Calibration de modèles financiers par minimisation d'entropie relative et modèles avec sauts

Laurent Nguyen

ECOLE NATIONALE DES PONTS ET CHAUSSEES

THESE

pour obtenir le grade de

DOCTEUR DE L'ECOLE NATIONALE DES PONTS ET CHAUSSEES

Spécialité : Mathématiques, Informatique

présentée et soutenue publiquement

par

M. Laurent NGUYEN

le 18 décembre 2003

Titre :

Calibration de Modèles Financiers par Minimisation d'Entropie Relative et Modèles avec Sauts

Directeur de thèse :

M. Benjamin JOURDAIN

JURY

M. Gilles PAGES	**Président**
M. Rama CONT	**Examinateur**
M. Jean-Franck JALLET	**Examinateur**
M. Bernard LAPEYRE	**Examinateur**
M. Jean-Pierre LARDANT	**Examinateur**
M. Nizar TOUZI	**Rapporteur**

RESUME : Calibration de Modèles Financiers par Minimisation d'Entropie Relative et Modèles avec Sauts

Le smile de volatilité implicite observé sur les marchés d'options traduit l'insuffisance du modèle de Black et Scholes. Avec la nécessité d'élaborer un modèle d'actif financier plus satisfaisant, vient celle de sa calibration, objet de cette thèse.

La calibration par minimisation de l'entropie relative a été proposée récemment dans le cadre de la méthode de Monte-Carlo. On a étudié la convergence et la stabilité de cette méthode et on l'a étendue à des critères plus généraux que l'entropie relative. Pour qu'il y ait absence d'opportunité d'arbitrage, il faut que le sous-jacent actualisé soit une martingale. La prise en compte de cette nécessité est abordée sous l'angle d'un problème de moments.

Dans la deuxième partie, on a considéré un modèle simple du phénomène de krach en introduisant en particulier des sauts dans la volatilité du sous-jacent. On a calculé le risque quadratique et effectué un développement approché du smile qui constitue un outil pour la calibration.

Finalement, dans la troisième partie, on utilise l'entropie relative afin de calibrer l'intensité des sauts d'un modèle de diffusion avec sauts et volatilité locale. La stabilité de la méthode est prouvée grâce à des techniques de contrôle optimal ainsi qu'au théorème des fonctions implicites.

MOTS-CLES :

smile de volatilité, calibration de modèles, entropie relative, méthode de Monte-Carlo, I-projection généralisée, problème de moments, contrôle optimal des processus de diffusion avec sauts

ABSTRACT : Calibration of Financial Models by Relative Entropy Minimization and Models with Jumps

The smile observed in option markets is the evidence of the deficiency of the Black and Scholes model. With the necessity to find a more relevant model of financial assets comes the requirement of its calibration. This is the subject of the present work.

A calibration technique based on the minimization of relative entropy has been recently suggested in the framework of Monte-Carlo methods. We have proved the convergence and stability of this technique and extended the results to criteria more general than relative entropy. The martingale constraint on the underlying necessary to insure No Free Lunch has been examined from the point of vew of moment problems.

In the second part we have considered a simple model of crashes by introducing jumps in the volatility process. The quadratic risk has been computed and approximate closed formulae of the smile have been obtained for the calibration.

Finally we have used the relative entropy criterion to calibrate the jump intensity of a jump diffusion model. The stability of this method has been proved by means of optimal control techniques combined with the implicit function theorem.

KEY WORDS :

volatility smile, calibration of models, relative entropy, Monte-Carlo method, generalized I-projection, moment problem, optimal control of jump diffusion processes

Remerciements

Cette thèse a été réalisée dans le cadre d'une convention CIFRE établie entre le Crédit Industriel et Commercial (CIC) et le CERMICS, laboratoire commun de l'Ecole Nationale des Ponts et Chaussées (ENPC) et de l'INRIA. Que les dirigeants de ces instituts reçoivent ici l'expression de ma gratitude.

Je remercie Monsieur Cyril Le Touzé ainsi que Monsieur Bernard Paget, Responsables de la Salle de Marché au CIC, d'avoir favorisé la réalisation de ma thèse.

Je remercie Monsieur Nicolas Bouleau, Président du Collège Doctoral de l'ENPC, d'en avoir suivi l'évolution.

Je remercie Monsieur Jean-Pierre Lardant, Sous-Directeur au CIC, de m'avoir accueilli dans son équipe et de m'avoir chaleureusement soutenu durant ce travail.

Monsieur Bernard Lapeyre, Directeur du CERMICS et Professeur à l'ENPC, m'a ouvert les portes de son laboratoire et je l'en remercie ainsi que du soutien et de la sympathie qu'il m'a témoignés.

Je remercie Monsieur Benjamin Jourdain, Professeur à l'ENPC d'avoir accepté de diriger ma thèse. Au cours de ces trois ans de recherche, il m'a toujours manifesté une attention éclairée et amicale et ce travail lui doit beaucoup.

Monsieur Jean-Franck Jallet, Responsable de l'activité Trading Convertibles au CIC, a suivi le déroulement de ma thèse. Son aisance et sa compétence dans le monde de la finance m'ont été d'un grand bénéfice et je l'en remercie.

Je remercie Monsieur Gilles Pagès, Professeur à l'Université Paris VI, ainsi que Monsieur Nizar Touzi, Professeur à l'ENSAE, d'avoir accepté d'établir les rapports sur mon travail. Leurs appréciations l'ont amplement récompensé.

Je remercie Monsieur Rama Cont, Chargé de recherche au CNRS de m'avoir fait l'honneur de participer à mon jury de thèse. J'ai tiré profit de son point de vue sur la calibration en finance qu'il a exposé à de nombreuses occasions.

Je remercie Monsieur David Perez, Responsable du système d'information au CIC, d'avoir trouvé le financement qui a permis l'achèvement de ma thèse.

Je remercie Madame Alice Tran, Responsable du suivi des doctorants à l'ENPC, de s'être occupée si aimablement de mon dossier de thèse et d'avoir assuré le bon déroulement de ma soutenance.

Je remercie Madame Fabienne Espitalier, Secrétaire au CIC, de s'être occupée avec gentillesse et efficacité de l'organisation de mes divers voyages : à Londres, Aussois, Barcelone, Agia Pelagia, Nice.

Ma recherche a bénéficié des conférences auxquelles j'ai assisté lors de groupes de travail hebdomadaires organisés à l'ENPC, à l'Université de Marne la Vallée ainsi qu'à l'INRIA, sur les thèmes de recherches du projet Mathfi. Je tiens donc à en remercier les organisateurs pour la qualité de ces séminaires.

Je remercie Monsieur Claude Martini, Chargé de recherche à l'INRIA, de m'avoir invité à exposer une partie de mes travaux dans la session qu'il organisait avec Monsieur Olivier Pironneau, Professeur à l'Université Paris 6, au cours d'une conférence internationale de mathématiques (AMAM 2003 à Nice).

L'enseignement de DEA de Madame Nicole El Karoui, Professeur à l'Université Paris 6, m'a été précieux pour aborder les mathématiques financières. Je tiens à l'en remercier, également pour une conversation plus personnelle, lors d'un séminaire à Aussois sur le thème de la calibration en finance.

J'ai reçu l'aide aimable de personnes auxquelles j'exprime ici ma reconnaissance : je remercie Madame Agnès Sulem, Directeur de recherche à l'INRIA, pour une discussion sur le contrôle des processus avec sauts et pour les références qu'elles m'a indiquées.
Monsieur Damien Lamberton, Professeur à l'Université de Marne la Vallée, Madame Marie-Claire Quenez, Maître de conférence à l'Université de Marne la Vallée, et Madame Anna-Lisa Amadori, Chercheur au IAC, CNR de Rome, ont eu la gentillesse de me donner des articles ainsi que des références bibliographiques. Je les en remercie donc. Je remercie Monsieur Jean-François Delmas, Professeur à l'ENPC, de m'avoir prêté un ouvrage indispensable en statistiques et de m'avoir indiqué certains logiciels pour la simulation ou la présentation. Je remercie Monsieur Jacques Printems, Maître de conférence à l'Université Paris 12, pour un avis fort utile concernant les espaces de Sobolev. Je remercie également Monsieur Michel Cohen de Lara, Professeur à l'ENPC, et Monsieur Roy Cerqueti, Chercheur à l'Université de Rome, pour une référence essentielle en analyse convexe. Je remercie Madame sandrine Hénon, Chercheur à l'Université de Marne la Vallée, d'avoir eu la gentillesse de me remettre un article sur le modèle SABR.

L'élaboration d'un pricer a été l'occasion de travailler avec Monsieur Julien Mosson, trader au CIC. Je le remercie pour cette réalisation commune, dont il a assumé l'essentiel de l'implémentation, et qui m'a été à la fois agréable et profitable.

Je remercie Monsieur Jacques Daniel, Ingénieur système au CERMICS, de m'avoir aidé à installer la version LaTeX utilisée pour la rédaction du présent document.

Je remercie Monsieur Claude Thomas, Responsable du projet d'automate au CIC, de m'avoir éclairé sur quelques difficultés du C++ que j'ai rencontrées en implémentant un algorithme de calibration.

J'ai particulièrement apprécié l'ambiance conviviale qui règne au CERMICS aussi bien qu'au sein de l'équipe de trading sur dérivés d'actions du CIC. Je remercie donc l'ensemble de mes collègues, opérateurs de marché, enseignants chercheurs, informaticiens, commerciaux et documentalistes, Madame Sylvie Berte, Secrétaire au CERMICS, avec lesquels il a été agréable de passer ces trois dernières années.

Je remercie les sympathiques collègues avec lesquels j'ai plaisir à travailler : Messieurs Jean-Franck Jallet, Gilles Morihain, Thomas Quillet. Pour le déjeuner du midi ou le café, j'ajoute Messieurs Claude Thomas, Edgar Mfoumoune, Nassim Mezouar et Philippe Savitch. J'inclue dans ces remerciements Monsieur Didier Mandin et Monsieur Julien Mosson pour les nombreux restaurants chinois ou autres, qu'il faudra renouveler.

Parce qu'il a été agréable de faire évoluer ma thèse en leur compagnie, je remercie Messieurs les traders sur dérivés d'actions : Eric Savard, Frédéric Herbette, Jean-Louis Hue, Olivier Cazeaux, Patrick Rolland ; Messieurs les arbitragistes : Arnaud Yvinec, David Lenfant, Eric Robbe, Gabriel Teodorescu ; Mesdames et Messieurs les commerciaux : Adolphe Baraderie, Alexandra Giancarli, Alexandra Turkman, Geoffroy de Bouillane ; Madame Agnès Bonafous, Responsable de la recherche à EIFB ; Monsieur Nicolas Robin, chargé du contrôle de risques au CIC ; Messieurs les traders sur produits structurés : Chafik Yassini, Mounir Bouba, Patrice Flambard ainsi que Monsieur Olivier Maigne, Responsable de leur équipe ; mon collègue chercheur de l'équipe de trading sur taux, Monsieur Hunor Albert-Lorincz, ainsi que Monsieur Pierre Yves Perret, Responsable de cette équipe.

Je remercie Messieurs Antonino Zanette et Bouhari Arouna avec lesquels j'ai connu des instants d'émotion intense mais également de chaleureuse détente à l'occasion de conférences internationales.

Je souhaite aussi remercier les camarades chercheurs que je n'ai pu citer jusqu'à présent et avec lesquels j'ai eu le plaisir de partager des moments d'amitié, à l'occasion de groupes de travail, de repas de restaurants et de pots, ou en tant que voisin de leurs bureaux : Mesdames et Messieurs Adel Ben Haj Yedder, Adrien Blanchet, Ahmed Kbaier, Bruno Deutsch, Christophe Chorro, Emmanuel Temam, Fabien Lejeune, Marie-Pierre Bavouzet, Maya Briani, Mohammed Ben-Alaya, Nicola Moreni, Pierre Cohort, Raffaella Carbone, Ralf Laviolette, Samuel Njoh, Tony Lelièvre, Vlad Bally et Yousra Gati.

J'exprime mon profond attachement à toutes les personnes dont la pensée m'est chère, ma famille en France, aux Etats-Unis, au Canada et au Viêt-Nam, ainsi que mes amis, mon Parrain ainsi que toute sa famille, Frédéric, Houcine et leurs familles, tous mes proches, tous ceux avec qui je passe de bons moments.

Je désire associer à mes remerciements la mémoire de mes Grands-Parents paternels et la pensée de ma Grand-Mère maternelle.

J'ai une pensée particulière pour Bertrand et Corinne, avec qui je partage tant de joies familiales, le bonheur qu'ils vivent avec Isabelle et Jean-Marie, celui d'avoir vu naître et de voir grandir Julie, Olivier et Léa.

Je termine ces remerciements par ceux que j'adresse affectueusement à ma mère Elisabeth et à mon père Hoe. Cette thèse leur est dédiée.

Table des matières

Introduction **4**

I Convergence d'une suite minimisante pour l'entropie relative **8**

I.1 Condition nécessaire d'existence au problème (P_n) 9

I.2 Convergence vers la solution de (PL) sous la condition (C) 10

I.3 Remarque sur l'entropie relative en tant que mesure de l'information et sur le principe du maximum d'entropie 17

II Extension des résultats à des critères de minimisation plus généraux **20**

II.1 Discussion des hypothèses sur le critère J 22

II.2 Problème contraint 23

II.2.1 Etude du problème limite sous la condition (C) 25

II.2.2 Convergence de la solution du problème discret sous la condition (C) . . . 28

II.2.3 Prise en compte de contraintes d'équivalence de la mesure de pricing et de la mesure a priori 36

II.2.4 Résultats connus de $\widehat{M}$-estimateurs 38

II.3 Problème pénalisé 39

II.3.1 Discussion des hypothèses sur le critère χ 41

II.3.2 Etude du problème limite 42

II.3.3 Convergence de la solution du problème discret 46

II.4 Calibration pénalisée sous une contrainte de type ensemble 49

II.5 Compléments : généralisation du choix de l'entropie dans le cas où la fonction de payoff est bornée 53

II.5.1 Problème contraint 54

II.5.2 Pénalisation du problème 55

II.5.3 Conclusion 57

III Minimisation de l'entropie relative pénalisée et solution généralisée **58**

III.1 Caractérisation de l'attracteur des suites minimisantes 59

III.2 Etude de la solution généralisée dans le cas d'un terme de pénalisation quadratique 64

III.2.1 Critère pour que la solution soit usuelle 64

III.2.2 Quelques exemples 66

III.2.3 Effet de la pénalisation sur la nature de la solution 68

III.2.4 Un contre-exemple à l'emploi d'une pénalisation forte 70

IV Théorème central limite associé au problème de minimisation d'une entropie **74**

IV.1 Etude générale de l'erreur Monte-Carlo 75

IV.1.1 Théorème central limite associé à la suite des $\widehat{M}$-estimateurs 75

IV.1.2 Vitesse de convergence de la suite des estimateurs du prix 77

IV.2 Cas de l'entropie relative 79

V Stabilité de la calibration **82**

VI Approximation de la probabilité martingale calibrée minimale **89**

VI.1 Convergence du problème de martingales approché 91

VI.2 Cas markovien 93

VI.2.1 Etude du problème discrétisé en temps 94

VI.2.2 Convergence par raffinement de la discrétisation en temps 97

VI.3 Approche pénalisée 99

VI.3.1 Problème de base de la probabilité martingale 99

VI.3.2 Cadre markovien 101

VI.4 Caractérisation de la probabilité martingale calibrée minimale 103

VII Prix d'options dans un modèle où la volatilité saute avec le sous-jacent **105**

VII.1 Une modélisation du krach 105

VII.2 Cas d'un retour progressif de la volatilité 106

VII.2.1 Présentation du modèle 106

VII.2.2 Valorisation d'une option européenne 107

VII.2.3 Couverture : calcul du delta 109

VII.3 Cas d'un retour brutal de la volatilité 115

VII.3.1 Le modèle . . . 115
VII.3.2 Valorisation . . . 116
VII.3.3 Un calcul de compensateur . . . 117
VII.3.4 Couverture . . . 121
VII.3.5 Expression développée du smile . . . 125

VIII Calibration d'une diffusion avec sauts **130**

VIII.1 Position du problème . . . 130
VIII.2 Choix d'un ensemble de probabilités martingales . . . 133
VIII.3 Pénalisation du problème . . . 140
VIII.3.1 Formulation générale . . . 140
VIII.4 Formulation HJB du problème . . . 142
VIII.4.1 Enoncé des résultats . . . 143
VIII.4.2 Preuve des résultats de l'étude précédente . . . 146
VIII.5 Régularisation du problème pénalisé et solution dans les espaces Hölderiens . . 155
VIII.5.1 Espaces fonctionnels utilisés . . . 156
VIII.5.2 Solution de l'équation de la programmation dynamique dans les espaces Hölderiens . . . 157
VIII.5.3 Preuve des résultats de l'étude précédente . . . 160
VIII.6 Cas où la mesure de Lévy est bornée . . . 175
VIII.6.1 Contrôle dépendant explicitement de la taille des sauts . . . 175
VIII.6.2 Contrôle ne dépendant pas explicitement de la taille des sauts . . . 175
VIII.7 Résolution du problème de calibration pour les fonctions de payoff non délocalisées 177
VIII.7.1 Formulation HJB pour les fonctions de payoff non délocalisées . . . 178
VIII.7.2 Régularité de la solution du système HJB . . . 180
VIII.7.3 Etude de la délocalisation des fonctions de payoff . . . 181
VIII.8 Perspective de solutions avec des hypothèses de régularité faibles . . . 183

Conclusion et perspectives **185**

Bibliographie **188**

Introduction

Le marché des options, instruments de spéculation mais également assurances contre le mouvement du prix des actions, s'est développé avec l'élaboration de modèles permettant de les valoriser. Les options doivent ainsi leur succès à la formule de Black et Scholes qui donne le prix d'un "Call" (option d'achat) et définit la couverture du vendeur par l'achat d'une certaine quantité de l'actif sous-jacent : la dérivée par rapport au prix de l'action ou "delta" est précisément cette quantité qui assure la réalisation d'un portefeuille insensible à la variation du prix du sous-jacent. Cette approche a été une étape révolutionnaire dans le monde de la finance, succédant à l'approche du portefeuille moyenne-variance introduite par Markowitz (voir la préface de [77] pour ce rappel historique).

Reprenant les idées de Bachelier émises en 1900 dans sa thèse [7], le modèle de Black et Scholes repose sur le mouvement Brownien dont l'importance avait été également reconnue dans les travaux de Samuelson [114]. Il a été l'outil mathématique indispensable à l'essor actuel des marchés d'options négociables, dont le premier, le Chicago Board Options Exchange, a été ouvert en 1973, l'année de la parution des deux articles fondateurs [16] et [92]. On renvoie à [66], [14] ainsi qu'au chapitre introductif de [96] pour avoir une idée précise du fonctionnement de ces marchés.
Ainsi qu'il est par exemple souligné au début du second chapitre de [14], la méthode de pricing proposée de façon indépendante par Black et Scholes d'une part et Merton d'autres part, offre une solution aux deux "problèmes majeurs" sur lesquels se heurtaient leurs prédecesseurs : d'une part, au lieu d'anticiper l'évolution future de l'action, on en détermine le comportement aléatoire en mesurant sa volatilité. D'autre part, l'actualisation ramenant au présent la valeur future de l'action doit se faire au taux sans risque : c'est le principe de la valorisation risque-neutre.
Ces deux principes sont la base des extensions de la théorie de Black, Scholes et Merton et permettent de les développer dans le cadre puissant des martingales selon [62] et [63] ainsi qu'il est exposé dans [96] par exemple. Le besoin de telles extensions s'est naturellement créé face à l'incapacité qu'a le modèle de Black-Scholes de rendre compte de phénomènes empiriquement observés tels que l'anticipation des fortes baisses des actions plutôt que celle des fortes hausses (le marché est pessimiste).
De manière précise, des tests statistiques ont été menés et ont abouti à contredire le fait que le logarithme du prix du sous-jacent suit une loi normale et que ses incréments sont indépendants, comme l'impliqueraient les hypothèses du modèle : ces aspects sont discutés au dernier chapitre de [98] ainsi qu'au sixième chapitre de [96] et l'on peut citer [91] dont l'étude montre que le prix d'une action est assez fidèlement distribué suivant une loi de Lévy tronquée, ainsi que le suggérait Mandelbrot dans [89] qui proposait ce type de lois dites à "queue épaisse" en remplacement de la loi de Gauss. De même, dans [57] a été établie une corrélation entre les incréments du logarithme du prix si le temps d'incrémentation était inférieur au quart d'heure.
Le marché constate en fait directement l'écart du comportement du prix d'une action au modèle de Black-Scholes par l'impossibilité de retrouver les prix des options du marché écrites sur cette

action en affectant une volatilité unique à la formule de Black-Scholes : au contraire, en inversant cette formule par rapport à la volatilité, on calcule pour chaque prix d'option la volatilité implicite correspondante et l'on obtient une courbe non constante dont la forme caractéristique lui donne le nom de smile. D'un point de vue mathématique, l'impossibilité de retrouver le smile du marché en utilisant la formule de Black-Scholes se comprend puisque l'on ne dispose que d'un seul paramètre avec la volatilité.
Pour reconstruire le smile, de nombreux modèles ont donc été proposés, rangés sommairement dans les deux grandes familles que sont les modèles à volatilité stochastique et les modèles à sauts, une combinaison des deux étant souvent préférée afin d' améliorer la structure du smile obtenu. Dans l'utilisation pratique de tels modèles, il est nécessaire de les calibrer, c'est-à-dire d'associer à une méthode numérique de calcul de prix une procédure d'intégration du smile afin de déterminer des prix en cohérence avec les données du marché.

La thèse se compose de trois parties indépendantes, présentant des approches différentes du problème de la calibration.
La première partie est constituée de six chapitres. Elle aborde une technique simple proposée par Avellaneda& all [6] afin de calibrer le pricing d'options par la méthode de Monte-Carlo. Cette dernière est intéressante parce que son implémentation est simple et elle s'impose dès que les méthodes déterministes se révèlent inopérantes, typiquement en grande dimension. Sa mise en oeuvre et ses champs d'applications sont exposés dans l'ouvrage élémentaire [61] et des techniques modernes de simulation sont présentées dans [84] et [78] par exemple.
Le développement d'une technique de calibration appropriée est donc crucial. Selon les idées de [6], il s'agit d'utiliser la minimisation de l'entropie relative afin de déterminer des poids à affecter aux trajectoires Monte-Carlo pour retrouver les contraintes de prix. La valorisation des options se fait alors grâce au Monte-Carlo pondéré.
Deux problèmes théoriques ont motivé l'étude de cette partie. D'une part, la convergence des prix était à prouver. D'autre part, la question de l'absence d'opportunité d'arbitrage se pose et nécessitait d'être approfondie.
Sous une condition géométrique que doivent vérifier les prix de calibration, pour que cette dernière soit possible, on a d'abord obtenu la convergence attendue en établissant un résultat de M-estimateurs. L'ensemble de ces résultats a fait l'objet de la publication [71].
L'extension des résultats précédents a été faite dans deux directions : d'abord, en relâchant les contraintes de calibration, on est conduit à la minimisation sans contrainte de l'entropie relative pénalisée, ce qui évite en particulier d'exiger une hypothèse sur les prix de calibration.
En second lieu, on peut préciser des conditions générales pour que les résultats de convergence subsistent si l'on adopte un critère convexe autre que l'entropie relative. La décision de changer de critère peut se justifier précisément avec des arguments de nature financière [10]. Dans le cas où on ne suppose pas bornée la fonction de payoff des options par rapport auxquelles on calibre, on a cependant commencé à étudié une famille de critères très semblables à l'entropie relative. Une discussion complémentaire autour de résultats de la littérature permet toutefois d'élargir la classe des critères convenables dès que le payoff des options de calibration est borné, ce qui permet d'envisager d'autres fonctionnelles convexes proposées dans la théorie des problèmes inverses.
Un des phénomènes intéressants mis en évidence par l'étude de la convergence du Monte-Carlo pondéré est la perte potentielle de calibration à la limite, du simple fait que, même dans le cas pénalisé, il n'y a pas toujours de solution au problème de calibration limite, formulé en remplaçant la loi empirique du sous-jacent par la loi sous laquelle les tirages sont faits.
Dans le cas de l'entropie relative avec des contraintes non relâchées, les travaux classiques de

Jupp & Mardia [72] et de Csiszar [38] permettent l'identification de la limite Monte-Carlo avec la I-projection généralisée introduite par Csiszar, ce qui rend encore plus naturelle la convergence obtenue. Ce phénomène est noté dans [71] et l'on en étend la portée au cas de l'entropie relative pénalisée en mettant l'accent sur ce qui peut empêcher le problème de calibration limite d'avoir une solution et donc la limite des solutions Monte-Carlo d'être calibrée.
Cette situation pathologique ne se rencontre cependant pas dans le cas où les options de calibration sont de payoff borné par exemple. Sous des conditions générales, on peut alors établir un théorème central limite permettant d'associer une erreur au prix, et l'on démontre ensuite la stabilité de la méthode par rapport aux variations des prix de calibration ainsi qu'aux erreurs numériques d'optimisation.
Quand on souhaite imposer la condition de parité Call-Put par exemple, il n'est toutefois plus gratuit de supposer que les contraintes sont bornées. C'est donc dans le cas d'un modèle d'actif borné que l'on termine l'étude sur la question de l'absence d'opportunité d'arbitrage. Cette dernière est motivée par la perte de la propriété de martingalité du sous-jacent actualisé quand sa loi est donnée par la solution du problème de calibration limite (sous réserve d'existence). En ajoutant des contraintes de martingalité, on est conduit à un problème de convergence de moments étudié notamment par Borwein & Lewis [21] [22] ainsi que Teboulle & Vajda [123]. Le lien avec les martingales restait à faire et l'on reconnait également le rôle joué par l'entropie relative quand la propriété de Markov est en jeu.

Dans la deuxième partie, on se propose d'expliquer le smile en modélisant d'un point de vue pratique le phénomène de krach. De façon assez réaliste, on peut considérer que ce dernier se traduit par une baisse brutale du sous-jacent accompagnée par un saut à la hausse de la volatilité dont le retour à la normale se fait brusquement, au bout d'un temps aléatoire, tandis que l'on adopte finalement le modèle Black-Scholes en dehors de ces périodes critiques. Ce modèle est l'un des plus simples intégrant une volatilité stochastique avec des sauts et l'on sait exprimer l'erreur quadratique de couverture des options standard, la difficulté technique étant essentiellement dûe aux temps de retour de la volatilité. On obtient enfin un développement approché du smile qui est utile pour calibrer le modèle.

Dans la dernière partie, on étudie la calibration d'un modèle où des sauts de l'actif sous-jacent se superposent à un terme de volatilité locale. De nombreux travaux ont été menés en vue de la calibration du modèle à volatilité locale qui, parmi les modèles à volatilité stochastique, est la variante la plus simple du modèle de Black-Scholes : elle l'enrichit en autorisant la volatilité à dépendre de façon déterministe du sous-jacent et du temps et elle a été popularisée par la formule de Dupire. Celle-ci n'est cependant qu'un outil théorique puisqu'elle exige de disposer d'une nappe continue et même lisse de prix de calibration. Si l'on veut s'affranchir de l'arbitraire de l'interpolation et du lissage de la nappe de ces prix, les méthodes auxquelles on peut recourir pour déterminer la fonction de volatilité locale conduisent à la minimisation d'un critère et l'on peut citer par exemple [82], [67], [17], [27] ainsi que [5]. Dans ce dernier article, le critère retenu dérive de l'entropie relative mais conduit à un comportement insatisfaisant de la surface de volatilité aux maturités des options de calibration. Par ailleurs, un travail récent [113] suggère de surmonter ce problème grâce à un procédé de régularisation des payoff de ces dernières.
Pour expliquer la structure du smile pour les maturités courtes, il est important d'inclure des sauts dans le modèle ainsi que cela a d'abord été proposé dans [93]. Les développements récents introduisent les sauts à partir des processus de Lévy et il est ainsi construit dans [35] une méthode de calibration par minimisation de l'entropie relative pénalisée pour les processus de

Lévy géométriques. Dans le cas d'une diffusion avec sauts plus générale, incorporant en particulier une volatilité locale, il apparaît également naturel de retenir le critère de l'entropie relative pénalisée pour la calibration du modèle, ce qui implique que cette dernière porte sur la mesure intensité des sauts au lieu de la fonction de volatilité locale alors déterminée à l'avance. En s'inspirant des techniques de [113] issues de la théorie HJB du contrôle optimal, on arrive à un résultat d'existence et de stabilité de la solution du problème de calibration. De façon précise, avec des hypothèses de type Holdérien sur les coefficients du modèle ainsi que sur les fonctions de payoff des options de calibration, la régularité de la calibration par minimisation de l'entropie relative est une conséquence du théorème des fonctions implicites.

Chapitre I

Convergence d'une suite minimisante pour l'entropie relative

Avellaneda & all [6] ont proposé une méthode de type Monte-Carlo pour calibrer des modèles d'actifs sous-jacents de manière à retrouver des prix d'options liquides. La finalité est de pouvoir donner un prix à des options peu liquides tout en restant cohérent avec les prix du marché.
Pour fixer le cadre de l'étude, on considère que le processus sous-jacent prend ses valeurs dans un espace S, supposé polonais pour la suite, suivant une loi a priori μ, selon laquelle on simule une suite de variables indépendantes $(X_i)_{i\geq 1}$ sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
La méthode proposée consiste alors à corriger la méthode de Monte-Carlo en construisant une probabilité a posteriori compatible avec les prix $C^1, ..., C^d$ observés d'un nombre d d'options définies par leurs fonctions de payoff $f^1, ..., f^d : S \to \mathbb{R}$, supposées boréliennes.
La correction se fait en affectant aux n premiers tirages $X_1, ..., X_n$ des poids $p_1, ..., p_n$ positifs et de somme 1, de façon à satisfaire à $\forall 1 \leq j \leq n,\ \sum_{i=1}^n p_i f^j(X_i) = C^j$, en supposant que les payoff ont été actualisés. Quitte à retrancher $C = (C^1, ..., C^d))$ à $f = (f^1, ..., f^d)$, on supposera également $C = \mathbf{0}$.
Puisque le choix de la mesure a priori reflète l'idée que l'on se fait du comportement du sous-jacent, il est souhaitable que la pondération retenue soit aussi proche que possible de la pondération uniforme $\forall 1 \leq i \leq n,\ p_i = 1/n$, ce qui impose le choix d'un critère qui est ici la minimisation de $\sum_{i=1}^n p_i \ln(p_i)$.
Il s'agit de la maximisation de l'entropie des p_i, définie par $-\sum_{i=1}^n p_i \ln(p_i)$, mais cela revient également à la minimisation de l'entropie relative $H(\nu_n \| \mu_n)$ de $\nu_n = \sum_{i=1}^n p_i \delta_{X_i}$ par rapport à la mesure empirique $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$, où l'on définit, pour deux mesures ν et ν' de l'espace $\mathcal{P}(S)$ des probabilités sur S :

$$H(\nu \| \nu') = \begin{cases} \int \ln(\frac{d\nu}{d\nu'}) d\nu & \text{si } \nu \text{ est absolument continue par rapport à } \nu' \\ +\infty & \text{sinon} \end{cases}$$

Avec ces notations, l'approche décrite plus haut revient à résoudre le problème

$$(P_n) \quad \text{trouver } \nu_n \in \mathcal{P}(S) \text{ qui minimise } H(\nu_n \| \mu_n) \text{ sous la contrainte } \int f d\nu_n = \mathbf{0}$$

On remarque que la convexité stricte de l'entropie implique que ce problème a au plus une solution.
La question qui se pose, qui n'est pas abordée dans [6], est donc de savoir à quelle condition avec une probabilité non nulle (P_n) admet une solution pour n assez grand, et également d'étudier le comportement asymptotique de cette solution quand n augmente.

A cet effet, il est naturel de faire apparaître le problème limite

$$(PL) \quad \text{trouver } \nu \in \mathcal{P}(S) \text{ qui minimise } H(\nu\|\mu) \text{ sous la contrainte } \int f d\nu = \mathbf{0}$$

Ici encore, avec la convexité stricte de l'entropie, l'unicité de la solution est acquise.
D'après Csiszar [38], s'il existe une probabilité $\eta \in \mathcal{P}(S)$ telle que $H(\eta\|\mu) < +\infty$ et $\int_S f d\eta = \mathbf{0}$, alors il existe un sous-espace E de $\mathbb{R}^d$ et $\theta \in E$ vérifiant $0 < \int_S 1_E(f(x))e^{\theta.f(x)}d\mu(x) < +\infty$ tels que toute suite minimisante pour le problème (PL) converge au sens de la norme en variation vers la distribution de Boltzmann $d\mu_E^\theta(x) = 1_E(f(x))e^{\theta.f(x)}d\mu(x)/\int_S 1_E(f(y))e^{\theta.f(y)}d\mu(y)$.
Si (PL) admet une solution alors elle est égale à μ_E^θ.
Mais il se peut que $\int_S f d\mu_E^\theta \neq \mathbf{0}$; on dit alors que μ_E^θ est la solution généralisée de (PL).
Cette présentation du problème reprend de manière développée la note [71] établissant la consistance de la méthode de calibration par minimisation d'entropie relative.
Dans une première partie, on donne une condition nécessaire pour qu'avec probabilité strictement positive, il existe n t.q. (P_n) admet une solution.
Dans une seconde partie cette condition est supposée satisfaite et l'on montre alors l'existence de la solution généralisée de (PL), en la caractérisant.
On montre ensuite que p.s. $\exists N$ t.q. $\forall n \geq N$ (P_n) admet une solution ν_n qui converge étroitement vers la solution généralisée de (PL) lorsque $n \to +\infty$.

I.1 Condition nécessaire d'existence au problème (P_n)

Le lemme suivant donne une condition nécessaire pour qu'avec probabilité non nulle, il exite n t.q. (P_n) admet une solution :

Lemme I.1.1 *Pour que $\mathbb{P}(\exists n \geq 1$ et $\eta_n \in \mathcal{P}(S)$ t.q. $H(\eta_n\|\mu_n) < +\infty$ et $\int_S f d\eta_n = \mathbf{0}) > 0$, il faut qu'il existe un sous-espace vectoriel E de $\mathbb{R}^d$ tel que si $\mu^f_{|E}$ désigne la restriction à E de l'image de μ par f,*

$$(C) \begin{cases} \mathbf{0} \in Int_E(Conv(Supp(\mu^f_{|E}))) \text{ l'intérieur dans } E \text{ de l'enveloppe convexe du support de } \mu^f_{|E} \\ \forall \eta \in \mathcal{P}(S) \text{ absolument continue par rapport à } \mu \text{ t.q. } \int \|f\| d\eta < +\infty \text{ et } \int_S f d\eta = \mathbf{0}, \\ \text{on a } \eta(f^{-1}(E)) = 1 \ . \end{cases}$$

Preuve : Soit $A = \{\exists n \geq 1$ et $\eta_n \in \mathcal{P}(S)$ t.q. $H(\eta_n\|\mu_n) < +\infty$ et $\int_S f d\eta_n = \mathbf{0}\}$.
On suppose donc $\mathbb{P}(A) > 0$.
Pour $\omega \in A$, η_n est absolument continue par rapport à μ_n et il existe donc des nombres p_i, $1 \leq i \leq n$ avec $p_i \geq 0$ et $\sum_{i=1}^n p_i = 1$, tels que $\eta_n = \sum_{i=1}^n p_i \delta_{X_i}$.
La construction de E s'effectue par une récurrence descendante sur la dimension d'un sous-espace F de $\mathbb{R}^d$ que l'on initialise en posant $F = \mathbb{R}^d$.
Supposons que pour $\mathbb{P}$-presque tout $\omega \in A$, $\eta_n = \sum_{i=1}^n 1_F(f(X_i))p_i\delta_{X_i}$, propriété qui est bien sûr vérifiée à l'initialisation.
Alors $\int_S f d\eta_n = \sum_{i=1}^n 1_F(f(X_i))p_i f(X_i) = \mathbf{0}$ ce qui implique que $\mathbf{0} \in Conv(\{f(X_i),\ i \in [1,n]\} \cap F)$.
Comme p.s. $\forall m \geq 1$, $Conv(\{f(X_i),\ i \in [1,m]\} \cap F) \subset Conv(Supp(\mu^f_{|F}))$, on en déduit que pour $\mathbb{P}$-presque tout $\omega \in A$

$$\mathbf{0} \in Conv(\{f(X_i),\ i \in [1,n]\} \cap F) \subset Conv(Supp(\mu^f_{|F})). \tag{I.1.1}$$

– Si $\mathbf{0} \in Int_F(Conv(Supp(\mu^f_{|F})))$, on pose $E = F$.

– Sinon $\mathbf{0} \in Fr_F(Conv(Supp(\mu^f_{|F})))$.
Dans ce cas, soit H un hyperplan d'appui de $Conv(Supp(\mu^f_{|F}))$ (dans F) en $\mathbf{0}$ d'équation $\alpha.y = 0$ avec $\alpha.y > 0$ pour tout $y \in Conv(Supp(\mu^f_{|F})) \setminus H$.
Pour $\omega \in A$ t.q. (I.1.1) et $\eta_n = \sum_{i=1}^n 1_F(f(X_i))p_i\delta_{X_i}$, comme $\alpha.\int_S fd\eta_n = 0$ ce qui s'écrit aussi $\sum_{i=1}^n p_i 1_F(f(X_i))\alpha.f(X_i) = 0$, on a :
$$\forall 1 \leq i \leq n,\ p_i > 0 \Rightarrow f(X_i) \in H$$
Ainsi pour presque tout $\omega \in A$, $\eta_n = \sum_{i=1}^n p_i 1_H(f(X_i))\delta_{X_i}$ et l'hypothèse de récurrence est vérifiée pour le sous-espace H avec $dim(H) = dim(F) - 1$.

Comme lorsque F est de dimension nulle, $Conv(Supp(\mu^f_{|F})) = Int_F(Conv(Supp(\mu^f_{|F})))$, E est bien défini par ce procédé.
Notons que nous avons également montré que pour $\mathbb{P}$-presque tout $\omega \in A$, $\eta_n = \sum_{i=1}^n 1_E(f(X_i))p_i\delta_{X_i}$ et
$$\eta_n(f^{-1}(E)) = 1. \tag{I.1.2}$$
Par une adaptation du raisonnement de récurrence descendante qui précède, on montre que $\eta(f^{-1}(E)) = 1$ si η est une probabilité absolument continue par rapport à μ t.q.$\int_S fd\eta = \mathbf{0}$ ∎

On verra plus loin que la condition (C) définit le sous-espace E de manière unique, ce qui n'est pas complètement évident puisque le choix de l'hyperplan relatif d'appui n'est pas forcément unique dans la construction récursive de la preuve précédente.
On peut cependant supposer que le choix de E est arbitraire pour l'instant.

