}\left[|X^n_{k\Delta}| + |Y^n_{k\Delta}|\right] < \infty$, then by the affine growth of ϕ, $\mathbb{E}\left[|\phi\left(X^n_{k\Delta}, Y^n_{k\Delta}\right)|\right] \leq K(1 + \mathbb{E}\left[|X^n_{k\Delta}| + |Y^n_{k\Delta}|\right]) < \infty$. Therefore $\mathbb{E}\left[\phi\left(X^n_{k\Delta}, Y^n_{k\Delta}\right)|X^n_{k\Delta}\right]$ is well defined and

$$\mathbb{E}\left[\left|\mathbb{E}\left[\phi\left(X^n_{k\Delta}, Y^n_{k\Delta}\right)\right| X^n_{k\Delta}\right]\right|\right] \leq K_\phi\left(1 + \mathbb{E}\left[|X^n_{k\Delta}| + |Y^n_{k\Delta}|\right]\right).$$

Moreover, as b_X, σ_X satisfy (SL), we obtain that $\mathbb{E}\left[|X^n_u| + |Y^n_u|\right] < \infty$, where for $u \in [k\Delta, (k+1)\Delta)$,

$$\begin{aligned}
X^n_u &= X^n_{k\Delta} + b_X\left(k\Delta, X^n_{k\Delta}, Y^n_{k\Delta}, \mathbb{E}\left[\phi\left(X^n_{k\Delta}, Y^n_{k\Delta}\right)|X^n_{k\Delta}\right]\right)(u - k\Delta) \\
&+ \sigma_X\left(k\Delta, X^n_{k\Delta}, Y^n_{k\Delta}, \mathbb{E}\left[\phi\left(X^n_{k\Delta}, Y^n_{k\Delta}\right)|X^n_{k\Delta}\right]\right)\left(W^1_u - W^1_{k\Delta}\right), \\
Y^n_u &= Y^n_{k\Delta} + b_Y\left(k\Delta, X^n_{k\Delta}, Y^n_{k\Delta}\right)(u - k\Delta) + \sigma_Y\left(k\Delta, X^n_{k\Delta}, Y^n_{k\Delta}\right)\left(W^2_u - W^2_{k\Delta}\right),
\end{aligned}$$

Hence the existence of the Euler scheme holds by induction, as (IC) is satisfied. Moreover, trajectorial uniqueness follows by construction. In Lemma 3.2.3 below, we show in addition that the moments of the process (X^n, Y^n) are bounded, uniformly in the discretization parameter $n \geq 1$.

Lemma 3.2.3. *Let $p \geq 1$ and assume that $\int (|x|^p + |y|^p)\,\mu_0(dx, dy) < \infty$. Under Assumption (SL), we have that*

$$\sup_{n\geq 1} \mathbb{E}\left[\sup_{s\leq T}|X^n_s|^p + \sup_{s\leq T}|Y^n_s|^p\right] < \infty.$$

Proof. For $n \geq 1, t \in [0, T]$, we have

$$\begin{aligned}
3^{1-p}\sup_{s\leq t}|X^n_s|^p &\leq |X_0|^p + \left(\int_0^t \left|b_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right)\right| du\right)^p \\
&+ \sup_{s\leq t}\left|\int_0^s \sigma_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right) dW_u\right|^p,
\end{aligned}$$

where we used the fact that for $(x, y, z) \in \mathbb{R}^3$, $|x + y + z|^p \leq 3^{p-1}\left(|x|^p + |y|^p + |z|^p\right)$. By the Burkholder-Davis-Gundy inequality on the last term of the r.h.s., and Jensen's inequality on the measure $\frac{1}{t}1_{[0,t]}(u)du$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\leq t}\left|\int_0^s \sigma_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right) dW_s\right|^p\right] \\
&\leq C_p T^{\frac{p}{2}-1}\int_0^t \mathbb{E}\left[\left|\left|\sigma_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right)\right|\right|^p\right] du,
\end{aligned}$$

where the Burkholder-Davis-Gundy constant $C_p > 0$ only depends on p. By Jensen's inequality, we also obtain that

$$\mathbb{E}\left[\left(\int_0^t \left|b_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right)\right| du\right)^p\right] \leq T^{p-1}\mathbb{E}\left[\int_0^t \left|b_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right)\right|^p du\right].$$

As (SL) is satisfied, we have that

$$\begin{aligned}
\frac{4^{1-p}}{K^p_X}\left|b_X\left(\tau_u, X^n_{\tau_u}, Y^n_{\tau_u}, \mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right)\right|^p &\leq 1 + \left|X^n_{\tau_u}\right|^p + \left|Y^n_{\tau_u}\right|^p + \left|\mathbb{E}\left[\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)|X^n_{\tau_u}\right]\right|^p \\
&\leq 1 + \sup_{s\leq u}|X^n_s|^p + \sup_{s\leq u}|Y^n_s|^p + \mathbb{E}\left[\left|\phi\left(X^n_{\tau_u}, Y^n_{\tau_u}\right)\right|^p|X^n_{\tau_u}\right], \\
&\leq 1 + \sup_{s\leq u}|X^n_s|^p + \sup_{s\leq u}|Y^n_s|^p \\
&+ 3^{p-1}K^p_\phi\mathbb{E}\left[\left(1 + \sup_{s\leq u}\left|X^n_{\tau_u}\right|^p + \sup_{s\leq u}\left|Y^n_{\tau_u}\right|^p\right)|X^n_{\tau_u}\right].
\end{aligned}$$

Using the same computation, we obtain that

$$\begin{aligned}\frac{4^{1-p}}{K_X^p}\left|\left|\sigma_X\left(\tau_u, X_{\tau_u}^n, Y_{\tau_u}^n, \mathbb{E}\left[\phi\left(X_{\tau_u}^n, Y_{\tau_u}^n\right)|X_{\tau_u}^n\right]\right)\right|\right|^p &\leq 1+\sup_{s\leq u}|X_s^n|^p+\sup_{s\leq u}|Y_s^n|^p \\ &+ 3^{p-1}K_\phi^p\mathbb{E}\left[\left(1+\sup_{s\leq u}\left|X_{\tau_u}^n\right|^p+\sup_{s\leq u}\left|Y_{\tau_u}^n\right|^p\right)|X_{\tau_u}^n\right].\end{aligned}$$

Similarly, we also have that

$$3^{1-p}\sup_{s\leq t}|Y_s^n|^p \leq |Y_0|^p+\left(\int_0^t\left|b_Y\left(\tau_u, X_{\tau_u}^n, Y_{\tau_u}^n\right)\right|du\right)^p+\sup_{s\leq t}\left|\int_0^s \sigma_Y\left(\tau_u, X_{\tau_u}^n, Y_{\tau_u}^n\right)dW_u^2\right|^p,$$

with

$$\begin{aligned}\frac{3^{1-p}}{K_Y^p}\left|b_Y\left(\tau_u, X_{\tau_u}^n, Y_{\tau_u}^n\right)\right|^p &\leq 1+\sup_{s\leq u}|X_s^n|^p+\sup_{s\leq u}|Y_s^n|^p,\\ \frac{3^{1-p}}{K_Y^p}\left|\left|\sigma_Y\left(\tau_u, X_{\tau_u}^n, Y_{\tau_u}^n\right)\right|\right|^p &\leq 1+\sup_{s\leq u}|X_s^n|^p+\sup_{s\leq u}|Y_s^n|^p.\end{aligned}$$

Using the tower property of the expectation, we obtain that for $t\in[0,T]$,

$$\mathbb{E}\left[\sup_{s\leq t}|X_s^n|^p+\sup_{s\leq t}|Y_s^n|^p\right]\leq \kappa_1+\kappa_2\int_0^t\mathbb{E}\left[\sup_{s\leq u}|X_s^n|^p+\sup_{s\leq u}|Y_s^n|^p\right]du,$$

where the coefficients κ_1 and κ_2 do not depend on n. We conclude that the bounds given by Gronwall's lemma on the function $t\to\mathbb{E}\left[\sup_{s\leq t}|X_s^n|^p+\sup_{s\leq t}|Y_s^n|^p\right]$ are independent from n, if that function is locally integrable. A standard localization procedure permits to overcome this difficulty. ☐

Let us now define for $(t,x)\in[0,T]\times\mathbb{R}^{d_1}$, $u(t,x)=\mathbb{E}\left[\varphi(Z_T)|\,Z_t=x\right]$ and the operator $\mathcal{L}$ defined for $v\in C^{0,2}\left([0,T]\times\mathbb{R}^{d_1}\right)$ by

$$\mathcal{L}v=\left(\nabla_x v\right)^* b+\frac{1}{2}\mathbf{Tr}\left(a\nabla_x^2 v\right).$$

The regularity of the function u and the PDE that it satisfies are consequences of [43, Ch. 5, Thm. 5.5] and [43, Ch. 5, Thm. 6.1].

Proposition 3.2.4 (Friedman)**.** *The following assertions hold under Assumption (A):*
i) The functions $\partial_t u$, $\nabla_x u$, $\nabla_x^2 u$ are continuous on $[0,T]\times\mathbb{R}^{d_1}$ and satisfy the backward Kolmogorov PDE

$$\begin{aligned}\partial_t u+\mathcal{L}u &= 0 \text{ on } [0,T]\times\mathbb{R}^{d_1},\\ u(T,x) &= \varphi(x) \text{ for } x\in\mathbb{R}^{d_1}.\end{aligned}\tag{3.2.3}$$

ii) The functions u and $\nabla_x^k u, 1\leq k\leq 4$ have polynomial growth in the space variable.
iii) Moreover, under Assumption (IC),

$$\mathbb{E}\left[u\left(T,Z_T\right)\right]=\mathbb{E}\left[\varphi\left(Z_T\right)\right]=\mathbb{E}\left[u\left(0,Z_0\right)\right].$$

We now prove Theorem 3.2.1 using the Talay-Tubaro technique developed in [95]. By Property (*iii*) of Proposition 3.2.4, and using the fact that the initial conditions X_0^n and Z_0 have the same distribution, we have that

$$\mathbb{E}\left[\varphi\left(X_T^n\right)\right]-\mathbb{E}\left[\varphi\left(Z_T\right)\right]=\mathbb{E}\left[u\left(T,X_T^n\right)\right]-\mathbb{E}\left[u\left(0,Z_0\right)\right]=\mathbb{E}\left[u\left(T,X_T^n\right)\right]-\mathbb{E}\left[u\left(0,X_0^n\right)\right]=\sum_{k=0}^{n-1}\mathcal{E}_k,\tag{3.2.4}$$

where for $0\leq k\leq n-1$, $\mathcal{E}_k=\mathbb{E}\left[u\left(t_{k+1},X_{t_{k+1}}^n\right)\right]-\mathbb{E}\left[u\left(t_k,X_{t_k}^n\right)\right]$ and $t_k=k\Delta$. It is sufficient to prove that the error terms $\mathcal{E}_k$ are of order $\frac{1}{n^2}$ uniformly in n and the index $0\leq k\leq n-1$ to conclude the proof.

Proposition 3.2.5. *Under Assumptions (A), (SL) and (IC), there exists a finite constant ζ, that does not depend on n and such that*

$$\forall 1 \leq k \leq n, \ \left|\mathbb{E}\left[\mathcal{E}_k\right]\right| \leq \frac{\zeta}{n^2}.$$

What follows is dedicated to the proof of Proposition 3.2.5. For notational simplicity, we define for $0 \leq k \leq n-1$,

$$b^n_{X,k} = b_X\left(t_k, X^n_{t_k}, Y^n_{t_k}, \mathbb{E}\left[\phi\left(X^n_{t_k}, Y^n_{t_k}\right)|X^n_{t_k}\right]\right), \tag{3.2.5}$$

$$\sigma^n_{X,k} = \sigma_X\left(t_k, X^n_{t_k}, Y^n_{t_k}, \mathbb{E}\left[\phi\left(X^n_{t_k}, Y^n_{t_k}\right)|X^n_{t_k}\right]\right), \tag{3.2.6}$$

$$a^n_{X,k} = \sigma^n_{X,k}\left(\sigma^n_{X,k}\right)^*, \tag{3.2.7}$$

and L^n_k the operator which writes for $v \in C^{0,2}\left([0,T] \times \mathbb{R}^{d_1}\right)$,

$$\forall\left(t,x\right) \in \left[t_k, t_{k+1}\right) \times \mathbb{R}^{d_1}, \ L^n_k v\left(t,x\right) = \left(\nabla_x v\right)^*\left(t,x\right) b^n_{X,k} + \frac{1}{2}\mathbf{Tr}\left(a^n_{X,k}\nabla^2_x v\left(t,x\right)\right).$$

By Ito's lemma, and using the fact that u is a solution to the PDE (3.2.3), we have that

$$u\left(t_{k+1}, X^n_{t_{k+1}}\right) - u\left(t_k, X^n_{t_k}\right) = \int_{t_k}^{t_{k+1}} \left(\nabla_x u\right)^*\left(t, X^n_t\right)\sigma^n_{X,n} dW^1_t + \int_{t_k}^{t_{k+1}} \left\{\partial_t u + L^n_k u\right\} dt. \tag{3.2.8}$$

Lemma 3.2.6. *Under Assumptions (A), (SL) and (IC), the expectation of the stochastic integral in the r.h.s. of (3.2.8) is equal to zero.*

Proof. It is sufficient to check that for $0 \leq k \leq n-1$,

$$\int_{t_k}^{t_{k+1}} \mathbb{E}\left[\left|\left(\nabla_x u\right)^*\left(t, X^n_t\right)\sigma^n_{X,k}\right|^2\right] dt < \infty. \tag{3.2.9}$$

We use the inequality $\left|\left(\nabla_x u\right)^*\left(t, X^n_t\right)\sigma^n_{X,k}\right|^2 \leq \frac{1}{2}\left(\left|\left|\sigma^n_{X,k}\right|\right|^4 + \left|\nabla_x u\left(t, X^n_t\right)\right|^4\right)$. As (SL) is satisfied, we have by Jensen's inequality that

$$\left|\left|\sigma^n_{X,k}\right|\right| \leq K_X\left(1 + |X^n_{t_k}| + |Y^n_{t_k}| + \mathbb{E}\left[\left|\phi\left(X^n_{t_k}, Y^n_{t_k}\right)\right||X^n_{t_k}\right]\right).$$

Moreover, as $\nabla_x u$ has polynomial growth in the space variable, we deduce by Lemma 3.2.3 that the l.h.s. in (3.2.9) is finite and this ends the proof. □

Let us define for $0 \leq k \leq n-1$, and $t \in [t_k, t_{k+1})$, $\delta X^n_{t,k} := X^n_t - X^n_{t_k} = b^n_{X,k}(t-t_k) + \sigma^n_{X,k}(W^1_t - W^1_{t_k})$.

Lemma 3.2.7. *Under Assumption (SL), for $l \in \mathbb{N}^*$, there exists a function $\eta_l : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q \to \mathbb{R}_+$ such that there exist $K_l, \alpha_l > 0$ satisfying*

$$\forall\left(x,y,z\right) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q, \eta_l\left(x,y,z\right) \leq K_l\left(1 + |x|^{\alpha_l} + |y|^{\alpha_l} + |z|^{\alpha_l}\right)$$

and for $n \in \mathbb{N}^*$, $0 \leq k \leq n-1$, $t \in (t_k, t_{k+1}]$, $1 \leq j \leq d_1$, we have almost surely

$$\left|\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^l_j |X^n_{t_k}, Y^n_{t_k}\right]\right| \leq \left|\eta_l\left(X^n_{t_k}, Y^n_{t_k}, \mathbb{E}\left[\left|\phi\left(X^n_{t_k}, Y^n_{t_k}\right)\right||X^n_{t_k}\right]\right)\right|\left(t-t_k\right)^{1\vee\frac{l}{2}}. \tag{3.2.10}$$

Proof. For the case $l=1$, we have that

$$\begin{aligned}\left|\mathbb{E}\left[\left(\delta X^n_{t,k}\right)_j |X^n_{t_k}, Y^n_{t_k}\right]\right| &\leq \left|b_{X,j}\right|\left(t_k, X^n_{t_k}, Y^n_{t_k}, \mathbb{E}\left[\phi\left(X^n_{t_k}, Y^n_{t_k}\right)|X^n_{t_k}\right]\right)\left(t-t_k\right)\\ &\leq K_X\left(1 + |X^n_{t_k}| + |Y^n_{t_k}| + \mathbb{E}\left[\left|\phi\left(X^n_{t_k}, Y^n_{t_k}\right)\right||X^n_{t_k}\right]\right)\left(t-t_k\right),\end{aligned}$$

so we can choose $\eta_1\left(x,y,z\right) = K_X\left(1+|x|+|y|+|z|\right)$ for $(x,y,z) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q$, $K_1 = K_X$ and $\alpha_1 = 1$. For $l \geq 2$, we have that

$$\left|\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^l_j |X^n_{t_k} = x, Y^n_{t_k} = y\right]\right| \leq \mathbb{E}\left[\left|\left(\delta X^n_{t,k}\right)_j\right|^l |X^n_{t_k} = x, Y^n_{t_k} = y\right]$$

Expanding the term $\left|\left(\delta X^n_{t,k}\right)_j\right|^l$ and using (SL), we have that

$$\left|\left(\delta X^n_{t,k}\right)_j\right|^l \leq \sum_{i=0}^{l}\binom{l}{i}\left(K_X\left(1+|X^n_{t_k}|+|Y^n_{t_k}|+\mathbb{E}\left[|\phi\left(X^n_{t_k},Y^n_{t_k}\right)||X^n_{t_k}\right]\right)\right)^l (t-t_k)^{l-i}\left(\sum_{p=1}^{d_1}|W^1_{t,p}-W^1_{t_k,p}|\right)^i.$$

For $i\geq 1$, as there exists $c_i>0$ such that $\mathbb{E}\left[|W^1_{t,p}-W^1_{t_k,p}|^i\right]=c_i\,(t-t_k)^{\frac{i}{2}}$ for $1\leq p\leq d_1$, and using the independence between $W^1_t-W^1_{t_k}$ and the $\left(X^n_{t_k},Y^n_{t_k}\right)-$measurable terms, we obtain (3.2.10), with the choice $\eta_l\left(x,y,z\right)=\left(K_X\left(1+|x|+|y|+|z|\right)\right)^l\sum_{i=0}^{l}d_1^i c_i\binom{l}{i}T^{\frac{l-i}{2}}$, $K_l=4^{l-1}K_X^l$ and $\alpha_l=l$, since as for $i\leq l$, we have that $(t-t_k)^{\frac{l-i}{2}}\leq T^{\frac{l-i}{2}}$. □

We now prove Proposition 3.2.5. By Lemma 3.2.6, we have that for $0\leq k\leq n-1$,

$$\mathcal{E}_k=\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left\{\partial_t u+L^n_k u\right\}(t,X^n_t)\,dt\right].$$

By Proposition 3.2.4, $\partial_t u+\mathcal{L}u=0$, so the error term $\mathcal{E}_k$ rewrites

$$\begin{aligned}\mathcal{E}_k &= \mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left(L^n_k-\mathcal{L}\right)u\left(t,X^n_t\right)dt\right]\\ &= \mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left(\nabla_x u\right)^*\left(t,X^n_t\right)\left(b^n_{X,k}-b\left(t,X^n_t\right)\right)dt\right]\\ &+ \frac{1}{2}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\mathbf{Tr}\left\{\left(a^n_{X,k}-a\left(t,X^n_t\right)\right)\nabla^2_x u\left(t,X^n_t\right)\right\}dt\right]=:E_{1,k}+\frac{1}{2}E_{2,k}.\end{aligned}\tag{3.2.11}$$

We only show that there exists a finite constant κ that does not depend on the index k nor the discretization parameter n such that $|E_{2,k}|\leq\frac{\kappa}{n^2}$. The same result on $E_{1,k}$ holds with similar computations. We use the following decomposition,

$$\begin{aligned}\mathbf{Tr}\left\{\left(a^n_{X,k}-a\left(t,X^n_t\right)\right)\nabla^2_x u\left(t,X^n_t\right)\right\} &= \mathbf{Tr}\left\{\left(a^n_{X,k}-a\left(t_k,X^n_{t_k}\right)\right)\nabla^2_x u\left(t,X^n_t\right)\right\} &(3.2.12)\\ &+ \mathbf{Tr}\left\{\left(a\left(t_k,X^n_{t_k}\right)-a\left(t,X^n_t\right)\right)\nabla^2_x u\left(t,X^n_t\right)\right\} &(3.2.13)\\ &=: \mathcal{T}_1(t)+\mathcal{T}_2(t)\end{aligned}$$

To deal with the first term in the r.h.s. of the Equality (3.2.12), we use the Taylor expansion with integral remainder:

$$\begin{aligned}\nabla^2_x u\left(t,X^n_t\right) &= \nabla^2_x u\left(t,X^n_{t_k}\right)+\left(\delta X^n_{t,k}\right)^*\nabla^3_x u\left(t,X^n_{t_k}\right) &(3.2.14)\\ &+\int_0^1(1-s)\left(\delta X^n_{t,k}\right)^*\nabla^4_x u\left(t,X^n_{t_k}+s\delta X^n_{t,k}\right)\delta X^n_{t,k}ds. &(3.2.15)\end{aligned}$$

We replace $\nabla^2_x u\left(t,X^n_t\right)$ by each term in the r.h.s. of (3.2.14) in $\mathcal{T}_1$. For the first term, taking the expectation and using (3.1.5), we obtain that

$$\mathbb{E}\left[\mathbf{Tr}\left\{\left(a^n_{X,k}-a\left(t_k,X^n_{t_k}\right)\right)\nabla^2_x u\left(t,X^n_{t_k}\right)\right\}\right]=\mathbb{E}\left[\mathbf{Tr}\left\{\left(\mathbb{E}\left[a^n_{X,k}\middle|X^n_{t_k}\right]-a\left(t_k,X^n_{t_k}\right)\right)\nabla^2_x u\left(t,X^n_{t_k}\right)\right\}\right]=0.\tag{3.2.16}$$

For the second term in the r.h.s. of (3.2.14), using the tower property of the expectation,

$$\begin{aligned}&\left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a^n_{X,k}-a\left(t_k,X^n_{t_k}\right)\right)\left(\delta X^n_{t,k}\right)^*\nabla^3_x u\left(t,X^n_{t_k}\right)\right\}\right]\right| &(3.2.17)\\ &\leq\mathbb{E}\left[\mathbf{Tr}\left\{\left|\left(a^n_{X,k}-a\left(t_k,X^n_{t_k}\right)\right)\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^*\middle|X^n_{t_k},Y^n_{t_k}\right]\nabla^3_x u\left(t,X^n_{t_k}\right)\right|\right\}\right]\end{aligned}$$

By Lemma 3.2.7 we have that for $1\leq j\leq d_1$,

$$\left|\mathbb{E}\left[\left(\delta X^n_{t,k}\right)_j\middle|X^n_{t_k},Y^n_{t_k}\right]\right|\leq\mathbb{E}\left[\left|\eta_1\left(X^n_{t_k},Y^n_{t_k},\mathbb{E}\left[|\phi\left(X^n_{t_k},Y^n_{t_k}\right)||X^n_{t_k}\right]\right)\right|\right](t-t_k).$$

Moreover, as $\nabla_x^3 u$ has polynomial growth, by Proposition 3.2.4 and Assumption (SL), there exists four constants $p_1, p_2, p_3 \geq 1$ and $\tilde{K} > 0$ such that

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{Tr}\left\{\left|\left(a_{X,k}^n - a\left(t_k, X_{t_k}^n\right)\right)\mathbb{E}\left[\left(\delta X_{t,k}^n\right)^* \middle| X_{t_k}^n, Y_{t_k}^n\right]\nabla_x^3 u\left(t, X_{t_k}^n\right)\right|\right\}\right] \\
&\leq \mathbb{E}\left[\tilde{K}\left(1 + |X_{t_k}^n|^{p_1} + |Y_{t_k}^n|^{p_2} + \mathbb{E}\left[\left|\phi\left(X_{t_k}^n, Y_{t_k}^n\right)\right| |X_{t_k}^n|^{p_3}\right]\right)\right](t - t_k) \\
&\leq \mathbb{E}\left[\tilde{K}\left(1 + |X_{t_k}^n|^{p_1} + |Y_{t_k}^n|^{p_2} + \left|\phi\left(X_{t_k}^n, Y_{t_k}^n\right)\right|^{p_3}\right)\right](t - t_k), \qquad (3.2.18)
\end{aligned}$$

using Jensen's inequality for the last inequality, as $p_3 \geq 1$. Moreover, as ϕ has polynomial growth, the expectation in the r.h.s of (3.2.18) is bounded uniformly in k, n by Lemma 3.2.3. For the last term in the r.h.s. of (3.2.15), we show that there exists a constant β_1 independent from k, n such that

$$\left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a_{X,k}^n - a\left(t_k, X_{t_k}^n\right)\right)\int_0^1 (1-s)\left(\delta X_{t,k}^n\right)^* \nabla_x^4 u\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right)\delta X_{t,k}^n ds\right\}\right]\right| \leq \beta_1 (t - t_k). \qquad (3.2.19)$$

Indeed, as $\nabla_x^4 u$ has polynomial growth by Proposition 3.2.4, there exists $\hat{K} > 0$ and $p_4 \geq 1$ such that for $a, b, c, d \in \{1, ..., d_1\}$,

$$\left|\partial_{abcd}^4 u\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right)\right| \leq \hat{K}\left(1 + \left|X_{t_k}^n\right|^{p_4} + \left|\delta X_{t,k}^n\right|^{p_4}\right),$$

so that

$$\left|\left|\left(\delta X_{t,k}^n\right)^* \nabla_x^4 u\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right)\delta X_{t,k}^n\right|\right| \leq \hat{K}\left(\left|\delta X_{t,k}^n\right|^2 + \left|X_{t_k}^n\right|^{p_4}\left|\delta X_{t,k}^n\right|^2 + \left|\delta X_{t,k}^n\right|^{p_4+2}\right).$$

As (SL) is satisfied, there exists $q_1, q_2, q_3 \geq 1$ and $\Gamma > 0$ such that

$$\begin{aligned}
&\left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a_{X,k}^n - a\left(t_k, X_{t_k}^n\right)\right)\int_0^1 (1-s)\left(\delta X_{t,k}^n\right)^* \nabla_x^4 u\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right)\delta X_{t,k}^n ds\right\}\right]\right| \\
&\leq \Gamma\mathbb{E}\left[\left(1 + |X_{t_k}^n|^{q_1} + |Y_{t_k}^n|^{q_2} + \left|\phi\left(X_{t_k}^n, Y_{t_k}^n\right)\right|^{q_3}\right)\right](t - t_k),
\end{aligned}$$

by Lemma 3.2.7. By Lemma 3.2.3, we deduce the existence of a constant $K_1 > 0$ independent from k, n and such that

$$\left|\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \mathcal{T}_1(t)\, dt\right]\right| \leq \frac{K_1}{n^2}.$$

We now deal with the term $\mathbb{E}[\mathcal{T}_2]$ through the decomposition

$$|\mathbb{E}[\mathcal{T}_2(t)]| \leq \left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a(t_k, X_t^n) - a(t, X_t^n)\right)\nabla_x^2 u(t, X_t^n)\right\}\right]\right| + \left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a\left(t_k, X_{t_k}^n\right) - a(t_k, X_t^n)\right)\nabla_x^2 u(t, X_t^n)\right\}\right]\right|. \qquad (3.2.20)$$

For the first term in the r.h.s. of (3.2.20), as σ has bounded positive derivatives and therefore sublinear growth, there exists a polynomial $Q_2 : \mathbb{R} \to \mathbb{R}$ such that

$$|a(t, X_t^n) - a(t_k, X_t^n)| \leq Q_2(|X_t^n|)(t - t_k).$$

As $\nabla_x^2 u$ also has polynomial growth, we conclude by Lemma 3.2.3 that there exists a constant β_2 independent from k, n such that

$$\left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a(t, X_t^n) - a(t_k, X_t^n)\right)\nabla_x^2 u(t, X_t^n)\right\}\right]\right| \leq \beta_2 (t - t_k).$$

For the second term in the r.h.s. of (3.2.20), we apply the Taylor expansion on $(a(t_k, X_t^n))_{t \in [t_k, t_{k+1}]}$ and $\left(\nabla_x^2 u(t, X_t^n)\right)_{t \in [t_k, t_{k+1}]}$:

$$\begin{aligned}
a(t_k, X_t^n) - a\left(t_k, X_{t_k}^n\right) &= \left(\delta X_{t,k}^n\right)^* \nabla_x a\left(t_k, X_{t_k}^n\right) \\
&+ \int_0^1 (1-s)\left(\delta X_{t,k}^n\right)^* \nabla_x^2 a\left(t_k, X_{t_k}^n + s\delta X_{t,k}^n\right)\delta X_{t,k}^n ds, \qquad (3.2.21) \\
\nabla_x^2 u(t, X_t^n) &= \nabla_x^2 u\left(t, X_{t_k}^n\right) + \int_0^1 \left(\delta X_{t,k}^n\right)^* \nabla_x^3 u\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right) ds. \qquad (3.2.22)
\end{aligned}$$

We replace the term $a\left(t_k, X_t^n\right) - a\left(t_k, X_{t_k}^n\right)$ (resp. $\nabla_x^2 u\left(t, X_t^n\right)$) by its expression from the r.h.s. of (3.2.21) (resp. 3.2.22) in the product $\left(a\left(t_k, X_t^n\right) - a\left(t_k, X_{t_k}^n\right)\right) \nabla_x^2 u\left(t, X_t^n\right)$. The only term of order 1 in $\delta X_{t,k}^n$ is multiplied by a $X_{t_k}^n$-measurable term. Taking the trace and then the expectation, we obtain the existence of β_3 independent from k, n such that

$$\left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(\delta X_{t,k}^n\right)^* \nabla_x a\left(t_k, X_{t_k}^n\right) \nabla_x^2 u\left(t, X_{t_k}^n\right)\right\}\right]\right| \le \beta_3 (t - t_k).$$

As the other terms that appear from the product $\left(a\left(t_k, X_t^n\right) - a\left(t_k, X_{t_k}^n\right)\right) \nabla_x^2 u\left(t, X_t^n\right)$ are of order larger than 1 in $\delta X_{t,k}^n$, we deduce similarly to (3.2.19) the existence of β_4, independent from k, n and such that

$$\left|\mathbb{E}\left[\mathbf{Tr}\left\{\left(a\left(t_k, X_t^n\right) - a\left(t_k, X_{t_k}^n\right)\right) \nabla_x^2 u\left(t, X_t^n\right)\right\}\right]\right| \le \beta_4 \left(t - t_k\right). \tag{3.2.23}$$

We then deduce the existence of K_2, independent from k, n such that $\left|\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \mathcal{T}_2\left(t\right) dt\right]\right| \le \frac{K_2}{n^2}$ and this concludes the proof of Proposition 3.2.5.

Comments on further expansions of the weak error. While we are able to bound the weak error by taking advantage of the property (3.1.5), it seems difficult to obtain an expansion of the weak error, even at order 1. This is due to the lack of a priori information on any relationship similar to (3.1.5) for products of the coefficients of the diffusion nonlinear in the sense of McKean. To illustrate this difficulty, let us consider the error term $E_{2,k}$ in the Equation (3.2.11), for $0 \le k \le n-1$. For notational simplicity, let us also assume that $d_1 = d_2 = q = 1$.

$$2E_{2,k} = \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a\left(t_k, X_{t_k}^n\right) - a\left(t, X_t^n\right)\right) \partial_x^2 u\left(t, X_t^n\right) dt\right] + \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right) \partial_x^2 u\left(t, X_t^n\right) dt\right].$$

Let us now assume that b, σ, φ have C^∞ regularity and that all their derivatives are bounded, so that u also belongs to C^∞ with bounded derivatives. Let us denote by Π_k the second term in the r.h.s.. In order to replace the term $\partial_x^2 u\left(t, X_t^n\right)$ by $\partial_x^2 u\left(t_k, X_{t_k}^n\right)$ in Π_k, we apply Ito's formula on $\partial_x^2 u$:

$$\begin{aligned}
\Pi_k &= \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right) \partial_x^2 u\left(t_k, X_{t_k}^n\right) dt\right] && (3.2.24)\\
&- \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right) \left(\int_{t_k}^{t} \left(2\partial_x b \partial_x^2 u + \partial_x^2 b \partial_x u + \partial_x a \partial_x^3 u + \frac{1}{2}\partial_x^2 a \partial_x^2 u\right)\left(s, X_s^n\right) ds\right) dt\right] && (3.2.25)\\
&+ \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right) \left(\int_{t_k}^{t} \left(b_{X,t_k}^n - b\left(s, X_s^n\right)\right) \partial_x^3 u\left(s, X_s^n\right) ds\right) dt\right] && (3.2.26)\\
&+ \frac{1}{2}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right) \left(\int_{t_k}^{t} \left(a_{X,t_k}^n - a\left(s, X_s^n\right)\right) \partial_x^4 u\left(s, X_s^n\right) ds\right) dt\right] && (3.2.27)
\end{aligned}$$

where the stochastic integral that appears from Ito's formula has null expectation and where we used the fact that $\partial_t \partial_x^2 u = -\partial_x^2 \mathcal{L} u$. The first term on the r.h.s of (3.2.24) is equal to zero due to (3.1.5). Taking advantage of the regularity of the function u, one would iterate Ito's formula in order to replace the position $\left(s, X_s^n\right)$ by $\left(t_k, X_{t_k}^n\right)$ whenever it appears in the expression of Π_k, which respectively gives at the lowest order for (3.2.25)-(3.2.26)-(3.2.27):

$$\begin{aligned}
&\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right) \left(\int_{t_k}^{t} \left(2\partial_x b \partial_x^2 u + \partial_x^2 b \partial_x u + \partial_x a \partial_x^3 u + \frac{1}{2}\partial_x^2 a \partial_x^4 u\right)\left(t_k, X_{t_k}^n\right) ds\right) dt\right],\\
&\frac{T^2}{2n^2}\mathbb{E}\left[\left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right)\left(b_{X,t_k}^n - b\left(t_k, X_{t_k}^n\right)\right) \partial_x^3 u\left(t_k, X_{t_k}^n\right)\right], && (3.2.28)\\
&\frac{T^2}{4n^2}\mathbb{E}\left[\left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right)^2 \partial_x^4 u\left(t_k, X_{t_k}^n\right)\right]. && (3.2.29)
\end{aligned}$$

Due to (3.1.5), the first expectation is equal to zero. It is however more difficult to assess whether the terms

$$n\sum_{k=0}^{n-1} \frac{T^2}{2n^2}\mathbb{E}\left[\left(a_{X,t_k}^n - a\left(t_k, X_{t_k}^n\right)\right)\left(b_{X,t_k}^n - b\left(t_k, X_{t_k}^n\right)\right) \partial_x^3 u\left(t_k, X_{t_k}^n\right)\right]$$

$$n\sum_{k=0}^{n-1}\frac{T^2}{4n^2}\mathbb{E}\left[\left(a^n_{X,t_k}-a\left(t_k,X^n_{t_k}\right)\right)^2\partial_x^4 u\left(t_k,X^n_{t_k}\right)\right] \tag{3.2.30}$$

admit a limit when $n\to\infty$. In the LSV model, a natural limit for the last term (3.2.30) would be

$$\frac{T}{4}\int_0^T\sigma^2_{Dup}\left(t,X_t\right)\mathbb{E}\left[\left(\frac{f^2\left(Y_t\right)}{\mathbb{E}\left[f^2\left(Y_t\right)|X_t\right]}-1\right)^2\partial_x^4 u\left(t,X_t\right)\middle|X_t\right]dt,$$

where (X,Y) is a solution to the SDE (3.1.1), but to our knowledge, no existence result is available in this framework so far. Thus, we are not able to obtain a limit of $n\left(\sum_{k=0}^{n-1}\Pi_k\right)$ as $n\to\infty$, nor an expansion of the weak error with our methodology.

3.2.3 Proof of Theorem 3.2.2

Let us denote by $\underline{\sigma},\overline{\sigma}$ two constants such that for $(t,x,y,z)\in[0,T]\times\mathbb{R}^3$,

$$0<\underline{\sigma}\le\sigma\left(t,x\right)\wedge\sigma_X\left(t,x,y,z\right)\le\sigma\left(t,x\right)\vee\sigma_X\left(t,x,y,z\right)\le\overline{\sigma}.$$

We also define $\bar{b}:=||b_X||_\infty\vee||b||_\infty$. We consider in what follows the function $\varphi:x\in\mathbb{R}\to\left(K-e^x\right)_+$, where x represents the log-spot. Here the terminal condition is less regular than in the framework of Theorem 3.2.1, so Proposition 3.2.4 does not apply to give polynomial growth to the function u and its derivatives. To overcome this difficulty, we first show the result when we regularise the payoff by convolution w.r.t. G_ϵ, for $\epsilon\in(0,1)$, where we recall G_ϵ is the density of the centered normal law with variance ϵ.

Proposition 3.2.8. *Under Assumptions (A'), (B) and (IC), there exists a constant $\zeta>0$ that such that*

$$\forall n\ge 2,\forall\epsilon\in(0,1),\ \left|\mathbb{E}\left[\varphi*G_\epsilon\left(X^n_T\right)\right]-\mathbb{E}\left[\varphi*G_\epsilon\left(Z_T\right)\right]\right|\le\zeta\frac{\log\left(n\right)}{n}. \tag{3.2.31}$$

If Proposition 3.2.8 holds, in order to prove Theorem 3.2.2, it is sufficient to let $\epsilon\to0$ in Inequality (3.2.31). Indeed we have that $||\varphi*G_\epsilon||_\infty\le K$, and as φ is Lipschitz, the function $\varphi*G_\epsilon$ converges uniformly to φ as $\epsilon\to0$, so we have that for $n\ge2$, $\mathbb{E}\left[\varphi*G_\epsilon\left(X^n_T\right)\right]\to\mathbb{E}\left[\varphi\left(X^n_T\right)\right]$ and $\mathbb{E}\left[\varphi*G_\epsilon\left(X_T\right)\right]\to\mathbb{E}\left[\varphi\left(X_T\right)\right]$ as $\epsilon\to0$. Finally, as the r.h.s. of Inequality (3.2.31) does not depend on ϵ, we obtain (3.2.2) and this concludes the proof.

Estimates

We first give a Gaussian upper bound for the density of the Euler scheme, as a consequence of [58, Theorem 4.6].

Proposition 3.2.9. *There exist positive constants K_A,α_A such that for any $x,y\in\mathbb{R}$ and any $0<t\le T$, one has*

$$p^n_x\left(t,y\right)\le K_A G_{\alpha_A t}(x-y),$$

where $p^n_x\left(t,y\right)$ is the density at y of the random variable X^n_t when $X^n_0=x$.

Proof. The Euler scheme X^n follows the dynamics

$$dX^n_t=b_X\left(\tau_t,X^n_{\tau_t},Y^n_{\tau_t},\mathbb{E}\left[\phi\left(X^n_{\tau_t},Y^n_{\tau_t}\right)|X^n_{\tau_t}\right]\right)dt+\sigma_X\left(\tau_t,X^n_{\tau_t},Y^n_{\tau_t},\mathbb{E}\left[\phi\left(X^n_{\tau_t},Y^n_{\tau_t}\right)|X^n_{\tau_t}\right]\right)dW^1_t.$$

Let us define, for $t\ge0$ and $x\in\mathbb{R}$,

$$\begin{aligned}\tilde{b}\left(t,x\right)&=\mathbb{E}\left[b_X\left(t,X^n_{\tau_t},Y^n_{\tau_t},\mathbb{E}\left[\phi\left(X^n_{\tau_t},Y^n_{\tau_t}\right)\middle|X^n_{\tau_t}\right]\right)\middle|X^n_t=x\right],\\ \tilde{a}\left(t,x\right)&=\mathbb{E}\left[a_X\left(t,X^n_{\tau_t},Y^n_{\tau_t},\mathbb{E}\left[\phi\left(X^n_{\tau_t},Y^n_{\tau_t}\right)\middle|X^n_{\tau_t}\right]\right)\middle|X^n_t=x\right].\end{aligned}$$

As b_X,σ_X are bounded and σ_X is uniformly elliptic, by [58, Theorem 4.6], we have that for $t\in[0,T]$, X^n_t has the same distribution as $\tilde{X}_t$, where $\tilde{X}$ is a solution to the SDE

$$d\tilde{X}_t=\tilde{b}\left(t,\tilde{X}_t\right)dt+\sqrt{\tilde{a}}\left(t,\tilde{X}_t\right)dW^1_t,$$

and we have that $\tilde{b},\tilde{\sigma}$ are bounded and $\tilde{\sigma}$ is elliptic. We conclude the proof by the Aronson estimates [7, Thm. 1]. □

Let us define $u_\epsilon(t,x) := \mathbb{E}\left[\varphi * G_\epsilon(Z_T)\middle| Z_t = x\right]$, for $(t,x) \in [0,T]\times\mathbb{R}$. We first give some rough estimates on $\partial_x^k u_\epsilon$ for $k = 0, ..., 4$.

Lemma 3.2.10. *For $\epsilon > 0$, the function u^ϵ belongs to $C^{1,6}([0,T)\times\mathbb{R})$ and is the solution to the Cauchy problem*

$$\begin{aligned} \partial_t u^\epsilon + \mathcal{L}u^\epsilon &= 0, && (3.2.32)\\ u^\epsilon(T,\cdot) &= \varphi * G_\epsilon(\cdot). && (3.2.33)\end{aligned}$$

Moreover, there exists K_ϵ and p_ϵ such that for $0 \le k \le 6$, $(t,x) \in [0,T)\times\mathbb{R}$, $|\partial_x^k u^\epsilon(t,x)| \le K_\epsilon(1+|x|^{p_\epsilon})$.

Proof. The coefficients b, σ of the operator $\mathcal{L}$ are bounded and regular, σ is uniformly elliptic and $\varphi * G_\epsilon$ and its derivatives have polynomial growth, so we have by [43, Ch. 5, Thm. 6.1] that u_ϵ is the solution to the Cauchy problem (3.2.32)-(3.2.33). Moreover by [43, Ch. 5, Thm. 5.5] and [43, Ch. 5, Thm. 6.1], u^ϵ belongs to $C^{1,6}([0,T)\times\mathbb{R})$ and its derivatives satisfy the polynonial growth property. □

We derive below uniform in ϵ estimates on $\partial_x^k u_\epsilon$ for $k = 0, ..., 4$.

Proposition 3.2.11. *The function $\partial_x u^\epsilon$ is uniformly bounded, uniformly in $\epsilon \in (0,1)$, and for $2 \le k \le 4$, there exists $\alpha_k, \beta_k > 0$, independent from ϵ, such that for $(t,x) \in [0,T)\times\mathbb{R}$,*

$$|\partial_x^k u^\epsilon(t,x)| \le \frac{\alpha_k}{(T-t)^{\frac{k-2}{2}}}\left(1 + G_{\frac{T-t}{\beta_k}}(x - \ln(K))\right).$$

To prove Proposition 3.2.11, let us successively compute the derivatives of Equations (3.2.32)-(3.2.33). We obtain that for $1 \le k \le 4$, and $\epsilon > 0$, $v_k = \partial_x^k u_\epsilon$ satisfies the Cauchy problem

$$\begin{aligned} \partial_t v_k + \mathcal{L}_k v_k &= f^{\epsilon,k} && (3.2.34)\\ v(T,\cdot) &= \partial_x^k(\varphi * G_\epsilon)(\cdot), && (3.2.35)\end{aligned}$$

where the operators $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \mathcal{L}_4$ are defined by

$$\begin{aligned}
\mathcal{L}_1 v &= \partial_x b v + \left(b + \frac{1}{2}\partial_x a\right)\partial_x v + \frac{1}{2}a\partial_x^2 v\\
\mathcal{L}_2 v &= \left(2\partial_x b + \frac{1}{2}\partial_x^2 a\right)v + (b + \partial_x a)\partial_x v + \frac{1}{2}a\partial_x^2 v\\
\mathcal{L}_3 v &= \left(3\partial_x b + \frac{3}{2}\partial_x^2 a\right)v + \left(b + \frac{3}{2}\partial_x a\right)\partial_x v + \frac{1}{2}a\partial_x^2 v\\
\mathcal{L}_4 v &= \left(4\partial_x b + 3\partial_x^2 a\right)v + (b + 2\partial_x a)\partial_x v + \frac{1}{2}a\partial_x^2 v,
\end{aligned}$$

for any function $v \in C^{0,2}([0,T)\times\mathbb{R})$ and

$$\begin{aligned}
f^{\epsilon,1} &= 0\\
f^{\epsilon,2} &= -\partial_x^2 b\partial_x u^\epsilon\\
f^{\epsilon,3} &= -\partial_x^3 b\partial_x u^\epsilon - \left(3\partial_x^2 b + \frac{1}{2}\partial_x^3 a\right)\partial_x^2 u^\epsilon,\\
f^{\epsilon,4} &= -\partial_x^4 b\partial_x u^\epsilon - \left(4\partial_x^3 b + \frac{1}{2}\partial_x^4 a\right)\partial_x^2 u^\epsilon - \left(6\partial_x^2 b + 2\partial_x^3 a\right)\partial_x^3 u^\epsilon.
\end{aligned}$$

For $1 \le k \le 4$, let Γ_k be the fundamental solution of the operator $\mathcal{L}_k$. As the coefficients in $\mathcal{L}_k$ are regular, bounded and a is elliptic, we have that by [42, Ch. 9, Thm. 7], for $0 \le q \le 2$, $0 \le t < s \le T$ and $x, y \in \mathbb{R}$,

$$\left|\partial_y^q \Gamma_k(t,x;s,y)\right| \le \frac{\Lambda_{q,k}}{(s-t)^{q/2}} G_{\frac{(s-t)}{\lambda_{q,k}}}(x-y), \qquad (3.2.36)$$

where $\Lambda_{q,k}, \lambda_{q,k}$, are positive constants. Let us establish the following representation.

Lemma 3.2.12. *For $1 \leq k \leq 4, \epsilon > 0$ and $(t,x) \in [0,T] \times \mathbb{R}$, we have that*

$$\begin{aligned}
\partial_x^k u^\epsilon(t,x) &= \int_{\mathbb{R}} \partial_x^k(\varphi * G_\epsilon)(y)\,\Gamma_k(t,x;T,y)\,dy - \int_{\mathbb{R}}\int_t^T f^{\epsilon,k}(s,y)\,\Gamma_k(t,x;s,y)\,dy \qquad (3.2.37)\\
&=: \mathcal{M}^{\epsilon,k}(t,x) - \mathcal{S}^{\epsilon,k}(t,x). \qquad (3.2.38)
\end{aligned}$$

Proof. The Cauchy problem (3.2.34)-(3.2.35) is solved for $1 \leq k \leq 4$ by $v = \partial_x^k u^\epsilon$. Moreover, by Lemma 3.2.10, $\partial_x^k u^\epsilon$ has polynomial growth. As by Corollary [43, Ch. 6, Cor. 4.4] there exists a unique solution to the Cauchy problem (3.2.34)-(3.2.35) with polynomial growth, it is sufficient to check that the r.h.s. of Equality (3.2.37) satisfies the same properties. By direct derivation under the integral sign it is easy to check that Equations (3.2.34)-(3.2.35) are satisfied for the choice $v_k = \mathcal{M}^{\epsilon,k} - \mathcal{S}^{\epsilon,k}$. Let us show that $\mathcal{M}^{\epsilon,k}$ and $\mathcal{S}^{\epsilon,k}$ have polynomial growth in the space variable. The function φ is bounded, so $\partial_x^k(\varphi * G_\epsilon) = \varphi * \partial_x^k G_\epsilon$ is uniformly bounded on $\mathbb{R}$, as $\partial_x^k G_\epsilon$ is integrable. Therefore $\mathcal{M}^{\epsilon,k}$ is bounded uniformly on $(t,x) \in [0,T] \times \mathbb{R}$. Moreover, as $\left(\partial_x^k u^\epsilon\right)_{1 \leq k \leq 4}$ has polynomial growth and the derivatives of b and a are bounded, $\left(f^{\epsilon,k}\right)_{1 \leq k \leq 4}$ has polynomial growth. Using Inequality (3.2.36), we have that for $l \in \mathbb{N}$,

$$\int_{\mathbb{R}} |y|^l \Gamma_k(t,x;T,y)\,dy \leq \int_{\mathbb{R}} |y|^l \Lambda_{0,k} G_{\frac{(s-t)}{\lambda_{0,k}}}(x-y) \leq \sum_{i=0}^{l} \binom{l}{i} \Lambda_{0,k} V_i |x|^i, \qquad (3.2.39)$$

where $V_i := \int_{\mathbb{R}} |y-x|^{l-i} G_{\frac{(s-t)}{\lambda_{0,k}}}(x-y)dy \leq C_i \left(\frac{s-t}{\lambda_{0,k}}\right)^{\frac{l-i}{2}} \leq C_i \left(\frac{T}{\lambda_{0,k}}\right)^{\frac{l-i}{2}}$ for $0 \leq i \leq l$, and C_i is a universal constant. It is then sufficient to integrate both sides of the Inequality (3.2.39) over the interval (t,T) to obtain that $\mathcal{S}^{\epsilon,k}$ has polynomial growth in the space variable and conclude the proof. $\square$

We now derive estimates on the functions $\mathcal{M}^{\epsilon,k}, \mathcal{S}^{\epsilon,k}$, for $1 \leq k \leq 4$ and $\epsilon \in (0,1)$, that are independent of ϵ.