I.2 Convergence vers la solution de (PL) sous la condition (C)

On suppose désormais que E est un sous-espace vectoriel de $\mathbb{R}^d$ t.q. la condition nécessaire (C) énoncée dans le Lemme I.1.1 est vérifiée.
La transformée de Laplace $\theta \in E \mapsto Z(\theta) = \int 1_E(f(x))e^{\theta\cdot f(x)}d\mu(x)$ est alors strictement convexe sur E, puisque, si ce n'était pas le cas, $\mu^f_{|E}$ serait concentrée sur un hyperplan relatif de E, en contradiction avec $\mathbf{0} \in Int_E(Conv(Supp(\mu^f_{|E})))$.
Notons que $\inf_{\theta\in E} Z(\theta) \leq Z(\mathbf{0}) \leq 1$.
De toute suite $(\theta_n)_n$ d'éléments de E non bornée on peut extraire une sous-suite $(\theta_{n'})_{n'}$ t.q. $\|\theta_{n'}\| \to +\infty$ et $\theta_{n'}/\|\theta_{n'}\| \to \alpha$.
Forcément, E n'est pas réduit à $\{\mathbf{0}\}$ et la condition $\mathbf{0} \in Int_E(Conv(Supp(\mu^f_{|E})))$ entraîne alors que $\lim_{n'} Z(\theta_{n'}) = +\infty$.
En effet, puisque E est de dimension non nulle, cette condition entraîne d'abord le fait que l'on peut trouver $a,\ \epsilon > 0$ tels que :
$$\mu^f_{|E}(\alpha\cdot x > 2\epsilon) > 2a$$
Il existe alors un borélien borné U de E tel que
$$\mu^f_{|E}(x \in U,\ \alpha\cdot x > 2\epsilon) > a$$
et l'on en déduit, pour n' assez grand :
$$\mu^f_{|E}\left(x \in U, \frac{\theta_{n'}}{\|\theta_{n'}\|}\cdot x > \epsilon\right) > a$$

et donc

$$Z(\theta_{n'}) \geq a e^{\epsilon \|\theta_{n'}\|} \to +\infty$$

Par conséquent, toute suite minimisante pour $Z(\theta)$ est bornée et on peut en extraire une sous suite qui converge vers un élément $\theta_* \in E$.
De plus, par le lemme de Fatou, $Z(\theta_*) = \inf_{\theta \in E} Z(\theta)$ et la stricte convexité de $Z(\theta)$ entraîne l'unicité de θ_*.
D'après Jupp & all [72], on en déduit que la solution généralisée du problème (PL) où μ est remplacée par $\mu_E = \mu_{|f^{-1}(E)}/\mu(f^{-1}(E))$ existe et qu'elle est donnée par

$$d\mu_E^{\theta_*}(x) = 1_E(f(x)) \frac{e^{\theta_* . f(x)} d\mu(x)}{\int_S 1_E(f(y)) e^{\theta_* . f(y)} d\mu(y)}$$

En particulier, il existe une probabilité γ t. q. $H(\gamma|\mu_E) < +\infty$ et $\int_S f d\gamma = \mathbf{0}$.
Comme pour toute probabilité η t. q. $\eta(f^{-1}(E)) = 1$, on a $H(\eta|\mu) = H(\eta|\mu_E) - \ln \mu(f^{-1}(E))$, on en déduit que $H(\gamma|\mu) < +\infty$, donc que (PL) admet une solution généralisée.
En outre, d'après le Lemme I.1.1, une suite minimisante pour le problème (PL) est constituée de probabilités η t.q. $\eta(f^{-1}(E)) = 1$.
C'est donc aussi une suite minimisante pour le problème (PL) où μ est remplacée par μ_E. Ainsi,

Proposition I.2.1 *Sous (C), il y a existence de la solution généralisée du problème (PL) et elle est égale à $\mu_E^{\theta_*}$.*
En outre, une condition suffisante pour que $\mu_E^{\theta_*}$ soit solution de (PL) i.e. vérifie $\int_S f d\mu_E^{\theta_*} = \mathbf{0}$ est que $\theta_* \in Int_E(\{\theta \in E,\ Z(\theta) < +\infty\})$.

Si $\vartheta \in Int_E(\{\theta \in E,\ Z(\theta) < +\infty\})$, alors Z_θ est C^∞ au voisinage de ϑ et $\nabla_\theta Z(\vartheta) = \int_S 1_E(f(x)) f(x) e^{\vartheta . f(x)} d\mu(x)$.
La deuxième assertion de la Proposition, s'obtient en écrivant la condition d'Euler d'optimalité en θ_*.
Pour finir sur la solution généralisée, on voit que l'expression de cette dernière règle le problème de l'unicité du sous-espace E puisque, avec la condition (C), c'est le sous-espace engendré par le support de l'image par f de $\mu_E^{\theta_*}$: l'unicité de la solution généralisée implique l'unicité de E.
La caractérisation de la solution généralisée de (PL) qui précéde va nous permettre d'obtenir le comportement asymptotique de la solution de (P_n) pour $n \to +\infty$:

Théorème I.2.2 *Sous (C), presque sûrement il existe un rang N à partir duquel le problème (P_n) admet une solution ν_n et la suite $(\nu_n)_{n \geq N}$ converge étroitement vers la solution généralisée $\mu_E^{\theta_*}$ du problème (PL).*
De manière générale, si $\varphi : S \to \mathbb{R}$ est mesurable bornée, on obtient $\mathbb{P}$ p.s.

$$\int_S \varphi d\nu_n \to \int_S \varphi d\mu_E^{\theta_*}$$

Preuve : La propriété $\mathbf{0} \in Int_E(Conv(Supp(\mu^f_{|E})))$ permet de montrer l'existence de A_0 t.q. $\mathbb{P}(A_0) = 1$ et $\forall \omega \in A_0$, $\exists N$ t.q. $\forall n \geq N$, $\mathbf{0} \in Int_E(Conv(\{f(X_i),\ 1 \leq i \leq n\} \cap E))$.
En effet, avec l'hypothèse (C), il existe un voisinage polyèdrique Δ de $\mathbf{0}$ tel que

$$\Delta \subset Conv(Supp(\mu^f_{|E}))$$

Les sommets de Δ étant des barycentres de points de $Supp(\mu^f_{|E})$, on en déduit qu'il existe une famille $(x_i)_{1 \leq i \leq p}$ de points de $Supp(\mu^f_E)$ telle que :

$$\mathbf{0} \in Int_E(Conv(x_i, 1 \leq i \leq p))$$

Puisque $\forall \epsilon > 0,\ \forall 1 \leq i \leq p,\ \mu^f(B(x_i,\epsilon)\bigcap E) > 0$, on obtient, presque-sûrement :

$$\forall 1 \leq i \leq p,\ \exists n_i > 0,\ f(X_{n_i}) \in B(x_i,\epsilon)\bigcap E$$

Pour ϵ assez petit et avec le Lemme I.2.3, énoncé après cette preuve, on a $\mathbf{0} \in Int_E(Conv(\{f(X_i),\ 1 \leq i \leq n\} \cap E))$, dès que n est supérieur à $N = \max(n_i,\ 1 \leq i \leq p)$, et cela prouve ce que l'on voulait.
En reprenant le raisonnement effectuée pour la Proposition I.2.1, on obtient que pour $\omega \in A_0$ et $n \geq N$ la solution généralisée du problème (P'_n) analogue à (P_n) mais avec μ_n remplacée par $\sum_{i=1}^n 1_E(f(X_i))\delta_{X_i}/\sum_{i=1}^n 1_E(f(X_i))$ est

$$\nu_n = \frac{\sum_{i=1}^n 1_E(f(X_i))e^{\theta_n \cdot f(X_i)}\delta_{X_i}}{\sum_{i=1}^n 1_E(f(X_i))e^{\theta_n \cdot f(X_i)}},$$

où θ_n est l'unique point où $Z_n : \theta \in E \to \sum_{i=1}^n 1_E(f(X_i))e^{\theta \cdot f(X_i)}$ atteint son minimum.
En adaptant la preuve du Lemme I.1.1 (voir (I.1.2)), on obtient $A_1 \subset A_0$ t.q. $\mathbb{P}(A_1) = 1$ et $\forall \omega \in A_1$, $\forall m \geq 1$, $\forall \eta_m \in \mathcal{P}(S)$ t.q. $H(\eta_m \| \mu_m) < +\infty$ et $\int_S f d\eta_m = \mathbf{0}$, $\eta_m(f^{-1}(E)) = 1$.
Donc $\forall \omega \in A_1$, $\forall n \geq N$, toute suite minimisante pour (P_n) est aussi minimisante pour (P'_n) ce qui implique que la solution généralisée de (P_n) est ν_n.
Comme $\forall \theta \in E,\ Z_n(\theta) < +\infty$, on a $\int_S f d\nu_n = \mathbf{0}$ i.e. ν_n est solution classique de (P_n).
Comme $\mu(f^{-1}(E)) > 0$ et que S est polonais, en combinant des résultats sur la méthode du rejet et la loi forte des grands nombres, on obtient $A_2 \subset A_1$ avec $\mathbb{P}(A_2) = 1$ t.q. $\forall \omega \in A_2$,

$$\frac{\sum_{i=1}^n 1_E(f(X_i))e^{\theta_* \cdot f(X_i)}}{\sum_{i=1}^n 1_E(f(X_i))} \longrightarrow_{n\to+\infty} \frac{\int_S 1_E(f(x))e^{\theta_* \cdot f(x)}d\mu(x)}{\mu(f^{-1}(E))} = \frac{Z(\theta_*)}{\mu(f^{-1}(E))} \tag{I.2.1}$$

$$\sum_{i=1}^n 1_E(f(X_i))\delta_{(X_i,f(X_i))}/\sum_{i=1}^n 1_E(f(X_i)) \text{ converge étroitement vers } \mu^{(id,f)}_{|f^{-1}(E)}/\mu(f^{-1}(E)) \tag{I.2.2}$$

où $\mu^{(id,f)}_{|f^{-1}(E)}$ désigne l'image de la restriction $\mu_{|f^{-1}(E)}$ de μ à $f^{-1}(E)$ par $x \in S \to (x, f(x))$.
Dans ce qui suit, nous fixons $\omega \in A_2$ et nous allons montrer par un raisonnement de type convergence de M-estimateurs que $\theta_n \to_{n\to+\infty} \theta_*$.
On commence par montrer par l'absurde que $(\theta_n)_n$ est bornée.
Si ce n'est pas le cas, on extrait une sous-suite $(\theta_{n'})_{n'}$ avec $|\theta_{n'}| \to +\infty$ et $\theta_{n'}/\|\theta_{n'}\| \to \alpha$.
En utilisant $\mathbf{0} \in Int_E(Conv(Supp(\mu^f_{|E})))$ et (I.2.2), on peut alors montrer que pour n' assez grand $\int_S \theta_{n'}.f d\nu_{n'} > 0$, ce qui constitue une contradiction.
De façon plus précise, la dimension de E est nécessairement non nulle dans ce cas et l'on peut reprendre un raisonnement précédent en considérant un borélien borné U de E tel que :

$$\mu^f_{|E}(x \in U,\ \alpha \cdot x > 2\epsilon) > a$$

où a et ϵ sont des nombres strictements positifs suffisamment petits.
Avec la convergence étroite (I.2.2), on a également, en prenant pour U un ouvert relatif de E :

$$\liminf \mu^f_{n'|E}(x \in U,\ \alpha \cdot x > 2\epsilon) > a$$

On en déduit que, pour n' assez grand, on a

$$\mu^f_{n'|E}\left(x \in U,\ \frac{\theta_{n'}}{\|\theta_{n'}\|} \cdot x > \epsilon\right) > a$$

Comme pour $x \leq 0$, $xe^x \geq -1/e$, on en déduit, pour n' assez grand

$$\int \theta_{n'} \cdot f e^{\theta_{n'} \cdot f} d\nu_{n'} \geq -\frac{1}{e} + a\epsilon\|\theta_{n'}\|e^{\epsilon\|\theta_{n'}\|} \to +\infty$$

et donc, pour n' assez grand pour que le deuxième membre de l'inégalité précédente soit strictement positif :

$$\int \theta_{n'} \cdot f d\nu_{n'} > 0$$

Le résultat est alors prouvé et $(\theta_n)_{n\in\mathbb{N}^*}$ est donc bornée.
Soit maintenant θ_∞ la limite d'une sous-suite convergente de $(\theta_n)_n$ que l'on indexe toujours par n pour simplifier les notations.
En appliquant le lemme de Skorokhod à (I.2.2), on obtient sur un espace de probabilité auxiliaire $(G, \mathcal{G}, Q)$ des variables (Y_n, W_n), $n \geq N$ (resp. (Y, W)) de loi $\sum_{i=1}^n 1_E(f(X_i))\delta_{(X_i, f(X_i))}$ $/ \sum_{i=1}^n 1_E(f(X_i))$ (resp. $\mu^{(id,f)}_{|f^{-1}(E)}/\mu(f^{-1}(E))$) t.q. Q p.s., $(Y_n, W_n) \to (Y, W)$.
L'égalité $\int_S f d\nu_n = \mathbf{0}$ entraîne $\mathbb{E}^Q(\theta_n.W_n e^{\theta_n.W_n}) = 0$.
Comme pour $w \leq 0$, $-we^w \leq 1/e$ on en déduit que $\sup_{n\geq N} \mathbb{E}^Q(|\theta_n.W_n| e^{\theta_n.W_n}) \leq 2/e$, puis que la suite de variables $(e^{\theta_n.W_n})_{n\geq N}$ est équiintégrable.
Donc pour $\varphi : S \to \mathbb{R}$ continue bornée,

$$\lim_{n\to+\infty} \mathbb{E}^Q(\varphi(Y_n)e^{\theta_n.W_n}) = \mathbb{E}^Q(\varphi(Y)e^{\theta_\infty.W}) = \int_S 1_E(f(x))\varphi(x)e^{\theta_\infty.f(x)}d\mu(x)/\mu(f^{-1}(E)). \tag{I.2.3}$$

Or pour $n \geq N$, $\dfrac{\sum_{i=1}^n 1_E(f(X_i))e^{\theta_*.f(X_i)}}{\sum_{i=1}^n 1_E(f(X_i))} = \dfrac{Z_n(\theta_*)}{\sum_{i=1}^n 1_E(f(X_i))} \geq \dfrac{Z_n(\theta_n)}{\sum_{i=1}^n 1_E(f(X_i))} = \mathbb{E}^Q(e^{\theta_n.Y_n})$.

En utilisant (I.2.1) et (I.2.3) avec $\varphi \equiv 1$ pour passer à la limite dans les membres extrêmes de cette inégalité, on obtient $Z(\theta_*) \geq Z(\theta_\infty)$.
Comme θ_* est l'unique point où Z atteint son minimum, $\theta_\infty = \theta_*$.
Pour $\varphi : S \to \mathbb{R}$ continue bornée, on conclut avec (I.2.3) que $\int_S \varphi d\nu_n = \mathbb{E}^Q(\varphi(Y_n)e^{\theta_n.W_n})/\mathbb{E}^Q(e^{\theta_n.W_n})$ $\to_{n\to+\infty} \int_S \varphi d\mu_E^{\theta_*}$.
Si $\varphi : S \to \mathbb{R}$ est mesurable bornée, on obtient que $\mathbb{P}$ p.s. $\int_S \varphi d\nu_n \to \int_S \varphi d\mu_E^{\theta_*}$, en raisonnant comme précédemment mais en remplaçant (I.2.2) par la convergence étroite de
$\sum_{i=1}^n 1_E(f(X_i))\delta_{(\varphi(X_i), f(X_i))}/\sum_{i=1}^n 1_E(f(X_i))$.

■

Au début de la preuve précédente, on a eu recours à un résultat intermédiaire qui repose sur le lemme suivant.

Lemme I.2.3 *Soit un polyèdre convexe Π de E, d'intérieur non vide et de sommets x_i, $1 \leq i \leq p$. Si*

$$\mathbf{0} \in Int_E(\Pi)$$

alors

$$\exists \eta > 0,\ \forall \epsilon \leq \eta,\ (\forall i,\ y_i \in \overline{B}(x_i, \epsilon) \bigcap E) \Rightarrow \mathbf{0} \in Int_E(Conv(y_i, 1 \leq i \leq p))$$

Preuve : On peut supposer ici $E = \mathbb{R}^d$, sans perte de généralité.
Si d=0, le lemme est immédiat.
Si $d \geq 1$, on choisit $\eta = \frac{1}{4}d(\mathbf{0}, Fr(\Pi)) > 0$ et $\epsilon \leq \eta$.
En introduisant le polyèdre Π' de sommets $(y_i, 1 \leq i \leq p)$ on a d'abord

$$\forall v \in \Pi',\ \exists u \in \Pi,\ \|u - v\| \leq \epsilon \tag{I.2.4}$$

En effet,

$$\exists(\alpha_i, 1 \le i \le p) \in \mathbb{R}_+^p,\ \sum_{i=1}^{p} \alpha_i = 1,\ \sum_{i=1}^{p} \alpha_i y_i = v$$

Soit

$$u = \sum_{i=1}^{p} \alpha_i x_i \in \Pi$$

Alors

$$\|u - v\| \le \sum_{i=1}^{p} \alpha_i \|y_i - x_i\| \le \epsilon$$

d'où l'affirmation.
De manière tout à fait symétrique, on a également :

$$\forall u \in \Pi,\ \exists v \in \Pi',\ \|u - v\| \le \epsilon \tag{I.2.5}$$

Montrons ensuite

$$\forall v \in Fr(\Pi'),\ d(v, Fr(\Pi)) \le \epsilon \tag{I.2.6}$$

Soit $v \in Fr(\Pi')$ et soit un hyperplan d'appui de Π' en v, d'équation :

$$\beta \cdot (x - v) = 0$$

avec β vecteur de $\mathbb{R}^d$ de norme 1 tel que

$$\forall x \in \Pi',\ \beta \cdot (x - v) \ge 0$$

Avec la relation (I.2.5), on voit alors que pour $\epsilon' > \epsilon$, Π est contenu dans le demi-espace défini par :

$$\beta \cdot (x - v) > -\epsilon'$$

et dont on note H l'hyperplan frontière.
Soit w la projection orthogonale de v sur H :

$$\overline{B}(v, \epsilon') \bigcap H = \{w\}$$

Alors $w \notin \Pi$.
On se donne $u \in \Pi$ vérifiant la relation (I.2.4) et donc a fortiori : $\|u - v\| \le \epsilon'$.
Ainsi :

$$[u, v] \subset \overline{B}(v, \epsilon')$$

Par connexité de Π, on a :

$$\exists z \in [u, w] \bigcap Fr(\Pi)$$

et l'on conclut

$$d(v, Fr(\Pi)) \le \epsilon'$$

On a donc prouvé (I.2.6) puisque ϵ' est arbitrairement proche de ϵ.
En appliquant (I.2.5), introduisons $A \in \Pi'$ tel que $d(\mathbf{0}, A) \le \epsilon$.
Alors, on a

$$\forall v \in Fr(\Pi'),\ \forall u \in Fr(\Pi),\ d(A, v) \ge d(A, Fr(\Pi)) - d(u, v)$$

La relation (I.2.6) donne

$$\begin{aligned} d(A, Fr(\Pi')) &\ge d(A, Fr(\Pi)) - \epsilon \\ &\ge d(\mathbf{0}, Fr(\Pi)) - d(\mathbf{0}, A) - \epsilon \\ &\ge d(\mathbf{0}, Fr(\Pi)) - 2\eta \ge \frac{d(\mathbf{0}, Fr(\Pi))}{2} \end{aligned} \tag{I.2.7}$$

d'où

$$B\left(A, \frac{d(\mathbf{0}, Fr(\Pi))}{2}\right) \subset \Pi'$$

Comme

$$B(\mathbf{0},\epsilon) \subset B(A, 2\epsilon) \subset B\left(A, \frac{d(\mathbf{0}, Fr(\Pi))}{2}\right)$$

on obtient la conclusion. ∎

Si l'on souhaite étendre la convergence à des fonctions de payoff non bornées, il faut imposer une condition d'intégrabilité supplémentaire.

Proposition I.2.4 *Supposons vérifiée la condition* (C).
Si ϕ *est une fonction réelle sur* S *telle que :*

$$\theta_* \in Int_E\left\{u \in E,\ \int 1_E(f(x)))e^{u\cdot f(x)}|\phi(x)|\mu(dx) < +\infty\right\}$$

alors $\mathbb{P}$*-presque sûrement*

$$\lim_{n\to\infty} \nu_n(\phi) = \mu_E^{\theta_*}(\phi)$$

Preuve : On a la majoration :

$$|\frac{\sum_{i=1}^n 1_E(f(X_i))\phi(X_i)e^{\theta_n\cdot f(X_i)}}{n} - \int 1_E(f(x))\phi(x)e^{\theta_*\cdot f(x)}\mu(dx)| \leq$$

$$|\frac{\sum_{i=1}^n 1_E(f(X_i))\phi(X_i)e^{\theta_*\cdot f(X_i)}}{n} - \int 1_E(f(x))\phi(x)e^{\theta_*\cdot f(x)}\mu(dx)|$$

$$+\frac{\sum_{i=1}^n 1_E(f(X_i))|\phi(X_i)||e^{\theta_n\cdot f(X_i)} - e^{\theta_*\cdot f(X_i)}|}{n} \tag{I.2.8}$$

La loi forte des grands nombres montre que le premier terme tend vers 0 $\mathbb{P}$-presque sûrement.
Si E est réduit à $\{\mathbf{0}\}$, on a $\forall n,\ \theta_n = \theta_* = \mathbf{0}$ et le deuxième terme est nul.
Sinon, pour pouvoir majorer le second terme, on considère $2^{d'}$ vecteurs ρ_l de E épuisant les relations

$$\rho_l = \theta_* + \epsilon\sum_{i=1}^{d'} \pm e_i$$

où $(e_i, 1 \leq i \leq d')$ est une base orthonormée de E, supposée de dimension $d' \geq 1$ et où $\epsilon > 0$ est assez petit pour satisfaire à :

$$\overline{B}(\theta_*, \sqrt{d'}\epsilon)\bigcap E \subset Int_E\{u \in E,\ \int 1_E(f(x)))e^{u\cdot f(x)}|\phi(x)|\mu(dx) < +\infty\}$$

On a alors la relation

$$\forall u \in \overline{B}(\theta_*,\epsilon)\bigcap E,\ \forall y \in E,\ e^{u\cdot y} \leq \sum_{i=1}^{2^{d'}} e^{\rho_i\cdot y}$$

et donc également, par l'inégalité des accroissements finis

$$\forall u \in \overline{B}(\theta_*, \epsilon) \bigcap E,\ \forall y \in E,\ |e^{u \cdot y} - e^{\theta_* \cdot y}| \leq \|u - \theta_*\| \|y\| \sum_{i=1}^{2^{d'}} e^{\rho_i \cdot y}$$

Ainsi, pour n assez grand, compte tenu de la convergence p.s. de θ_n vers θ_*, le deuxième terme de (I.2.8) est majoré par :

$$\|\theta_n(\omega) - \theta_*\| \sum_{j=1}^{2^{d'}} \frac{\sum_{i=1}^n 1_E(f(X_i)) |\phi(X_i)| \|f(X_i)\| e^{\rho_j \cdot f(X_i)}}{n}$$

dont la limite est nulle.
On a donc prouvé la convergence p.s.

$$\lim_{n \to \infty} \frac{\sum_{i=1}^n 1_E(f(X_i)) \phi(X_i) e^{\theta_n \cdot f(X_i)}}{n} = \int 1_E(f(x)) \phi(x) e^{\theta \cdot f(x)} \mu(dx)$$

et finalement, on prouve le résultat avec la limite p.s.

$$\lim_{n \to \infty} \frac{\sum_{i=1}^n 1_E(f(X_i)) \phi(X_i) e^{\theta_n \cdot f(X_i)}}{\sum_{i=1}^n 1_E(f(X_i)) e^{\theta_n \cdot f(X_i)}} = \frac{\int 1_E(f(x)) \phi(x) e^{\theta \cdot f(x)} \mu(dx)}{\int 1_E(f(x)) e^{\theta \cdot f(x)} \mu(dx}$$

puisque la convergence p.s. du dénominateur est établie avec le Théorème I.2.2. ∎

Remarque I.2.5 *Le cas où μ est la loi lognormale sur $\mathbb{R}$ (densité $1_{\{x>0\}} \exp(-\ln(x)^2/2)/x\sqrt{2\pi}$) et la contrainte est donnée par $f(x) = x - c$ avec $c > E(X_1) = \sqrt{e}$ fournit un exemple (inspiré du Call dans le modèle de Black-Scholes en finance) où la solution généralisée de (PL) qui est égale à μ, ne satisfait pas la contrainte.*

Pour contourner le problème soulevé dans cette dernière remarque, il est donc indiqué de calibrer par rapport à des Put plutôt qu'à des Call.
Il n'est cependant pas possible de se contenter de Put dès que l'on souhaite que la relation de parité Call-Put soit vérifiée. La perte potentielle de martingalité du sous-jacent actualisé sous la mesure μ_E^θ est un véritable problème auquel on peut répondre partiellement en ajoutant des contraintes de martingalité approchée. Si l'on ne borne pas ces contraintes, on voit qu'il est possible de rencontrer le phénomène de solution généralisée. L'étude de l'approximation de la martingalité sera donc approfondie ultérieurement dans le cas où le sous-jacent est borné.

Exemple numérique

Donnons à présent une illustration du résultat de convergence établi au Théorème I.2.2. On considère le modèle de Heston où l'actif unidimensionel $(S_t)_{t \geq 0}$ évolue avec une volatilité stochastique. Ce modèle incorpore le phénomène de retour à la moyenne que l'on cherche en général à capturer en finance. Dans ce cadre, on résout numériquement le système d'équations différentielles stochastiques

$$\begin{aligned} dS_t &= S_t(\mu dt + \sqrt{V_t} dW_t^S) \\ dV_t &= \alpha(m - V_t)dt + \sqrt{V_t}\left(\rho dW_t^S + \sqrt{1-\rho^2} dW_t^V\right) \end{aligned} \tag{I.2.9}$$

où W^S et W^V sont deux processus de Wiener indépendants. La loi a priori μ est une version discrétisée par le schéma d'Euler de ce modèle.

	K=87.5	K=100.0	K=112.5	K=125.0
T=0.5	**0.567** *0.515*	**3.517** *3.393*	**10.284** *10.159*	**19.982** *19.906*
T=1.0	**1.043** *0.967*	**3.928** *3.778*	**9.439** *9.248*	**17.228** *17.046*

FIG. I.1 – Valeurs des contraintes

Ici, on se fixe $\mu = 0.1$, $\alpha = 0.1$, $m = 0.1$, $\rho = 0.5$, $S_0 = 100$ et $\sqrt{V_0} = 0.2$. Le modèle a priori est de plus supposé être une probabilité risque neutre de référence, c'est-à-dire que le taux d'actualisation r vaut μ.
On calibre ensuite le modèle en se donnant des prix d'options. Afin de réduire le temps des simulations, on se limite à huit prix de Put artificiellement générés sous un modèle proche du modèle a priori. On a donc repris le modèle (I.2.9) avec des valeurs de paramètres identiques sauf celle de la volatilité initiale fixée à $\sqrt{V_0} = 0.205$. Les contraintes ont été obtenue par la méthode de Monte-Carlo. Remarquons que l'on aurait également pu utiliser une formule fermée [65].
La faible différence entre les deux volatilités initiales suffit à produire des écarts de prix appréciables. On peut le constater sur le tableau de la Figure I.1 où les contraintes sur les huit Put sont consignées en gras. Des prix calculés par Monte-Carlo sous la loi a priori pour 100 000 tirages apparaissent en italique et montrent la nécessité de la calibration.

Pour tester la méthode de calibration, on choisit un Put de strike $K = 105$ et de maturité $T = 1$. Sur la Figure I.2, on constate la convergence du prix calibré lorsque le nombre de tirages Monte-Carlo augmente. C'était l'illustration cherchée pour le résultat principal de ce chapitre.
Il est intéressant de comparer cette convergence avec celle obtenue sans la calibration. Cette comparaison est faite à la Figure I.3 et l'on constate que la calibration accélère la méthode de Monte-Carlo. Cela s'explique bien du fait de la présence des contraintes. On reviendra sur cette observation dans un chapitre futur où l'on analysera la vitesse de convergence du Monte-Carlo calibré.
A titre indicatif, le prix du Put testé vaut 5.813 (par Monte-Carlo) si la volatilité initiale est égale à 0.205. De manière générale, on peut espérer que les prix d'options calculés sous la probabilité minimisant l'entropie relative restent proches de ceux obtenus sous la mesure ayant servi à générer les contraintes.

I.3 Remarque sur l'entropie relative en tant que mesure de l'information et sur le principe du maximum d'entropie

L'article [31] est l'un des premiers à avoir suggéré l'utilisation de l'entropie relative pour calibrer les modèles financiers tandis que [6] place le problème dans le cadre de la méthode de Monte-Carlo. Notre article [71] donne les résultats de convergence que l'on a développés dans ce chapitre. Par ailleurs, les auteurs de [56] étudient un problème de convergence d'estimateurs empiriques mettant en jeu le problème de la mauvaise spécification du facteur d'actualisation. L'entropie relative y intervient car c´est le point de vue des grandes déviations qui est adopté.
L'entropie relative est utilisée par de nombreux autres auteurs pour les problèmes de prix d'options, en citant par exemple [52], [122], [103]. C'est à l'étude des marchés incomplets que l'on doit son apparition en finance [51] tandis que son introduction en économie est discutée dans [85], bien après évidemment que cette notion ne soit couramment utilisée en physique statistique.
Dans la critique qu'il faisait au sujet d'un ouvrage relatif à ce dernier domaine, E.T. Jaynes

définissait les deux aspects de la théorie permettant d'une part de construire un ensemble d'états à partir de l'information connue, et, réciproquement, de déduire cette information à partir de la connaissance d'un ensemble d'états en déterminant la probabilité d'entropie maximale. C'est cet aspect que Jaynes considérait comme l'un des plus beaux dans la théorie de l'information qu'il a lui même contribué à introduire en physique dans [69] grâce aux travaux de Shannon [117] donnant ainsi le cadre moderne de la théorie de Gibbs. L'impact de ces travaux en sciences physiques a été également profondément discuté dans l'ouvrage classique [30].
L'un des succès les mieux établis de l'entropie relative comme mesure de l'information réside dans le théorème fondamental de la théorie de l'information qui donne la borne supérieure pour le taux de transmission de l'information entre un émetteur et un récepteur. Cette borne supérieure est la capacité du canal définie comme la valeur maximale de l'information mutuelle (construite à partir de l'entropie relative) entre l'ensemble des symboles émis servant à coder les messages et ceux qui sont reçus. En l'absence de bruit par exemple, la capacité est simplement l'entropie maximale de la source. Pour une référence sur le sujet, on peut consulter [102] par exemple.
Dans l'article fondateur [117], la forme de l'entropie de Shannon est mathématiquement déduite à partir de propriétés simples dans le cas d'un espace probabilisé fini et en l'absence d'information a priori (les atomes sont a priori équiprobables). Le passage au cas des probabilités continues a été à l'origine d'une erreur de Shannon et Jaynes y fait brièvement référence dans son cours élémentaire de physique statistique [70]. L'entropie relative y est introduite comme la limite naturelle de l'entropie de Shannon lorsque l'on passe des probabilités discrètes aux probabilités à densités grâce à un procédé convenable de discrétisation des états, à condition de se donner une loi a priori. Jaynes ne précise pas la façon d'obtenir cette discrétisation mais il est parfaitement légitime de la construire comme un échantillon de la loi a priori. Il restait alors à démontrer que le principe du maximum d'entropie était préservé à la limite et c'est ce qui constitue le résultat principal de ce chapitre.
La référence [60] offre un panorama très riche des applications du principe du maximum d'entropie. Pour approfondir davantage son lien avec la méthode de Monte-Carlo, on peut citer des résultats concernant les échantillons de variables indépendantes et identiquement distribuées. On sait en effet que conditionnellement à une contrainte de moment, la distribution empirique s'obtient en vertu du principe du maximum d'entropie ([126], [125]), résultat généralisé dans [38] à une contrainte convexe et mis en liaison avec la propriété de Sanov. L'entropie relative y intervient classiquement en tant que fonction de taux de grandes déviations ([115], [119], [8]). On a déjà cité [56] pour un exemple d'application à la finance.
Pour voir qu'à la limite des grands tirages les distributions empiriques vérifiant les contraintes se concentrent autour de la distribution obtenue par le principe du maximum d'entropie, une façon intuitive est d'utiliser la méthode du maximum de vraisemblance.
Afin de préciser cela, quitte à opérer une discrétisation de l'espace S, on fait l'hypothèse que ce dernier est fini : $S = \{x_i,\ 1 \leq i \leq M\}$ et que de plus la mesure a priori μ attribue à tout x_i de S une probabilité non nulle. Alors la probabilité empirique s'écrit pour n fixé :

$$\mu_n = \sum_{i=1}^{M} \frac{N_{i,n}}{n} \delta_{x_i}$$

en posant $N_{i,n} = \sum_{j=1}^{n} 1_{X_j = x_i}$.
Pour des entiers $n_1, ..., n_M$ de somme n, si on définit la mesure :

$$\nu = \sum_{i=1}^{M} \frac{n_i}{n} \delta_{x_i}$$

on a :

$$\mathbb{P}(\mu_n = \nu) = \frac{n!}{n_1!...n_M!} \mu(x_1)^{n_1} ... \mu(x_M)^{n_M}$$

Lorsque chacun des n_i est grand, la formule de Stirling donne :

$$\mathbb{P}(\mu_n = \nu) \sim \frac{1}{\sqrt{2\pi}^{M-1}} e^{-\sum_{i=1}^{M} n_i \ln(\frac{n_i}{n\mu(x_i)}) + \frac{1}{2}(\ln(n) - \sum_{i=1}^{M} \ln(n_i))} \tag{I.3.1}$$

Suivant la méthode du maximum de vraisemblance, on rend maximale l'expression ci-dessus pour ν définie par les familles d'entiers $(n_i)_{1 \leq i \leq M}$ de somme n et vérifiant la contrainte $\| \sum_{i=1}^{M} \frac{n_i}{n} f(x_i) \| \leq \epsilon$, ϵ étant un niveau d'erreur fixé compte tenu qu'ici la contrainte ne peut être satisfaite en général. En négligeant $\ln(n_i)$ par rapport à $n_i \ln(n_i)$ dans l'expression (I.3.1), on est alors conduit à minimiser sous contrainte

$$\sum_{i=1}^{M} n_i \ln \left(\frac{n_i}{n\mu(x_i)} \right) = nH(\nu \| \mu)$$

Adapté du chapitre XVIII de [100], le développement élémentaire ci-dessus, proche du raisonnement original de Boltzmann, donne ainsi un des fondements du principe du maximum d'entropie.

Chapitre II

Extension des résultats à des critères de minimisation plus généraux

Le problème de calibration étudié dans la section précédente est un problème dit de programmation partiellement finie, c'est-à-dire que l'on cherche à optimiser une fonctionnelle convexe sur un espace de dimension infinie (ou bien grande pour le cas discret) sous un nombre fini de contraintes linéaires.
Ce type de problème surgit dans des domaines variés et une discussion importante dans la littérature concerne le choix de la fonctionnelle convexe et l'on se reportera notamment à l'introduction de [26] pour une présentation concise du sujet.
S'il est justifié d'employer l'entropie relative en tant que mesure de l'information afin d'estimer une densité de probabilité à partir de ses moments, comme au chapitre précédent, d'autres critères ont également été proposés.
Ainsi qu'il est discuté dans [11] par exemple, la généralisation naturelle de l'entropie de Shannon conduit à la classe des f-entropies, où f est une fonction concave définie en adoptant l'opposée d'un critère convexe J tel que $J(u) = u \ln u$ dans le cas de l'entropie de Shannon, $J(u) = u(1-u)$ dans celui de l'entropie quadratique ou encore $J(u) = -\frac{u^\alpha - u}{2^{1-\alpha}-1}$ avec $\alpha \neq 1$ pour définir l'entropie de Darocry et l'on peut également citer l'entropie de Burg définie par le choix $J(u) = -\ln(u)$. La possibilité de modifier l'entropie relative en tant que mesure de l'information est également abordée dans [88] et [37] donne un point de vue axiomatique du sujet.
Comme autre exemple, la méthode du maximum d'entropie sur la moyenne introduite dans [39] et [54] est une extension de la méthode du maximum d'entropie qui conduit à considérer la famille des fonctions convexes définies comme transformées de Cramer (fonction conjuguée de la transformée de log-Laplace) des mesures de probabilité sur $\mathbb{R}_+$.
D'un point de vue financier, dans le cadre des marchés incomplets, il est désormais usuel de déterminer une probabilité martingale par minimisation d'un critère convexe. Récemment [10] a donné une interprétation économique précise au choix du critère de minimisation. En effet, sous des conditions générales la probabilité martingale minimale est la mesure minimax pour une fonction d'utilité de type HARA. L'intérêt d'une telle fonction est bien connu et est exposé au chapitre IV de [127] par exemple. Il y a donc des raisons financières à valoriser les options sous la probabilité minimisant une pseudo-distance, ce qu'exprime précisément le Lemme 3.4 de [74] pour les cas courants d'utilité. Le choix de l'entropie relative correspond ainsi classiquement à celui d'une utilité exponentielle [42] [46]. Si l'on souhaite cependant une couverture moyenne-variance, il faut choisir le critère de minimisation quadratique. Dans le cadre d'un modèle particulier, la comparaison de ces critères est faite dans [64]. On y envisage la famille des critères usuels qui,

en plus de l'entropie relative comprend l'entropie de Burg et, pour $p \in \mathbb{R}\backslash\{0,1\}$,

$$J(u) = \begin{cases} \frac{p}{p-1}u^p & \text{si } u \geq 0 \\ +\infty & \text{sinon} \end{cases}$$

avec $0^p = +\infty$ si $p < 0$.
En revenant à la calibration par Monte-Carlo, ainsi qu'il est indiqué dans [6], il est donc légitime d'étendre la méthode de minimisation d'entropie relative à des critères convexes plus généraux. Il faut souligner cependant qu'en minimisant le critère sur l'ensemble de toutes les probabilités et non sur celles qui sont martingales, on sort du cadre de [10] et [74]. Une façon de rattraper cet écart est d'ajouter des contraintes de martingalité, ce qui sera étudié avec plus de précision dans un chapitre futur.
On commencera à choisir le critère de manière à ce que son comportement reste assez proche de celui de l'entropie relative. Les résultats reposent alors essentiellement sur des techniques fondamentales de dualité dont l'intérêt est de résoudre le problème de calibration en restreignant le choix de la solution à une famille paramétrée de densités, le paramètre étant déterminé en résolvant le problème dual qui est un problème de minimisation simple. Une référence classique sur le sujet est [23] à condition de borner les contraintes par exemple (voir aussi [12]).

Dans la suite, on adopte comme critère une fonction réelle J strictement convexe sur son domaine $[0,+\infty[$, dérivable sur $]0,+\infty[$ et de dérivée vérifiant $J'(]0,+\infty[) = \mathbb{R}$. On peut alors poser pour toute probabilité ν

$$H_J(\nu\|\mu) = \begin{cases} \int J(\frac{d\nu}{d\mu})d\mu & \text{si } \nu \text{ est absolument continue par rapport à } \mu \\ +\infty & \text{sinon} \end{cases}$$

Précisons ici la nature des hypothèses faites sur J et introduisons la fonction conjuguée J^* définie par

$$\forall u \in \mathbb{R},\ J^*(u) = \sup_{x\in\mathbb{R}}(xu - J(x)) \tag{II.0.1}$$

On verra plus loin que les conditions imposées à J sont nécessaires et suffisantes pour que J^* soit finie, strictement convexe, croissante, dérivable et minorée sur $\mathbb{R}$ et qu'elle vérifie de plus la condition :

$$\lim_{u\to+\infty} J^{*\prime}(u) = +\infty \tag{II.0.2}$$

En particulier, on peut préciser la minoration de J^* par :

$$\forall u \in \mathbb{R},\ J^*(u) \geq -J(0) \tag{II.0.3}$$

En anticipant sur la suite, il faudra d'abord attaquer un problème de minimisation "dual" associé à J^* : le fait que J^* soit strictement convexe, minorée et qu'elle vérifie (II.0.2) permet d'assurer l'unicité et l'existence du minimum. Enfin, au moins dans le cas de contraintes bornées, la résolution du problème dual fournit la solution du problème de calibration si J^* est croissante et dérivable sur $\mathbb{R}$.
Après cette précision sur le choix du critère J, donnons la forme sous laquelle interviendra la dualité. Si l'on commence par supposer que J est s.c.i., ce qui revient à dire ici que J est continue à droite en 0, alors grâce au théorème de Fenchel-Moreau (Théorème 1.10 p.10 de [29] par exemple), J est la fonction conjuguée de J^*. Alors, l'inégalité

$$\forall x \geq 0,\ \forall u \in \mathbb{R},\ J(x) \geq xu - J^*(u) \tag{II.0.4}$$

est réalisée avec égalité si et seulement si :

$$x = J^{*\prime}(u)$$

C'est essentiellement de cette manière que l'on exploitera la relation de conjugaison entre J et J^*. Remarquons que la valeur de J en 0 n'intervient pas, ce qui fait qu'on ne perd pas en généralité en supposant que J est continue en 0 et donc que J est sci.
Avec les propriétés indiquées, J est très proche de l'entropie de Shannon. Comme au chapitre précédent, les résultats de convergence sont alors valables, au risque de faire apparaître une solution généralisée à la limite. C'est également dans ce cadre qu'il a été agréable d'examiner le problème de calibration où l'on relâche les contraintes, soit par une pénalisation du critère de minimisation, soit en considérant des contraintes de type ensemble. Un avantage théorique à choisir J proche de l'entropie relative est que cela permet de traiter des contraintes non bornées (sous la probabilité a priori μ).
Toutefois, lorsque l'on se contente de contraintes bornées, il est possible de prendre des alternatives classiques à l'entropie de Shannon comme l'entropie quadratique ou l'entropie de Burg par exemple, exclues avec les hypothèses présentes. On a déjà remarqué que l'utilisation de tels critères présentait un intérêt en finance. On envisagera donc les extensions qui s'imposent dans la dernière partie du chapitre.

II.1 Discussion des hypothèses sur le critère J

On se propose d'abord de caractériser le critère J de façon à assurer les propriétés sur sa conjuguée J^* précisées ci dessus. On retrouvera le jeu d'hypothèses énoncé ci dessus, sous lequel J garde une forme proche du choix entropique.
Il est tout d'abord possible de reformuler les hypothèses sur J^*. En effet, J^* est de classe $\mathcal{C}^1$ puisque sa dérivée est croissante et vérifie, comme toute dérivée, la propriété des valeurs intermédiaires. L'ensemble des conditions sur J^* équivaut donc à se donner une fonction J^* minorée et dérivable sur $\mathbb{R}$, de dérivée $J^{*\prime}$ strictement croissante et telle que

$$J^{*\prime}(\mathbb{R}) =]0, +\infty[\tag{II.1.1}$$

Rappelons ici que, sans sacrifier à la généralité, on a supposé que J est sci, si bien que J est la fonction convexe conjuguée de J^*. On va alors montrer que J est finie et strictement convexe sur $[0, +\infty[$, dérivable sur $]0, +\infty[$, avec :

$$J'(]0, +\infty[) = \mathbb{R}$$

On peut prouver cela à l'aide de résultats généraux exposés dans [106]. On va en effet montrer que la fonction convexe J^* est du "type de Legendre" introduit dans [106] (p.258) qui définit ainsi la classe des fonctions convexes $K : \mathbb{R}^m \to]-\infty, +\infty]$, $m \in \mathbb{N}$, strictement convexes sur l'intérieur de leur domaine $domK = \{x \in \mathbb{R}^m,\ K(x) < +\infty\}$ et "essentiellement régulières" (p.251 de [106]) au sens :

1. $Int(domK) \neq \emptyset$
2. K est différentiable sur $Int(domK)$
3. $\forall x \in Fr(domK),\ \lim_{y \in domK \to x} \|K'(y)\| = +\infty$

D'abord, J^* étant finie et dérivable sur $\mathbb{R}$, elle est essentiellement régulière. Elle est également strictement convexe par hypothèse, ce qui montre que J^* est du type de Legendre. C'est aussi une fonction fermée, ce qui revient à dire (v. p. 52 de [106]) qu'elle est sci puisqu'elle est propre, en rappelant qu'une fonction convexe est propre si elle n'admet pas $-\infty$ pour valeur et qu'elle admet des valeurs finies.
Avec le Théorème 26.5 p.258 de [106], on en déduit que la conjuguée J de J^* est également du type de Legendre, donc en particulier strictement convexe et dérivable sur l'intérieur, non

vide, de son domaine $domJ$, et que de plus $J^{*\prime}$ réalise une bijection de $Int(domJ^*) = \mathbb{R}$ sur $Int(domJ)$, avec J' pour bijection réciproque.
Avec la relation (II.1.1), on obtient :

$$Int(domJ) =]0, +\infty[$$

Puisque J est la fonction convexe conjuguée de J^* et puisque J^* est minorée, on a également :

$$J(0) = -\inf_{u\in\mathbb{R}} J^*(u) < +\infty$$

et, par conséquent :

$$domJ = [0, +\infty[$$

Réciproquement, si l'on part des hypothèses que J est strictement convexe sur son domaine $[0, +\infty[$, dérivable sur $]0, +\infty[$, de dérivée vérifiant

$$J'(]0, +\infty[) = \mathbb{R} \tag{II.1.2}$$

on peut en déduire que sa conjuguée J^* est minorée, strictement convexe et dérivable sur $\mathbb{R}$, de dérivée $J^{*\prime}$ strictement croissante et telle que

$$J^{*\prime}(\mathbb{R}) =]0, +\infty[$$

En effet, avec la relation (II.1.2), la conjuguée J^* est définie comme une transformée de Legendre par

$$\forall u \in \mathbb{R},\ J^*(u) = uJ'^{-1}(u) - J(J'^{-1}(u)) \tag{II.1.3}$$

ce qui prouve que la valeur de J en 0 ne joue pas de rôle, si bien qu'on peut supposer J continue en 0 et donc sci.
On peut alors montrer que J est du type de Legendre pour appliquer une nouvelle fois le Théorème 26.5 p.258 de [106]. En effet, J est dérivable sur $Int(domf) =]0, +\infty[$, avec $\lim_{x\to 0} J'(x) = +\infty$, ce qui prouve que J est essentiellement régulière, au sens précisé plus haut et, puisque J est strictement convexe sur $]0, +\infty[$, elle est bien du type de Legendre. Enfin, J étant fermée car propre et sci, le Théorème 26.5 de [106] montre que sa conjuguée J^* est strictement convexe et dérivable sur l'intérieur de son domaine $domJ^*$, et que de plus J' réalise une bijection de $Int(domJ) =]0, +\infty[$ sur $Int(domJ^*)$, avec $J^{*'}$ pour bijection réciproque.
Avec la relation (II.1.2), on obtient :

$$Int(domJ^*) = \mathbb{R}$$

Enfin, J^* étant la conjuguée de J, on a :

$$\forall u \in \mathbb{R},\ J^*(u) \geq -J(0) > +\infty$$

ce qui prouve que J^* est minorée.

II.2 Problème contraint

En se donnant un critère convexe du type discuté plus haut, on cherche à satisfaire les contraintes de manière exacte. Pour n tirages Monte-Carlo, le problème discret consiste donc à résoudre, en se ramenant à une contrainte $\mathbf{C}$ nulle :

$$(P_n^J) \quad \text{trouver } \nu_n \in \mathcal{P}(S) \text{ qui minimise } H_J(\nu_n\|\mu_n) \text{ sous la contrainte } \int f d\nu_n = \mathbf{0}$$

Reprenons ici un commentaire fait au chapitre précédent ainsi qu'au début de celui-ci : pour que les prix donnés par cette méthode aient une certaine validité, il est important que parmi les contraintes on impose des conditions de martingalité approchée du sous-jacent actualisé.
On reformule le problème précédent comme la recherche d'un vecteur $(p_i)_{1\leq i\leq n}$ de $[0,1]^n$ tel que $\sum_{i=1}^n p_i = 1$ et $\sum_{i=1}^n f(X_i)p_i = \mathbf{0}$, minimisant

$$\sum_{i=1}^{n} J(np_i)$$

Il s'agit donc de trouver le minimum d'une fonction strictement convexe sur un compact. L'existence de la solution, forcément unique, en résulte dès que l'on prouve la non vacuité de ce compact puisque J est de plus sci. On rappelle qu'on l'a supposé en particulier continue à droite en 0 mais cette dernière condition est en fait inutile. En effet, J a une dérivée à droite infinie en 0, excluant que la solution ait une composante p_i nulle.
Tout d'abord, le résultat suivant, établi dans le cas de l'entropie, tient toujours :

Lemme II.2.1 *Pour que* $\mathbb{P}(\exists n \geq 1$ *et* $\eta_n \in \mathcal{P}(S)$ *t.q.* $H_J(\eta_n\|\mu_n) < +\infty$ *et* $\int f d\eta_n = \mathbf{0}) > 0$, *il faut qu'il existe un sous-espace vectoriel* E *de* $\mathbb{R}^d$ *tel que tienne la condition*

$$(C)\left\{\begin{array}{l} \mathbf{0} \in Int_E(Conv(Supp(\mu^f_{|E}))) \\ \forall \eta \in \mathcal{P}(S) \textit{ absolument continue par rapport à } \mu \textit{ t.q. } \int \|f\| d\eta < +\infty \textit{ et } \int f d\eta = \mathbf{0}, \\ \textit{on a } \eta(f^{-1}(E)) = 1 \end{array}\right.$$

La convergence passe alors naturellement par l'étude du problème limite

$$(PL^J) \quad \text{trouver } \nu \in \mathcal{P}(S) \text{ qui minimise } H_J(\nu\|\mu) \text{ sous la contrainte } \int f d\nu = \mathbf{0}$$

Par stricte convexité de J, l'unicité de la solution est acquise et seul le problème de l'existence se pose.
La condition (C) est également nécessaire pour que le problème limite ait une solution, comme on peut le voir en reprenant le Lemme II.2.1.
Pour la suite, il sera commode de noter l'ensemble

$$\mathcal{C} = \left\{\nu \in \mathcal{P}(S),\ \nu \ll \mu,\ \int \|f\| d\nu < +\infty \text{ et } \int f d\nu = \mathbf{0}\right\} \tag{II.2.1}$$

Dans le cas où le vecteur de payoff f est supposé borné par exemple, on connaît des résultats permettant de résoudre le problème limite, reposant essentiellement sur l'absence de saut de dualité dès qu'une condition dite de qualification des contraintes est vérifiée (Corollaire 2.6 de [23]).
Dans ce cas en effet, grâce au Théorème 1 de [123], la solution du problème de calibration existe puis, en appliquant le Théorème 4.8 de [23]), on en déduit que la solution est donnée en résolvant le problème dual.
A titre de compléments, on reviendra plus précisément sur l'utilisation de ces résultats dans la dernière partie de ce chapitre.
Un premier objectif de l´étude qui suit est en fait d'étendre le cadre des résultats cités à celui d'une fonction de payoff non bornée, en montrant en particulier qu'il n'y a pas de saut de dualité.
Alors, si le problème de calibration admet une solution, cette dernière s'obtient grâce à celle du problème dual tandis que de manière générale, on peut définir une solution généralisée du problème de calibration dans le sens où elle a vocation à être la limite de la solution du problème de calibration discret, ce qui constitue le deuxième et principal objectif de l'étude.

L'approche de la solution d'un problème de moments par celle d'une suite de problèmes discrets n'est pas nouvelle, et la conclusion du chapitre dernier en donne un exemple classique.
Dans la littérature récente, on peut ainsi citer [39] où les auteurs discrétisent le problème de minimisation d'entropie afin de pouvoir traiter des contraintes compliquées tandis que dans [54], cette étude est reprise de manière systématique avec pour objet de montrer la convergence d'estimateurs bayésiens. Dans ces deux références, la discrétisation est cependant déterministe, et il en est donc aussi de la suite des solutions discrètes minimisant l'entropie relative, si bien que leurs problématiques ne relèvent pas des techniques Monte-Carlo.
La convergence Monte-Carlo des estimateurs minimisant l'entropie relative, ou un critère convexe plus général, est à rattacher aux résultats de la théorie classique des M-estimateurs présentée par exemple dans l'ouvrage général [19].
On consacrera donc une courte partie à discuter dans quelle mesure on peut utiliser ces résultats connus pour établir la convergence souhaitée.

II.2.1 Etude du problème limite sous la condition (C)

Une première remarque concernant la solution est le cas particulier suivant, montrant que le critère est cohérent avec le choix de la mesure a priori. :

Lemme II.2.2 *Si μ est un élément de $\mathcal{C}$, alors c'est la solution du problème (PL^J).*

Preuve : Il suffit d'écrire pour toute probabilité $\nu \in \mathcal{P}(S)$ absolument continue par rapport à μ :

$$J\left(\frac{d\nu}{d\mu}\right) - J(1) \geq J'(1)\left(\frac{d\nu}{d\mu} - 1\right)$$

d'où,

$$\int J\left(\frac{d\nu}{d\mu}\right) d\mu - J(1) \geq 0$$

ce qui prouve le résultat. ■

Il se peut que le problème (PL^J) n'admette pas de solution, et, dans le cas particulier de l'entropie, on a vu que cela conduit à l'accepter en un sens généralisé.
De façon classique grâce à la technique des multiplicateurs de Lagrange, puisque J^* est la conjuguée de J, l'idée est d'attaquer le problème de minimisation du critère dual défini comme la fonction à valeurs dans $]-\infty, +\infty]$

$$(u, u_0) \in E \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0) 1_E(f) d\mu - u_0 \tag{II.2.2}$$

Typiquement, l'hypothèse (C) est une condition de "qualification" des contraintes qui assure l'existence du minimum du critère dual. De façon plus précise, on va reprendre des arguments utilisés dans le cas de l'entropie relative.
On rappelle d'abord que E est un espace vectoriel puisque l'on s'est ramené à une contrainte nulle.
L'intégrale posée dans (II.2.2) est parfaitement définie puisque d'une part J^* est une fonction convexe donc mesurable et, d'autre part elle est minorée par hypothèse. Avec la stricte convexité de J^* et la condition (C), la fonction définie en (II.2.2) est strictement convexe et le minimum de cette fonction est donc unique.
Pour l'existence du minimum, on voit d'une part que la fonction à minimiser est sci grâce au lemme de Fatou, valable puisque J^* est minorée par hypothèse. Dès que J^* croît assez vite (plus

que linéairement de façon précise), et sous la condition (C), on peut d'autre part montrer que la fonction

$$(u,u_0)\in E\times\mathbb{R}\longmapsto\int J^*(u\cdot f+u_0)1_E(f)d\mu-u_0$$

tend vers $+\infty$ en l'infini. C'est une conséquence du Lemme II.2.10 en prenant $\eta_n=\mu$ pour tout n, et c'est ici qu'intervient la condition (II.0.2).
Le lemme est renvoyé avec sa démonstration pour l'étude de la convergence du problème discret dans laquelle il joue un rôle fondamental. La preuve de l'existence du minimum est alors achevée et l'on peut énoncer :

Proposition II.2.3 *Sous l'hypothèse (C), la fonction*

$$(u,u_0)\in E\times\mathbb{R}\longmapsto\int J^*(u\cdot f+u_0)1_E(f)d\mu-u_0$$

atteint son minimum en un point unique (θ,θ_0).

En cas d'existence de la solution du problème primal, la technique de dualité permet de l'identifier. En effet, si l'on introduit la distance

$$H_J(\mu\|\mathcal{C})=\inf_{\nu\in\mathcal{C}}H_J(\nu\|\mu)\tag{II.2.3}$$

où l'ensemble $\mathcal{C}$ est défini par (II.2.1), on peut la relier à la valeur minimale associée au problème dual par :

$$H_J(\mu\|\mathcal{C})=-\left(\int J^*(\theta\cdot f+\theta_0)1_E(f)d\mu-\theta_0\right)+J(0)\mu(f^{-1}(E^c))\tag{II.2.4}$$

Ceci se prouve par le procédé d'approximation habituel de Jupp & Mardia [72] dont on peut montrer la convergence bien que l'on ne dispose pas forcément d'une solution généralisée.

Proposition II.2.4 *Supposons vérifiée la condition (C).*
La valeur minimale associée au problème (PL^J) est alors donnée par la relation (II.2.4) et si le problème (PL^J) admet une solution, celle ci est

$$\mu_{E,J}^{(\theta,\theta_0)}\equiv J^{*\prime}(\theta\cdot f+\theta_0)1_E(f)\mu\tag{II.2.5}$$

(Si $E=\mathbb{R}^d$, on notera plus simplement $\mu_J^{(\theta,\theta_0)}$ au lieu de $\mu_{\mathbb{R}^d,J}^{(\theta,\theta_0)}$)
Réciproquement, si $\mu_{E,J}^{(u,u_0)}$ est un élément de $\mathcal{C}$, avec $(u,u_0)\in E\times\mathbb{R}$, alors c'est la solution du problème (PL^J) et l'on a $(u,u_0)=(\theta,\theta_0)$.

Preuve : Commençons par la partie réciproque de la proposition en posant :

$$D(u,u_0)=-\left(\int J^*(u\cdot f+u_0)1_E(f)d\mu-u_0\right)+J(0)\mu(f^{-1}(E^c))$$

Si ν est un élément de $\mathcal{C}$, alors puisque le support de l'image par f de ν est forcément dans E, on a :

$$H_J(\nu\|\mu)-D(u,u_0)=\int\underbrace{\left(J\left(\frac{d\nu}{d\mu}\right)+J^*(u\cdot f+u_0)\right)}_{\geq(u\cdot f+u_0)\frac{d\nu}{d\mu}}1_E(f)d\mu-u_0\geq 0\tag{II.2.6}$$

L'inégalité vient de la relation (II.0.4) dans laquelle on a le cas d'égalité si et seulement si $\nu=\mu_{E,J}^{(u,u_0)}$.

On en déduit d'abord que $\mu_{E,J}^{(u,u_0)}$ est la solution (PL^J) si c'est un élément de $\mathcal{C}$ puis, en admettant la relation (II.2.4), on obtient l'égalité $D(u,u_0) = D(\theta,\theta_0)$ d'où $(u,u_0) = (\theta,\theta_0)$ par unicité du minimum.
Pour la partie directe de la proposition, admettons aussi la relation (II.2.4) qui fait l'objet du Lemme II.2.13 et que l'on reporte dans l'étude du problème discret.
Cela se traduit par l'égalité $H_J(\mu\|\mathcal{C}) = D(\theta,\theta_0)$, et donc, si ν est un élément de $\mathcal{C}$, la relation (II.2.6) donne, en remplaçant (u,u_0) par (θ,θ_0)

$$H_J(\nu\|\mu) - H_J(\mu\|\mathcal{C}) \geq 0 \tag{II.2.7}$$

avec égalité si et seulement si $\nu = \mu_{E,J}^{(\theta,\theta_0)}$, ce qui prouve la proposition. ∎

Remarque II.2.5 *1. Dans le cas où f est bornée, l'absence de saut de dualité traduite par la relation (II.2.4) ainsi que l'existence de la solution du problème dual s'obtiennent avec le Corollaire 4.8 de [24], sous des hypothèses plus faibles sur J. On affirme ici que ces résultats se maintiennent lorsque le payoff f n'est pas bornée, à condition de spécifier certaines hypothèses sur J.*

2. *On verra dans la suite que $\mu_{E,J}^{(\theta,\theta_0)}$ est une probabilité si bien que pour montrer que c'est la solution usuelle de (PL^J), il suffit donc de vérifier que $\int \|f\| d\mu_{E,J}^{(\theta,\theta_0)} < +\infty$ et $\int f d\mu_{E,J}^{(\theta,\theta_0)} = \mathbf{0}$.*
3. *Si $\mu_{E,J}^{(u,u_0)}$ est une probabilité telle que $\int \|f\| d\mu_{E,J}^{(u,u_0)} < +\infty$, alors pour $\nu = \mu_{E,J}^{(u,u_0)}$ l'inégalité tient avec égalité dans le membre de droite de (II.2.6) et ce dernier est alors fini avec l'hypothèse d'intégrabilité de f, sans préjuger de sa nullité.*
Le membre de gauche de (II.2.6) est donc fini et comme J^* est minoré on en déduit :

$$H_J(\mu_{E,J}^{(u,u_0)}\|\mu) < +\infty$$

En particulier, cette relation est vraie pour la solution généralisée $\mu_{E,J}^{(\theta,\theta_0)}$ sous la seule condition $\int \|f\| d\mu_{E,J}^{(\theta,\theta_0)} < +\infty$.
C'est cependant un résultat plus faible que celui obtenu dans le cas de l'entropie relative où l'on a de manière inconditionnelle $H(\mu_E^{\theta,\theta_0}\|\mu) \leq H(\mu\|\mathcal{C})$
4. *On peut remarquer enfin que μ appartient à la famille $(\mu_J^{\mathbf{0},u_0})_{u_0\in\mathbb{R}}$.*
En effet, avec l'hypothèse (II.0.2) en particulier, la fonction $u_0 \longmapsto J^*(u_0) - u_0$ admet un minimum et donc :

$$\exists u_0 \in \mathbb{R},\ J^{*\prime}(u_0) = 1$$

ce qui prouve le résultat.
La partie réciproque de la proposition précédente permet alors de retrouver le Lemme II.2.2.

Si f est bornée, la dualité entre les problèmes (PL^J) et celui de minimisation du critère dual est établie de façon générale en notant par exemple avec le Corollaire du Théorème 2 de [104], que les fonctions $g \in L^\infty(S) \mapsto \int J^*(g) 1_E(f) d\mu$ et $\phi \in L^1(S) \mapsto \int J(\phi) 1_E(f) d\mu$ sont conjuguées l'une de l'autre dès que J est également sci.
Avec les hypothèses de régularité faites sur J, ce résultat peut se retrouver et, de plus, la solution du problème dual fournit effectivement la solution du problème primal.

Proposition II.2.6 *Dans le cas où f est bornée, et sous la condition (C), le problème (PL^J) admet pour solution $\mu_{E,J}^{(\theta,\theta_0)}$.*

Preuve : J^* étant supposée croissante, sa dérivée est positive et $\mu_J^{(\theta,\theta_0)}$ est donc une mesure positive.
D'autre part, $J^{*\prime}$ étant elle même croissante puisque J^* est convexe, le théorème de dérivation s'applique car f est bornée et l'on conclut que la fonction $(u, u_0) \in E \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0)1_E(f)d\mu - u_0$ est partout finie et de classe $\mathcal{C}^1$.
La condition d'Euler au minimum (θ, θ_0) s'écrit donc :

$$\int d\mu_{E,J}^{(\theta,\theta_0)} = 1$$

et

$$\int f d\mu_{E,J}^{(\theta,\theta_0)} = \mathbf{0}$$

la mesure $\mu_{E,J}^{(\theta,\theta_0)}$ est donc une probabilité qui vérifie les contraintes.
On a ainsi prouvé que $\mu_{E,J}^{(\theta,\theta_0)}$ est un élément de l'ensemble $\mathcal{C}$ intoduit par (II.2.1) et l'on peut conclure grâce à la proposition précédente. ∎

Remarque II.2.7 *La relation (II.2.4) , admise pour l'instant, peut être prouvée dans le cas où f est bornée. D'abord, en appliquant la relation (II.2.6) à (θ, θ_0) on a en toute généralité l'inégalité, valable pour tout $\nu \in \mathcal{C}$*

$$H_J(\nu \| \mu) \geq D(\theta, \theta_0) = -\left(\int J^*(\theta \cdot f + \theta_0)1_E(f)d\mu - \theta_0 \right) + J(0)\mu(f^{-1}(E^c)) \qquad \text{(II.2.8)}$$

Il suffit ensuite de remarquer que l'appartenance de $\mu_{E,J}^{(\theta,\theta_0)}$ à $\mathcal{C}$ est assurée grâce à la Proposition II.2.6 et qu'avec le cas d'égalité dans la relation (II.2.6) on obtient :

$$H_J(\mu_{E,J}^{(\theta,\theta_0)} \| \mu) = D(\theta, \theta_0) \qquad \text{(II.2.9)}$$

II.2.2 Convergence de la solution du problème discret sous la condition (C)

Pour la convergence du problème discret, il est naturel de passer par le problème dual qui s'envisage comme un problème de $\widehat{M}$-estimateurs suivant la terminologie de [19], qui est la recherche du minimum d'une fonction du type

$$\int \psi(x, u)\mu(dx)$$

Puisque le problème limite admet une solution, la convergence est établie dans [19] par exemple, avec des conditions sur J assez générales. On peut cependant les améliorer dans notre cadre particulier.

Remarque II.2.8 *Il est également possible d'envisager le problème comme un problème de M-estimateurs, toujours suivant la terminologie de [19] puisqu'il s'agit de trouver $(u, u_0) \in E \times \mathbb{R}$ tel que*

$$\int J^{*\prime}(u \cdot f + u_0)1_E(f)d\mu = 1$$

et

$$\int f J^{*\prime}(u \cdot f + u_0)1_E(f)d\mu = \mathbf{0}$$

avec la condition que le membre de gauche de cette dernière égalité existe.
En effet, d'après la réciproque de la Proposition II.2.4, la réciproque de ce problème donne celle du problème primal. en remplaçant (θ, θ_0) par (u, u_0).
Le point de vue des M-estimateurs est cependant moins pratique puisqu'il n'y a pas toujours de solution pour le cas limite.

Proposition II.2.9 *Sous la condition (C), presque sûrement, il existe un rang N à partir duquel le problème (P_n^J) admet une solution ν_n et la suite $(\nu_n)_{n\geq N}$ converge étroitement vers $\mu_{E,J}^{(\theta,\theta_0)}$.*
En particulier, la mesure $\mu_{E,J}^{(\theta,\theta_0)}$ est une probabilité.
De manière générale, si $\phi : S \to \mathbb{R}$ est mesurable bornée, on a la convergence presque sûre

$$\int \phi d\nu_n \longrightarrow_{n\to+\infty} \int \phi d\mu_{E,J}^{(\theta,\theta_0)}$$

Preuve : En reprenant le résultat de convergence établi pour l'entropie relative, on sait que la condition (C) permet d'assurer que presque sûrement à partir d'un certain rang N, elle est vérifiée en remplaçant μ par la distribution empirique μ_n, E étant l'espace défini par (C) :

$$(C_n) \begin{cases} \mathbf{0} \in Int_E(Conv(Supp(\mu^f_{n|E}))) \\ \forall \eta \in \mathcal{P}(S) \text{ absolument continue par rapport à } \mu_n \ t.q. \ \int \|f\| d\eta < +\infty \text{ et } \int f d\eta = \mathbf{0}, \\ \text{on a } \eta(f^{-1}(E)) = 1 \end{cases}$$

Alors la proposition II.2.6 fournit la solution :

$$\nu_n = J^{*'}(\theta_n \cdot f + \theta_{0,n}))1_E(f)\mu_n$$

où $(\theta_n, \theta_{0,n})$ est le minimum sur $E \times \mathbb{R}$ de

$$(u, u_0) \in E \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0) 1_E(f) d\mu_n - u_0$$

Montrons que $(\theta_n, \theta_{0,n})$ est bornée presque sûrement.
En effet, $(\theta_n, \theta_{0,n})$ réalisant un minimum, on obtient la majoration :

$$\int J^*(\theta_n \cdot f + \theta_{0,n}) 1_E(f) d\mu_n - \theta_{0,n} \leq \int J^*(0) 1_E(f) d\mu_n \leq |J^*(0)|$$

En appliquant le lemme II.2.10 énoncé à la fin de la preuve, avec $\eta_n = \mu_n$, on conclut que $(\theta_n, \theta_{0,n})$ est bornée presque sûrement.
Pour passer à la convergence, on considère ensuite, pour ω fixé, une valeur d'adhérence $(\theta_\infty, \theta_{0,\infty})$ de la suite $(\theta_n, \theta_{0,n})$.
Quitte à extraire une sous-suite, on peut supposer que la suite converge vers cette valeur limite et il s'agit d'identifier $(\theta_\infty, \theta_{0,\infty})$ à (θ, θ_0).
Avec le lemme de Skorokhod, comme on l'a vu pour l'entropie, quitte à choisir ω en dehors d'un ensemble négligeable, on obtient sur un espace de probabilité auxiliaire $(G, \mathcal{G}, Q)$ des variables (Y_n, W_n) et (Y, W) de lois respectives $\sum_{i=1}^n 1_E(f(X_i))\delta_{(X_i, f(X_i))} / \sum_{i=1}^n 1_E(f(X_i))$ et $\mu^{(id,f)}_{|f^{-1}(E)} / \mu(f^{-1}(E))$ telles que Q-ps (Y_n, W_n) converge vers (Y, W).
Avec la convergence supposée de $(\theta_n, \theta_{0,n})$, la continuité ainsi que la minoration de J^* on a donc,

en appliquant le lemme de Fatou et la loi forte des grands nombres :

$$\begin{aligned}
\liminf_{n\to+\infty}\left(\int J^*(\theta_n\cdot f+\theta_{0,n})1_E(f)d\mu_n-\theta_{0,n}\right) \\
&= \liminf_{n\to+\infty}\left(\frac{\sum_{i=1}^n 1_E(f(X_i))}{n}\mathbb{E}^Q J^*(\theta_n\cdot W_n+\theta_{0,n})-\theta_{0,n}\right) \\
&\geq \mu(f^{-1}(E))\mathbb{E}^Q J^*(\theta_\infty\cdot W+\theta_{0,\infty})-\theta_{0,\infty} \\
&= \int J^*(\theta_\infty\cdot f+\theta_{0,\infty})1_E(f)d\mu-\theta_{0,\infty} \qquad \text{(II.2.10)}
\end{aligned}$$

D'autre part, en rappelant que le couple $(\theta_n, \theta_{0,n})$ réalise le minimum du problème discret, on a :

$$\int J^*(\theta_n\cdot f+\theta_{0,n})1_E(f)d\mu_n-\theta_{0,n} \leq \int J^*(\theta\cdot f+\theta_0)1_E(f)d\mu_n-\theta_0 \to_{n\to\infty} \int J^*(\theta\cdot f+\theta_0)1_E(f)d\mu-\theta_0 \qquad \text{(II.2.11)}$$

en appliquant la loi forte des grands nombres.
Avec (II.2.10), on en déduit l'inégalité :

$$\int J^*(\theta_\infty\cdot f+\theta_{0,\infty})1_E(f)d\mu-\theta_{0,\infty} \leq \int J^*(\theta\cdot f+\theta_0)1_E(f)d\mu-\theta_0$$

ce qui prouve, par unicité du minimum du problème dual limite :

$$(\theta_\infty, \theta_{0,\infty}) = (\theta, \theta_0)$$

Le choix de ω se faisant en dehors d'un ensemble négligeable, on voit qu'ainsi, presque sûrement, $(\theta_n, \theta_{0,n})$ est une suite relativement compacte qui converge vers son unique valeur d'adhérence (θ, θ_0).
Pour terminer, la convergence étroite se prouve en montrant que $(\phi(Y_n)J^{*\prime}((\theta_n\cdot W_n+\theta_{0,n})))$ est une suite équi-intégrable, où $\phi : S \to \mathbb{R}$ est une fonction continue bornée.
C'est évident si f est bornée μ-ps et l'on supposera donc le contraire.
Avec la contrainte

$$\int f d\nu_n = \mathbf{0}$$

on a l'égalité

$$\int (\theta_n\cdot f+\theta_{0,n})J^{*'}(\theta_n\cdot f+\theta_{0,n})1_E(f)d\mu_n = \theta_{0,n} \qquad \text{(II.2.12)}$$

qui se réécrit :

$$\mathbb{E}^Q\left((\theta_n\cdot W_n+\theta_{0,n})J^{*'}(\theta_n\cdot W_n+\theta_{0,n})\right) = \frac{\theta_{0,n}}{\sum_{i=1}^n 1_E(f(X_i))/n} \qquad \text{(II.2.13)}$$

Le membre de gauche est le terme d'une suite bornée puisque le membre de droite converge vers $\frac{\theta_{0,\infty}}{\mu(f^{-1}(E))}$.
Le résultat repose alors essentiellement sur la propriété :

$$uJ^{*'}(u) \text{ est borné en } -\infty \qquad \text{(II.2.14)}$$

Pour prouver ce résultat, il suffit de prouver que la quantité $uJ^{*\prime}(u)$, négative en $-\infty$, est minorée.
En effet, on a :

$$\forall u\in\mathbb{R},\ J^*(0)-J^*(u) \geq -uJ^{*\prime}(u)$$

d'où

$$\forall u\in\mathbb{R},\ uJ^{*\prime}(u) \geq J^*(u)-J^*(0)$$

et l'on obtient la minoration voulue car J^* est minorée d'après (II.0.3).
La relation (II.2.13), la propriété (II.2.14) et le fait que $\lim_{u\to+\infty} J^{*\prime}(u) = +\infty$ (avec (II.0.2)) entrainent immédiatement l'uniforme intégrabilité de $(J^{*\prime}((\theta_n \cdot W_n + \theta_{0,n})))$ et donc l'équi-intégrabilité souhaitée puisque ϕ est bornée.
On en déduit

$$\mathbb{E}^Q \left(\phi(Y_n) J^{*'}(\theta_n \cdot W_n + \theta_{0,n})\right) \longrightarrow_{n\to+\infty} \mathbb{E}^Q \left(\phi(Y) J^{*'}(\theta \cdot W + \theta_0)\right)$$

puisque les intégrants convergent Q-ps (avec en particulier la continuité de ϕ). La convergence ci-dessus signifie exactement que l'on a pour toute fonction ϕ continue bornée

$$\int \phi d\nu_n \longrightarrow_{n\to+\infty} \int \phi d\mu_{E,J}^{(\theta,\theta_0)}$$

De manière plus générale, si $\phi : S \to \mathbb{R}$ est mesurable bornée, la convergence ci-dessus reste valable, p.s., en appliquant le lemme de Skorokhod à la convergence étroite p.s. de $\sum_{i=1}^n 1_E(f(X_i))\delta_{(\phi(X_i),f(X_i))}/\sum_{i=1}^n 1_E(f(X_i))$ vers $\mu_{|f^{-1}(E)}^{(\phi,f)}/\mu(f^{-1}(E))$. ∎

Il est intéressant de noter que c'est un procédé Monte-Carlo qui a permis ici de caractériser la mesure $\mu_{E,J}^{(\theta,\theta_0)}$ comme une probabilité.
Le lemme suivant établit la compacité relative de la suite $(\theta_n, \theta_{0,n})$ qui manque dans la preuve de la proposition précédente. Ce résultat sert également à la preuve de l'existence de la solution du problème dual et c'est également ici que la condition (II.0.2) est essentielle.

Lemme II.2.10 *Soit $(\eta_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{P}(S)$ telle que $\eta_{n|E}^f$ converge étroitement vers $\mu_{|E}^f$.*
Sous la condition (C), pour que $(u_n, u_{0,n})$ soit une suite bornée de $E \times \mathbb{R}$, il suffit que la suite

$$\int J^*(u_n \cdot f + u_{0,n}) 1_E(f) d\eta_n - u_{0,n} \tag{II.2.15}$$

soit majorée.