- For $k = 1$, we have that $\mathcal{S}^{\epsilon,1} = 0$ and as the first order derivative of $\partial_x \varphi$ in the sense of distributions is equal to $x \in \mathbb{R} \to -e^x 1_{\{e^x \leq K\}}$, we get that $||\partial_x(\varphi * G_\epsilon)||_\infty \leq K$. Therefore, using (3.2.36) for $q = 0$, $\mathcal{M}^{\epsilon,1}$ is uniformly bounded on $(t,x) \in [0,T) \times \mathbb{R}$ and uniformly in ϵ, by $K\Lambda_{0,1}$.

- For $k = 2$, we have that $f^{\epsilon,2}$ is uniformly bounded on $[0,T) \times \mathbb{R}$, uniformly in ϵ and so is $\mathcal{S}^{\epsilon,2}$. Moreover, for $(t,x) \in [0,T) \times \mathbb{R}$,
$$\mathcal{M}^{\epsilon,2}(t,x) \leq \int_{\mathbb{R}} \left|\partial_x^2(\varphi * G_\epsilon)(y)\right| \Gamma_2(t,x;T,y)dy.$$
As $\partial_x^2 \varphi$ is equal to the function $x \in \mathbb{R} \to -e^x 1_{\{e^x \leq K\}} + K\delta_{\ln K}(x)$ in the sense of distributions, where $\delta_{\ln K}$ is the Dirac measure on the point $\ln K$, we have that
$$\forall y \in \mathbb{R},\ \left|\partial_x^2(\varphi * G_\epsilon(y))\right| \leq K\left(1 + G_\epsilon(y - \ln K)\right). \qquad (3.2.40)$$
We obtain that $\int_{\mathbb{R}} K\Gamma_2(t,x;T,y)\,dy$ is uniformly bounded on $[0,T) \times \mathbb{R}$ by $K\Lambda_{0,2}$ and
$$\int_{\mathbb{R}} G_\epsilon(y - \ln K)\, G_{\frac{T-t}{\lambda_{0,2}}}(x-y)\,dy \leq \sqrt{\frac{\lambda_{0,2}}{2\pi(T-t)}} \exp\left(-\frac{(x - \ln(K))^2}{2\frac{T-t}{\lambda_{0,2}} + 2}\right) dy, \qquad (3.2.41)$$
as the l.h.s. of the inequality is the density at the point $x - \ln(K)$ of a centered Gaussian variable with variance $\frac{T-t}{\lambda_{0,2}} + \epsilon$, hence the result for $k = 2$ and $\epsilon \in (0,1)$.

- For $k = 3$, to study $\mathcal{M}^{\epsilon,3}$, we use the integration by parts formula:
$$\mathcal{M}_t^{\epsilon,3} = -\int_{\mathbb{R}} \partial_x^2(\varphi * G_\epsilon)(y)\,\partial_y \Gamma_3(t,x;T,y)\,dy.$$
Using Inequalities (3.2.36), (3.2.40) and the same arguments as in the case $k = 2$, there exists constants $\alpha_{3,1}, \beta_{3,1} > 0$, independent from ϵ, such that
$$|\mathcal{M}_t^{\epsilon,3}| \leq \frac{\alpha_{3,1}}{\sqrt{T-t}} \left(1 + G_{\frac{T-t}{\beta_{3,1}}}(x - \ln(K))\right). \qquad (3.2.42)$$

To control $\mathcal{S}_t^{\epsilon,3}$, it is sufficient to bound $\tilde{\mathcal{S}}_t^{\epsilon,3} := \int_{\mathbb{R}}\int_t^T |\partial_x^2 u^\epsilon(s,y)|\Gamma_3(t,x;s,y)dy$, as the terms $\partial_x^3 b\partial_x u^\epsilon$ and $3\partial_x^2 b + \frac{1}{2}\partial_x^3 a$ are bounded. To do so, we use the bound on $\partial_x^2 u^\epsilon$ derived previously, and we obtain

$$\left|\tilde{\mathcal{S}}_t^{\epsilon,3}\right| \leq \int_t^T \int_{\mathbb{R}} \alpha_2\Lambda_{0,3}\left(1+G_{\frac{T-s}{\beta_2}}(y-\ln(K))\right)G_{\frac{s-t}{\lambda_{0,3}}}(x-y)dsdy \leq \alpha_2\Lambda_{0,3}\left(T+\sqrt{\frac{2}{\pi}\beta_2 T}\right),$$

where we used the fact that $G_{\frac{T-s}{\beta_2}}(y-\ln(K)) \leq \sqrt{\frac{\beta_2}{2\pi(T-s)}}$ for $y\in\mathbb{R}$ and for $x\in\mathbb{R}$, $\int_{\mathbb{R}} G_{\frac{s-t}{\lambda_{0,3}}}(x-y)dy = 1$. Combining the upper bounds on $\mathcal{M}^{\epsilon,3}$ and $\mathcal{S}^{\epsilon,3}$, we obtain the result for $k=3$.

- For $k=4$, also use the integration by parts $\mathcal{M}_t^{\epsilon,4} = \int \partial_x^2(\varphi * G_\epsilon)(y)\,\partial_y^2\Gamma_4(t,x;T,y)\,dy$, and with the same arguments used to obtain Inequality (3.2.42), there exist constants $\alpha_{4,1},\beta_{4,1}$, independent from ϵ such that
$$|\mathcal{M}_t^{\epsilon,4}| \leq \frac{\alpha_{4,1}}{T-t}\left(1+G_{\frac{T-t}{\beta_{4,1}}}(x-\ln(K))\right).$$
Finally, to estimate $\mathcal{S}_t^{\epsilon,4}$, it is sufficient to bound $\tilde{\mathcal{S}}_t^{\epsilon,4} := \int_{\mathbb{R}}\int_t^T \partial_x^3 u^\epsilon(s,y)\Gamma_4(t,x;s,y)dy$, as the estimation of the other terms in $f^{\epsilon,4}$ involving $\partial_x u^\epsilon, \partial_x^2 u^\epsilon$ in $f^{\epsilon,4}$ can be performed as previously. As we have that
$$\int_{\mathbb{R}} G_{\frac{T-s}{\beta_3}}(y-\ln(K))G_{\frac{s-t}{\lambda_{0,4}}}(x-y)dy = G_{\frac{T-s}{\beta_3}+\frac{s-t}{\lambda_{0,4}}}(x-\ln(K)) \leq \frac{1}{\sqrt{2\pi}}\left(\frac{\max(\beta_3,\lambda_{0,4})}{T-t}\right)^{\frac{1}{2}}, \tag{3.2.43}$$
we obtain that
$$\left|\tilde{\mathcal{S}}_t^{\epsilon,4}\right| \leq \int_t^T \frac{\alpha_3\Lambda_{0,4}}{\sqrt{T-s}}\int_{\mathbb{R}}\left(1+G_{\frac{T-s}{\beta_3}}(y-\ln(K))\right)G_{\frac{s-t}{\lambda_{0,4}}}(x-y)dy \leq \alpha_3\Lambda_{0,4}\left(2\sqrt{T}+\sqrt{\frac{2}{\pi}\max(\beta_3,\lambda_{0,4})}\right),$$
and conclude the proof, gathering the upper bounds on $\mathcal{M}^{\epsilon,4}$ and $\tilde{\mathcal{S}}^{\epsilon,4}$.

Proof of Proposition 3.2.8

Similarly to the proof in Section 3.2.2, we have that

$$\mathbb{E}\left[\varphi * G_\epsilon(X_T^n)\right] - \mathbb{E}\left[\varphi * G_\epsilon(Z_T)\right] = \mathbb{E}\left[u^\epsilon(T,X_T^n)\right] - \mathbb{E}\left[u^\epsilon(0,X_0^n)\right] = \sum_{k=0}^{n-1}\mathcal{E}_k^\epsilon,$$

where

$$\mathcal{E}_k^\epsilon = \mathbb{E}\left[u^\epsilon\left(t_{k+1},X_{t_{k+1}}^n\right) - u^\epsilon\left(t_k,X_{t_k}^n\right)\right] = \mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left\{\partial_t u^\epsilon + L_k^n u^\epsilon\right\}(t,X_t^n)\,dt\right].$$

and the stochastic integral $\int_{t_k}^{t_{k+1}}\partial_x u(t,X_t^n)\,\sigma_{X,n}^n dW_t$ from Ito's formula has zero expectation due to the fact that σ_X is bounded and by Lemma 3.2.11, $\partial_x u^\epsilon$ is also bounded. By Lemma 3.2.10, we have that

$$\mathcal{E}_k^\epsilon = \mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left(b_{X,k}^n - b(t,X_t^n)\right)\partial_x u^\epsilon(t,X_t^n)\,dt\right] + \frac{1}{2}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left(a_{X,k}^n - a(t,X_t^n)\right)\partial_x^2 u^\epsilon(t,X_t^n)\,dt\right] =: E_{k,1}^\epsilon + \frac{1}{2}E_{k,2}^\epsilon$$

We only explain how to deal with the term $E_{k,2}^\epsilon$ for $0\leq k\leq n-1$, as the term $E_{k,1}^\epsilon$ is treated similarly. When $k=n-1$, by Proposition 3.2.11, $\left|\partial_x^2 u^\epsilon(t,x)\right| \leq \alpha_2\left(1+G_{\frac{T-t}{\beta_2}}(x-\ln(K))\right)$. Moreover, as $\left(a_{X,n-1}^n - a\right)$ is bounded uniformly in n, we obtain that for $n\geq 2$,

$$\left|E_{n-1,2}^\epsilon\right| \leq 2\overline{\sigma}^2\int_{T\left(1-\frac{1}{n}\right)}^T\int_{\mathbb{R}^2}\alpha_2\left(1+G_{\frac{T-t}{\beta_2}}(x-\ln(K))\right)K_A G_{\alpha_A t}(x-\xi)dx\mu_{Z_0}(d\xi)dt,$$

so that with computations similar to (3.2.43), we obtain that $\left|E_{n-1,2}^\epsilon\right| \leq 2\overline{\sigma}^2\alpha_2 K_A\left(1+\frac{1}{\sqrt{2\pi\min\left(\frac{1}{\beta_2},\alpha_A\right)T}}\right)\frac{T}{n}$.

We then study the case where $k<n-1$. We decompose $E_{k,2}^\epsilon$ in the following way,

$$E_{k,2}^\epsilon = \mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left(a_{X,k}^n - a\left(t_k,X_{t_k}^n\right)\right)\partial_x^2 u^\epsilon(t,X_t^n)\,dt + \int_{t_k}^{t_{k+1}}\left(a\left(t_k,X_{t_k}^n\right) - a(t,X_t^n)\right)\partial_x^2 u^\epsilon(t,X_t^n)\,dt\right]$$

$$= \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \mathcal{T}^{\epsilon}_{1,k}(t)dt + \int_{t_k}^{t_{k+1}} \mathcal{T}^{\epsilon}_{2,k}(t)dt\right]$$

We use the Taylor expansion with integral remainder on $\partial_x^2 u^\epsilon(t, X_t^n)$, similarly to (3.2.14), and we replace $\partial_x^2 u^\epsilon(t, X_t^n)$ by each term of the expansion in $E^\epsilon_{k,2}$. We first obtain that

$$\mathbb{E}\left[\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right)\partial_x^2 u^\epsilon\left(t, X^n_{t_k}\right)\right] = \mathbb{E}\left[\left(\mathbb{E}\left[a^n_{X,k}\middle| X^n_{t_k}\right] - a\left(t_k, X^n_{t_k}\right)\right)\partial_x^2 u^\epsilon\left(t, X^n_{t_k}\right)\right] = 0.$$

Then we have that

$$\begin{aligned}\left|\mathbb{E}\left[\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right)\delta X^n_{t,k}\partial_x^3 u^\epsilon\left(t, X^n_{t_k}\right)\right]\right| &\leq \mathbb{E}\left[\left|\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right)\mathbb{E}\left[\delta X^n_{t,k}\middle| X^n_{t_k}, Y^n_{t_k}\right]\partial_x^3 u^\epsilon\left(t, X^n_{t_k}\right)\right|\right] \\ &\leq 2\bar{b}\bar{\sigma}^2\mathbb{E}\left[\left|\partial_x^3 u^\epsilon\left(t, X^n_{t_k}\right)\right|\right](t - t_k).\end{aligned}$$

We now bound $\mathbb{E}\left[\left|\partial_x^3 u^\epsilon\left(t, X^n_{t_k}\right)\right|\right]$ uniformly in ϵ. By Propositions 3.2.11 and 3.2.9, we have that

$$\forall 0 \leq k \leq n-2, t \in [t_k, t_{k+1}),\ \mathbb{E}\left[\left|\partial_x^3 u^\epsilon\left(t, X^n_{t_k}\right)\right|\right] \leq \frac{M}{\sqrt{T-t}},$$

where

$$\begin{aligned}M &:= \int_{\mathbb{R}^2} \alpha_3\left(1 + G_{\frac{T-t}{\beta_3}}(x - \ln(K))\right) K_A G_{\alpha_A t_k}(x - \xi)dx\mu_{Z_0}(d\xi) \\ &\leq \alpha_3 K_A \int_{\mathbb{R}}\left(1 + G_{\frac{T-t}{\beta_3}+\alpha_A t_k}(\xi - \ln K)\right)\mu_{Z_0}(d\xi) \leq \alpha_3 K_A + \frac{\alpha_3 K_A}{\sqrt{2\pi\min\left(\frac{1}{\beta_3}, \alpha_A\right)T\left(1 - \frac{1}{n}\right)}}\end{aligned} \tag{3.2.44}$$

We then obtain that for $n \geq 2$,

$$\sum_{k=0}^{n-2}\int_{t_k}^{t_{k+1}}\left|\mathbb{E}\left[\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right)\delta X^n_{t,k}\partial_x^3 u^\epsilon\left(t, X^n_{t_k}\right)\right]\right|dt \leq 2\bar{b}\bar{\sigma}^2\ \frac{T}{n}\int_0^{t_{n-1}}\frac{M}{\sqrt{T-t}} \leq 4\bar{b}\bar{\sigma}^2 M\frac{T^{\frac{3}{2}}}{n}.$$

We study the integral remainder from the Taylor expansion replacing $\partial_x^2 u^\epsilon$ in $E^\epsilon_{k,2}$ to obtain a constant $\kappa_2 > 0$ such that

$$\left|\mathbb{E}\left[\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right)\int_0^1(1-s)\left(\delta X^n_{t,k}\right)^2\partial_x^4 u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right)ds\right]\right| \leq \kappa_2\frac{t - t_k}{T - t}. \tag{3.2.45}$$

As a_X, a are bounded, we have that

$$\left|\mathbb{E}\left[\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right)\int_0^1(1-s)\left(\delta X^n_{t,k}\right)^2\partial_x^4 u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right)ds\right]\right| \leq 2\bar{\sigma}^2\sup_{s\in[0,1]}\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2\left|\partial_x^4 u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right)\right|\right].$$

We then use the estimate on $\partial_x^4 u^\epsilon$ given by Proposition 3.2.11. For $s \in [0,1]$,

$$\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2\left|\partial_x^4 u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right)\right|\right] \leq \mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2\frac{\alpha_4}{T-t}\left(1 + G_{\frac{T-t}{\beta_4}}\left(X^n_{t_k} + s\delta X^n_{t,k} - \ln(K)\right)\right)\right] \tag{3.2.46}$$

We have that $\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2\right] \leq C(t - t_k)$ for a constant $C > 0$ which does not depend on ϵ, n, as the functions b_X, σ_X are bounded. We now study the second term in the r.h.s. of (3.2.46). Let $\beta \in \left(0, \min\left(\beta_4, \frac{1}{2\bar{\sigma}^2}\right)\right)$ and let us use the inequality for $x, z \in \mathbb{R}$, $-(x+z)^2 \leq -\frac{x^2}{2} + z^2$. We have that

$$\begin{aligned}J^n_k := &\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2 G_{\frac{T-t}{\beta_4}}\left(X^n_{t_k} + s\delta X^n_{t,k} - \ln(K)\right)\right] \leq \mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2\sqrt{\frac{\beta_4}{2\pi(T-t)}}\exp\left(-\beta\frac{\left(X^n_{t_k} + s\delta X^n_{t,k} - \ln(K)\right)^2}{2(T-t)}\right)\right] \\ &\leq \mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2\sqrt{\frac{\beta_4}{2\pi(T-t)}}\exp\left(-\frac{\beta}{4(T-t)}\left(X^n_{t_k} - \ln(K)\right)^2 + \frac{\beta s^2}{2(T-t)}\left(\delta X^n_{t,k}\right)^2\right)\right].\end{aligned}$$

As b_X, σ_X are bounded, we have that $\left|\delta X^n_{t,k}\right|^2 \leq 2\bar{b}^2 (t-t_k)^2 + 2\bar{\sigma}^2 \left(W^1_t - W^1_{t_k}\right)^2 =: \Delta X^2$ and as $W_t - W_{t_k}$ is independent from $X^n_{t_k}$, we have that

$$J^n_k \leq \mathbb{E}\left[\sqrt{\frac{\beta_4}{2\pi(T-t)}} \exp\left(-\frac{\beta}{4(T-t)}\left(X^n_{t_k} - \ln(K)\right)^2\right)\right] \mathbb{E}\left[\Delta X^2 \exp\left(\frac{\beta}{2(T-t)}\Delta X^2\right)\right], \tag{3.2.47}$$

for $s \in [0,1]$. By Proposition 3.2.9 and similarly to (3.2.44), we obtain the following bound for first term in the r.h.s. of (3.2.47),

$$\mathbb{E}\left[\sqrt{\frac{\beta_4}{2\pi(T-t)}} \exp\left(-\frac{\beta}{4(T-t)}\left(X^n_{t_k} - \ln(K)\right)^2\right)\right] \leq \sqrt{\frac{\beta_4}{\beta}} \frac{K_A}{\sqrt{\pi \min\left(\frac{2}{\beta}, \alpha_A\right) T \left(1-\frac{1}{n}\right)}}.$$

We now estimate the second term in the r.h.s. of (3.2.47) by

$$\begin{aligned}\mathbb{E}\left[\Delta X^2 \exp\left(\frac{\beta}{2(T-t)}\Delta X^2\right)\right] &\leq 2\bar{b}^2 \exp\left(\beta\bar{b}^2 \frac{(t-t_k)^2}{(T-t)}\right) \mathbb{E}\left[\exp\left(\beta\bar{\sigma}^2 \frac{(t-t_k)}{(T-t)} Q^2\right)\right](t-t_k)^2 \\ &+ 2\bar{\sigma}^2 \exp\left(\beta\bar{b}^2 \frac{(t-t_k)^2}{(T-t)}\right) \mathbb{E}\left[Q^2 \exp\left(\beta\bar{\sigma}^2 \frac{(t-t_k)}{(T-t)} Q^2\right)\right](t-t_k),\end{aligned} \tag{3.2.48}$$

where $Q \sim \mathcal{N}(0,1)$. As $t \leq t_{n-1}$, we have that $\frac{t-t_k}{T-t} \in [0,1)$, so $\exp\left(\beta\bar{b}^2 \frac{(t-t_k)^2}{(T-t)}\right) \leq \exp\left(\beta\bar{b}^2 \frac{T}{n}\right) \leq \exp\left(\beta\bar{b}^2 T\right)$. We now estimate the two expectations in the r.h.s. of (3.2.48), in uniformly in $t \in [0, t_{n-1}]$. For $k = 0, 2$, we have that

$$\begin{aligned}\mathbb{E}\left[Z^k \exp\left(\beta\bar{\sigma}^2 \frac{(t-t_k)}{(T-t)} Z^2\right)\right] &\leq \mathbb{E}\left[Z^k \exp\left(\beta\bar{\sigma}^2 Z^2\right)\right] \\ &\leq \frac{1}{\sqrt{2\pi}} \int z^k \exp\left(\left(\beta\bar{\sigma}^2 - \frac{1}{2}\right) z^2\right) dz.\end{aligned}$$

As $\beta\bar{\sigma}^2 < \frac{1}{2}$, we have that $\mathbb{E}\left[Z^k \exp\left(\beta\bar{\sigma}^2 Z^2\right)\right] < \infty$. This is sufficient to obtain (3.2.45). Integrating and summing the bounds over the index $1 \leq k < n-1$, we obtain that

$$\sum_{k=0}^{n-2} \int_{t_k}^{t_{k+1}} \left|\mathbb{E}\left[\left(a^n_{X,k} - a\left(t_k, X^n_{t_k}\right)\right) \int_0^1 (1-s)\left(\delta X^n_{t,k}\right)^2 \partial^4_x u\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right) ds\right]\right| \leq \kappa_2 \log(n) \frac{T}{n}.$$

Gathering the previous estimates, we conclude that there exists $\zeta_1 > 0$, independent from $n \geq 2$ and $\epsilon \in (0,1)$ and such that

$$\sum_{k=0}^{n-2} \left|\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \mathcal{T}^\epsilon_{1,k}(t)\, dt\right]\right| \leq \zeta_1 \frac{\log(n)}{n}.$$

The terms $\mathcal{T}^\epsilon_{2,k}$ for $0 \leq k \leq n-2$ are treated similarly to the terms $\mathcal{T}_{2,k}$ in the proof of Proposition 3.2.5, by introducing the Taylor expansion at the second and the first order on respectively a and $\partial^2_x u^\epsilon$, as in (3.2.21)-(3.2.22). As a has bounded derivatives, we obtain the following bounds on each term

$$\mathbb{E}\left[\delta X^n_{t,k} \partial_x a\left(t_k, X^n_{t_k}\right) \partial^2_x u^\epsilon\left(t, X^n_{t_k}\right)\right] \leq \bar{b}\, ||\partial_x a||_\infty \mathbb{E}\left[\left|\partial^2_x u^\epsilon\left(t, X^n_{t_k}\right)\right|\right](t-t_k),$$

$$\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2 \left(\int_0^1 (1-s)\partial^2_x a\left(t_k, X^n_{t_k} + s\delta X^n_{t,k}\right) ds\right) \partial^2_x u^\epsilon\left(t, X^n_{t_k}\right)\right] \leq \frac{||\partial^2_x a||_\infty}{2} \mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2 \partial^2_x u^\epsilon\left(t, X^n_{t_k}\right)\right],$$

$$\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2 \partial_x a\left(t_k, X^n_{t_k}\right) \int_0^1 \partial^3_x u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right) ds\right] \leq ||\partial_x a||_\infty \mathbb{E}\left[\left(\delta X^n_{t,k}\right)^2 \int_0^1 \partial^3_x u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right) ds\right],$$

$$\mathbb{E}\left[\left(\delta X^n_{t,k}\right)^3 \left(\int_0^1 (1-s)\partial^2_x a\left(t_k, X^n_{t_k} + s\delta X^n_{t,k}\right) ds\right) \int_0^1 \partial^3_x u^\epsilon\left(t, X^n_{t_k} + s\delta X^n_{t,k}\right) ds\right]$$

$$\leq \frac{||\partial_x^2 a||_\infty}{2}\mathbb{E}\left[\left(\delta X_{t,k}^n\right)^3 \int_0^1 \partial_x^3 u^\epsilon\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right) ds\right].$$

By Propositions 3.2.9 and 3.2.11, $\mathbb{E}\left[\left|\partial_x^2 u^\epsilon\left(t, X_{t_k}^n\right)\right|\right]$ is bounded uniformly in ϵ, t, k and $n \geq 2$ and there exists $\Pi_1 > 0$ independent of ϵ, t, k, and $n \geq 2$, such that

$$\mathbb{E}\left[\left(\delta X_{t,k}^n\right)^2 \partial_x^2 u^\epsilon\left(t, X_{t_k}^n\right)\right] \leq \Pi_1\left(t - t_k\right).$$

Moreover, with the same technique used to control J_k^n, we obtain the existence of Π_2 and Π_3 independent of ϵ, t, k such that

$$\mathbb{E}\left[\left(\delta X_{t,k}^n\right)^2 \int_0^1 \partial_x^3 u^\epsilon\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right) ds\right] \leq \Pi_2 \frac{t - t_k}{\sqrt{T-t}},$$
$$\mathbb{E}\left[\left(\delta X_{t,k}^n\right)^3 \int_0^1 \partial_x^3 u^\epsilon\left(t, X_{t_k}^n + s\delta X_{t,k}^n\right) ds\right] \leq \Pi_3 \frac{\left(t - t_k\right)^{\frac{3}{2}}}{\sqrt{T-t}} \leq \Pi_3\left(t - t_k\right),$$

as $t \leq t_{n-1}$. Therefore, after integration, we obtain $\zeta_2 > 0$, independent from n and ϵ such that

$$\sum_{k=0}^{n-2}\left|\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \mathcal{T}_{2,k}^\epsilon(t)\, dt\right]\right| \leq \frac{\zeta_2}{n},$$

so there exists $\zeta_3 > 0$ such that $\sum_{k=0}^{n-2}|E_{k,2}^\epsilon| \leq \zeta_3 \frac{\log(n)}{n}$, for $n \geq 2$. This is sufficient to obtain that there exists $\zeta > 0$ such that

$$\forall \epsilon \in (0,1), \forall n \geq 2, \sum_{k=0}^{n-1}|E_{k,2}^\epsilon| \leq \zeta \frac{\log(n)}{n}.$$

This concludes the proof of Proposition 3.2.8.

3.3 The interacting particles system

3.3.1 Half-step scheme and representation of the conditional expectation

In this section we assume that σ_X is uniformly elliptic, which means that there exists $\underline{\sigma} > 0$ such that $\left(a_X - \underline{\sigma}^2 I_{d_1}\right) \in \mathcal{S}_{d_1}^+(\mathbb{R})$. Let us introduce the half-step algorithm. Let $\left(Z_k^1, Z_{k+\frac{1}{2}}^1\right)_{k\geq 0}$, $\left(Z_k^2\right)_{k\geq 0}$ be two families of i.i.d. random variables with $\left(Z_0^1, Z_{\frac{1}{2}}^1\right) \sim \mathcal{N}_{2d_1}(0, I_{2d_1})$ and $\left(Z_0^2\right) \sim \mathcal{N}_{d_2}(0, I_{d_2})$ where for $m \geq 1$, $\mathcal{N}_m(0, I_m)$ is the standard centered Gaussian law in dimension m. The half-step algorithm is initialized with $\left(\hat{X}_0, \hat{Y}_0\right) \sim \mu_0$, independent of $\left(Z_k^1, Z_{k+\frac{1}{2}}^1\right)_{k\geq 0}$, $\left(Z_k^2\right)_{k\geq 0}$ and evolves inductively according to

$$\begin{aligned}
\hat{X}_{t_{k+\frac{1}{2}}}^n &= \hat{X}_{t_k}^n + \hat{b}_{X,t_k}^n \Delta + \left(\hat{a}_{X,t_k}^n - \underline{\sigma}^2 I_{d_1}\right)^{\frac{1}{2}} \sqrt{\Delta} Z_k^1, \\
\hat{X}_{t_{k+1}}^n &= \hat{X}_{t_{k+\frac{1}{2}}}^n + \underline{\sigma}\sqrt{\Delta} Z_{k+\frac{1}{2}}^1, \\
\hat{Y}_{t_{k+1}}^n &= \hat{Y}_{t_k}^n + b_Y\left(t_k, X_{t_k}^n, Y_{t_k}^n\right)\Delta + \sigma_Y\left(t_k, \hat{X}_{t_k}^n, \hat{Y}_{t_k}^n\right)\sqrt{\Delta} Z_k^2,
\end{aligned} \tag{3.3.1}$$

where we define for $n \geq 1$ and $0 \leq k \leq n-1$, $\hat{b}_{X,t_k}^n = b_X\left(t_k, \hat{X}_{t_k}^n, \hat{Y}_{t_k}^n, \mathbb{E}\left[\phi\left(\hat{X}_{t_k}^n, \hat{Y}_{t_k}^n\right)|\hat{X}_{t_k}^n\right]\right)$ and $\hat{a}_{X,k}^n = a_X\left(t_k, \hat{X}_{t_k}^n, \hat{Y}_{t_k}^n, \mathbb{E}\left[\phi\left(\hat{X}_{t_k}^n, \hat{Y}_{t_k}^n\right)|\hat{X}_{t_k}^n\right]\right)$.

As $\left(Z_k^1, Z_{k+\frac{1}{2}}^1, Z_k^2\right)$ and $\left(\hat{X}_{t_k}^n, \hat{Y}_{t_k}^n\right)$ are independent, the conditional law of the term

$$\left(\hat{a}_{X,t_k}^n - \underline{\sigma}^2 I_{d_1}\right)^{\frac{1}{2}} \sqrt{\Delta} Z_k^1 + \underline{\sigma}\sqrt{\Delta} Z_{k+\frac{1}{2}}^1,$$

w.r.t. $\left(\hat{X}_{t_k}^n, \hat{Y}_{t_k}^n\right)$ is the normal centered distribution with variance matrix $\Delta \hat{a}_{X,t_k}^n$, so the half step scheme (3.3.1) and the explicit Euler scheme (3.1.6) are equivalent in the sense that the vectors $\left(X_{t_k}^n, Y_{t_k}^n\right)_{0\leq k\leq n}$ and

$\left(\hat{X}^n_{t_k},\hat{Y}^n_{t_k}\right)_{0\le k\le n}$ have the same law. Splitting the Euler scheme defined in Section 3.2 and using the ellipticity of σ_X, the half-step algorithm helps us obtain a representation of the conditional expectation as a ratio of convolutions. For $\rho>0$ and $d\in\mathbb{N}^*$, let us define $G_{\rho,d}:x\in\mathbb{R}^d\to\frac{e^{-\frac{|x|^2}{2\rho}}}{(2\pi\rho)^{\frac{d}{2}}}$, the density of a centered gaussian variable with variance ρ in dimension d.

Proposition 3.3.1. *Let (ξ,Z,γ) be a random variable with values in $\mathbb{R}^{d_1}\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$. Let us assume that $Z\sim\mathcal{N}_{d_1}(0,I_{d_1})$ and is independent of (ξ,γ). Let $\alpha>0$ and let us define $\chi:=\xi+\alpha Z$. The following assertions hold:*

(i) for any measurable and bounded function $\psi:\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\to\mathbb{R}$, we have that

$$\mathbb{E}\left[\psi(\chi,\gamma)\right]=\int_{\mathbb{R}^{d_1}}\mathbb{E}\left[\psi(x,\gamma)\,G_{\alpha^2,d_1}(x-\xi)\right]dx.$$

(ii) Let $\left(\tilde{\xi},\tilde{\gamma}\right)$ be a copy of (ξ,γ) independent of χ. For $\psi:\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\to\mathbb{R}$ a measurable function such that $\psi(\chi,\gamma)$ is integrable, we have almost surely that

$$\mathbb{E}\left[\psi(\chi,\gamma)|\,\chi\right]=\frac{\mathbb{E}\left[\psi(\chi,\tilde{\gamma})\,G_{\alpha^2,d_1}\left(\chi-\tilde{\xi}\right)\Big|\,\chi\right]}{\mathbb{E}\left[G_{\alpha^2,d_1}\left(\chi-\tilde{\xi}\right)\Big|\,\chi\right]}.$$

Proof. Let $\psi:\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}\to\mathbb{R}$ be a measurable and bounded function. As Z is independent of (ξ,γ), we have by Fubini's theorem

$$\mathbb{E}\left[\psi(\chi,\gamma)\right]=\mathbb{E}\left[\psi(\xi+\alpha Z,\gamma)\right]=\mathbb{E}\left[\int_{\mathbb{R}^{d_1}}\psi(\xi+z,\gamma)\,G_{\alpha^2,d_1}(z)\,dz\right]=\mathbb{E}\left[\int_{\mathbb{R}^{d_1}}\psi(x,\gamma)\,G_{\alpha^2,d_1}(x-\xi)\,dx\right]$$

where we made the change of variable $x:=\xi+z$, and this proves (i).

By (i), we have that χ has the density w.r.t. the Lebesgue measure $p(x)=\mathbb{E}\left[G_{\alpha^2,d_1}(x-\xi)\right]$ for $x\in\mathbb{R}^{d_1}$. For $\Gamma:\mathbb{R}^{d_1}\to\mathbb{R}$ measurable and bounded, we have that

$$\begin{aligned}\mathbb{E}\left[\Gamma(\chi)\frac{\mathbb{E}\left[\psi(\chi,\tilde{\gamma})\,G_{\alpha^2,d_1}\left(\chi-\tilde{\xi}\right)\Big|\,\chi\right]}{\mathbb{E}\left[G_{\alpha^2,d_1}\left(\chi-\tilde{\xi}\right)\Big|\,\chi\right]}\right]&=\mathbb{E}\left[\frac{\Gamma(\chi)\,\psi(\chi,\tilde{\gamma})\,G_{\alpha^2,d_1}\left(\chi-\tilde{\xi}\right)}{p(\chi)}\right]\\&=\int_{\mathbb{R}^{d_1}}\Gamma(x)\,\mathbb{E}\left[\psi(x,\tilde{\gamma})\,G_{\alpha^2,d_1}\left(x-\tilde{\xi}\right)\right]dx\\&=\int_{\mathbb{R}^{d_1}}\mathbb{E}\left[\Gamma(x)\,\psi(x,\gamma)\,G_{\alpha^2,d_1}(x-\xi)\right]dx=\mathbb{E}\left[\Gamma(\chi)\,\psi(\chi,\gamma)\right],\end{aligned}$$

where we used the independence between $\left(\tilde{\xi},\tilde{\gamma}\right)$ and χ for the second equality and (i) for the last one, and this concludes the proof. □

By Proposition 3.3.1, setting $\xi=\hat{X}^n_{t_{k-\frac{1}{2}}}$, $Z=Z^1_{k-\frac{1}{2}}$, $\alpha=\underline{\sigma}\sqrt{\Delta}$ and $\chi=\hat{X}^n_{t_k}$, let us denote by $p^n_X(t_k,x)$ the density of $\hat{X}^n_{t_k}$ at the point $x\in\mathbb{R}^{d_1}$. The following result holds.

Lemma 3.3.2. *For $k\ge1$, let $\left(\tilde{X}^n_{t_{k-\frac{1}{2}}},\tilde{Y}^n_{t_k}\right)$ be a copy of $\left(\hat{X}^n_{t_{k-\frac{1}{2}}},\hat{Y}^n_{t_k}\right)$ and independent of $\hat{X}^n_{t_k}$. $\hat{X}^n_{t_k}$ admits the density $p^n_X(t_k,x)=\mathbb{E}\left[G_{\underline{\sigma}^2\Delta,d_1}\left(x-\hat{X}^n_{t_{k-\frac{1}{2}}}\right)\right]$ w.r.t. the Lebesgue measure. Moreover, the following representation holds*

$$\mathbb{E}\left[\phi\left(\hat{X}^n_{t_k},\hat{Y}^n_{t_k}\right)\Big|\,\hat{X}^n_{t_k}\right]=\frac{\mathbb{E}\left[\phi\left(\hat{X}^n_{t_k},\tilde{Y}^n_{t_k}\right)G_{\underline{\sigma}^2\Delta,d_1}\left(\hat{X}^n_{t_k}-\tilde{X}^n_{t_{k-\frac{1}{2}}}\right)\Big|\,\hat{X}^n_{t_k}\right]}{\mathbb{E}\left[G_{\underline{\sigma}^2\Delta,d_1}\left(\hat{X}^n_{t_k}-\tilde{X}^n_{t_{k-\frac{1}{2}}}\right)\Big|\,\hat{X}^n_{t_k}\right]}.\tag{3.3.2}$$

In the following section, we replace the joint law of $\left(\hat{X}^n_{t_k},\hat{Y}^n_{t_k}\right)$ that is involved in (3.3.2) by its empirical law, in the context of a particles system, as the schemes 3.1.6 and 3.3.1 are not directly implementable. Beforehand, let us slightly modifiy the Gaussian kernel in (3.3.2), so that the ratio in the r.h.s. satisfies the Lipschitz property w.r.t. $\hat{X}^n_{t_k}$.

3.3.2 Lipschitz modification and weak error

We modify the heat kernel, so that the ratio in the r.h.s. of (3.3.2) becomes a Lipschitz function of $\hat{X}^n_{t_k}$, under some regularity condition on the function ϕ that will be discussed in the following subsection. For $\alpha > 0, \rho \in (0,1)$ and $d \in \mathbb{N}^*$ let us define,

$$G_{\rho,d,\alpha} = \alpha + G_{\rho,d}.$$

Lemma 3.3.3. *Let $\alpha > 0$ and $\rho \in (0,1)$, then $||G_{\rho,d} - G_{\rho,d,\alpha}||_\infty \leq \alpha$ and there exists a constant K_d that only depends on d and such that*

$$||\nabla_x G_{\rho,d,\alpha}||_\infty \leq \frac{K_d}{\rho^{\frac{d+1}{2}}}.$$

Proof. It is sufficient to prove the second inequality. We have that for $x \in \mathbb{R}^d$,

$$|\nabla_x G_{\rho,d,\alpha}(x)| \leq \frac{|x|}{\rho} G_{\rho,d}(x) \leq \frac{1}{\rho} \times \frac{1}{(2\pi\rho)^{\frac{d-1}{2}}} \sup_{y\in\mathbb{R}} \{|y| G_{1,1}(y)\},$$

using the change of variable $y = \frac{|x|}{\sqrt{\rho}}$ and we conclude the proof with the choice $K_d = \frac{1}{(2\pi)^{\frac{d-1}{2}}} \sup_{y\in\mathbb{R}} \{|y| G_{1,1}(y)\}$. □

Let us introduce additional assumptions on the coefficients b, b_X, σ, σ_X.

Assumption (B'). *There exist constants $L, \kappa > 0$ such that for $t \geq 0$, and $(x, y, z_1, z_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times (\mathbb{R}^q)^2$,*

$$|b_X(t,x,y,z_1) - b_X(t,x,y,z_2)| \vee ||a_X(t,x,y,z_1) - a_X(t,x,y,z_2)|| \leq L|z_1 - z_2|,$$
$$|b_X(t,x,y,z_1)| \vee |b(t,x)| \leq \kappa\sqrt{1+|x|^2}.$$

Moreover, the coefficients σ_X, σ are bounded and there exists $\underline{\sigma} > 0$ such that $\sigma_X \geq \underline{\sigma}$. Finally, ϕ is bounded by a constant $\overline{\phi} > 0$.

Assumption (IC'). *There exists a constant $\lambda_0 > 0$ such that $\mathbb{E}\left[e^{\lambda_0|X_0|}\right] < \infty$.*

We define the modified half step scheme. It is initialized with $(\mathbf{X}_0, \mathbf{Y}_0) \sim \mu_0$, independent of $\left(Z^1_k, Z^1_{k+\frac{1}{2}}\right)_{k\geq 0}$, $\left(Z^2_k\right)_{k\geq 0}$ and evolves inductively according to

$$\begin{aligned} \mathbf{X}^n_{t_{k+\frac{1}{2}}} &= \mathbf{X}^n_{t_k} + \mathbf{b}^n_{X,k}\Delta + \left(\mathbf{a}^n_{X,t_k} - \sigma_0^2 I_{d_1}\right)^{\frac{1}{2}} \sqrt{\Delta} Z^1_k, \\ \mathbf{X}^n_{t_{k+1}} &= \mathbf{X}^n_{t_{k+\frac{1}{2}}} + \sigma_0\sqrt{\Delta} Z^1_{k+\frac{1}{2}}, \\ \mathbf{Y}^n_{t_{k+1}} &= \mathbf{Y}^n_{t_k} + b_Y\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)\Delta + \sigma_Y\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)\sqrt{\Delta} Z^2_k, \end{aligned} \tag{3.3.3}$$

where $\sigma_0 \in (0, \underline{\sigma})$, for $n \geq 1$, $0 \leq k \leq n-1$, $\mathbf{b}^n_{X,k} = b_X\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}, \mathbf{E}_k\left(\mathbf{X}^n_{t_k}\right)\right)$, $\mathbf{a}^n_{X,k} = a_X\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}, \mathbf{E}_k\left(\mathbf{X}^n_{t_k}\right)\right)$, and for $\alpha > 0$,

$$\mathbf{E}_k\left(\mathbf{X}^n_{t_k}\right) = \frac{\mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \tilde{\mathbf{Y}}^n_{t_k}\right) G_{\sigma_0^2\Delta, d_1, \alpha}\left(\mathbf{X}^n_{t_k} - \tilde{\mathbf{X}}^n_{t_{k-\frac{1}{2}}}\right) \middle| \mathbf{X}^n_{t_k}\right]}{\mathbb{E}\left[G_{\sigma_0^2\Delta, d_1, \alpha}\left(\mathbf{X}^n_{t_k} - \tilde{\mathbf{X}}^n_{t_{k-\frac{1}{2}}}\right) \middle| \mathbf{X}^n_{t_k}\right]}, \tag{3.3.4}$$

with $\left(\tilde{\mathbf{X}}^n_{t_{k-\frac{1}{2}}}, \tilde{\mathbf{Y}}^n_{t_k}\right)$ a copy of $\left(\mathbf{X}^n_{t_{k-\frac{1}{2}}}, \mathbf{Y}^n_{t_k}\right)$, independent of $\mathbf{X}^n_{t_k}$.

Under Assumptions (B) and (IC'), the existence of the process $(\mathbf{X}^n, \mathbf{Y}^n)$ is given by the boundedness of the coefficient σ_X, the sublinear growth of b_X and arguments similar to those used in the beginning of Section 3.2.2. Moreover, adapting the proofs of Lemmas 3.2.3 and 3.2.7, the two following results hold.

Lemma 3.3.4. *Under Assumptions (B) and (IC'), for any $p \geq 0$, the moments of $\mathbf{X}^n_{t_k}$ and $\mathbf{Y}^n_{t_k}$ at order p are bounded uniformly in $n \in \mathbb{N}^*$ and $0 \leq k \leq n-1$.*

Let us denote $\mathbf{X}^n_{t_{k+1}} - \mathbf{X}^n_{t_k}$ by $\mathbf{D}_k$ for $0 \leq k \leq n-1$.

Lemma 3.3.5. *For $l \in \mathbb{N}^*$, there exists a positive fonction $\eta_l : \mathbb{R}^{d_1} \to \mathbb{R}$ with polynomial growth such that for $n \in \mathbb{N}^*$, $0 \leq k \leq n-1$, $1 \leq j \leq d_1$, we have almost surely*

$$\left|\mathbb{E}\left[(\mathbf{D}_k)_j^l \,|\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right]\right| \leq \eta_l\left(\mathbf{X}^n_{t_k}\right)\Delta^{1\vee\frac{l}{2}}. \tag{3.3.5}$$

Proposition 3.3.6. *Under Assumptions $(A),(B),(IC')$, there exist two constants $\zeta, q > 0$ such that*

$$\forall n \geq 2,\ \left|\mathbb{E}\left[\varphi\left(\mathbf{X}^n_T\right)\right] - \mathbb{E}\left[\varphi\left(Z_T\right)\right]\right| \leq \zeta\left(\Delta + \frac{\alpha}{h} + h + h|\log(h)|^{d_1} + h|\log(h)|^{q+d_1}\right),$$

for all $h > 0$.

The proof of Proposition 3.3.6 relies on the two following lemmas, that we prove in the end of this section.

Lemma 3.3.7. *Under Assumptions (B') and (IC'), we have that*

$$\sup_{n\geq 1, 0\leq k\leq n} \mathbb{E}\left[e^{\lambda|\mathbf{X}^n_{t_k}|}\right] < \infty,$$

for $\lambda \leq \frac{\lambda_0}{2}e^{-\left(\kappa T + \kappa^2\frac{T}{2}\right)}$.

Lemma 3.3.8. *Let Z be a random variable with values in $\mathbb{R}^d$ and assume that there exists $\lambda > 0$ such that $\mathbb{E}\left[e^{\lambda|Z|}\right] < \infty$. If Z has a density p w.r.t. the Lebesgue measure, then we have the inequality*

$$\forall h > 0, \forall k \in \mathbb{N},\ \int_{\mathbb{R}^d} |x|^k p(x) 1_{\{p(x)\leq h\}} dx \leq h\left(\frac{k!2^k}{\lambda^k}\mathbb{E}\left[e^{\lambda|Z|}\right] + \frac{2\pi^{\frac{d}{2}}}{(k+d)\,\Gamma(\frac{d}{2})}\left(\frac{2\log(h)}{\lambda}\right)^{k+d}\right),$$

where Γ is the Gamma function defined by $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$ for $a > 0$.