Preuve : Tout d'abord, il est facile de remarquer avec la convexité de J^* que la condition (II.0.2) équivaut à :

$$\lim_{u\to+\infty} \frac{J^*(u)}{u} = +\infty \tag{II.2.16}$$

Supposons ensuite, par contraposée que $(u_n, u_{0,n})$ ne soit pas bornée et montrons que la suite définie par (II.2.15) n'est pas majorée. Quitte à se limiter à une sous-suite, on peut considérer que l'on a

$$\|u_n\| + |u_{0,n}| \longrightarrow_{n\to\infty} +\infty$$

et que l'on peut trouver un élément non nul (α, α_0) de $E \times \mathbb{R}$ tel que :

$$\frac{(u_n, u_{0,n})}{\|u_n\| + |u_{0,n}|} \longrightarrow_{n\to\infty} (\alpha, \alpha_0)$$

Le lemme se prouve en examinant les différents cas possibles.
D'abord, avec la minoration (II.0.3), on a :

$$\int J^*(u_n \cdot f + u_{0,n}) 1_E(f) d\eta_n - u_{0,n} \geq -|J(0)| - u_{0,n}$$

Si $(u_{0,n})$ n'est pas minorée, le lemme est donc acquis, sinon on a :

$$\alpha_0 \geq 0 \tag{II.2.17}$$

Supposons alors que α est non nul. Dans ce cas, E n'est pas réduit à $\{\mathbf{0}\}$ et la condition (C) montre l'existence de ϵ, $a > 0$ tels que :

$$\mu(f \in E,\ \alpha \cdot f > 3\epsilon) > 3a$$

On peut alors trouver un ouvert borné U de E tel que

$$\mu(f \in U,\ \alpha \cdot f > 3\epsilon) > 2a$$

et donc, à partir d'un certain rang, compte tenu de la convergence étroite des restrictions à E des images par f des probabilités η_n, on a la minoration

$$\eta_n(f \in U,\ \alpha \cdot f > 3\epsilon) > a$$

Comme pour n assez grand, on a, U étant supposé borné :

$$\forall y \in U,\ \left| \frac{u_n}{\|u_n\| + |u_{0,n}|} \cdot y - \alpha \cdot y \right| < \epsilon$$

on en déduit qu'à partir d'un certain rang, on a :

$$\eta_n\left(f \in U,\ \frac{u_n}{\|u_n\| + |u_{0,n}|} \cdot f > 2\epsilon \right) > a$$

Donc, pour n assez grand pour assurer également :

$$\left| \frac{u_{0,n}}{\|u_n\| + |u_{0,n}|} - \alpha_0 \right| < \epsilon$$

on aurait

$$\eta_n(f \in E, u_n \cdot f + u_{0,n} > (\alpha_0 + \epsilon)(\|u_n\| + |u_{0,n}|)) > a$$

Avec la croissance et la minoration de J^*, on en déduit :

$$\int J^*(u_n \cdot f + u_{0,n}) 1_E(f) d\eta_n - u_{0,n} \geq -|J(0)| + aJ^*((\|u_n\| + |u_{0,n}|)(\alpha_0 + \epsilon)) - (\|u_n\| + |u_{0,n}|)(\alpha_0 + \epsilon)$$

Puisque $\alpha_0 + \epsilon$ est strictement positif à cause de (II.2.17), la relation (II.2.16) montre que le membre de droite diverge vers $+\infty$, ce qui prouve le lemme dans ce cas.
Finalement, toujours avec la condition (II.2.17), examinons le dernier cas où α est nul et l'on a alors

$$u_{0,n} \to +\infty$$

et

$$\frac{u_n}{u_{0,n}} \to \mathbf{0}$$

Soit U' un ouvert borné de E tel que :

$$\mu(f \in U') = 2a' > 0$$

Alors, pour n assez grand, on a :

$$\forall y \in U',\ |u_n \cdot y| < \epsilon u_{0,n}$$

avec $\epsilon > 0$.
On a pour n assez grand, comme ce qui a été fait précédemment :

$$\int J^*(u_n \cdot f + u_{0,n}) 1_E(f) d\eta_n - u_{0,n} \geq -|J(0)| + a' J^*(u_{0,n}(1-\epsilon)) - u_{0,n}$$

En prenant ϵ strictement plus petit que 1, on voit donc que le membre de droite diverge encore vers $+\infty$, ce qui achève la preuve du lemme. ∎

Avec des hypothèses supplémentaires, on peut étendre le résultat de convergence de la Proposition II.2.9 à des fonctions non bornées.

Corollaire II.2.11 *Supposons que J^* soit de classe C^2.*
Si $\phi : S \to \mathbb{R}$ est une fonction mesurable tel que :

$$\int |\phi| d\mu_{E,J}^{(\theta,\theta_0)} < +\infty$$

et s'il existe un voisinage Δ de (θ, θ_0) dans $E \times \mathbb{R}$ tel que :

$$\int \sup_{(u,u_0)\in\Delta} J^{*\prime\prime}(u \cdot f + u_0) \max(1, \|f\|) |\phi| 1_E(f) d\mu < +\infty \tag{II.2.18}$$

alors, sous la condition (C), on a presque sûrement la convergence

$$\int \phi d\nu_n \longrightarrow_{n\to+\infty} \int \phi d\mu_{E,J}^{(\theta,\theta_0)}$$

Preuve : Avec la convergence établie dans la Proposition II.2.9, presque sûrement la suite $(\theta_n, \theta_{0,n})$ se trouve dans Δ à partir d'un rang fini N_0.
En supposant Δ convexe, sans perte de généralité, et en appliquant l'inégalité des accroissements finis, on a p.s., pour $n \geq N_0$:

$$\begin{aligned} &|\int \phi J^{*\prime}(\theta_n \cdot f + \theta_{0,n}) 1_E(f) d\mu_n - \int \phi J^{*\prime}(\theta \cdot f + \theta_0) 1_E(f) d\mu_n| \\ &\leq \ (\|\theta_n - \theta\| + |\theta_{0,n} - \theta_0|) \int \sup_{(u,u_0)\in\Delta} J^{*\prime\prime}(u \cdot f + u_0) \max(1, \|f\|) |\phi| 1_E(f) d\mu_n \end{aligned}$$

En appliquant la loi forte des grands nombres, avec la condition (II.2.18), et en utilisant la convergence de la suite $(\theta_n, \theta_{0,n})$, on obtient le résultat. ∎

On peut citer en complément un résultat relatif à la convergence des estimateurs $\int J^{*\prime}(\theta_n \cdot f + \theta_{0,n}) 1_E(f) d\mu_n$ de $\int J^{*\prime}(\theta \cdot f + \theta_0) 1_E(f) d\mu$.
Pour cela, il suffit de se replacer dans la situation de la preuve de la Proposition II.2.9 où l'on rappelle que l'on a obtenu presque sûrement, avec le lemme de Skorokhod, sur un espace de probabilité auxiliaire $(G, \mathcal{G}, Q)$ des variables W_n et W de lois respectives $\sum_{i=1}^n 1_E(f(X_i))\delta_{f(X_i)} / \sum_{i=1}^n 1_E(f(X_i))$ et $\mu^f_{|f^{-1}(E)} / \mu(f^{-1}(E))$ telles que Q-ps W_n converge vers W.
En combinant les relations (II.2.11) et (II.2.10) de la Proposition II.2.9, on obtient, compte tenu

de la convergence de la suite $(\theta_n, \theta_{0,n})$

$$\begin{aligned}
\int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0 &= \lim_{n\to+\infty} \left(\int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu_n - \theta_0 \right) \\
&\geq \limsup_{n\to+\infty} \left(\int J^*(\theta_n \cdot f + \theta_{0,n}) 1_E(f) d\mu_n - \theta_{0,n} \right) \\
&= \limsup_{n\to+\infty} \left(\frac{\sum_{i=1}^n 1_E(f(X_i))}{n} \mathbb{E}^Q J^*(\theta_n \cdot W_n + \theta_{0,n}) - \theta_{0,n} \right) \\
&\geq \liminf_{n\to+\infty} \left(\frac{\sum_{i=1}^n 1_E(f(X_i))}{n} \mathbb{E}^Q J^*(\theta_n \cdot W_n + \theta_{0,n}) - \theta_{0,n} \right) \\
&\geq \mu(f^{-1}(E)) \mathbb{E}^Q J^*(\theta \cdot W + \theta_0) - \theta_0 \\
&= \int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0 \qquad \text{(II.2.19)}
\end{aligned}$$

Les inégalités précédentes sont donc des égalités et l'on énonce :

Corollaire II.2.12 *Sous la condition (C), on a la convergence presque sûre :*

$$\int J^*(\theta_n \cdot f + \theta_{0,n}) 1_E(f) d\mu_n \longrightarrow_{n\to\infty} \int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu$$

Avec le résultat précédent, on sait que la valeur minimale associée au problème dual discret, $\int J^*(\theta_n \cdot f + \theta_{0,n}) 1_E(f) d\mu_n - \theta_{0,n}$, converge vers $\int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0$ qui est la valeur minimale associée au problème dual limite.
Cette valeur est naturellement reliée à la distance associée au problème primal grâce à la relation (II.2.4) de la Proposition II.2.4. On en avait reporté la preuve que l'on se propose de présenter ici.

Lemme II.2.13 *Sous l'hypothèse (II.0.2) et avec la condition (C), on a l'égalité*

$$H_J(\mu \| \mathcal{C}) = -\left(\int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0 \right) + J(0)\mu(f^{-1}(E^c))$$

Le corollaire II.2.12 s'énonce ainsi comme la convergence presque sûre :

$$\inf_{\nu \in \mathcal{P}(S), \int \|f\| d\nu < +\infty, \int f d\nu = \mathbf{0}} H_J(\nu \| \mu_n) \longrightarrow_{n\to\infty} H_J(\mu \| \mathcal{C})$$

Preuve : Compte tenu de la relation (II.2.8), il suffit de prouver l'inégalité

$$H_J(\mu \| \mathcal{C}) \leq -\left(\int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0 \right) + J(0)\mu(f^{-1}(E^c)) \qquad \text{(II.2.20)}$$

Comme dans Jupp & Mardia [72], on va procéder par approximation et donc considérer une suite croissante (S_n) de sous-ensembles mesurables de S, supposés de mesure non nulle et de réunion S, et tels que f soit bornée sur chaque S_n.
On pose

$$\tilde{\mu}_n = \frac{1_{S_n}}{\mu(S_n)} \mu$$

On vérifie alors qu'à partir d'un certain rang on a :

$$\mathbf{0} \in Int_E(Conv(Supp(\tilde{\mu}^f_{n|E}))) \tag{II.2.21}$$

En effet, avec la relation (C), il existe un hypercube Π de E tel que

$$\mathbf{0} \in Int_E(\Pi) \subset \Pi \subset Conv(Supp(\mu^f_{|E}))$$

Comme

$$Supp(\mu^f_{|E}) = \overline{\bigcup Supp(\tilde{\mu}^f_{n|E})}$$

on a

$$\begin{aligned}
\Pi &\subset Int_E(Conv(\overline{\bigcup Supp(\tilde{\mu}^f_{n|E})})) \\
&= Int_E(Conv(\bigcup Supp(\tilde{\mu}^f_{n|E}))) \\
&\subset Conv(\bigcup Supp(\tilde{\mu}^f_{n|E})) \\
&= \bigcup Conv(Supp(\tilde{\mu}^f_{n|E}))
\end{aligned}$$

L'égalité de la deuxième ligne résulte d'une propriété classique de l'enveloppe convexe tandis que la deuxième vient de la croissance des supports. On en déduit en particulier que les sommets de Π appartiennent à $Conv(Supp(\tilde{\mu}^f_{n|E}))$ pour n assez grand, ce qui suffit à prouver (II.2.21).
Puisque f est bornée sur S_n, on sait, en vertu de ce qui précède, et grâce à la remarque qui suit la Proposition II.2.6, qu'à partir d'un certain rang, on peut appliquer la relation (II.2.4) en remplaçant μ par $\tilde{\mu}_n$:

$$H_J(\tilde{\mu}_n \| \mathcal{C}) = -\left(\int J^*(\tilde{\theta}_n \cdot f + \tilde{\theta}_{0,n}) 1_E(f) d\tilde{\mu}_n - \tilde{\theta}_{0,n}\right) + J(0)\tilde{\mu}_n(f^{-1}(E^c))$$

où la suite $(\tilde{\theta}_n, \tilde{\theta}_{0,n})$ réalise le minimum de

$$\int J^*(u \cdot f + u_0) 1_E(f) d\tilde{\mu}_n - u_0$$

On en déduit :

$$\begin{aligned}
H_J(\mu \| \mathcal{C}) &\leq \lim_{n\to+\infty} H_J(\tilde{\mu}_n \| \mathcal{C}) \\
&= -\lim_{n\to+\infty} \left(\int J^*(\tilde{\theta}_n \cdot f + \tilde{\theta}_{0,n}) 1_E(f) d\tilde{\mu}_n - \tilde{\theta}_{0,n}\right) + J(0)\tilde{\mu}_n(f^{-1}(E^c))
\end{aligned} \tag{II.2.22}$$

Pour étudier la convergence du membre de droite de l'égalité précédente, on commence par montrer que la suite $(\tilde{\theta}_n, \tilde{\theta}_{0,n})$ est bornée en appliquant le Lemme II.2.10. Il suffit en effet de noter que la suite $1_E \tilde{\mu}^f_n / \tilde{\mu}_n(f^{-1}(E))$ converge en variation, et donc étroitement, vers $1_E \mu^f / \mu(f^{-1}(E))$, ce qui permet d'appliquer le lemme de Skorokhod.
On a donc :

$$\int J^*(\tilde{\theta}_n \cdot f + \tilde{\theta}_{0,n}) d\tilde{\mu}_n - \tilde{\theta}_{0,n} + J(0)\tilde{\mu}_n(f^{-1}(E^c)) \longrightarrow_{n\to+\infty} \int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0 J(0)\mu(f^{-1}(E^c)$$

On obtient finalement l'inégalité (II.2.20) grâce à la relation (II.2.22). ∎

II.2.3 Prise en compte de contraintes d'équivalence de la mesure de pricing et de la mesure a priori

Puisque la probabilité a priori a pour fonction principale de caractériser les évènements négligeables, on peut souhaiter que la solution soit à rechercher parmi les probabilités qui lui sont équivalentes. Du point de vue discret, une condition nécessaire à l'existence d'une probabilité calibrée équivalente à la mesure empirique μ_n, à partir d'un rang n assez grand, est la condition (C) avec pour E le sous-espace affine engendré par $Supp(\mu^f)$, c'est-à-dire :

$$(C_e) \ \ \mathbf{0} \in ri(Conv(Supp(\mu^f)))$$

où ri désigne l'intérieur relatif. Alors, pour n assez grand, le problème P_n^J admet la solution ν_n équivalente à μ_n qui est également la solution du problème

$(P_{e,n}^J)$ trouver $\nu_n \in \mathcal{P}(S)$ avec $\nu_n \sim \mu_n$ et qui minimise $H_J(\nu_n \| \mu_n)$ sous la contrainte $\int f d\nu_n = \mathbf{0}$

Le problème limite associé à ce dernier problème est :

(PL_e^J) trouver $\nu \in \mathcal{P}(S)$, $\nu \sim \mu$ qui minimise $H_J(\nu \| \mu)$ sous la contrainte $\int f d\nu = \mathbf{0}$

On sait que si le problème (PL^J) admet une solution, cette dernière vaut $\mu_J^{(\theta,\theta_0)}$ qui est équivalente à μ et c'est aussi la solution du problème (PL_e^J). En particulier, on a :

$$\inf_{\nu \in \mathcal{C}_e} H_J(\nu \| \mu) = H_J(\mu \| \mathcal{C}) \tag{II.2.23}$$

en introduisant l'ensemble :

$$\mathcal{C}_e = \left\{ \nu \in \mathcal{P}(S),\ \nu \sim \mu,\ \int \|f\| d\nu < +\infty \text{ et } \int f d\nu = \mathbf{0} \right\} \tag{II.2.24}$$

Réciproquement, le prochain résultat exclut le cas où le problème (PL_e^J) admet une solution au sens usuel mais pas (PL^J) : autrement, la solution ν_n du problème discret $(PL_{e,n}^J)$, également solution du problème (PL_n^J) à partir d'un certain rang, convergerait vers la solution généralisée du problème (PL^J) qui, n'étant pas calibrée, différerait de la solution usuelle du problème (PL_e^J).

Lemme II.2.14 *Sous la condition (C_e), on a l'identité (II.2.23).*
En particulier le problème (PL_e^J) admet une solution si et seulement s'il en est ainsi du problème (PL^J) et leurs solutions sont alors identiques.

Preuve : Il s'agit de trouver une suite de probabilités λ_n de $\mathcal{C}_e$ telle que $H_J(\lambda_n \| \mu)$ approche $H_J(\mu \| \mathcal{C})$.
Pour cela, on reprend le procédé d'approximation de la preuve du Lemme II.2.13 en considérant au lieu des probabilités $\tilde{\mu}_n = \frac{1_{S_n}}{\mu(S_n)} \mu$ les mesures $1_{S_n}\mu$, ce qui permet d'avoir en particulier à partir d'un certain rang n la relation $\mathbf{0} \in Int(Conv(Supp((1_{S_n}\mu)^f)))$. On peut alors vérifier que les développements faits dans l'étude du problème limite restent entièrement valides si l'on remplace la probabilité μ par la mesure positive bornée $1_{S_n}\mu$, et en particulier

$$H_J(1_{S_n}\mu \| \mathcal{C}) = -\left(\int J^*(\tilde{\theta}_n \cdot f + \tilde{\theta}_{0,n}) 1_{S_n} d\mu - \tilde{\theta}_{0,n} \right)$$

où la suite $(\tilde{\theta}_n, \tilde{\theta}_{0,n})$ de $\mathbb{R} \times E$ (en rappelant ici que le sous-espace E est engendré par $Supp(\mu^f)$) réalise le minimum de

$$\int J^*(u \cdot f + u_0) 1_{S_n} d\mu - u_0$$

En tenant compte alors de la convergence de $\mu(S_n)$ et de la convergence en variation de $\tilde{\mu}_n^f$, on utilise une technique identique à la preuve du Lemme II.2.10 pour montrer que la suite $(\tilde{\theta}_n, \tilde{\theta}_{0,n})$ est bornée. Avec le lemme de Skorokhod, on obtient ensuite la convergence suivante

$$\int J^*(\tilde{\theta}_n \cdot f + \tilde{\theta}_{0,n}) 1_{S_n} \mu - \tilde{\theta}_{0,n} = \mu(S_n) \int J^*(\tilde{\theta}_n \cdot f + \tilde{\theta}_{0,n}) d\tilde{\mu}_n - \tilde{\theta}_{0,n} \longrightarrow \int J^*(\theta \cdot f + \theta_0) 1_E(f) d\mu - \theta_0$$

D'autre part, puisque f est bornée sur S_n, on peut également voir par adaptation des résultats du cas limite que $H_J(1_{S_n}\mu \| \mathcal{C})$ est effectivement atteint pour $\tilde{\nu}_n = J^{*'}(\tilde{\theta}_n + \tilde{\theta}_{0,n}) 1_{S_n}\mu$.
Avec la relation (II.2.4), la convergence précédente s'écrit :

$$H_J(\tilde{\nu}_n \| 1_{S_n}\mu) \to H_J(\mu \| \mathcal{C}) \tag{II.2.25}$$

Posons alors pour tout $n \in \mathbb{N}$:

$$\alpha_n = \frac{1}{n^2} \frac{1}{1 + \int \|f\| 1_{S_n} d\tilde{\nu}_n}$$

et pour tout $i \geq n$:

$$\alpha_i^n = \frac{\alpha_i}{\sum_{j \geq n} \alpha_j}$$

Définissons :

$$\lambda_n = \sum_{i \geq n} \alpha_i^n \tilde{\nu}_i$$

Avec le théorème de Fubini et le théorème de la convergence dominée, c'est clairement une suite de probabilités calibrées et de plus équivalentes à μ. On a en particulier :

$$\forall n \geq 0, \; H_J(\lambda_n \| \mu) \geq H_J(\mu \| \mathcal{C})$$

Avec le lemme de Fatou et la convexité de J, on a par ailleurs pour tout n

$$\begin{aligned}
H_J(\lambda_n \| \mu) &\leq \liminf_p H_J\left(\sum_{i=n}^p \alpha_i^n \tilde{\nu}_i \| \mu \right) \\
&\leq \liminf_p \left(\sum_{i=n}^p \alpha_i^n H_J(\tilde{\nu}_i \| \mu) + \left(1 - \sum_{i=n}^p \alpha_i^n \right) J(0) \right) \\
&= \liminf_p \left(\sum_{i=n}^p \alpha_i^n H_J(\tilde{\nu}_i \| 1_{S_i}\mu) + \sum_{i=n}^p \alpha_i^n (1 - \mu(S_i)) J(0) + \left(1 - \sum_{i=n}^p \alpha_i^n \right) J(0) \right) \\
&= \liminf_p \sum_{i=n}^p \alpha_i^n H_J(\tilde{\nu}_i \| 1_{S_i}\mu)
\end{aligned} \tag{II.2.26}$$

Avec la convergence (II.2.25), on en déduit :

$$\liminf H_J(\lambda_n \| \mu) \leq \limsup H_J(\lambda_n \| \mu) \leq H_J(\mu \| \mathcal{C})$$

Ainsi

$$\lim H_J(\lambda_n \| \mu) = H_J(\mu \| \mathcal{C})$$

ce qui prouve l'inégalité

$$\inf_{\nu \in \mathcal{C}_e} H(\nu \| \mu) \leq H_J(\mu \| \mathcal{C})$$

et l'égalité voulue en raison de l'inégalité inverse. ∎

II.2.4 Résultats connus de $\widehat{M}$-estimateurs

Au cours de la preuve de la Proposition II.2.9, on a établi la convergence de la suite $(\theta_n, \theta_{0,n})$ sous les conditions (II.0.2) et (C). Au prix d'un renforcement de ces conditions, on aurait cependant pu établir cette convergence en utilisant des résultats connus concernant les $\widehat{M}$-estimateurs et exposés dans [19]. On va montrer cela précisément en rajoutant l'hypothèse d'intégrabilité

$$\forall (t, t_0) \in E \times \mathbb{R}, \int J^*(t \cdot f + t_0) 1_E(f) d\mu < +\infty \tag{II.2.27}$$

La condition (II.2.27) est une hypothèse assez forte que l'on pose pour pouvoir appliquer sans problème les conditions de [19] et est vérifiée si f est bornée par exemple. On va se ramener à la situation du Théorème 1 p.60 de [19]. Afin d'adapter ce résultat, on note

$$\forall (t, t_0) \in E \times \mathbb{R},\ \forall x \in S,\ \psi(x, t, t_0) = J^*(t \cdot f(x) + t_0) 1_E(f(x)) - t_0$$

et on pose pour tout $\Delta > 0$:

$$\forall (t, t_0) \in E \times \mathbb{R},\ \forall x \in S,\ \psi^\Delta(x, t, t_0) = \inf_{||u||^2 + |u_0|^2 \le \Delta^2} \psi(x, t + u, t_0 + u_0)$$

Avec le Théorème 1 p.60 de [19], on a la convergence des $\widehat{M}$-estimateurs sous les conditions suivantes, écrites avec des modifications mineures :

- (A_c) Les $\widehat{M}$-estimateurs sont recherchés dans K, compact fixé de $E \times \mathbb{R}$.
- (A_0) La fonction $(t, t_0) \longmapsto \int \psi(x, t, t_0) \mu(dx)$ atteint un minimum unique (θ, θ_0) sur K.
- (A_0^Δ) $\forall \delta > 0,\ \exists \Delta > 0, \epsilon > 0,\ \inf_{||t-\theta||^2 + |t_0 - \theta_0|^2 \ge \delta^2} \int (\psi^\Delta(x, t, t_0) - \psi(x, \theta, \theta_0)) \mu(dx) > \epsilon$

Il est cependant nécessaire d'affaiblir la condition de compacité (A_c), qui est faite dans [19] par souci de simplicité, à condition de remplacer le jeu de conditions précédents par le suivant :

- (A'_c) La suite des $\widehat{M}$-estimateurs est presque sûrement dans un compact de $E \times \mathbb{R}$.
- (A'_0) La fonction $(t, t_0) \longmapsto \int \psi(x, t, t_0) \mu(dx)$ atteint un minimum unique (θ, θ_0) sur $E \times \mathbb{R}$.
- $(A_0'^\Delta)$ $\forall \gamma, \delta$ t.q. $\gamma \ge \delta > 0,\ \exists \Delta > 0, \epsilon > 0,$
  $\inf_{\gamma^2 \ge ||t-\theta||^2 + |t_0 - \theta_0|^2 \ge \delta^2} \int (\psi^\Delta(x, t, t_0) - \psi(x, \theta, \theta_0)) \mu(dx) > \epsilon$

En adaptant la preuve du Théorème 1 p. 60 de [19], on obtient en effet la convergence des $\widehat{M}$-estimateurs sous ces trois hypothèses. Avant de le démontrer, on doit vérifier la validité de ces hypothèses.

Proposition II.2.15 *Sous les hypothèses (C) et (II.2.27), on obtient les conditions (A'_c), (A'_0) et $(A_0'^\Delta)$.*

Preuve : En effet, avec le lemme II.2.10, on a d'abord la condition (A'_c) de relative compacité p.s. de la suite des $\widehat{M}$-estimateurs. Ensuite, la condition (A'_0) est le résultat d'existence et d'unicité de la solution du problème dual, donné par la Proposition II.2.3. Enfin, la condition $(A_0'^\Delta)$ s'obtient en adaptant la preuve du Théorème 2 p.61 de [19]. Supposons en effet, par l'absurde que $(A_0'^\Delta)$ ne soit pas vérifiée et qu'il existe ainsi des nombres γ, δ avec $\gamma \ge \delta > 0$ ainsi qu'une suite $(t_k, t_{0,k}) \in E \times \mathbb{R}$ avec $\gamma^2 \ge \|t_k - \theta\|^2 + |t_{0,k} - \theta_0|^2 \ge \delta^2$ et une suite $\Delta_k > 0$ tendant vers 0 tels que

$$\lim_{k \to +\infty} \int (\psi^{\Delta_k}(x, t_k, t_{0,k}) - \psi(x, \theta, \theta_0)) \mu(dx) = 0 \tag{II.2.28}$$

La suite $(t_k, t_{0,k})$ étant bornée, on peut supposer qu'elle converge vers $(t, t_0) \in E \times \mathbb{R}$ avec $\gamma^2 \geq \|t-\theta\|^2 + |t_0 - \theta_0|^2 \geq \delta^2$. D'autre part, la fonction J^* étant minorée, on peut appliquer le lemme de Fatou et écrire, compte tenu de la continuité de J^* et du fait que $(t, t_0) \neq (\theta, \theta_0)$:

$$\liminf_{k\to\infty} \int (\psi^{\Delta_k}(x, t_k, t_{0,k}) - \psi(x, \theta, \theta_0))\mu(dx) \geq \int (\psi(x, t, t_0) - \psi(x, \theta, \theta_0))\mu(dx) > 0$$

On obtient donc une contradiction avec la relation (II.2.28). ∎

Il reste à montrer la convergence des $\widehat{M}$-estimateurs.

Proposition II.2.16 *Sous l'ensemble des hypothèses* $(A'_c), (A'_0)$ *et* $(A_0'^{\Delta})$*, la suite des* $\widehat{M}$*-estimateurs* $(\theta_n, \theta_{0,n})$ *converge vers* (θ, θ_0).

Preuve : En effet, avec la condition (A'_c), on peut trouver un ensemble A d'épreuves de complémentaire négligeable sur lequel la suite des $\widehat{M}$-estimateurs $(\theta_n, \theta_{0,n})$ est bornée en tout point et sur lequel on a, avec la loi forte des grands nombres, en choisissant un ensemble dénombrable dense $\tilde{E}$ de E :

$$\forall \gamma, \delta \in \mathbb{Q}_+^*,\ \forall (t, t_0) \in \tilde{E} \times \mathbb{Q},\ \int \psi^{\Delta(\gamma,\delta)}(x, t, t_0)\mu_n(dx) \longrightarrow_{n\to\infty} \int \psi^{\Delta(\gamma,\delta)}(x, t, t_0)\mu(dx)$$

où $\Delta(\gamma, \delta)$ est associé au couple (γ, δ) de manière à satisfaire la condition $(A_0'^{\Delta})$. Notons que c'est ici qu'intervient la condition d'intégrabilité (II.2.27)
Fixons à présent un élément ω de A et soit $\delta \in \mathbb{Q}_+^*$ et $\gamma \in \mathbb{Q}_+^*$ assez grand pour que l'ensemble $\overline{B}((\theta, \theta_0), \gamma) \bigcap E \times \mathbb{R}$ contienne la suite $(\theta_n(\omega), \theta_{0,n}(\omega))$. On peut recouvrir l'ensemble $\big(\overline{B}((\theta, \theta_0), \gamma) B((\theta, \theta_0), \delta)\big) \bigcap E \times \mathbb{R}$ par des voisinages

$$U_k = \{(t, t_0) \in E \times \mathbb{R} \mid \|t - t_k\|^2 + |t_0 - t_{0,k}|^2 \leq \Delta^2\},\ 1 \leq k \leq p < \infty$$

avec $(t_k, t_{0,k}) \in \tilde{E} \times \mathbb{Q},\ (t_k, t_{0,k}) \notin B((\theta, \theta_0), \delta) \bigcap E \times \mathbb{R}$ et $\Delta = \Delta(\gamma, \delta)$.

$$\begin{aligned} \inf_{\gamma^2 \geq \|t-\theta\|^2 + |t_0-\theta_0|^2 \geq \delta^2} \textstyle\int (\psi(x, t, t_0) - \psi(x, \theta, \theta_0))\mu_n(dx) & \\ &\geq \min_k \textstyle\int (\psi^{\Delta}(x, t_k, t_{0,k}) - \psi(x, \theta, \theta_0))\mu_n(dx) \\ &\longrightarrow \min_k \textstyle\int (\psi^{\Delta}(x, t_k, t_{0,k}) - \psi(x, \theta, \theta_0))\mu(dx) > \epsilon \end{aligned}$$

où $\epsilon > 0$ est associé au couple (γ, δ) par la condition $(A_0'^{\Delta})$.
Il apparaît alors que le minimum $(\theta_n, \theta_{0,n})$ de $(t, t_0) \mapsto \int \psi(x, t, t_0)$ sur $B((\theta, \theta_0), \gamma) \bigcap E \times \mathbb{R}$ se trouve dans $B((\theta, \theta_0), \delta)$ pour n assez grand, ce qui prouve la convergence souhaitée car δ est arbitrairement petit. ∎

II.3 Problème pénalisé

D'un point de vue financier, il est souhaitable de relâcher les contraintes puisque les prix ne sont connus que dans une fourchette. Dans cette partie, on présente une généralisation de la méthode de pénalisation adoptée dans [6] : les auteurs proposent ainsi d'ajouter un terme quadratique à l'entropie relative de manière à considérer le problème de minimisation sur $\nu \in \mathcal{P}(S)$

$$H(\nu\|\mu_n) + \frac{1}{2}\sum_{i=1}^{d} \frac{1}{\gamma_i}\Big(\int_S f^i d\nu\Big)^2$$

où les γ^i, $1 \le i \le d$ sont d nombres strictement positifs.
L'objectif que l'on se fixe ici est d'étudier la convergence Monte-Carlo de la solution du problème de calibration discret dans un cadre toutefois plus général que [6] où l'aspect de convergence n'est d'ailleurs pas abordé. Comme dans le cas de la calibration avec contraintes, on obtient des résultats plus généraux que ceux fournis par les résultats classiques de M-estimateurs.

On reprend donc le critère J précédent et l'on se donne en outre une fonction $\chi : \mathbb{R}^d \longrightarrow \mathbb{R}$, finie, strictement convexe, sci et minimale en $\mathbf{0}$. On introduit la fonction convexe conjuguée $\chi^* : \mathbb{R}^d \longrightarrow]-\infty, +\infty]$ définie par

$$\forall u \in \mathbb{R}^d,\ \chi^*(u) = \sup_{x \in \mathbb{R}^d} (u \cdot x - \chi(x)) \tag{II.3.1}$$

Avec le théorème de Fenchel-Moreau, les fonctions χ et χ^* sont conjuguées l'une de l'autre, on a donc :

$$\forall x \in \mathbb{R}^d,\ \chi(x) = \sup_{u \in \mathbb{R}^d} (x \cdot u - \chi^*(u)) \tag{II.3.2}$$

C'est essentiellement sous cette dernière forme que l'on exploitera la dualité.
Précisons ici les propriétés de χ^* dont on aura besoin. Bien sûr, χ^* est sci et de plus elle est minorée. On montre en effet qu'elle atteint son minimum en $\mathbf{0}$

$$\begin{aligned} \chi^*(\mathbf{0}) &= \sup_{x \in \mathbb{R}^d} -\chi(x) \\ &= -\chi(\mathbf{0}) \\ &= -\sup_{u \in \mathbb{R}^d} -\chi^*(u) = \inf_{u \in \mathbb{R}^d} \chi^*(u) \end{aligned}$$

En plus de ces conditions automatiquement vérifiées, on supposera que χ^* est strictement convexe sur son domaine $dom\chi^* \equiv \{u \in \mathbb{R}^d,\ \chi^*(u) < +\infty\}$ et que l'on a enfin

$$\frac{\chi^*(u)}{\|u\|} \longrightarrow_{\|u\| \to +\infty} +\infty \tag{II.3.3}$$

On donnera plus loin un jeu de conditions portant uniquement sur χ ou sur sa conjuguée χ^* et permettant d'être dans le cadre de nos hypothèses. On verra qu'il suffit d'imposer des conditions de régularité supplémentaires, ce qui est vérifié en pratique mais, afin de traiter le cas de contraintes de type ensemble, objet de la section prochaine, on n'envisage pas ici une telle régularité qui par ailleurs n'apporte pas de simplification notable pour les démonstrations.
Commençons par préciser où interviennent les différentes hypothèses. D'abord en supposant que χ est finie, on s'assure que le problème de pénalisation défini ci-dessous est non trivial, de plus, le fait qu'elle soit minimale en $\mathbf{0}$ permet la cohérence du choix du critère avec la mesure a priori μ (voir le Lemme II.3.1) et finalement, les conditions portant sur χ^* imposent l'unicité et l'existence de la solution du problème dual introduit ultérieurement.
Ces précisions étant faites, le problème consiste à résoudre, pour n tirages Monte-Carlo, le vecteur de contraintes $\mathbf{C}$ étant ramené à $\mathbf{0}$:

$$(P_n^{J,\chi}) \quad \text{trouver } \nu_n \in \mathcal{P}(S) \text{ qui minimise } H_J(\nu_n \| \mu_n) + \chi(\textstyle\int f d\nu_n)$$

Rappelons que l'intérêt de la pénalisation est de pouvoir relâcher les contraintes et que l'existence d'une solution au problème $(P_n^{J,\chi})$ peut aussi être assurée sans condition particulière. En effet, le problème se reformule comme la recherche d'un vecteur $(p_i)_{1 \le i \le n}$ de $[0,1]^n$ tel que $\sum_{i=1}^n p_i = 1$, minimisant la fonction

$$\frac{1}{n} \sum_{i=1}^n J(np_i) + \chi\left(\sum_{i=1}^n f(X_i) p_i\right)$$

Il s'agit donc de trouver le minimum d'une fonction strictement convexe sur un compact et l'existence de la solution, forcément unique, est acquise car J est sci puisqu'on l'a supposée en particulier continue à droite en 0. Comme dans le cas contraint, cette dernière hypothèse est gratuite.
Afin d'étudier la convergence de cette solution, on introduit naturellement le problème limite :

$$(PL^{J,\chi}) \quad \text{trouver } \nu \in \mathcal{P}(S) \text{ qui minimise } H_J(\nu\|\mu) + Q_\chi(\nu)$$

en posant pour tout $\nu \in \mathcal{P}(S)$

$$Q_\chi(\nu) = \begin{cases} \chi(\int f d\nu) & \text{si } \int_S \|f\| d\nu < +\infty \\ +\infty & \text{sinon} \end{cases}$$

Avant de résoudre le problème $PL^{J,\chi}$ il convient au préalable de s'assurer la non vacuité de l'ensemble

$$\mathcal{E}_{J,\chi} = \{\nu \in \mathcal{P}(S),\ H_J(\nu\|\mu) + Q_\chi(\nu) < +\infty\} \tag{II.3.4}$$

Il suffit en effet de considérer la mesure

$$\mu_a = \frac{1_{\{\|f\| \le a\}}}{\mu(\{\|f\| \le a\})}\mu$$

où a est un réel positif assez grand pour que le dénominateur de l'expression précédente soit non nul. Puisque χ est finie sur $\mathbb{R}^d$, la probabilité μ_a appartient à l'ensemble $\mathcal{E}_{J,\chi}$.

II.3.1 Discussion des hypothèses sur le critère χ

Conditions portant sur χ^* seulement

Au début de cette section, on a donné un jeu de conditions où l'on impose essentiellement à χ d'être convexe, finie partout et minimale en $\mathbf{0}$, tandis que l'ensemble des autres hypothèses portent sur χ^*. Au prix d'un renforcement des hypothèses, on peut aboutir au fait que ces dernières ne portent plus que sur χ^*.
En imposant des conditions de régularité supplémentaires sur χ^*, on peut en effet s'assurer que sa conjuguée χ est finie partout. On part donc de la fonction convexe $\chi^* : \mathbb{R}^d \longrightarrow]-\infty, +\infty]$, sci, strictement convexe sur son domaine, minimale en $\mathbf{0}$ et l'on suppose vérifiée la condition (II.3.3). Faisons l'hypothèse supplémentaire que χ^* est différentiable sur l'intérieur supposé non vide de $dom\chi^*$ et imposons la condition

$$\chi^{*'}(Int(dom\chi^*)) = \mathbb{R}^d \tag{II.3.5}$$

que l'on complètera par la condition supplémentaire suivante, vacuitement réalisée si $dom\chi^*$ vaut $\mathbb{R}^d$:

$$\forall u \in Fr(dom\chi^*),\ \lim_{v \in dom\chi^* \to u} \|\chi^{*'}(v)\| = +\infty \tag{II.3.6}$$

Montrons alors que χ possède nécessairement les propriétés énoncées.
Il s'agit de montrer d'abord que χ^* est du type de Legendre dont on a donné la signification dans la première section du chapitre. Elle vérifie les trois conditions pour être essentiellement régulière :

1. $Int(dom\chi^*) \neq \emptyset$
2. χ^* est différentiable sur $Int(dom\chi^*)$

3. La condition (II.3.6) est vérifiée.

Etant strictement convexe, χ^* est donc du type de Legendre. Etant propre et semi-continue inférieurement, elle est également fermée. Avec le Théorème 26.5 p.258 de [106], on en déduit que la conjuguée χ de χ^* est également du type de Legendre, donc en particulier strictement convexe et dérivable sur l'intérieur de son domaine $dom\chi$, et que de plus $\chi^{*\prime}$ réalise une bijection de $Int(dom\chi^*)$ sur $Int(dom\chi)$, avec χ' pour bijection réciproque. La condition (II.3.5) montre alors :

$$dom\chi = \mathbb{R}^d$$

Finalement la condition de minimum est facile obtenir, comme au début de cette section.

Conditions portant sur χ seulement

Si l'on se donne réciproquement une fonction χ finie, strictement convexe et différentiable sur $\mathbb{R}^d$, la fonction χ^*, définie comme sa conjuguée, est strictement convexe et différentiable sur l'intérieur, non vide, de son domaine, avec de plus la condition (II.3.6) : c'est encore une conséquence du Théorème 26.5 p.258 de [106].
Cependant, la convexité stricte de χ^* n'est pas assurée sur l'ensemble du domaine de χ^*, ce qui pose un problème pour l'unicité de la solution du problème dual limite. Une façon de spécifier la fonction χ est alors d'imposer en plus la condition :

$$\lim_{x\in\mathbb{R}^d\to+\infty} \|\chi'(x)\| = +\infty \tag{II.3.7}$$

En effet, avec le Lemme 26.7 et le Théorème 26.6 de [106], on en déduit que :

$$dom\chi^* = \mathbb{R}^d$$

De plus, la condition (II.3.3) est alors équivalente à :

$$\lim_{x\in\mathbb{R}^d\to+\infty} \frac{\chi'(x)\cdot x - \chi(x)}{\|\chi'(x)\|} = +\infty \tag{II.3.8}$$

II.3.2 Etude du problème limite

Comme dans le cas contraint, on peut vérifier la cohérence du choix du critère de minimisation avec celle de la mesure a priori. En effet, avec le Lemme II.2.2 et le fait que χ est minimale en $\mathbf{0}$, on tire immédiatement

Lemme II.3.1 *Si μ est un élément de $\mathcal{C}$, alors c'est la solution du problème $(PL^{J,\chi})$.*

On ne dispose pas toujours d'une solution à ce problème, même en un sens généralisé comme dans le cas de l'entropie pénalisée et on est conduit à examiner une nouvelle fois un problème dual, introduit ici comme la minimisation de la fonction à valeurs dans $]-\infty,+\infty]$

$$(u,u_0)\in\mathbb{R}^d\times\mathbb{R}\longmapsto \int J^*(u\cdot f+u_0)d\mu - u_0 + \chi^*(-u) \tag{II.3.9}$$

L'unicité du minimum de (II.3.9) est assurée par la stricte convexité de χ^* et de J^* ainsi que le fait que (II.3.9) est finie en au moins un point (en notant que le minimum $\chi^*(\mathbf{0})$ est fini).
Pour l'existence, on utilise d'abord le fait que (II.3.9) définit une fonction sci, en utilisant en particulier le lemme de Fatou, comme dans le cas contraint. Pour conclure, il suffit de montrer la divergence de cette fonction vers $+\infty$ pour $\|u\|+|u_0|$ tendant vers l'infini, ce qui est est assuré par la condition (II.3.3) : on obtient ce résultat en prenant $\eta_n=\mu$ pour tout n dans le Lemme II.3.9 énoncé plus loin.

Proposition II.3.2 *La fonction*

$$(u, u_0) \in \mathbb{R}^d \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0) d\mu - u_0 + \chi^*(-u)$$

atteint son minimum en un point unique $(\theta, \theta_0) \in \mathbb{R}^d \times \mathbb{R}$.

L'intérêt d'introduire la fonction précédente apparaît dans l'égalité

$$\inf_{\nu \in \mathcal{P}(S)} (H_J(\nu \| \mu) + Q_\chi(\nu)) = -\left(\int J^*(\theta \cdot f + \theta_0) d\mu - \theta_0 + \chi^*(-\theta) \right) \tag{II.3.10}$$

Elle permettra en particulier l'identification de la solution du problème primal, mais on en renvoie aussi la démonstration dans l'étude du cas discret.

On a ainsi un énoncé presque identique à la Proposition II.2.4 :

Proposition II.3.3 *La valeur minimale associée au problème* $(PL^{J,\chi})$ *est donnée par la relation (II.3.10) et si le problème* $(PL^{J,\chi})$ *admet une solution, celle ci est*

$$\mu_J^{(\theta,\theta_0)}$$

en reprenant la notation (II.2.5), le couple (θ, θ_0) *étant le minimum obtenu dans la proposition précédente.*

Réciproquement, $\mu_J^{(\theta,\theta_0)}$ *est la solution du problème* $(PL^{J,\chi})$ *si et seulement si c'est une probabilité telle que* $\int \|f\| d\mu_J^{(\theta,\theta_0)} < +\infty$ *et* $\int f d\mu_J^{(\theta,\theta_0)} \in \partial\chi^*(-\theta)$*, le sous-différentiel de* χ^* *en* $-\theta$.

De façon plus générale pour $(u, u_0) \in \mathbb{R}^d \times \mathbb{R}$*, sous la condition supplémentaire que* $Supp(\mu^f)$ *engendre linéairement* $\mathbb{R}^d$*, la mesure* $\mu_J^{(u,u_0)}$ *est solution usuelle du problème* $(PL^{J,\chi})$ *si et seulement si c'est une probabilité telle que* $\int \|f\| d\mu_J^{(u,u_0)} < +\infty$ *et* $\int f d\mu_J^{(u,u_0)} \in \partial\chi^*(-u)$*, le sous-différentiel de* χ^* *en* $-u$ *et alors* $(u, u_0) = (\theta, \theta_0)$.

Preuve :

Si $\nu \in \mathcal{P}(S)$ est une probabilité telle que $\int \|f\| d\nu < +\infty$, on a la suite d'inégalités pour $(u, u_0) \in \mathbb{R}^d \times \mathbb{R}$:

$$\begin{aligned} \int J\left(\frac{d\nu}{d\mu}\right) d\mu + \chi\left(\int f d\nu\right) &\geq \int (u \cdot f + u_0) \frac{d\nu}{d\mu} d\mu - \int J^*(u \cdot f + u_0) d\mu \\ &\quad + \chi\left(\int f d\nu\right) \\ &\geq u_0 - \int J^*(u \cdot f + u_0) d\mu - \chi^*(-u) \end{aligned} \tag{II.3.11}$$

On prouve la partie directe de la proposition en prenant $(u, u_0) = (\theta, \theta_0)$ dans (II.3.11) : alors, si $\nu \in \mathcal{E}_{J,\chi}$, l'ensemble non vide défini par (II.3.4), est la solution du problème primal et si l'on admet la relation (II.3.10), les inégalités (II.3.11) sont des égalités. Avec la relation (II.0.4) et son cas d'égalité, et en notant que tous les termes de (II.3.11) sont finis, la première inégalité de (II.3.11) n'a lieu avec égalité que si $\nu = \mu_J^{(\theta,\theta_0)}$, ce qui donne la caractérisation attendue de la solution en renvoyant la preuve de la relation (II.3.10) au Lemme II.3.12.

Passons à la réciproque avec $(u, u_0) \in \mathbb{R}^d \times \mathbb{R}$ en faisant l'hypothèse que $\mu_J^{(u,u_0)}$ est une probabilité telle que $\int \|f\| d\mu_J^{(u,u_0)} < +\infty$ et prenons dans la relation (II.3.11) $\nu = \mu_J^{(u,u_0)}$ si bien que grâce au cas d'égalité de la relation (II.0.4), la première inégalité de (II.3.11) est une égalité.

Montrons d'abord que si $\int f d\mu_J^{(u,u_0)}$ est un sous-gradient de χ^* en $-u$, $\mu_J^{(u,u_0)}$ est la solution usuelle.

En effet, avec la relation de conjugaison (II.3.2), la condition de sous-gradient s'écrit de manière équivalente :

$$\chi\left(\int f d\mu_J^{(u,u_0)}\right) = -u \cdot \int f d\mu_J^{(u,u_0)} - \chi^*(-u) \tag{II.3.12}$$

Puisque $\mu_J^{(u,u_0)}$ est une probabilité, cela implique que la seconde inégalité de (II.3.11) est également une égalité, puis en admettant encore la relation (II.3.10), comme (θ, θ_0) est le maximum unique du membre de droite de la dernière inégalité de (II.3.11), on en déduit que $(u, u_0) = (\theta, \theta_0)$ et que $\mu_J^{(u,u_0)} = \mu_J^{(\theta,\theta_0)}$ est la solution usuelle du problème $(PL^{J,\chi})$ car réalisant le minimum de l'entropie généralisée pénalisée, et c'est ce que l'on voulait établir.

Pour voir ensuite la nécessité de la condition de sous-gradient, on suppose que $\mu_J^{(u,u_0)}$ est la solution usuelle du problème $(PL^{J,\chi})$, avec $(u, u_0) \in \mathbb{R}^d \times \mathbb{R}$ et alors on a $\mu_J^{(u,u_0)} = \mu_J^{(\theta,\theta_0)}$ d'après la partie directe de la preuve, d'où $(u, u_0) = (\theta, \theta_0)$ si l'on suppose que $Supp(\mu^f)$ engendre linéairement $\mathbb{R}^d$.
En admettant de nouveau la relation (II.3.10), les deux inégalités (II.3.11) sont réalisées avec des égalités et en particulier, puisque la solution usuelle $\mu_J^{(u,u_0)}$ est nécessairement une probabilité telle que $\int J^*(u \cdot f + u_0) d\mu < +\infty$ d'après (II.3.10), la deuxième égalité de (II.3.11) implique la relation (II.3.12), c'est-à-dire la condition de sous-gradient.

$\blacksquare$

Remarque II.3.4 1. *Dans la réciproque de la proposition précédente, il est en fait inutile de supposer que $\mu_J^{(\theta,\theta_0)}$ est une probabilité car on verra dans la Proposition II.3.8 que ce dernier fait est automatiquement réalisé, comme dans le cas contraint.*

2. *On note alors que si $\mu_J^{(u,u_0)}$ est une probabilité vérifiant $\int \|f\| d\mu_J^{(u,u_0)} < +\infty$, le premier membre de (II.3.11) est fini puisque la première inégalité de (II.3.11) est une égalité tandis que J^* est minoré et que χ est fini.*
On en déduit donc, sous la condition précédente :

$$H_J(\mu_J^{(u,u_0)} \| \mu) + Q_\chi(\mu_J^{(u,u_0)}) < +\infty$$

C'est en particulier vrai pour $\mu_J^{(\theta,\theta_0)}$ dès que l'on vérifie condition d'intégrabilité $\int \|f\| d\mu_J^{(\theta,\theta_0)} < +\infty$

Remarque II.3.5 1. *La condition de sous-gradient dans la partie réciproque de la proposition précédente est une caractérisation de la solution usuelle du problème de calibration pénalisé dans la famille des mesures $\mu_J^{(u,u_0)}$ et est une généralisation de celle donnée dans le cas contraint à la Proposition II.2.4.*
L'hypothèse que $Supp(\mu^f)$ engendre linéairement $\mathbb{R}^d$ est alors naturelle car elle assure que le couple (u, u_0) détermine de manière univoque la mesure $\mu_J^{(u,u_0)}$ et l'on peut s'y ramener, quitte à diminuer la dimension d, c'est-à-dire le nombre de contraintes en supprimant celles qui sont redondantes, à condition toutefois que la conjuguée χ^*, modifiée par cette perte de dimension, conservent les propriétés spécifiées au début de cette partie (c'est-à-dire la stricte convexité et la condition de croissance (II.3.3)).

2. *Si l'on veut cependant se passer de l'hypothèse que $Supp(\mu^f)$ engendre linéairement $\mathbb{R}^d$, la dernière assertion de la proposition II.3.3 devient :*

Si $\mu_J^{(u,u_0)}$ est une probabilité telle que $\int \|f\| d\mu_J^{(u,u_0)} < +\infty$, c'est la solution du problème $(PL^{J,\chi})$ si et seulement s'il existe $v \in Vect^{\perp}(Supp(\mu^f))$, le sous-espace des vecteurs orthogonaux à la direction de la variété affine engendrée $Supp(\mu^f)$, tel que $\int f d\mu_J^{(u,u_0)} \in \partial\chi^*(-(u+v))$, le sous-différentiel de χ^* en $-(u+v)$ et alors $(u, u_0) = (\theta - v, \theta_0 + a \cdot v)$ avec $a \in Supp(\mu^f)$ arbitraire.

Pour voir en effet que la condition de gradient est suffisante, on pose :

$$\begin{cases} w = u + v \\ w_0 = u_0 - a \cdot v \end{cases} \tag{II.3.13}$$

avec v et a tels que spécifiés ci dessus, si bien que $\mu_J^{(w,w_0)} = \mu_J^{(u,u_0)}$ et qu'en particulier la condition de sous-gradient se réécrit $\int f d\mu_J^{(w,w_0)} \in \partial\chi^*(-w)$ et implique que $\mu_J^{(w,w_0)}$ est solution usuelle du problème $(PL^{J,\chi})$ puisque l'on a établi précédemment le caractère suffisant de la condition simple de sous-gradient sans recourir à l'hypothèse sur le support de μ^f.
Réciproquement, si $\mu_J^{(u,u_0)}$ est solution usuelle de $(PL^{J,\chi})$, alors on a déjà établi $\mu_J^{(u,u_0)} = \mu_J^{(\theta,\theta_0)}$ si bien que les relations (II.3.13) sont valables avec $(w, w_0) = (\theta, \theta_0)$ et avec v et a tels que spécifiés ci-dessus et puisque nécessairement $\int f d\mu_J^{(\theta,\theta_0)} \in \partial\chi^*(-\theta)$, d'après la proposition précédente, on en déduit la nécessité de la condition de sous-gradient.

En bornant f, on obtient l'existence de la solution, ce qui est à la base du procédé d'approximation qui établira (II.3.10).

Proposition II.3.6 *Dans le cas où f est bornée, le problème $(PL^{J,\chi})$ admet pour solution $\mu_J^{(\theta,\theta_0)}$.*

Preuve : Il s'agit de vérifier les conditions du critère de la proposition précédente. On montre d'abord que la mesure positive $\mu_J^{(\theta,\theta_0)}$ est une probabilité en appliquant la condition de minimum en θ_0 de la fonction $u_0 \in \mathbb{R} \longrightarrow \int J^*(\theta \cdot f + u_0) d\mu - u_0 + \chi^*(-\theta)$, dérivable sur $\mathbb{R}$ puisque f est bornée :

$$\int d\mu_J^{(\theta,\theta_0)} = 1 \tag{II.3.14}$$

La relation d'intégrabilité $\int \|f\| d\mu_J^{(\theta,\theta_0)} < +\infty$ est ensuite trivialement vérifiée puisque f est bornée et il reste à s'assurer du fait que $\int f d\mu_J^{(\theta,\theta_0)}$ est un sous-gradient de χ^* en $-\theta$.

Pour cela, on applique la condition de minimum en θ de la fonction $u \in \mathbb{R}^d \longmapsto \Lambda(u) + \chi^*(-u)$ en posant pour tout u de $\mathbb{R}^d$ $\Lambda(u) = \int J^*(u \cdot f + \theta_0) d\mu - \theta_0$.
Cela se traduit par le fait que $\mathbf{0}$ est dans le sous-différentiel $\partial(\Lambda + \chi^*(-.))(\theta)$ de la fonction convexe Λ en θ :

$$\mathbf{0} \in \partial(\Lambda + \chi^*(-.))(\theta) = \int f d\mu_J^{\theta,\theta_0} - \partial\chi^*(-\theta) \tag{II.3.15}$$

L'appartenance vient donc avec l'égalité qui est une application du théorème 23.8 de [106] : en effet, compte tenu du fait que Λ est finie sur $\mathbb{R}^d$, puisque f est bornée, on en déduit trivialement que les intérieurs relatifs des domaines de Λ et de $\chi^*(-.)$ ont une intersection non vide, et l'on se trouve dans les conditions d'application de ce théorème. ■

Remarque II.3.7 *Avec la démonstration précédente, on a en fait prouvé que la valeur donnée par (II.3.10) est atteinte pour $\nu = \mu_J^{(\theta,\theta_0)}$, dans le cas où f est bornée.*
La relation (II.3.11) permet alors d'obtenir dans ce cas l'égalité (II.3.10) , admise jusqu'à présent.

II.3.3 Convergence de la solution du problème discret

La convergence du problème discret est un problème de $\widehat{M}$-estimateurs, du point de vue du problème dual (II.3.9).
On énonce le résultat principal :

Proposition II.3.8 *Presque sûrement la solution ν_n du problème $(P_n^{J,\chi})$ converge étroitement vers $\mu_J^{(\theta,\theta_0)}$.*
En particulier, la mesure $\mu_J^{(\theta,\theta_0)}$ est une probabilité.
De manière générale, si $\phi : S \to \mathbb{R}$ est mesurable bornée, on a la convergence presque sûre

$$\int \phi d\nu_n \longrightarrow_{n\to+\infty} \int \phi d\mu_J^{(\theta,\theta_0)}$$

Preuve : Il s'agit d'adapter la preuve de la Proposition II.2.9 en introduisant, avec la Proposition II.3.6, la solution du problème discret

$$\nu_n = J^{*\prime}(\theta_n \cdot f + \theta_{0,n}))\mu_n$$

où $(\theta_n, \theta_{0,n})$ est le minimum de

$$(u, u_0) \in \mathbb{R}^d \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0) d\mu_n - u_0 + \chi^*(-u)$$

Grâce au Lemme II.3.9 prochain, avec $\eta_n = \mu_n$, et la majoration obtenue pour le choix $(u, u_0) = (\mathbf{0}, 0)$

$$\int J^*(\theta_n \cdot f + \theta_{0,n}) d\mu_n - \theta_{0,n} + \chi^*(-\theta_n) \le \int J^*(0) d\mu_n + \chi^*(\mathbf{0}) \le |J^*(0)| + |\chi^*(\mathbf{0})|$$

on en déduit que la suite $(\theta_n, \theta_{0,n})$ est bornée.
Le lemme de Skorokhod et le lemme de Fatou conduisent ensuite, par des développements similaires à la relation (II.2.10), à la convergence presque sûre de $(\theta_n, \theta_{0,n})$ vers (θ, θ_0).
Passons à la convergence étroite, en considérant une fonction continue bornée $\phi : S \to \mathbb{R}$ et en prouvant l'équi-intégrabilité de la suite $(\phi(Y_n) J^{*\prime}((\theta_n \cdot W_n + \theta_{0,n})))$.
En combinant les conditions (II.3.14) et (II.3.15), avec μ_n au lieu de μ, on obtient la relation

$$\int (\theta_n \cdot f + \theta_{0,n}) J^{*'}(\theta_n \cdot f + \theta_{0,n}) d\mu_n \in \theta_{0,n} + \theta_n \cdot \partial\chi^*(-\theta_n) \tag{II.3.16}$$

Mais pour tout a de $\partial\chi^*(-\theta_n)$, on a :

$$\chi^*(\mathbf{0}) - \chi^*(-\theta_n) \ge a \cdot \theta_n$$

ce qui prouve :

$$\sup_{n\in\mathbb{N}, a\in\partial\chi^*(-\theta_n)} a \cdot \theta_n \le \chi^*(\mathbf{0}) - \inf \chi^* = 0$$

Donc, d'après (II.3.16),

$$\int (\theta_n \cdot f + \theta_{0,n}) J^{*'} (\theta_n \cdot f + \theta_{0,n}) d\mu_n \leq \theta_{0,n}$$

L'équi-intégrabilité cherchée s'obtient donc de manière presque identique au cas contraint, en utilisant cette inégalité au lieu de (II.2.12). ∎

Pour que la preuve soit complète, on énonce le résultat suivant, et c'est ici qu'intervient la condition (II.3.3)

Lemme II.3.9 *Soit $(\eta_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{P}(S)$ telle que η_n^f converge étroitement vers μ^f.*
Pour que $(u_n, u_{0,n})$ soit une suite bornée de $\mathbb{R}^d \times \mathbb{R}$, il suffit que la suite

$$\int J^*(u_n \cdot f + u_{0,n}) d\eta_n - u_{0,n} + \chi^*(-u_n) \tag{II.3.17}$$

soit majorée.

Preuve : On suit les grandes lignes de la preuve du Lemme II.2.10 en raisonnant par contraposée et, quitte à se ramener à une sous-suite, en considérant que

$$\|u_n\| + |u_{0,n}| \longrightarrow_{n\to\infty} +\infty$$

et que l'on peut trouver un élément non nul (α, α_0) de $\mathbb{R}^d \times \mathbb{R}$ tel que :

$$\frac{(u_n, u_{0,n})}{\|u_n\| + |u_{0,n}|} \longrightarrow_{n\to\infty} (\alpha, \alpha_0)$$

Il s'agit de prouver que l'expression (II.3.17) diverge vers $+\infty$. Les fonctions J^* et χ^* étant minorées, c'est acquis si $u_{0,n}$ ne l'est pas. Dans le cas contraire, on a :

$$\alpha_0 \geq 0 \tag{II.3.18}$$

Si α est non nul, alors il existe un réel A>0 tel que, pour n assez grand on ait :

$$|u_{0,n}| \leq A\|u_n\|$$

Alors, compte tenu du fait que J^* est minoré, le résultat vient de :

$$\chi^*(-u_n) - u_{0,n} \geq \chi^*(-u_n) - A\|u_n\| \to +\infty$$

en vertu de la condition (II.3.3).
Dans le dernier cas, α est nul et, compte tenu de (II.3.18), on a alors

$$u_{0,n} \to +\infty$$

et

$$\frac{u_n}{u_{0,n}} \to \mathbf{0}$$

Ce cas se traite exactement comme dans la preuve du Lemme II.2.10 en utilisant la convergence étroite de η_n^f, et l'on a :

$$\int J^*(u_n \cdot f + u_{0,n}) d\eta_n - u_{0,n} \to +\infty$$

L'expression (II.3.17) tend une fois de plus vers $+\infty$ et la preuve est achevée. ∎

Citons le résultat suivant, dont la preuve est identique au Corollaire II.2.11 :

Corollaire II.3.10 *Supposons que J^* soit de classe C^2.*
Si $\phi : S \to \mathbb{R}$ est une fonction mesurable telle que

$$\int |\phi| d\mu_J^{(\theta,\theta_0)} < +\infty \tag{II.3.19}$$

et s'il existe un voisinage Δ de (θ, θ_0) dans $\mathbb{R}^d \times \mathbb{R}$ tel que :

$$\int \sup_{(u,u_0)\in\Delta} J^{*\prime\prime}(u \cdot f + u_0) \max(1, \|f\|)|\phi| d\mu < +\infty \tag{II.3.20}$$

alors on a presque sûrement la convergence

$$\int \phi d\nu_n \longrightarrow_{n\to+\infty} \int \phi d\mu_J^{(\theta,\theta_0)}$$

De plus, par une suite d'inégalités semblable à (II.2.19), on en déduit la convergence

Corollaire II.3.11 *On a la convergence presque sûre :*

$$\int J^*(\theta_n \cdot f + \theta_{0,n}) d\mu_n \longrightarrow_{n\to\infty} \int J^*(\theta \cdot f + \theta_0) d\mu$$

Avec la convergence de la suite $(\theta_n, \theta_{0,n})$, on a en fait montré que la valeur minimale associée au problème dual discret $\int J^*(\theta_n \cdot f + \theta_{0,n}) d\mu_n - \theta_{0,n} + \chi^*(-\theta_n)$ converge vers la valeur minimale associée au problème dual limite.
Afin de faire le lien avec le problème primal, on peut prouver la relation (II.3.10) exactement comme dans le Lemme II.2.13.

Lemme II.3.12 *On a l'égalité*

$$\inf_{\nu\in\mathcal{P}(S)} (H_J(\nu\|\mu) + Q_\chi(\nu)) = -\left(\int J^*(\theta \cdot f + \theta_0) d\mu - \theta_0 + \chi^*(-\theta)\right)$$

On en déduit, avec ce qui précède

$$\inf_{\nu\in\mathcal{P}(S)} (H_J(\nu\|\mu_n) + Q_\chi(\nu)) \longrightarrow_{n\to\infty} \inf_{\nu\in\mathcal{P}(S)} (H_J(\nu\|\mu_n) + Q_\chi(\nu))$$

Avec des considérations similaires qui ont conduit au Lemme II.2.14, on peut également prouver :

$$\inf_{\nu\in\mathcal{P}(S),\nu\sim\mu} (H_J(\nu\|\mu) + Q_\chi(\nu)) = \inf_{\nu\in\mathcal{P}(S)} (H_J(\nu\|\mu) + Q_\chi(\nu)) \tag{II.3.21}$$

qui permet de régler l'étude du problème suivant :

$$(PL^{J,\chi}) \quad \text{trouver } \nu \in \mathcal{P}(S),\ \nu \sim \mu \text{ qui minimise } H_J(\nu\|\mu) + Q_\chi(\nu)$$

On énonce en effet

Lemme II.3.13 *La relation (II.3.21) est valide.*
Le problème $(PL^{J,\chi})$ admet une solution si et seulement s'il en est de même du problème $(PL^{J,\chi})$ et alors les solutions sont identiques.

La preuve se calque sur celle du lemme II.2.14. En particulier, pour adapter la suite d'inégalités (II.2.26), on utilise la convexité, le caractère sci de la fonction χ et le fait qu'elle soit finie en $\mathbf{0}$.

II.4 Calibration pénalisée sous une contrainte de type ensemble

Pour les besoins d'une étude ultérieure, on envisage un problème de calibration mixte où l'on sépare les contraintes en envisageant la calibration exacte par rapport à une partie d'entre elles et en l'effectuant de manière pénalisée par rapport au reste. De manière plus générale, les contraintes exactes sont remplacées par l'assujettissement à rester dans un ensemble convexe, ce qui est une extension naturelle du cas contraint déjà envisagé si bien qu'en gardant le cadre Monte-Carlo habituel, il s'agit de résoudre

$$(P_n^{J,\chi,K}) \quad \text{trouver } \nu_n \in \mathcal{P}(S) \text{ qui minimise } H_J(\nu_n \| \mu_n) + \chi\Big(\int f_2 d\nu_n\Big) \text{ sous la contrainte } \int f_1 d\nu_n \in K$$

où $(f_1, f_2) = f$ est l'écriture du payoff en deux composantes vectorielles f_1 et f_2 à valeurs dans $\mathbb{R}^{d_1}$ et $\mathbb{R}^{d_2}$ avec $d_1 + d_2 = d$, K est un ensemble convexe fermé de $\mathbb{R}^{d_1}$ et $\chi : \mathbb{R}^{d_2} \to \mathbb{R}$ est une fonction strictement convexe, sci et minimale en $\mathbf{O}$ et ayant en outre toutes les autres propriétés énoncées dans la partie concernant la pénalisation.
De façon formelle, on peut se ramener au cas pénalisé en introduisant la fonction indicatrice (au sens de Rockafellar) de K $\delta(.|K)$, définie pour tout x de $\mathbb{R}^{d_1}$ par :

$$\delta(x|K) = \begin{cases} 0 & \text{si } x \in K \\ +\infty & \text{sinon} \end{cases}$$

Le problème contraint s'identifie alors au problème pénalisé avec $(x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \mapsto \delta(x_1|K) + \chi(x_2)$ comme fonction de pénalisation.
Cependant, comme dans le cas contraint élémentaire, on doit imposer une condition de nature géométrique pour avoir une solution à partir d'un certain rang :

Lemme II.4.1 *Pour que* $\mathbb{P}(\exists n \geq 1 \text{ et } \eta_n \in \mathcal{P}(S) \text{ t.q. } H_J(\eta_n \| \mu_n) < +\infty \text{ et } \int f_1 d\eta_n \in K) > 0$, *il faut qu'il existe un sous-espace affine* E *de* $\mathbb{R}^{d_1}$ *tel que*

$$(C_K) \begin{cases} K \bigcap Int_E(Conv(Supp(\mu^{f_1}_{|E}))) \neq \emptyset \\ \forall \eta \in \mathcal{P}(S) \text{ absolument continue par rapport à } \mu \text{ t.q. } \int \|f_1\| d\eta < +\infty \text{ et } \int f_1 d\eta \in K, \\ \text{on a } \eta(f_1^{-1}(E)) = 1 \end{cases}$$

Preuve : Il suffit d'adapter la preuve du Lemme II.2.1 et on se contente ainsi de donner une idée de la démonstration. On part d'une probabilité $\eta_n = \sum_{i=1}^n p_i \delta_{X_i}$ telle que $\int f_1 d\eta_n \in K$ et on construit E par récurrence descendante sur la dimension d'un sous espace affine F de $\mathbb{R}^{d_1}$ que l'on initialise en posant $F = \mathbb{R}^{d_1}$.
On suppose donc que $\eta_n = \sum_{i=1}^n 1_F(X_i) p_i \delta_{X_i}$ p.s. ce qui est forcément vérifié à l'initialisation si le problème admet une solution. Comme p.s. $Conv(\{f_1(X_i),\ i \in [1,n]\} \cap F) \subset Conv(Supp(\mu^{f_1}_{|F}))$, de $\int f_1 d\eta_n \in K$ on déduit $K_F \equiv K \bigcap Conv(Supp(\mu^{f_1}_{|F})) \neq \emptyset$.

- Si $K \bigcap Int_F(Conv(Supp(\mu^{f_1}_{|F}))) \neq \emptyset$, on pose $E = F$.

- Sinon, il existe dans F un hyperplan affine d'appui H de $Conv(Supp(\mu^{f_1}_{|F}))$ contenant K_F.
  Alors, puisque p.s. le support de $\eta_n^{f_1}$ est inclus dans celui de $\mu^{f_1}_{|F}$, par l'hypothèse de récurrence, on a :
  $$\int f_1 d\eta_n \in K \Rightarrow \int f_1 d\eta_n \in K_F \Rightarrow Supp(\eta_n^{f_1}) \subset H$$
  la dernière implication se validant par une équation de l'hyperplan relatif H.
  On poursuit alors la récurrence avec H au lieu de F.

De manière identique, on montre que $\eta(f_1^{-1}(E)) = 1$ si η est une probabilité absolument continue par rapport à μ t.q.$\int_S f_1 d\eta \in K$ ∎

Pour donner les résultats d'existence de la solution et de convergence de cette dernière, on introduit bien sûr le problème limite :

$$(PL^{J,\chi,K}) \quad \text{trouver } \nu \in \mathcal{P}(S) \text{ qui minimise } H_J(\nu\|\mu) + \chi\left(\int f_2 d\nu \right) \text{ sous la contrainte } \int f_1 d\nu \in K$$

En reprenant les notations de la section précédente, le problème limite est donc $(PL^{J,\delta(.|K)+\chi})$ et il s'agit de minimiser sans contrainte le critère sur $\nu \in \mathcal{P}(S)$ $H_J(\nu\|\mu) + Q_{\delta(.|K)+\chi}(\nu)$ en rappelant

$$Q_{\delta(.|K)+\chi}(\nu) = \begin{cases} \delta\left(\int f_1 d\nu | K\right) + \chi\left(\int f_2 d\nu\right) & \text{si } \int_S \|f\| d\nu < +\infty \\ +\infty & \text{sinon} \end{cases}$$

On est alors conduit à résoudre le problème en introduisant la fonction χ^* conjuguée de χ ainsi que la fonction support $\delta^*(|K)$ du convexe K, conjuguée de $\delta(.|K)$, et définie pour tout u de $\mathbb{R}^{d_1}$ par

$$\delta^*(u|K) = \sup_{x \in K} u \cdot x$$

On suppose alors que le convexe K est fermé, ce qui assure que la fonction indicatrice convexe $\delta(.|K)$ est sci et qu'elle est donc la conjuguée de la fonction support $\delta^*(|K)$ par le théorème de Fenchel-Moreau.
Sous l'hypothèse (C_K), quitte à remplacer f_1 par $f_1 - \mathbf{C}_1$ où $\mathbf{C}_1 \in Int_E(Conv(Supp(\mu^{f_1}_{|E}))) \bigcap K$, on ne perd pas de généralité en supposant que K contient $\mathbf{0}$ et que soit ainsi satisfaite la condition :

$$(C'_K) \begin{cases} \mathbf{0} \in Int_E(Conv(Supp(\mu^{f_1}_{|E}))) \bigcap K \\ \forall \eta \in \mathcal{P}(S) \text{ absolument continue par rapport à } \mu \; t.q. \; \int \|f_1\| d\eta < +\infty \text{ et } \int f_1 d\eta \in K, \\ \text{on a } \eta(f_1^{-1}(E)) = 1 \end{cases}$$

En particulier, du fait que K contient $\mathbf{0}$, on s'assure que la fonction support est positive. En résumé, c'est sous la condition (C'_K) et en supposant que K est fermé que l'on étudie le problème contraint.
D'abord, en reprenant la notation (II.3.4), on vérifie la non vacuité de l'ensemble

$$\mathcal{E}_{J,\delta(.|K)+\chi} = \{\nu \in \mathcal{P}(S), \; H_J(\nu\|\mu) + Q_{\delta(.|K)+\chi}(\nu) < +\infty\}$$

En effet, d'une part la relation (C'_K) reste valable en remplaçant μ par

$$\mu_a = \frac{1_{\{\|f_2\| \leq a\}}}{\mu(\{\|f_2\| \leq a\})}\mu$$

si a est un réel positif assez grand, ce qui permet de se ramener au cas où f_2 est bornée et donc où le terme de pénalisation $\chi\left(\int f_2 d\mu \right)$ est fini.
Alors, puisque $\mathbf{0} \in K$, l'ensemble $\mathcal{E}_{J,\delta(.|K)} + \chi$ contient $\{\nu \in \mathcal{C}, \; H_J(\nu\|\mu)) < +\infty\}$ qui est non vide puisque la propriété (C'_K) permet d'appliquer les résultats du cas contraint élémentaire avec en particulier $H_J(\mu\|\mathcal{C}) < +\infty$. On voit donc que la condition (C'_K) permet de se passer de l'hypothèse de finitude de la fonction de pénalisation faite dans la section précédente. Il est également immédiat que si la loi a priori μ vérifie les contraintes, c'est la solution du problème.

Lemme II.4.2 *Si μ est un élément de $\mathcal{C}_K$, alors c'est la solution du problème $(PL^{J,K,\chi})$.*

en notant

$$\mathcal{C}_K = \left\{ \nu \in \mathcal{P}(S),\ \nu \ll \mu,\ \int \|f\| d\nu < +\infty,\ \int f_1 d\nu \in K \text{ et } \int f_2 d\mu = \mathbf{0} \right\} \tag{II.4.1}$$

Maintenant, le problème dual peut s'écrire comme la minimisation de la fonction à valeurs dans $]-\infty,+\infty]$

$$(u_1,u_2,u_0) \in E \times \mathbb{R}^{d_2} \times \mathbb{R} \longmapsto \int J^*(u_1 \cdot f_1 + u_2 \cdot f_2 + u_0) 1_E(f_1) d\mu - u_0 + \delta^*(-u_1|K) + \chi^*(-u_2) \tag{II.4.2}$$

La fonction support $\delta^*(.|K)$, conjuguée de la fonction de pénalisation, jouit des propriétés énoncées dans la section précédente à l'exception de la stricte convexité et de la condition (II.3.3). On va voir que la condition (C'_K) permet encore de s'en passer. En effet, avec cette condition, le premier terme du second membre de (II.4.2) est strictement convexe, ce qui assure l'unicité du minimum sans requérir la stricte convexité de la fonction support. Pour le résultat d'existence, on prouve le lemme suivant :

Lemme II.4.3 *Soit $(\eta_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{P}(S)$ telle que $\eta^f_{n|E\times\mathbb{R}^{d_2}}$ converge étroitement vers $\mu^f_{|E\times\mathbb{R}^{d_2}}$.*
Pour que $(u_{1,n},u_{2,n},u_{0,n})$ soit une suite bornée de $E\times\mathbb{R}^{d_2}\times\mathbb{R}$, il suffit que la suite

$$\int J^*(u_{1,n} \cdot f_1 + u_{2,n} \cdot f_2 + u_{0,n}) 1_{E\times\mathbb{R}^{d_2}}(f) d\eta_n - u_{0,n} + \delta^*(-u_{1,n}|K) + \chi^*(-u_{2,n}) \tag{II.4.3}$$

soit majorée.

En effet, avec en particulier le fait que $\delta^*(.|K)$ est minorée car positive et avec la condition (C'_K) au lieu de (C), la preuve se fait en reprenant celle du Lemme II.2.10 et du Lemme II.3.9. A ces corrections près les techniques du cas pénalisé sont valables dans ce cas et l'on énonce alors ;

Proposition II.4.4 *Sous l'hypothèse (C'_K), la fonction*

$$(u_1,u_2,u_0) \in E \times \mathbb{R}^{d_2} \times \mathbb{R} \longmapsto \int J^*(u_1 \cdot f_1 + u_2 \cdot f_2 + u_0) 1_E(f_1) d\mu - u_0 + \delta^*(-u_1|K) + \chi^*(-u_2)$$

atteint son minimum en un point unique $(\theta_1,\theta_2,\theta_0)$.

Ensuite, en adaptant l'ensemble des Propositions II.3.3 et II.3.6, on énonce :

Proposition II.4.5 *Supposons vérifiée la condition (C'_K).*
On a

$$\inf_{\nu\in\mathcal{P}(S)} (H_J(\nu\|\mu) + Q_{\delta(.|K)+\chi}(\nu)) \ = \tag{II.4.4}$$

$$-\left(\int J^*(\theta_1 \cdot f_1 + \theta_2 \cdot f_2 + \theta_0) 1_E(f_1) d\mu - \theta_0 + \delta^*(-\theta_1|K) + \chi^*(-\theta_2) \right) + J(0)\mu(f_1^{-1}(E^c)) \tag{II.4.5}$$

Si le problème $(PL^{J,K,\chi})$ admet une solution, celle ci est

$$\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$$

en reprenant la notation (II.2.5).
De plus, sous la condition supplémentaire que $Supp(\mu^{f_2})$ engendre linéairement $\mathbb{R}^{d_2}$, alors pour $(u_1,u_2,u_0) \in E\times\mathbb{R}^{d_2}\times\mathbb{R}$, $\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J}$ est solution usuelle de $(PL^{J,K,\chi})$ si et seulement si c'est une probabilité telle que $\int \|f\| d\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J} < +\infty$, $\int f_1 d\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J} \in \partial\delta^*(-u_1|K)$ et $\int f_2 d\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J} \in \partial\chi^*(-u_2)$, et alors $(u_1,u_2,u_0) = (\theta_1,\theta_2,\theta_0)$.
Dans le cas où f est bornée, le problème $(PL^{J,K,\chi})$ admet effectivement la solution $\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$.