We now prove Proposition 3.3.6. Let us use the Talay-Tubaro technique and write

$$\mathbb{E}\left[\varphi\left(\mathbf{X}^n_T\right)\right] - \mathbb{E}\left[\varphi\left(Z_T\right)\right] = \sum_{k=0}^{n-1}\mathfrak{E}_k,$$

where for $0 \leq k \leq n-1$, $\mathfrak{E}_k = \mathbb{E}\left[u\left(t_{k+1}, \mathbf{X}^n_{t_{k+1}}\right)\right] - \mathbb{E}\left[u\left(t_k, \mathbf{X}^n_{t_k}\right)\right]$, where we recall that the function u is defined by $u(t,x) = \mathbb{E}\left[\varphi(Z_T)|Z_t = x\right]$ for $(t,x) \in [0,T]\times\mathbb{R}^{d_1}$. By Proposition 3.2.4, we obtain that u belongs to $C^{2,4}([0,T]\times\mathbb{R})$ and write

$$\begin{aligned} u\left(t_{k+1}, \mathbf{X}^n_{t_{k+1}}\right) - u\left(t_k, \mathbf{X}^n_{t_k}\right) = {}& \Delta\partial_t u\left(t_k, \mathbf{X}^n_{t_k}\right) + \left(\nabla_x u\right)^*\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{D}^n_k \\ & + \frac{1}{2}\left(\mathbf{D}^n_k\right)^*\nabla_x^2 u\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{D}^n_k + \mathcal{R}_k, \end{aligned} \tag{3.3.6}$$

where the remainder term $\mathcal{R}_k$ is given by

$$\mathcal{R}_k = \Delta^2\int_0^1 (1-s)\,\partial_t^2 u\left(t_k + s\Delta, \mathbf{X}^n_{t_k} + s\mathbf{D}^n_k\right)ds \tag{3.3.7}$$

$$+ 2\Delta\int_0^1 (1-s)\,\partial_t\left(\nabla_x u\right)^*\left(t_k + s\Delta, \mathbf{X}^n_{t_k} + s\mathbf{D}^n_k\right)\mathbf{D}^n_k ds \tag{3.3.8}$$

$$+ \int_0^1 (1-s)\left(\mathbf{D}^n_k\right)^*\left(\nabla_x^2 u\left(t_k + s\Delta, \mathbf{X}^n_{t_k} + s\mathbf{D}^n_k\right) - \nabla_x^2 u\left(t_k, \mathbf{X}^n_{t_k}\right)\right)\mathbf{D}^n_k ds \tag{3.3.9}$$

By Proposition 3.2.4,

$$-\partial_t u\left(t_k, \mathbf{X}^n_{t_k}\right) = \left(\nabla_x u\right)^*\left(t_k, \mathbf{X}^n_{t_k}\right)b\left(t_k, \mathbf{X}^n_{t_k}\right) + \frac{1}{2}\mathbf{Tr}\left(a\left(t_k, \mathbf{X}^n_{t_k}\right)\nabla_x^2 u\left(t_k, \mathbf{X}^n_{t_k}\right)\right).$$

Moreover, conditioning w.r.t. $\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)$, we have that $\mathbb{E}\left[\left(\nabla_x u\right)^*\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{D}^n_k\right] = \Delta\mathbb{E}\left[\left(\nabla_x u\right)^*\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{b}^n_{X,k}\right]$ and

$$\mathbb{E}\left[\left(\mathbf{D}^n_k\right)^*\nabla_x^2 u\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{D}^n_k\right] = \Delta^2\mathbb{E}\left[\left(\mathbf{b}^n_{X,k}\right)^*\nabla_x^2 u\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{b}^n_{X,k}\right] + \Delta\mathbb{E}\left[\mathbf{Tr}\left(\mathbf{a}^n_{X,t_k}\nabla_x^2 u\left(t_k, \mathbf{X}^n_{t_k}\right)\right)\right].$$

We can now rewrite

$$\mathfrak{E}_k = \Delta \mathbb{E}\left[\left(\mathbf{b}^n_{X,k} - b\left(t_k, \mathbf{X}^n_{t_k}\right)\right)^* \nabla_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right] + \frac{\Delta}{2}\mathbb{E}\left[\mathbf{Tr}\left(\mathbf{a}^n_{X,t_k} - a\left(t_k, \mathbf{X}^n_{t_k}\right)\right)\nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right] + \tilde{\mathcal{R}}_k, \quad (3.3.10)$$

where for $0 \le k \le n-1$,

$$\tilde{\mathcal{R}}_k = \Delta^2 \mathbb{E}\left[\left(\mathbf{b}^n_{X,k}\right)^* \nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\mathbf{b}^n_{X,k}\right] + \mathbb{E}\left[\mathcal{R}_k\right].$$

We only show how to deal with the second term in (3.3.10). The first term is treated in the same way. We will study $\tilde{\mathcal{R}}_k$ in the following Lemma. As we have that a.s.,

$$a\left(t_k, \mathbf{X}^n_{t_k}\right) = \mathbb{E}\left[a_X\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}, \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]\right)|\mathbf{X}^n_{t_k}\right],$$

let us replace a by that representation in (3.3.10) in order to take advantage of the Lipschitz property given by Assumption (B), and study

$$\mathbb{E}\left[\mathbf{Tr}\left(\mathbf{a}^n_{X,t_k} - a\left(t_k, \mathbf{X}^n_{t_k}\right)\right)\nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right] = \mathbb{E}\left[\mathbf{Tr}\left(\mathbf{a}^n_{X,t_k} - a_X\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}, \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]\right)\right)\nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right].$$

By Assumption (B) and the fact that $\nabla^2_x u$ has polynomial growth in the space variable, there exist $K, p > 0$ such that

$$\begin{aligned}
&\left|\mathbb{E}\left[\mathbf{Tr}\left(\mathbf{a}^n_{X,t_k} - a_X\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}, \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]\right)\right)\nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right]\right| \\
&\le K\mathbb{E}\left[\left|\mathbf{E}_k\left(\mathbf{X}^n_{t_k}\right) - \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]\right|\left(1 + \left|\mathbf{X}^n_{t_k}\right|^p\right)\right].
\end{aligned}$$

Let us introduce the notation $\delta\mathbf{E}_k = \mathbf{E}_k\left(\mathbf{X}^n_{t_k}\right) - \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]$ for $0 \le k \le n-1$. Let $h > 0$ be a parameter to be fixed later and consider the event

$$A_h = \left\{\mathbb{E}\left[G_{\sigma_0^2\Delta}\left(\tilde{\mathbf{X}}^n_{t_{k-\frac{1}{2}}} - \mathbf{X}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right] > h\right\}.$$

By Proposition 3.3.1 and the definition (3.3.4),

$$|\delta\mathbf{E}_k|\,1_{\{A_h\}} \le \frac{2\bar{\phi}}{h}\left|\mathbb{E}\left[\left(G_{\sigma_0^2\Delta,d_1,\alpha} - G_{\sigma_0^2\Delta,d_1}\right)\left(\mathbf{X}^n_{t_k} - \tilde{\mathbf{X}}^n_{t_{k-\frac{1}{2}}}\right)\middle|\mathbf{X}^n_{t_k}\right]\right| \le 2\bar{\phi}\frac{\alpha}{h},$$

where we used Lemma 3.3.3 in the last inequality and the fact that the function ϕ is bounded. Hence the existence of $K > 0$ such that

$$\left|\mathbb{E}\left[\mathbf{Tr}\left(\left(\mathbf{a}^n_{X,t_k} - a_X\left(t_k, \mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}, \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]\right)\right)\nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right)1_{A_h}\right]\right| \le K\frac{\alpha}{h}. \quad (3.3.11)$$

Moreover, by Lemma 3.3.7, let $0 < \lambda \le \frac{\lambda_0}{2}e^{-\left(\kappa T + \frac{\kappa^2 T^2}{2}\right)}$ and let us define

$$\sup_{n\ge 1, 0\le k\le n}\mathbb{E}\left[e^{\lambda|\mathbf{X}^n_{t_k}|}\right] =: \eta.$$

We use the fact that $|\delta\mathbf{E}_k| \le 2\bar{\phi}$ and apply Lemma 3.3.8 to obtain that on A_h^c, as $x \in \mathbb{R}^{d_1} \to \mathbb{E}\left[G_{\sigma_0^2\Delta}\left(x - \tilde{\mathbf{X}}^n_{t_{k-\frac{1}{2}}}\right)\right]$ is the density of the variable $\mathbf{X}^n_{t_k}$ by Proposition 3.3.1,

$$\begin{aligned}
&\mathbb{E}\left[\left|\mathbf{E}_k\left(\mathbf{X}^n_{t_k}\right) - \mathbb{E}\left[\phi\left(\mathbf{X}^n_{t_k}, \mathbf{Y}^n_{t_k}\right)|\mathbf{X}^n_{t_k}\right]\right|\left(1 + \left|\mathbf{X}^n_{t_k}\right|^p\right)1_{A_h^c}\right] \le 2\bar{\phi}\mathbb{E}\left[\left(1 + \left|\mathbf{X}^n_{t_k}\right|^p\right)1_{A_h^c}\right] \\
&\le 2\bar{\phi}h\left(\left(1 + \frac{p!2^p}{\lambda^p}\right)\eta + \frac{2\pi^{\frac{d_1}{2}}}{\Gamma\left(\frac{d_1}{2}\right)}\left(\frac{2}{\lambda}|\log(h)|\right)^{d_1}\left(\frac{1}{d_1} + \frac{1}{d_1+p}\left(\frac{2}{\lambda}|\log(h)|\right)^p\right)\right). \quad (3.3.12)
\end{aligned}$$

Dealing in the same way with the first term in (3.3.10), we have the existence of $\zeta, q > 0$ such that

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(\mathbf{b}^n_{X,k} - b\left(t_k, \mathbf{X}^n_{t_k}\right)\right)^* \nabla_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right]\right| + \left|\mathbb{E}\left[\mathbf{Tr}\left(\mathbf{a}^n_{X,t_k} - a\left(t_k, \mathbf{X}^n_{t_k}\right)\right)\nabla^2_x u\left(t_k, \mathbf{X}^n_{t_k}\right)\right]\right| \\
&\le \zeta\left(\frac{\alpha}{h} + h + h|\log(h)|^{d_1} + h|\log(h)|^{q+d_1}\right).
\end{aligned}$$

From Lemma 3.3.9 below, we obtain that $\sum_{k=0}^{n-1}\left|\tilde{\mathcal{R}}_k\right| = O\left(\Delta\right)$ and we conclude that

$$\sum_{k=0}^{n-1}\left|\mathfrak{E}_k\right| \leq \Pi\left(\Delta + \frac{\alpha}{h} + h + h|\log(h)|^{d_1} + h|\log(h)|^{p+d_1}\right),$$

where Π is a constant that does not depend of Δ.

Lemma 3.3.9. *Under Assumptions (A) and (IC), there exists $\Gamma > 0$ such that*

$$\forall 0 \leq k \leq n-1,\ |\tilde{\mathcal{R}}_k| \leq \Gamma\Delta^2.$$

Proof. As b, b_X have sublinear growth and $\nabla_x^2 u$ has polynomial growth in the variable x, by Lemma 3.3.4, there exists Γ_1 such that for $n \geq 0$ and $0 \leq k \leq n-1$,

$$\left|\mathbb{E}\left[\left(\mathbf{b}_{X,k}^n\right)^* \nabla_x^2 u\left(t_k, \mathbf{X}_{t_k}^n\right)\mathbf{b}_{X,k}^n\right]\right| \leq \Gamma_1$$

It is then sufficient to show that $\mathbb{E}\left[\mathcal{R}_k\right] \leq \Gamma_2\Delta^2$ where Γ_2 does not depend on k, n. In the following of the proof, the constants K, p could vary at each different argument, but are always assumed to be independent of n and k. Given the regularity of u and the coefficients b, σ, we have that by Proposition 3.2.4,

$$\partial_t^2 u = \partial_t \mathcal{L}u = \partial_t b^* \nabla_x u + \frac{1}{2}\mathbf{Tr}\left(\partial_t a \nabla_x^2 u\right) + b^*\partial_t \nabla_x u + \frac{1}{2}\mathbf{Tr}\left(a\partial_t \nabla_x^2 u\right),$$

where $\partial_t \nabla_x u = \nabla_x \mathcal{L}u$, $\partial_t \nabla_x^2 u = \nabla_x^2 \mathcal{L}u$, and those derivatives have polynomial growth in the space variable. This is sufficient to obtain that the expectation of the term in (3.3.7) is bounded by $\Pi_1\Delta^2$ for a constant $\Pi_1 > 0$, as there exist $K, p > 0$ such that

$$\mathbb{E}\left[|\partial_t^2 u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n + s\mathbf{D}_k^n\right)|\right] \leq \mathbb{E}\left[K\left(1 + \left|\mathbf{X}_{t_k}^n\right|^p + |\mathbf{D}_k^n|^p\right)\right]$$

and the r.h.s. is uniformly bounded in n, k by Lemmas 3.3.4 and 3.3.5. For the term in (3.3.8), we use the Taylor expansion at order one:

$$\partial_t \nabla_x u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n + s\mathbf{D}_k^n\right) = \partial_t \nabla_x u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n\right) + s\int_0^1 \partial_t \nabla_x^2 u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n + us\mathbf{D}_k^n\right)\mathbf{D}_k^n du.$$

By independence of $W_{t_{k+1}} - W_{t_k}$ with $\left(\mathbf{X}_{t_k}^n, \mathbf{Y}_{t_k}^n\right)$, conditioning w.r.t. $\left(\mathbf{X}_{t_k}^n, \mathbf{Y}_{t_k}^n\right)$ and using the sublinearity of b by Assumption (B), we have that

$$\left|\mathbb{E}\left[\Delta\left(\mathbf{D}_k^n\right)^* \partial_t \nabla_x u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n\right)\right]\right| \leq \Delta^2 \kappa \mathbb{E}\left[\sqrt{1 + |\mathbf{X}_{t_k}^n|^2}|\partial_t \nabla_x u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n\right)|\right],$$

and the expectation in the r.h.s. is bounded uniformly. Moreover, using the polynomial growth of $\partial_t \nabla_x^2 u$, we check by Lemmas 3.3.4 and 3.3.5 that there exists $K > 0$ such that

$$\left|\mathbb{E}\left[\Delta\int_0^1 (1-s)\, s\int_0^1 \left(\mathbf{D}_k^n\right)^* \partial_t \nabla_x^2 u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n + us\mathbf{D}_k^n\right)\mathbf{D}_k^n du ds\right]\right| \leq K\Delta^2.$$

Therefore we obtain that the expectation of the term in (3.3.8) is bounded by $\Pi_2\Delta^2$ for a constant $\Pi_2 > 0$. Finally, for the term in (3.3.9), let us apply the Taylor expansion at first order on $u \in (0,1) \to \nabla_x^2 u\left(t_k + u\Delta, \mathbf{X}_{t_k}^n + u\mathbf{D}_k^n\right)$. For $s \in (0,1)$,

$$\begin{aligned}\nabla_x^2 u\left(t_k + s\Delta, \mathbf{X}_{t_k}^n + s\mathbf{D}_k^n\right) - \nabla_x^2 u\left(t_k, \mathbf{X}_{t_k}^n\right) =& s\Delta\int_0^1 \partial_t \nabla_x^2 u\left(t_k + us\Delta, \mathbf{X}_{t_k}^n + us\mathbf{D}_k^n\right)du \\ &+ s\mathbf{D}_k^n \int_0^1 \nabla_x^3 u\left(t_k + us\Delta, \mathbf{X}_{t_k}^n + us\mathbf{D}_k^n\right)du.\end{aligned}$$

Similarly to the computations of the previous terms, there exists $\Pi_3 > 0$ such that

$$\mathbb{E}\left[\left|\Delta\int_0^1 (1-s)\, s\int_0^1 \left(\mathbf{D}_k^n\right)^* \partial_t \nabla_x^2 u\left(t_k + us\Delta, \mathbf{X}_{t_k}^n + us\mathbf{D}_k^n\right)\mathbf{D}_k^n du ds\right|\right] \leq \Pi_3\Delta^2,$$

and using the Taylor expansion on $s \in (0,1) \to \nabla_x^3 u\left(t_k + u\Delta, \mathbf{X}^n_{t_k} + us\mathbf{D}^n_k\right)$, we also obtain $\Pi_4 > 0$ such that

$$\mathbb{E}\left[\left|(\mathbf{D}^n_k)^3 \int_0^1 (1-s)\, s \int_0^1 \nabla_x^3 u\left(t_k + us\Delta, \mathbf{X}^n_{t_k} + us\mathbf{D}^n_k\right) duds\right|\right] \leq \Pi_4 \Delta^2,$$

and this concludes the proof. □

We now give the proofs of Lemmas 3.3.7 and 3.3.8 to conclude this section.

Proof of Lemma 3.3.7

Proof. To estimate the exponential moments of $|\mathbf{X}^n_{t_k}|$ for $0 \leq k \leq n$, it is sufficient to estimate the exponential moments of $\sqrt{1+|\mathbf{X}^n_{t_k}|^2}$. Using the fact that $\sqrt{1+x} \leq 1 + \frac{x}{2}$ for $x \geq 0$, and Assumption (B'), let us notice that

$$\begin{aligned}\sqrt{1+|\mathbf{X}^n_{t_k} + \mathbf{b}^n_{X,k}\Delta|^2} &= \sqrt{1+|\mathbf{X}^n_{t_k}|^2}\left(1 + 2\frac{\left(\mathbf{b}^n_{X,k}\right)^* \mathbf{X}^n_{t_k}}{1+|\mathbf{X}^n_{t_k}|^2}\Delta + \frac{|\mathbf{b}^n_{X,k}|^2}{1+|\mathbf{X}^n_{t_k}|^2}\Delta^2\right)^{\frac{1}{2}} \\ &\leq \sqrt{1+|\mathbf{X}^n_{t_k}|^2}\left(1 + \kappa\Delta + \kappa^2\frac{\Delta^2}{2}\right). \end{aligned} \tag{3.3.13}$$

Let us consider the process Z defined on $[0,T]$ by $Z_0 = \sqrt{1+|X_0|^2}$ and by

$$Z_t = \sqrt{1+|\mathbf{X}^n_{t_k} + \mathbf{b}^n_{X,k}\Delta + (\mathbf{a}^n_{X,k})^{\frac{1}{2}}(W^1_t - W^1_{t_k})|^2}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(k+1)},$$

for $t \in (t_k, t_{k+1}]$ and $0 \leq k \leq n-1$. Notice that for $0 \leq k \leq n$,

$$Z_{t_k} = \sqrt{1+|\mathbf{X}^n_{t_k}|^2}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-k}.$$

Using the fact that $\Delta = \frac{T}{n}$ and for $x \geq 0$, $\log(1+x) \leq x$, we have that for $0 \leq k \leq n$,

$$\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^k \leq \left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^n = e^{n\log(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2})} \leq e^{\kappa T + \kappa^2\frac{T^2}{2}} =: e_T.$$

We then obtain that for $0 \leq k \leq n$,

$$|\mathbf{X}^n_{t_k}| \leq \sqrt{1+|\mathbf{X}^n_{t_k}|^2} \leq e_T Z_{t_k}. \tag{3.3.14}$$

Let us define for $0 \leq i \leq n-1$, $Z_{t_i+} = \lim_{t\to t_i, t>t_i} Z_t$. The jumps of the process Z only occur at the discretization times $(t_k)_{1\leq k\leq n-1}$, and by Inequality (3.3.13), they are nonpositive. Therefore we have that for $1 \leq k \leq n$,

$$0 \leq Z_{t_k} \leq Z_0 + \sum_{i=0}^{k-1}\left(Z_{t_{i+1}} - Z_{t_i+}\right). \tag{3.3.15}$$

Let $z \in \mathbb{R}^{d_1}$ and let us define for $x \in \mathbb{R}^{d_1}$, $R_z(x) = \sqrt{1+|z+x|^2}$. Moreover, for $0 \leq i \leq n-1$,

$$\begin{aligned} Z_{t_{i+1}} - Z_{t_i+} &= \int_{t_i}^{t_{i+1}}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)} \nabla_x^* R_{\mathbf{X}^n_{t_i}+\mathbf{b}^n_{X,i}\Delta}((\mathbf{a}^n_{X,i})^{\frac{1}{2}}(W^1_t - W^1_{t_i}))(\mathbf{a}^n_{X,i})^{\frac{1}{2}} dW_t \\ &+ \frac{1}{2}\int_{t_i}^{t_{i+1}}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)} \mathbf{Tr}\left(\mathbf{a}^n_{X,i}\nabla_x^2 R_{\mathbf{X}^n_{t_i}+\mathbf{b}^n_{X,i}\Delta}((\mathbf{a}^n_{X,i})^{\frac{1}{2}}(W^1_t - W^1_{t_i}))\right) dt. \end{aligned}$$

As $\mathbf{a}^n_{X,i}$ is bounded and the derivatives $\nabla_x R_z$, $\nabla_x^2 R_z$ are bounded, uniformly in $z, x \in \mathbb{R}^{d_1}$, we have that

$$\left|(\mathbf{a}^n_{X,i})^{\frac{1}{2}}\nabla_x R_{\mathbf{X}^n_{t_i}+\mathbf{b}^n_{X,i}\Delta}((\mathbf{a}^n_{X,i})^{\frac{1}{2}}(W^1_t - W^1_{t_i}))\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)}\right| \leq d||a_X||_\infty^{\frac{1}{2}}||\nabla_x R||_\infty, \tag{3.3.16}$$

$$\left|\mathbf{Tr}\left(\mathbf{a}_{X,i}^n\nabla_x^2 R_{\mathbf{X}_{t_i}^n+\mathbf{b}_{X,i}^n\Delta}((\mathbf{a}_{X,i}^n)^{\frac{1}{2}}(W_t^1-W_{t_i}^1))\right)\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)}\right|\leq d^2||a_X||_\infty||\nabla_x^2 R||_\infty.$$

Using Inequality (3.3.15), we have that

$$\begin{aligned}0\leq Z_{t_k}\leq&\sqrt{1+|X_0|^2}+\frac{d^2}{2}||a_X||_\infty||\nabla_x^2R||_\infty T\\&+\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)}\nabla_x^* R_{\mathbf{X}_{t_i}^n+\mathbf{b}_{X,i}^n\Delta}((\mathbf{a}_{X,i}^n)^{\frac{1}{2}}(W_t^1-W_{t_i}^1))(\mathbf{a}_{X,i}^n)^{\frac{1}{2}}dW_t^1.\end{aligned}$$

Using (3.3.14) and the Cauchy-Schwarz inequality, we obtain that for $0\leq k\leq n$,

$$\begin{aligned}&\mathbb{E}\left[e^{\mu|\mathbf{X}_{t_k}^n|}\right]\leq\mathbb{E}\left[e^{\mu e_T Z_{t_k}}\right]\\&\leq\exp\left(\mu e_T\frac{d^2}{2}||a_X||_\infty||\nabla_x^2R||_\infty T\right)\mathbb{E}\left[\exp\left(2\mu e_T\sqrt{1+|X_0|^2}\right)\right]^{\frac{1}{2}}\qquad(3.3.17)\\&\times\mathbb{E}\left[\exp\left(2\mu e_T\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)}\nabla_x^*R_{\mathbf{X}_{t_i}^n+\mathbf{b}_{X,i}^n\Delta}((\mathbf{a}_{X,i}^n)^{\frac{1}{2}}(W_t^1-W_{t_i}^1))(\mathbf{a}_{X,i}^n)^{\frac{1}{2}}dW_t^1\right)\right]^{\frac{1}{2}}.\qquad(3.3.18)\end{aligned}$$

Using Inequality (3.3.16), the integrand of the stochastic integral in (3.3.18) is uniformly bounded so we have that

$$\begin{aligned}&\mathbb{E}\left[\exp\left(2\mu e_T\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\left(1+\kappa\Delta+\kappa^2\frac{\Delta^2}{2}\right)^{-(i+1)}\nabla_x^*R_{\mathbf{X}_{t_i}^n+\mathbf{b}_{X,i}^n\Delta}((\mathbf{a}_{X,i}^n)^{\frac{1}{2}}(W_t^1-W_{t_i}^1))(\mathbf{a}_{X,i}^n)^{\frac{1}{2}}dW_t^1\right)\right]^{\frac{1}{2}}\\&\leq\exp\left(T\left(\mu e_T d||a_X||_\infty^{\frac{1}{2}}||\nabla_xR||_\infty\right)^2\right)\end{aligned}$$

Finally, we have that $\mathbb{E}\left[e^{2\mu e_T\sqrt{1+|X_0|^2}}\right]<\infty$ for $2\mu e_T\leq\lambda_0$ and this concludes the proof. □

Proof of Lemma 3.3.8

Proof. If $h>1$, as $1_{\{p(x)\leq h\}}\leq he^{\frac{\lambda}{2}|x|}$ for $x\in\mathbb{R}^d$,

$$\int_{\mathbb{R}^d}|x|^kp(x)1_{\{p(x)\leq h\}}dx=h\int_{\mathbb{R}^d}|x|^kp(x)e^{\frac{\lambda}{2}|x|}dx=h\mathbb{E}\left[|Z|^ke^{\frac{\lambda}{2}|Z|}\right]\leq h\frac{k!2^k}{\lambda^k}\mathbb{E}\left[e^{\lambda|Z|}\right],$$

as for $z>0$, $z^k\leq\frac{k!2^k}{\lambda^k}e^{\frac{\lambda}{2}z}$, and $\mathbb{E}\left[e^{\lambda|Z|}\right]<\infty$. Now, let us assume that $h\leq1$. For $x\in\mathbb{R}^d$ such that $|x|\geq\frac{-2\log(h)}{\lambda}$, we have that $\exp\left(-\frac{\lambda}{2}|x|\right)\leq h$. We then write

$$\int_{\mathbb{R}^d}|x|^kp(x)1_{\{p(x)\leq h\}}dx=\int_{\mathbb{R}^d}|x|^kp(x)1_{\{p(x)\leq h,|x|\geq\frac{-2\log(h)}{\lambda}\}}dx+\int_{\mathbb{R}^d}|x|^kp(x)1_{\{p(x)\leq h,|x|\leq\frac{-2\log(h)}{\lambda}\}}dx.$$

For the first term of the r.h.s.,

$$\begin{aligned}\int_{\mathbb{R}^d}|x|^kp(x)1_{\{p(x)\leq h,|x|\geq\frac{-2\log(h)}{\lambda}\}}dx&=\int_{\mathbb{R}^d}|x|^kp(x)e^{\frac{\lambda}{2}|x|}e^{-\frac{\lambda}{2}|x|}1_{\{p(x)\leq h,|x|\geq\frac{-2\log(h)}{\lambda}\}}dx\\&\leq h\mathbb{E}\left[|Z|^ke^{\frac{\lambda}{2}|Z|}\right]\leq h\frac{k!2^k}{\lambda^k}\mathbb{E}\left[e^{\lambda|Z|}\right].\end{aligned}$$

For the second term of the r.h.s, we have that

$$\int_{\mathbb{R}^d}|x|^kp(x)1_{\{p(x)\leq h,|x|\leq\frac{-2\log(h)}{\lambda}\}}dx\leq h\int_{\mathbb{R}^d}|x|^k1_{\{|x|\leq\frac{-2\log(h)}{\lambda}\}}dx=h\frac{2\pi^{\frac{d}{2}}}{(k+d)\,\Gamma(\frac{d}{2})}\left(\frac{2\log(h)}{\lambda}\right)^{k+d},$$

where we integrated the function $x\in\mathbb{R}^d\to|x|^k$ over the d-dimensional ball with radius $\frac{-2\log(h)}{\lambda}$ in $\mathbb{R}^d$ to obtain the equality. This concludes the proof. □

3.3.3 Interacting particles system and rough error bounds

In this section, we moreover assume Lipschitz regularity on the functions $\varphi, \phi, b_X, a_X, b_Y, a_Y$.

Assumption (Lipf). *There exist constants L_ϕ and L_φ such that*

$$\forall x_1, x_2 \in \mathbb{R}^{d_1}, |\varphi(x_1) - \varphi(x_2)| \leq L_\varphi |x_1 - x_2|,$$
$$\forall x_1, x_2 \in \mathbb{R}^{d_1}, \forall y_1, y_2 \in \mathbb{R}^{d_2}, |\phi(x_1, y_1) - \phi(x_2, y_2)| \leq L_\phi(|x_1 - x_2| + |y_1 - y_2|).$$

Moreover, $|\phi|$ is bounded by $\overline{\phi}$.

Assumption (Lipc). *There exists a constant L_c such that for all $t \geq 0$ and $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^q$,*

$$|b_X(t, x_1, y_1, z_1) - b_X(t, x_2, y_2, z_2)| \vee |a_X(t, x_1, y_1, z_1) - a_X(t, x_2, y_2, z_2)| \leq L_c(|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|),$$
$$|b_Y(t, x_1, y_1) - b_Y(t, x_2, y_2)| \vee |a_Y(t, x_1, y_1) - a_Y(t, x_2, y_2)| \leq L_c(|x_1 - x_2| + |y_1 - y_2|).$$

Let $N \in \mathbb{N}^*$ be the number of particles and let the function F^N be defined for $z, x_1, ..., x_n \in \mathbb{R}^{d_1}$ and $y_1, ..., y_N \in \mathbb{R}^{d_2}$ by

$$F^N(z, x_1, ..., x_n, y_1, ..., y_n) = \frac{\frac{1}{N}\sum_{j=1}^N \phi(z, y_j)\, G_{\sigma_0^2\Delta, d_1, \alpha}(z - x_j)}{\frac{1}{N}\sum_{j=1}^N G_{\sigma_0^2\Delta, d_1, \alpha}(z - x_j)}.$$

Using the fact that ϕ is Lipschitz and bounded, $G_{\rho, d_1, \alpha} \geq \alpha$ and its Lipschitz constant is given by Lemma 3.3.3, it is easy to check the following properties.

Lemma 3.3.10. *For $z, x_1, ..., x_n, \tilde{z}, \tilde{x} \in \mathbb{R}^{d_1}$ and $y_1, ..., y_n, \tilde{y} \in \mathbb{R}^{d_2}$, we have that*

$$\left|F^N(z, x_1, ..., x_n, y_1, ..., y_n) - F^N(\tilde{z}, x_1, ..., x_n, y_1, ..., y_n)\right| \leq K_Z |z - \tilde{z}|$$
$$\left|F^N(z, x_1, ..., x_n, y_1, ..., y_n) - F^N(z, \tilde{x}, x_2..., x_n, y_1, ..., y_n)\right| \leq \frac{K_X}{N}|x_1 - \tilde{x}|,$$
$$\left|F^N(z, x_1, ..., x_n, y_1, ..., y_n) - F^N(z, x_1, ..., x_n, \tilde{y}, y_2..., y_n)\right| \leq \frac{K_Y}{N}|y_1 - \tilde{y}|,$$

where $K_Z = \left(L_\phi + \frac{K_d\overline{\phi}}{\alpha\left(\sigma_0^2\Delta\right)^{\frac{d+1}{2}}}\right)$, $K_X = \frac{K_d\overline{\phi}}{\alpha\left(\sigma_0^2\Delta\right)^{\frac{d+1}{2}}}$, and $K_Y = \frac{L_\phi}{\alpha\left(2\pi\sigma_0^2\Delta\right)^{\frac{d}{2}}}$ and the constant K_d is introduced in Lemma 3.3.3.

Remark. *Let us remark that replacing the regularizing function $G_{\rho,d,\alpha}$ by the Lipschitz modification*

$$G^\beta_{\rho,d} = \begin{cases} G_{\rho,d}(x) & \textit{if } |x| \leq \sqrt{2\beta\rho\log(\frac{1}{\rho})}, \\ \frac{1}{(2\pi)^{d/2}}\rho^{\beta - d/2} & \textit{otherwise}, \end{cases}$$

for $\beta > 0$, it is possible to lower the exponent on Δ in the denominator of K_Z, but it does not improve the exponent of Δ in the denominators of K_X or K_Y.

We introduce a particle system associated to the modified half-step scheme (3.3.3). Let $\left(\left(Z_k^{1,i}, Z_{k+\frac{1}{2}}^{1,i}, Z_k^{2,i}\right)_{k\geq 0}\right)_{1\leq i\leq N}$ be an i.i.d. family with the same law as $\left(Z_k^1, Z_{k+\frac{1}{2}}^1, Z_k^2\right)_{k\geq 0}$. Let $\left(\mathbf{X}_0^{i,N}, \mathbf{Y}_0^{i,N}\right)_{1\leq i\leq N}$ be i.i.d. random variables with the same law as $(\mathbf{X}_0, \mathbf{Y}_0)$ and which are independent of $\left(Z_k^{1,i}, Z_{k+\frac{1}{2}}^{1,i}, Z_k^{2,i}\right)_{k\geq 0, 1\leq i\leq N}$. The dynamics is given by

$$\begin{aligned}
\mathbf{X}^{n,i,N}_{t_{k+\frac{1}{2}}} &= \mathbf{X}^{n,i,N}_{t_k} + \mathbf{b}^{n,i,N}_{X,k}\Delta + \left(\mathbf{a}^{n,i,N}_{X,t_k} - \sigma_0^2 I_{d_1}\right)^{\frac{1}{2}}\sqrt{\Delta} Z_k^{1,i}, \\
\mathbf{X}^{n,i,N}_{t_{k+1}} &= \mathbf{X}^{n,i,N}_{t_{k+\frac{1}{2}}} + \sigma_0\sqrt{\Delta} Z^{1,i}_{k+\frac{1}{2}},
\end{aligned}$$

$$\mathbf{Y}_{t_{k+1}}^{n,i,N} = \mathbf{Y}_{t_k}^{n,i,N} + b_Y\left(t_k, \mathbf{X}_{t_k}^{n,i,N}, \mathbf{Y}_{t_k}^{n,i,N}\right)\Delta + \sigma_Y\left(t_k, \mathbf{X}_{t_k}^{n,i,N}, \mathbf{Y}_{t_k}^{n,i,N}\right)\sqrt{\Delta}Z_k^{2,i}. \tag{3.3.19}$$

Here, for $n \geq 1$, $0 \leq k \leq n-1$, $N \geq 1$, $1 \leq i \leq N$ and $\alpha > 0$, $\mathbf{b}_{X,k}^{n,i,N} = b_X\left(t_k, \mathbf{X}_{t_k}^{n,i,N}, \mathbf{Y}_{t_k}^{n,i,N}, \mathbf{E}_k^N\left(\mathbf{X}_{t_k}^{n,i,N}\right)\right)$, $\mathbf{a}_{X,k}^{n,i,N} = a_X\left(t_k, \mathbf{X}_{t_k}^{n,i,N}, \mathbf{Y}_{t_k}^{n,i,N}, \mathbf{E}_k^N\left(\mathbf{X}_{t_k}^{n,i,N}\right)\right)$ and

$$\mathbf{E}_k^N\left(\mathbf{X}_{t_k}^{n,i,N}\right) = F_N\left(\mathbf{X}_{t_k}^{n,i,N}, \mathbf{X}_{t_{k-\frac{1}{2}}}^{n,1,N}, ..., \mathbf{X}_{t_{k-\frac{1}{2}}}^{n,N,N}, \mathbf{Y}_{t_k}^{n,1,N}, ..., \mathbf{Y}_{t_k}^{n,N,N}\right),$$

Given a test function $\varphi : \mathbb{R}^{d_1} \to \mathbb{R}$, we are interested in the mean square error

$$\mathbf{MSE}(\varphi) = \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N \varphi(\mathbf{X}_T^{n,i,N}) - \mathbb{E}\left[\varphi(Z_T)\right]\right)^2\right].$$

For $1 \leq i \leq N$, let $\overline{\mathbf{X}}_T^{n,i}$ be constructed according to the modified half-step scheme (3.3.3) and driven by the variables $\left(Z_k^{1,i}, Z_{k+\frac{1}{2}}^{1,i}, Z_k^{2,i}\right)_{k\geq 0}$. The error $\mathbf{MSE}(\varphi)$ can be separated into three terms using

$$\begin{aligned}\left(\frac{1}{N}\sum_{i=1}^N \varphi(\mathbf{X}_T^{n,i,N}) - \mathbb{E}\left[\varphi(Z_T)\right]\right)^2 &\leq 3\left(\frac{1}{N}\sum_{i=1}^N \varphi(\mathbf{X}_T^{n,i,N}) - \frac{1}{N}\sum_{i=1}^N \varphi(\overline{\mathbf{X}}_T^{n,i})\right)^2 + 3\left(\frac{1}{N}\sum_{i=1}^N \varphi(\overline{\mathbf{X}}_T^{n,i}) - \mathbb{E}\left[\varphi(\overline{\mathbf{X}}_T^{n,1})\right]\right)^2 \\ &+ 3\left(\mathbb{E}\left[\varphi(\overline{\mathbf{X}}_T^{n,1})\right] - \mathbb{E}\left[\varphi(Z_T)\right]\right)^2 =: 3\left(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\right),\end{aligned}$$

The term $\mathcal{E}_2$ is a statistical error term and the term $\mathcal{E}_3$ is the weak error between Z_T and the Euler scheme (3.3.3). To estimate $\mathbf{MSE}(\varphi)$, it is thus sufficient to estimate $\mathbb{E}\left[\mathcal{E}_1\right]$, $\mathbb{E}\left[\mathcal{E}_2\right]$ and $\mathcal{E}_3$. By Proposition 3.3.6 from the analysis in the previous section, there exist two constants $\zeta_1, q > 0$, independent of n such that for all $h > 0$,

$$\mathcal{E}_3 \leq \zeta_1\left(\Delta + \frac{\alpha}{h} + h + h|\log(h)|^{d_1} + h|\log(h)|^{q+d_1}\right)^2,$$

Moreover, as $\overline{\mathbf{X}}_T^{n,i}$ and $\overline{\mathbf{X}}_T^{n,j}$ are independent for $1 \leq i \neq j \leq N$, we have that

$$\mathbb{E}\left[\mathcal{E}_2\right] = \frac{\mathbf{Var}(\varphi(\overline{\mathbf{X}}_T^{n,1}))}{N} \leq \frac{\mathbb{E}\left[\varphi^2(\overline{\mathbf{X}}_T^{n,1})\right]}{N}.$$

By Assumption (Lipf), φ has sublinear growth so by Lemma 3.3.4 there exists $\zeta_2 > 0$ such that $\sup_n \mathbb{E}\left[\varphi^2\left(\overline{\mathbf{X}}_T^{n,1}\right)\right] < \zeta_2$ and therefore $\mathbb{E}\left[\mathcal{E}_2\right] \leq \frac{\zeta_2}{N}$. The term $\mathcal{E}_1$ is the propagation of chaos error term. By Assumption (Lipf),

$$\left|\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N\left(\varphi\left(\mathbf{X}_T^{n,i,N}\right) - \varphi\left(\mathbf{X}_T^{n,i}\right)\right)\right)^2\right]\right| \leq \frac{L_\varphi^2}{N}\sum_{i=1}^N \mathbb{E}\left[\left|\mathbf{X}_T^{n,i,N} - \mathbf{X}_T^{n,i}\right|^2\right].$$

Theorem 3.3.11. *Under Assumptions (Lipf), (Lipc), (IC') and (B), there exists a constant $K > 0$ independent of Δ, N and such that*

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\left|\mathbf{X}_T^{n,i,N} - \overline{\mathbf{X}}_T^{n,i}\right|^2 + \left|\mathbf{Y}_T^{n,i,N} - \overline{\mathbf{Y}}_T^{n,i}\right|^2\right] \leq \frac{1}{N}e^{\frac{K}{\alpha^2\Delta^{d_1+1}}}.$$

Proof. To simplify notations, let us drop the index n denoting the total number of time steps. We have that

$$\mathbb{E}\left[\left|\mathbf{X}_T^{i,N} - \overline{\mathbf{X}}_T^{i}\right|^2\right] \leq 2\mathbb{E}\left[\left|\sum_{k=0}^{n-1}\left(\mathbf{b}_k^{i,N} - \overline{\mathbf{b}}_k^{i}\right)\Delta\right|^2\right] + 2\mathbb{E}\left[\left|\sum_{k=0}^{n-1}\left(\sqrt{\mathbf{a}_k^{i,N} - \sigma_0^2 I_{d_1}} - \sqrt{\overline{\mathbf{a}}_k^{i} - \sigma_0^2 I_{d_1}}\right)\sqrt{\Delta}Z_k^{1,i}\right|^2\right]. \tag{3.3.20}$$

To deal with the first term in the r.h.s., we use that

$$\mathbb{E}\left[\left|\sum_{k=0}^{n-1}\left(\mathbf{b}_k^{i,N}-\overline{\mathbf{b}}_k^i\right)\Delta\right|^2\right]\le T\Delta\sum_{k=0}^{n-1}\mathbb{E}\left[\left|\mathbf{b}_k^{i,N}-\overline{\mathbf{b}}_k^i\right|^2\right]. \tag{3.3.21}$$

By Assumption (Lipc),

$$\mathbb{E}\left[\left|\mathbf{b}_k^{i,N}-\overline{\mathbf{b}}_k^i\right|^2\right]\le 3L_c^2\mathbb{E}\left[\left|\mathbf{X}_k^{i,N}-\overline{\mathbf{X}}_k^i\right|^2+\left|\mathbf{Y}_k^{i,N}-\overline{\mathbf{Y}}_k^i\right|^2+\left|\mathbf{E}_k^N\left(\mathbf{X}_{t_k}^{i,N}\right)-\mathbf{E}_k\left(\overline{\mathbf{X}}_{t_k}^i\right)\right|^2\right],$$

We then study the difference

$$\begin{aligned}
&\left|\mathbf{E}_k^N\left(\mathbf{X}_{t_k}^{i,N}\right)-\mathbf{E}_k\left(\overline{\mathbf{X}}_{t_k}^i\right)\right|^2\\
&\le 2\left|F_N\left(\mathbf{X}_{t_k}^{i,N},\mathbf{X}_{t_{k-\frac12}}^{1,N},...,\mathbf{X}_{t_{k-\frac12}}^{N,N},\mathbf{Y}_{t_k}^{1,N},...,\mathbf{Y}_{t_k}^{N,N}\right)-F_N\left(\overline{\mathbf{X}}_{t_k}^i,\overline{\mathbf{X}}_{t_{k-\frac12}}^1,...,\overline{\mathbf{X}}_{t_{k-\frac12}}^N,\overline{\mathbf{Y}}_{t_k}^1,...,\overline{\mathbf{Y}}_{t_k}^N\right)\right|^2\\
&\quad+2\left|F_N\left(\overline{\mathbf{X}}_{t_k}^i,\overline{\mathbf{X}}_{t_{k-\frac12}}^1,...,\overline{\mathbf{X}}_{t_{k-\frac12}}^N,\overline{\mathbf{Y}}_{t_k}^1,...,\overline{\mathbf{Y}}_{t_k}^N\right)-\mathbf{E}_k\left(\overline{\mathbf{X}}_{t_k}^i\right)\right|^2\\
&\le 2T_1+2T_2
\end{aligned}$$

For the term T_1, by Assumption (Lipf) and using Lemma 3.3.10, we have that

$$\begin{aligned}
T_1&\le 3K_Z^2|\mathbf{X}_{t_k}^{i,N}-\overline{\mathbf{X}}_{t_k}^i|^2+3K_X^2\left(\frac1N\sum_{j=1}^N|\mathbf{X}_{t_k}^{i,N}-\overline{\mathbf{X}}_{t_k}^i|\right)^2+3K_Y^2\left(\frac1N\sum_{j=1}^N|\mathbf{Y}_{t_k}^{i,N}-\overline{\mathbf{Y}}_{t_k}^i|\right)^2,\\
&\le 3K_Z^2|\mathbf{X}_{t_k}^{i,N}-\overline{\mathbf{X}}_{t_k}^i|^2+3K_X^2\frac1N\sum_{j=1}^N|\mathbf{X}_{t_k}^{i,N}-\overline{\mathbf{X}}_{t_k}^i|^2+3K_Y^2\frac1N\sum_{j=1}^N|\mathbf{Y}_{t_k}^{i,N}-\overline{\mathbf{Y}}_{t_k}^i|^2.
\end{aligned}$$

Moreover, let us introduce the notation for $1\le j\le N$,

$$\begin{aligned}
\psi_j&=G_{\underline{\sigma}^2\Delta,d_1,\alpha}\left(\overline{\mathbf{X}}_{t_k}^i-\overline{\mathbf{X}}_{t_{k-\frac12}}^j\right)-\mathbb{E}\left[G_{\underline{\sigma}\Delta,d_1,\alpha}\left(\overline{\mathbf{X}}_{t_k}^i-\tilde{\mathbf{X}}_{t_{k-\frac12}}\right)\Big|\overline{\mathbf{X}}_{t_k}^i\right],\\
\chi_j&=\phi\left(\overline{\mathbf{X}}_{t_k}^i,\overline{\mathbf{Y}}_{t_k}^j\right)G_{\underline{\sigma}^2\Delta,d_1,\alpha}\left(\overline{\mathbf{X}}_{t_k}^i-\overline{\mathbf{X}}_{t_{k-\frac12}}^j\right)-\mathbb{E}\left[\phi\left(\overline{\mathbf{X}}_{t_k}^i,\tilde{\mathbf{Y}}_{t_k}\right)G_{\underline{\sigma}\Delta,d_1,\alpha}\left(\overline{\mathbf{X}}_{t_k}^i-\tilde{\mathbf{X}}_{t_{k-\frac12}}\right)\Big|\overline{\mathbf{X}}_{t_k}^i\right],
\end{aligned}$$

where $\left(\tilde{\mathbf{X}}_{t_{k-\frac12}},\tilde{\mathbf{Y}}_{t_k}\right)$ is a copy of $\left(\overline{\mathbf{X}}_{t_{k-\frac12}}^i,\overline{\mathbf{Y}}_{t_k}^i\right)$, independent of $\overline{\mathbf{X}}_{t_k}^i$. As $\alpha\le G_{\rho,d,\alpha}$ and ϕ is bounded, we obtain that

$$\mathbb{E}\left[\left|F_N\left(\overline{\mathbf{X}}_{t_k}^i,\overline{\mathbf{X}}_{t_{k-\frac12}}^1,...,\overline{\mathbf{X}}_{t_{k-\frac12}}^N,\overline{\mathbf{Y}}_{t_k}^1,...,\overline{\mathbf{Y}}_{t_k}^N\right)-\mathbf{E}_k\left(\overline{\mathbf{X}}_{t_k}^i\right)\right|^2\right]\le\frac{2}{\alpha^2N^2}\mathbb{E}\left[\overline{\phi}^2\left|\sum_{j=1}^N\psi_j\right|^2+\left|\sum_{j=1}^N\chi_j\right|^2\right].$$

By conditional independence w.r.t. $\overline{\mathbf{X}}_{t_k}^i$, we have that that $\mathbb{E}\left[\psi_{j_1}\psi_{j_2}\right]=0$ and $\mathbb{E}\left[\chi_{j_1}\chi_{j_2}\right]=0$ for $1\le j_1\ne j_2\le N$. As moreover, $G_{\rho,d,\alpha}\le\alpha+(2\pi\underline{\sigma}\Delta)^{-\frac d2}$, we obtain that

$$\mathbb{E}\left[\left|\sum_{j=1}^N\psi_j\right|^2\right]=\sum_{j=1}^N\mathbb{E}\left[\psi_j^2\right]\le 4N\left(\alpha+(2\pi\underline{\sigma}\Delta)^{-\frac{d_1}{2}}\right)^2,\ \mathbb{E}\left[\left|\sum_{j=1}^N\chi_j\right|^2\right]=\sum_{j=1}^N\mathbb{E}\left[\chi_j^2\right]\le 4\overline{\phi}^2N\left(\alpha+(2\pi\underline{\sigma}\Delta)^{-\frac{d_1}{2}}\right)^2,$$

and we conclude that

$$\mathbb{E}\left[T_2\right]\le\frac{16\overline{\phi}^2}{N\alpha^2}\left(\alpha+(2\pi\underline{\sigma}\Delta)^{-\frac{d_1}{2}}\right)^2.$$

The second term in (3.3.20) is treated in the same way as the first term. Indeed, for $0 \le k_1 < k_2 \le n$, as Z_{k_2} is independent of

$$\left(\left(\sqrt{\mathbf{a}_{k_1}^{i,N} - \sigma_0^2 I_{d_1}} - \sqrt{\overline{\mathbf{a}}_{k_1}^{i} - \sigma_0^2 I_{d_1}}\right) Z_{k_1}^{1,i}, \sqrt{\mathbf{a}_{k_2}^{i,N} - \sigma_0^2 I_{d_1}} - \sqrt{\overline{\mathbf{a}}_{k_2}^{i} - \sigma_0^2 I_{d_1}}\right)$$

we obtain that

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{k=0}^{n-1}\left(\sqrt{\mathbf{a}_{k}^{i,N} - \sigma_0^2 I_{d_1}} - \sqrt{\overline{\mathbf{a}}_{k}^{i} - \sigma_0^2 I_{d_1}}\right)\sqrt{\Delta} Z_k^{1,i}\right|^2\right] &= T\Delta \sum_{k=0}^{n-1}\mathbb{E}\left[\left|\left(\sqrt{\mathbf{a}_{k}^{i,N} - \sigma_0^2 I_{d_1}} - \sqrt{\overline{\mathbf{a}}_{k}^{i} - \sigma_0^2 I_{d_1}}\right) Z_k^{1,i}\right|^2\right] \\ &\le \frac{T}{4(\underline{\sigma}^2 - \sigma_0^2)}\Delta \sum_{k=0}^{n-1}\mathbb{E}\left[\left|\mathbf{a}_k^{i,N} - \overline{\mathbf{a}}_k^{i}\right|^2\right],\end{aligned}$$

and $\mathbf{a}_k^{i,N} - \overline{\mathbf{a}}_k^{i}$ satisfies the same Lipschitz property as $\mathbf{b}_k^{i,N} - \overline{\mathbf{b}}_k^{i}$. Let us introduce the notation

$$\Pi_k^X = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\left(\mathbf{X}_{t_k}^{i,N} - \overline{\mathbf{X}}_{t_k}^{i}\right)^2\right], \Pi_k^Y = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\left(\mathbf{Y}_{t_k}^{i,N} - \overline{\mathbf{Y}}_{t_k}^{i}\right)^2\right],$$

for $0 \le k \le n$. So far, we have obtained the existence of a constant $C > 0$ independent of Δ and N such that

$$\Pi_n^X \le C\sum_{k=0}^{n-1}\Delta(1 + K_X^2 + K_Z^2)\Pi_k^X + C\sum_{k=0}^{n-1}\Delta(1 + K_Y^2)\Pi_k^Y + C\frac{16\overline{\phi}^2}{N\alpha^2}\left(\alpha + \left(2\pi\underline{\sigma}^2\Delta\right)^{-\frac{d_1}{2}}\right)^2, \tag{3.3.22}$$

Moreover, by Assumption (Lipc), we also have that

$$\Pi_n^Y \le 2L_c^2\sum_{k=0}^{n-1}\left(\Pi_k^X + \Pi_k^Y\right). \tag{3.3.23}$$

Then summing Inequalities (3.3.22) and (3.3.23), we conclude the proof by the discrete version of Gronwall's Lemma. □

By Theorem 3.3.11, the error $\mathbf{MSE}(\varphi)$ can be estimated by

$$\mathbf{MSE}(\varphi) \le 3\zeta_1\left(\Delta + \frac{\alpha}{h} + h + h|\log(h)|^{d_1} + h|\log(h)|^{q+d_1}\right)^2 + 3\frac{\zeta_2}{N} + \frac{3}{N}e^{\frac{K}{\alpha^2\Delta^{d+1}}}. \tag{3.3.24}$$

Remark. *Let us set $h \sim \Delta^{\gamma}$, $\alpha \sim \Delta^{\gamma+\eta}$ where $\gamma, \eta \in (0,1)$. Given a tolerance $\epsilon^2 > 0$ on the error, it is sufficient to choose Δ and N such that*

$$\Delta^{\eta} \vee \Delta^{\gamma}|\log(\Delta)|^{q+d_1} = \epsilon, \;\; N = \epsilon^{-2}e^{\frac{K}{\Delta^{d+1+2\gamma+2\eta}}},$$

to ensure that $\mathbf{MSE}(\varphi) = \mathcal{O}(\epsilon^2)$.