La preuve de la Proposition II.3.3 s'adapte bien à la proposition précédente avec $\delta(.|K)+\chi$ comme fonction de pénalisation puisque dans cette preuve l'on n'utilise le fait que $\chi\left(\int f d\nu\right)$ est fini que dans la partie directe où cela est réalisé car on suppose alors que ν est solution usuelle du problème de calibration.
Notons aussi que la précaution prise avant la condition de sous-gradient implique que le triplet de paramètres $(u_1,u_2,u_0)\in E\times\mathbb{R}^{d_2}\times\mathbb{R}$ détermine de manière univoque la mesure $\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$.

Remarque II.4.6 *1. En anticipant une nouvelle fois sur la suite, on peut ajouter qu'en prenant $(u_1,u_2,u_0)=(\theta_1,\theta_2,\theta_0)$ dans la proposition précédente, il est automatiquement vérifié que $\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$ est une probabilité.*

2. On peut également reprendre les arguments du numéro 2 de la Remarque II.3.4 et énoncer que si $\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J}$ est une probabilité telle que $\int\|f\|d\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J}<+\infty$, alors

$$H_J(\mu^{(u_1,u_2,u_0)}_{E\times\mathbb{R}^{d_2},J}\|\mu)<+\infty$$

et qu'en particulier, cette relation est vraie pour la solution généralisée $\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$ sous la seule condition $\int\|f\|d\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}<+\infty$.

Voici également une conséquence intéressante lorsque l'on envisage le problème de calibration purement contraint :

Corollaire II.4.7 *Supposons $d_2=0$ (donc $d_1=d$) ainsi que $E=\mathbb{R}^d$.*
Si le problème $(PL^{J,K,\chi})$ admet une solution distincte de μ, alors $\int f d\mu_J^{(\theta_1,\theta_0)}\in FrK$.

Cet énoncé signifie qu'en absence de pénalisation, si μ ne vérifie pas les contraintes et est équivalente à la solution calibrée, alors les prix des options de calibration obtenus sous cette solution se situent à la frontière de l'ensemble de contraintes.
Preuve : Remarquons d'abord qu'avec l'hypothèse $E=\mathbb{R}^d$ et la condition (C'_K), on déduit que K est d'intérieur non vide, ce qui rend l'énoncé non trivial.
La condition $\int f d\mu_J^{(\theta_1,\theta_0)}\in\partial\delta^*(-\theta|K)$ équivaut à $-\theta_1\in\partial\delta\left(\int f d\mu_J^{(\theta_1,\theta_0)}|K\right)$: c'est une conséquence du Théorème 23.5 de [106]. Ainsi, $\int f d\mu_J^{(\theta_1,\theta_0)}$ appartient au domaine K de $\delta(.|K)$ et de plus ne se trouve à l'intérieur de K que si θ_1 est nul, donc que si $\mu_J^{(\theta_1,\theta_0)}$ vaut μ. Comme, on exclut ce cas, on en déduit que l'intégrale $\int f d\mu_J^{(\theta_1,\theta_0)}$ se trouve bien sur la frontière de K. ∎

Remarque II.4.8 *De manière plus générale, en absence de pénalisation, si le problème $(PL^{J,K})$ admet une solution distincte de $\frac{\mu_{f^{-1}(E)}}{\mu(f^{-1}(E))}$, alors $\int f d\mu_{E,J}^{(\theta_1,\theta_0)}\in Fr_E K\bigcap E$, la frontière relative de $K\bigcap E$ dans E.*

Les résultats de convergence s'adaptent enfin parfaitement, avec principalement l'énoncé suivant :

Proposition II.4.9 *Sous la condition (C_K), presque sûrement, il existe un rang N à partir duquel le problème $(P_n^{J,K,\chi})$ admet une solution ν_n et la suite $(\nu_n)_{n\geq N}$ converge étroitement vers $\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$.*
En particulier, la mesure $\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$ est une probabilité.
De manière générale, si $\phi:S\to\mathbb{R}$ est mesurable bornée, on a la convergence presque sûre

$$\int\phi d\nu_n\longrightarrow_{n\to+\infty}\int\phi d\mu^{(\theta_1,\theta_2,\theta_0)}_{E\times\mathbb{R}^{d_2},J}$$

Naturellement, on pourrait aussi introduire le problème discret

$$(P_{e,n}^{J,K,\chi}) \quad \begin{cases} \text{trouver } \nu_n \in \mathcal{P}(S),\ \nu_n \sim \mu_n \text{ qui minimise } H_J(\nu_n\|\mu_n) + \chi\Big(\int f_2 d\nu_n\Big) \\ \text{sous la contrainte } \int f_1 d\nu_n \in K \end{cases}$$

auquel on associe le problème limite

$$(PL_e^{J,K,\chi}) \quad \begin{cases} \text{trouver } \nu \in \mathcal{P}(S),\ \nu \sim \mu \text{ qui minimise } H_J(\nu\|\mu) + \chi\Big(\int f_2 d\nu\Big) \\ \text{sous la contrainte } \int f_1 d\nu \in K \end{cases}$$

Une condition nécessaire à l'existence de la solution du problème discret est la condition (C_K) avec pour E le sous-espace affine engendré par $Supp(\mu^{f_1})$:

$$(C_{e,K}) \ \ K \bigcap ri(Conv(Supp(\mu^{f_1}))) \neq \emptyset$$

où ri désigne l'intérieur relatif.

Lemme II.4.10 *Sous la condition $(C_{e,K})$, la quantité*

$$\inf_{\nu\sim\mu}(H_J(\nu\|\mu) + Q_{\delta(.|K)+\chi}(\nu))$$

est donnée par la relation (II.4.5).
Le problème $(PL_e^{J,K})$ admet une solution si et seulement si c'est le cas du problème $(PL^{J,K})$ et alors leurs solutions sont identiques.

Il suffit par exemple de suivre les considérations faites pour l'identité (II.3.21) en notant en particulier que l'on s'est ramené à $\chi(\mathbf{0})+\delta(\mathbf{0}|K)$ fini, ce qui permet d'adapter la suite d'inégalités (II.2.26).

Remarque II.4.11 *Notons que l'on pourrait remplacer la fonction indicatrice $\delta(.|K)$ par une fonction convexe ψ sci de domaine K et minorée.*
les résultats précédents sont valables à l'exception en général du Lemme II.4.2 ainsi que du Corollaire II.4.7 et de la Remarque II.4.8.
On peut cependant adapter le Lemme II.4.2 en énonçant que si le terme de pénalisation global $\psi+\chi$ atteint son minimum au prix des options sous la mesure a priori μ, alors cette dernière est la solution du problème limite.

II.5 Compléments : généralisation du choix de l'entropie dans le cas où la fonction de payoff est bornée

On a supposé jusqu'à présent que la fonction convexe J vérifie les propriétés données au début de ce chapitre. Un tel choix permet d'obtenir des résultats de convergence sans que la solution du problème limite existe, en particulier sans supposer que la fonction de payoff f est bornée.
On peut cependant regretter d'avoir eu à imposer la condition

$$\lim_{u\to 0+} J'(u) = -\infty \tag{II.5.1}$$

ce qui ne permet pas de considérer par exemple le cas de normes L^p avec $p>1$ où l'on prend $J(u) = \begin{cases} u^p & \text{si } u \geq 0 \\ +\infty & \text{sinon} \end{cases}$.

Cette condition assure la convexité stricte de J^* alors que l'on va montrer que l'on peut s'en passer si f est bornée, si on suppose toujours que J est une fonction sci de domaine $[0,+\infty[$ où elle est strictement convexe, et qu'elle vérifie la condition :

$$\lim_{u\to+\infty} \frac{J(u)}{u} = +\infty \tag{II.5.2}$$

D'abord, avec le Théorème 26.3 de [106] et la stricte convexité de J, sa convexe conjuguée J^* est "essentiellement régulière" selon la définition donnée dans la première section du chapitre. Elle est donc en particulier dérivable sur l'intérieur non vide de son domaine.
Mais le Lemme 4.2 de [23] et l'hypothèse (II.5.2) impliquent que $domJ^* =]-\infty,+\infty[$ tandis qu'avec le fait que J est fini au voisinage de $+\infty$ la condition (II.0.2) est automatiquement vérifiée :

$$\lim_{u\to+\infty} J^{*\prime}(u) = +\infty \tag{II.5.3}$$

De plus, puisque J est la conjuguée de la fonction dérivable J^* et puisqu'elle est égale à $+\infty$ sur $]-\infty,0[$, on en déduit que $J^{*'} \geq 0$. On obtient ainsi un jeu de conditions plus faible que celui examiné au début de ce chapitre et l'on va montrer que les résultats principaux restent valables dans le cas où f est bornée.

II.5.1 Problème contraint

On commence par examiner le cas contraint de base en supposant que la condition géométrique (C) est vérifiée. Dans le cas où le vecteur de payoff f est borné, on peut alors remplacer les hypothèses discutées au début du chapitre par les hypothèses ci-dessus. Ainsi qu'on l'a remarqué plus haut, sans la condition (II.5.1) on perd la stricte convexité de J^* et donc l'unicité de la solution du problème de calibration dual. En fait, cela n'empêche pas de résoudre le problème primal.
En effet, en utilisant seulement le fait que J est propre sci et strictement convexe, le Théorème 4.8 de [23] permet d'affirmer que la Proposition II.2.6 est valable , à condition toutefois de vérifier la condition de qualification des contraintes du Corollaire 2.6 de [23]. Suivant ce dernier résultat, puisque $]0,+\infty[\subset domJ$, cela revient à s'assurer l'existence de $\nu \in \mathcal{C}$ telle que $\nu \sim \mu 1_E(f)$ et $\int J(\frac{d\nu}{d\mu}) 1_E(f) d\mu < +\infty$.
Or, sous la condition géométrique (C), l'existence d'une telle probabilité ν a déjà été obtenue sous la forme de la solution $\mu^{(u,u_0)}_{E,J_0}$, à densité bornée, du problème de calibration correspondant au choix de l'entropie relative pour J_0, ou tout autre critère convexe satisfaisant hypothèses générales requises, y compris (II.5.1).
On énonce donc :

Proposition II.5.1 *Supposons que J soit strictement convexe sur son domaine $[0,+\infty[$, sci et que de plus les conditions (II.5.2) et (II.0.2) tiennent.*
Dans le cas où f est bornée, et sous la condition (C), le problème (PL^J) admet pour solution $\mu^{(\theta,\theta_0)}_{E,J}$ où le critère dual (II.2.2) est minimal en (θ,θ_0).

La résolution du problème discret repose sur le résultat précédent l'obstacle majeur à la convergence étant encore une fois la non unicité de la solution du problème dual. On va voir que cet obstacle n'est qu'apparent et l'on introduit la solution discrète

$$\nu_n = J^{*'}(\theta_n \cdot f + \theta_{0,n}) 1_E(f) \mu_n$$

où $(\theta_n, \theta_{0,n})$ minimise sur $E \times \mathbb{R}$ la fonction

$$(u, u_0) \in E \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0) 1_E(f) d\mu_n - u_0$$

Pour établir la convergence de ν_n, on reprend la preuve de la Proposition II.2.9 en supposant que le vecteur de payoff f est borné afin d'assurer l'existence de la solution du problème limite primal (PL^J), en vertu de la Proposition II.5.1.
La non-unicité de $(\theta_n, \theta_{0,n})$ empêche la convergence de ce dernier mais comme on peut montrer que le Lemme II.2.10 est vrai car on dispose de la relation (II.0.2) si $domJ^*$ n'est pas majoré, on en déduit que la suite $(\theta_n, \theta_{0,n})$ est bornée.
En fixant une épreuve ω de Ω en dehors d'un ensemble négligeable, on considère une sous-suite (ν_{i_n}) extraite de (ν_n) et dont on va montrer qu'il existe une sous-suite qui converge vers la solution $\mu_{E,J}^{(\theta,\theta_0)}$ du problème limite.
D'abord, en adaptant la preuve de la Proposition II.2.9, on montre que toute valeur d'adhérence $(\tilde{\theta}, \tilde{\theta}_{0,n})$ de la suite $(\theta_{i_n}, \theta_{0,i_n})$ minimise $\int J^*(u \cdot f + u_0) 1_E(f) d\mu - u_0$ et donc, une nouvelle fois avec la Proposition II.2.6, on a $\mu_{E,J}^{(\tilde{\theta},\tilde{\theta}_0)} = \mu_{E,J}^{(\theta,\theta_0)}$ l'unique solution usuelle du problème de calibration limite.
Quitte à prendre une sous-suite de (i_n), on peut supposer que $(\theta_{i_n}, \theta_{0,i_n})$ converge vers $(\tilde{\theta}, \tilde{\theta}_{0,n})$ et, en reprenant l'argument d'équi-intégrabilité de la preuve de la Proposition II.2.9, la suite $\nu_{i_n}(\omega)$ converge étroitement vers $\mu_{E,J}^{(\tilde{\theta},\tilde{\theta}_0)} = \mu_{E,J}^{(\theta,\theta_0)}$ et cela prouve la convergence souhaitée puisque l'on est parti d'une sous-suite quelconque.
On a donc prouvé :

Proposition II.5.2 *Si J vérifie les hypothèses de la Proposition II.5.1, sous la condition (C) et en supposant que le payoff f est borné, presque sûrement, il existe un rang N à partir duquel le problème (P_n^J) admet une solution ν_n et la suite $(\nu_n)_{n \geq N}$ converge étroitement vers $\mu_{E,J}^{(\theta,\theta_0)}$.*
En particulier, la mesure $\mu_{E,J}^{(\theta,\theta_0)}$ est une probabilité.
De manière générale, si $\phi : S \to \mathbb{R}$ est mesurable bornée, on a la convergence presque sûre

$$\int \phi d\nu_n \longrightarrow_{n \to +\infty} \int \phi d\mu_{E,J}^{(\theta,\theta_0)}$$

Remarque II.5.3 *Remarquons que la proposition précédente est encore valable si f n'est pas bornée à condition de supposer que le problème limite primal (PL^J) admet une solution. Pour se passer de cette hypothèse, qui n'est pas systématiquement réalisée, et pour affirmer la convergence de la suite (ν_n), il faut en principe imposer la convexité stricte de J^* qui implique l'unicité de la solution (θ, θ_0) du problème limite dual : en effet, la suite $(\theta_n, \theta_{0,n})$ converge alors vers son unique valeur d'adhérence (θ, θ_0) puis, avec l'argument d'équi-intégrabilité, ν_n converge vers $\mu_{E,J}^{(\theta,\theta_0)}$.*
Mais, avec le Théorème 26.3 de [106] ainsi que la convexité stricte de J sur son domaine, on en déduit que J est du "type de Legendre", ce qui nous ramène aux conditions discutées dans la première section du chapitre au fait près que $Int(domJ^*) =]-\infty, \lim_{u \to +\infty} J'(u)[$ au lieu d'être $\mathbb{R}$. Cela constitue quand même une généralisation des résultats obtenus lorsque f n'est pas bornée, avec en particulier l'absence de saut de dualité (II.2.4) car la preuve du Lemme II.2.13 est inchangée : on a un résultat qui sort du cadre d'application des résultats classiques (Corollaire 4.8 de [24] et Corollaire 2.6 de [23]) et c'est une généralisation intéressante.
Evidemment, dans ce cas la relation (II.5.1) tient comme conséquence de la régularité essentielle de J et exclut donc d'adopter la norme L^p dans le cas où f n'est pas bornée.

II.5.2 Pénalisation du problème

On reprend ici le problème pénalisé limite $(P^{J,\chi})$ ainsi que sa version discrétisée $(P^{J,\chi})_n$ où J a les propriétés données précédemment tandis que χ a les propriétés précisées au début de la

deuxième partie de ce chapitre à l'exception du fait que χ^* n'est plus supposé strictement convexe sur son domaine.
Pour obtenir la Proposition II.3.3 sous ces hypothèses faibles, dans le cas où f est bornée, il faut adapter le Théorème 4.8 de [23] au cas d'un critère pénalisé.
Avec le Théorème 4.2 de [24], compte tenu du fait que χ est finie sur $\mathbb{R}^d$, la relation (II.3.10) est valable, c'est-à-dire que l'on a absence de saut de dualité et qu'en plus le problème dual admet une solution (θ, θ_0) (non unique a priori car ni J^* ni χ^* ne sont supposés strictement convexes). Dans ces conditions, la Proposition II.3.6 est encore valable :

Proposition II.5.4 *Sous les hypothèses précédentes sur J et χ, dans le cas où f est bornée, la relation (II.3.10) est vérifiée et le problème $(PL^{J,\chi})$ admet pour solution $\mu_J^{(\theta,\theta_0)}$.*

Preuve : On vient de justifier l'absence de saut de dualité et, pour établir le résultat qui reste, on reprend avec de légères modifications la preuve du Théorème 4.8 de [23]).
Puisque f est bornée, l'application $A : \varphi \in L^1(\mu) \mapsto \int (f,1)\varphi d\mu \in \mathbb{R}^{d+1}$ est continue, et on la prolonge avec une opération de double adjonction par $A^{TT} : L^{\infty'}(\mu) \to \mathbb{R}^{d+1}$.
Présentée de façon approximative, l'idée est alors de résoudre le problème dans $L^{\infty'}(\mu)$ puis de vérifier que la solution obtenue est en fait un élément de $L^1(\mu)$.
En utilisant le Théorème 19 de [107], la condition d'optimalité (II.3.15) de (θ, θ_0) s'écrit, en notant $I_{J^*} : \varphi \in L^\infty(\mu) \mapsto \int J^*(\varphi) d\mu$:

$$\mathbf{0} \in A^{TT} \partial I_{J^*}(A^T(\theta, \theta_0)) - \partial \chi^*(-\theta) - (\mathbf{0}_{\mathbb{R}^d}, 1)$$

si bien qu'il existe $\overline{\nu} \in \partial I_{J^*}(A^T(\theta, \theta_0)) \subset L^{\infty'}(\mu)$ tel que

$$A^{TT}(\overline{\nu}) \in \partial \chi^*(-\theta) + (\mathbf{0}_{\mathbb{R}^d}, 1) \tag{II.5.4}$$

Avec le Corollaire 1B de [105], il existe $\varphi \in L^1(\mu)$ ainsi qu'une composante singulière $\nu_* \in L^{\infty'}$ $\overline{\nu} = \varphi + \nu_*$ avec $\varphi \in \partial J^*(A^T(\theta, \theta_0))$ p.s. et où le maximum de ν_* sur $domI_{J^*}$ est atteint en $A^T(\theta, \theta_0)$.
Mais d'après le Lemme 4.3 de [23] ainsi que la condition (II.5.2), I_{J^*} est continue en $A^T(\theta, \theta_0)$ et donc $A^T(\theta, \theta_0) \in Int(domI_{J^*})$: on en déduit $\nu_* = 0$ en utilisant le Corollaire 2C ainsi que le Théorème 2 de [105].
De plus $A^T(\theta, \theta_0) \in Int(domJ^*)$ p.s. et, puisque J^* est essentiellement régulière, il vient que $\varphi = J^{*'}(A^T(\theta, \theta_0))$ p.s. et en particulier $\varphi \geq 0$ car on peut voir facilement que J^* est croissante. La condition d'optimalité s'écrit alors

$$A\varphi = \int (f,1)\varphi \in \partial \chi^*(-\theta) + (\mathbf{0}_{\mathbb{R}^d}, 1)$$

ce qui achève de montrer que $\varphi\mu$ est une probabilité tandis qu'avec l'expression de φ on a la condition

$$\int f J^{*'}(\theta \cdot f + \theta_0) \in \partial \chi^*(-\theta)$$

et, avec l'égalité (II.3.10), comme dans la preuve de la Proposition II.3.3, l'examen de la double inégalité (II.3.11) (avec le Théorème 23.5 de [106]) démontre que la condition de sous-gradient précédente implique que $\mu_J^{(\theta,\theta_0)}$ est la solution usuelle de $(PL^{J,\chi})$. ■

Remarque II.5.5 *Afin de résoudre le problème primal, on n'a pas eu besoin de l'étendre mais cette voie peut aboutir de manière rigoureuse si S est supposé compact et f est continue, auquel cas on sait exploiter la dualité entre l'ensemble des mesures de Borel-Radon sur S et celui des fonctions continues sur S.*

Cette méthode est présentée dans [26] et, sous certaines conditions, permet une généralisation des solutions du problème primal en autorisant qu'elle ne soit pas absolument continue par rapport à la mesure a priori μ (Théorème 4.1 de la référence précédemment citée).

En adaptant alors les commentaires faits pour établir la Proposition II.5.2, la Proposition II.3.8 reste valable en se restreignant à une fonction de payoff bornée :

Proposition II.5.6 *Sous les hypothèses précédentes sur J et χ et en supposant que f est bornée, presque sûrement la solution ν_n du problème $(P_n^{J,\chi})$ converge étroitement vers $\mu_J^{(\theta,\theta_0)}$.*
En particulier, la mesure $\mu_J^{(\theta,\theta_0)}$ est une probabilité.
De manière générale, si $\phi : S \to \mathbb{R}$ est mesurable bornée, on a la convergence presque sûre

$$\int \phi d\nu_n \longrightarrow_{n\to+\infty} \int \phi d\mu_J^{(\theta,\theta_0)}$$

II.5.3 Conclusion

A condition de se restreindre à la classe des fonctions de payoff bornées, on peut relâcher la condition (II.5.1) et justifier la convergence Monte-Carlo pour des critères dont le comportement s'éloigne de celui de l'entropie relative. Cela nous autorise en particulier à adopter un critère construit sur une norme L^p, $p > 1$.
La condition (II.5.2) demeure cependant une restriction non artificielle et exclut de cette étude les critères définis à partir d'une fonction puissance d'exposant $p < 1$

$$J(u) = \begin{cases} \frac{p}{p-1}u^p & \text{si } u \geq 0 \\ +\infty & \text{sinon} \end{cases}$$

avec $0^p = +\infty$ si $p < 0$. De même, il n'a pas été possible d'inclure dans l'étude l'entropie de Burg :

$$J(u) = \begin{cases} -\ln(u) & \text{si } u \geq 0 \\ +\infty & \text{sinon} \end{cases}$$

Dans ce cas, il est cependant possible d'envisager une extension du problème selon la Remarque II.5.5, conduisant sous certaines conditions à le résoudre au moins en un sens généralisé.

Chapitre III

Minimisation de l'entropie relative pénalisée et solution généralisée

Sous des hypothèses convenables, il a été démontré au chapitre précédent que la suite des solutions Monte-Carlo du problème de calibration pénalisé convergeait vers l'attracteur $\mu_J^{(\theta,\theta_0)}$, obtenu en résolvant un problème de minimisation dual.
Si l'on considère le cas usuel où le critère convexe est l'entropie relative, on peut en fait obtenir une caractérisation plus forte de la solution généralisée $\mu_J^{(\theta,\theta_0)}$ en terme de convergence de suites. Ce phénomène est bien connu et a été rappelé dans le cas de la calibration exacte où l'on dispose de la notion de I-projection généralisée selon [38]. On se propose donc d'étendre le résultat au cas de la calibration pénalisée.
En réalité, on traitera de façon unifiée le cas de la calibration pénalisée ainsi que le cas de la calibration avec contraintes de type ensemble. On ne considère pas ici le mélange de ces deux types de calibration mais cela ne nuit pas à l'intérêt des résultats obtenus. En particulier, au résultat de [72] traitant les contraintes exactes on donnera une extension pour cette situation plus générale où l'on a l'existence de la I-projection généralisée de la mesure μ sur un ensemble convexe.

On fait donc les hypothèses alternatives suivantes :

$$(\tilde{C}) \begin{cases} \text{K est un ensemble convexe fermé de } \mathbb{R}^d\text{, la condition } (C'_K) \text{ est vérifiée,} \\ \text{E est le sous-espace affine de } \mathbb{R}^d \text{ qui lui est associé dans le cas de contraintes de type} \\ \text{ensemble et l'on pose alors } \chi = \delta(.|K) \\ \\ \text{ou bien, dans le cas pénalisé :} \\ E = \mathbb{R}^d \text{ et } \chi : \mathbb{R}^d \longrightarrow \mathbb{R} \text{ est une fonction finie, strictement convexe, sci et minimale en } \mathbf{0} \\ \text{ et sa conjuguée } \chi^* \text{ est strictement convexe sur le domaine où elle est finie et vérifie :} \\ \frac{\chi^*(u)}{\|u\|} \longrightarrow_{\|u\|\to+\infty} +\infty \end{cases}$$

On redonne ici la condition géométrique

$$(C'_K) \begin{cases} \mathbf{0} \in Int_E(Conv(Supp(\mu_{|E}^{f_1}))) \bigcap K \\ \forall \eta \in \mathcal{P}(S) \text{ absolument continue par rapport à } \mu \ t.q. \int \|f_1\| d\eta < +\infty \text{ et } \int f_1 d\eta \in K, \\ \text{on a } \eta(f_1^{-1}(E)) = 1 \end{cases}$$

Ainsi qu'on l'a déjà remarqué, c'est sans perte de généralité que l'on peut adopter cette hypothèse à la place de (C_K).

Rappelons alors le problème de calibration :

$$(PL^\chi) \quad \text{trouver } \nu \in \mathcal{P}(S) \text{ qui minimise } H(\nu\|\mu) + Q_\chi(\nu)$$

en posant pour tout $\nu \in \mathcal{P}(S)$

$$Q_\chi(\nu) = \begin{cases} \chi(\int f d\nu) & \text{si } \int_S \|f\| d\nu < +\infty \\ +\infty & \text{sinon} \end{cases}$$

III.1 Caractérisation de l'attracteur des suites minimisantes

On commence par étendre la notion de I-projection généralisée dans le cadre du problème de minimisation (PL^χ) :

Lemme III.1.1 *Il existe une probabilité unique $\nu^\chi \in \mathcal{P}(S)$, absolument continue par rapport à μ, telle que toute suite minimisante pour le problème (PL^χ) converge en variation vers ν^χ.*
On dira ici que ν^χ est la (PL^χ)-projection généralisée de la mesure μ.

Dans le cas de la calibration avec contraintes, ν^χ est simplement la I-projection généralisée de la mesure μ sur l'ensemble

$$\mathcal{C}_K = \left\{ \nu \in \mathcal{P}(S),\ \nu \ll \mu,\ \int \|f\| d\nu < +\infty \text{ et } \int f d\nu \in K \right\} \tag{III.1.1}$$

Dans ce cas, la preuve de l'existence de la I-projection, forcément unique, se résume à montrer qu'il existe $\nu \in \mathcal{C}_K,\ H(\nu\|\mu) < +\infty$.
La démonstration suivante redonne alors ce résultat connu et établit aussi l'existence de l'attracteur dans le cas pénalisé.

Preuve : Commençons par remarquer que

$$\exists \alpha \in \mathcal{P}(S),\ H(\alpha\|\mu) + Q_\chi(\alpha) < \infty \tag{III.1.2}$$

C'est une conséquence des relations (II.3.10) et (II.4.5) mais on peut le retrouver de manière plus directe.

Dans le cas contraint, avec la condition géométrique (C'_K) se trouve satisfaite la condition (C) du cas contraint élémentaire : avec les résultats concernant ce dernier, on obtient l'existence de la projection généralisée de μ sur $\{\nu \in \mathcal{P}(S),\ \nu \ll \mu,\ \int \|f\| d\nu < +\infty \text{ et } \int f d\nu = \mathbf{0}\}$ et l'assertion (III.1.2) en découle en rappelant qu'avec (C'_K) on s'est ramené à $\mathbf{0} \in K$. Dans ce cas, on obtient directement du théorème 2.1. de [38] l'existence de la projection généralisée ν^χ.
Dans le cas pénalisé, la relation (III.1.2) est immédiate à établir car, f étant μ-ps finie, on peut trouver $a \geq 0$ tel que

$$\mu\{\|f\| \leq a\} = p > 0$$

On pose alors $h = \frac{1}{p} 1_{\{\|f\| \leq a\}}$ et on a le résultat avec $\alpha = h\mu$. On déduit alors de (III.1.2) l'existence d'une suite minimisante pour le problème (PL^χ) et il suffit alors de reprendre la technique du théorème 2.1. de [38] en considérant une telle suite minimisante (α_n) et en montrant qu'elle est de Cauchy pour la norme de la variation. On en conclura par complétude de $L^1(\mu)$ que la limite existe et est absolument continue par rapport à μ.
L'identité du parallélogramme pour l'entropie s'écrit :

$$H(\alpha_m\|\mu) + H(\alpha_n\|\mu) = 2H\left(\frac{\alpha_m + \alpha_n}{2}\middle\|\mu\right) + H\left(\alpha_m\middle\|\frac{\alpha_m + \alpha_n}{2}\right) + H\left(\alpha_n\middle\|\frac{\alpha_m + \alpha_n}{2}\right)$$

L'inégalité de convexité pour le terme de pénalisation donne par ailleurs :

$$\chi\left(\int f d\alpha_m\right) + \chi\left(\int f d\alpha_n\right) \geq 2\chi\left(\int f d\frac{\alpha_m + \alpha_n}{2}\right)$$

En additionnant membre à membre les deux équations précédentes, et en tenant compte du fait que la suite (α_n) est minimisante, on obtient en passant à la limite

$$0 \geq \limsup_{n,m} \left[H\left(\alpha_m \middle\| \frac{\alpha_m + \alpha_n}{2}\right) + H\left(\alpha_n \middle\| \frac{\alpha_m + \alpha_n}{2}\right) \right]$$

Le second membre étant positif, on en déduit que l'inégalité est une égalité, ce qui permet d'établir la convergence vers 0 de $H\left(\alpha_m \| \frac{\alpha_m+\alpha_n}{2}\right) + H\left(\alpha_n \| \frac{\alpha_m+\alpha_n}{2}\right)$.
En utilisant la relation entre norme en variation et entropie relative

$$|\alpha - \beta| \leq 2\sqrt{H(\alpha\|\beta)}$$

vraie pour α et β de $\mathcal{P}(S)$, ainsi que l'inégalité

$$|\alpha_m - \alpha_n| \leq \left|\alpha_m - \frac{\alpha_m + \alpha_n}{2}\right| + \left|\alpha_n - \frac{\alpha_m + \alpha_n}{2}\right|$$

on établit alors que (α_n) est une suite de Cauchy. Finalement, en intercalant les termes de deux suites minimisantes, on obtient une suite minimisante qui converge d'après ce qui précède. La limite est donc indépendante de la suite minimisante, ce qui achève la preuve. ∎

Remarque III.1.2 *Si l'on se donne un sous-ensemble convexe $\mathcal{C}$ de $\mathcal{P}(S)$, tel que*

$$\inf_{\nu \in \mathcal{C}} \left(H(\nu\|\mu) + Q_\chi(\nu)\right) < +\infty$$

la preuve précédente montre qu'au problème d'entropie pénalisée restreint aux probabilités de $\mathcal{C}$ on peut faire correspondre la notion de (PL^χ)-projection généralisée de la mesure μ sur l'ensemble $\mathcal{C}$, limite en variation de toute suite minimisante éventuellement en dehors de $\mathcal{C}$ si ce dernier n'est pas fermé pour la norme de la variation.
Dans le cas de la calibration avec contraintes, ce n'est autre que la I-projection généralisée de la mesure μ sur $\mathcal{C}_K \bigcap \mathcal{C}$.

Ceci étant acquis, on va montrer que la (PL^χ)-projection généralisée du problème (PL^χ) s'identifie à la solution généralisée exhibée à la dernière section, et dit autrement, sa recherche se ramène à la minimisation de la fonction

$$\theta \in E \longmapsto \ln Z(\theta) + \chi^*(-\theta) \tag{III.1.3}$$

où l'on rappelle que la transformée de Laplace de $1_E(f)\mu^f$ est notée

$$Z(\theta) = \int 1_E(f) e^{\theta \cdot f} d\mu$$

Avec la Proposition II.3.2 dans le cas pénalisé ou bien la Proposition II.4.4 dans le cas contraint, on peut d'abord énoncer :

Lemme III.1.3 *La fonction $\ln Z + \chi^*(-.)$ admet un minimum unique $\theta_* \in E$.*

Il suffit de noter que le critère entropique $J(x) = x \ln x$ admet pour conjuguée $J^*(x) = e^{x-1}$ de telle sorte que le critère dual à minimiser est la fonction $(u, u_0) \in E \times \mathbb{R} \longmapsto \int e^{u \cdot f + u_0 - 1} 1_E(f) d\mu - u_0 + \chi^*(-u)$.
La minimisation par rapport à $u_0 \in \mathbb{R}$ pour $u \in E$ fixé donne alors l'optimum $u_0 = 1 - \ln \int 1_E(f) e^{u \cdot f} d\mu$ et l'on aboutit bien au critère (III.1.3) à minimiser sur E.
On donne cependant ici une preuve plus directe que celles qui ont conduit aux deux propositions citées ci-dessus.
Preuve : L'unicité du minimum est acquise si l'on prouve que la fonction $\ln Z + \chi^*(-.)$ est strictement convexe sur E.
La convexité de $\ln Z$ est en effet un résultat classique relatif à la transformée de Laplace et s'obtient en appliquant l'inégalité d'Hölder. Avec la stricte convexité de χ^* dans le cas pénalisé ou bien la stricte convexité de $\ln Z$ sous la condition (C_K) dans le cas contraint, on en déduit la stricte convexité cherchée.
Pour conclure à l'existence d'un minimum, on va montrer que $\ln Z + \chi^*(-.)$ est semi-continue inférieurement et tend vers $+\infty$ avec $\|\theta\|$.
On a déjà remarqué que Z est semi-continue inférieurement grâce au lemme de Fatou et il en est de même de $\ln Z$, par continuité et croissance du logarithme. $\ln Z + \chi^*(-.)$ est donc semi-continue inférieurement.
Pour étudier le comportement à l'infini, on tient pour acquis que $\ln Z$ converge vers l'infini sous la condition géométrique (C'_K), car on l'a déjà obtenu sous (C) et cela règle le cas de la calibration sous contraintes.
Dans le cas pénalisé, l'idée est de se ramener au cas où $f$est μ-intégrable et même bornée en choisissant a>0 tel que $\mu(\{f \in E,\ \|f\| \leq a\}) > 0$ et en posant

$$\mu_a = \frac{1_{\{f \in E,\ \|f\| \leq a\}}}{\mu(\{f \in E,\ \|f\| \leq a\})} \mu$$

On a alors

$$\ln Z(\theta) \geq \ln \int e^{\theta \cdot f} d\mu_a + \ln \mu(\{f \in E,\ \|f\| \leq a\})$$

On en déduit en utilisant en particulier l'inégalité de Jensen pour le premier terme du second membre

$$\ln Z(\theta) + \chi^*(-\theta) \geq \chi^*(-\theta) + \theta \cdot \int f d\mu_a + \ln \mu(\{f \in E,\ \|f\| \leq a\})$$

ce qui permet d'en déduire la convergence vers $+\infty$ avec $\|\theta\|$, grâce à la condition de croissance imposée sur χ^*. ∎

Si l'on note θ_* le mininimum de $\ln Z + \chi^*(-.)$, la proposition suivante montre que la (PL^χ)-projection généralisée a pour densité par rapport à μ la distribution de Boltzmann de paramètre θ_* et est donc également la solution généralisée du problème (PL^χ) introduite dans la section précédente.

Proposition III.1.4 *On a*

$$\inf_{\nu \in \mathcal{P}(S)} H(\nu|\mu) + Q_\chi(\nu) = -(\ln Z(\theta_*) + \chi^*(-\theta_*)) \qquad \text{(III.1.4)}$$

La (PL^χ)-projection généralisée de μ est donnée par

$$\mu_E^{\theta_*} = \frac{1_E(f) e^{\theta_* \cdot f}}{Z(\theta_*)} \mu$$

C'est la solution usuelle du problème (PL^χ) si et seulement si $\int \|f\| d\mu_E^{\theta_*} < +\infty$ et $\int f d\mu_E^{\theta_*} \in \partial \chi^*(-\theta_*)$

.

Preuve : L'égalité (III.1.4) ainsi que la condition de sous-gradient ont été déjà obtenues dans un cadre général et il s'agit uniquement de montrer que la (PL^{χ})-projection généralisée de μ est de la forme annoncée en adaptant pour cela la preuve due à Jupp & Mardia dans [72].

Pour la suite il est pratique de noter pour tout ν de $\mathcal{P}(S)$

$$K_{\chi}(\nu) = H(\nu\|\mu) + Q_{\chi}(\nu) \tag{III.1.5}$$

si bien que l'égalité (III.1.4) s'écrit

$$\inf_{\nu\in\mathcal{P}(S)} K_{\chi}(\nu) = -(\ln Z(\theta_*) + \chi^*(-\theta_*)) \tag{III.1.6}$$

On procède ensuite par approximation en considérant une suite croissante (S_n) de sous-ensembles mesurables de S, de mesure non nulle et de réunion S, telle que $f$est bornée sur chaque S_n.
En particulier

$$\forall n,\ 0 < \mu(S_n) \to_{n\to\infty} 1$$

Posons

$$\tilde{\mu}_n = \frac{1_{S_n}}{\mu(S_n)}\mu$$

et introduisons les sous-ensembles de $\mathcal{P}(S)$:

$$\tilde{\mathcal{C}}_n = \{\nu \in \mathcal{P}(S),\ \nu \ll \tilde{\mu}_n\}$$

et

$$\tilde{\mathcal{C}}_\infty = \bigcup \tilde{\mathcal{C}}_n$$

En formulant le problème $(\tilde{PL}_n^{\chi})$ comme le problème (PL^{χ}) avec μ remplacé par $\tilde{\mu}_n$, on vérifie facilement que l'ensemble des hypothèses $(\tilde{C})$ est vérifié à partir d'un rang n assez grand ($n = 1$ dans le cas pénalisé), avec le même E que le problème original.
Puisque l'on est dans le cas où $f$est bornée, alors d'après des résultats vus dans la section dernière (Proposition II.3.6 et dernière assertion de la Proposition II.4.5), le problème $(\tilde{PL}_n^{\chi})$ admet la solution usuelle

$$\tilde{\nu}_n = 1_E(f)e^{\tilde{\theta}_n\cdot f - \ln \tilde{Z}_n(\tilde{\theta}_n)}\tilde{\mu}_n \tag{III.1.7}$$

avec

$$\tilde{Z}_n(\theta) = \int 1_E(f)e^{\theta\cdot f}d\tilde{\mu}_n$$

et où $\tilde{\theta}_n$ réalise le minimum de

$$\ln \tilde{Z}_n(\theta) + \chi^*(-\theta)$$

En appliquant de plus l'égalité (III.1.4) déjà prouvée au problème $(\tilde{PL}_n^{\chi})$, on a :

$$\inf_{\nu\in\tilde{\mathcal{C}}_n} (H(\nu\|\tilde{\mu}_n) + Q_{\chi}(\nu)) = H(\tilde{\nu}_n\|\tilde{\mu}_n) + Q_{\chi}(\tilde{\nu}_n) = -\left(\ln \tilde{Z}_n(\tilde{\theta}_n) + \chi^*(-\tilde{\theta}_n)\right) \tag{III.1.8}$$

En reprenant la notation introduite dans (III.1.5), $\tilde{\nu}_n$ est également solution du problème de minimisation de K_{χ} sur $\tilde{\mathcal{C}}_n$ puisque l'on a :

$$\forall \nu \in \tilde{\mathcal{C}}_n,\ K_{\chi}(\nu) = H(\nu\|\tilde{\mu}_n) + Q_{\chi}(\nu) - \ln \mu(S_n) \tag{III.1.9}$$

D'autre part, $\tilde{\mathcal{C}}_\infty$ est convexe et l'on a clairement :

$$\inf_{\nu \in \tilde{\mathcal{C}}_\infty} K_\chi(\nu) = \lim_{n\to\infty} \inf_{\nu \in \tilde{\mathcal{C}}_n} K_\chi(\nu) < +\infty \tag{III.1.10}$$

Alors, grâce à la Remarque III.1.2, on peut définir la (PL^χ)-projection généralisée $\tilde{\nu}_\infty$ de μ sur l'ensemble convexe $\tilde{\mathcal{C}}_\infty$. $(\tilde{\nu}_n)$ étant une suite minimisante dans $\tilde{\mathcal{C}}_\infty$, on en déduit, par définition de la solution généralisée et comme cette dernière est absolument continue par rapport à μ :

$$\frac{d\tilde{\nu}_n}{d\mu} \to \frac{d\tilde{\nu}_\infty}{d\mu} \tag{III.1.11}$$

dans $L^1(\mu)$.
Quitte à extraire une sous-suite, on peut supposer que la convergence est μ-ps, ce qui implique la convergence μ-ps de $\tilde{\theta}_n \cdot f - \ln \tilde{Z}_n(\tilde{\theta}_n)$.

Afin de préparer la caractérisation de $\tilde{\nu}_\infty$, on va montrer que les suites $(\tilde{\theta}_n)$ et $(\ln \tilde{Z}_n(\tilde{\theta}_n))$ sont bornées.
En utilisant les relations (III.1.9) et (III.1.8), on écrit d'abord pour tout n

$$K_\chi(\tilde{\nu}_n) = -\Big(\ln \tilde{Z}_n(\tilde{\theta}_n) + \chi^*(-\tilde{\theta}_n) + \ln \mu(S_n)\Big) \tag{III.1.12}$$

Puisque $K_\chi(\tilde{\nu}_n) = \min_{\nu\in\tilde{\mathcal{C}}_n} K_\chi(\nu)$ est minorée en vertu de (III.1.6), on déduit de la relation précédente le fait que $\ln \tilde{Z}_n(\tilde{\theta}_n) + \chi^*(-\tilde{\theta}_n)$ est majorée : comme $\tilde{\mu}_n^f$ converge étroitement vers μ^f, on en déduit bien que $(\tilde{\theta}_n)$ est bornée, avec les Lemmes II.3.9 et II.4.3, en prenant $u_n = \tilde{\theta}_n$ et $u_{0,n} = 1 - \ln \tilde{Z}_n(\tilde{\theta}_n)$.
On peut cependant donner ici une simplification des étapes qui ont permis d'établir les lemmes de la section précédente.
En effet, compte tenu de la croissance de la suite (S_n), on a la minoration :

$$\tilde{Z}_n(\tilde{\theta}_n) \geq \int 1_E(f) e^{\tilde{\theta}_n \cdot f} d\tilde{\mu}_1 \frac{\mu(S_1)}{\mu(S_n)}$$

On en déduit, grâce à l'inégalité de Jensen, en supposant $\mu(f^{-1}(E) \bigcap S_1) > 0$:

$$\ln \tilde{Z}_n(\tilde{\theta}_n) \geq \tilde{\theta}_n \cdot \int \frac{1_E(f) f}{\tilde{\mu}_1(f^{-1}(E))} d\tilde{\mu}_1 + \ln \tilde{\mu}_1(f^{-1}(E)) + \ln \mu(S_1) - \ln \mu(S_n)$$

Alors avec (III.1.12), on obtient la majoration, pour tout n

$$K_\chi(\tilde{\nu}_n) \leq -\left(\chi^*(-\tilde{\theta}_n) + \tilde{\theta}_n \cdot \int \frac{1_E(f) f}{\tilde{\mu}_1(f^{-1}(E))} d\tilde{\mu}_1 + \ln \tilde{\mu}_1(f^{-1}(E)) + \ln \mu(S_1)\right)$$

La suite $(\tilde{\theta}_n)$ est donc bornée puisque le premier membre est positif ou nul tandis que l'expression $\chi^*(-\theta) + \theta \cdot \int \frac{1_E(f) f}{\tilde{\mu}_1(f^{-1}(E))} d\tilde{\mu}_1 + \ln \tilde{\mu}_1(f^{-1}(E)) + \ln \mu(S_1)$ tend vers $+\infty$ avec $\|\theta\|$.
Avec la convergence de $\tilde{\theta}_n \cdot f - \ln \tilde{Z}_n(\tilde{\theta}_n)$, on voit enfin que la suite $(\ln \tilde{Z}_n(\tilde{\theta}_n))$ est également bornée.
Ceci étant établi, quitte à procéder de nouveau par sous-extraction de suites, on peut supposer que $(\tilde{\theta}_n)$ et $(\ln \tilde{Z}_n(\tilde{\theta}_n))$ ont pour limites $\tilde{\theta}_\infty \in E$ et $\tilde{L}_\infty \in \mathbb{R}$:

$$\tilde{\theta}_n \to \tilde{\theta}_\infty \tag{III.1.13}$$

et

$$\ln \tilde{Z}_n(\tilde{\theta}_n) \to \tilde{L}_\infty \tag{III.1.14}$$

On en déduit :

$$\frac{d\tilde{\nu}_\infty}{d\mu} = 1_E(f) e^{\tilde{\theta}_\infty \cdot f - \tilde{L}_\infty} \tag{III.1.15}$$

Puisque $\tilde{\nu}_\infty$ est une probabilité, on a

$$Z(\tilde{\theta}_\infty) < +\infty$$

et

$$\tilde{L}_\infty = \ln Z(\tilde{\theta}_\infty) \tag{III.1.16}$$

Montrons alors que $\tilde{\nu}_\infty$ est la solution généralisée du problème (PL^χ).
Du fait de l'inclusion de $\tilde{\mathcal{C}}_\infty$ dans $\mathcal{P}(S)$, on a

$$\begin{aligned}
\inf_{\nu \in \mathcal{P}(S)} K_\chi(\nu) \leq \inf_{\nu \in \tilde{\mathcal{C}}_\infty} K_\chi(\nu) &= \lim_{n\to\infty} \inf_{\nu \in \tilde{\mathcal{C}}_n} K_\chi(\nu) \quad \text{d'après (III.1.10)} \\
&= \lim_{n\to\infty} \left(\inf_{\nu \in \tilde{\mathcal{C}}_n} (H(\nu \| \tilde{\mu}_n) + Q_\chi(\nu)) - \ln \mu(S_n) \right) \quad \text{d'après (III.1.9)} \\
&= \lim_{n\to\infty} \left(-(\ln \tilde{Z}_n(\tilde{\theta}_n) + \chi^*(-\tilde{\theta}_n)) - \ln \mu(S_n) \right) \quad \text{d'après (III.1.8)} \\
&= -(\ln Z(\tilde{\theta}_\infty) + \chi^*(-\tilde{\theta}_\infty)) \quad \text{d'après (III.1.14) et (III.1.16)} \\
&\leq -(\ln Z(\theta_*) + \chi^*(-\theta_*)) \\
&= \inf_{\nu \in \mathcal{P}(S)} K_\chi(\nu) \quad \text{d'après (III.1.6)} \qquad \text{(III.1.17)}
\end{aligned}$$

Les relations précédentes sont donc toutes des égalités, ce qui va permettre de conclure.
En effet, on a d'abord

$$\inf_{\nu \in \mathcal{P}(S)} K_\chi(\nu) = \inf_{\nu \in \tilde{\mathcal{C}}_\infty} K_\chi(\nu)$$

ce qui prouve que $\tilde{\nu}_\infty$ est la (PL^χ)-projection généralisée de μ (sur $\mathcal{P}(S)$) par unicité de cette dernière.
On tire également de (III.1.17) l'égalité

$$\ln Z(\tilde{\theta}_\infty) + \chi^*(-\tilde{\theta}_\infty) = \ln Z(\theta_*) + \chi^*(-\theta_*)$$

et cela montre, par unicité du minimum de $\ln Z + \chi^*$, l'égalité

$$\tilde{\theta}_\infty = \theta_*$$

Avec (III.1.15) et (III.1.16), en reprenant la notation de la proposition, on en déduit

$$\tilde{\nu}_\infty = \mu_E^{\theta_*}$$

et la solution généralisée est donc de la forme annoncée, ce qui achève la preuve. ∎

III.2 Etude de la solution généralisée dans le cas d'un terme de pénalisation quadratique

III.2.1 Critère pour que la solution soit usuelle

Après avoir identifié la solution généralisée du problème de calibration pénalisé comme étant la limite en variation des suites minimisantes, il est intéressant de regarder les situations où cette limite peut différer de la solution usuelle. Ces situations ont justifié l'étude précédente et l'on désire également examiner l'effet de la pénalisation. L'objet de ce numéro est d'établir une condition permettant de reconnaître si la solution est usuelle ou seulement généralisée, dans le cas de la pénalisation quadratique.

De façon précise, on va se placer dans le cas pratique proposé dans [6] où la pénalisation est choisie du type :

$$\chi_\gamma : x \longmapsto \frac{1}{2}\sum_{i=1}^{d}\frac{1}{\gamma_i}(x^i)^2$$

où l'on s'est donné un vecteur $\gamma = (\gamma^i)_{1\leq i\leq d}$ de d nombres strictement positifs et l'on supposera que $Supp(\mu^f)$ engendre linéairement $\mathbb{R}^d$.
On peut toujours se ramener à cette situation en partant d'une forme quadratique définie positive puis en la restreignant au sous-espace engendrée par $Supp(\mu^f)$: en effet si χ est une telle forme quadratique sur $\mathbb{R}^d$ et si F est le sous-espace affine engendré par $Supp\mu^f$, on peut considérer le χ-projeté orthogonal a du point $\mathbf{0}$ sur F, et noter l'égalité

$$\forall x \in F,\ \chi(x) = \chi(a) + \chi(x-a)$$

Finalement, quitte à remplacer f par $f-a$ et à diagonaliser χ sur une base χ-orthogonale de l'espace vectoriel directeur de F, on est bien ramené à une constante additive près à un critère de pénalisation du type de χ_γ avec en plus la condition que $Supp(\mu^f)$ engendre $\mathbb{R}^d$, quitte à diminuer d.

Remarque III.2.1 *Le terme quadratique de pénalisation a une interprétation naturelle si l'on reprend le dernier commentaire du premier chapitre.*
Considérons que l'observation $\mathbf{C}$*, non ramenée à* $\mathbf{0}$ *ici, s'obtient à partir du vrai modèle à une erreur gaussienne près :*

$$\mathbf{C} = \int_S f d\nu + \epsilon$$

où ϵ est un vecteur Gaussien centré de matrice de variances-covariances V.
Ainsi, conditionellement à ν, la vraisemblance en $\mathbf{C}$ *est supposée être :*

$$\mathbb{P}^{\mathbf{C}|\nu} = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\mathbf{det}V}}e^{-\chi(\int f d\nu - \mathbf{C})} \tag{III.2.1}$$

en posant :

$$\forall u \in \mathbb{R}^d,\ \chi(u) = \frac{1}{2}{}^t u V^{-1} u$$

En faisant l'hypothèse que l'inconnue ν est distribuée suivant une loi approchée par (I.3.1), la méthode du maximum de vraisemblance conduit donc à la minimisation de

$$H(\nu\|\mu) + \chi\bigg(\int f d\nu - \mathbf{C}\bigg)$$

Avec le choix d'un tel critère de pénalisation, l'ensemble des conditions $(\tilde{C})$ est satisfait et en particulier χ_γ^* est strictement convexe et vérifie la condition de croissance désirée, ce qui permet d'appliquer les résultats généraux.
On notera alors θ_γ le minimum de $\ln Z + \chi_\gamma^*$ et on regroupe en un seul énoncé la dernière assertion de la Proposition II.3.3 ainsi que le numéro 2 de la Remarque II.3.4 en rappelant que $\theta_\gamma \in D_\mu$ désigne l'ensemble de définition de la transformée de Laplace de μ^f :

Lemme III.2.2 *1. Si $\int \|f\| d\mu_\gamma^\theta < +\infty$ alors*

$$+\infty > H(\mu^{\theta_\gamma} \| \mu) + Q_{\chi_\gamma}(\mu^{\theta_\gamma}) \geq \inf_{\nu \in \mathcal{P}(S)} H(\nu | \mu) + Q_{\chi_\gamma}(\nu)$$

avec égalité à droite si et seulement si μ^{θ_γ} est la solution usuelle du problème (PL^{χ_γ}).

2. Supposons que $Supp(\mu^f)$ engendre $\mathbb{R}^d$.
Alors la solution est usuelle si et seulement s'il existe $\theta \in D_\mu$ avec $\int \|f\| d\mu^\theta < +\infty$ tel que l'égalité suivante soit réalisée

$$\forall i \in [1, d] \int f^i d\mu^\theta = -\gamma_i \theta^i \tag{III.2.2}$$

Alors on a $\theta = \theta_\gamma$

On peut remarquer qu'avec la condition d'égalité (III.2.2) on retrouve la satisfaction des contraintes dans le cas où γ tend vers $\mathbf{0}$ et on remarque aussi que cette relation généralise la condition du premier ordre dans le cas où le minimum θ_γ est intérieur à l'ensemble de définition D_μ de la transformée de Laplace de μ^f.

III.2.2 Quelques exemples

Examinons quelques situations où la solution n'existe qu'au sens généralisé. D'après ce qui précède, cela n'est possible que si D_μ n'est pas un ouvert. On précisera pour finir le lien entre la position du paramètre θ_γ dans D_μ, celle de la contrainte $\mathbf{0}$ dans $Conv\left(Supp\left(\mu^f\right)\right)$ et la nature de la solution μ^{θ_γ}.

1. Il peut d'abord se faire que f ne soit pas intégrable par rapport à μ^{θ_γ}, ce qui empêche cette dernière d'être solution usuelle.
   Par exemple, c'est le cas lorsque f est l'identité sur $\mathbb{R}$, μ est une loi sur $\mathbb{R}$ n'ayant pas de moyenne et dont la transformée de Laplace n'est définie qu'en 0 (prendre la loi de Cauchy $\mu = \frac{1}{\pi} \frac{1}{1+x^2} dx$).
   Dans la situation où μ est pourvue d'une moyenne, et où transformée de Laplace n'est définie qu'en 0 , on ne vérifie la condition d'égalité du Lemme précédent que si la moyenne est nulle, ce qui fait encore que la solution μ n'est pas usuelle en général.

2. Le cas de figure précédent n'est pas pertinent en finance car la transformée de Laplace de μ^f est définie au moins sur $\mathbb{R}_-^d$.

   Pour se placer dans cette situation, on se ramène au cas où $f$est l'identité sur $\mathbb{R}$ et l'on considère la loi sur $\mathbb{R}$
   $$\tilde{\mu} = A \frac{e^{-ax}}{1 + |x|^3} 1_{[M, +\infty[}$$
   avec $a \geq 0$ et où $A > 0$ est choisi de telle sorte que $\tilde{\mu}$ soit une probabilité. Le réel M est a priori quelconque mais, dans un contexte financier, on le choisira positif ou nul.
   Prenons alors comme contrainte un nombre m tel que
   $$m > \frac{\int x e^{ax} \tilde{\mu}(dx)}{\int e^{ax} \tilde{\mu}(dx)} + \gamma a \tag{III.2.3}$$
   et ramenons cette contrainte à 0, suivant l'habitude, en posant
   $$\mu = \tilde{\mu} * \delta_{-m}$$

Grâce à la convexité stricte de $\ln \int e^{\theta x}\mu(dx) + \frac{1}{2}\gamma\theta^2$, on a la croissance stricte de sa dérivée $\int x\mu^\theta(dx) + \gamma\theta$ sur $]-\infty, a]$.
En vertu de la convergence dominée, le support de μ étant limité à gauche, cette expression est continue à gauche en a, ce qui fait que l'on a :

$$\begin{aligned}
\forall\theta \leq a, \int x\mu^\theta(dx) + \gamma\theta &\leq \int x\mu^a(dx) + \gamma a \\
&\leq \int x\tilde{\mu}^a(dx) + \gamma a - m \\
&< 0
\end{aligned}$$

avec la condition (III.2.3).
Ainsi, la solution n'existe pas au sens usuel puisque la condition d'égalité (III.2.2) du Lemme précédent ne peut être vérifiée en aucun point de l'ensemble de définition $D_\mu =]-\infty, a]$ de la transformée de Laplace de μ.
Dans ce cas, le paramètre θ_γ vaut forcément a, car sinon il serait intérieur à D_μ et la solution serait usuelle.

3. Comme dans le cas contraint, la condition sur la contrainte $\mathbf{C} = \mathbf{0}$ du type

$$\mathbf{0} \in Int\left(Conv\left(Supp\left(\mu^f\right)\right)\right) \tag{III.2.4}$$

ne suffit pas à assurer que la solution est usuelle, et a fortiori que le paramètre θ_γ est intérieur à D_μ.
Pour le voir, on remarque dans l'exemple précédent que m est strictement plus grand que M, ce qui fait que 0 est un point intérieur à $Supp(\mu) = [M - m, +\infty[$, alors que la solution n'existe qu'au sens généralisé d'après l'étude précédente.
Inversement, il se peut que θ_γ soit intérieur à D_μ sans que (III.2.4) soit vérifiée. On reprend encore la mesure μ définie dans l'exemple du 2 avec cette fois m un nombre tel que $m \leq M$, ce qui fait que l'inégalité (III.2.3) tient en sens inverse :

$$\frac{\int xe^{ax}\tilde{\mu}(dx)}{\int e^{ax}\tilde{\mu}(dx)} + \gamma a > m \tag{III.2.5}$$

On voit d'abord que 0 n'est pas un point intérieur à $Supp(\mu)$ et la condition (III.2.4) n'est pas réalisée.
Cependant, la solution μ^{θ_γ} est usuelle, et θ_γ est même intérieur à D_μ. En effet, reprenons la fonction $\Lambda : \theta \to \int x\mu^\theta(dx) + \gamma\theta$ associée à la condition (III.2.2).
On a vu au numéro 2 que Λ est continue sur $]-\infty, a]$.
Alors $\Lambda(a) > 0$ à cause de (III.2.5) tandis que $\lim_{\theta\to-\infty} \Lambda(\theta) = -\infty$, ce qui se justifie du fait que $\Lambda(\theta)$ est la somme du terme $\int x\mu^\theta(dx)$ qui est croissant par convexité de $\ln \int e^{\theta x}\mu(dx)$ et du terme $\gamma\theta$ qui tend vers $-\infty$.
L'application Λ s'annule donc en un point de l'ouvert $]-\infty, a[$, qui est donc le minimum θ_γ de $\ln \int e^{\theta x}\mu(dx) + \frac{1}{2}\gamma\theta^2$

4. Notons enfin que la solution peut être usuelle, c'est-à-dire la condition (III.2.2) peut être réalisée, sans que θ_γ soit intérieur à D_μ. Il suffit de reprendre l'exemple du numéro précédent en remplaçant l'inégalité stricte (III.2.5) par une égalité en posant :

$$m = \frac{\int xe^{ax}\tilde{\mu}(dx)}{\int e^{ax}\tilde{\mu}(dx)} + \gamma a \tag{III.2.6}$$

La condition (III.2.2) est donc vérifiée en a.
On en déduit, avec le Lemme III.2.2, que la solution est usuelle et que $\theta_\gamma = a$.

III.2.3 Effet de la pénalisation sur la nature de la solution

Il est également légitime de se demander si on peut contraindre la solution à devenir usuelle en pénalisant davantage, c'est à dire en augmentant les γ_i.
La proposition suivante apporte une réponse à cette question.

Proposition III.2.3 *Si l'ensemble de définition D_μ de la transformée de Laplace de μ^f est un voisinage de l'origine $\mathbf{0}$, alors, dès que $\min_i \gamma_i$ est assez grand, le minimum θ_γ est dans l'intérieur de D_μ et par conséquent la solution du problème (PL^{χ_γ}) est usuelle .*
Plus précisément, si l'on autorise le vecteur de contraintes $\mathbf{C}$ à varier dans un ensemble compact arbitraire K de $\mathbb{R}^d$, alors θ_γ tend vers $\mathbf{0}$ uniformément par rapport à $\mathbf{C}$ dans K, lorsque $\min_i \gamma_i$ tend vers $+\infty$.
On en déduit aussi que μ^{θ_γ} tend en variation vers μ et que $\int f d\mu^{\theta_\gamma}$ tend vers $\int f d\mu$ uniformément par rapport à $\mathbf{C}$ dans K, lorsque $\min_i \gamma_i$ tend vers $+\infty$.

Preuve : Posons d'abord pour tout $\theta \in \mathbb{R}^d$

$$\xi(\theta) = \int f d\mu^\theta$$

et

$$\Lambda_\gamma(\theta) = (\gamma_i \theta^i)_{1 \leq i \leq d}$$

Avec ces notations, en autorisant $\mathbf{C}$ à ne plus être nulle, c'est-à-dire en remplaçant f par $f - \mathbf{C}$, on remarque que la condition (III.2.2) revient à

$$(\xi + \Lambda_\gamma)(\theta_\gamma) = \mathbf{C}$$

Ceci étant, soit ϵ assez petit pour avoir

$$\overline{B}(\mathbf{0}, \epsilon) \subset D_\mu$$

Alors, montrons que l'ensemble $(\xi + \Lambda_\gamma)(B(\mathbf{0}, \epsilon))$ contient K dès que $\min_i \gamma_i$ est assez grand. Grâce à la remarque précédente et au Lemme III.2.2 cela prouvera que $\theta_\gamma \in B(\mathbf{0}, \epsilon)$ et la preuve sera achevée.
D'abord $\xi + \Lambda_\gamma$ est un homéomorphisme sur $Int(D_\mu)$. En effet, d'une part la différentielle de $\xi + \Lambda_\gamma$ est inversible en tout point de $Int(D_\mu)$ puisque $D\xi$ est symétrique positive et que $D\Lambda_\gamma = \Lambda_\gamma$ est symétrique définie positive. On peut alors appliquer le théorème de l'inverse locale et conclure que $\xi + \Lambda_\gamma$ est un homéomorphisme local sur $Int(D_\mu)$.
D'autre part, $\xi + \Lambda_\gamma$ est injectif sur $Int(D_\mu)$. En effet, si l'on se donne θ de $Int(D_\mu)$ et $\mathbf{C}$ de $\mathbb{R}^d$ tels que

$$(\xi + \Lambda_\gamma)(\theta) = \mathbf{C}$$

alors, le Lemme III.2.2 et la remarque faite en début de preuve montre que θ est le minimum θ_γ du problème pénalisé (PL^{χ_γ}) correspondant à la contrainte $\mathbf{C}$, d'où l'unicité requise.
L'application $\xi + \Lambda_\gamma$ est donc un homéomorphisme sur $Int(D_\mu)$ et on a donc

$$Fr\Big((\xi + \Lambda_\gamma)(\overline{B}(\mathbf{0}, \epsilon))\Big) = (\xi + \Lambda_\gamma)(S(\mathbf{0}, \epsilon))$$

Alors, si l'on pose

$$\delta = d\Big(\xi(\mathbf{0}), Fr\Big((\xi + \Lambda_\gamma)(B(\mathbf{0}, \epsilon))\Big)\Big)$$

on peut trouver $u \in \mathbb{R}^d$, avec $\|u\| = \epsilon$ tel que

$$\delta = d(\xi(\mathbf{0}), \xi(u) + \Lambda_\gamma(u))$$

En appliquant l'inégalité triangulaire, on en tire

$$\delta \geq d(\xi(u) + \Lambda(u), \xi(u)) - d(\xi(\mathbf{0}), \xi(u)) \geq \epsilon \min_i \gamma_i - \delta'$$

avec

$$\delta' = \max_{\|v\|=\epsilon} d(\chi(\mathbf{0}), \chi(S(\mathbf{0}, \epsilon)))$$

Quitte à augmenter $\min_i \gamma_i$, on peut donc supposer que le membre de droite de l'inégalité précédente est positif. On écrit d'une part

$$B(\xi(\mathbf{0}), \epsilon \min_i \gamma_i - \delta') \bigcap Fr\bigg((\xi + \Lambda_\gamma)(B(\mathbf{0}, \epsilon))\bigg) = \emptyset$$

D'autre part, on a de façon évidente

$$\xi(\mathbf{0}) \in B(\xi(\mathbf{0}), \epsilon \min_i \gamma_i - \delta') \bigcap (\xi + \Lambda_\gamma)(B(\mathbf{0}, \epsilon)) \neq \emptyset$$

Par connexité de l'ensemble $(\xi + \Lambda_\gamma)(B(\mathbf{0}, \epsilon))$, on déduit des deux relations précédentes l'inclusion :

$$B(\xi(\mathbf{0}), \epsilon \min_i \gamma_i - \delta') \subset (\xi + \Lambda_\gamma)(B(\mathbf{0}, \epsilon))$$

Comme pour $\min_i \gamma_i$ assez grand, K est inclus dans le membre de gauche, on a prouvé l'inclusion souhaitée dans le membre de droite, ce qui prouve la convergence uniforme de θ_γ.
La preuve est achevée grâce à la continuité en $\mathbf{0}$ de $\theta \mapsto \mu^\theta$ pour la norme de la variation, ainsi que de celle de $\theta \mapsto \int f d\mu^\theta$ pour la norme de $\mathbb{R}^d$. ∎

Remarque III.2.4 *Donnons une illustration de cette proposition en dimension un.*
En fait, il apparaît que dans les cas pratiques, la notion de solution usuelle se conserve si l'on augmente la pénalisation.
Prenons ainsi une loi μ sur $\mathbb{R}$, dont le support est inclus dans une demi-droite $[M, +\infty[$ et dont la transformée de Laplace a pour domaine $]-\infty, a]$ avec $a \geq 0$.
On reprend les notations introduites au début de la démonstration précédente.
En raisonnant comme dans le numéro 2 des exemples précédents, la continuité et la croissance de $\xi + \Lambda_\gamma$ permettent d'écrire

$$Im(\xi + \Lambda_\gamma) =]-\infty, (\xi + \Lambda_\gamma)(a)]$$

On en déduit ainsi que

$$\forall \gamma' > \gamma,\ Im(\xi + \Lambda_{\gamma'}) \supset Im(\xi + \Lambda_\gamma)$$

avec égalité si et seulement si $a = 0$.
Cela signifie que l'ensemble des contraintes pour lesquelles la solution du problème pénalisé est usuelle croît avec la pénalisation, la croissance étant stricte si 0 est intérieur à D_μ (i.e. $a > 0$).
Enfin, si $a = 0$, l'existence d'une solution usuelle, pour une contrainte donnée est indépendante de la pénalisation.

III.2.4 Un contre-exemple à l'emploi d'une pénalisation forte

Dans le cas multidimensionnel ($d > 1$), si $\mathbf{0}$ n'est pas intérieur à D_μ, il se peut même qu'une augmentation de la pénalisation empêche parfois la solution d'être usuelle.
En effet, $Im(\xi + \Lambda_\gamma)$ ne croît pas forcément avec la pénalisation γ.
Prenons par exemple $S = \mathbb{R}^2$, $f = Id_{\mathbb{R}^2}$ et

$$\mu = \frac{1}{4}\left(\delta_{(1,0)} + \delta_{(0,1)}\right) + \frac{1}{2}\alpha$$

où α est la probabilité concentrée sur la diagonale positive $\Delta_+ = \{(x,x) \mid x \in \mathbb{R}_+\}$ de $\mathbb{R}^2$ et définie par

$$\forall A \in \mathcal{B}(\mathbb{R}^2),\ \alpha(A) = \tilde{\alpha}(\{x \mid (x,x) \in A\})$$

$\tilde{\alpha}$ étant une probabilité sur $\mathbb{R}_+$ dont la transformée de Laplace a pour ensemble de définition $D_{\tilde{\alpha}} =]-\infty, 0]$ et dont la moyenne m existe.
Avant de poursuivre, donnons à cela une interprétation financière.
Avec $K > 0$, on définit deux payoff sur $\mathbb{R}^3$:

$$f^1(x,y,z) = 1_{\{y<x\le K\}} + 1_{\{\min(x,y)>K\}}z$$

et

$$f^2(x,y,z) = f^1(y,x,z)$$

On considère trois sous-jacents $X_i,\ 1 \le i \le 3$ tels que la loi du couple (X_1, X_2) soit diffuse, symétrique et vérifie

$$P(\min(X_1,X_2) > K) = \frac{1}{2}$$

que celle de X_3 soit $\tilde{\alpha}$ et qu'enfin X_3 soit indépendante du couple (X_1, X_2).
Alors, il est aisé de vérifier :

$$P^{(f^1,f^2)(X_1,X_2,X_3)} = \mu$$

Avec ces sous-jacents, l'option définie par le payoff f^1 paye 1 si $X_2 < X_1 \le K$, une action X_3 si $\min(X_1,X_2) > K$ et 0 sinon.
Après avoir apporté cette interprétation, on se propose de prouver l'existence d'un point $\mathcal{C} \in Im(\xi)$ tel que $\mathbf{C}$ ne soit pas dans l'image de $\xi + \Lambda_\gamma$ à partir d'une certaine valeur de γ.
Dans ce cas, on voit que le problème non pénalisé associé à la mesure μ et à la contrainte $\mathbf{C}$ admet une solution usuelle alors qu'une pénalisation poussée fait perdre à la solution ce caractère usuel.
On doit donc étudier l'image de la fonction $\xi + \Lambda_\gamma$.
D'abord, l'application

$$\xi\ :\ \theta \longmapsto \int x d\mu^\theta(x) = \frac{\int x e^{\theta\cdot x} d\mu(x)}{\int e^{\theta\cdot x} d\mu(x)}$$

est définie sur $D_\mu = \{\theta \in \mathbb{R}^2 \mid \theta^1 + \theta^2 \le 0\}$.
Un morceau de la frontière de $\xi(D_\mu)$ est donné par le paramétrage suivant :

$$\mathcal{C}\ :\ u \in [-\infty, +\infty] \longmapsto \frac{\frac{1}{2}\left(e^u\begin{pmatrix}1\\0\end{pmatrix} + e^{-u}\begin{pmatrix}0\\1\end{pmatrix}\right) + m\begin{pmatrix}1\\1\end{pmatrix}}{\frac{1}{2}(e^u + e^{-u}) + 1}$$

Admettons provisoirement que l'image de $\xi + \Lambda_\gamma$ est sous le morceau de frontière définie par le paramétrage :

$$\mathcal{C}^\gamma\ :\ u \in]-\infty, +\infty[\longmapsto \frac{\frac{1}{2}\left(e^u\begin{pmatrix}1\\0\end{pmatrix} + e^{-u}\begin{pmatrix}0\\1\end{pmatrix}\right) + m\begin{pmatrix}1\\1\end{pmatrix}}{\frac{1}{2}(e^u + e^{-u}) + 1} + u\begin{pmatrix}\gamma_1\\-\gamma_2\end{pmatrix}$$

Il est facile de vérifier que pour γ_1 assez grand l'abscisse du paramétrage de $\mathcal{C}^\gamma$

$$A(u) \equiv \frac{\frac{1}{2}e^u + m}{\frac{1}{2}(e^u + e^{-u}) + 1} + \gamma_1 u$$

est une bijection de $\mathbb{R}$ sur $\mathbb{R}$, et qu'en particulier il existe un unique u_{γ_1} tel que

$$A(u_{\gamma_1}) = 0$$

Si l'on introduit la valeur de l'ordonnée du paramétrage de $\mathcal{C}^\gamma$ en u_{γ_1}

$$B(u_{\gamma_1}) \equiv \frac{\frac{1}{2}e^{-u_{\gamma_1}} + m}{\frac{1}{2}(e^{u_{\gamma_1}} + e^{-u_{\gamma_1}}) + 1} - \gamma_2 u_{\gamma_1}$$

On a

$$B(u_{\gamma_1}) - 1 = B(u_{\gamma_1}) + A(u_{\gamma_1}) - 1 = \frac{2m - 1}{\frac{1}{2}(e^{u_{\gamma_1}} + e^{-u_{\gamma_1}}) + 1} + (\gamma_1 - \gamma_2)u_{\gamma_1}$$

Supposons

$$m < \frac{1}{2}$$

Puisque l'on a aussi

$$u_{\gamma_1} < 0$$

on en déduit que pour $\gamma_1 \geq \gamma_2$ assez grand, on a

$$B(u_{\gamma_1}) < 1$$

et donc la courbe $\mathcal{C}^\gamma$ est strictement en dessous du point $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Comme $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ est un point frontière de $Im(\xi)$, on peut choisir un point $\mathbf{C}$ de $Im(\xi)$ assez proche de $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ pour que $\mathcal{C}^\gamma$ soit strictement en dessous de $\mathbf{C}$. Cette situation est illustrée à la Figure III.1 où l'on a fait la représentation graphique de $\mathcal{C}^\gamma$.

Il reste à montrer que l'image de $\xi + \Lambda_\gamma$ est elle même située en dessous de la courbe $\mathcal{C}^\gamma$ pour déduire que $\mathbf{C}$ n'est pas dans cette image. Amettons provisoirement qu'avec le paramétrage précédent de $\mathcal{C}^\gamma$, l'abscisse $u \in]-\infty, +\infty[\mapsto A(u)$ est croissante et que l'ordonnée $u \in [-\infty, +\infty] \mapsto B(u)$ est décroissante. L'idée est alors de prouver qu'il est possible de suivre un chemin partant de la frontière $\mathcal{C}^\gamma$ et contenu dans l'image de $\xi + \Lambda_\gamma$ tel que l'abscisse et l'ordonnée du point courant décroissent : cela signifiera précisément que l'image de $\xi + \Lambda_\gamma$ est située en dessous de $\mathcal{C}^\gamma$.

Dans cette perspective, si $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$ est dans D_μ, on peut relier $\mathcal{C}^\gamma$ à $(\xi + \Lambda_\gamma)(\theta)$ grâce à l'image par $\zeta + \Lambda_\gamma$ du segment tracé dans D_μ

$$\xi \in [0, 1] \longmapsto \delta_- \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \zeta\delta_+ \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

avec

$$\delta_- = \frac{\theta_1 - \theta_2}{2} \text{ et } \delta_+ = \frac{\theta_1 + \theta_2}{2}$$

Si $x(\zeta)$ et $y(\zeta)$ sont les coordonnées de cette image, on va alors montrer que ce sont des fonctions décroissantes.

On a

$$x(\zeta) = \frac{\frac{1}{2}e^{\delta_- + \zeta\delta_+} + \int xe^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)}{D}$$

en posant

$$D(\zeta) = \frac{1}{2}(e^{\delta_-} + e^{-\delta_-})e^{\zeta\delta_+} + \int e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)$$

Si $(\xi + \Lambda_\gamma)(\theta)$ n'est pas un point de $\mathcal{C}^\gamma$, on a aussi

$$\delta_+ < 0$$

Alors, en dérivant, on obtient pour $\zeta > 0$

$$\begin{aligned}
\frac{x'(\zeta)D(\zeta)^2}{\delta_+} &= \left(\frac{1}{2}e^{\delta_- + \zeta\delta_+} + 2\int x^2 e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\right) D(\zeta) \\
&\quad - \left(\frac{1}{2}e^{\delta_- + \zeta\delta_+} + \int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\right)\left(\frac{1}{2}(e^{\delta_-} + e^{-\delta_-})e^{\zeta\delta_+} + 2\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\right) \\
&\quad + \gamma_1 D(\zeta)^2 \\
&= \frac{1}{2}e^{\delta_- + \zeta\delta_+}\int e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x) \\
&\quad + (e^{\delta_-} + e^{-\delta_-})e^{\zeta\delta_+}\int x^2 e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x) \\
&\quad - \frac{1}{2}(e^{\delta_-} + e^{-\delta_-})e^{\zeta\delta_+}\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x) \\
&\quad + 2\left(\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x) - \left(\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\right)^2\right) \\
&\quad + \gamma_1 D(\zeta)^2
\end{aligned}$$

Il s'agit de prouver que le membre de gauche de l'expression précédente est positif.
Dans le membre de droite, seul le troisième terme est négatif, ce qui conduit à étudier par exemple le signe de

$$\begin{aligned}
E &\equiv \gamma_1 D(\zeta)^2 - \frac{1}{2}(e^{\delta_-} + e^{-\delta_-})e^{\zeta\delta_+}\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x) \\
&\geq \frac{1}{2}(e^{\delta_-} + e^{-\delta_-})\int e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\gamma_1 - \frac{1}{2}(e^{\delta_-} + e^{-\delta_-})\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x) \\
&\geq \frac{1}{2}(e^{\delta_-} + e^{-\delta_-})\int e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)\left(\gamma_1 - \frac{\int x e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)}{\int e^{2\zeta\delta_+ + x} d\tilde{\alpha}(x)}\right)
\end{aligned}$$

Finalement, compte tenu de la croissance de la fonction

$$u \longmapsto \frac{\int x e^{ux} d\tilde{\alpha}(x)}{\int e^{ux} d\tilde{\alpha}(x)}$$

sur $]-\infty, 0]$, en tant que dérivée de la fonction convexe $\ln \int e^{ux} d\tilde{\alpha}(x)$, on en déduit que E est positif dès que $\gamma_1 \geq m$, et la décroissance de $x(\zeta)$ suit.
De manière identique, la décroissance de $y(\zeta)$ est assurée dès que γ_2 est assez grand.
La preuve est alors achevée si l'on montre que $u \in]-\infty, +\infty[\mapsto A(u)$ est croissante et que $u \in]-\infty, +\infty[\mapsto B(u)$ est décroissante.
On calcule :

$$A'(u) = \frac{\frac{1}{2} + \frac{1}{2}e^u - \frac{m}{2}(e^u - e^{-u})}{(\frac{1}{2}(e^u + e^{-u}) + 1)^2} + \gamma_1$$

et on a $A'(u) > 0$ puisque $m < 1$.
De même :

$$B'(u) = -\frac{\frac{1}{2} + \frac{1}{2}e^{-u} - \frac{m}{2}(e^{-u} - e^{u})}{(\frac{1}{2}(e^{u} + e^{-u}) + 1)^2} - \gamma_2$$

et on a $B'(u) < 0$.