3.3.4 Numerical results

In this section, we take $d_1 = d_2 = 1$. We consider the LSV model in the Black-Scholes setting: $\sigma_{Dup} = 1, r = 0, T = 1, S_0 = 1$. Moreover, $b_Y = 1, \sigma_Y = 1$ and for $y \in \mathbb{R}$, $f(y) = 1 + \min(1, y^2)$.

Weak error check. To investigate numerically the dependence of the weak error on the number of time steps, we use a scheme called kernel approximation (KA) scheme which is close to the one proposed by Guyon and Henry-Labordère. More precisely, it is initialized for $1 \le i \le N$ with $\mathcal{X}_0^{n,i,N} = \mathcal{Y}_0^{n,i,N} = 0$ and evolves according to

$$\begin{aligned}\mathcal{X}_{t_{k+1}}^{n,i,N} &= \mathcal{X}_{t_k}^{n,i,N} + B_{X,k}^{n,i,N}\Delta + A_{X,t_k}^{n,i,N}\sqrt{\Delta}Z_k^{1,i}, \\ \mathcal{Y}_{t_{k+1}}^{n,i,N} &= \mathcal{Y}_{t_k}^{n,i,N} + \Delta + \sqrt{\Delta}Z_k^{2,i},\end{aligned}$$

where for $n \geq 1$, $0 \leq k \leq n-1$, $1 \leq i \leq N$, $B^{n,i,N}_{X,k} = -\frac{1}{2}\frac{f^2(\mathcal{Y}^{n,i,N}_{t_k})}{\mathcal{E}^N_k(\mathcal{X}^{n,i,N}_{t_k})}$, $A^{n,i,N}_{X,k} = \frac{f(\mathcal{Y}^{n,i,N}_{t_k})}{\sqrt{\mathcal{E}^N_k(\mathcal{X}^{n,i,N}_{t_k})}}$ and

$$\mathcal{E}^N_k\left(\mathcal{X}^{n,i,N}_{t_k}\right) = \frac{\sum_{j=1}^N f^2(\mathcal{Y}^{n,j,N}_{t_k})G_\epsilon(\mathcal{X}^{n,i,N}_{t_k} - \mathcal{X}^{n,j,N}_{t_k})}{\sum_{j=1}^N G_\epsilon(\mathcal{X}^{n,i,N}_{t_k} - \mathcal{X}^{n,j,N}_{t_k})},$$

with $\epsilon > 0$ a fixed regularization parameter and G_ϵ is the gaussian density with variance ϵ. As a large number of terms in the approximation $\mathcal{E}^N_k\left(\mathcal{X}^{n,i,N}_{t_k}\right)$ of the conditional expectation are negligible, we accelerate the computation by sorting the particle system at each time step and fix a threshold $\eta = 10^{-3}$, so that the terms $G_\epsilon(\mathcal{X}^{n,i,N}_{t_k} - \mathcal{X}^{n,j,N}_{t_k})$ in the numerator and denominator of $\mathcal{E}^N_k\left(\mathcal{X}^{n,i,N}_{t_k}\right)$ are neglected whenever $G_\epsilon(\mathcal{X}^{n,i,N}_{t_k} - \mathcal{X}^{n,j,N}_{t_k}) < \eta$. For a large number of particles $N = 10000$ and a small regularization parameter $\epsilon = 0.01$, we give results for the put price estimated by $\frac{1}{N}\sum_{i=1}^N \left(K - e^{\mathcal{X}^{n,i,N}_T}\right)_+$ at each of the $M = 1000$ independent Monte-Carlo runs. We set $K = 0.8, 1, 2$, that is when the option is respectively out-of-the-money, at-the-money and in-the-money. According to the numerical results, the weak error behaves as $O(\frac{1}{n})$ where n is the number of time steps.

- Strike $K = 0.8$

Time steps	Estimated weak error	95% Confidence interval	$n\times$ Estimated weak error
$n = 5$	-0.001843	[-0.002001, -0.001686]	-0.009217
$n = 10$	-0.000944	[-0.001100, -0.000787]	-0.009438
$n = 15$	-0.000589	[-0.000751, -0.000426]	-0.008830
$n = 20$	-0.000476	[-0.000632, -0.000320]	-0.009524
$n = 25$	-0.000328	[-0.000488, -0.000169]	-0.008212

- Strike $K = 1$

Time steps	Estimated weak error	95% Confidence interval	$n\times$ Estimated weak error
$n = 5$	-0.002195	[-0.002396, -0.001995]	-0.010977
$n = 10$	-0.001121	[-0.001323, -0.000920]	-0.011214
$n = 15$	-0.000693	[-0.000903, -0.000484]	-0.010400
$n = 20$	-0.000580	[-0.000784, -0.000377]	-0.011608
$n = 25$	-0.000405	[-0.000610, -0.000199]	-0.010116

- Strike $K = 2$

Time steps	Estimated weak error	95% Confidence interval	$n\times$ Estimated weak error
$n = 5$	-0.001290	[-0.001659, -0.000920]	-0.006448
$n = 10$	-0.000529	[-0.000907, -0.000151]	-0.005291
$n = 15$	-0.000149	[-0.000546, 0.000249]	-0.002230
$n = 20$	-0.000154	[-0.000536, 0.000227]	-0.003089
$n = 25$	0.000007	[-0.000382, 0.000396]	0.000174

Part II

Option Pricing under Initial Margin Requirements

Chapter 4

Numerical approximation of anticipative McKean BSDEs arising in initial margin requirements

Ce chapitre est un article écrit avec Ankush Agarwal, Stefano De Marco, Emmanuel Gobet, José Lopez-Salas et Fanny Noubiagain dans le cadre de l'édition 2017 du CEMRACS. Il a été soumis à ESAIM: Proceedings and Surveys.

Abstract

We introduce a new class of anticipative backward stochastic differential equations with a dependence of McKean type on the law of the solution, that we name MKABSDE. We provide existence and uniqueness results in a general framework with relaxed regularity assumptions on the parameters. We show that such stochastic equations arise within the modern paradigm of derivative pricing where a central counterparty (CCP) demands each member to deposit variation and initial margins to cover their exposure. In the case when the initial margin is proportional to the Conditional Value-at-Risk (CVaR) of the contract price, we apply our general result to obtain existence and uniqueness of the price as a solution of a MKABSDE. We also provide several linear and non-linear approximations, which we solve using different numerical methods.

Keywords: non-linear pricing, CVaR initial margins, anticipative BSDE, weak non-linearity.

4.1 Initial margin and McKean Anticipative BSDE (MKABSDE)

4.1.1 Financial context and motivation

The paradigm of linear risk-neutral pricing of financial contracts has changed in the last few years, influenced by the regulators. Nowadays, banks and financial institutions have to post collateral to a central counterparty (CCP, also called clearing house) in order to secure their positions. Everyday, the CCP asks each member to post a certain amount according to the exposure of their Over-the-Counter (OTC) contracts. The variation and initial margin deposits correspond to collaterals in order to cover respectively a new contract at inception and the daily change in its market value on the one hand, and the possible mark-to-market loss during the liquidation period in case of default on the other hand (see, for example, [12] for details). In this work we focus only on the initial margin requirement (IM for short), and we investigate how it affects the valuation and hedging of the contract. As stated in [12, p.11 3(d)], *"IM protects the transacting parties from the potential future exposure that could arise from future changes in the mark-to-market value of the contract during the time it takes to close out and replace the position in the event that one or more counterparties default. The amount of initial margin reflects the size of the potential future exposure. It depends on a variety of factors, [...] the expected duration of the contract closeout and replacement period, and can change over time."* In this work, we will consider IM deposits that are proportional to the Conditional Value-at-Risk (CVaR) of the contract price over a future period of length Δ (typically $\Delta = 1$ week or 10 days, standing for the closeout and replacement period). We focus on CVaR rather than Value-at-Risk (VaR) due to its pertinent properties; it is indeed well established that CVaR is a coherent risk measure whereas VaR is not [8].

We make some distinctions in our analysis according to the way the *contract price* is computed in the presence of IM. While [12] refers to a mark-to-market value of the contract that can be seen as an *exogenous* value, we investigate the case where this value is *endogenous* and is given by the value of the hedging portfolio including the additional IM costs. By doing so, we introduce a new non-linear pricing rule, that is: the value of the hedging portfolio V_t together with its hedging component π_t solve a stochastic equation including a term depending on the law of the solution (due to the CVaR). We justify that this problem can be seen as a new type of anticipative Backward Stochastic Differential Equation (BSDE) with McKean interaction [79]. From now on, we refer to this kind of equation as MKABSDE, standing for McKean Anticipative BSDE; the subsection 4.1.2 gives a toy example of such a model. In Section 4.2, we derive stability estimates for these MKABSDEs, under general Lipschitz conditions, and prove existence and uniqueness results. In Section 4.3, we verify that these results can be applied to a general complete Itô market [71], accounting for IM requirements. Then, we derive some approximations based on classical non-linear BSDEs whose purpose is to quantify the impact of choosing the reference price for IM as exogenous or endogenous, and to compare with the case without IM. Essentially, in Theorem 4.3.1 we prove that the hedging portfolio with exogenous or endogenous reference price for IM coincide up to order 2 in Δ when Δ is small (which is compatible with Δ equal to few days). Section 4.4 is devoted to numerical experiments: we solve the different approximating BSDEs using finite difference methods in dimension 1, and nested Monte Carlo and regression Monte Carlo methods in higher dimensions.

4.1.2 An example of anticipative BSDE with dependence in law

We start with a simple financial example with IM requirements, in the case of a single tradable asset. A more general version with a multidimensional Itô market will be studied in Section 4.3. Let us assume that the price of a tradable asset, denoted S, evolves accordingly to a geometric Brownian motion

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t, \tag{4.1.1}$$

where $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$ and W is an one-dimensional Brownian motion.

In the classical financial setting (see, for example, [83]), consider the situation where a trader wants to sell a European option with maturity $T > 0$ and payoff $\Phi(S_T)$, and to hedge it dynamically with risky and riskless assets S and S^0, where $S^0_t = e^{rt}$ for $t \in [0, T]$ and r is a risk-free interest rate. We denote by (V, π), the value of the self-financing portfolio and π the amount of money invested in the risky asset, respectively. In order to ensure the replication of the payoff at maturity, the couple (V, π) should solve the following stochastic equation

$$\begin{cases} \mathrm{d}V_t &= r\left(V_t - \pi_t\right)\mathrm{d}t + \pi_t \dfrac{\mathrm{d}S_t}{S_t},\ t \in [0, T], \\ V_T &= \Phi(S_T). \end{cases} \tag{4.1.2}$$

Eq (4.1.2) is a BSDE since the terminal condition of V is imposed. Because all the coefficients are linear in V and π, (4.1.2) is a linear BSDE (see [72] for a broad overview on BSDEs and their applications in finance).

Accounting for IM requirement will introduce an additional cost in the above self-financing dynamics. We assume that the required deposit is proportional to the CVaR of the portfolio over Δ days (typically $\Delta = 10$ days) at the risk-level α (typically $\alpha = 99\%$). The funding cost for this deposit is determined by an interest rate R.[1] Therefore, the IM cost can be modelled as an additional term in the dynamics of the self-financing portfolio as

$$\mathrm{d}V_t = \big(r\,(V_t - \pi_t) - R\;\mathbf{CVaR}^{\alpha}_{\mathcal{F}_t}\,(V_t - V_{t+\Delta})\big)\,\mathrm{d}t + \pi_t \frac{\mathrm{d}S_t}{S_t}, \tag{4.1.3}$$

where the CVaR of a random variable L, conditional on the underlying sigma-field $\mathcal{F}_t$ at time t, is defined by (see [91])

$$\mathbf{CVaR}^{\alpha}_{\mathcal{F}_t}(L) = \inf_{x\in\mathbb{R}} \mathbb{E}\left[\frac{(L-x)^+}{1-\alpha} + x \middle| \mathcal{F}_t\right]. \tag{4.1.4}$$

Since $V_{t+\Delta}$ may be meaningless as t gets close to T, in (4.1.3) one should consider $V_{(t+\Delta)\wedge T}$ instead. Rewriting (4.1.3) in integral form together with the replication constraint, we obtain a BSDE

$$V_t = \Phi(S_T) + \int_t^T \big(-r\,(V_s - \pi_s) - \mu\pi_s + R\;\mathbf{CVaR}^{\alpha}_{\mathcal{F}_s}\,\big(V_s - V_{(s+\Delta)\wedge T}\big)\big)\,\mathrm{d}s - \int_t^T \pi_s \sigma \mathrm{d}W_s,\; t\in[0,T]. \tag{4.1.5}$$

The conditional CVaR term is anticipative and non-linear in the sense of McKean [79], for it involves the law of future variations of the portfolio conditional to the knowledge of the past. This is an example of McKean Anticipative BSDE, which we study in broader generality in Section 4.2.

Coming back to the financial setting, (V, π) stands for a valuation rule which treats the IM adjustment as endogenous (in the sense that CVaR is computed on V itself). One could alternatively consider that CVaR is related to an exogenous valuation (the so-called *mark-to-market*), for instance the one due to (4.1.2) (assuming that (4.1.2) models the market evolution of the option price). Later in Section 4.3, we give quantitative error bounds between these different valuation rules. Without advocating one with respect to the other, we rather compare their values and estimate (theoretically and numerically) how well one of their output prices approximates the others. As a consequence, these results may serve as a support for banks and regulators for improving risk management and margin requirement rules.

4.1.3 Literature review on anticipative BSDEs and comparison with our contribution

BSDEs were introduced by Pardoux and Peng [85]. Since then, the theoretical properties of BSDEs with different generators and terminal conditions have been extensively studied. The link between Markovian BSDEs and partial differential equations (PDEs) was studied in [86]. Under some smoothness assumptions, [86] established that the solution of the Markovian BSDE corresponds to the solution of a semi-linear parabolic PDE. In addition, several applications in finance have been proposed, in particular by El Karoui and co-authors [72] who considered the application to European option pricing in the constrained case. In fact, [72] showed that, under some constraints on the hedging strategy, the price of a contingent claim is given by the solution of a non-linear convex BSDE.

Recently, a new class of BSDEs called anticipated BSDEs (ABSDEs for short) was introduced by Peng and Yang [88]. The main feature of this class is that the generator includes not only the value of the solution at the present, but also at a future date. As in the classical theory of BSDEs, there exists a duality between these ABSDEs and stochastic differential delay equations. In [88] the existence, uniqueness and a comparison theorem for the solution is provided under a kind of Lipschitz condition which depends on the conditional expectation. One can also find more general formulations of ABSDE in Cheredito and Nam [29]. As in the case of classical BSDEs, the question of weakening the Lipschitz condition considered in [88] has been tackled by Yang and Elliott [99], who extended the existence theorem for ABSDEs from Lipschitz to continuous coefficients, and proved that the comparison theorem for anticipated BSDEs still holds. They also established a minimal solution.

[1] This interest rate corresponds to the difference of a funding rate minus the interest rate paid by the CCP for the deposit, typically $R \approx 3\%$

At the same time, Buckdahn and Imkeller [22] introduced the so-called time-delayed BSDEs (see also Delong and Imkeller [31, 32]). As opposed to the ABSDEs of [88], in this case the generator depends on the values of the solution at the present and at past dates, weighted with a time delay function. Assuming that the generator satisfies a certain kind of Lipschitz assumption depending on a probability measure, Delong and Imkeller [32] proved the existence and uniqueness of a solution for a sufficiently small time horizon or for a sufficiently small Lipschitz constant of the generator. These authors also showed that, when the generator is independent of y and for a small delay, existence and uniqueness hold for an arbitrary Lipschitz constant. Later, Delong and Imkeller [33] provided an application of time-delayed BSDEs to problems of pricing and hedging, and portfolio management. This work focuses on participating contracts and variable annuities, which are worldwide life insurance products with capital protections, and on claims based on the performance of an underlying investment portfolio.

More recently, Crépey *et al.* [30] have worked in a setting which is close to the problem we tackle here, introducing an application of ABSDEs to the problem of computing different types of valuation adjustments (XVAs) for derivative prices, related to funding (X=F), capital (X=K) and credit risk (X=C). In particular, they focus on the case where the initial margins of an OTC contract can be funded directly with the economic capital of the bank involved in the trade, giving rise to different terms in the price evolution equation. The connection of economic capital and funding valuation adjustment leads to an ABSDE, whose anticipated part consists of a conditional risk measure of the martingale increment of the solution over a future time period. These authors have showed that the system of ABSDEs formed by the FVA and the KVA processes is well-posed. Mathematically, the existence and uniqueness of the solution to the system is established through the convergence of Picard iterations.

Inspired by the dynamics of the self-financial portfolio in 4.1.5, we consider a new type of ABSDEs (the McKean ABSDEs) where the generator depends on the value of the solution, but also on the law of the whole trajectory between the present and a future date, possibly up to maturity. We state a priori estimates on the differences between the solutions of two such MKABSDEs. Based on these estimates, we derive existence and uniqueness of the solution to a MKABSDE via a fixed-point theorem.

4.2 A general McKean Anticipative BSDE

In order to give meaning to (4.1.5) and to more general (multidimensional) cases such as eq (4.3.2) below, we now introduce a general mathematical setup for studying existence and uniqueness of solutions.

4.2.1 Notation

Let $T > 0$ be the finite time horizon and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a d-dimensional Brownian motion, where $d \geq 1$. We denote $(\mathcal{F}_t)_{t\in[0,T]}$ the filtration generated by W, completed with the $\mathbb{P}$-null sets of $\mathcal{F}$. Let $t \in [0,T], \beta \geq 0$ and $m \in \mathbb{N}^*$. We will make use of the following notations:

- For any $a = (a_1, ..., a_m) \in \mathbb{R}^m$, $|a| = \sqrt{\sum_{i=1}^m a_i^2}$.
- Given a process $(x_s)_{s\in[0,T]}$, we set $x_{t:T} := (x_s)_{s\in[t,T]}$.
- $L^2_T(\mathbb{R}^m) = \left\{\mathbb{R}^m\text{-valued } \mathcal{F}_T\text{-measurable random variables } \xi \text{ such that } \mathbb{E}\left[|\xi|^2\right] < \infty\right\}$.
- $\mathbb{H}^2_{\beta,T}(\mathbb{R}^m) = \left\{\mathbb{R}^m\text{-valued and } \mathcal{F}\text{-adapted stochastic processes } \varphi \text{ such that } \mathbb{E}\left[\int_0^T e^{\beta t}|\varphi_t|^2 \mathrm{d}t\right] < \infty\right\}$. For $\varphi \in \mathbb{H}^2_{\beta,T}(\mathbb{R}^m)$, we define $||\varphi||_{\mathbb{H}^2_{\beta,T}} = \sqrt{\mathbb{E}\left[\int_0^T e^{\beta t}|\varphi_t|^2 \mathrm{d}t\right]}$.
- $\mathbb{S}^2_{\beta,T}(\mathbb{R}^m) = \left\{\text{Continuous processes } \varphi \in \mathbb{H}^2_{\beta,T}(\mathbb{R}^m) \text{ such that } \mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\varphi_t|^2\right] < \infty\right\}$. For $\varphi \in \mathbb{S}^2_{\beta,T}(\mathbb{R}^m)$, we define $||\varphi||_{\mathbb{S}^2_{\beta,T}} = \sqrt{\mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\varphi_t|^2\right]}$.

Note that $\mathbb{H}^2_{\beta,T}(\mathbb{R}^m) = \mathbb{H}^2_{0,T}(\mathbb{R}^m)$ and $\mathbb{S}^2_{\beta,T}(\mathbb{R}^m) = \mathbb{S}^2_{0,T}(\mathbb{R}^m)$, for any $\beta \geq 0$. The additional degree of freedom given by the parameter β in the definition of the space norm will be useful when deriving a priori estimates (see Lemma 4.2.2).

4.2.2 Main result

Our aim is to find a pair of processes $(Y,Z) \in \mathbb{S}^2_{0,T}(\mathbb{R}) \times \mathbb{H}^2_{0,T}\left(\mathbb{R}^d\right)$ satisfying

$$Y_t = \xi + \int_t^T f\left(s, Y_s, Z_s, \Lambda_s\left(Y_{s:T}\right)\right) \mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s, \qquad t \in [0,T], \tag{4.2.1}$$

for a certain mapping $\Lambda_t(\cdot)$ to be defined below. We call Equation (4.2.1) McKean Anticipative BSDE (MKABSDE) with parameters (f, Λ, ξ). In order to obtain existence and uniqueness of solutions, we require that the mappings f and Λ satisfy some suitable Lipschitz properties (specified below), and that the terminal condition ξ be square integrable.

Assumption (S). *For any $y, z, \lambda \in \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}$, $f(\cdot, y, z, \lambda)$ is a $\mathcal{F}$-adapted stochastic process with values in $\mathbb{R}$ and there exists a constant $C_f > 0$ such that almost surely, for all $(s, y_1, z_1, \lambda_1), (s, y_2, z_2, \lambda_2) \in [0,T]\times\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}$,*

$$\left|f\left(s, y_1, z_1, \lambda_1\right) - f\left(s, y_2, z_2, \lambda_2\right)\right| \le C_f\left(\left|y_1 - y_2\right| + \left|z_1 - z_2\right| + \left|\lambda_1 - \lambda_2\right|\right).$$

Moreover, $\mathbb{E}\left[\int_0^T \left|f(s,0,0,0)\right|^2 \mathrm{d}s\right] < \infty$.

Assumption (A). *For any $X \in \mathbb{S}^2_{0,T}(\mathbb{R})$, $\left(\Lambda_t\left(X_{t:T}\right)\right)_{t\in[0,T]}$ defines a stochastic process that belongs to $\mathbb{H}^2_{0,T}(\mathbb{R})$. There exist a constant $C_\Lambda > 0$ and a family of measures $\left(\nu_t\right)_{t\in[0,T]}$ on $\mathbb{R}$ such that for every $t \in [0,T]$, ν_t has support included in $[t,T]$, $\nu_t([t,T]) = 1$, and for any $y^1, y^2 \in \mathbb{S}^2_{0,T}(\mathbb{R})$, we have*

$$\left|\Lambda_t\left(y^1_{t:T}\right) - \Lambda_t\left(y^2_{t:T}\right)\right| \le C_\Lambda \mathbb{E}\left[\int_t^T \left|y^1_s - y^2_s\right| \nu_t(\mathrm{d}s) \,\middle|\, \mathcal{F}_t\right], \mathrm{d}t \otimes \mathrm{d}\mathbb{P} \text{ a.e. } .$$

Moreover, there exists a constant $\kappa > 0$ such that for every $\beta \ge 0$ and every continuous path $x : [0,T] \to \mathbb{R}$,

$$\int_0^T e^{\beta s} \int_s^T \left|x_u\right| \nu_s(\mathrm{d}u)\, \mathrm{d}s \le \kappa \sup_{t\in[0,T]} e^{\beta t}\left|x_t\right|.$$

We will say that a function $\tilde{f}$ (resp. a mapping $\tilde{\Lambda}$) satisfies Assumption (S) (resp. (A)) if that assumption holds for the choice $f = \tilde{f}$ (resp. $\Lambda = \tilde{\Lambda}$). We can now give the main result of this section.

Theorem 4.2.1. *Under Assumptions (S) and (A), for any terminal condition $\xi \in L^2_T(\mathbb{R})$ the BSDE (4.2.1) has a unique solution $(Y,Z) \in \mathbb{S}^2_{0,T}(\mathbb{R}) \times \mathbb{H}^2_{0,T}\left(\mathbb{R}^d\right)$.*

4.2.3 Proof of Theorem 4.2.1

The proof uses classical arguments. We first establish *a priori* estimates in the same spirit as in [72] on the solutions to the BSDE. Then for a suitable constant $\beta \ge 0$, we use Picard's fixed point method in the space $\mathbb{S}^2_{\beta,T}(\mathbb{R}) \times \mathbb{H}^2_{\beta,T}\left(\mathbb{R}^d\right)$ to obtain existence and uniqueness of a solution to Equation (4.2.1).

Lemma 4.2.2. *Let $\left(Y^1, Z^1\right), \left(Y^2, Z^2\right) \in \mathbb{S}^2_{0,T}(\mathbb{R}) \times \mathbb{H}^2_{0,T}\left(\mathbb{R}^d\right)$ be solutions to MKABSDE (4.2.1) associated respectively to the parameters $\left(f^1, \Lambda^1, \xi^1\right)$ and $\left(f^2, \Lambda^2, \xi^2\right)$. We assume that f^1 satisfies Assumption (S) and that Λ^1 satisfies Assumption (A). Let us define $\delta Y := Y^1 - Y^2$, $\delta Z := Z^1 - Z^2$, $\delta\xi := \xi^1 - \xi^2$. Finally, let us define for $s \in [0,T]$,*

$$\delta_2 f_s = f^1\left(s, Y^2_s, Z^2_s, \Lambda^2\left(Y^2_{s:T}\right)\right) - f^2\left(s, Y^2_s, Z^2_s, \Lambda^2\left(Y^2_{s:T}\right)\right), \quad \text{and} \quad \delta_2\Lambda_s = \Lambda^1_s\left(Y^2_{s:T}\right) - \Lambda^2_s\left(Y^2_{s:T}\right).$$

Then there exists a constant $C > 0$ such that for $\mu > 0$, we have for β large enough

$$||\delta Y||^2_{\mathbb{S}^2_{\beta,T}} \le C\left(e^{\beta T}\mathbb{E}\left[|\delta\xi|^2\right] + \frac{1}{\mu^2}\left(||\delta_2 f||^2_{\mathbb{H}^2_{\beta,T}} + C_{f^1}||\delta_2\Lambda||^2_{\mathbb{H}^2_{\beta,T}}\right)\right),$$
$$||\delta Z||^2_{\mathbb{H}^2_{\beta,T}} \le C\left(e^{\beta T}\mathbb{E}\left[|\delta\xi|^2\right] + \frac{1}{\mu^2}\left(||\delta_2 f||^2_{\mathbb{H}^2_{\beta,T}} + C_{f^1}||\delta_2\Lambda||^2_{\mathbb{H}^2_{\beta,T}}\right)\right).$$

Proof. The proof is based on similar arguments used in [72]. Let us use the decomposition:

$$
\begin{aligned}
&|f^1(s,Y_s^1,Z_s^1,\Lambda_s^1(Y_{s:T}^1))-f^2(s,Y_s^2,Z_s^2,\Lambda_s^2(Y_{s:T}^2))|\\
&\qquad\le \left|f^1(s,Y_s^1,Z_s^1,\Lambda_s^1(Y_{s:T}^1))-f^1(s,Y_s^2,Z_s^2,\Lambda_s^2(Y_{s:T}^2))\right|\\
&\qquad+\left|f^1(s,Y_s^2,Z_s^2,\Lambda_s^2(Y_{s:T}^2))-f^2(s,Y_s^2,Z_s^2,\Lambda_s^2(Y_{s:T}^2))\right|\\
&\qquad\le C_{f^1}\left(|\delta Y_s|+|\delta Z_s|+|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^2(Y_{s:T}^2)|\right)+|\delta_2 f_s|\\
&\qquad\le C_{f^1}\left(|\delta Y_s|+|\delta Z_s|+|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|+|\delta_2\Lambda_s|\right)+|\delta_2 f_s|.
\end{aligned}
$$

By Itô's lemma on the process $t\to e^{\beta t}|\delta Y_t|^2$, where $\beta\ge 0$, and using the previous inequality, we have that

$$
\begin{aligned}
&e^{\beta t}|\delta Y_t|^2+\beta\int_t^T e^{\beta s}|\delta Y_s|^2\mathrm{d}s+\int_t^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\\
&=e^{\beta T}|\delta\xi|^2+2\int_t^T e^{\beta s}\delta Y_s\left(f^1(s,Y_s^1,Z_s^1,\Lambda_s^1(Y_{s:T}^1))-f^2(s,Y_s^2,Z_s^2,\Lambda_s^2(Y_{s:T}^2))\right)\mathrm{d}s-2\int_t^T e^{\beta s}\delta Y_s\delta Z_s\mathrm{d}W_s\\
&\le e^{\beta T}|\delta\xi|^2+2\int_t^T e^{\beta s}|\delta Y_s|\left(C_{f^1}\left(|\delta Y_s|+|\delta Z_s|+|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|+|\delta_2\Lambda_s|\right)+|\delta_2 f_s|\right)\mathrm{d}s\\
&-2\int_t^T e^{\beta s}\delta Y_s\delta Z_s\mathrm{d}W_s.
\end{aligned}\tag{4.2.2}
$$

Applying Young's inequality with $\lambda,\mu\neq 0$, we have

$$
\begin{aligned}
&2|\delta Y_s|\left(C_{f^1}(|\delta Z_s|+|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|+|\delta_2\Lambda_s|)+|\delta_2 f_s|\right)\\
&\quad\le\frac{C_{f^1}}{\lambda^2}|\delta Z_s|^2+\lambda^2C_{f^1}|\delta Y_s|^2+\frac{C_{f^1}}{\lambda^2}|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|^2+\lambda^2C_{f^1}|\delta Y_s|^2\\
&\quad+\frac{C_{f^1}}{\mu^2}|\delta_2\Lambda_s|^2+\mu^2C_{f^1}|\delta Y_s|^2+\frac{1}{\mu^2}|\delta_2 f_s|^2+\mu^2|\delta Y_s|^2\\
&\quad\le\left(\mu^2+C_{f^1}(\mu^2+2\lambda^2)\right)|\delta Y_s|^2+\frac{C_{f^1}}{\lambda^2}|\delta Z_s|^2+\frac{C_{f^1}}{\lambda^2}|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|^2\\
&\quad+\frac{C_{f^1}}{\mu^2}|\delta_2\Lambda_s|^2+\frac{1}{\mu^2}|\delta_2 f_s|^2.
\end{aligned}
$$

Then plug this bound into (4.2.2) to get

$$
\begin{aligned}
&e^{\beta t}|\delta Y_t|^2+\beta\int_t^T e^{\beta s}|\delta Y_s|^2\mathrm{d}s+\int_t^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\\
&\quad\le e^{\beta T}|\delta\xi|^2+\left(\mu^2+C_{f^1}(2+\mu^2+2\lambda^2)\right)\int_t^T e^{\beta s}|\delta Y_s|^2\mathrm{d}s+\frac{C_{f^1}}{\lambda^2}\int_t^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\\
&\quad+\frac{C_{f^1}}{\lambda^2}\int_t^T e^{\beta s}|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|^2\mathrm{d}s+\frac{C_{f^1}}{\mu^2}\int_t^T e^{\beta s}|\delta_2\Lambda_s|^2\mathrm{d}s+\frac{1}{\mu^2}\int_t^T e^{\beta s}|\delta_2 f_s|^2\mathrm{d}s\\
&\quad-2\int_t^T e^{\beta s}\delta Y_s\delta Z_s\mathrm{d}W_s.
\end{aligned}\tag{4.2.3}
$$

Choosing $\lambda^2>C_{f^1}$ and

$$
\beta\ge\mu^2+C_{f^1}(2+\mu^2+2\lambda^2),\tag{4.2.4}
$$

we get from (4.2.3) that

$$
\begin{aligned}
\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\right]&\le\frac{\lambda^2}{\lambda^2-C_{f^1}}\mathbb{E}\left[e^{\beta T}|\delta\xi|^2+\frac{C_{f^1}}{\lambda^2}\int_t^T e^{\beta s}|\Lambda_s^1(Y_{s:T}^1)-\Lambda_s^1(Y_{s:T}^2)|^2\mathrm{d}s\right]\\
&+\frac{\lambda^2}{\lambda^2-C_{f^1}}\mathbb{E}\left[\frac{C_{f^1}}{\mu^2}\int_t^T e^{\beta s}|\delta_2\Lambda_s|^2\mathrm{d}s+\frac{1}{\mu^2}\int_t^T e^{\beta s}|\delta_2 f_s|^2\mathrm{d}s\right].
\end{aligned}
$$

Here we have used that the stochastic integral in (4.2.3) is a true martingale, by invoking $\delta Y \in \mathbb{S}^2_{0,T}(\mathbb{R})$, $\delta Z \in \mathbb{H}^2_{0,T}(\mathbb{R}^d)$, the computations like in (4.2.7) and a localization procedure. From (4.2.3) we also have that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\delta Y_t|^2 + \left(1-\frac{C_{f^1}}{\lambda^2}\right)\int_0^T e^{\beta s}|\delta Z_s|^2 \mathrm{d}s\right]\\
&\le \mathbb{E}\Bigg[e^{\beta T}|\delta\xi|^2 + \frac{C_{f^1}}{\lambda^2}\int_0^T e^{\beta s}|\Lambda^1_s(Y^1_{s:T}) - \Lambda^1_s(Y^2_{s:T})|^2\mathrm{d}s + \frac{C_{f^1}}{\mu^2}\int_0^T e^{\beta s}|\delta_2\Lambda_s|^2\mathrm{d}s + \frac{1}{\mu^2}\int_0^T e^{\beta s}|\delta_2 f_s|^2\mathrm{d}s\\
&\quad + 2\sup_{t\in[0,T]}\left|\int_t^T e^{\beta s}\delta Y_s \delta Z_s \mathrm{d}W_s\right|\Bigg]. \qquad (4.2.5)
\end{aligned}$$

As Λ^1 satisfies Assumption (A), the Jensen inequality yields that

$$\begin{aligned}
\mathbb{E}\left[\int_0^T e^{\beta s}|\Lambda^1_s(Y^1_{s:T}) - \Lambda^1_s(Y^2_{s:T})|^2\mathrm{d}s\right] &\le C_\Lambda^2\mathbb{E}\left[\int_0^T e^{\beta s}\int_s^T |\delta Y_u|^2\nu_s(\mathrm{d}u)\mathrm{d}s\right]\\
&\le \kappa C_\Lambda^2 \mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\delta Y_t|^2\right]. \qquad (4.2.6)
\end{aligned}$$

By the Burkholder-Davis-Gundy inequality, there exists a positive constant C_1 such that

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T e^{\beta s}\delta Y_s\delta Z_s\mathrm{d}W_s\right|\right] &\le C_1\mathbb{E}\left[\left(\int_0^T e^{2\beta s}|\delta Y_s|^2|\delta Z_s|^2\mathrm{d}s\right)^{1/2}\right]\\
&\le C_1\mathbb{E}\left[\left(\sup_{s\in[0,T]} e^{\beta s}|\delta Y_s|^2\right)^{1/2}\left(\int_0^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\right)^{1/2}\right]. \qquad (4.2.7)
\end{aligned}$$

Therefore, by Young's inequality with $\gamma > 0$, we have

$$\begin{aligned}
2\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T e^{\beta s}\delta Y_s\delta Z_s\mathrm{d}W_s\right|\right] &\le \frac{C_1}{\gamma^2}\mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\delta Y_t|^2\right] + \gamma^2 C_1\mathbb{E}\left[\int_0^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\right]\\
&\le \frac{C_1}{\gamma^2}\mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\delta Y_t|^2\right] + \frac{\gamma^2 C_1\lambda^2}{\lambda^2 - C_{f^1}}\mathbb{E}\Bigg[e^{\beta T}|\delta\xi|^2 + \frac{C_{f^1}}{\mu^2}\int_t^T e^{\beta s}|\delta_2\Lambda_s|^2\mathrm{d}s\\
&\quad + \frac{1}{\mu^2}\int_t^T e^{\beta s}|\delta_2 f_s|^2\mathrm{d}s + \frac{C_{f^1}}{\lambda^2}\int_t^T e^{\beta s}|\Lambda^1_s(Y^1_{s:T}) - \Lambda^1_s(Y^2_{s:T})|^2\mathrm{d}s\Bigg]. \qquad (4.2.8)
\end{aligned}$$

Combining Inequalities (4.2.5)–(4.2.8) leads to

$$\begin{aligned}
&\left(1 - \frac{C_1}{\gamma^2} - \frac{\kappa C_{f^1}C_\Lambda^2}{\lambda^2} - \frac{\kappa C_{f^1}C_1C_\Lambda^2\gamma^2}{\lambda^2 - C_{f^1}}\right)\mathbb{E}\left[\sup_{t\in[0,T]} e^{\beta t}|\delta Y_t|^2\right] + \left(1 - \frac{C_{f^1}}{\lambda^2}\right)\mathbb{E}\left[\int_0^T e^{\beta s}|\delta Z_s|^2\mathrm{d}s\right]\\
&\le \left(1 + \frac{\gamma^2 C_1\lambda^2}{\lambda^2 - C_{f^1}}\right)\left(\mathbb{E}\left[e^{\beta T}|\delta\xi|^2\right] + \frac{1}{\mu^2}\mathbb{E}\left[C_{f^1}\int_0^T e^{\beta s}|\delta_2\Lambda_s|^2\mathrm{d}s + \int_0^T e^{\beta s}|\delta_2 f_s|^2\mathrm{d}s\right]\right).
\end{aligned}$$

Let us define the continuous function Γ by

$$\Gamma(\gamma,\lambda) = 1 - \frac{C_1}{\gamma^2} - \frac{\kappa C_{f^1}C_\Lambda^2}{\lambda^2} - \frac{\kappa C_{f^1}C_1C_\Lambda^2\gamma^2}{\lambda^2 - C_{f^1}}$$

for any $\gamma > 0$ and any $\lambda > 0$ with $\lambda^2 > C_{f^1}$. Observe that if we set $\gamma(\lambda) = \sqrt{\lambda}$ with $\lambda > 0$, we have $\lim_{\lambda\to\infty}\Gamma(\gamma(\lambda),\lambda) = 1$, so there exist λ,γ large enough such that $\Gamma(\gamma,\lambda) > 0$. For such a choice of γ and λ, we then obtain the announced result with the constant

$$C = \frac{1 + \frac{C_1\gamma^2\lambda^2}{\lambda^2 - C_{f^1}}}{\min\left(\Gamma(\gamma,\lambda), 1 - \frac{C_{f^1}}{\lambda^2}\right)}.$$

Recall that β is large enough according to λ (see inequality (4.2.4)). $\square$

Proof of Theorem 4.2.1. We use the previous apriori estimates in the case where (Y^1, Z^1) and (Y^2, Z^2) solve respectively the BSDEs

$$Y_t^1 = \xi + \int_t^T f_s\left(U_s^1, V_s^1, \Lambda_s\left(U_{s:T}^1\right)\right) \mathrm{d}s - \int_t^T Z_s^1 \mathrm{d}W_s,$$
$$Y_t^2 = \xi + \int_t^T f_s\left(U_s^2, V_s^2, \Lambda_s\left(U_{s:T}^2\right)\right) \mathrm{d}s - \int_t^T Z_s^2 \mathrm{d}W_s.$$

Here, $(U^1, V^1), (U^2, V^2) \in \mathbb{S}_{0,T}^2(\mathbb{R}) \times \mathbb{H}_{0,T}^2(\mathbb{R}^d)$ are given processes. Therefore $f_s\left(U_s^1, V_s^1, \Lambda_s\left(U_{s:T}^1\right)\right)$ and $f_s\left(U_s^2, V_s^2, \Lambda_s\left(U_{s:T}^2\right)\right)$ define processes in $\mathbb{H}_{0,T}^2(\mathbb{R})$ owing to the Assumptions (S) and (A). Therefore, the existence and uniqueness of (Y^1, Z^1) and (Y^2, Z^2) in $\mathbb{S}_{0,T}^2(\mathbb{R}) \times \mathbb{H}_{0,T}^2(\mathbb{R}^d)$ as solutions of standard BSDEs is automatic (see [72, Theorem 2.1, Proposition 2.2]).

In addition, the process $Y^1 - Y^2$ is then solution to the BSDE $Y_t^1 - Y_t^2 = \int_t^T \delta_2 f_s \mathrm{d}s - \int_t^T \left(Z_s^1 - Z_s^2\right) \mathrm{d}W_s$, where the driver $\delta_2 f_s = f_s\left(U_s^1, V_s^1, \Lambda_s\left(U_{s:T}^1\right)\right) - f_s\left(U_s^2, V_s^2, \Lambda_s\left(U_{s:T}^2\right)\right)$ does not depend on Y_s^1 nor Y_s^2. Using Lemma 4.2.2 for $C_f = 0$ and $\mu > 0$, we have that for $\beta > 0$ large enough,

$$||\delta Y||_{\mathbb{S}_{\beta,T}^2}^2 + ||\delta Z||_{\mathbb{H}_{\beta,T}^2}^2 \le \frac{C}{\mu^2} ||\delta_2 f||_{\mathbb{H}_{\beta,T}^2}^2.$$

Moreover,

$$\begin{aligned} ||\delta_2 f||_{\mathbb{H}_{\beta,T}^2}^2 &= \mathbb{E}\left[\int_0^T e^{\beta s} |f_s\left(U_s^1, V_s^1, \Lambda_s\left(U_{s:T}^1\right)\right) - f_s\left(U_s^2, V_s^2, \Lambda_s\left(U_{s:T}^2\right)\right)|^2 \mathrm{d}s\right] \\ &\le 3C_f^2 \mathbb{E}\left[\int_0^T e^{\beta s}\left(|\delta U_s|^2 + |\delta V_s|^2 + |\Lambda_s\left(U_{s:T}^1\right) - \Lambda_s\left(U_{s:T}^2\right)|^2\right) \mathrm{d}s\right]. \end{aligned}$$

As we have that $||\delta U||_{\mathbb{H}_{\beta,T}^2}^2 \le T ||\delta U||_{\mathbb{S}_{\beta,T}^2}^2$, and

$$\mathbb{E}\left[\int_0^T e^{\beta s} |\Lambda_s\left(U_{s:T}^1\right) - \Lambda_s\left(U_{s:T}^2\right)|^2 \mathrm{d}s\right] \le C_\Lambda^2 \mathbb{E}\left[\int_0^T e^{\beta s} \int_s^T |\delta U_u|^2 \nu_s(\mathrm{d}u) \, \mathrm{d}s\right] \le \kappa C_\Lambda^2 ||\delta U||_{\mathbb{S}_{\beta,T}^2}^2.$$

We obtain that

$$||\delta Y||_{\mathbb{S}_{\beta,T}^2}^2 + ||\delta Z||_{\mathbb{H}_{\beta,T}^2}^2 \le \frac{3CC_f^2}{\mu^2}\left(\left(\kappa C_\Lambda^2 + T\right) ||\delta U||_{\mathbb{S}_{\beta,T}^2} + ||\delta V||_{\mathbb{H}_{\beta,T}^2}^2\right).$$

We now choose $\mu^2 > 3CC_f^2(\kappa C_\Lambda^2 + T + 1)$, and obtain that for β large enough, the mapping $\phi : (U, V) \to (Y, Z)$ is a contraction in the space $\mathbb{S}_{\beta,T}^2(\mathbb{R}) \times \mathbb{H}_{\beta,T}^2(\mathbb{R}^d)$. Hence, we get existence and uniqueness of a solution to the BSDE (4.2.1). □

4.3 The Case of CVaR initial margins

In this section, we apply the previous results on MKABSDE to equation (4.1.5) and to its generalizations (with respect to the dimension of S, the underlying dynamic model and the terminal condition) that will be defined below. Beyond usual existence and uniqueness results, our aim is to analyse related approximations, obtained when CVaR is evaluated using Gaussian expansions (justified as $\Delta \to 0$, see Theorem 4.3.1).

4.3.1 A well posed problem

Let us consider a general Itô market with d tradable assets [71, Chapter 1]. The riskless asset S^0 (money account) follows the dynamics $\frac{\mathrm{d}S_t^0}{S_t^0} = r_t \mathrm{d}t$, and we have d risky assets $\left(S^1, ..., S^d\right)$ following

$$\frac{\mathrm{d}S_t^i}{S_t^i} = \mu_t^i \mathrm{d}t + \sum_{j=1}^d \sigma_t^{ij} \mathrm{d}W_t^j, \; S_0^i = s_0^i \in \mathbb{R}, \; 1 \le i \le d. \tag{4.3.1}$$

The processes $r, \mu := \left(\mu^i\right)_{1\leq i\leq d}, \sigma := \left(\sigma^{ij}\right)_{1\leq i,j\leq d}$ are $\mathcal{F}$-adapted stochastic processes with values respectively in $\mathbb{R}, \mathbb{R}^d$, and the set of matrices of size $d\times d$. Moreover, we assume that $\mathrm{d}t\otimes\mathrm{d}\mathbb{P}$ a.e., the matrix σ_t is invertible and the processes r and $\sigma^{-1}(\mu - r\mathbf{1})$ are uniformly bounded, where we define the column vector $\mathbf{1} := (1, ..., 1)^\top \in \mathbb{R}^d$. For a path-dependent payoff ξ paid at maturity T, the dynamics of the hedging portfolio (V, π) with CVaR initial margin requirement (over a period $\Delta > 0$) is given by

$$V_t = \xi + \int_t^T \left(-r_s V_s + \pi_s\left(r_s\mathbf{1} - \mu_s\right) + R\ \mathbf{CVaR}^\alpha_{\mathcal{F}_s}\left(V_s - V_{(s+\Delta)\wedge T}\right)\right)\mathrm{d}s - \int_t^T \pi_s\sigma_s\mathrm{d}W_s.$$

Here π is a row vector whose ith coordinate consists of the amount invested in ith asset. The derivation is analogous to that of Section 4.1.2. This equation rewrites, with the variables $(V, Z = \pi\sigma)$,

$$V_t = \xi + \int_t^T \left(-r_s V_s + Z_s\sigma_s^{-1}\left(r_s\mathbf{1} - \mu_s\right) + R\ \mathbf{CVaR}^\alpha_{\mathcal{F}_s}\left(V_s - V_{(s+\Delta)\wedge T}\right)\right)\mathrm{d}s - \int_t^T Z_s\mathrm{d}W_s. \tag{4.3.2}$$

Existence and uniqueness of a solution to the above MKABSDE are consequences of Theorem 4.2.1.

Corollary 4.3.1. *For any square integrable terminal condition ξ, the CVaR initial margin problem (4.3.2) is well posed with a unique solution $(V, Z) \in \mathbb{S}^2_{\beta,T}(\mathbb{R}) \times \mathbb{H}^2_{\beta,T}(\mathbb{R}^d)$ for any $\beta \geq 0$.*

Proof. The driver of the BSDE has the form

$$f(t, v, z, \lambda) = -r_t v + z\sigma_t^{-1}\left(r_t\mathbf{1} - \mu_t\right) + \lambda,\ t \geq 0,\ v, \lambda \in \mathbb{R},\ z \in \mathbb{R}^d,$$

and we also introduce the functional

$$\Lambda_t\left(X_{t:T}\right) := R\ \mathbf{CVaR}^\alpha_{\mathcal{F}_t}\left(X_t - X_{(t+\Delta)\wedge T}\right) = R\inf_{x\in\mathbb{R}}\mathbb{E}\left[\frac{\left(X_t - X_{(t+\Delta)\wedge T} - x\right)^+}{1-\alpha} + x \mid \mathcal{F}_t\right],\ t \in [0,T],\ X \in \mathbb{S}^2_{0,T}(\mathbb{R}).$$

Since r and $\sigma^{-1}(\mu - r\mathbf{1})$ are uniformly bounded, f clearly satisfies Assumption (S). We now check that Λ satisfies Assumption (A). For $X \in \mathbb{S}^2_{0,T}(\mathbb{R})$ and $x \in \mathbb{R}$, we have

$$\mathbb{E}\left[X_t - X_{(t+\Delta)\wedge T}\middle|\mathcal{F}_t\right] \leq \inf_{x\in\mathbb{R}}\mathbb{E}\left[\frac{\left(X_t - X_{(t+\Delta)\wedge T} - x\right)^+}{1-\alpha} + x\middle|\mathcal{F}_t\right] \leq \mathbb{E}\left[\frac{\left(X_t - X_{(t+\Delta)\wedge T}\right)^+}{1-\alpha}\middle|\mathcal{F}_t\right], \tag{4.3.3}$$

where for the left hand side (l.h.s.) we use the fact that as $\alpha \in (0,1)$, for $z, x \in \mathbb{R}$, $\frac{(z-x)^+}{1-\alpha} + x \geq z$, and for the right hand side (r.h.s.), we upper bound the infimum with the value taken at $x = 0$. As it is easy to check that both the l.h.s. and the r.h.s. of (4.3.3) belong to $\mathbb{H}^2_{0,T}$, we conclude that $\Lambda_t(X) \in \mathbb{H}^2_{0,T}(\mathbb{R})$. Now, let $X^1, X^2 \in \mathbb{S}^2_{0,T}$. Then, we have that

$$\begin{aligned}
\left|\Lambda_t\left(X^1\right) - \Lambda_t\left(X^2\right)\right| &\leq R\mathbb{E}\left[\left|\frac{X^1_t - X^2_t - \left(X^1_{(t+\Delta)\wedge T} - X^2_{(t+\Delta)\wedge T}\right)}{1-\alpha}\right| \mid \mathcal{F}_t\right] \\
&\leq \frac{2R}{1-\alpha}\mathbb{E}\left[\int_t^T \left|X^1_s - X^2_s\right|\nu_t(\mathrm{d}s) \mid \mathcal{F}_t\right],
\end{aligned}$$

where for the first inequality, we use the fact that $\left|\inf_{x\in\mathbb{R}} g^1(x) - \inf_{x\in\mathbb{R}} g^2(x)\right| \leq \sup_{x\in\mathbb{R}}\left|g^1(x) - g^2(x)\right|$ for any functions $g^1, g^2 : \mathbb{R} \to \mathbb{R}$ and the 1-Lipschitz property of the positive part function, and for the second inequality, $\nu_t(\mathrm{d}s) := \frac{1}{2}\delta_t(\mathrm{d}s) + \frac{1}{2}\delta_{(t+\Delta)\wedge T}(\mathrm{d}s)$, where for $u \geq 0$, δ_u is the Dirac measure on $\{u\}$. Moreover, for $\beta \geq 0$,

$$\int_0^T e^{\beta s}\int_s^T \left|X^1_u\right|\nu_s(\mathrm{d}u) = \frac{1}{2}\int_0^T e^{\beta t}\left(\left|X^1_t\right| + \left|X^1_{(t+\Delta)\wedge T}\right|\right)\mathrm{d}t \leq \sup_{t\in[0,T]} e^{\beta t}\left|X^1_t\right|,$$

so Assumption (S) holds with $\kappa = 1$ and $C_\Lambda = \frac{2R}{1-\alpha}$. We finally apply Theorem 4.2.1 to complete the proof. $\square$

4.3.2 Approximation by standard BSDEs when $\Delta \ll 1$

The numerical solution of (4.3.2) is challenging in full generality. In fact, it is a priori more difficult than solving a standard BSDE, for which we can employ, for example, regression Monte-Carlo methods (see e.g. [51] and references therein). In this work, we take advantage of the fact that Δ is small (recall Δ = one week or 10 days) in order to provide handier approximations of (V, Z), given in terms of standard non-linear or linear BSDEs. Below we define these different BSDEs and provide the error estimates of such approximations.

At the lowest order in the parameter $\sqrt{\Delta}$, for $s \in [0, T]$, formally we have that, conditionally to $\mathcal{F}_s$,

$$V_s - V_{(s+\Delta)\wedge T} \approx -\int_s^{(s+\Delta)\wedge T} Z_u \mathrm{d}W_u \overset{(d)}{=} -|Z_s|\sqrt{(s+\Delta)\wedge T - s} \times G,$$

where we freeze the process Z at current time s and $G \overset{(d)}{=} \mathcal{N}(0,1)$ is independent from $\mathcal{F}_s$. This is an approximation of CVaR using the "Delta" of the portfolio (see [48, Section 2]). Plugging this approximation into (4.3.2), and defining

$$C_\alpha := \mathbf{CVaR}^\alpha(\mathcal{N}(0,1)) = \left.\frac{e^{-x^2/2}}{(1-\alpha)\sqrt{2\pi}}\right|_{x=\mathcal{N}^{-1}(\alpha)}, \tag{4.3.4}$$

we obtain a standard non-linear BSDE

$$V_t^{NL} = \xi + \int_t^T \left(-r_s V_s^{NL} + Z_s^{NL}\sigma_s^{-1}(r_s\mathbf{1} - \mu_s) + R\, C_\alpha \sqrt{(s+\Delta)\wedge T - s}|Z_s^{NL}|\right) \mathrm{d}s - \int_t^T Z_s^{NL} \mathrm{d}W_s. \tag{4.3.5}$$

Seeing V^{NL} as a function of the small parameter Δ appearing in the driver, and making an expansion at the orders 0 and 1 w.r.t. $\sqrt{\Delta}$ by following the expansion procedure in [50], we obtain two linear BSDEs, respectively (V^{BS}, Z^{BS}) and (V^L, Z^L) where

$$V_t^{BS} = \xi + \int_t^T \left(-r_s V_s^{BS} + Z_s^{BS}\sigma_s^{-1}(r_s\mathbf{1} - \mu_s)\right) \mathrm{d}s - \int_t^T Z_s^{BS} \mathrm{d}W_s, \tag{4.3.6}$$

$$V_t^{L} = \xi + \int_t^T \left(-r_s V_s^{L} + Z_s^{L}\sigma_s^{-1}(r_s\mathbf{1} - \mu_s) + RC_\alpha \sqrt{(s+\Delta)\wedge T - s}|Z_s^{BS}|\right) \mathrm{d}s - \int_t^T Z_s^{L} \mathrm{d}W_s. \tag{4.3.7}$$

Let us comment on these different models.

- The simplest equation is (V^{BS}, Z^{BS}), corresponding to the usual linear valuation rule [72, Theorem 1.1] without IM requirement. When the model is a one-dimensional geometric Brownian motion and $\xi = (S_T - K)_+$, the solution is given by the usual Black-Scholes formula.
- The second simplest equation is (V^L, Z^L) where the IM cost is computed using the "Delta" of an exogenous reference price given by the simplest pricing rule (V^{BS}, Z^{BS}) without IM. This is still a linear BSDE but its simulation is not simple though, since one needs to know Z^{BS} to simulate (V^L, Z^L). We use a nested Monte-Carlo procedure in our experiments.
- The third equation is (V^{NL}, Z^{NL}) where the IM cost is computed using the "Delta" of the endogenous price (V^{NL}, Z^{NL}) itself.

Existence and uniqueness of a solution to the BSDEs (4.3.5), (4.3.6) and (4.3.7) are direct consequences of [85], as the respective drivers satisfy standard Lipschitz properties and the processes r and $\sigma^{-1}(r\mathbf{1} - \mu)$ are bounded.

Proposition 4.3.1. *The standard BSDEs (4.3.5), (4.3.6) and (4.3.7) have a unique solution in the L_2-space $\mathbb{S}^2_{0,T} \times \mathbb{H}^2_{0,T}$, and their norms are uniformly bounded in $\Delta \le T$.*

The main result of this part is the following theorem.

Theorem 4.3.1. *Define the L_2 time-regularity index of Z^{NL} by*

$$\mathcal{E}^{NL}(\Delta) := \frac{1}{\Delta}\mathbb{E}\left[\int_0^T \int_t^{(t+\Delta)\wedge T} \left|Z_s^{NL} - Z_t^{NL}\right|^2 \mathrm{d}s\mathrm{d}t\right]. \tag{4.3.8}$$

We always have $\sup_{0<\Delta\leq T}\mathcal{E}^{NL}(\Delta) < +\infty$*. Moreover, there exist constants* $K_1, K_2, K_3 > 0$*, independent from* Δ*, such that*

$$||V^L - V^{BS}||^2_{\mathbb{S}^2_{0,T}} + ||Z^L - Z^{BS}||^2_{\mathbb{H}^2_{0,T}} \leq K_1\Delta, \tag{4.3.9}$$

$$||V^{NL} - V^L||^2_{\mathbb{S}^2_{0,T}} + ||Z^{NL} - Z^L||^2_{\mathbb{H}^2_{0,T}} \leq K_2\Delta^2, \tag{4.3.10}$$

$$||V - V^{NL}||^2_{\mathbb{S}^2_{0,T}} + ||Z - Z^{NL}||^2_{\mathbb{H}^2_{0,T}} \leq K_3\Delta\left(\Delta + \mathcal{E}^{NL}(\Delta)\right). \tag{4.3.11}$$

In addition, we have

$$\mathcal{E}^{NL}(\Delta) = O(\Delta), \tag{4.3.12}$$

and thus $||V - V^{NL}||^2_{\mathbb{S}^2_{0,T}} + ||Z - Z^{NL}||^2_{\mathbb{H}^2_{0,T}} = O(\Delta^2)$ *provided that the additional sufficient conditions below are fulfilled:*

(i) the terminal condition is a Lipschitz functional of S*, that is,* $\xi = \Phi(S_{0:T})$ *for some functional* Φ *satisfying*

$$|\Phi(x_{0:T}) - \Phi(x'_{0:T})| \leq C_\Phi \sup_{t\in[0,T]} |x_t - x'_t|,$$

for any continuous paths $x, x' : [0,T] \to \mathbb{R}^d$*;*

(ii) the coefficients r, σ, μ *are constant.*

Let us remark that the results from [102] used in the proof of estimate (4.3.12) and consequently the estimate (4.3.12) itself should also hold under (i) and the following more general assumptions:

(iii) the processes r, σ, μ are Markovian, i.e. $r_t = \hat{r}(t, S_t)$, $\sigma^{ij}_t = \hat{\sigma}^{ij}(t, S_t)$ and $\mu^i_t = \hat{\mu}^i(t, S_t)$ for some deterministic functions $\hat{r}, \hat{\mu}^i, \hat{\sigma}^{ij}$;

(iv) the functions $x \to \hat{\mu}^i(x)x_i$, $x \to \hat{\sigma}^{ij}(x)x_i$ are globally Lipschitz in $(t,x) \in [0,T]\times\mathbb{R}^d$, for any $1 \leq i,j \leq d$;

(v) the functions $\hat{r}$ and $\hat{\sigma}^{-1}\left(\hat{r}\mathbf{1} - \hat{\mu}\right)$ are globally Lipschitz in $(t,x) \in [0,T]\times\mathbb{R}^d$.

As mentioned above, we may expect that $\mathcal{E}^{NL}(\Delta) = O(\Delta)$ also under (i)-(iii)-(iv)-(v), so that (V, Z) and (V^{NL}, Z^{NL}) are very close to each other. These approximation results illustrate that there is a significative difference (at the order of $\sqrt{\Delta}$) between valuation with or without initial margin cost (see (4.3.9)); however, the other valuation rules yield comparable values as soon as $\Delta \ll 1$ (see (4.3.10)-(4.3.11)).

4.3.3 Proof of Theorem 4.3.1

$\triangleright$ *Estimate on* $\mathcal{E}^{NL}(\Delta)$. We start with a deterministic inequality. For any positive function Ψ and any $\beta \geq 0$, we have

$$\int_0^T e^{\beta t}\left(\int_t^{(t+\Delta)\wedge T}\Psi_s \mathrm{d}s\right)\mathrm{d}t \leq \Delta\int_0^T e^{\beta s}\Psi_s\mathrm{d}s. \tag{4.3.13}$$

Indeed the left hand side of (4.3.13) can be written as

$$\int_0^T\int_0^T e^{\beta t}\Psi_s\mathbf{1}_{t\leq s\leq(t+\Delta)\wedge T}\mathrm{d}s\mathrm{d}t = \int_0^T \Psi_s\left(\int_0^T e^{\beta t}\mathbf{1}_{t\leq s\leq(t+\Delta)\wedge T}\mathrm{d}t\right)\mathrm{d}s, \tag{4.3.14}$$

which readily gives the announced result. Using $(a+b)^2 \leq 2a^2 + 2b^2$ and (4.3.13) with $\beta = 0$ gives

$$\mathcal{E}^{NL}(\Delta) \leq \frac{2}{\Delta}\mathbb{E}\left[\int_0^T\int_t^{(t+\Delta)\wedge T}(|Z^{NL}_s|^2 + |Z^{NL}_t|^2)\mathrm{d}s\mathrm{d}t\right] \leq 4\mathbb{E}\left[\int_0^T\left|Z^{NL}_t\right|^2\mathrm{d}t\right], \tag{4.3.15}$$

which is uniformly bounded in Δ (Proposition 4.3.1).

We now derive finer estimates that reveal the L_2 time-regularity of Z^{NL} under the extra assumptions (i)-(ii). In this Markovian setting, we know that Z^{NL} has a càdlàg version (see [102, Remark (ii) after Lemma 2.5]). Then, introduce the equidistant times $t_i = i\Delta$ for $0 \le i \le n := \lfloor \frac{T}{\Delta} \rfloor$ and $t_{n+1} = T$. We claim that

$$\mathcal{E}^{NL}(\Delta) \le 4 \sum_{i=0}^{n} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \left|Z_s^{NL} - Z_{t_i}^{NL}\right|^2 + \left|Z_s^{NL} - Z_{t_{i+1}}^{NL}\right|^2 \mathrm{d}s\right]. \tag{4.3.16}$$

With this result at hand, the estimate (4.3.12) directly follows from an application of [102, Theorem 3.1]. To get (4.3.16), set $\varphi^-(s)$ and $\varphi^+(s)$ for the grid times before and after s. Then, we write

$$\begin{aligned}
\int_0^T \int_t^{(t+\Delta)\wedge T} |Z_s^{NL} - Z_t^{NL}|^2 \mathrm{d}s\mathrm{d}t &\le 2 \sum_{i=0}^{n} \int_{t_i}^{t_{i+1}} \int_t^{(t+\Delta)\wedge T} (|Z_s^{NL} - Z_{t_{i+1}}^{NL}|^2 + |Z_t^{NL} - Z_{t_{i+1}}^{NL}|^2)\mathrm{d}s\mathrm{d}t \\
&\le 2 \int_0^T \int_t^{(t+\Delta)\wedge T} \left(|Z_s^{NL} - Z_{\varphi^-(s)}^{NL}|^2 + |Z_s^{NL} - Z_{\varphi^+(s)}^{NL}|^2\right) \mathrm{d}s\mathrm{d}t + 2\Delta \int_0^T |Z_t^{NL} - Z_{\varphi^+(t)}^{NL}|^2 \mathrm{d}t \\
&\le 4\Delta \left(\int_0^T |Z_t^{NL} - Z_{\varphi^-(t)}^{NL}|^2 \mathrm{d}t + \int_0^T |Z_t^{NL} - Z_{\varphi^+(t)}^{NL}|^2 \mathrm{d}t\right)
\end{aligned}$$

where we have used (4.3.13) with $\beta = 0$. The inequality (4.3.16) readily follows.
▷ *Proof of (4.3.9).* This error estimate is related to the difference of two linear BSDEs. The drivers of V^{BS} and V^L are respectively $f^{BS}(s, y, z, \lambda) = -r_s y + z\sigma_s^{-1}(r_s\mathbf{1} - \mu_s)$, and

$$f^L(s, y, z, \lambda) = -r_s y + z\sigma_s^{-1}(r_s\mathbf{1} - \mu_s) + RC_\alpha\sqrt{(s+\Delta)\wedge T - s}|Z_s^{BS}|,$$

for $s \in [0, T]$, $v, \lambda \in \mathbb{R}$ and $z \in \mathbb{R}^d$, hence

$$\left(f^L - f^{BS}\right)(s, y, z, \lambda) = RC_\alpha\sqrt{(s+\Delta)\wedge T - s}|Z_s^{BS}|.$$

By Lemma 4.2.2, we obtain that for $\mu > 0$, β large enough and $K_1 = \frac{C}{\mu^2}(RC_\alpha)^2 ||Z^{BS}||_{\mathbb{H}^2_{\beta,T}}$,

$$||V^L - V^{BS}||^2_{\mathbb{S}^2_{0,T}} + ||Z^L - Z^{BS}||^2_{\mathbb{H}^2_{0,T}} \quad \le \quad ||V^L - V^{BS}||^2_{\mathbb{S}^2_{\beta,T}} + ||Z^L - Z^{BS}||^2_{\mathbb{H}^2_{\beta,T}} \le K_1\Delta.$$

We are done with (4.3.9).

▷ *Proof of (4.3.10).* Then, as $\xi \in L^2$, as the processes $r, \sigma^{-1}(\mu - r\mathbf{1})$ are bounded and as the non-linear term

$$t, z \in [0, T] \times \mathbb{R}^d \to RC_\alpha\sqrt{(t+\Delta)\wedge T - t}|z|$$

is Lipschitz in the variable z, uniformly in time, we obtain Inequality (4.3.10) as an application of [50, Theorem 2.4], for which assumptions **H.1** $-$ **H.3** are satisfied.

▷ *Proof of (4.3.11).* Using computations similar to those in the proof of Lemma 4.2.2, we obtain existence of a constant $C > 0$ such that for $\mu > 0$ and β large enough,

$$||V - V^{NL}||^2_{\mathbb{S}^2_{\beta,T}} + ||Z - Z^{NL}||^2_{\mathbb{H}^2_{\beta,T}} \le \frac{C}{\mu^2}\left|\left| \mathbf{CVaR}^\alpha_{\mathcal{F}_\cdot}\left(V_\cdot^{NL} - V_{(\cdot+\Delta)\wedge T}^{NL}\right) - \mathbf{CVaR}^\alpha_{\mathcal{F}_\cdot}\left(-\int_\cdot^{(\cdot+\Delta)\wedge T} Z_\cdot^{NL} \mathrm{d}W_s\right)\right|\right|^2_{\mathbb{H}^2_{\beta,T}}.$$

As the CVaR function is subadditive [91], we have that given A, B two random variables, and $t \in [0, T]$, $\mathbf{CVaR}^\alpha_{\mathcal{F}_t}(A) \le \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(B) + \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(A - B)$. Inverting the roles of A and B, we obtain that

$$\begin{aligned}
0 \le \left| \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(A) - \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(B)\right| &\le \max\left(\mathbf{CVaR}^\alpha_{\mathcal{F}_t}(A - B), \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(B - A)\right) \\
&\le \frac{1}{1-\alpha}\left(\mathbb{E}\left[(A - B)^+ \mid \mathcal{F}_t\right] + \mathbb{E}\left[(B - A)^+ \mid \mathcal{F}_t\right]\right) = \frac{\mathbb{E}\left[|A - B| \mid \mathcal{F}_t\right]}{1-\alpha},
\end{aligned}$$

where for the last inequality, we have used that for $U \in \{A - B, B - A\}$, $\inf_{x\in\mathbb{R}} \mathbb{E}\left[\frac{(U-x)^+}{1-\alpha} + x \mid \mathcal{F}_t\right] \le \mathbb{E}\left[\frac{U^+}{1-\alpha} \mid \mathcal{F}_t\right]$. We then have that

$$\left| \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(A) - \mathbf{CVaR}^\alpha_{\mathcal{F}_t}(B)\right|^2 \le \frac{1}{(1-\alpha)^2}\mathbb{E}\left[(A - B)^2 \mid \mathcal{F}_t\right].$$

Setting, for $t \in [0,T]$, $A_t = -\int_t^{(t+\Delta)\wedge T} Z_t^{NL} \mathrm{d}W_s$ and $B_t = V_t^{NL} - V_{(t+\Delta)\wedge T}^{NL}$ and using the previous inequality, we obtain that

$$||V - V^{NL}||^2_{\mathbb{S}^2_{\beta,T}} + ||Z - Z^{NL}||^2_{\mathbb{H}^2_{\beta,T}} \leq \frac{C}{\mu^2 (1-\alpha)^2} \mathbb{E}\left[\int_0^T e^{\beta t} \left(V_t^{NL} - V_{(t+\Delta)\wedge T}^{NL} + \int_t^{(t+\Delta)\wedge T} Z_t^{NL} \mathrm{d}W_s\right)^2 \mathrm{d}t\right].$$

We use the following decomposition,

$$\begin{aligned} V_t^{NL} - V_{(t+\Delta)\wedge T}^{NL} + \int_t^{(t+\Delta)\wedge T} Z_t^{NL} \mathrm{d}W_s &= \int_t^{(t+\Delta)\wedge T} \left(-r_s V_s^{NL} + Z_s^{NL} \sigma_s^{-1} (r_s \mathbf{1} - \mu_s) + R\lambda C_\alpha \sqrt{(s+\Delta)\wedge T - s} |Z_s^{NL}|\right) \mathrm{d}s \\ &\quad - \int_t^{(t+\Delta)\wedge T} \left(Z_s^{NL} - Z_t^{NL}\right) \mathrm{d}W_s =: \Pi_1(t) - \Pi_2(t), \end{aligned}$$

so that $||V - V^{NL}||^2_{\mathbb{S}^2_{\beta,T}} + ||Z - Z^{NL}||^2_{\mathbb{H}^2_{\beta,T}} \leq \frac{2C}{\mu^2(1-\alpha)^2} \mathbb{E}\left[\int_0^T e^{\beta t} \left(\Pi_1^2(t) + \Pi_2^2(t)\right) \mathrm{d}t\right]$. By Jensen's inequality and the inequality (4.3.13), we get

$$\begin{aligned} \mathbb{E}\left[\int_0^T e^{\beta t} \Pi_1^2(t) \mathrm{d}t\right] &\leq 3\Delta \mathbb{E}\left[\int_0^T e^{\beta t} \int_t^{(t+\Delta)\wedge T} \left(\left(-r_s V_s^{NL}\right)^2 + \left(Z_s^{NL} \sigma_s^{-1} (r_s \mathbf{1} - \mu_s)\right)^2 + \left(R\lambda C_\alpha \sqrt{\Delta} |Z_s^{NL}|\right)^2\right) \mathrm{d}s \mathrm{d}t\right] \\ &\leq 3\Delta^2 \left(|r|^2_\infty + |\sigma^{-1}(r\mathbf{1} - \mu)|^2_\infty + (RC_\alpha)^2\right) \mathbb{E}\left[\int_0^T e^{\beta t} (|V_t^{NL}|^2 + |Z_t^{NL}|^2) \mathrm{d}t\right]. \end{aligned}$$

By invoking the uniform estimate of Proposition 4.3.1, we finally obtain that $\mathbb{E}\left[\int_0^T e^{\beta t} \left(\Pi_1(t)\right)^2 \mathrm{d}t\right] \leq \tilde{K}_1 \Delta^2$ for some $\tilde{K}_1$. Moreover, using Ito's isometry, we have that

$$\mathbb{E}\left[\int_0^T e^{\beta t} \Pi_2^2(t) \mathrm{d}t\right] = \mathbb{E}\left[\int_0^T e^{\beta t} \int_t^{(t+\Delta)\wedge T} \left|Z_s^{NL} - Z_t^{NL}\right|^2 \mathrm{d}s \mathrm{d}t\right] \leq e^{\beta T} \Delta \mathcal{E}^{NL}(\Delta).$$

Gathering all the previous arguments leads to the (4.3.11). The proof of Theorem 4.3.1 is completed. □

4.4 Numerical Examples

In the absence of numerical methods to estimate the solution of the McKean Anticipative BSDE (4.3.2) in full generality, we rather solve numerically the BSDE approximations (4.3.5) or (4.3.7) as discussed in Section 4.3.2. For this purpose, when the dimension d is greater than one, we use the Stratified Regression Multistep-forward Dynamical Programming (SRMDP) scheme developed in [49]. In our numerical tests in this section, we set the coefficients of the model (4.3.1) to be constant (multi-dimensional geometric Brownian motion) and we take $\mu^i = r$. Observe that setting $R = 0$ reduces the original BSDE to the linear equation (4.3.6). This will serve us as a benchmark value in order to measure the impact of Initial Margins (IM).

4.4.1 Finite difference method for (V^{NL}, Z^{NL}) in dimension 1

In order to check the validity of our results, we first obtain a benchmark when $d = 1$ by solving the semi-linear parabolic PDE related to the BSDE (4.3.5) when $\xi = \Phi(S_T)$, see [87]. By an application of Itô's lemma, the semi-linear PDE is given by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + C_\alpha R \sigma \sqrt{(t+\Delta)\wedge T - t} \left|S\frac{\partial V}{\partial S}\right| - r V = 0, \quad (t,S) \in [0,T) \times \mathbb{R}^+, \tag{4.4.1}$$

$$V(T,S) = \Phi(S), \quad S \in \mathbb{R}^+, \tag{4.4.2}$$

and $(V_t^{NL}, Z_t^{NL}) = (V(t,S_t), \frac{\partial V}{\partial S}(t,S_t)\sigma S_t)$.

Remark 4.4.1. *If $\Phi(S) = \max(S-K,0)$ or $\max(K-S,0)$ for some $K > 0$, i.e., either a call or a put option payoff, we expect the gradient $\frac{\partial V}{\partial S}$ to have a constant sign. In such a case, the PDE (4.4.1)–(4.4.2) becomes linear and in fact has an explicit solution, given by a Black-Scholes formula with time-dependent continuous dividend yield $d(t) = -C_\alpha R \sigma \sqrt{(t+\Delta)\wedge T - t}\ \mathrm{sign}(\frac{\partial V}{\partial S})$.*

We use a classical finite difference methods to solve (4.4.1)-(4.4.2) (see, for example, [3]). First, we perform a change of variable, $x = \ln S$, so that the PDE can be rewritten in the following form for the function $v(t,x) := V(t, e^x)$:

$$\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial v}{\partial x} + C_\alpha R\sigma\sqrt{(t+\Delta)\wedge T - t}\left|\frac{\partial v}{\partial x}\right| - r\,v = 0, \quad (t,x) \in [0,T) \times \mathbb{R}, \tag{4.4.3}$$

$$v(T,x) = \Phi(e^x), \quad x \in \mathbb{R}. \tag{4.4.4}$$

We denote the finite difference domain by $D = [0,T] \times [x_{\min}, x_{\max}]$ with $-\infty < x_{\min} < x_{\max} < \infty$. The domain D is approximated with a uniform mesh $\mathcal{D} = \{(t^n, x_i) : n = 0,1,\ldots,N,\ i = 0,1,\ldots,M\}$, where $t^n := n\Delta t$ and $x_i := x_{\min} + i\Delta x$. Here, for N time intervals, $\Delta t = T/N$ and $\Delta x = (x_{\max} - x_{\min})/M$ for M spatial steps. Furthermore, we denote $v(t^n, x_i) = v_i^n$. Next, consider the following finite difference derivative approximations under the well-known ω-scheme, i.e., we replace v_i^n by $\omega v_i^n + (1-\omega)v_i^{n+1}$, where $\omega \in [0,1]$ is a constant parameter, such that

$$\frac{\partial v}{\partial t}(t^n, x_i) \approx \frac{v_i^{n+1} - v_i^n}{\Delta t}, \quad \frac{\partial v}{\partial x}(t_n, x_i) \approx \omega\frac{v_{i+1}^n - v_i^n}{\Delta x} + (1-\omega)\frac{v_{i+1}^{n+1} - v_i^{n+1}}{\Delta x},$$
$$\frac{\partial^2 v}{\partial x^2}(t_n, x_i) \approx \omega\frac{v_{i+1}^n - 2v_i^n + v_{i-1}^n}{(\Delta x)^2} + (1-\omega)\frac{v_{i+1}^{n+1} - 2v_i^{n+1} + v_{i-1}^{n+1}}{(\Delta x)^2}.$$

The choice $\omega = 0.5$ corresponds to Crank-Nicolson method. We also "linearize" the non-linear term by treating it as explicit, i.e., at any time t_n we take $\left|\frac{\partial v}{\partial x}(t^n, x_i)\right| \approx \left|\frac{\partial v}{\partial x}(t^{n+1}, x_i)\right|$.

The substitution of finite difference derivative approximations in (4.4.3)-(4.4.4) along with the "linearization" step, leads to the following tridiagonal linear system at each time step $n = N-1, \ldots, 0$ which can be solved by Thomas algorithm [101]: $Av^n = b^{n+1}$, with nonzero coefficients of the tridiagonal matrix $A = (a_{i,j})$ given by

$$\begin{aligned} &a_{0,0} = 1, && a_{0,1} = 0, && a_{M,M-1} = 0, \quad a_{M,M} = 1,\\ &a_{i,i} = 1 + 2\theta\omega + \kappa\omega + \rho\omega, && a_{i,i+1} = -\theta\omega - \kappa\omega, && a_{i-1,i} = -\theta\omega, \quad i = 1,\ldots,M-1, \end{aligned}$$

and the time dependent vector b^{n+1} as:

$$\begin{aligned} b_0^{n+1} = v_0^{n+1}, \quad b_M^{n+1} = v_M^{n+1}, b_i^{n+1} &= \theta(1-\omega)v_{i-1}^{n+1} + (1 - 2\theta(1-\omega) - \kappa(1-\omega) - \rho(1-\omega))v_i^{n+1}\\ &+ (\theta(1-\omega) + \kappa(1-\omega))v_{i+1}^{n+1} + \beta^n|v_{i+1}^{n+1} - v_i^{n+1}|, \quad i = 1,\ldots,M-1, \end{aligned}$$

where v_0^{n+1}, v_M^{n+1} are given by the boundary conditions and the remaining constants are defined as below

$$\theta = \frac{\sigma^2\Delta t}{2(\Delta x)^2}, \quad \kappa = \frac{(r - \frac{1}{2}\sigma^2)\Delta t}{\Delta x}, \quad \rho = r\Delta t, \quad \beta^n = \frac{C_\alpha R\sigma\Delta t\sqrt{(t_n+\Delta)\wedge T - t_n}}{\Delta x}.$$

The ith coordinate of vector v^n is the approximation of the value $v(t^n, x_i)$.

We set the model parameters as $T = 1$, $\sigma = 0.25$, $r = 0.02$, $\alpha = 0.99$ and $\Delta = 0.02$ (1 week) and consider three different options – call, put and butterfly, for different strikes. We set $R = 0.02$ when accounting for IM and $R = 0$ otherwise. The finite difference space domain is taken as $[\ln(10^{-6}), \ln(4K)]$ while for SRMDP algorithm we take $[-5,5]$ to be the space domain. Furthermore, for finite difference scheme, $N = 10^3$ and $M = 10^6$. For **LP0** version of SRMDP algorithm, the number of hypercubes are 2800, the number of time steps are 50 and the number of simulations per hypercube are 2500. In Figure 4.1 and Table 4.1, we present the results for implied volatilities, prices and deltas of several call options including not only *at the money* strike but also *in* and *out of the money* strikes. First, we compute the values using the classical Black-Scholes formula (B-S $R = 0$) in order to allow the reader to assess the impact of taking into account IM. Next, we solve the non-linear BSDE using the three discussed method:; the exact Black-Scholes formula where IM is considered as a time dependent dividend yield (BS $R = 0.02$) (see Remark 4.4.1), the finite difference method (FD) and SRMDP algorithm. For the last method, we compute 95% confidence intervals for price and delta of the options. We also present the results of several put options in Figure 4.2 and Table 4.2. In any case, we observe that the IM has a significant impact on the Implied Volatility of option prices (about 20-30 bps for usual prices). Finally, we consider butterfly options with payoff function

$$\Phi(S_T) = (S_T - (K-2))^+ - 2\,(S_T - K)^+ + (S_T - (K+2))^+.$$

This derivative product involves three options with different strikes, the investor buys a call option with low strike price $K-2$, buys a call option with high strike price $K+2$ and sells two call options with strike price K. Note that the sign of the first derivative of option price (delta) of a butterfly option varies with the value of the underlying asset, therefore explicit Black-Scholes formula is not available when IM is also taken into consideration ($R = 0.02$). The results are presented in Figure 4.3 and Table 4.3. In all the three cases, we observe that SRMDP algorithm provides good accuracy when compared to the true values and finite difference estimates.

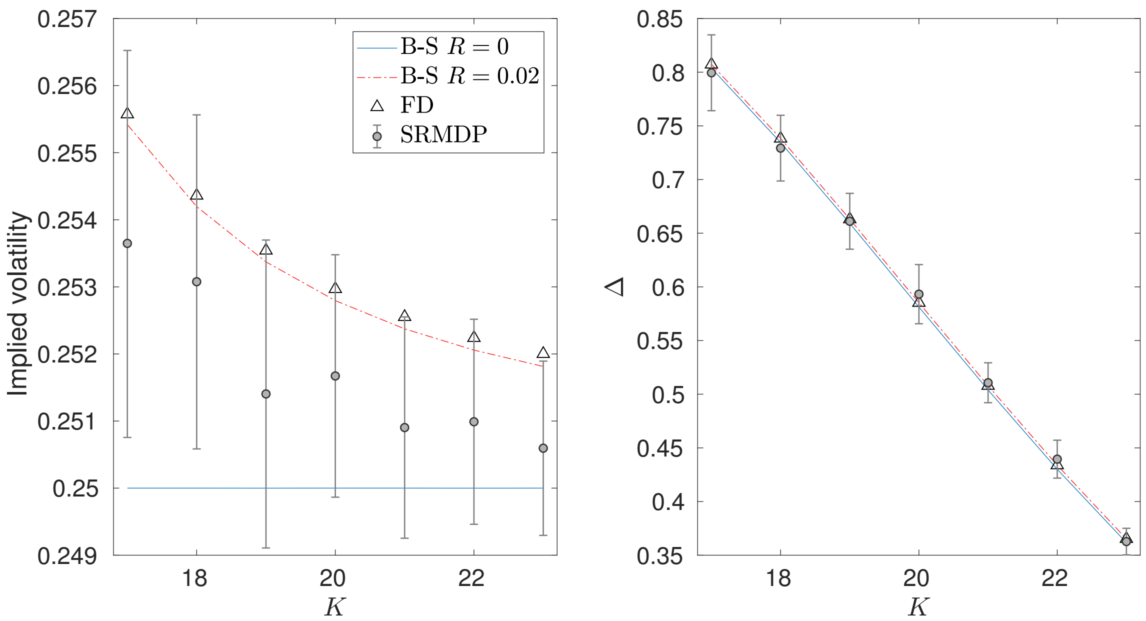


Figure 4.1: Implied volatility and delta for call options with spot value $S_0 = 20$ and different strikes K

	B-S $R=0$		B-S $R=0.02$		FD		SRMDP	
K	B-S$(0,S_0)$	$\Delta(0,S_0)$	B-S$(0,S_0)$	$\Delta(0,S_0)$	$V(0,S_0)$	$\nabla V(0,S_0)$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)/(\sigma S_0)$
17	3.9534	0.8037	3.9835	0.8073	3.9844	0.8072	$[3.9575, 3.9897]$	$[0.7641, 0.8347]$
18	3.2795	0.7345	3.3071	0.7383	3.3082	0.7382	$[3.2833, 3.3161]$	$[0.6986, 0.7598]$
19	2.6863	0.6592	2.7111	0.6631	2.7123	0.6630	$[2.6797, 2.7134]$	$[0.6350, 0.6871]$
20	2.1741	0.5812	2.1959	0.5852	2.1973	0.5852	$[2.1730, 2.2012]$	$[0.5656, 0.6207]$
21	1.7398	0.5039	1.7587	0.5079	1.7601	0.5078	$[1.7338, 1.7601]$	$[0.4920, 0.5292]$
22	1.3777	0.4301	1.3939	0.4338	1.3953	0.4338	$[1.3734, 1.3975]$	$[0.4218, 0.4571]$
23	1.0805	0.3617	1.0941	0.3651	1.0954	0.3652	$[1.0752, 1.0946]$	$[0.3503, 0.3750]$

Table 4.1: Price and delta for call options with spot value $S_0 = 20$ and different strikes K.

4.4.2 Variance reduction for solving (V^{NL}, Z^{NL}) using (V^{BS}, Z^{BS})

In order to asses the impact of using $R > 0$ on the solution of the BSDE (4.3.5), in the case of European call and put options in one dimension, it is better to solve the BSDE difference $(V_t^{DF}, Z_t^{DF}) = (V_t^{NL} - V_t^{BS}, Z_t^{NL} - Z_t^{BS})$ which has a reduced variance in the algorithm. Note that for a call option

$$Z_t^{BS} = \sigma S_t \Phi(d_1), \quad d_1 = \frac{\ln(S_t/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

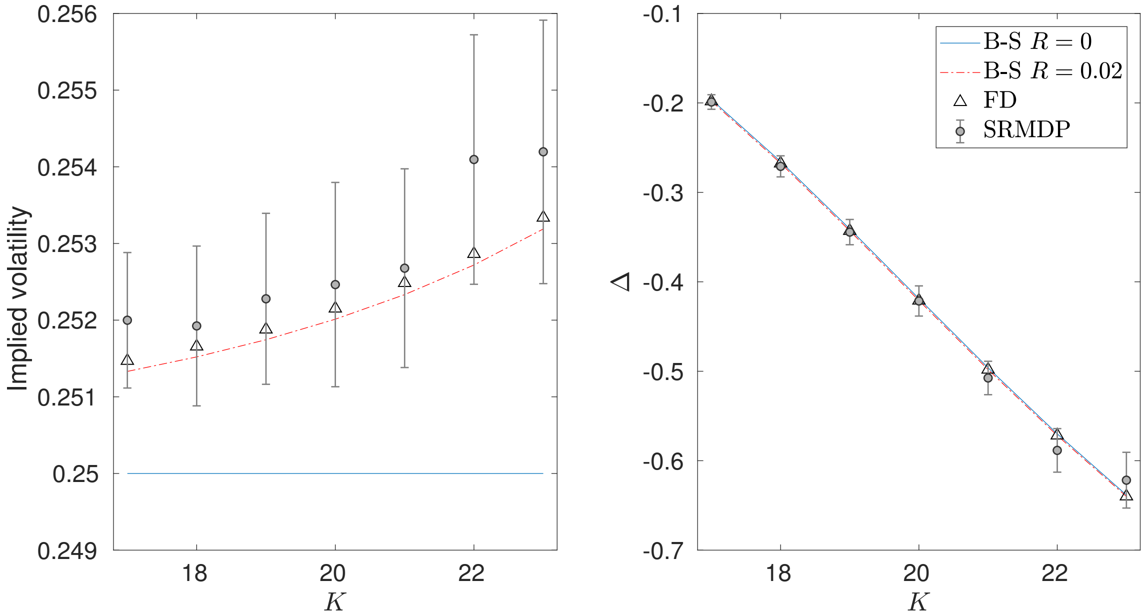


Figure 4.2: Implied volatility and delta for put options with spot value $S_0 = 20$ and different strikes K.

	B-S $R = 0$		B-S $R = 0.02$		FD		SRMDP	
K	B-S$(0, S_0)$	$\Delta(0, S_0)$	B-S$(0, S_0)$	$\Delta(0, S_0)$	$V(0, S_0)$	$\nabla V(0, S_0)$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)/(\sigma S_0)$
17	0.6168	-0.1963	0.6241	-0.1980	0.6249	-0.1981	$[0.6229, 0.6328]$	$[-0.2071, -0.1908]$
18	0.9231	-0.2655	0.9331	-0.2675	0.9340	-0.2676	$[0.9289, 0.9426]$	$[-0.2827, -0.2592]$
19	1.3101	-0.3408	1.3229	-0.3429	1.3239	-0.3430	$[1.3186, 1.3350]$	$[-0.3585, -0.3303]$
20	1.7781	-0.4188	1.7938	-0.4209	1.7949	-0.4209	$[1.7869, 1.8077]$	$[-0.4383, -0.4046]$
21	2.3240	-0.4961	2.3426	-0.4981	2.3438	-0.4981	$[2.3350, 2.3557]$	$[-0.5262, -0.4888]$
22	2.9421	-0.5699	2.9635	-0.5718	2.9646	-0.5717	$[2.9615, 2.9871]$	$[-0.6128, -0.5641]$
23	3.6251	-0.6383	3.6490	-0.6400	3.6501	-0.6398	$[3.6436, 3.6695]$	$[-0.6529, -0.5906]$

Table 4.2: Price and delta for put options with spot value $S_0 = 20$ and different strikes K.

where Φ is the standard Gaussian cumulative distribution function. Therefore $|Z_t^{NL}| = |Z_t^{DF} + \sigma S_t \Phi(d_1)|$. Then, the BSDE for the difference (V_t^{DF}, Z_t^{DF}) in the case of a call option[2] is given by:

$$V_t^{DF} = 0 + \int_t^T \left(-rV_s^{DF} + C_\alpha R\sqrt{(s+\Delta)\wedge T - s}|Z_s^{DF} + \sigma S_s \Phi(d_1)|\right) \mathrm{d}s - \int_t^T Z_s^{DF} \mathrm{d}W_s.$$

In Table 4.4, the BSDE (V_t^{DF}, Z_t^{DF}) is solved for several call and put options using the SRMDP algorithm. Besides, exact solutions (ES) are computed through the difference between Black-Scholes formula where IM's contribution is considered as a time-dependent dividend yield and the classical Black-Scholes formula with $R = 0$.

Once again these tests allow us to demonstrate that SRMDP algorithm provides accurate results in one dimension.

4.4.3 Nested Monte Carlo for computing (V^L, Z^L) in dimension 1

As discussed in Section 4.3.2, we can further approximate the solution of non-linear BSDE V^{NL} with a linear BSDE V^L with portfolio weight $Z = Z^{BS}$. In this case, we have an explicit stochastic representation for Z_t^{BS}

[2]For a put option an analogous BSDE can be written taking into account that $Z_t^{BS} = \sigma S_t(\Phi(d_1) - 1)$.

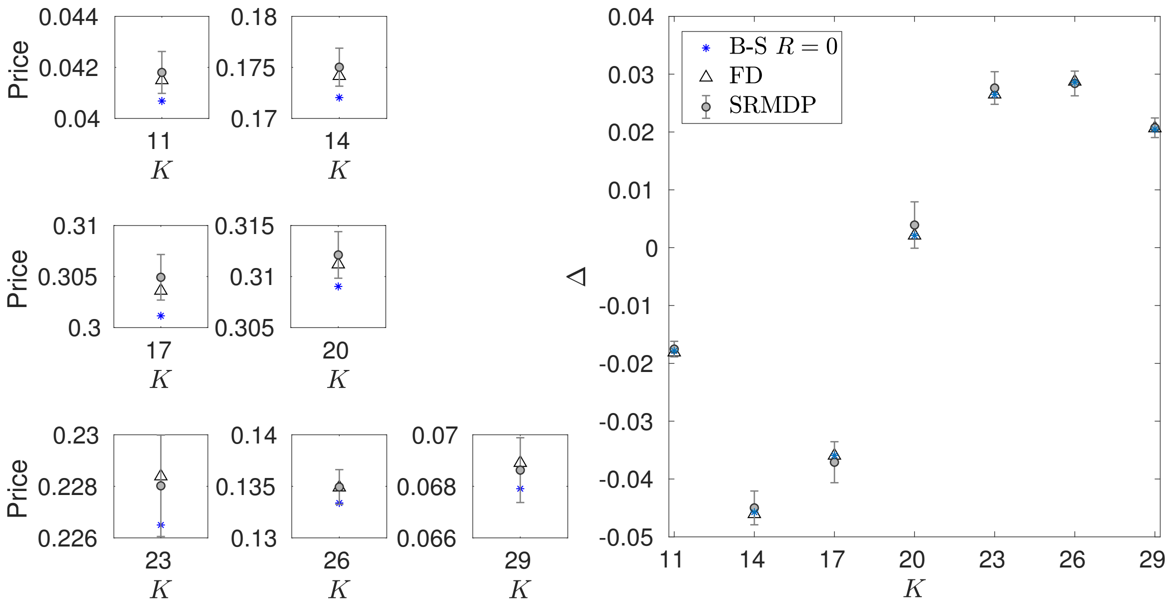


Figure 4.3: Price and delta for butterfly options with spot value $S_0 = 20$ and different strikes K.

	B-S $R=0$		FD		SRMDP	
K	B-S$(0, S_0)$	$\Delta(0, S_0)$	$V(0, S_0)$	$\nabla V(0, S_0)$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)/(\sigma S_0)$
11	0.0407	−0.0178	0.0415	−0.0181	[0.0410, 0.0426]	[−0.0189, −0.0162]
14	0.1720	−0.0457	0.1742	−0.0461	[0.1732, 0.1769]	[−0.0479, −0.0421]
17	0.3012	−0.0359	0.3036	−0.0359	[0.3027, 0.3072]	[−0.0406, −0.0336]
20	0.3090	0.0021	0.3112	0.0021	[0.3098, 0.3144]	[−0.0001, 0.0079]
23	0.2265	0.0265	0.2284	0.0265	[0.2261, 0.2300]	[0.0248, 0.0304]
26	0.1334	0.0286	0.1349	0.0287	[0.1333, 0.1366]	[0.0263, 0.0305]
29	0.0679	0.0205	0.0689	0.0207	[0.0674, 0.0699]	[0.0190, 0.0224]

Table 4.3: Price and delta for butterfly options with spot value $S_0 = 20$ and different strikes K.

	ES: B-S $R = 0.02$ − B-S $R = 0$		SRMDP	
	B-S$(0, S_0)$	$\Delta(0, S_0)$	95% CI $V_0^{DF}(S_0)$	95% CI $Z_0^{DF}(S_0)/(\sigma S_0)$
Call, $K = 17$	0.0302	0.0036	[0.0302, 0.0304]	[0.0033, 0.0037]
Call, $K = 20$	0.0218	0.0040	[0.0218, 0.0219]	[0.0038, 0.0042]
Call, $K = 23$	0.0136	0.0035	[0.0136, 0.0137]	[0.0033, 0.0036]
Put, $K = 17$	0.0074	−0.0017	[0.0074, 0.0075]	[−0.0020, −0.0016]
Put, $K = 20$	0.0157	−0.0021	[0.0157, 0.0158]	[−0.0023, −0.0019]
Put, $K = 23$	0.0239	−0.0016	[0.0239, 0.0240]	[−0.0018, −0.0015]

Table 4.4: SRMDP algorithm for BSDE (V_t^{DF}, Z_t^{DF}).

given as follows

$$Z_t^{BS} = \frac{\partial V_t^{BS}}{\partial S}(\sigma S_t)^{-1},$$

where $V_t^{BS}(s) := \mathbb{E}[e^{-r(T-t)}\Phi(S_T)|S_t = s]$. Then, we use the likelihood ratio method of Broadie and Glasserman [21] to find out the derivative and get

$$Z_t^{BS}(s) = \frac{\partial}{\partial s}\mathbb{E}[e^{-r(T-t)}\Phi(X_T)|S_t = s](\sigma s)^{-1} = \mathbb{E}\left[e^{-r(T-t)}\Phi(S_T)\frac{W_T - W_t}{(T-t)}(\sigma s)\Big|S_t = s\right](\sigma s)^{-1}$$
$$= \mathbb{E}\left[e^{-r(T-t)}\left(\Phi(S_T) - \Phi(S_t)\right)\frac{W_T - W_t}{(T-t)}\Big|S_t = s\right]. \tag{4.4.5}$$

Therefore, in linear BSDE V^L with portfolio weight Z^{BS}, we have

$$V_0^L = \mathbb{E}\left[e^{-rT}\Phi\left(S_T\right) + \int_0^T e^{-rs}\left(RC_\alpha|Z_s^{BS}|\sqrt{(s+\Delta)\wedge T - s}\right)ds\right]$$
$$= \mathbb{E}\left[e^{-rT}\Phi\left(S_T\right) + Te^{-rU}\left(RC_\alpha|Z_U^{BS}|\sqrt{(U+\Delta)\wedge T - U}\right)\right],$$

where U is a uniformly distributed independent random variable on $[0,T]$ and Z_s^{BS} is given as in (4.4.5). By once again using the likelihood ratio method, we get the following formula

$$Z_0^L = \mathbb{E}\left[e^{-rT}\Phi\left(S_T\right)\frac{W_T}{T} + Te^{-rU}\left(RC_\alpha|Z_U^{BS}|\sqrt{(U+\Delta)\wedge T - U}\right)\frac{W_U}{U}\right].$$

We solve the linear BSDE V^L by taking advantage of finite difference method and Nested Monte Carlo algorithm (Nested MC) for different payoffs (calls, puts and butterfly options) where we use the same model parameters as earlier. In Nested Monte Carlo algorithm, we estimate Z^{BS} as in (4.4.5) using 100 independent inner sample paths for each outer Monte Carlo sample path. The results are presented in Table 4.5. We observe that as the Nested Monte Carlo algorithm results are accurate in one dimension, the algorithm provides an alternative to compute the estimates for (V^L, Z^L) in higher dimensions.

	FD		Nested MC	
			95% CI	95% CI
	$V(0,S_0)$	$\nabla V(0,S_0)$	$V_0^L(S_0)$	$Z_0^L(S_0)/(\sigma S_0)$
Call, $K=17$	3.9843	0.8072	$[3.9796, 3.9856]$	$[0.8059, 0.8082]$
Call, $K=20$	2.1971	0.5852	$[2.1931, 2.1979]$	$[0.5843, 0.5862]$
Call, $K=23$	1.0953	0.3653	$[1.0924, 1.0958]$	$[0.3644, 0.3659]$
Put, $K=17$	0.6249	-0.1981	$[0.6233, 0.6249]$	$[-0.1983, -0.1977]$
Put, $K=20$	1.7950	-0.4209	$[1.7937, 1.7966]$	$[-0.4216, -0.4205]$
Put, $K=23$	3.6502	-0.6398	$[3.6468, 3.6511]$	$[-0.6407, -0.6393]$
Butterfly, $K=11$	0.0414	-0.0181	$[0.0412, 0.0415]$	$[-0.0181, -0.0180]$
Butterfly, $K=20$	0.3112	0.0021	$[0.3112, 0.3119]$	$[0.0021, 0.0022]$
Butterfly, $K=29$	0.0689	0.0206	$[0.0686, 0.0690]$	$[0.0206, 0.0207]$

Table 4.5: Nested MC algorithm for BSDE (V_t^L, Z_t^L).

4.4.4 Basket options in higher dimensions

In this section we solve the non-linear BSDE in high dimensions using SRMDP algorithm. In this setting, traditional full grid methods like finite difference are not able to tackle the problem for dimension greater than 3.

We consider call option on a basket of d assets where the asset process is modelled by multi-dimensional geometric Brownian motion with constant correlation $\rho_{ij} = \rho = 0.75$ for $i \neq j$ and constant volatility $\sigma_0 = 0.25$. The full-rank volatility matrix σ in model (4.3.1) is then given by

$$\sigma\sigma^\top = \Sigma \text{ where } \Sigma := (\Sigma_{ij})_{1\le i,j\le d} \text{ with } \Sigma_{ij} = \sigma_0^2\rho, i \neq j \text{ and } \Sigma_{ii} = \sigma_0^2.$$

Then, $A_0 := \left(\left((\sigma^1 S_0^1)^\top, \dots, (\sigma^d S_0^d)^\top\right)^\top\right)^{-1}$ where σ^i is the ith row of σ. The payoff is given by

$$\Phi(S_T^1, \dots, S_T^d) = \left(\sum_i p^i S_T^i - K\right)^+.$$

The option expiration is set to $T = 1$ year and the interest rate $r = 0.02$. We suppose that weights $p_i = \frac{1}{d}$ for all i. The strike price K equals 20 and the initial values of the assets $S_0 = (S_0^1, \dots, S_0^d)$ are specified in Table 4.6. The rest of the model parameters are the same as earlier. In this table, we present prices and deltas for different basket options with several underlyings. In the first column, classical crude Monte Carlo values are shown (MC $R = 0$, IM was not considered). In the second column SRMDP values are displayed taking into account IM.

	MC ($R=0$)		SRMDP ($R=0.02$)	
S_0	95% CI $V_0^{BS}(S_0)$	95% CI $Z_0^{BS}(S_0)A_0$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)A_0$
$(18, 20)$	$[1.5102, 1.5113]$	$[-0.0685, -0.0682]$	$[1.5015, 1.5468]$	$[-0.0772, -0.0649]$
		$[0.6237, 0.6245]$		$[0.6297, 0.6556]$
$(18, 20, 22)$	$[2.0067, 2.0081]$	$[-0.4676, -0.4671]$	$[1.9915, 2.0447]$	$[-0.4756, -0.4641]$
		$[0.3813, 0.3817]$		$[0.3873, 0.4167]$
		$[0.7435, 0.7443]$		$[0.7725, 0.7882]$
$(16, 18, 20, 22)$	$[1.4470, 1.4481]$	$[-0.6589, -0.6582]$	$[1.4677, 1.5090]$	$[-0.6689, -0.6182]$
		$[0.1062, 0.1064]$		$[0.0962, 0.1374]$
		$[0.4334, 0.4338]$		$[0.4234, 0.4628]$
		$[0.6093, 0.6100]$		$[0.5943, 0.6310]$
$(16, 18, 20, 22, 24)$	$[1.9672, 1.9676]$	$[-1.0467, -1.0455]$	$[1.9928, 2.0692]$	$[-1.0767, -1.0242]$
		$[-0.0855, -0.0852]$		$[-0.1155, -0.0752]$
		$[0.3342, 0.3347]$		$[0.3042, 0.3467]$
		$[0.5655, 0.5662]$		$[0.5355, 0.5762]$
		$[0.7039, 0.7047]$		$[0.6839, 0.7167]$

Table 4.6: Prices and deltas for the basket call option.

Chapter 5

Nested Monte Carlo

Ce chapitre est un travail réalisé avec Florian Bourgey, Stefano De Marco et Emmanuel Gobet.

Abstract

Using the MLMC method, we improve the computational cost of estimating nested expectations involving functions possibly possessing a finite number of singular points. Alternatively, we also approximate these nested expectations using upper and lower bounds, similarly to American option pricing methods. We apply these techniques on option pricing under initial margin requirements and illustrate numerically the efficiency.

Keywords: Nested Monte-Carlo, Multilevel Monte-Carlo, Initial margin.

5.1 Introduction

The paradigm of linear risk-neutral pricing of financial contracts has changed in the last few years, influenced by the regulators. Nowadays, banks and financial institutions have to post collateral to a central counterparty (CCP, also called clearing house) in order to secure their positions. Everyday, the CCP asks each member to post a certain amount according to the exposure of their Over-the-Counter (OTC) contracts becoming "the buyer of every seller and the seller of every buyer". In this setting, the CCP will ask for Initial Margin deposit which corresponds to collateral whose aim is to cover potential mark-to-market losses during the liquidation period in case of default one of the members. In [4], the authors consider a model for Initial Margin computation where the risk measure is endogeneous w.r.t. the portfolio loss and its law, over a period Δ. As the period Δ is small (typically one week), an expansion of that corrected price at the first order can be interpreted as the price of option with collateralization w.r.t. an exogenous risk measure. More precisely, they solve the linear BSDE for the couple (V_t^L, Z_t^L) :

$$V_t^L = \xi + \int_t^T \left(-r_s V_s^L + Z_s^L \sigma_s^{-1} \left(r_s \mathbf{1} - \mu_s\right) + RC_\alpha \sqrt{(s+\Delta)\wedge T - s}\,\left|Z_s^{BS}\right|\right) \mathrm{d}s - \int_t^T Z_s^L \mathrm{d}W_s,$$

where the price V^L corresponds to the price of a self-financing portfolio and Z^L the quantity invested in the risky asset (delta). Furthermore, the second term of the integral corresponds to the Initial Margin additional cost that is computed using the delta of an exogenous model (here Black-Scholes model). More precisely, (V_t^{BS}, Z_t^{BS}) is the solution of the following BSDE:

$$V_t^{BS} = \xi + \int_t^T \left(-r_s V_s^{BS} + Z_s^{BS} \sigma_s^{-1} \left(r_s \mathbf{1} - \mu_s\right)\right) \mathrm{d}s - \int_t^T Z_s^{BS} \mathrm{d}W_s.$$

The price V^L can be seen as a nested expectation. To compute it efficiently, we tackle the generic problem of estimating the expectation

$$I := \mathbb{E}\left[g\left(\mathbb{E}\left[f\left(X,Y\right)|X\right]\right)\right],$$

using the multilevel Monte-Carlo (MLMC) method, introduced by Giles in [46]. Here, X, Y are two independent random variables with values in $\mathbb{R}^d$ and the functions $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are measurable.

The efficiency of the MLMC technique is determined through the classical bias-variance decomposition: the number of levels is given by the biais between the estimator and the expected value and the complexity is given by the variance of one outer simulation at each level. In the literature, efforts have been made to compute I in the case where g is a step function. This problem is important whenever it comes to compute the probability of a large loss for a financial portfolio. Let us define for $n \geq 1$, the natural estimator $\hat{E}_{f,n}(X) = \frac{1}{n}\sum_{i=1}^n f(X, Y_i)$ where $(Y_i)_{1\leq i\leq n}$ are i.i.d. samplings of Y and independent of X. To our knowledge, in early papers dealing with this problem using the Monte-Carlo method to approximate the inner and outer expectations, such as [20] or [53], the authors assume that for $n \geq 1$, the couple $(\mathbb{E}[f(X,Y)|X], \sqrt{n}(\hat{E}_{f,n}(X) - \mathbb{E}[f(X,Y)|X]))$ has a density w.r.t . the Lebesgue measure which admits moments of order 4 as well as for its partial derivatives w.r.t. the first coordinate up to the order 2, uniformly in n. Although the authors achieve to obtain expansions of the bias at order 1, one can easily check that in the case where Y is a discrete random variable where $\mathbb{P}(Y = 1) = p = 1 - \mathbb{P}(Y = -1)$ with $p > 0.5$, X is gaussian and $f(X, Y) = XY$, we have that $\mathbb{E}[f(X,Y)|X] = (2p-1)X$, $\hat{E}_{f,1}(X) - \mathbb{E}[f(X,Y)|X] = (Y - 2p + 1)X$ and the couple $((2p-1)X, (Y - 2p + 1)X)$ does not have a density w.r.t. the Lesbesgue measure. Moreover, even under existence of the joint density, the control over the moments of the density of the joint law unifom in n does not seem easy to verify either.

For other functions g, in [23], the authors consider the problem of estimating $\mathbb{E}[g(\mathbb{E}(Z_1|Z_2))]$ in the case where g is piecewise linear and Z_1 is a Bernoulli random variable with random parameter $Z_2 \in [0,1]$, and in [46], the author considers the case where g is twice differentiable, thus allowing a more general setting for X. In both papers, the tolerance ϵ^2 on the mean square error (MSE) is achieved with complexity $\mathcal{O}(\epsilon^{-2})$ using an antithetic estimator.

In our setting, the function g would be typically the absolute value. As this function does not belong to C^2 but is still more regular than the step function, it will be sufficient for us to only study the regularity of the law of $\mathbb{E}[f(X,Y)|X]$ at the neighbourhood of the singularities of g instead of considering the joint law of $\mathbb{E}[f(X,Y)|X]$ and its estimator. The arguments used in this work are close to those in [45], where g is a step function and the authors assume that at the neighbourhood of 0, the singular point of the Heaviside function, the random variable $\frac{\mathbb{E}[f(X,Y)|X]}{\sqrt{\mathbf{Var}(f(X,Y)|X)}}$ has a bounded density. Our work also extends the results obtained in [47].

In Section 5.2, we study the generic problem of estimating I. We show, by adapting arguments from [46], that using the MLMC method, the tolerance ϵ^2 on the MSE can be achieved with complexity $\mathcal{O}(\epsilon^{-2})$, even if the test function g does not belong to C^2, provided that the probability that the conditional expectation $\mathbb{E}\left[f\left(X,Y\right)|X\right]$ is close to the singular points of g is controled polynomially. In Section 5.2.2, following the philosophy of American options we introduce an algorithm computing lower and upper bounds using non nested Monte Carlo simulations. In Section 5.3.1 we show that the problem of computing V^L fits the framework of our generic problem and apply the MLMC method on it. We then numerically compare the results obtained using the two methods.

Notations

- For a random variable A and $p > 0$ such that $\mathbb{E}\left[|A|^p\right] < \infty$, we define $||A||_p := \mathbb{E}\left[|A|^p\right]^{\frac{1}{p}}$. This is a norm for $p \geq 1$.
- For $x \in \mathbb{R}^d$, $E_f\left(x\right) := \mathbb{E}\left[f\left(x, Y\right)\right]$.
- For $a > 0, T > 0, K > 0, r > 0, \forall t \in [0, T), \forall s > 0$

$$A\left(t\right) := \frac{\log\left(\frac{K}{K-a}\right)}{\sigma\sqrt{T-t}}, \quad B\left(t\right) := \frac{\log\left(\frac{K+a}{K}\right)}{\sigma\sqrt{T-t}}. \tag{5.1.1}$$

Remark 5.1.1. Notice that $A\left(t\right) > B\left(t\right)$ since $\frac{K^2}{K^2-a^2} > 1$.

$$d_1\left(t, T, s, K\right) := \frac{\log\left(\frac{s}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\left(T-t\right)}{\sigma\sqrt{T-t}}, \quad d_2\left(t, T, s, K\right) = d_1\left(t, T, s, K\right) - \sigma\sqrt{T-t}. \tag{5.1.2}$$

$$\delta_t : x \in \mathbb{R} \to \mathcal{N}\left(x - B\left(t\right)\right) + \mathcal{N}\left(x + A\left(t\right)\right) - 2\mathcal{N}\left(x\right). \tag{5.1.3}$$

$$H_t : x \in \mathbb{R} \to e^{-\frac{A(t)^2}{2}} e^{-A(t)x} + e^{-\frac{B(t)^2}{2}} e^{B(t)x} - 2. \tag{5.1.4}$$

5.2 Theoretical methodology

5.2.1 Multilevel Monte-Carlo

Let us consider two independent families of random variables $\left(X^l_m\right)_{l,m\in\mathbb{N}^*}$ and $\left(Y^{l,m}_j\right)_{j,l,m\in\mathbb{N}^*}$. The terms inside each family are i.i.d. and have the same distribution as X and Y respectively. For $L \in \mathbb{N}^*$, $\mathbf{M} = \left(M_0, ..., M_L\right)$, $\mathbf{n} = \left(n_0, ..., n_L\right) \in \left(\mathbb{N}^*\right)^{L+1}$, we consider the antithetic multilevel estimator of I,

$$\begin{aligned}\hat{I}^{ML}_{\mathbf{M},\mathbf{n}} =& \frac{1}{M_0}\sum_{m=1}^{M_0} g\left(\frac{1}{n_0}\sum_{j=1}^{n_0} f\left(X^0_m, Y^{0,m}_j\right)\right) + \sum_{l=1}^{L}\frac{1}{M_l}\sum_{m=1}^{M_l}\left\{g\left(\frac{1}{n_l}\sum_{j=1}^{n_l} f\left(X^l_m, Y^{l,m}_j\right)\right)\right.\\ &\left. - \frac{1}{2}\left(g\left(\frac{1}{n_{l-1}}\sum_{j=1}^{n_{l-1}} f\left(X^l_m, Y^{l,m}_j\right)\right) + g\left(\frac{1}{n_{l-1}}\sum_{j=n_{l-1}+1}^{n_l} f\left(X^l_m, Y^{l,m}_j\right)\right)\right)\right\}.\end{aligned}$$

In what follows, we will set $n_l = n_0 2^l$, for $0 \leq l \leq L$. The MSE of the approximation of I by $\hat{I}^{ML}_{\mathbf{M},\mathbf{n}}$ is given by the classical bias-variance decomposition,

$$\mathbf{MSE} := \mathbb{E}\left[\left(\hat{I}^{ML}_{\mathbf{M},\mathbf{n}} - I\right)^2\right] = \left(\mathbb{E}\left[\hat{I}^{ML}_{\mathbf{M},\mathbf{n}}\right] - I\right)^2 + \mathbf{Var}\left(\hat{I}^{ML}_{\mathbf{M},\mathbf{n}}\right).$$

Assumption 1. *Let us assume that g is continuous and that there exists a (finite) subdivision of the real line $-\infty =: \mathfrak{d}_0 < \mathfrak{d}_1 < ... < \mathfrak{d}_\theta < \mathfrak{d}_{\theta+1} := \infty$, such that on each of the $\theta + 1$ intervals $\left(\mathfrak{d}_i, \mathfrak{d}_{i+1}\right)$, for $0 \leq i \leq \theta$, g belongs to C^1, g' is bounded and is uniformly Holder with exponent $\eta \in (0, 1]$.*

Assumption 2. *There exists $p > 2$ such that $\mathbb{E}\left[|f(X,Y)|^p\right] < \infty$.*

Assumption 3. *There exist constants $\nu, K_\nu > 0$ such that*

$$\forall z > 0,\ \mathbb{P}\left(\min_{1\le i\le\theta}|E_f(X) - d_i| \le z\right) \le K_\nu z^\nu.$$

Under Assumption 1, 2 and 3, Propositions 5.2.1 and 5.2.2 below help us study respectively the bias and variance terms. Their proof will be given in the following subsections.

Proposition 5.2.1. *Under Assumption 1, 2 and 3, there exists a positive constant κ such that*

$$\forall n \ge 1, \qquad \left|\mathbb{E}\left[g\left(\frac{1}{n}\sum_{j=1}^{n} f(X,Y_j)\right) - g\left(E_f(X)\right)\right]\right| \le \frac{\kappa}{n^{\frac{1}{2}\left(1+\frac{(p-1)\nu}{p+\nu}\wedge\eta\right)}}.$$

Let us also define, for $1 \le l \le L$,

$$V_l := \mathbf{Var}\left(g\left(\frac{1}{n_l}\sum_{j=1}^{n_l} f(X,Y_j)\right) - \frac{1}{2}g\left(\frac{1}{n_{l-1}}\sum_{j=1}^{n_{l-1}} f(X,Y_j)\right) - \frac{1}{2}g\left(\frac{1}{n_{l-1}}\sum_{j=n_{l-1}+1}^{n_l} f(X,Y_j)\right)\right).$$

Proposition 5.2.2. *Under Assumption 1, 2 and 3, there exists a positive constant $\tilde{\kappa}$, independent of $1 \le l \le L$ and $n_l \in \mathbb{N}^*$, such that*

$$V_l \le \tilde{\kappa} n_l^{-\left(1+\frac{(p-2)\nu}{2(p+\nu)}\wedge\eta\right)}.$$

Theorem 5.2.3. *Under Assumptions 1, 2 and 3, given a tolerance on the error $\epsilon > 0$, the bound* $\mathbf{MSE} = \mathcal{O}(\epsilon^2)$ *is achieved with a computational complexity $\mathcal{O}\left(\epsilon^{-2}\right)$ with the choice $M_l = M_0 2^{-\left(1+\frac{(p-2)\nu}{4(p+\nu)}\wedge\frac{\eta}{2}\right)l}$, $M_0 = \mathcal{O}(\epsilon^{-2})$, $n_l = n_0 2^l$, $n_0 = \mathcal{O}(1)$, $L = \left\lceil \frac{1}{\alpha}\log(\frac{2\kappa}{n_0^\alpha \epsilon^2})\right\rceil$, where $\alpha = \frac{1}{2}\left(1 + \frac{(p-1)\nu}{p+\nu}\wedge\eta\right)$.*

Proof. We first analyze the bias term. We have that $\mathbb{E}\left[\hat{I}^{ML}_{\mathbf{M},\mathbf{n}}\right] = \mathbb{E}\left[g\left(\frac{1}{n_L}\sum_{j=1}^{n_L} f(X,Y_j)\right)\right]$. By Proposition 5.2.1, there exists κ independent of $\mathbf{n}$ and such that

$$\left|\mathbb{E}\left[\hat{I}^{ML}_{\mathbf{M},\mathbf{n}} - I\right]\right| = \left|\mathbb{E}\left[g\left(\frac{1}{n_L}\sum_{j=1}^{n_L} f(X,Y_j)\right)\right] - I\right| \le \frac{\kappa}{n_L^\alpha} = \kappa n_0^\alpha 2^{-\alpha L}.$$

Moreover, by Proposition 5.2.2, we have that for $1 \le l \le L$,

$$V_l \le \tilde{\kappa} n_0^{-\beta} 2^{-\beta l},$$

where $\beta = \left(1 + \frac{(p-2)\nu}{2(p+\nu)}\wedge\eta\right)$. Finally, as the computational cost C_l for one sampling in the layer $0 \le l \le L$ is bounded by $2^{\gamma l}$, where $\gamma = 1$, we apply [46, Theorem 1]. It is easy to check that $\alpha \ge \frac{1}{2}\min(\beta,\gamma)$ and that for any $p > 2$, $\beta > \gamma$, hence the complexity. Concerning the optimal outer layers, by [46, Section 1.3] we choose $L = \left\lceil \frac{1}{\alpha}\log(\frac{2\kappa}{n_0^\alpha\epsilon^2})\right\rceil$ so that $\frac{\kappa}{n_L^\alpha} < \frac{\epsilon^2}{2}$. Moreover, it is sufficient to choose $M_l = M_0 2^{\frac{-(\beta+\gamma)l}{2}}$ with $\frac{\beta+\gamma}{2} = 1 + \frac{(p-2)\nu}{4(p+\nu)}\wedge\frac{\eta}{2}$ and $M_0 = \mathcal{O}(\epsilon^{-2})$ to conclude the proof. □

5.2.2 Non nested upper and lower bounds

Following the philosophy of American option pricing, another way of estimating nested expectations is to approach its value through lower and upper bounds which are computed using non nested Monte Carlo algorithms.

Lemma 5.2.4. *Let $K \in \mathbb{R}$ and $\mathcal{R}$ a real integrable random variable. For any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$ and random variable $\mathcal{O}$, we have that*

$$\mathbb{E}[(\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)_+] \ge \mathbb{E}[(\mathcal{R} - K)1_{\varphi(\mathcal{O})\ge K}] =: C_\varphi(K).$$
$$\mathbb{E}[(K - \mathbb{E}[\mathcal{R}|\mathcal{O}])_+] \ge \mathbb{E}[(K - \mathcal{R})1_{\varphi(\mathcal{O})\le K}] =: P_\varphi(K).$$

Proof. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a measurable function and $\mathcal{O}$ a random variable. As we have that

$$\mathbb{E}[(\mathcal{R} - K)1_{\varphi(\mathcal{O})\geq K}] = \mathbb{E}[(\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)1_{\varphi(\mathcal{O})\geq K}]$$

and $(\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)_+ = (\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)1_{\mathbb{E}[\mathcal{R}|\mathcal{O}]\geq K}$, we obtain that

$$\mathbb{E}[(\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)_+] - \mathbb{E}[(\mathcal{R} - K)1_{\varphi(\mathcal{O})\geq K}] = \mathbb{E}[(\mathbb{E}[\mathcal{R}|\mathcal{O}] - K)(1_{\mathbb{E}[\mathcal{R}|\mathcal{O}]\geq K} - 1_{\varphi(\mathcal{O})\geq K})].$$

and we obtain the first inequality as the term inside the expectation in the r.h.s. is nonnegative. The second inequality is obtained with similar arguments. □

Given a convex function $g : \mathbb{R} \to \mathbb{R}$, g is differentiable except on a countable set of points. On a differentiability point $z \in \mathbb{R}$ of g, we have that

$$\forall x \in \mathbb{R},\ g(x) = g(z) + g'(z)(x - z) + \int_z^\infty (x - u)^+\mu(du) + \int_{-\infty}^z (u - x)^+\mu(du), \tag{5.2.1}$$

where $\mu(dx)$ corresponds to the second derivative of g in the sense of distributions. Let us define for any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$J_\varphi(z) = g(z) + g'(z)(\mathbb{E}[\mathcal{R}] - z) + \int_z^\infty C_\varphi(u)\mu(du) + \int_{-\infty}^z P_\varphi(u)\mu(du). \tag{5.2.2}$$

Proposition 5.2.5. *Let $\mathcal{R}$ be a real integrable random variable and $g : \mathbb{R} \to \mathbb{R}$ a convex function such that $g(\mathcal{R})$ is integrable. The following equality holds*

$$\sup_\varphi J_\varphi = \mathbb{E}[g(\mathbb{E}[\mathcal{R}|\mathcal{O}])] = \inf_\epsilon \mathbb{E}[g(\mathcal{R} - \epsilon)], \tag{5.2.3}$$

where $\epsilon \in \{\Theta$ integrable and such that $\mathbb{E}[\Theta|\mathcal{O}] = 0\}$. The equality is reached for $\varphi(\mathcal{O}) = \mathbb{E}[\mathcal{R}|\mathcal{O}]$ and $\epsilon = \mathcal{R} - \mathbb{E}[\mathcal{R}|\mathcal{O}]$.

Proof. To obtain the equality on the l.h.s. of (5.2.3), it is sufficient to replace x by $\mathbb{E}[\mathcal{R}|\mathcal{O}]$ in (5.2.1) and conclude by Lemma 5.2.4 and Fubini's theorem. We now prove the equality in the r.h.s. of (5.2.3). For ϵ integrable and such that $\mathbb{E}[\epsilon|\mathcal{O}] = 0$, we have by the tower property of the expectation and by Jensen's inequality that $\mathbb{E}[g(R-\epsilon)] \geq \mathbb{E}[g([\mathbb{E}[\mathcal{R}-\epsilon|\mathcal{O}])] = \mathbb{E}[g(\mathbb{E}[\mathcal{R}|\mathcal{O}])]$. This concludes the proof as it is easy to check that the equality is reached for $\epsilon = \mathcal{R} - \mathbb{E}[\mathcal{R}|\mathcal{O}]$. □

Estimators

Using Proposition 5.2.5, we can derive estimators giving upper and lower bounds to the computation of $\mathbb{E}[g(\mathbb{E}[f(X,Y)|X])]$, which corresponds to the previous paragraph with the particular choice $\mathcal{R} = f(X,Y)$ and $\mathcal{O} = X$. Let $(p_i)_{i\geq 1}$ a basis of $L^2(X)$, let $(q_i)_{i\geq 1}$ a basis of $L^2(Y)$. For $k \geq 1$ let $L_{X,k} = \text{span}((p_i)_{1\leq i\leq k})$ and $L_{Y,k} = \text{span}((q_i)_{1\leq i\leq k})$. Given $k, N \in \mathbb{N}^*$, we solve the following minimization problem

$$(l_1^*, ..., l_k^*) = \underset{l_1,...,l_k\in\mathbb{R}}{\text{argmin}} \frac{1}{N}\sum_{i=1}^N \left(f(\tilde{X}_i, \tilde{Y}_i) - \sum_{j=1}^k l_j p_j(\tilde{X}_i)\right)^2. \tag{5.2.4}$$

$$(u_{ab}^*)_{1\leq a,b\leq k} = \underset{u_{ab}\in\mathbb{R}, 1\leq a,b\leq k}{\text{argmin}} \frac{1}{N}\sum_{i=1}^N \left(f(\tilde{X}_i, \tilde{Y}_i) - \sum_{a,b=1}^k u_{ab}p_a(\tilde{X}_i)(q_b(\tilde{Y}_i) - \mathbb{E}[q_b(\tilde{Y}_i)])\right)^2. \tag{5.2.5}$$

where for $1 \leq i \leq N$, each $(\tilde{X}_i, \tilde{Y}_i)$ is a independent copy of (X, Y). We then define:

$$\varphi_k^*(X) = \sum_{j=1}^k l_j^* p_j(X). \tag{5.2.6}$$

$$\epsilon_k^* = \sum_{a,b=1}^k u_{ab}^* p_a(X)(q_b(Y) - \mathbb{E}[q_b(Y)]). \tag{5.2.7}$$

Let us notice that the coefficients $(l_1^*, ..., l_k^*)$ and $(u_{ab}^*)_{1\le a,b\le k}$ are random, as functions of the sampling $(\tilde{X}_i, \tilde{Y}_i)_{1\le i\le k}$. As the coefficients l^* and u^* are independent of (X,Y), and X and Y are independent, it is easy to check that $\mathbb{E}[\epsilon_k^*|X] = 0$ and that we moreover have the following result, using the same arguments as in the proof of Proposition 5.2.5.

Lemma 5.2.6. *The following inequality holds*

$$\mathbb{E}[J_{\varphi_k^*}] \le \mathbb{E}[g(\mathbb{E}[f(X,Y)|X])] \le \mathbb{E}[g(f(X,Y) - \epsilon_k^*)].$$

5.2.3 Proof of Proposition 5.2.1

Let us define for $x \in \mathbb{R}^d$ and $n \ge 1$, $\delta E_n(x) = \hat{E}_{f,n}(x) - E_f(x)$. We first control the moments of $\delta E_n(X)$.

Lemma 5.2.7. *Let us assume that there exists $p \ge 2$, such that $\mathbb{E}[|f(X,Y)|^p] < \infty$. Then we have that*

$$\forall n \ge 1,\ \forall q \in (0,p],\ ||\delta E_n(X)||_q \le \frac{C_p}{\sqrt{n}},$$

where $C_p > 0$ is a constant, independent of n.

Proof. By Jensen's inequality, we have that for $q \in (0,p]$, $||\delta E_n(X)||_q \le ||\delta E_n(X)||_p$. It is then sufficient to deal with the case $q = p$. Let us define for $j \ge 1$, $Z_j = f(X,Y_j) - E_f(X)$ and for $n \ge 1$, $S_n = \sum_{i=1}^n Z_i$. Conditionally to X, S is a martingale, as the variables Y_j, for $j \ge 1$, are i.i.d. and independent from X. By Burkholder's inequality [60, Theorem 2.10], there exists a constant c_p which only depends on p and such that a.s.,

$$\mathbb{E}[|S_n|^p|X] \le c_p \mathbb{E}\left[\left|\sum_{i=1}^n Z_i^2\right|^{p/2}|X\right] \le c_p n^{p/2}\mathbb{E}[|Z_1|^p|X],$$

where in the last inequality we used the fact that as $\frac{p}{2} \ge 1$, $\left|\sum_{i=1}^n Z_i^2\right|^{p/2} \le n^{\frac{p}{2}-1}\sum_{i=1}^n |Z_i|^p$. Moreover, as $\delta E_n(X) = \frac{S_n}{n}$, we obtain that for $C_p = c_p^{\frac{1}{p}}||Z_1||_p$,

$$||\delta E_n(X)||_p = \frac{\mathbb{E}[|S_n|^p]^{\frac{1}{p}}}{n} \le \frac{C_p}{\sqrt{n}}.$$

□

We will also use the following Taylor expansion with integral remainder. Let us introduce the generalized derivative of g which is equal to 0 on the set $\{d_1, ..., d_\theta\}$ and to g' elsewhere. By a slight abuse of notation, we will also denote this function by g'.

Lemma 5.2.8. *Under Assumption 1, for any $x \notin \{\mathfrak{d}_1, ..., \mathfrak{d}_\theta\}$ and any $y \in \mathbb{R}$, we have*

$$g(y) - g(x) = g'(x)(y-x) + (y-x)\int_0^1 [g'(x + \lambda(y-x)) - g'(x)]\mathrm{d}\lambda, \tag{5.2.8}$$

Moreover, for $x, y \in \mathbb{R}$ and $0 \le k \le \theta$ such that $x, y \in (\mathfrak{d}_k, \mathfrak{d}_{k+1})$, we have that

$$|g(y) - g(x) - g'(x)(y-x)| \le \frac{[g']_\eta}{1+\eta}|y-x|^{1+\eta}. \tag{5.2.9}$$

where $[g']_\eta = \max_{0\le k\le\theta} \sup_{z_1,z_2\in(\mathfrak{d}_k,\mathfrak{d}_{k+1}), z_1\ne z_2} \frac{|g'(z_1)-g'(z_2)|}{|z_1-z_2|}$.

Proof. Without loss of generality, let us assume that $x < y$ and let us denote the finite set $(e_1, ...e_k) = [x,y] \cap \{\mathfrak{d}_1, ..., \mathfrak{d}_\theta\}$, ordered in ascending order. With a slight abuse of notation, let us also denote $e_0 = x, e_{k+1} = y$. As g is continuous on $\mathbb{R}$ and moreover C^1 on each interval (e_i, e_{i+1}) for $1 \le i \le k$, we have that $g(e_{i+1}) - g(e_i) = \int_{e_i}^{e_{i+1}} g'(s)\mathrm{d}s$, so that summing over the index i, we have $g(y) - g(x) = \int_x^y g'(s)\mathrm{d}s$. We then obtain (5.2.8) with the change of variable $x + \lambda(y-x) = s$. To obtain (5.2.9), it is sufficient to use (5.2.8) and the Holder property of g' on the interval $(\mathfrak{d}_k, \mathfrak{d}_{k+1})$. □

We now prove Proposition 5.2.1. By Lemma 5.2.8, using (5.2.8), we have that

$$g\left(\hat{E}_{f,n}(X)\right) - g\left(E_f(X)\right) = \delta E_n(X)\int_0^1 \left(g'\left(E_f(X) + \lambda\delta E_n(X)\right) - g'\left(E_f(X)\right)\right)\mathrm{d}\lambda + \delta E_n(X)\, g'\left(E_f(X)\right) =: T_2 + T_1.$$

Let us notice that by Assumption 3, the variable $g'(E_f(X))$ is a.s. well defined. Using the tower property of the conditional expectation, we have that

$$\mathbb{E}[T_1] = \mathbb{E}\left[\delta E_n(X)\, g'\left(E_f(X)\right)\right] = 0,$$

as we have that $\mathbb{E}\left[\hat{E}_{f,n}(X)\middle| X\right] - E_f(X) = 0$ a.s. and g' is bounded.

We now control the term $\mathbb{E}[T_2]$. Let us define $h = \frac{1}{2}\min_{1\le i\neq j\le\theta}\left(|\mathfrak{d}_i - \mathfrak{d}_j|\right)$ and introduce a parameter $z\in(0,h)$ to be fixed later. We consider the following partition:

$$\Omega_1 = \left\{\min_{1\le j\le\theta}|E_f(X) - \mathfrak{d}_j| \le z\right\},\quad \Omega_2 = \Omega_1^c\cap\left\{|\delta E_n(X)| > z\right\},\quad \Omega_3 = \Omega_1^c\cap\left\{|\delta E_n(X)|\le z\right\}.$$

We study the quantities $\mathbb{E}[T_2 1_{\Omega_1}], \mathbb{E}[T_2 1_{\Omega_2}]$ and $\mathbb{E}[T_2 1_{\Omega_3}]$. For the first term, as g' is bounded, we have that for $q, q^*\ge 1$ such that $\frac{1}{q}+\frac{1}{q^*} = 1$, and $q^*\le p$, by Holder's inequality,

$$\left|\mathbb{E}[T_2 1_{\Omega_1}]\right| \le 2\|g'\|_\infty \mathbb{E}\left[|\delta E_n(X)|\, 1_{\Omega_1}\right] \le 2\|g'\|_\infty \mathbb{P}(\Omega_1)^{\frac{1}{q}}\|\delta E_n(X)\|_{q^*}. \tag{5.2.10}$$

By Assumption 3, we have that $\mathbb{P}(\Omega_1)\le K_\nu z^\nu$. Moreover, to deal with the q^* norm in (5.2.10) where $q^*\le p$, we have by Lemma 5.2.7 that

$$\left|\mathbb{E}[T_2 1_{\Omega_1}]\right| \le 2\|g'\|_\infty K_\nu^{\frac{1}{q}} z^{\frac{\nu}{q}}\frac{C_p}{\sqrt{n}}.$$

Noticing that $1_{\Omega_2}\le\frac{|\delta E_n(X)|^r}{z^r}$ for any $r\in[1,p-1]$,

$$\left|\mathbb{E}[T_2 1_{\Omega_2}]\right| \le 2\frac{\|g'\|_\infty}{z^r}\mathbb{E}\left[|\delta E_n(X)|^{r+1}\right] = 2\frac{\|g'\|_\infty}{z^r}\|\delta E_n(X)\|_{r+1}^{r+1} \le 2\frac{\|g'\|_\infty C_p^{r+1}}{z^r n^{\frac{r+1}{2}}},$$

where we applied Lemma 5.2.7 in the last inequality with $r+1\le p$. In the space Ω_3, both $E_f(X)$ and $\hat{E}_{f,n}(X)$ belong to the same interval among $((\mathfrak{d}_i,\mathfrak{d}_{i+1}))_{0\le i\le\theta}$. Using the Holder property of the function g', we obtain for the last term

$$\begin{aligned}\left|\mathbb{E}[T_2 1_{\Omega_3}]\right| &= \left|\mathbb{E}\left[\left(\delta E_n(X)\right)\int_0^1\left(g'\left(E_f(X)+\lambda\delta E_n(X)\right) - g'(Z)\right)1_{\Omega_3}\mathrm{d}\lambda\right]\right|\\ &\le [g']_\eta\mathbb{E}\left[\left(\delta E_f(X)\right)^{1+\eta}1_{\Omega_3}\right] \le [g']_\eta\frac{C_p^{1+\eta}}{n^{\frac{1+\eta}{2}}},\end{aligned}$$

where we used Lemma 5.2.7 in the last inequality. The biais is now bounded by a function of z, as

$$\left|\mathbb{E}\left[g\left(\hat{E}_{f,n}(X)\right) - g\left(E_f(X)\right)\right]\right| \le 2\|g'\|_\infty K_\nu^{\frac{1}{q}} z^{\frac{\nu}{q}}\frac{C_p}{\sqrt{n}} + 2\frac{\|g'\|_\infty C_p^{r+1}}{z^r n^{\frac{r+1}{2}}} + [g']_\eta\frac{C_p^{1+\eta}}{n^{\frac{1+\eta}{2}}}.$$

The minimum of the r.h.s. corresponds to the minimum of the function

$$z\in\mathbb{R}\to\Psi(z) := \Gamma_q z^{\frac{\nu}{q}} n^{-\frac{1}{2}} + \Lambda_r z^{-r} n^{-\frac{r+1}{2}},$$

where for $r, q\ge 1$, $\Gamma_q = 2\|g'\|_\infty K_\nu^{\frac{1}{q}} C_p$, $\Lambda_r = 2\|g'\|_\infty C_p^{r+1}$.

As $z\le h\wedge z_0$, we have that the minimum is achieved for n large enough and then $z = \left(\frac{\Lambda_r rq}{\nu\Gamma_q}\right)^{\frac{q}{rq+\nu}} n^{-\frac{1}{2}\frac{rq}{rq+\nu}}$. In order to maximize $\frac{r\nu}{rq+\nu}$ under the constraints $\frac{1}{q}+\frac{1}{q^*} = 1, r+1\le p$, we set $q = \frac{p}{p-1}, r = p-1$, and obtain that for $n\ge h^{-2\left(1+\frac{\nu}{rq}\right)}\left(\frac{\Lambda_r rq}{\nu\Gamma_q}\right)^{\frac{2}{r}} =: N$,

$$\left|\mathbb{E}\left[g\left(\hat{E}_{f,n}(X)\right) - g\left(E_f(X)\right)\right]\right| \le \frac{\kappa_1}{n^{\frac{1}{2}\left(1+\frac{(p-1)\nu}{p+\nu}\wedge\eta\right)}},$$

for $\kappa_1 = \Gamma_q\left(\frac{\Lambda_r rq}{\nu\Gamma_q}\right)^{\frac{\nu}{rq+\nu}} + \Lambda_r\left(\frac{\Lambda_r rq}{\nu\Gamma_q}\right)^{\frac{-rq}{rq+\nu}} + [g']_\eta C_p^{1+\eta}$. Moreover let κ_2 be a uniform bound on the bias for $1\le n\le N$. We conclude the proof by choosing $\kappa = \kappa_1 + \kappa_2 N^{\frac{1}{2}\left(1+\frac{(p-1)\nu}{p+\nu}\wedge\eta\right)}$.

5.2.4 Proof of Proposition 5.2.2

For notational simplicity, let us define the following estimators

$$\hat{Z}_1 = \tfrac{1}{n_{l-1}} \textstyle\sum_{j=1}^{n_{l-1}} f(X, Y_j), \quad \hat{Z}_2 = \tfrac{1}{n_{l-1}} \textstyle\sum_{j=n_{l-1}+1}^{n_l} f(X, Y_j), \quad \hat{Z} = \tfrac{\hat{Z}_1 + \hat{Z}_2}{2} = \tfrac{1}{n_l} \textstyle\sum_{j=1}^{n_l} f(X, Y_j),$$

so that V_l rewrites $V_l = \mathbf{Var}\left(g\left(\hat{Z}\right) - \tfrac{1}{2} g\left(\hat{Z}_1\right) - \tfrac{1}{2} g\left(\hat{Z}_2\right)\right)$. Let us also define,

$$\delta\hat{Z}_1 = \hat{Z}_1 - E_f(X), \quad \delta\hat{Z}_2 = \hat{Z}_2 - E_f(X), \quad \delta\hat{Z} = \tfrac{1}{2}(\delta\hat{Z}_1 + \delta\hat{Z}_2) = \hat{Z} - E_f(X).$$

Let us recall that $h = \frac{1}{2} \min_{1 \le i \ne j \le \theta} (|\mathfrak{d}_i - \mathfrak{d}_j|)$ and introduce again the parameter $z \in (0, h)$. We use introduce a partition which is similar to the one used in the proof of Proposition 5.2.1:

$$\begin{aligned}
\tilde{\Omega}_1 &= \left\{ \min_{1 \le j \le \theta} |E_f(X) - \mathfrak{d}_j| \le z \right\} = \Omega_1, \\
\tilde{\Omega}_2 &= \tilde{\Omega}_1^c \cap \left\{ \max\left(\left|\delta\hat{Z}_1\right|, \left|\delta\hat{Z}_2\right| \right) > z \right\}, \\
\tilde{\Omega}_3 &= \tilde{\Omega}_1^c \cap \left\{ \max\left(\left|\delta\hat{Z}_1\right|, \left|\delta\hat{Z}_2\right| \right) \le z \right\}.
\end{aligned}$$

We have that $V_l \le \mathbb{E}\left[\left(g\left(\hat{Z}\right) - \tfrac{1}{2} g\left(\hat{Z}_1\right) - \tfrac{1}{2} g\left(\hat{Z}_2\right)\right)^2\right]$. We use the Taylor expansion at the first order to control the terms $\mathbb{E}\left[\left(g\left(\hat{Z}\right) - \tfrac{1}{2} g\left(\hat{Z}_1\right) - \tfrac{1}{2} g\left(\hat{Z}_2\right)\right)^2 1_{\tilde{\Omega}_i}\right]$, for $i = 1, 2$. For $A \in \{Z, Z_1, Z_2\}$, we have that by (5.2.8),

$$g\left(\hat{A}\right) = g\left(E_f(X)\right) + \delta\hat{A} \int_0^1 g'\left(E_f(X) + \lambda\delta\hat{A}\right) \mathrm{d}\lambda,$$

We then obtain

$$\begin{aligned}
\left(g\left(\hat{Z}\right) - \frac{1}{2} g\left(\hat{Z}_1\right) - \frac{1}{2} g\left(\hat{Z}_2\right)\right)^2 &\le \frac{1}{2}\left(\delta\hat{Z}_1 \int_0^1 \left(g'\left(E_f(X) + s\delta\hat{Z}\right) - g'\left(E_f(X) + \lambda\delta\hat{Z}_1\right)\right) \mathrm{d}\lambda\right)^2 \\
&\quad + \frac{1}{2}\left(\delta\hat{Z}_2 \int_0^1 \left(g'\left(E_f(X) + s\delta\hat{Z}\right) - g'\left(E_f(X) + \lambda\delta\hat{Z}_2\right)\right) \mathrm{d}\lambda\right)^2 \\
&\le 2\|g'\|_\infty^2 \left(\left(\delta\hat{Z}_1\right)^2 + \left(\delta\hat{Z}_2\right)^2\right),
\end{aligned}$$

so that for $i = 1, 2$,

$$\mathbb{E}\left[\left(g\left(\hat{Z}\right) - \frac{1}{2} g\left(\hat{Z}_1\right) - \frac{1}{2} g\left(\hat{Z}_2\right)\right)^2 1_{\tilde{\Omega}_i}\right] \le 4\|g'\|_\infty^2 \mathbb{E}\left[\left(\delta\hat{Z}_1\right)^2 1_{\tilde{\Omega}_i}\right],$$

In the space $\tilde{\Omega}_1$, by Holder's inequality, for $q, q^* \ge 1$ such that $\frac{1}{q} + \frac{1}{q^*} = 1$, $q, 2q^* \le p$ and Lemma 5.2.7,

$$\mathbb{E}\left[\left(\delta\hat{Z}_1\right)^2 1_{\tilde{\Omega}_1}\right] \le \|\delta\hat{Z}_1\|_{2q^*}^2 \mathbb{P}\left(\tilde{\Omega}_1\right)^{\frac{1}{q}} \le \frac{C_p^2}{n_{l-1}} K_\nu^{\frac{1}{q}} z^{\frac{\nu}{q}}.$$

In the space $\tilde{\Omega}_2$, as we have that $1_{\tilde{\Omega}_2} < \frac{\max\left(|\delta\hat{Z}_1|, |\delta\hat{Z}_2|\right)^r}{z^r}$ for any $r \in (0, p-2]$, we obtain that

$$\mathbb{E}\left[\left(\delta\hat{Z}_1\right)^2 1_{\tilde{\Omega}_2}\right] \le \frac{1}{z^r} \mathbb{E}\left[\max\left(\left|\delta\hat{Z}_1\right|, \left|\delta\hat{Z}_2\right|\right)^{r+2}\right] \le \frac{2}{z^r} \frac{C_p^{r+2}}{n_{l-1}^{\frac{r+2}{2}}}.$$

To control the term $\mathbb{E}\left[\left(g\left(\hat{Z}\right) - \tfrac{1}{2} g\left(\hat{Z}_1\right) - \tfrac{1}{2} g\left(\hat{Z}_2\right)\right)^2 1_{\tilde{\Omega}_3}\right]$, we use the fact that in the space $\tilde{\Omega}_3$, the values $\hat{Z}_1$, $\hat{Z}_2$, $\hat{Z}$ and $E_f(X)$ belong to the same interval, where the inequality (5.2.9) holds. Therefore, by (5.2.9),

$$\left|g\left(\hat{Z}\right) - \frac{1}{2} g\left(\hat{Z}_1\right) - \frac{1}{2} g\left(\hat{Z}_2\right)\right| \le \frac{[g']_\eta}{1+\eta}\left(\left(\delta\hat{Z}\right)^{1+\eta} + \frac{1}{2}\left(\delta\hat{Z}_1\right)^{1+\eta} + \frac{1}{2}\left(\delta\hat{Z}_2\right)^{1+\eta}\right).$$

Finally, using the fact that $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$ for $a,b,c \in \mathbb{R}$,

$$\mathbb{E}\left[\left(g\left(\hat{Z}\right)-\frac{1}{2}g\left(\hat{Z}_1\right)-\frac{1}{2}g\left(\hat{Z}_2\right)\right)^2 1_{\tilde{\Omega}_3}\right] \leq \frac{3}{(1+\eta)^2}[g']_\eta^2 C_p^{2(1+\eta)}\left(\frac{1}{n_l^{1+\eta}}+\frac{1}{2n_{l-1}^{1+\eta}}\right)$$
$$=\frac{3(1+2^\eta)}{(1+\eta)^2}[g']_\eta^2\left(\frac{C_p^2}{n_l}\right)^{1+\eta},$$

as $n_l = 2n_{l-1}$. Gathering the previous results, we obtain that

$$V_l \leq \mathbb{E}\left[\left(g\left(\hat{Z}\right)-\frac{1}{2}g\left(\hat{Z}_1\right)-\frac{1}{2}g\left(\hat{Z}_2\right)\right)^2\right]$$
$$\leq 8||g'||_\infty^2\frac{C_p^2}{n_l}K_\nu^{\frac{1}{q}}z^{\frac{\nu}{q}}+8||g'||_\infty^2\frac{1}{z^r}\frac{2^{\frac{r+2}{2}}C_p^{r+2}}{n_l^{\frac{r+2}{2}}}+\frac{3(1+2^\eta)}{(1+\eta)^2}[g']_\eta^2\left(\frac{C_p^2}{n_l}\right)^{1+\eta},$$

and we conclude in the same way we did in the proof of Proposition 5.2.1, setting $r = p-2, q = \frac{p/2}{p/2-1}$.

5.2.5 An alternative set of hypotheses

We could relax the assumption on the regularity of g, to obtain the following result which may present interest in its own, but as we can check that the density of $E_f(X)$ in the context of the computation of V^L is not bounded, so one could not apply [9, Thm. 2.4 (i)].

Theorem 5.2.9. *Let us assume that g admits a derivative g' in the sense of distributions which has bounded variation. Let us assume that $E_f(X)$ has a bounded density χ_E w.r.t. the Lebesgue measure and that there exists $p>3$ such that $\mathbb{E}[|f(X,Y)|^p] < \infty$. The tolerance $\epsilon^2 > 0$ on the MSE can be achieved with complexity $\mathcal{O}(\epsilon^{-2})$.*

The proof is a direct adaptation of the one of Theorem 5.2.3 using the two propositions below. Having the density of $E_f(X)$ bounded permits to relax the regularity hypothesis on the function g', by applying [9, Thm. 2.4 (i)].

Proposition 5.2.10. *Under the assumptions in Theorem 5.2.9, there exists a constant $\kappa > 0$ such that*

$$\left|\mathbb{E}\left[g\left(\hat{E}_{f,n}(X)\right)-g\left(E_f(X)\right)\right]\right| \leq \frac{\kappa}{n^{1-\frac{2}{p+1}}}.$$

Proof. We use the Taylor expansion with integral remainder:

$$g\left(\hat{E}_{f,n}(X)\right)-g\left(E_f(X)\right) = \delta E_n(X)\int_0^1 g'\left(E_f(X)+\lambda\delta E_n(X)\right)d\lambda$$
$$= \delta E_n(X)\, g'\left(E_f(X)\right)+\delta E(X)\int_0^1\left(g'\left(E_f(X)+\lambda\delta E_n(X)\right)-g'\left(E_f(X)\right)\right)d\lambda$$
$$=: T_1 + T_2.$$

Using the tower property of the conditional expectation, we have that

$$\mathbb{E}[T_1] = \mathbb{E}\left[g'\left(E_f(X)\right)\delta E_n(X)\right] = 0,$$

as we have that $\mathbb{E}[\delta E_n(X)|X] = 0$ a.s.. We now apply Holder's inequality on $\mathbb{E}[T_2]$. Let $q, q^* \in [1,p]$ such that $\frac{1}{q}+\frac{1}{q^*} = 1$. We obtain that,

$$|\mathbb{E}[T_2]| \leq \int_0^1 \|\delta E_n(X)\|_{q^*}\left\|\left(g'\left(E_f(X)+\lambda\delta E_n(X)\right)-g'\left(E_f(X)\right)\right)\right\|_q d\lambda.$$

By Lemma 5.2.7, we have that $\|\delta E_n(X)\|_{q^*} \leq \frac{C_p}{\sqrt{N}}$. For the second term, we have by [9, Thm. 2.4 (i)] that for $0 \leq \lambda \leq 1$, and $r \leq p$

$$\left\|g'\left(E_f(X)+\lambda\delta E_n(X)\right)-g'\left(E_f(X)\right)\right\|_q \leq 3^{1+\frac{1}{q}}||g'||_{TV}\,\|\chi_E\|_\infty^{\frac{r}{q(r+1)}}\,\|\delta E_n(X)\|_r^{\frac{r}{q(r+1)}}.$$

We then maximize the ratio $\frac{r}{q(r+1)}$ with the choice $q = \frac{p}{p-1}$ and $r = p$ to obtain that:

$$\left|\mathbb{E}\left[g\left(\hat{E}_{f,n}(X)\right) - g\left(E_f(X)\right)\right]\right| \le \frac{\kappa}{n^{1-\frac{2}{p+1}}},$$

for $\kappa = 3^{1+\frac{1}{q}} ||g'||_{TV} \, \|\chi_E\|_\infty^{\frac{r}{q(r+1)}} \, C_{q^*}$ and this concludes the proof. □

Proposition 5.2.11. *Under the assumptions of Theorem 5.2.9, there exists a constant $\tilde{\kappa}$, independent of $1 \le l \le L$ and $n_l \in \mathbb{N}^*$, such that*

$$V_l \le \frac{\tilde{\kappa}}{n_l^{1+\frac{p-2}{p+1}}}.$$

Proof. We consider the antithetic estimator defined in 5.2.1 and keep the notation of Proposition 5.2.2. For $\hat{A} \in \{\hat{Z}_1, \hat{Z}_2, \hat{Z}\}$, we have the Taylor decomposition

$$g(\hat{A}) = g(E_f(X)) + \delta\hat{A}g'(E_f(X)) + (\delta\hat{A}) \int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{A}) - g'(E_f(X))\right) d\lambda.$$

As $\delta\hat{Z} = \frac{1}{2}\left(\delta\hat{Z}_1 + \delta\hat{Z}_2\right)$, we have that

$$\begin{aligned}\left(g(\hat{Z}) - \frac{1}{2}(g(\hat{Z}_1) + g(\hat{Z}_2))\right)^2 &\le 3\left(\delta\hat{Z}\int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{Z}) - g'(E_f(X))\right) d\lambda\right)^2 \\ &+ \frac{3}{4}\left(\delta\hat{Z}_1\int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{Z}_1) - g'(E_f(X))\right) d\lambda\right)^2 \\ &+ \frac{3}{4}\left(\delta\hat{Z}_2\int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{Z}_2) - g'(E_f(X))\right) d\lambda\right)^2. \end{aligned} \tag{5.2.11}$$

Taking the expectation of the first term in the r.h.s. and using the Holder inequality for $2q^* \le p$ and $q > 1$ so that $\frac{1}{q} + \frac{1}{q^*} = 1$, we obtain that

$$\begin{aligned}&\mathbb{E}\left[\left(\delta\hat{Z}\int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{Z}) - g'(E_f(X))\right) d\lambda\right)^2\right] \\ &\le \mathbb{E}\left[\left(\int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{Z}) - g'(E_f(X))\right) d\lambda\right)^{2q}\right]^{\frac{1}{q}} \mathbb{E}\left[(\delta\hat{Z})^{2q^*}\right]^{\frac{1}{q^*}}.\end{aligned}$$

By Lemma 5.2.7, the second term in the r.h.s. is bounded by $\frac{C_p}{n_l}$. Moreover, by Jensen's inequality and [9, Thm. 2.4 (i)],

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^1 \left(g'(E_f(X) + \lambda\delta\hat{Z}) - g'(E_f(X))\right) d\lambda\right)^{2q}\right] &\le \int_0^1 \mathbb{E}\left[\left(g'(E_f(X) + \lambda\delta\hat{Z}) - g'(E_f(X))\right)^{2q}\right] d\lambda \\ &\le 3^{2q+1}||g'||_{TV}^{2q}||\chi_E||_\infty^{\frac{p}{p+1}}||\delta Z||_p^{\frac{p}{p+1}}.\end{aligned}$$

With similar computations on the two other terms in the r.h.s. of (5.2.11), we conclude the proof using Lemma 5.2.7 for the choice $q = \frac{p/2}{p/2-1}$,which optimizes the ratio $\frac{p}{q(p+1)}$. □

5.3 Application on the computation of V^L

5.3.1 Application on V_L

In the following, we work in a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and consider the filtration $\mathcal{F}$ of the market defined as the augmented filtration of W i.e $\mathcal{F}_t = \mathcal{F}_t^W := \sigma\left(W_s, 0 \le s \le t, \mathcal{N}_\mathbb{P}\right)$ where $\mathcal{N}_\mathbb{P}$ denotes the family of $\mathbb{P}-$negligible sets of $\mathcal{A}$. In this section, we consider a concrete case of IM computation where Theorem 5.2.3 applies. More precisely, let us consider a Black & Scholes model for which the dynamics of one single tradable asset $(S_t)_{t\geq 0}$ is given by :

$$S_t = S_0 e^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma W_t}, \quad S_0 > 0, \tag{5.3.1}$$

with $r > 0$ a constant interest rate and $\sigma > 0$ the volatility. In [4], the IM is computed w.r.t. the CVaR of the future evolution of the value of the portoflio over a small duration $\Delta > 0$, and an approximation of correction induced by the IM is given by the following quantity:

$$\mathbb{E}\left[\int_0^T e^{-rt}\left(RC_\alpha \left|Z_t^{BS}\right| \sqrt{(t+\Delta)\wedge T - t}\right)\mathrm{d}t\right],$$

where T is the option's maturity, Δ is the period over which the **CVaR** is computed, R the funding cost interest rate, $C_\alpha := \mathbf{CVaR}^\alpha\left(\mathcal{N}(0,1)\right) = \left.\frac{e^{-\frac{x^2}{2}}}{(1-\alpha)\sqrt{2\pi}}\right|_{x=\mathcal{N}^{-1}(x)}$ and $Z_t^{BS} := z^{BS}(t, S_t)$ where $z^{BS}(t,s)$ is defined, for a payoff function Φ, as :

$$z^{BS}(t,s) := \sigma s \frac{\partial \mathbb{E}\left[e^{-r(T-t)}\Phi(S_T)\middle| S_t = s\right]}{\partial s} = \mathbb{E}\left[e^{-r(T-t)}\left(\Phi(S_T) - \Phi(S_t)\right)\frac{W_T - W_t}{T-t}\middle| S_t = s\right]. \tag{5.3.2}$$

Assuming that every hedging operation is performed before $\tilde{T} := T - \Delta$, we will consider the modified but simpler quantity:

$$I = RC_\alpha\sqrt{\Delta}\mathbb{E}\left[\int_0^{\tilde{T}} e^{-rt}\left(\left|Z_t^{BS}\right|\right)\mathrm{d}t\right] = \mathbb{E}\left[e^{-rU}\left|Z_U^{BS}\right|\right], \tag{5.3.3}$$

where we have assumed for notational simplicity in the last equality that $\sqrt{\Delta}RC_\alpha\tilde{T} = 1$ and defined $U \sim \mathcal{U}\left[0,\tilde{T}\right] \perp \left(Z_t^{BS}\right)_{0\leq t\leq \tilde{T}}$.

Proposition 5.3.1. *Assume that we have one single tradable asset $(S_t)_{t\geq 0}$ following (5.3.1). Let $Y \sim \mathcal{N}(0,1) \perp \tilde{Z} \sim \mathcal{N}(0,1) \perp U \sim \mathcal{U}\left[0,\tilde{T}\right]$ and consider the function $g : x \in \mathbb{R} \to |x|$. Then $X := (U, S_0 e^{(r-\frac{\sigma^2}{2})U+\sigma\sqrt{U}\tilde{Z}}) \perp Y$ and*

$$I = \mathbb{E}\left[g\left(E_f(X)\right)\right] = \mathbb{E}\left[\left|\mathbb{E}\left[f(X,Y)|X\right]\right|\right],$$

where f is defined as :

$$f : ((t,s),y) \in \left(\left[0,\tilde{T}\right]\times(0,\infty)\right)\times\mathbb{R} \to e^{-rT}\frac{\left(\Phi\left(se^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}y}\right) - \Phi(s)\right)y}{\sqrt{T-t}}. \tag{5.3.4}$$

Proof. We rewrite I, using the fact that S_t and $W_T - W_t$ are independent:

$$\begin{aligned} I &= \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{E}\left[\left|e^{-rt}Z_t^{BS}\right|\right]\mathrm{d}t \\ &= \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{E}\left[\left|\mathbb{E}\left[e^{-rT}\left(\Phi\left(S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T-W_t)}\right) - \Phi(s)\right)\frac{W_T-W_t}{T-t}|S_t\right]\right|\right]\mathrm{d}t. \end{aligned}$$

Now, since $Y \perp \tilde{Z} \perp U$ and $W_T - W_t \overset{\mathcal{L}}{=} \sqrt{T-t}\times Y$, we obtain that

$$I = \mathbb{E}\left[\left|\mathbb{E}\left[e^{-rT}\frac{\left(\Phi\left(S_U e^{(r-\frac{\sigma^2}{2})(T-U)+\sigma\sqrt{T-U}Y}\right) - \Phi(S_U)\right)Y}{\sqrt{T-U}}|U, S_U\right]\right|\right] = \mathbb{E}\left[g\left(\mathbb{E}\left[f\left((U,S_U),Y\right)|U,S_U\right]\right)\right],$$

using the definitions of f, g. □

5.3.2 Call and put options

The premium at time $0 \leq t \leq T$ of a call or put option of maturity $T > 0$ and strike $K > 0$ is defined as :

$$v_{T,K}(t, S_t) := \mathbb{E}\left[e^{-r(T-t)}\Phi(S_T)\middle|\mathcal{F}_t\right],, \tag{5.3.5}$$

where $\Phi^{call} := (x-K)_+$ for a call option and $\Phi^{put}(x) := (K-x)_+$for a put option. From now on, we will use the notation $v(t,S_t) := v_{T,K}(t,S_t)$ for an option of maturity $T > 0$ and $K > 0$ and use the notation $v_{\bar{T},\bar{K}}$ when $\bar{T} \neq T$ and $\bar{K} \neq K$. What's more, we write explicitely if the option is either a call or put. Further, when considering the model in (5.3.1), the delta of a call option is positive i.e :

$$\Delta_t^{call}(S_t) := \left.\frac{\partial v^{call}(t,x)}{\partial x}\right|_{x=S_t} = \mathcal{N}(d_1(t,T,S_t,K))) > 0, \tag{5.3.6}$$

while for a put option, the delta is negative i.e :

$$\Delta_t^{put}(S_t) := \left.\frac{\partial v^{put}(t,x)}{\partial x}\right|_{x=S_t} = -\mathcal{N}(-d_1(t,T,S_t,K))) < 0. \tag{5.3.7}$$

Taking advantage of these properties, the function g, which is the absolute function here, induces no nonlinearity and we can even obtain closed formulas for the quantity I.

Proposition 5.3.2. *Denote $I^{call} := \mathbb{E}\left[\left|E_f^{call}(X)\right|\right]$ and $I^{put} := \mathbb{E}\left[\left|E_f^{put}(X)\right|\right]$ the quantity I associated to each option defined in (5.3.3) then :*

$$I^{call} = S_0\sigma\mathcal{N}(d_1(0,T,S_0,K)), \qquad I^{put} = -S_0\sigma\mathcal{N}(-d_1(0,T,S_0,K)). \tag{5.3.8}$$

Proof. We first prove the Proposition for a call option. Using the fact that $\left(e^{-rt}v^{call}(t,S_t)\right)_{0\leq t\leq\tilde{T}}$ is a martingale and taking the derivative w.r.t S_0 in $e^{-rt}v^{call}(t,S_t) = \mathbb{E}\left[e^{-rT}v^{call}(T,S_T)\middle|\mathcal{F}_t\right]$ yields:

$$e^{-rt}\Delta_t^{call}(S_t)\frac{\partial S_t}{\partial S_0} = \mathbb{E}\left[e^{-rT}\left(\Phi^{call}\right)'(S_T)\frac{\partial S_T}{\partial S_0}\middle|\mathcal{F}_t\right].$$

This proves that $e^{-rt}\Delta_t^{call}(S_t)\frac{\partial S_t}{\partial S_0}$ is a martingale closed by the terminal value $e^{-rT}\left(\Phi^{call}\right)'(S_T)\frac{\partial S_T}{\partial S_0}$ (the argument holds even in a model more general than in the Black & Scholes model). Now since S_t is linear in S_0 (see (5.3.1)), we have $\frac{\partial S_t}{\partial S_0} = \frac{S_t}{S_0}$ and therefore $\left(e^{-rt}S_t\Delta_t^{call}(S_t)\right)_{0\leq t\leq\tilde{T}} = \left(e^{-rt}Z_t^{BS}/\sigma\right)_{0\leq t\leq\tilde{T}}$ is a martingale. Going back to I^{call}, as the delta is positive, we have that:

$$I^{call} = \frac{1}{\tilde{T}}\int_0^{\tilde{T}}\mathbb{E}\left[e^{-rt}Z_t^{BS}\right]\mathrm{d}t = \sigma S_0\mathcal{N}(d_1(0,T,S_0,K)).$$

Similarly, for a put option, using that $-Z_0^{BS} = \sigma S_0\mathcal{N}(-d_1(0,T,S_0,K))$ and that the delta is negative, we prove the expression for I^{put}. □

5.3.3 The butterfly case

In the following section, we deal with a non-trivial case for which Theorem 5.2.3 applies. More precisely, we focus on a butterfly option whose delta, as a function of the spot value, changes its sign. Let us recall that the butterfly payoff, price and delta are respectively defined as:

$$\Phi(x) := (x-(K+a))_+ + (x-(K-a))_+ - 2\times(x-K)_+, \tag{5.3.9}$$

$$v_{T,K,a}(t,S_t) := v_{T,K+a}^{call}(t,S_t) + v_{T,K-a}^{call}(t,S_t) - 2\times v_{T,K}^{call}(t,S_t), \tag{5.3.10}$$

$$\Delta_t(S_t) := \left.\frac{\partial v_{T,K,a}(t,x)}{\partial x}\right|_{x=S_t} = \mathcal{N}(d_1(t,T,S_t,K+a)) + \mathcal{N}(d_1(t,T,S_t,K-a)) - 2\times\mathcal{N}(d_1(t,T,S_t,K)). \tag{5.3.11}$$

for a given strike $K > 0$ and parameter $a > 0$.

The aim of this section will be to check Assumptions 1, 2, and 3 for such an option. As the sign of the delta is not constant (see Figure 5.1), and the absolute value function has a singular point in 0, we study more precisely the values of the delta around 0 to check that Assumption 3 is verified.

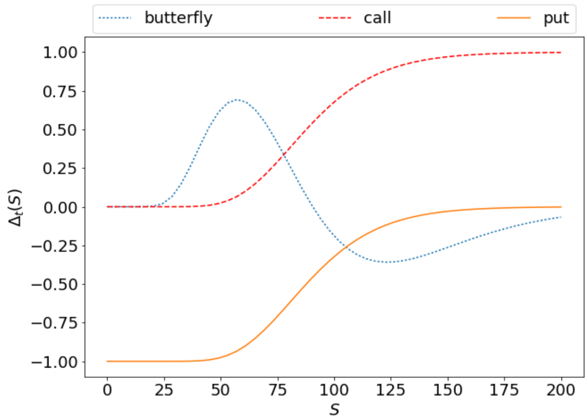


Figure 5.1: Delta for a call, put and butterfly option w.r.t spot S

Choosing the following parameters $S_0 = K = 100, a = \frac{K}{2}, T = 1, \Delta = \frac{T}{5}, r = 0.1, \sigma = 0.3$, we plot in Figure 5.1 the delta of a call, put and butterfly option and observe that while the call and put have a constant sign, the butterfly's sign changes.

Theorem 5.3.3. *Assumptions 1, 2, 3 hold true in the butterfly case. Therefore, Theorem 5.2.3 applies.*

The proof of Theorem relies on the following technical results.

Technical results

Remark 5.3.4. Notice that we can rewrite the butterfly delta as (see (5.1.3)) :

$$\Delta_t(S_t) = \delta_t(d_1(t, T, S_t, K)). \tag{5.3.12}$$

Proposition 5.3.5. *For every* $t \in \left[0, \tilde{T}\right]$,

1. *The function H_t defined in (5.1.4) has exactly two zeros denoted $\alpha(t), \beta(t)$ s.t $t \to \alpha(t)$ and $t \to \beta(t) \in \mathcal{C}^\infty([0,T))$ and $\alpha(t) < 0 < \beta(t)$.*

2. *The gamma function $\Gamma_t : x \in (0, \infty) \to \Delta_t'(x)$ has two zeroes denoted $\tilde{\alpha}(t), \tilde{\beta}(t)$ defined as :*

$$\tilde{\alpha}(t) := Ke^{\sigma\alpha(t)\sqrt{T-t} - \left(r + \frac{\sigma^2}{2}\right)(T-t)}, \quad \tilde{\beta}(t) := Ke^{\sigma\beta(t)\sqrt{T-t} - \left(r + \frac{\sigma^2}{2}\right)(T-t)}, \tag{5.3.13}$$

   s.t $t \to \tilde{\alpha}(t)$ and $t \to \tilde{\beta}(t) \in \mathcal{C}^\infty([0,T))$.

3. *The delta $x \in (0,\infty) \to \Delta_t(x)$ is increasing for $x \in (0, \tilde{\alpha}(t)) \cup \left(\tilde{\beta}(t), \infty\right)$, decreasing for $x \in \left(\tilde{\alpha}(t), \tilde{\beta}(t)\right)$ and has a unique zero denoted $\gamma(t)$. Furthermore, $t \to \gamma(t) \in \mathcal{C}^\infty([0,T))$ and we have the control:*

$$\forall t \in [0,T), \quad \tilde{\alpha}(t) < Ke^{-\left(r + \frac{\sigma^2}{2}\right)(T-t)} < \gamma(t) < \tilde{\beta}(t). \tag{5.3.14}$$

Proof. Let us justify 1. For notational simplicity, we omit the dependence of t in A and B. Taking the derivative of H_t yields:

$$H_t'(x) = -Ae^{-\frac{A^2}{2}}e^{-Ax} + Be^{-\frac{B^2}{2}}e^{Bx}.$$

Now $H_t'(x)=0 \iff x=x_0(t):=\frac{\log\left(\frac{A}{B}\right)}{A+B}-\frac{1}{2}(A-B)$, $H_t'(x)>0$ for $x>x_0(t)$ and $H_t'(x)<0$ for $x<x_0(t)$. Using that $H_t(x)\xrightarrow[x\to\pm\infty]{}+\infty$, H_t is decreasing on $(-\infty,x_0(t))$ and increasing on $(x_0(t),+\infty)$. Further, let us show that:

$$H_t(x_0(t))=e^{-A\frac{\log\left(\frac{A}{B}\right)}{A+B}-\frac{AB}{2}}+e^{B\frac{\log\left(\frac{A}{B}\right)}{A+B}-\frac{AB}{2}}-2<0.$$

Let us define $y:=\frac{A}{B}>1$ so that $H_t(x_0(t))=e^{-\frac{AB}{2}}\left(e^{-\frac{y\log(y)}{1+y}}+e^{\frac{\log(y)}{1+y}}-2\right):=e^{-\frac{AB}{2}}(h(y)-2)$. Since $e^{\frac{-AB}{2}}<1$, it is now sufficient to prove that

$$\forall y>1,\quad h(y)<2.$$

Set $g(y):=\frac{y\log(y)}{1+y}$ so that $h(y)=\frac{1+y}{e^{\frac{y\log(y)}{1+y}}}=\frac{1+y}{e^{g(y)}}$. A direct computation yields $g'(y)=\frac{\log(y)}{(1+y)^2}+\frac{1}{(1+y)}$ and $h'(y)=e^{-g(y)}(1-(1+y)g'(y))=\frac{-\log(y)e^{-g(y)}}{(1+y)}<0$ for $y>1$. Hence h is continuous and decreasing on $(1,\infty)$, thus strictly upper bounded by $h(1)=2$, which is the announced claim.

Using now the intermediate value theorem, there exist two unique roots $\alpha(t),\beta(t)\in\mathbb{R}$ to H_t. To summarize, we have :

$$\alpha(t)<x_0(t)<\beta(t),\qquad H_t(\alpha(t))=H_t(\beta(t))=0.$$

Using also that $H_t(0)=e^{-\frac{A^2}{2}}+e^{-\frac{B^2}{2}}-2<0$, we finally get:

$$\alpha(t)<0<\beta(t). \tag{5.3.15}$$

Let us now prove that $\alpha\in\mathcal{C}^\infty([0,T))$ (similar arguments can be used for β). Note that H which is $\mathcal{C}^\infty$ can be defined for negative t as well, say on $(-\eta,T)$ with $\eta>0$ and we keep writing H for this extended defintion. The roots α and β are also well defined on $(-\eta,T)$. Now let $t_1\in(-\eta,T)$ and using $\forall t\in(-\eta,T),\quad x_0(t)\neq\alpha(t)$, we have:

$$H(t_1,\alpha(t_1))=0,\quad \frac{\partial H}{\partial x}(t_1,\alpha(t_1))=-A(t_1)e^{-\frac{A(t_1)^2}{2}}e^{-A(t_1)\alpha(t_1)}+B(t_1)e^{-\frac{B(t_1)2}{2}}e^{B(t_1)\alpha(t_1)}\neq 0.$$

Using the implicit function theorem (see Thereom 10.2.2 in [34]), there exist neighborhoods $\mathcal{U}_{t_1}$ and $\mathcal{U}_{\alpha(t_1)}$ of t_1 and $\alpha(t_1)$, a function $\chi:\mathcal{U}_{t_1}\to\mathcal{U}_{\alpha(t_1)}$ of class $\mathcal{C}^\infty$ s.t.:

$$\forall t\in\mathcal{U}_{t_1},\ H(t,\chi(t))=0,\quad \chi(t)\in\mathcal{U}_{\alpha(t_1)}$$

Hence $\alpha(t_1)=\chi(t_1)$ and α is $\mathcal{C}^\infty$ on t_1. Thus α is $\mathcal{C}^\infty$ for every $t_1\in(-\eta,T)$ hence $\alpha\in\mathcal{C}^\infty((-\eta,T))$ and consequently $\alpha\in\mathcal{C}^\infty([0,T))$.

We now justify 2. Let us denote $d_1(t,x)$ instead of $d_1(t,T,x,K)$ and notice that $\forall t\in[0,T)$, $x\in(0,\infty)\to d_1(t,x)=\frac{\log\left(\frac{x}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$ is increasing and:

$$y=d_1(t,x)\iff x=Ke^{\sigma\sqrt{T-t}y-\left(r+\frac{\sigma^2}{2}\right)(T-t)}. \tag{5.3.16}$$

Moreover, taking the derivative of δ_t yields:

$$\delta_t'(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\left(e^{-\frac{A(t)^2}{2}}e^{-A(t)y}+e^{-\frac{B(t)^2}{2}}e^{B(t)y}-2\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}H_t(y).$$

Using 1 in the current Proposition, we have that the function δ_t' has two unique zeros $\alpha(t),\beta(t)$. Thus from (5.3.16) the function Γ_t has two unique zeros $\tilde{\alpha}(t),\tilde{\beta}(t)$ defined as :

$$\tilde{\alpha}(t):=Ke^{\sigma\sqrt{T-t}\alpha(t)-\left(r+\frac{\sigma^2}{2}\right)(T-t)},\quad \tilde{\beta(t)}:=Ke^{\sigma\sqrt{T-t}\beta(t)-\left(r+\frac{\sigma^2}{2}\right)(T-t)}.$$

Also still from 1, $t\to\tilde{\alpha}(t)$ and $t\to\tilde{\beta}(t)\in\mathcal{C}^\infty([0,T))$ which proves the announced claim for 2.

Finally, we justify 3. The function $y\to\delta_t(y)$ is increasing on $(-\infty,\alpha(t))\cup(\beta(t),\infty)$ and decreasing on $(\alpha(t),\beta(t))$ since H_t is positive on $(-\infty,\alpha(t))\cup(\beta(t),\infty)$, negative on $(\alpha(t),\beta(t))$ and has two unique zeros $\alpha(t),\beta(t)$. Using that $x\to d_1(t,x)$ is an increasing function and by Remark 5.3.4, we conclude that $x\to\Delta_t(x)$

is increasing on $(0, \tilde{\alpha}(t)) \cup \left(\tilde{\beta}(t), \infty\right)$ and decreasing on $\left(\tilde{\alpha}(t), \tilde{\beta}(t)\right)$. Noticing that, $\forall t \in [0,T)$, $\Delta_t(x) \xrightarrow[x \to 0^+]{} 0$ and $\Delta_t(x) \xrightarrow[x \to +\infty]{} 0$, by the intermediate value theorem, there exists a unique function $t \in [0,T) \to \gamma(t)$ s.t.:

$$\forall t \in [0,T), \quad \tilde{\alpha}(t) < \gamma(t) < \tilde{\beta}(t), \quad \Delta_t(\gamma(t)) = 0.$$

Also, using the fact that $\delta_t(0) = \Delta_t\left(Ke^{-\left(r+\frac{\sigma^2}{2}\right)(T-t)}\right) = \mathcal{N}(A(t)) - \mathcal{N}(B(t)) > 0$, $x \in \mathbb{R} \to \mathcal{N}(x)$ is a non-decreasing function and (5.1.1), we obtain:

$$\forall t \in [0,T), \ \tilde{\alpha}(t) < Ke^{-\left(r+\frac{\sigma^2}{2}\right)(T-t)} < \gamma(t).$$

Finally, let us prove that : $t \to \gamma(t) \in \mathcal{C}^\infty([0,T))$ using the implicit function theorem as for 1. More precisely, note that Δ which is $\mathcal{C}^\infty$ can be defined for negative t as well, say on $(-\eta, T)$ with $\eta > 0$ and we keep writing Δ for this extended defintion. Also, denote $\Delta(t,x)$ for $\Delta_t(x)$ and let $t_1 \in (-\eta, T)$. Noticing that:

$$\forall t \in [0,T), \ \gamma(t) \neq \tilde{\alpha}(t) \neq \tilde{\beta}(t), \quad \Delta(t_1, \gamma(t_1)) = 0, \quad \frac{\partial \Delta}{\partial x}(t_1, \gamma(t_1)) \neq 0,$$

we use the implicit function theorem and get that there exist neighborhoods $\mathcal{U}_{t_1}$ and $\mathcal{U}_{\gamma(t_1)}$ of t_1 and $\gamma(t_1)$, a function $\chi : \mathcal{U}_{t_1} \to \mathcal{U}_{\gamma(t_1)}$ of class $\mathcal{C}^\infty$ s.t :

$$\forall t \in \mathcal{U}_{t_1}, \ \Delta(t, \chi(t)) = 0, \quad \chi(t) \in \mathcal{U}_{\gamma(t_1)}.$$

Hence $\gamma(t_1) = \chi(t_1)$ and γ is $\mathcal{C}^\infty$ on t_1. Thus γ is $\mathcal{C}^\infty$ on every $t_1 \in (-\eta, T)$ hence $\gamma \in \mathcal{C}^\infty((-\eta,T))$ and consequently $\gamma \in \mathcal{C}^\infty([0,T))$. □

Remark 5.3.6. Let us consider a geometric butterfly option defined as the purchase of two calls with strike Ke^{-a} and Ke^{a} $(a > 0)$ and the sale of two puts with strike K. In other words, K is the geometric mean of Ke^{-a} and Ke^{a} while for a plain butterfly option, they are the arithmetic mean of K. Then, we can obtain an explicit formula for the delta's zero (keeping on with the notation Δ_t) :

$$\Delta_t(T,K,a) := \mathcal{N}\left(d_1(t,T,S_t,K) + \tilde{A}(t)\right) + \mathcal{N}\left(d_1(t,T,S_t,K) - \tilde{B}(t)\right) - 2\mathcal{N}(d_1(t,T,S_t,K)).$$

$$\tilde{A}(t) = -\tilde{B}(t) = \frac{a}{\sigma\sqrt{T-t}}, \qquad \gamma(t) = Ke^{-\left(r+\frac{\sigma^2}{2}\right)(T-t)}.$$

$$\tilde{\alpha}(t) = \gamma(t)\left(e^{\frac{a^2}{2\sigma^2(T-t)}} - \sqrt{e^{\frac{a^2}{\sigma^2(T-t)}} - 1}\right)^{\frac{(T-t)\sigma^2}{a}}.$$

$$\tilde{\beta}(t) = \gamma(t)\left(e^{\frac{a^2}{2\sigma^2(T-t)}} + \sqrt{e^{\frac{a^2}{\sigma^2(T-t)}} - 1}\right)^{\frac{(T-t)\sigma^2}{a}}.$$

Yet, such options do not reflect the reality of the markets as plain butterfly options have arithmetic strikes for which we do not have closed formulas.

Proposition 5.3.7. *For every $t \in [0,\tilde{T}]$, denote $\mathcal{V}(t)$ a neighborhood of $\gamma(t)$ defined as $\mathcal{V}(t) := (\gamma(t) - \epsilon(t), \gamma(t) + \epsilon(t))$ where $\epsilon(t) := \left(\frac{\gamma(t)-\tilde{\alpha}(t)}{2}\right) \wedge \left(\frac{\tilde{\beta}(t)-\gamma(t)}{2}\right)$ and χ_t the density function of the distribution $\Delta_t(S_t)$. Then, χ_t is uniformly bounded on $\mathcal{W}(t) := \Delta_t(\mathcal{V}(t))$ i.e :*

$$\forall y \in \mathcal{W}(t), \quad \chi_t(y) \leq C(t) < \infty,$$

where $C \in L^1\left(\left[0,\tilde{T}\right]\right)$.

Proof. By Proposition 5.3.5, $s \in \mathcal{V}(t) \to \Delta_t(s)$ is decreasing and takes its values in $\mathcal{W}(t)$ hence we can define Δ_t^{-1} the inverse function of Δ_t and Δ_t is a C^1- diffeomorphism from $\mathcal{V}(t)$ to $\mathcal{W}(t)$. Keeping on with the notation p_t used in Lemma (5.3.9), the density χ_t on $\mathcal{W}(t)$ is defined as:

$$\forall t \in [0,T),\ \forall y \in \mathcal{W}(t), \quad \chi_t(y) = \left.\frac{p_t(s)}{|\Delta_t'(s)|}\right|_{s=\Delta_t^{-1}(y)},$$

where:

$$|\Delta_t'(s)| = \frac{e^{-\frac{d_1(t,s)^2}{2}}}{s\sigma\sqrt{T-t}\sqrt{2\pi}} |H_t(d_1(t,s))|.$$

Using Proposition 5.3.5 and the fact that $s \to d_1(t,s)$ is increasing on $(0,\infty)$, we have for $t \in [0,\tilde{T}]$ that the function $s \to H_t(d_1(t,s))$ is increasing on $(\tilde{\alpha}(t), \tilde{x_0}(t))$ and decreasing on $\left(\tilde{x_0}(t), \tilde{\beta}(t)\right)$ where $\tilde{x}_0(t) := Ke^{\sigma\sqrt{T-t}x_0(t)-\left(r+\frac{\sigma^2}{2}\right)(T-t)}$ and $x_0(t) = \frac{\log\left(\frac{A(t)}{B(t)}\right)}{A(t)+B(t)} - \frac{1}{2}(A(t)-B(t))$ (see (5.3.5)).

Now if $\tilde{x}_0(t) \in \mathcal{V}(t)$ then:

$$\forall s \in \mathcal{V}(t),\ |H_t(d_1(t,s))| \geq |H_t(d_1(t,\gamma(t)+\epsilon(t)))| \wedge |H_t(d_1(t,\gamma(t)-\epsilon(t)))|,$$

which holds also if $\tilde{x}_0(t) \notin \mathcal{V}(t)$. Further, using that $s \in (0,\infty) \to e^{-\frac{d_1(t,s)^2}{2}}$ is increasing on $\left(0, Ke^{-(r+\frac{\sigma^2}{2})(T-t)}\right)$ and decreasing on $\left(Ke^{-(r+\frac{\sigma^2}{2})(T-t)}, \infty\right)$, we have:

$$\forall s \in \mathcal{V}(t), e^{-\frac{d_1(t,s)^2}{2}} \geq e^{-\frac{d_1(t,\gamma(t)+\epsilon(t))^2}{2}} \wedge e^{-\frac{d_1(t,\gamma(t)-\epsilon(t))^2}{2}}.$$

Consequently, for every $s \in \mathcal{V}(t)$,

$$|\Delta_t'(s)| \geq \frac{e^{-\frac{d_1(t,\gamma(t)+\epsilon(t))^2}{2}} \wedge e^{-\frac{d_1(t,\gamma(t)-\epsilon(t))^2}{2}} |H_t(d_1(t,\gamma(t)+\epsilon(t)))| \wedge |H_t(d_1(t,\gamma(t)-\epsilon(t)))|}{(\gamma(t)+\epsilon(t))\sigma\sqrt{T}\sqrt{2\pi}}.$$

Using Lemma 5.3.9, we finally obtain the following upper bound for χ_t, $\forall y \in \mathcal{W}(t)$:

$$\chi_t(y) \leq \frac{(\gamma(t)+\epsilon(t))\sqrt{T}\frac{e^{\left(\sigma^2-r\right)t}}{S_0\sqrt{t}}}{e^{-\frac{d_1(t,\gamma(t)+\epsilon(t))^2}{2}} \wedge e^{-\frac{d_1(t,\gamma(t)-\epsilon(t))^2}{2}} |H_t(d_1(t,\gamma(t)+\epsilon(t)))| \wedge |H_t(d_1(t,\gamma(t)-\epsilon(t)))|} := C(t),$$

where $C \in \mathcal{C}^0\left(\left[0,\tilde{T}\right]\right)$ since $\gamma, \epsilon, H \in \mathcal{C}^0\left(\left[0,\tilde{T}\right]\right)$ by Proposition 5.3.5. Also, notice that

$$\min_{t\in\left[0,\tilde{T}\right]} |H_t(d_1(t,\gamma(t)-\epsilon(t)))| \wedge |H_t(d_1(t,\gamma(t)+\epsilon(t)))| > 0.$$

Indeed, suppose $\min_{t\in\left[0,\tilde{T}\right]}(|H_t(d_1(t,\gamma(t)+\epsilon(t)))|) = 0$ then since H is continuous and by the extreme value theorem, there exists $t_0 \in \left[0,\tilde{T}\right]$ s.t. $|H_{t_0}(d_1(t_0,\gamma(t_0)+\epsilon(t_0)))| = 0$ or $|H_{t_0}(d_1(t_0,\gamma(t_0)-\epsilon(t_0)))| = 0$, which is impossible as $\tilde{\alpha}(t_0)$ and $\tilde{\beta}(t_0)$ are the only two zeros of the function $H_{t_0}(d_1(t_0,.))$ by Proposition 5.3.5.

Now, for $\tilde{T} < T$, we obtain:

$$\int_0^{\tilde{T}} \chi_t(y) 1_{y\in\mathcal{W}(t)} dt \leq \int_0^{\tilde{T}} C(t)\,dt < \infty,$$

using that $C \in \mathcal{C}^0\left([0,\tilde{T}]\right)$ and $C(t) \underset{0+}{\sim} \frac{\bar{C}}{\sqrt{t}} \in L^1\left(\left[0,\tilde{T}\right]\right)$ where :

$$\bar{C} := \frac{(\gamma(0)+\epsilon(0))\sqrt{T}}{S_0 e^{-\frac{d_1(0,\gamma(0)+\epsilon(0))^2}{2}} \wedge e^{-\frac{d_1(0,\gamma(0)-\epsilon(0))^2}{2}} |H_0(d_1(0,\gamma(0)+\epsilon(0)))| \wedge |H_0(d_1(0,\gamma(0)-\epsilon(0)))|} > 0.$$

$\square$

Proposition 5.3.8. *Let V be a normal r.v with mean $\mu \in \mathbb{R}$ and finite variance σ^2 then there exist $z_0 > 0, \rho > 0$ s.t :*

$$\forall 0 < z < z_0, \qquad \mathbb{P}\left(\mathcal{N}\left(V\right) < z\right) < z^{\rho}.$$

Proof. For every $z > 0$,

$$\mathbb{P}\left(\mathcal{N}\left(V\right) < z\right) = \mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(z\right) - \mu}{\sigma}\right),$$

and consequently for $0 < \sigma < \sigma' < \infty$, using (5.3.19):

$$\begin{aligned}\frac{\mathcal{N}\left(\frac{\mathcal{N}^{-1}(z)-\mu}{\sigma}\right)}{\mathcal{N}\left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right)} &\underset{z\to 0}{\sim} \frac{\sigma}{\sigma'}\frac{\exp\left(-\frac{1}{2}\left(\frac{\mathcal{N}^{-1}(z)-\mu}{\sigma}\right)^2\right)}{\exp\left(-\frac{1}{2}\left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right)^2\right)}\frac{1}{1-\frac{\mu}{\mathcal{N}^{-1}(z)}}\\ &\underset{z\to 0}{\sim} \frac{\sigma}{\sigma'}\exp\left(-\frac{1}{2}\left(\frac{\sigma'-\sigma}{\sigma\sigma'}\mathcal{N}^{-1}\left(z\right)-\frac{\mu}{\sigma}\right)\left(\frac{\sigma'+\sigma}{\sigma\sigma'}\mathcal{N}^{-1}\left(z\right)-\frac{\mu}{\sigma}\right)\right).\end{aligned}$$

Then using $\sigma' > \sigma > 0$ and $\mathcal{N}^{-1}\left(z\right) \xrightarrow[z\to 0]{} -\infty$, we have : $\frac{\mathcal{N}\left(\frac{\mathcal{N}^{-1}(z)-\mu}{\sigma}\right)}{\mathcal{N}\left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right)} \xrightarrow[z\to 0]{} 0$. Thus, there exists $\lambda_0 > 0$ s.t for every $0 < z < \lambda_0$:

$$\mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(z\right)-\mu}{\sigma}\right) < \mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(z\right)}{\sigma'}\right).$$

Now, let $\rho > 0$ and notice that for $z < 2^{-\frac{1}{\rho}}$, $\mathcal{N}\left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right) \leq z^{\rho} \iff \left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right)^2 \geq \left(\mathcal{N}^{-1}\left(z^{\rho}\right)\right)^2$ and recall (see [35]):

$$\left(\mathcal{N}^{-1}\left(z\right)\right)^2 \underset{z\to 0}{\sim} \log\left(\frac{1}{2\pi z^2}\right) - \log\left(\log\left(\frac{1}{2\pi z^2}\right)\right) \underset{z\to 0}{\sim} -2\log\left(z\right).$$

Choosing ρ s.t. $0 < \rho < \frac{1}{\left(\sigma'\right)^2}$, we have that $\frac{\left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right)^2}{\left(\mathcal{N}^{-1}(z^{\rho})\right)^2} \underset{z\to 0}{\sim} \frac{1}{\rho\left(\sigma'\right)^2} > 1$ and there exists λ_1 s.t for every $0 < z < \lambda_1$, one has $\left(\frac{\mathcal{N}^{-1}(z)}{\sigma'}\right)^2 \geq \left(\mathcal{N}^{-1}\left(z^{\rho}\right)\right)^2$. Finally, choosing $z_0 := \lambda_0 \wedge \lambda_1 \wedge 2^{-\frac{1}{\rho}}$, we prove the announced claim. □

Lemma 5.3.9. *For $t \in (0, \tilde{T}]$, denote $s \in (0,\infty) \to p_t\left(s\right)$ the density function of S_t for $t \geq 0$ defined in (5.3.1). Then*

$$\forall t \in (0,\tilde{T}], \forall s \in (0,\infty), p_t\left(s\right) := \frac{e^{-\frac{\left(\log(s)-\left(\log(S_0)+\mu t\right)\right)^2}{2\sigma^2 t}}}{s\sigma\sqrt{t}\sqrt{2\pi}} \leq p_0\left(t\right) := \frac{e^{\left(\sigma^2-r\right)t}}{S_0\sigma\sqrt{t}\sqrt{2\pi}} \in L^1\left(\left[0,\tilde{T}\right]\right). \tag{5.3.17}$$

Proof. For $t \in (0, \tilde{T}]$, let us take the derivative of p_t i.e :

$$\forall s \in (0,\infty), p_t'\left(s\right) = \frac{1}{\sigma\sqrt{t}\sqrt{2\pi}}\frac{e^{-\frac{\left(\log(s)-\left(\log(S_0)+\mu t\right)\right)^2}{2\sigma^2 t}}\left[-\frac{\left(\log(s)-\left(\log(S_0)+\mu t\right)\right)}{\sigma^2 t}-1\right]}{s^2}.$$

Now, $p_t'\left(s\right) = 0 \iff s = s_0\left(t\right) := S_0 e^{\left(\mu-\sigma^2\right)t}$, $p_t'\left(s\right) > 0$ if $s < s_0\left(t\right)$ and $p_t'\left(s\right) < 0$ if $s > s_0\left(t\right)$. Hence $s \to p_t\left(s\right)$ is increasing on $\left(0, s_0\left(t\right)\right)$ and decreasing on $\left(s_0\left(t\right), \infty\right)$. Finally :

$$p_0\left(t\right) := p\left(t, s_0\right) = \frac{e^{-\frac{\left(-\sigma^2 t\right)^2}{2\sigma^2 t}}}{S_0 e^{\left(\mu-\sigma^2\right)t}\sigma\sqrt{t}\sqrt{2\pi}} = \frac{e^{\left(\sigma^2-r\right)t}}{S_0\sigma\sqrt{t}\sqrt{2\pi}} \underset{0+}{\sim} \frac{1}{S_0\sigma\sqrt{2\pi}}\frac{1}{\sqrt{t}} \in L^1\left(\left[0,\tilde{T}\right]\right).$$

□

Proof of Theorem 5.3.3

We start with Assumption 1, we have $g(x) = |x|$ and $g'(x) = \operatorname{sgn}(x)$ hence Assumption 1 holds true with $\eta = 1$ and g has only one singular point, at 0.

We now check Assumption 2. First recall that if $\log(S) \sim \mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma > 0$ and $Y \sim \mathcal{N}(0,1)$ then (see [67]):

$$\forall n \in \mathbb{N}^*, \mathbb{E}[S^n] = e^{n\mu + \frac{n^2\sigma^2}{2}}, \quad \mathbb{E}\left[Y^{2n}\right] = \frac{(2n)!}{2^n n!}, \quad \mathbb{E}\left[Y^{2n+1}\right] = 0. \tag{5.3.18}$$

From the definition of f (see (5.3.4)), one can notice that : $\frac{e^{-rT}}{\sqrt{T-U}} \leq \frac{e^{-rT}}{\sqrt{\Delta}} < \infty$ where $U \sim \mathcal{U}\left[0, \tilde{T}\right]$. Consequently, we only focus on:

$$\mathbb{E}\left[\left(\left(\Phi\left(\tilde{S_U}\right) - \Phi(S_U)\right)Y\right)^4\right] = \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{E}\left[\left(\left(\Phi\left(\tilde{S_t}\right) - \Phi(S_t)\right)Y\right)^4\right] \mathrm{d}t,$$

where $\mu := r - \frac{\sigma^2}{2}, \tilde{S}_t = S_t e^{\mu(T-t)+\sigma\sqrt{T-t}Y}, Y \sim \mathcal{N}(0,1), U \sim \mathcal{U}\left[0, \tilde{T}\right]$. Further, for every $x > 0, \Phi(x) \leq x$ hence :

$$\mathbb{E}\left[\left(\left(\Phi\left(\tilde{S_U}\right) - \Phi(S_U)\right)Y\right)^4\right] \leq \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{E}\left[\tilde{S_t^4}Y^4\right] \mathrm{d}t + \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{E}\left[S_t^4 Y^4\right] \mathrm{d}t.$$

By independence of S_t and Y, we have first that :

$$\int_0^{\tilde{T}} \mathbb{E}\left[Y^4\right]\mathbb{E}\left[S_t^4\right] \mathrm{d}t = 3\int_0^{\tilde{T}} e^{4(\log(S_0)+\mu t)+4\sigma^2 t} \mathrm{d}t < \infty,$$

also by Cauchy-Schwarz's inequality, equations (5.3.18) and observing that $\log\left(\tilde{S}_t\right) \sim \mathcal{N}\left(\log(S_0) + \mu T, \sigma^2 T\right)$:

$$\begin{aligned}\int_0^{\tilde{T}} \mathbb{E}\left[\tilde{S_t^4}Y^4\right] \mathrm{d}t &\leq \mathbb{E}\left[Y^8\right]^{\frac{1}{2}} \int_0^{\tilde{T}} \mathbb{E}\left[\tilde{S_t^8}\right]^{\frac{1}{2}} \mathrm{d}t \\ &\leq \tilde{T} e^{4(\log(S_0)+\mu T)+16\sigma^2 T}\mathbb{E}\left[Y^8\right]^{\frac{1}{2}} < \infty,\end{aligned}$$

which proves the announced claim.

Finally let us deal with Assumption 3. From (5.3.3), we have that $|E_f(X)| = e^{-rU} S_U |\Delta_U(S_U)|$ and since $U \sim \mathcal{U}\left[0, \tilde{T}\right]$, one has that $e^{-rU} \geq e^{-r\tilde{T}}$ and it is sufficient to study

$$\mathbb{P}\left(S_U |\Delta_U(S_U)| < z\right).$$

Moreover, we have that:

$$\begin{aligned}\mathbb{P}\left(S_U |\Delta_U(S_U)| < z\right) &\leq \mathbb{P}\left(S_U < \sqrt{z}\right) + \mathbb{P}\left(|\Delta_U(S_U)| < \sqrt{z}\right) \\ &\leq \frac{1}{\tilde{T}}\int_0^{\tilde{T}}\int_0^{\sqrt{z}} p_t(s)\,\mathrm{d}s\mathrm{d}t + \mathbb{P}\left(|\Delta_U(S_U)| < \sqrt{z}\right).\end{aligned}$$

By Lemma 5.3.9, we have that

$$\begin{aligned}\frac{1}{\tilde{T}}\int_0^{\tilde{T}}\int_0^{\sqrt{z}} p_t(s)\,dsdt &\leq \frac{\sqrt{z}}{\tilde{T}}\int_0^{\tilde{T}} \frac{e^{\left(\sigma^2 - r\right)t}}{S_0\sigma\sqrt{t}\sqrt{2\pi}}\mathrm{d}t \\ &\leq \frac{2e^{\left(\sigma^2-r\right)\tilde{T}}}{\sqrt{\tilde{T}}S_0\sigma\sqrt{2\pi}}\sqrt{z}.\end{aligned}$$

Further, defining $\mathcal{S}_1(t) := (0, \tilde{\alpha}(t)), \mathcal{S}_2(t) := \left(\tilde{\alpha}(t), \tilde{\beta}(t)\right), \mathcal{S}_3(t) := \left(\tilde{\beta}(t), \infty\right)$ one has :

$$\begin{aligned}\mathbb{P}\left(|\Delta_U(S_U)| < \sqrt{z}\right) = \mathbb{P}\left(\Delta_U(S_U) < \sqrt{z}, S_U \in \mathcal{S}_1(U)\right) + \mathbb{P}\left(|\Delta_U(S_U)| < \sqrt{z}, S_U \in \mathcal{S}_2(U)\right) \\ + \mathbb{P}\left(-\Delta_U(S_U) < \sqrt{z}, S_U \in \mathcal{S}_3(U)\right) =: p_1 + p_2 + p_3.\end{aligned}$$

Let us first start with p_2. Let us notice that: $p_2 = \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{P}\left(|\Delta_t(S_t)| < \sqrt{z}, S_t \in \mathcal{S}_2(t)\right) \mathrm{d}t$. Also choosing $z < \min_{t\in\left[0,\tilde{T}\right]} \left\{-\Delta_t\left(\frac{\tilde{\beta}(t)-\gamma(t)}{2}\right) \wedge \Delta_t\left(\frac{\gamma(t)-\tilde{\alpha}(t)}{2}\right)\right\} := z_2$, we have for every $t \in \left[0, \tilde{T}\right], S_t \in \mathcal{V}(t) \subset \mathcal{S}_2(t)$ and using Proposition 5.3.7:

$$\begin{aligned} p_2 &= \frac{1}{\tilde{T}}\int_0^{\tilde{T}} \mathbb{P}\left(|\Delta_t(S_t)| < \sqrt{z}, S_t \in \mathcal{S}_2(t)\right) \mathrm{d}t \\ &\leq \frac{1}{\tilde{T}}\int_0^{\tilde{T}}\int_{-\sqrt{z}}^{\sqrt{z}} \chi_t(y)\,\mathrm{d}y\mathrm{d}t \leq \left(\frac{2}{\tilde{T}}\int_0^{\tilde{T}} C(t)\,dt\right)\sqrt{z}. \end{aligned}$$

We now focus on p_1. We recall that for $t \in \left[0, \tilde{T}\right]$, $\Delta_t(S_t) = \delta_t(d_1(t, S_t))$ and moreover by [52], we have that:

$$\mathcal{N}(x) \underset{x\to-\infty}{\sim} -\frac{\exp\left(-\frac{x^2}{2}\right)}{x\sqrt{2\pi}}. \tag{5.3.19}$$

Hence, we have:

$$\begin{aligned} \frac{\mathcal{N}(y-B(t))}{\mathcal{N}(y+A(t))} &\underset{y\to-\infty}{\sim} \frac{\exp\left(-\frac{(y-B(t))^2}{2}\right)}{\exp\left(-\frac{(y+A(t))^2}{2}\right)}\frac{y+A(t)}{y-B(t)} \\ &\underset{y\to-\infty}{\sim} \exp\left(\frac{1}{2}(A(t)+B(t))(2y-B(t)+A(t))\right) \xrightarrow[y\to-\infty]{} 0, \end{aligned}$$

and similarly:

$$\frac{\mathcal{N}(y)}{\mathcal{N}(y+A(t))} \underset{y\to-\infty}{\sim} \exp\left(\frac{A(t)}{2}(2y+A(t))\right) \xrightarrow[y\to-\infty]{} 0.$$

Finally, $\delta_t(y) \underset{y\to-\infty}{\sim} \mathcal{N}(y+A(t))$ and as A, B are uniformly bounded for $t \in [0, \tilde{T}]$, we deduce the existence of $\theta_1 < 0$ s.t.:

$$\forall t \in [0, \tilde{T}], \quad \forall y < \theta_1, \quad \delta_t(y) > \frac{1}{2}\mathcal{N}(y+A(t)) > \frac{1}{2}\mathcal{N}(y),$$

and $d_1(U, S_U) < \theta_1 \iff 0 < \delta_U(d_1(U, S_U)) = \Delta_U(S_U) < \delta_U(\theta_1)$. Now choosing:

$$\sqrt{z} < \sqrt{z_1} := \min_{t\in\left[0,\tilde{T}\right]} \delta_t(\theta_1) \leq \delta_U(\theta_1) < \delta_U(\alpha(U)),$$

we have for every $0 < z < z_1$, $p_1 \leq \mathbb{P}(\mathcal{N}(d_1(U, S_U)) < 2\sqrt{z})$. Further, noticing that $d_1(U, S_U)|\,U \sim \mathcal{N}\left(\mu_U, \sigma_U^2\right)$ where:

$$\underline{\mu} := \frac{\log\left(\frac{S_0}{K}\right) + rT + \sigma^2\left(T/2 - \tilde{T}\right)}{\sigma\sqrt{T}} < \mu_U := \frac{\log\left(\frac{S_0}{K}\right) + \left(r+\sigma^2/2\right)T - \sigma^2 U}{\sigma\sqrt{T-U}} < \bar{\mu} := \frac{\log\left(\frac{S_0}{K}\right) + \left(r+\sigma^2/2\right)T}{\sigma\sqrt{T-\tilde{T}}},$$

$$0 < \sigma_U := \sqrt{\frac{U}{T-U}} < \sigma_{\tilde{T}} := \sqrt{\frac{\tilde{T}}{T-\tilde{T}}} < \infty,$$

we obtain:

$$\begin{aligned} \mathbb{P}\left(\mathcal{N}(d_1(U, S_U)) < 2\sqrt{z}\right) &= \mathbb{P}\left(\mathcal{N}(\mu_U + \sigma_U N) < 2\sqrt{z}, N > 0\right) + \mathbb{P}\left(\mathcal{N}(\mu_U + \sigma_U N) < 2\sqrt{z}, N < 0\right) \\ &\leq \mathbb{P}\left(\mathcal{N}(\underline{\mu} - \sigma_{\tilde{T}} N) < 2\sqrt{z}, N > 0\right) + \mathbb{P}\left(\mathcal{N}(\underline{\mu} + \sigma_{\tilde{T}} N) < 2\sqrt{z}, N < 0\right) \\ &\leq 2\mathbb{P}\left(\mathcal{N}(\underline{\mu} + \sigma_{\tilde{T}} N) < 2\sqrt{z}\right), \end{aligned}$$

with $N \sim \mathcal{N}(0,1) \perp U$ and $N \overset{\mathcal{L}}{=} -N$. Hence, from Proposition 5.3.8 there exist $z_0 > 0, \rho > 0$ s.t.:

$$\forall 0 < z < z_0 \wedge z_2, \quad p_1 \leq 2^{\rho+1} z^{\frac{\rho}{2}}.$$

Finally, let us consider p_3. Let us notice that for every $x \in \mathbb{R}$, $\mathcal{N}(x) + \mathcal{N}(-x) = 1$, getting that:

$$-\delta_t(y) = \mathcal{N}(-y - A(t)) + \mathcal{N}(-y + B(t)) - 2\mathcal{N}(-y).$$

Moreovre, we observe that :

$$\frac{\mathcal{N}(-y - A(t))}{\mathcal{N}(-y + B(t))} \underset{y\to+\infty}{=\sim} \exp\left(\frac{1}{2}(A(t) + B(t))(-2y - A(t) + B(t))\right) \xrightarrow[y\to+\infty]{} 0,$$

$$\frac{\mathcal{N}(-y)}{\mathcal{N}(-y + B(t))} \underset{y\to+\infty}{=\sim} \exp\left(\frac{1}{2}B(t)(-2y + B(t))\right) \xrightarrow[y\to+\infty]{} 0.$$

Then as for p_1, there exists $-\theta_3 < 0$ s.t $\forall y > \theta_3$, $-\delta_t(y) > \frac{1}{2}\mathcal{N}(-y + B(t)) > \frac{1}{2}\mathcal{N}(-y) > 0$ and we thus choose:

$$\sqrt{z} < \sqrt{z_3} := \min_{t\in[0,\tilde{T}]} -\delta_t(\theta_3) \leq -\delta_U(\theta_3) < -\delta_U(\beta(U)).$$

Then, for every $z < z_3$ we obtain that has:

$$\begin{aligned} p_3 &= \mathbb{P}\left(-\delta_U(d_1(U, S_U)) < \sqrt{z}, S_U \in \mathcal{S}_3(U)\right) \leq \mathbb{P}\left(\mathcal{N}(-d_1(U, S_U)) < 2\sqrt{z}\right) \\ &= \mathbb{P}\left(\mathcal{N}(-\mu_U - \sigma_U N) < 2\sqrt{z}\right) = \mathbb{P}\left(\mathcal{N}(-\mu_U + \sigma_U N) < 2\sqrt{z}\right), \end{aligned}$$

and we conclude like p_1 using now $\mu_U < \bar{\mu}$. Finally, choosing $z < \bar{z} := \frac{z_0\wedge z_1\wedge z_2\wedge z_3}{e^{-r\tilde{T}}}$, we obtain that :

$$\begin{aligned} \mathbb{P}(|E_f(X)| < z) &\leq \left(\frac{2}{\tilde{T}}\int_0^{\tilde{T}} C(t)\,\mathrm{d}t + \frac{2e^{\left(\sigma^2 - r\right)\tilde{T}}\sqrt{\tilde{T}}}{\tilde{T}S_0\sigma\sqrt{2\pi}}\right) z^{\frac{1}{2}} + 2^{\rho+2}z^{\frac{\rho}{2}} \\ &\leq K_\nu z^\nu, \end{aligned}$$

with $K_\nu := 2\left(\frac{\int_0^{\tilde{T}} C(t)\mathrm{d}t}{\tilde{T}} + \frac{e^{\left(\sigma^2-r\right)\tilde{T}}\sqrt{\tilde{T}}}{\tilde{T}S_0\sigma\sqrt{2\pi}} + 2^{\rho+1}\right) > 0$ and $\nu := \frac{1}{2}(1 \wedge \rho) > 0$.

5.3.4 Algorithms and numerical results

In the following section, we test numerically the antithetic multilevel Monte-Carlo (MLMC) method exposed in the previous section and we compare the obtained results with a naive nested Monte-Carlo (NMC) method and a standard multilevel Monte-Carlo (MLMC2). More precisely, we focus on the plain butterfly option with payoff defined in (5.3.9) and compare all three methods. In our numerical tests, we will make use of both closed formulas for I^{call} and I^{put} defined in (5.3.8) as benchmark values for our methods. Unfortunately, we do not possess any analogous formula for the quantity $I^{butterfly}$ associated to the plain butterfly option. However, we can partly overcome this issue using a brute force Monte Carlo method on the butterfly's delta since we have an explicit expression for $E_f(X)$.

Estimators

We define the NMC estimator $\hat{I}_{M,N}$ by:

$$\hat{I}_{M,N} = \frac{1}{M}\sum_{m=1}^{M} g\left(\frac{1}{N}\sum_{j=1}^{N} f\left(X_m, Y_j^m\right)\right), \tag{5.3.20}$$

where $M, N \in \mathbb{N}^*$. One can also consider both the antithetic MLMC estimator $\hat{I}^{ML}_{\mathbf{M},\mathbf{n}}$ defined in Section 5.2.1 and the non-antithetic MLMC2 estimator $\hat{I}^{ML2}_{\mathbf{M},\mathbf{n}}$ defined by:

$$\begin{aligned} \hat{I}^{ML2}_{\mathbf{M},\mathbf{n}} =& \frac{1}{M_0}\sum_{m=1}^{M_0} g\left(\frac{1}{n_0}\sum_{j=1}^{n_0} f\left(X_m^0, Y_j^{0,m}\right)\right) + \sum_{l=1}^{L}\frac{1}{M_l}\sum_{m=1}^{M_l}\left\{g\left(\frac{1}{n_l}\sum_{j=1}^{n_l} f\left(X_m^l, Y_j^{l,m}\right)\right)\right. \\ &\left. - g\left(\frac{1}{n_{l-1}}\sum_{j=1}^{n_{l-1}} f\left(X_m^l, Y_j^{l,m}\right)\right)\right\}. \end{aligned}$$

In order to estimate the **RMSE** $:= \sqrt{\mathbf{MSE}}$, we use once again a MC simulation with typically $n_{rmse} = 50$ simulations and consider the estimator :

$$R\hat{M}SE := \sqrt{\frac{1}{n_{rmse}} \sum_{j=1}^{n_{rmse}} \left(\hat{I}_j - I\right)^2}, \tag{5.3.21}$$

where $(\hat{I}_j)_{1\leq j\leq n_{RMSE}}$ are independent copies of NMC, MLMC or MLMC2 estimators.

Remark 5.3.10. The optimal parameters obtained in Theorem 5.2.3 depend on some parameters $\nu, \eta, p, \alpha, \beta, \gamma$. Here, we have that $p = \infty, \eta = 1$. Moreover, as by the end of the proof of Theorem 5.3.3, we only have existence of $\nu > 0$, we make the arbitrary choice $\nu = \frac{1}{2}$. That gives: $\alpha = \frac{1}{2}(1 + \nu \wedge 1) = \frac{3}{4}$, $\gamma + \beta = 1 + \frac{\nu}{4} = \frac{9}{8}$ and we have:

$$M_l = M_0 2^{-\frac{9}{8}l}, M_0 = \mathcal{O}(\epsilon^{-2}), n_l = n_0 2^l, n_0 = \mathcal{O}(1), L = \left\lceil \frac{-\log\left(n_0^{\frac{3}{4}}\epsilon\right)}{\log\left(2^{\frac{3}{4}}\right)} \right\rceil. \tag{5.3.22}$$

Optimal choice of layers and complexity

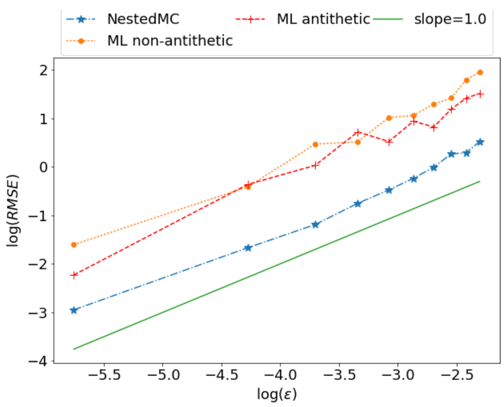
Recall also that for $\epsilon > 0$, if we want to achieve $\mathbf{MSE}^{nested} := \mathbb{E}\left[(I_{M,N} - I)^2\right] = \mathcal{O}\left(\epsilon^2\right)$, we choose: $M = \mathcal{O}\left(\epsilon^{-2}\right)$ and $N = \mathcal{O}\left(\epsilon^{-1}\right)$ and the computation cost is $\mathbf{COST}^{nested} := M \times N = \mathcal{O}\left(\epsilon^{-3}\right)$.
Finally to achieve $\mathbf{MSE}^{ML2} := \mathbb{E}\left[\left(I_{\mathbf{M},\mathbf{n}}^{ML2} - I\right)^2\right] = \mathcal{O}\left(\epsilon^2\right)$ for the MLMC2 estimator, we choose $\tilde{n}_l = \tilde{n}_0 2^l, \tilde{M}_l = \tilde{M}_0 2^{-l}, \tilde{M}_0 = \mathcal{O}\left(-\log(\epsilon)\epsilon^{-2}\right), \tilde{L} = \left\lceil \frac{-\log(\tilde{n}_0\epsilon)}{\log(2)} \right\rceil$ and the computational cost is $\mathbf{COST}^{ML2} = \sum_{l=0}^{\tilde{L}} \tilde{M}_l \tilde{n}_l = \tilde{M}_0 \tilde{n}_0 \left(\tilde{L} + 1\right) = \mathcal{O}\left((\log(\epsilon))^2 \epsilon^{-2}\right)$ (see [46, Theorem 1 and Section3] with $\gamma = \beta$).

Parameters :

We fix the following parameters $S_0 = K = 100, a = \frac{K}{2}, T = 1, r = 0.1, \sigma = 0.3$. We obtain using (5.3.8) : $I_{call} = 20.567, I_{put} = 9.432$ and by a brute force Monte Carlo $I_{butterfly} = 11.964$.

Graphs of $\log(\mathbf{RMSE})$ w.r.t $\log(\epsilon)$ and $\log(\mathbf{COST})$

Let us first plot the graph of $\log(\mathbf{RMSE})$ w.r.t $\log(\epsilon)$ in order to retrieve the complexity of $\mathcal{O}(\epsilon)$ for the **RMSE**. More precisely, for a fixed threshold $\epsilon > 0$, we set the parameters $M, N, \mathbf{M}, \mathbf{n}, L, \tilde{\mathbf{M}}, \tilde{\mathbf{n}}, \tilde{L}$ as detailed above. The log-log plot of **RMSE** w.r.t ϵ should therefore show a slope of approximately 1. We choose a range of 10 values of ϵ from $10^{-2.5}$ and 10^{-1} and compute the **RMSE** associated to each method.



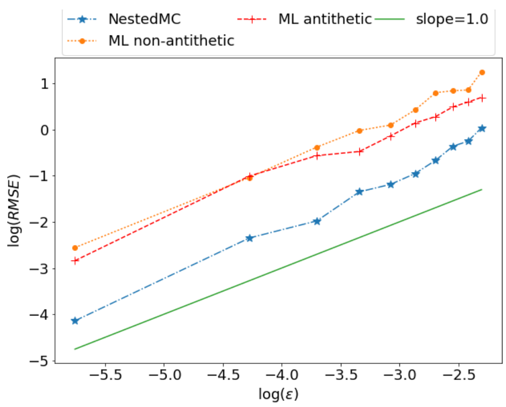


Figure 5.2: $\log(\mathbf{RMSE})$ w.r.t $\log(\epsilon)$ for a call (left) and put option (right).

As expected, we observe in Figures (5.2) and (5.3) a slope of approximately one for each method. In all figures, the NMC method seems to be better than both MLMC methods as its **RMSE** is always lower. However,

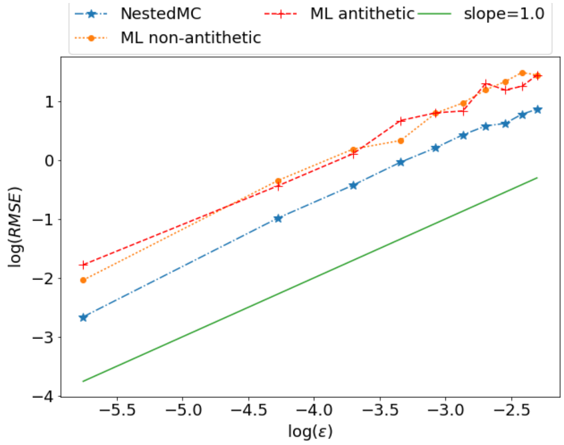


Figure 5.3: $\log(\mathbf{RMSE})$ w.r.t $\log(\epsilon)$ for the butterfly option

one needs to keep in mind that as we have fixed a $\epsilon > 0$ and chosen the associated parameters, we did not consider the computational cost which is always higher in the NMC method compares to MLMC methods.

We now show the plot of $\log(\mathbf{RMSE})$ w.r.t $\log(\mathbf{COST})$ in order to compare each method and see which one works best. More precisely, for a fixed **COST** we choose the optimal parameters using a grid-search algorithm on the choice of (M, N) for NMC, (M_0, n_0, L) for MLMC and $\left(\tilde{M}_0, \tilde{n_0}, \tilde{L}\right)$ for MLMC2. We take a range of 5 values for **COST** from 5×10^5 to 5×10^7, 5 value for M from 10^3 to 10^5, 5 values for $M_0, \tilde{M_0}$ from 10^3 to 10^5 and $L, \tilde{L} \in \{3, 4, 5, 6\}$. We compute the corresponding parameters $N, n_0, \tilde{n_0}$ using that:

$$N = \frac{\mathbf{COST}}{M},\ n_0 = \frac{(1 - 2^{-1/8})\mathbf{COST}}{(1 - 2^{-(L+1)/8})M_0},\ \tilde{n_0} = \frac{\mathbf{COST}}{(L+1)\tilde{M_0}} \tag{5.3.23}$$

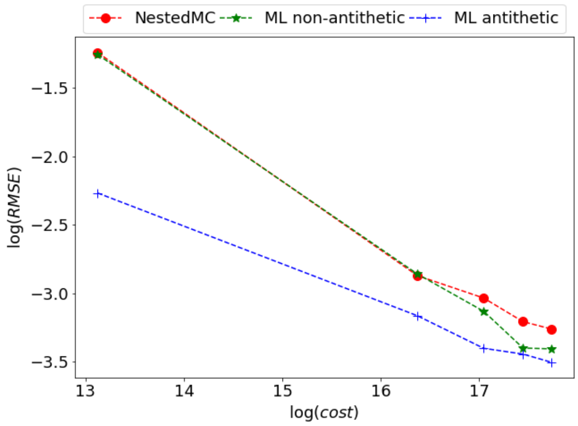


Figure 5.4: $\log(\mathbf{RMSE})$ w.r.t $\log(\mathbf{COST})$ for the butterfly option

As expected, we see that the MLMC (antithetic) gives the best results as it returns the smallest **RMSE** for a fixed **COST** while the NMC gives the highest **RMSE**. We have chosen here to only plot the butterfly case since for both the call and put options, the optimal parameters returned (e.g 1 inner simulation for the NMC after a Gridsearch algorithm) are those of a Monte Carlo simulation as in that case, the absolute function induces no nonlinearity.

COST	NMC	(M,N)	MLMC	(M_0,n_0,L)	MLMC2	$\left(\tilde{M_0},\tilde{n_0},\tilde{L}\right)$
500×10^3	0.289	$(1000,500)$	0.104	$(75250,1,5)$	0.285	$(1000,125,3)$
12875×10^3	0.0566	$(25750,500)$	0.0422	$(75250,35,5)$	0.0573	$(50500,42,5)$
25250×10^3	0.0480	$(25750,980)$	0.0333	$(100000,52,5)$	0.0436	$(50500,83,5)$
37625×10^3	0.0404	$(50500,745)$	0.0319	$(100000,77,5)$	0.0333	$(75250,83,5)$
50000×10^7	0.0384	$(75250,644)$	0.0300	$(100000,118,5)$	0.0331	$(75250,132,3)$

Table 5.1: COST/RMSE/Optimal parameters

Non-nested biased lower and upper estimators

Following ideas in Section 5.2.2, we choose for the basis of $X\in\mathbb{R}^2$ the tensor product of a Chebychev and Hermite polynomial basis i.e for $k_U\geq 0, k_Z\geq 0, L_{X,k_X}:=L_{U,k_U}\otimes L_{Z,k_Z}$ where:

$$k_X:=k_U\times k_Z, L_{U,k_U}:=\text{span}\left(T_0\left(U\right),T_1\left(U\right),...,T_{k_U}\left(U\right)\right), L_{Z,k_Z}:=\text{span}\left(H_0\left(Z\right),H_1\left(Z\right),...,H_{k_Z}\left(Z\right)\right),$$

with the polynomials defined as:

$$\begin{aligned} H_0\left(X\right)=1, H_1\left(X\right)=X, &\quad \forall i\geq 1, H_{i+1}\left(X\right)=XH_i\left(X\right)-iH_{i-1}\left(X\right).\\ T_0\left(X\right)=1, T_1\left(X\right)=X, &\quad \forall i\geq 1, T_{i+1}\left(X\right)=2XT_i\left(X\right)-T_{i-1}\left(X\right).\end{aligned}$$

Then, we solve the following least-squares problem associated:

$$\left(l^*_{ab}\right)=\underset{l_{ab}\in\mathbb{R},0\leq a\leq k_U,0\leq b\leq k_Z}{\text{argmin}}\frac{1}{N}\sum_{i=1}^{N}\left(f(\tilde{X}_i,\tilde{Y}_i)-\sum_{0\leq a\leq k_U,0\leq b\leq k_Z}l_{ab}T_a(\tilde{U}_i)H_b\left(\tilde{Z}_i\right)\right)^2. \tag{5.3.24}$$

For ϵ, we keep the same basis for X i.e $L_{X,d_X}=L_{U,d_U}\otimes L_{Z,d_Z}, d_U\geq 0, d_Z\geq 0$ and choose for $Y: L_{Y,d_Y}=\text{span}\left(H_0\left(Y\right),H_1\left(Y\right),...,H_{d_Y}\left(Y\right)\right), d_Y\geq 0$. Taking advantage of the fact that $\mathbb{E}\left[H_b\left(Y\right)\right]=1_{\{b=0\}}$ (see [2, 22.2.15 p.775]), we now solve the following least-squares problem :

$$\left(u^*_{abc}\right)=\underset{u_{abc}\in\mathbb{R},0\leq a\leq k_U,0\leq b\leq k_Z,0<c\leq k_Y}{\text{argmin}}\frac{1}{N}\sum_{i=1}^{N}\left(f(\tilde{X}_i,\tilde{Y}_i)-\sum_{0\leq a\leq k_U,0\leq b\leq k_Z,0<c\leq k_Y}u_{abc}T_a(\tilde{U}_i)H_b\left(\tilde{Z}_i\right)H_c\left(\tilde{Y}_i\right)\right)^2. \tag{5.3.25}$$

To calibrate our estimators, we generate $N=10^5$ independent copies $\left(\tilde{X}_i,\tilde{Y}_i\right)_{1\leq i\leq N}$ of (X,Y) in order to solve both (5.3.24) and (5.3.25) and get $\varphi^*_k\left(X\right)$ and ϵ^*_d (see (5.2.6) and (5.2.7)). From here, we generate another set of $\tilde{N}=10^6$ independent copies of (X,Y) and proceed with the Monte Carlo simulations associated to both estimators.

Impact and optimal choice of the polynomial degrees

Once the choice of the basis is fixed, one needs to find the optimal degrees choice for (k_U,k_Z) and (d_U,d_Z,d_Y). To do so, we proceed with a grid-search algorithm i.e we choose a range of $k_{max}=8$ for the lower bound, $d_{max}=8$ degrees for the upper bound and choose the optimal degrees for which the estimate is the closest to I.

For the call option, it seems that the lower bound degree has no impact on the value estimated since the delta is positive. Concerning the upper bound degree, the optimal is for $d_U=3, d_Z=6, d_Y=6$. For the butterfly option, the optimal lower bound degrees are $k_U=4, k_Z=4$ while for the upper bound $d_U=3, d_Z=6, d_Y=7$. We see that the higher k or d, the better the approximation up to a certain degree for which we start overfitting the data.

(k_U, k_Z)	Lower	CI
$(1,1)$	20.511	[20.437, 20.584]
$(2,2)$	20.52	[20.447, 20.593]
$(3,3)$	20.548	[20.475, 20.621]
$(4,4)$	20.518	[20.445, 20.591]
$(5,5)$	20.523	[20.450, 20.596]
$(6,6)$	20.532	[20.459, 20.605]
$(7,7)$	20.511	[20.438, 20.584]
$(8,8)$	20.618	[20.544, 20.691]
$(9,9)$	20.507	[20.434, 20.58]

(d_U, d_Z, d_Y)	Upper	CI
$(1,1,1)$	22.760	[22.698 , 22.822]
$(2,2,2)$	22.420	[22.393 , 22.447]
$(3,3,3)$	21.092	[21.071, 21.113]
$(4,4,4)$	20.934	[20.913 , 20.955]
$(5,5,5)$	20.879	[20.856 , 20.901]
$(6,6,6)$	20.851	[20.821, 20.881]
$(7,7,7)$	21.154	[21.011, 21.296]
$(8,8,8)$	21.95	[21.441, 22.459]
$(3,6,6)$	20.81	[20.787, 20.834]

Table 5.2: Lower and upper estimators for a call option and different degrees

(k_U, k_Z)	Lower	CI
$(1,1)$	11.781	[11.728, 11.834]
$(2,2)$	11.155	[11.102 , 11.209]
$(3,3)$	11.846	[11.792, 11.899]
$(4,4)$	11.743	[11.69, 11.796]
$(5,5)$	11.835	[11.781, 11.888]
$(6,6)$	11.821	[11.768, 11.874]
$(7,7)$	12.006	[11.953, 12.059]
$(8,8)$	11.957	[11.904 , 12.01]
$(9,9)$	11.915	[11.904 , 12.01]

(d_U, d_Z, d_Y)	Upper	CI
$(1,1,1)$	16.053	[16.015, 16.091]
$(2,2,2)$	16.146	[16.119, 16.172]
$(3,3,3)$	14.583	[14.559, 14.607]
$(4,4,4)$	13.811	[13.789, 13.832]
$(5,5,5)$	13.178	[13.144, 13.211]
$(6,6,6)$	12.948	[12.883, 13.013]
$(7,7,7)$	12.887	[12.781, 12.993]
$(8,8,8)$	13.165	[12.95, 13.379]
$(3,6,7)$	12.702	[12.678, 12.727]

Table 5.3: Lower and upper estimators for a butterfly option and different degrees

Bibliography

[1] F. Abergel and R. Tachet. A non-linear partial integro-differential equation from mathematical finance. *Discrete Contin. Dyn. Syst.(3)*, 907-917, 2010.

[2] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1964.

[3] Y. Achdou and O. Pironneau. *Computational Methods for Option Pricing.* SIAM series, Frontiers in Applied Mathematics, Philadelphia, 2005.

[4] A. Agarwal, S. De Marco, E. Gobet, J. G. López-Salas, F. Noubiagain, and A. Zhou. Numerical approximations of McKean Anticipative Backward Stochastic Differential Equations arising in Initial Margin requirements. Preprint, January 2018.

[5] J. M. P. Albin. A continuous non-Brownian motion martingale with Brownian motion marginal distributions. *Statistics and Probability Letters*, 78(6):682, January 2010.

[6] A. Alfonsi, C. Labart, and J. Lelong. Stochastic local intensity loss models with interacting particle systems. *Mathematical Finance*, 26(2):366–394, 2016.

[7] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73(6):890–896, 11 1967.

[8] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[9] R. Avikainen. On irregular functionals of SDEs and the Euler scheme. *Finance and Stochastics*, 13(3):381–401, Sep 2009.

[10] D. Baker, C. Donati-Martin, and M. Yor. *A Sequence of Albin Type Continuous Martingales with Brownian Marginals and Scaling*, pages 441–449. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[11] V. Bally and D. Talay. The law of the euler scheme for stochastic differential equations. *Probability Theory and Related Fields*, 104(1):43–60, Mar 1996.

[12] Basel Committee on Banking Supervision. Margin requirements for non-centrally-cleared derivatives. *http://www.bis.org/bcbs/publ/d317.pdf*, 2015.

[13] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov. Positive densities of transition probabilities of diffusion processes. *Theory of Probability & Its Applications*, 53(2):194–215, 2009.

[14] M. Bossy and J.-F. Jabir. Particle approximation for lagrangian stochastic models with specular boundary condition. *Electron. Commun. Probab.*, 23:14 pp., 2018.

[15] M. Bossy, J.-F. Jabir, and D. Talay. On conditional mckean lagrangian stochastic models. *Probability Theory and Related Fields*, 151(1):319–351, Oct 2011.

[16] M. Bossy and D. Talay. Convergence rate for the approximation of the limit law of weakly interacting particles: application to the burgers equation. *Ann. Appl. Probab.*, 6(3):818–861, 08 1996.

[17] M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Mathematics of Computation*, 66:157–192, 1997.

[18] D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 51(4):621–51, 1978.

[19] H. Brezis. *Analyse fonctionnelle, Théorie et applications.* Masson, Paris, 1983.

[20] M. Broadie, Y. P. Du, and C. C. Moallemi. Risk estimation via regression. *Operations Research*, 63(5):1077–1097, 2015.

[21] M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management science*, 42(2):269–285, 1996.

[22] R. Buckdahn and P. Imkeller. Backward stochastic differential equations with time delayed generator. *prepint*, 2009.

[23] K. Bujok, B. M. Hambly, and C. Reisinger. Multilevel simulation of functionals of bernoulli random variables with application to basket credit derivatives. *Methodology and Computing in Applied Probability*, 17(3):579–604, Sep 2015.

[24] P. Cannarsa and T. D'Aprile. Lecture notes on measure theory and functional analysis, 2006.

[25] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II.* Springer.

[26] R. Carmona, J.P. Fouque, and L. H. Sun. Mean field games and systemic risk: a toy model. *Communications in Mathematical Sciences*, 13:911–933, 2014.

[27] L. Cattiaux, P.and Pédèches. The 2-D stochastic Keller-Segel particle model : existence and uniqueness. working paper or preprint, January 2016.

[28] P. Chaudru De Raynal. Strong well-posedness of mckean-vlasov stochastic differential equation with hölder drift. *ArXiv Preprint.*

[29] P. Cheredito and K. Nam. BSE's, BSDE's and fixed point problems. *The Annals of Probability*, 45(6A):3795–3828, 2017.

[30] S. Crépey, R. Élie, W. Sabbagh, and S. Song. When capital is a funding source: The XVA anticipated BSDEs. *preprint*, 2017.

[31] L. Delong and P. Imkeller. On Malliavin's differentiability of time delayed BSDEs driven by Brownian motions and Poisson random measures. *preprint*, 2009.

[32] L. Delong and P. Imkeller. Backward Stochastic Differential Equations with time delayed generators - Results and counterexamples. *The Annals of Probability*, 20(4):1512–1536, 2010.

[33] L. Delong and P. Imkeller. Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management in insurance and finance. *Applicationes Mathematicae*, 39:463–488, 2012.

[34] J. Dieudonné, D. Huet, and G. Julia. Eléments d'analyse. tome i, fondements de l'analyse moderne. 1979.

[35] D. E. Dominici. The inverse of the cumulative standard normal probability function. *Integral Transforms and Special Functions*, 14(4):281–292, 2003.

[36] B. Dupire. Pricing with a Smile. *Risk*, 7, 18-20, 1994.

[37] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence.* Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley, 1986.

[38] J. Y. Fan, K. Hamza, and F. Klebaner. Mimicking self-similar processes. *Bernoulli*, 21(3):1341–1360, 08 2015.

[39] A. Figalli. Existence and uniqueness of martingale solutions for sdes with rough or degenerate coefficients. *Journal of Functional Analysis*, 254(1):109 – 153, 2008.

[40] N. Fournier and M. Hauray. Propagation of chaos for the Landau equation with moderately soft potentials. *ArXiv e-prints*, January 2015.

[41] N. Fournier and B. Jourdain. Stochastic particle approximation of the keller–segel equation and two-dimensional generalization of bessel processes. *Ann. Appl. Probab.*, 27(5):2807–2861, 10 2017.

[42] A. Friedman. *Partial differential equations of parabolic type.* Prentice-Hall, 1964.

[43] A. Friedman. *Stochastic Differential Equations and Applications.* Academic Press, London, 1975.

[44] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[45] M. Giles and A.-L. Haji-Ali. Multilevel nested simulation for efficient risk estimation. *arXiv:1802.05016.*

[46] M. B. Giles. Multilevel monte carlo methods. *Acta Numerica*, 24:259–328, 2015.

[47] D. Giorgi, V. Lemaire, and G. Pagès. Weak error for nested Multilevel Monte Carlo. working paper or preprint, June 2018.

[48] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Variance reduction techniques for estimating Value-at-Risk. *Management Science*, 46:1349–1364, 2000.

[49] E. Gobet, J. G. López-Salas, P. Turkedjiev, and C. Vázquez. Stratified Regression Monte-Carlo Scheme for Semilinear PDEs and BSDEs with Large Scale Parallelization on GPUs. *SIAM Journal on Scientific Computing*, 38(6):C652–C677, 2016.

[50] E. Gobet and S. Pagliarani. Analytical approximations of BSDEs with nonsmooth driver. *SIAM Journal on Financial Mathematics*, 6(1):919–958, 2015.

[51] E. Gobet and P. Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Math. Comp.*, 85(299):1359–1391, 2016.

[52] R. D. Gordon. Values of mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *The Annals of Mathematical Statistics*, 12(3):364–366, 1941.

[53] M. B. Gordy and S. Juneja. Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848, 2010.

[54] H. Guennoun and P. Henry-Labordère. Local volatility models enhanced with jumps. *Available at SSRN: https://ssrn.com/abstract=2781102*, May. 2016.

[55] J. Guyon. Path-dependent volatility. *Risk*, October, 2014.

[56] J. Guyon. Cross-dependent volatility. *Risk*, April, 2016.

[57] J. Guyon and P. Henry-Labordère. Being particular about calibration. *Risk magazine*, pages 88–93, Jan. 2012.

[58] I. Gyöngy. Mimicking the one-dimensional marginal distribution of processes having an Ito differential. *Probab. Theory Relat. Fields*, 71(4), 501-516, 1986.

[59] J.K. Hale. *Ordinary Differential Equations.* Dover Books on Mathematics Series. Dover Publications, 2009.

[60] P. Hall and C. C. Heyde. Martingale limit theory and its application. Probability and Mathematical Statistics: A Series of Monographs and Textbooks, pages ix – xi. Academic Press, 1980.

[61] K. Hamza and F. C. Klebaner. A family of non-gaussian martingales with gaussian marginals. *J. Appl. Math. and Stochastic Analysis*, 2007.

[62] J. Haškovec and C. Schmeiser. Stochastic particle approximation for measure valued solutions of the 2d keller-segel system. *Journal of Statistical Physics*, 135(1):133–151, Apr 2009.

[63] P. Henry-Labordere, X. Tan, and N. Touzi. An explicit martingale version of the one-dimensional Brenier's theorem with full marginals constraint. *Stochastic Processes and their Applications*, 126(9):2800 – 2834, 2016.

[64] F. Hirsch, C. Profeta, B. Roynette, and M. Yor. *Constructing Self-Similar Martingales via Two Skorokhod Embeddings*, pages 451–503. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[65] D. G. Hobson. Fake exponential brownian motion. *Statistics & Probability Letters*, 83(10):2386 – 2390, 2013.

[66] J.-F. Jabir, D. Talay, and M. Tomasevic. Mean–field limit of a particle approximation of the one-dimensional parabolic–parabolic Keller-Segel model without smoothing. working paper or preprint, January 2018.

[67] L Johnson Norman, S Kotz, and N Balakrishnan. Lognormal distributions, continuous univariate distributions. vol. 1, 1994.

[68] B. Jourdain, T. Lelièvre, and R. Roux. Existence, uniqueness and convergence of a particle approximation for the adaptive biasing force process. *ESAIM: M2AN*, 44(5):831–865, 2010.

[69] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 34(6):727 – 766, 1998.

[70] B. Jourdain, S. Méléard, and W. A. Woyczynski. Probabilistic approximation and inviscid limits for one-dimensional fractional conservation laws. *Bernoulli*, 11(4):689–714, 08 2005.

[71] I. Karatzas and S.E. Shreve. *Methods of mathematical finance*. Springer-Verlag, New York, 1998.

[72] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7:1–71, 1997.

[73] T. G. Kurtz. *Equivalence of Stochastic Equations and Martingale Problems*, pages 113–130. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[74] J. Lasry and P.-L. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, Mar 2007.

[75] T. Lelievre, F. Otto, M. Rousset, and G. Stoltz. Long-time convergence of an Adaptive Biasing Force method. working paper or preprint, May 2007.

[76] A. Lipton. The vol smile problem. *Risk Magazine*, 15:61–65, 2002.

[77] D. B. Madan and M. Yor. Making markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536, 08 2002.

[78] C. Marchioro and M. Pulvirenti. Hydrodynamics in two dimensions and vortex theory. *Comm. Math. Phys.*, 84(4):483–503, 1982.

[79] H.P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA*, 56:1907–1911, 1966.

[80] S. Méléard. Stochastic particle approximations for two-dimensional navier-stokes equations. In Alejandro Maass, Servet Martínez, and Jaime San Martín, editors, *Dynamics and Randomness II*, pages 147–197, Dordrecht, 2004. Springer Netherlands.

[81] J. L. Menaldi. *Stochastic differential equations with jumps*. 2014.

[82] S. Menozzi and V. Lemaire. On some non asymptotic bounds for the euler scheme. *Electron. J. Probab.*, 15:1645–1681, 2010.

[83] M. Musiela and M. Rutkowski. *Martingale methods in financial modelling*. Springer Verlag, second edition, 2005.

[84] K. Oleszkiewicz. On fake Brownian motions. *Statistics and Probability Letters*, 78(11):1251, March 2010.

[85] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equations. *Systems & Control Letters*, 14(1):55–61, 1990.

[86] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equation. *Lecture Notes in CIS, Springer-Verlag*, 176:200–217, 1992.

[87] E. Pardoux and A. Rascanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, volume 69 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2014.

[88] S. Peng and Z. Yang. Anticipated backward stochastic differential equations. *The Annals of Probability*, 37(3):877–902, 2009.

[89] V. Piterbarg. Markovian projection method for volatility calibration. *Risk Magazine*, 20:84–89, 2007.

[90] Y. Ren, D. Madan, and M. Qian. Calibrated and pricing with embedded local volatility models. *Risk*, September:138–143, 2007.

[91] R.T. Rockafellar and S. Uryasev. Optimization of conditional Value-at-Risk. *Journal of risk*, 2:21–42, 2000.

[92] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979. Reprinted in 2006.

[93] A.-S Sznitman. Topics in propagation of chaos. In Paul-Louis Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*, pages 165–251, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[94] D. Talay and M. Tomasevic. A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The one-dimensional case. working paper or preprint, February 2018.

[95] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):483–509, 1990.

[96] D. Talay and O. Vaillant. A stochastic particle method with random weights for the computation of statistical solutions of mckean-vlasov equations. *Ann. Appl. Probab.*, 13(1):140–180, 01 2003.

[97] R. Temam. *Navier-Stokes Equations*. North-Holland, Amsterdam, 1979.

[98] V. C. Tran. A wavelet particle approximation for mckean–vlasov and 2d-navier–stokes statistical solutions. *Stochastic Processes and their Applications*, 118(2):284 – 318, 2008.

[99] Z. Yang and R.J. Elliott. Anticipated backward stochastic differential equations with continuous coefficients. *Communications on Stochastic Analysis*, 7(2):303–319, 2013.

[100] G. Yin and C. Zhu. Properties of solutions of stochastic differential equations with continuous-state-dependent switching. *Journal of Differential Equations*, 249(10):2409 – 2439, 2010.

[101] D.M. Young and R.T. Gregory. *A survey on Numerical Mathematics*, volume 1. Dover Publications, 1973.

[102] J. Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, 14(1):459–488, 02 2004.

[103] X. Zhang. Degenerate irregular sdes with jumps and application to integro-differential equations of fokker-planck type. *Electron. J. Probab.*, 18:25 pp., 2013.