Chapitre IV

Théorème central limite associé au problème de minimisation d'une entropie

Avellaneda & all [6] suggèrent que la méthode de calibration par la minimisation de l'entropie relative permet une réduction de la variance de l'estimateur Monte-Carlo du prix d'une option. On a pu faire cette constatation dans l'expérience numérique du premier chapitre. Pour expliquer ce phénomène, l'idée avancée dans [6] est que la méthode effectue implicitement une régression linéaire de la fonction de payoff de l'option évaluée sur les contraintes.
Bien qu'une telle interprétation ne soit possible que dans le cas où la calibration est exacte, il est cependant intéressant de faire l'étude en regardant le cas pénalisé, le but fixé pour cette étude étant finalement de déterminer l'erreur Monte-Carlo de la méthode de calibration.
En suivant les généralisations antérieures, on considérera le cas d'un critère convexe convenable plutôt que le seul cas de l'entropie relative. Cependant, on verra aussi que c'est dans ce dernier cas que l'on peut faire apparaître la régression linéaire évoquée plus haut.
Pour reprendre le cadre général, on se donne une fonction strictement convexe $J^* : \mathbb{R} \longrightarrow \mathbb{R}$, croissante, dérivable et minorée ainsi qu'une fonction convexe sci $\chi^* : \mathbb{R} \longrightarrow]-\infty, +\infty]$ que l'on supposera nulle dans le cas contraint ou bien, dans le cas pénalisé, strictement convexe sur l'ensemble $dom\chi^* \equiv \{u \in \mathbb{R}^d,\ \chi^*(u) < +\infty\}$ (le "domaine" de χ^*), dérivable sur l'intérieur, supposé non vide de $dom\chi^*$, et admettant un minimum en $\mathbf{0}$.
Ceci permet un traitement simultané de ces deux cas en présentant le problème dual comme la recherche du minimum $(\theta_n, \theta_{0,n})$ de la fonction à valeurs dans $]-\infty, +\infty]$

$$\Psi : (u, u_0) \in E \times \mathbb{R} \longmapsto \int J^*(u \cdot f + u_0) 1_E(f) d\mu - u_0 + \chi^*(-u) \tag{IV.0.1}$$

avec l'hypothèse supplémentaire

$$(\tilde{C}) \begin{cases} \text{La condition } (C) \text{ est vérifiée et } E \text{ est le sous-espace qui lui est associé, dans le cas contraint} \\ \text{ou bien, dans le cas pénalisé :} \\ E = \mathbb{R}^d \text{ et } \chi^* \text{ est deux fois différentiable sur un voisinage de } -\theta \text{ où } D^2\chi^* \text{ est non dégénérée.} \end{cases}$$

où l'on rappelle que (θ, θ_0) est la solution du problème dual limite.
Dans la pratique, le voisinage est le plus grand possible, égal à $Int(dom\chi^*)$.

IV.1 Etude générale de l'erreur Monte-Carlo

IV.1.1 Théorème central limite associé à la suite des $\widehat{M}$-estimateurs

Commençons alors par établir un résultat relatif à la convergence de la suite $(\theta_n, \theta_{0,n})$. Pour espérer l'obtenir, il faut des hypothèses d'intégrabilité supplémentaires concernant les critères J^* et χ^*.

$$(H) \left\{ \begin{array}{l} J^* \text{ admet des dérivées jusqu'à l'ordre 3 sur un voisinage } \mathcal{U} \times \mathcal{U}_0 \text{ de } (\theta, \theta_0) \text{ dans } E \times \mathbb{R} \\ \int J^{*(k)}(\theta \cdot f + \theta_0) 1_E(f)(1 + \|f\|^k) d\mu < +\infty, \ k = 1, \ 2 \\ \int \sup_{(u,u_0) \in \mathcal{U} \times \mathcal{U}_0} |J^{*(3)}(u \cdot f + u_0)| 1_E(f)(1 + \|f\|^3) d\mu < +\infty \\ \chi^* \text{ admet des dérivées partielles bornées jusqu'à l'ordre 3 sur } -\mathcal{U} \end{array} \right.$$

Pour utiliser sans ambiguité des notations matricielles, on considèrera l'injection canonique de E dans $\mathbb{R}^d$ et l'on identifiera également les vecteurs de E ou de $E \times \mathbb{R}$ avec leur représentation matricielle dans la base canonique. On peut finalement considérer que χ^* est défini sur $\mathbb{R}^d \times \mathbb{R}$.

Théorème IV.1.1 *Supposons vérifiées les conditions $(\tilde{C})$, (H) ainsi que la condition d'intégrabilité supplémentaire*

$$\int \left(J^{*'}(\theta \cdot f + \theta_0) \right)^2 (1 + \|f\|^2) 1_E(f) d\mu < +\infty \tag{IV.1.1}$$

Alors, le vecteur de $E \times \mathbb{R}$

$$\sqrt{n} \begin{pmatrix} \theta_n - \theta \\ \theta_{0,n} - \theta_0 \end{pmatrix}$$

converge vers la loi normale de support $E \times \mathbb{R}$

$$\mathcal{N}\left(\mathbf{0}_{\mathbb{R}^d \times \mathbb{R}}, V(\theta, \theta_0) U(\theta, \theta_0) V(\theta, \theta_0) \right)$$

où $U(\theta, \theta_0)$ est défini par

$$U(\theta, \theta_0) = COV\left(J^{*'}(\theta \cdot f(X_1) + \theta_0) 1_{\{f(X_1) \in E\}} \begin{pmatrix} f(X_1) \\ 1 \end{pmatrix} \right)$$

et où $V(\theta, \theta_0)$ est la forme bilinéaire sur $\mathbb{R}^d \times \mathbb{R}$ de support $(E \times \mathbb{R}) \times (E \times \mathbb{R})$ et dont la restriction à $E \times \mathbb{R}$ est définie par $\left(\nabla^{(2)}_{\theta, \theta_0} \Psi \right)^{-1}_{|E \times \mathbb{R}}$, avec pour expression du Hessien

$$\nabla^{(2)}_{\theta, \theta_0} \Psi = E\left(J^{*(2)}(\theta \cdot f(X_1) + \theta_0) 1_{\{f(X_1) \in E\}} \begin{pmatrix} f(X_1) \\ 1 \end{pmatrix} \otimes \begin{pmatrix} f(X_1) \\ 1 \end{pmatrix} \right) + \nabla^{(2)}_{-\theta} \chi^*$$

Preuve : On peut vérifier que le Hessien de Ψ a bien l'expression donnée, ce qui revient à pouvoir dériver deux fois en (θ, θ_0) sous l'intégrale $\int J^*(u \cdot f + u_0) 1_E(f) d\mu$.
On utilise en effet la formule de Taylor avec reste intégral :

$$\begin{aligned} J^{*'}(u \cdot f + u_0) \quad &= \quad J^{*'}(\theta \cdot f + \theta_0)) + \frac{1}{2} J^{*(2)}(\theta \cdot f + \theta_0)((u - \theta) \cdot f + (u_0 - \theta_0)) \\ &\quad + \frac{1}{2} \int_0^1 J^{*(3)}((\theta + t(u - \theta)) \cdot f + (\theta_0 + t(u_0 - \theta_0))) dt ((u - \theta)) \cdot f + (u_0 - \theta_0))^2 \end{aligned}$$

et l'hypothèse (H) permet d'appliquer le théorème de dérivation, ce qui montre que l'on peut dériver une première fois sous l'intégrale $\int J^*(u \cdot f + u_0) 1_E(f) d\mu$.
On peut de même dériver une seconde fois en écrivant

$$\begin{aligned} J^{*(2)}(u \cdot f + u_0) &= J^{*(2)}(\theta \cdot f + \theta_0)) \\ &+ \int_0^1 J^{*(3)}((\theta + t(u-\theta)) \cdot f + (\theta_0 + t(u_0 - \theta_0)))dt((u-\theta)) \cdot f + (u_0 - \theta_0)) \end{aligned} \quad \text{(IV.1.2)}$$

Ce premier point étant établi, on va montrer que le Hessien de Ψ en (θ, θ_0) est non dégénéré sur $E \times \mathbb{R}$.
Avec la positivité de $J^{*(2)}$, conséquence de la convexité de J^*, ainsi que la convexité de χ^*, le Hessien est une forme bilinéaire positive dont il suffit de montrer qu'elle est de plus définie sur $E \times \mathbb{R}$.
Soit donc $(\alpha, \alpha_0) \in E \times \mathbb{R}$ tel que

$$\nabla^{(2)}_{\theta,\theta_0} \Psi \cdot (\alpha, \alpha_0) \cdot (\alpha, \alpha_0) = 0$$

On en déduit :

$$\int J^{*(2)}(\theta \cdot f + \theta_0)(\theta \cdot f + \theta_0)(\alpha \cdot f + \alpha_0)^2 1_E(f) d\mu + \nabla^{(2)}_{-\theta} \chi^* \cdot \alpha \cdot \alpha = 0$$

La condition $(\tilde{C})$ montre facilement la nullité de (α, α_0) et donc la non dégénérescence de $\nabla^{(2)}_{\theta,\theta_0} \Psi$.
Passons maintenant à la convergence en loi.
On pose

$$\forall (u, u_0) \in E \times \mathbb{R},\ \Psi_n(u, u_0) = \frac{\sum_{i=1}^n J^*(u \cdot f(X_i) + u_0) 1_{\{f(X_i) \in E\}}}{n} - u_0 + \chi^*(-u)$$

et l'on applique la formule de Taylor avec reste intégral à la différentielle $\nabla \Psi_n$ au minimum $(\theta_n, \theta_{0,n})$, qui est défini p.s. à partir d'un rang $N(\omega)$ (égal à 1 dans le cas pénalisé) :

$$\begin{aligned} \mathbf{0} &= \nabla \Psi_n(\theta_n, \theta_{0,n}) = \nabla \Psi_n(\theta, \theta_0) \\ &+ \int_0^1 (\nabla^{(2)} \Psi_n(\theta + t(\theta_n - \theta)), \theta_0 + t(\theta_{0,n} - \theta_0)) dt \cdot \begin{pmatrix} \theta_n - \theta \\ \theta_{0,n} - \theta_0 \end{pmatrix} \end{aligned} \quad \text{(IV.1.3)}$$

Montrons que

$$\nabla^{(2)} \Psi_n(\theta + t(\theta_n - \theta, \theta_{0,n} - \theta_0)) \to \nabla^{(2)} \Psi(\theta, \theta_0)\ ps$$

uniformément par rapport à $t \in [0, 1]$.
Si l'on écrit l'expression du Hessien de Ψ_n, un développement de Taylor semblable à celui de la relation (IV.1.3) permet d'établir

$$\begin{aligned} \nabla^{(2)} \Psi_n(\theta + t(\theta_n - \theta), \theta_0 + t(\theta_{0,n} - \theta_0)) &= \frac{1}{n} \sum_{k=1}^n J^{*(2)}(\theta \cdot f(X_k) + \theta_0) 1_{\{f(X_k) \in E\}} \begin{pmatrix} f(X_k) \\ 1 \end{pmatrix}^{\otimes 2} \\ &+ \nabla^{(2)} \chi^*(-\theta) + \sigma_n \cdot t(\theta_n - \theta, \theta_{0,n} - \theta_0) \end{aligned} \quad \text{(IV.1.4)}$$

où l'on a défini l'opérateur σ_n par

$$\begin{aligned} \sigma_n &= \frac{1}{n} \sum_{k=1}^n \int_0^1 J^{*(3)}((\theta + st(\theta_n - \theta)) \cdot f(X_k) + (\theta_0 + st(\theta_{0,n} - \theta_0)) ds 1_{\{f(X_k) \in E\}} \begin{pmatrix} f(X_k) \\ 1 \end{pmatrix}^{\otimes 3} \\ &- \int_0^1 \nabla^{(3)} \chi^*(-\theta - st(\theta_n - \theta)) ds \end{aligned}$$

Alors, pour n assez grand, on a :

$$\|\sigma_n\| \leq \frac{\sum_{k=1}^{n} \sup_{(u,u_0)\in\mathcal{U}} J^{*(3)}(u\cdot f(X_k)+u_0)1_{\{f(X_k)\in E\}}(1+\|f(X_k)\|^3)}{n} + \sup_{u\in\mathcal{U}} \| \bigtriangledown^{(3)} \chi^*(-u)\|$$

D'après la loi forte des grands nombres, et la troisième condition de l'hypothèse (H), on en déduit $\limsup \|\sigma_n\| < +\infty$, ce qui entraine bien la convergence uniforme souhaitée, grâce à (IV.1.4). C'est dire que le reste intégral de la formule (IV.1.3) converge ps vers $\bigtriangledown^{(2)}\Psi(\theta,\theta_0)$.
Avec la non dégénérescence de cette limite, on obtient ps pour $n \geq N$:

$$\sqrt{n}\begin{pmatrix} \theta_n - \theta \\ \theta_{0,n} - \theta_0 \end{pmatrix} = V(\theta,\theta_0)\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left(J^{*'}(\theta\cdot f(X_k)+\theta_0)1_{\{f(X_k)\in E\}}\begin{pmatrix} f(X_k) \\ 1 \end{pmatrix} - \begin{pmatrix} \nabla_{-\theta}\chi^* \\ 1 \end{pmatrix}\right) + \epsilon_n \tag{IV.1.5}$$

avec $\epsilon_n \to 0$ en probabilité, grâce au théorème central classique, valable en vertu de (IV.1.1). En effet, la condition de centrage découle de l'optimalité supposée de (θ,θ_0).
Comme $P(n \geq N)$ tend vers 1, on en déduit que le membre de gauche converge en loi comme le membre de droite avec la limite énoncée. ■

IV.1.2 Vitesse de convergence de la suite des estimateurs du prix

Après avoir établi ce premier résultat, passons à l'étude du comportement du prix. Pour une fonction test réelle φ donnée, on sait que la convergence a lieu presque sûrement dans le cas où se trouve réalisée la condition suivante :

$$\int \sup_{(u,u_0)\in\Delta\times U_0} J^{*(2)}(u\cdot f+u_0)(1+\|f\|)|\varphi|1_E(f)d\mu < +\infty \tag{IV.1.6}$$

où Δ est un voisinage de θ,θ_0 dans $E\times\mathbb{R}$.
Cette hypothèse est réalisée sous la condition plus forte

$$(H')\left\{\begin{array}{l} \int J^{*(k)}(\theta\cdot f+\theta_0)1_E(f)(1+\|f\|^k)|\varphi|d\mu < +\infty,\ k=1,\ 2 \\ \int \sup_{(u,u_0)\in\mathcal{U}\times\mathcal{U}_0} |J^{*(3)}(u\cdot f+u_0)|1_E(f)(1+\|f\|^3)|\varphi|d\mu < +\infty \end{array}\right.$$

Toutefois, pour pouvoir appliquer le théorème central classique, on est encore conduit à faire une hypothèse d'intégrabilité supplémentaire.

Théorème IV.1.2 *Supposons vérifiées les conditions* $(\tilde{C})$*, (H), (H'),* (IV.1.1) *ainsi que*

$$\int \left(J^{*'}(\theta\cdot f+\theta_0)\right)^2 \varphi^2 1_E(f)d\mu < +\infty \tag{IV.1.7}$$

Alors, la suite

$$\sqrt{n}\left(\int \varphi d\nu_n - \int \varphi d\mu_{J,E}^{\theta,\theta_0}\right)$$

converge vers la loi normale centrée

$$\mathcal{N}(0, B(\theta,\theta_0)A(\theta,\theta_0)^t B(\theta,\theta_0))$$

en définissant la matrice

$$A(\theta,\theta_0) = COV\left(J^{*'}(\theta\cdot f(X_1)+\theta_0)1_{\{f(X_1)\in E\}}\begin{pmatrix} f(X_1)\\ 1\\ \varphi(X_1)\end{pmatrix}\right)$$

et le vecteur ligne de $(\mathbb{R}^d\times\mathbb{R})\times\mathbb{R}$

$$B(\theta,\theta_0) = \begin{pmatrix} -{}^tF(\theta,\theta_0)V(\theta,\theta_0), & 1\end{pmatrix}$$

avec

$$F(\theta,\theta_0) = \int J^{*(2)}(\theta\cdot f+\theta_0)\varphi\begin{pmatrix} f\\ 1\end{pmatrix}1_{\{f\in E\}}d\mu$$

$V(\theta,\theta_0)$ *est l'opérateur défini dans le théorème précédent que l'on identifie avec sa matrice dans la base canonique de* $\mathbb{R}^d\times\mathbb{R}$.

Preuve : En introduisant l'erreur Δ_n sur le prix, on écrit

$$\begin{aligned}\sqrt{n}\Delta_n &= \int\varphi d\nu_n - \int\varphi d\mu^{\theta,\theta_0}_{J,E}\\ &= \sqrt{n}\left(\int J^{*'}(\theta\cdot f+\theta_0)\varphi d\mu_n - \int\varphi d\mu^{\theta,\theta_0}_{J,E}\right)\\ &\quad+\sqrt{n}\left(\int J^{*'}(\theta_n\cdot f+\theta_{0,n})\varphi d\mu_n - \int J^{*'}(\theta\cdot f+\theta_0)\varphi d\mu_n\right)\end{aligned}\tag{IV.1.8}$$

Le premier terme du second membre relève du théorème central limite classique.
Pour le second, en raisonnant comme pour la relation (IV.1.5), en utilisant l'hypothèse (H'), on établit, φ étant scalaire :

$$\begin{aligned}\sqrt{n}\bigg(\int J^{*'}(\theta_n\cdot f+\theta_{0,n})\varphi d\mu_n - \int J^{*'}(\theta\cdot f+\theta_0)\varphi d\mu_n\bigg) &=\\ &\sqrt{n}(\theta_n-\theta)\cdot\int J^{*(2)}(\theta\cdot f+\theta_0)\varphi f 1_{\{f\in E\}}d\mu\\ &+\sqrt{n}(\theta_{0,n}-\theta_0)\cdot\int J^{*(2)}(\theta\cdot f+\theta_0)\varphi 1_{\{f\in E\}}d\mu\\ &+\epsilon'_n\end{aligned}\tag{IV.1.9}$$

avec $\epsilon'_n\to 0$ en probabilité.
La relation (IV.1.8) s'écrit alors, compte tenu de (IV.1.5)

$$\begin{aligned}\sqrt{n}\Delta_n &= -{}^tF(\theta,\theta_0)V(\theta,\theta_0)\frac{1}{\sqrt{n}}\sum_{k=1}^n\left(J^{*'}(\theta\cdot f(X_k)+\theta_0)\begin{pmatrix} f(X_k)\\ 1\end{pmatrix}1_{\{f(X_k)\in E\}} - \begin{pmatrix}\nabla_{-\theta}\chi^*\\ 1\end{pmatrix}\right)\\ &\quad+\frac{1}{\sqrt{n}}\sum_{k=1}^n\left((J^{*'}(\theta\cdot f(X_k)+\theta_0)\varphi(X_k)1_{\{f(X_k)\in E\}} - \int\varphi d\mu^{\theta,\theta_0}_{J,E})\right)+\epsilon_n\end{aligned}\tag{IV.1.10}$$

avec $\epsilon_n\to 0$ en probabilité.
Cela s'écrit

$$\sqrt{n}\Delta_n = B(\theta,\theta_0)Z_n+\epsilon_n\tag{IV.1.11}$$

avec $\epsilon_n\to 0$ en probabilité et en considèrant le vecteur aléatoire de $\mathbb{R}^{d+2}$

$$Z_n = \frac{1}{\sqrt{n}}\sum_{k=1}^n\begin{pmatrix} J^{*'}(\theta\cdot f(X_k)+\theta_0)f(X_k)1_{\{f(X_k)\in E\}} - \nabla_{-\theta}\chi^*\\ J^{*'}(\theta\cdot f(X_k)+\theta_0)1_{\{f(X_k)\in E\}}-1\\ J^{*'}(\theta\cdot f(X_k)+\theta_0)1_{\{f(X_k)\in E\}}\varphi(X_k) - \int\varphi d\mu^{\theta,\theta_0}_{J,E})\end{pmatrix}$$

Avec (IV.1.1) et (IV.1.7) le théorème central limite s'applique au vecteur Z_n qui converge en loi vers la loi normale centrée $\mathcal{N}(\mathbf{0}_{\mathbb{R}^{d+2}}, A(\theta, \theta_0))$, ce qui permet d'établir le résultat. ∎

Remarque IV.1.3 *On étend facilement le théorème au cas où φ est vectoriel, de dimension p. En effet, l'expression (IV.1.10) de l'erreur est valable à condition de poser*

$$F(\theta, \theta_0) = \int J^{*(2)}(\theta \cdot f(x) + \theta_0) \begin{pmatrix} f(x) \\ 1 \end{pmatrix} \otimes \varphi(x) 1_{\{f(x) \in E\}} \mu(dx)$$

On aboutit alors à la généralisation de (IV.1.11) :

$$\sqrt{n}\Delta_n = B(\theta, \theta_0) Z_n + \epsilon_n \tag{IV.1.12}$$

avec

$$B(\theta, \theta_0) = \begin{pmatrix} -{}^tF(\theta, \theta_0) V(\theta, \theta_0) & I_p \end{pmatrix}$$

Avec ces modifications, la loi limite est donnée par $\mathcal{N}(\mathbf{0}_{\mathbb{R}^p}, B(\theta, \theta_0) A(\theta, \theta_0) {}^tB(\theta, \theta_0))$ où $A(\theta, \theta_0))$ est défini comme précédemment.

IV.2 Cas de l'entropie relative

Lorsque l'on adopte comme critère J l'entropie relative, les résultats de convergence précédents sont plus maniables compte tenu de la forme de la conjuguée du critère :

$$\forall u \in \mathbb{R},\ J^*(u) = e^{u-1}$$

et cette expression reste en particulier inchangée lors des dérivations successives qui interviennent. On a de plus une simplification des conditions qui permettent d'assurer la validité de ces résultats de convergence. On peut en effet remplacer la condition (H) par

$$(\tilde{H}) \begin{cases} \theta \in I = Int_E\{u \in E \mid \int e^{u \cdot f(x)} 1_{\{f(x) \in E\}} \mu(dx) < +\infty\} \\ \chi^* \text{ admet des dérivées partielles bornées jusqu'à l'ordre 3 sur un voisinage } -\mathcal{U} \text{ de } -\theta \text{ dans } E \end{cases}$$

Montrons en effet que (H) tient sous cette dernière hypothèse. En suivant un raisonnement déjà établi, on considère $2^{d'}$ vecteurs ρ_l de E épuisant les relations

$$\rho_l = \theta + \epsilon \sum_{i=1}^{d'} \pm e_i$$

où $(e_i, 1 \leq i \leq d')$ est une base orthonormée de E, supposée de dimension $d' \geq 1$ et où $\epsilon > 0$ est assez petit pour satisfaire à :

$$\overline{B}(\theta, \sqrt{d'}\epsilon) \bigcap E \subset I$$

On a alors la relation

$$\forall u \in \overline{B}(\theta, \epsilon) \bigcap E,\ \forall y \in E,\ e^{u \cdot y} \leq \sum_{i=1}^{2^{d'}} e^{\rho_i \cdot y}$$

ce qui permet d'obtenir (H).
Exactement de la même façon, l'hypothèse (H') peut être remplacée par la condition plus commode :

$$(\tilde{H}') \quad \theta \in I_\varphi = Int_E \left\{ u \in E \;\middle|\; \int e^{u \cdot f(x)} |\varphi(x)| 1_{\{f(x) \in E\}} \mu(dx) < +\infty \right\}$$

Montrons à présent que la calibration par entropie relative effectue implicitement la régression de la fonction de payoff sur les contraintes ainsi qu'il est suggéré dans [6]. On se place donc dans le cadre de la calibration exacte ($\chi^* = 0$).
Dans ce cas, compte tenu de la forme particulière de J^* ainsi que de la contrainte, la restriction de l'opérateur $V(\theta, \theta_0)$ à son support $(E \times \mathbb{R}) \times (E \times \mathbb{R})$ a pour expression particulière

$$V(\theta, \theta_0)_{|E\times\mathbb{R}} = \begin{pmatrix} \left(\int f \otimes f d\mu_E^\theta\right)^{-1}_{E\times E} & \mathbf{0}_{\mathbb{R}^d} \\ {}^t\mathbf{0}_{\mathbb{R}^d} & 1 \end{pmatrix} \tag{IV.2.1}$$

En injectant d'abord cette expression dans (IV.1.5), on obtient :

$$\begin{aligned} \sqrt{n}(\theta_n - \theta) &= V(\theta)\frac{1}{\sqrt{n}}\sum_{k=1}^n e^{\theta\cdot f(X_k)+\theta_0-1}1_{\{f(X_k)\in E\}}f(X_k) \\ &\quad +\epsilon_n \end{aligned} \tag{IV.2.2}$$

avec :

$$V(\theta) = \left(\int f \otimes f d\mu_E^\theta\right)^{-1}_{E\times E}$$

On peut alors énoncer :

Théorème IV.2.1 *Supposons vérifiées l'hypothèse (C) ainsi que les conditions*

$$\theta \in I$$

et

$$\int e^{2\theta\cdot f(x)}\|f(x)\|^2 1_{\{f(x)\in E\}}\mu(dx) < +\infty$$

Alors, la suite

$$\sqrt{n}(\theta_n - \theta)$$

converge vers la loi normale centrée de support E

$$N(0, V(\theta)U(\theta)V(\theta))$$

où $U(\theta)$ est défini par

$$U(\theta) = COV(e^{\theta\cdot f(X_1)+\theta_0-1}1_{\{f(X_1)\in E\}}f(X_1))$$

En utilisant encore l'expression (IV.2.1), on obtient de même, pour le prix :

Théorème IV.2.2 *Supposons*

$$\theta \in I \bigcap I_\varphi$$

et

$$\int e^{2\theta\cdot f(x)}(1 + \|f(x)\|^2 + \|\varphi(x)\|^2)1_{\{f(x)\in E\}}\mu(dx) < +\infty$$

Alors, la suite

$$\sqrt{n}\left(\int \varphi d\nu_n - \int \varphi d\mu_E^\theta\right)$$

converge vers la loi normale centrée de support E

$$N(0, B(\theta)A(\theta)^t B(\theta))$$

en définissant la matrice

$$A(\theta) = COV\left(e^{\theta\cdot f(X_1)+\theta_0-1}1_{\{f(X_1)\in E\}}\begin{pmatrix} f(X_1) \\ \varphi(X_1) \\ 1 \end{pmatrix}\right)$$

et le vecteur ligne

$$B(\theta) = \left(\ -^tF(\theta)V(\theta), -\int \varphi d\mu_E^\theta, 1\ \right)$$

avec :

$$F(\theta) = \int \varphi f d\mu_E^\theta$$

Il est intéressant de faire apparaitre la possibilité d'une réduction de variance.
Pour ce faire, on injecte l'expression (IV.2.1) dans celle de l'erreur (IV.1.10), toujours avec $\chi^* = 0$, et l'on obtient, après simplification

$$\begin{aligned}\Delta_n &= \frac{1}{n}\sum_{k=1}^n \left(-\int \varphi f d\mu_E^\theta \cdot \left(\int f\otimes f d\mu_E^\theta\right)^{-1}_{E\times E} f(X_k) - \int \varphi d\mu_E^\theta + \varphi(X_k)\right) \\ &\quad e^{\theta\cdot f(X_k)+\theta_0-1}1_{\{f(X_k)\in E\}} + \frac{\epsilon_n}{\sqrt{n}} \\ &= \frac{1}{n}\sum_{k=1}^n e^{\theta\cdot f(X_k)+\theta_0-1}1_{\{f(X_k)\in E\}}\left(\varphi(X_k) - \Pi_f^a(\varphi)(X_k))\right) + \frac{\epsilon_n}{\sqrt{n}}\end{aligned}$$

avec $\epsilon_n \to 0$ en probabilité et en faisant apparaître la régression affine $\Pi_f^a(\varphi)$ de φ sur les $f^i, 1 \leq i \leq d$ sous la probabilité μ_E^θ.
En appliquant le théorème central limite, on trouve comme variance associée celle, sous μ de la variable centrée

$$e^{\theta\cdot f+\theta_0-1}1_{\{f\in E\}}(\varphi - \Pi_f^a(\varphi))$$

ou bien celle, sous μ_E^θ de

$$\left(e^{\theta\cdot f+\theta_0-1}1_{\{f\in E\}}\right)^{\frac{1}{2}}(\varphi - \Pi_f^a(\varphi))$$

Sous les hypothèses du théorème IV.1.2, la variance associée au théorème central limite est donc proche de celle de $\varphi - \Pi_f^a(\varphi)$ sous μ lorsque θ est faible. On pourra ainsi espérer une valeur inférieure à la variance de φ sous la mesure a priori μ, proche de la variance sous μ_E^θ si θ est faible.
En conclusion, si on part d'une mesure a priori proche de la calibration, la minimisation de l'entropie relative permet bien d'améliorer la convergence du Monte-Carlo.

Pour terminer, on illustre numériquement le théorème central limite dans le cadre des simulations présenté au premier chapitre. Les tirages Monte-Carlo se font donc sous le modèle a priori (I.2.9) discrétisé suivant le schéma d'Euler. Avec les contraintes données par le tableau de la Figure I.1, on lance 1500 Monte-Carlo calibrés par minimisation de l'entropie relative. Chacune des simulations s'effectue avec 5000 tirages. On donne la densité normalisée du multiplicateur de Lagrange θ^1 (Figure IV.1) ainsi que celle du prix calibré du Put de strike $K = 105$ et de maturité $T = 1$ (Figure IV.2).

Chapitre V

Stabilité de la calibration

Dans cette partie, on s'intéresse à la stabilité de la méthode de calibration face aux erreurs numériques dans le programme de minimisation (déterminant les multiplicateurs de Lagrange) et également face à la variation des prix d'options qui fixent les contraintes.
Pour simplifier le cadre de l'étude, on se restreint à l'entropie relative usuelle en envisageant seulement le cas de contraintes exactes.
Dans la pratique, à un rang n donné, le minimiseur θ_n est approché numériquement par un estimateur $\tilde{\theta_n}$ que l'on recherche dans un compact, voisinage supposé de θ, où la transformation de Laplace Ψ^e atteint son minimum, en définissant :

$$\forall u \in \mathbb{R}^d, \; \Psi^e(u) = \int e^{u \cdot f(x)} 1_{\{f(x) \in E\}} \mu(dx)$$

Le théorème central limite reste valable si l'erreur d'approximation décroît plus vite que l'inverse de $\sqrt{n}$.
De manière effective, l'erreur que l'on contrôle est celle qui porte sur la valeur minimale de la transformée de Laplace discrète Ψ_n^e définie pour tout $u \in \mathbb{R}^d$ par :

$$\Psi_n^e(u) = \frac{\sum_{i=1}^n e^{u \cdot f(X_i)} 1_{\{f(X_i) \in E\}}}{n}$$

On peut encore démontrer la validité du théorème central limite si l'erreur d'approximation tend assez rapidement vers 0.
On doit commencer par s'assurer de la convergence des estimateurs, ce qu'établit le résultat suivant :

Proposition V.0.3 *Supposons que*

$$\theta \in I = Int_E\{u \in E \mid \Psi^e(u) < +\infty\}$$

et supposons aussi que K est un voisinage compact de θ dans E, inclus dans I. Alors, si $(\tilde{\theta}_n)$ est une suite d'estimateurs de θ dans K vérifiant

$$\Psi_n^e(\tilde{\theta}_n) - \Psi_n^e(\theta_n) \to 0 \;\; P\text{-}p.s.$$

on a :

$$\tilde{\theta}_n \to \theta \;\; P\text{-}p.s.$$

De plus, si φ est une fonction mesurable réelle telle que

$$\theta \in I_\varphi = Int_E\left\{u \in E \mid \int e^{u \cdot f(x)} \|\varphi(x)\| 1_{\{f(x) \in E\}} \mu(dx) < +\infty\right\}$$

alors

$$\int \varphi d\tilde{\nu}_n \to \int \varphi d\mu_E^\theta \ P\text{-}p.s.$$

où $\tilde{\nu}_n$ est défini en substituant $\tilde{\theta}_n$ à θ_n dans l'expression de ν_n.

Cette dernière convergence appliquée à une suite dénombrable déterminante de fonctions continues bornées implique la convergence étroite de $\tilde{\nu}_n$ vers μ_θ presque sûrement.
Preuve : I étant convexe, on peut supposer que K l'est également.
Remarquons que $D^2\Psi^e$ est définie positive sur tout I, en tant que forme bilinéaire sur E, puisque $Int_E(Conv(Supp(\mu^f))) \neq \emptyset$ et qu'elle y est également continue.
Avec la compacité de K, cela prouve qu'il existe $\sigma_0 > 0$ tel que

$$(A0)\ \forall u \in K,\ \forall v \in E,\ D_u^2\Psi^e.v.v \geq \sigma_0\|v\|^2$$

L'essentiel de la preuve sera acquis si l'on prouve qu'il existe $\sigma > 0$ tel que p.s. pour n assez grand, la relation suivante soit valable :

$$(A1)\ \ \forall u \in K,\ \Psi_n^e(u) - \Psi_n^e(\theta_n) \geq \sigma\|u - \theta_n\|^2$$

sachant également que θ_n se trouve dans K à partir d'un certain rang.
Pour montrer $(A1)$, on commence par appliquer la formule de Taylor à l'ordre 2 à Ψ_n^e en θ_n, ps pour n assez grand :

$$\forall u \in K,\ \Psi_n^e(u) = \Psi_n^e(\theta_n) + \int_0^1 D^2\Psi_n^e(\theta_n + t(u-\theta_n))(1-t)dt \cdot (u-\theta_n) \cdot (u-\theta_n)$$

Il suffit alors de prouver $D^2\Psi_n^e$ converge vers $D^2\Psi^e$ uniformément sur K ps.
Cela prouvera en effet, grâce à la convexité de K que l'intégrale apparaissant dans le second membre de la formule de Taylor converge vers $\int_0^1 D^2\Psi^e(\theta+t(u-\theta))(1-t)dt$, et ceci uniformément en $u \in K$, compact par hypothèse.
Pour finir, la compacité de K et la majoration $(A0)$ permettent la minoration uniforme de l'intégrale, ce qui établit $(A1)$ avec $0 < \sigma < \frac{\sigma_0}{2}$.
Prouvons donc la convergence uniforme souhaitée.
Grâce à la compacité de K, pour $\alpha > 0$ fixé assez petit, on peut se donner une suite de points $u_i \in K,\ 1 \leq i \leq N_\alpha$ telle que les boules fermées de centres u_i et de rayon α recouvrent K et soient incluses dans I.
On écrit ps, pour n assez grand, par la formule de Taylor et une majoration facile :

$$\forall i,\ 1 \leq i \leq N_\alpha,\ \forall u, v \in \overline{B}(u_i, \alpha),\ \|D^2\Psi_n^e(v) - D^2\Psi_n^e(u)\| \leq L_{\alpha,i,n}\|v-u\|$$

avec

$$L_{\alpha,i,n} = \sum_{l=1}^{l=2^{d'}} \frac{\sum_{k=1}^n e^{(u_i+\alpha\rho_l)\cdot f(X_k)} 1_{\{f(X_k)\in E\}}\|f(X_k)\|^3}{n}$$

où les $2^{d'}$ vecteurs ρ_l épuisent les relations

$$\rho_l = d'^{-\frac{1}{2}} \sum_{i=1}^{d'} \pm e_i$$

$(e_i, 1 \leq i \leq d')$ étant une base orthonormée de E.
Par la loi forte des grands nombres, $L_{\alpha,i,n}$ converge p.s. avec n, vers une limite majorée par :

$$L = 2^{d'} \max_{u \in K_0} \int e^{u\cdot f} 1_{f\in E}\|f\|^3 d\mu < \infty$$

où K_0 est un voisinage compact de K dans I.

On a alors la majoration :

$$\begin{aligned}\max_{u\in K}\|D^2\Psi_n^e(u)-D^2\Psi^e(u)\| &\leq \max_{u,v\in K,\|v-u\|\leq\alpha}\|D^2\Psi^e(v)-D^2\Psi^e(u)\| \\ &\quad+\max_{1\leq i\leq N_\alpha}\|D^2\Psi_n^e(u_i)-D^2\Psi^e(u_i)\|+\alpha\max_{1\leq i\leq N_\alpha}L_{\alpha,i,n}\alpha\end{aligned}$$

et l'on en déduit :

$$\limsup_n \max_{u\in K}\|D^2\Psi_n^e(u)-D^2\Psi^e(u)\|\leq \max_{u,v\in K,\|v-u\|\leq\alpha}\|D^2\Psi^e(v)-D^2\Psi^e(u)\|+L\alpha$$

En utilisant la continuité uniforme de $D^2\Psi^e$ sur K, la convergence uniforme est ainsi établie et $(A1)$ est prouvée.
La convergence de $\Psi_n^e(\tilde\theta_n)$ permet d'en déduire alors celle de $\tilde\theta_n$ et, ce résultat étant établi, la convergence de $\int\varphi d\tilde\nu_n$ s'obtient alors avec les arguments employés antérieurement pour celle de $\int\varphi d\nu_n$, à condition de justifier l'argument d'équi-intégrabilité par le fait que $\Psi_n^e(\tilde\theta_n)$ est bornée presque sûrement. ■

Proposition V.0.4 *Supposons comme précédemment que K est un voisinage compact de θ inclus dans I et que*

$$\Psi_n^e(\tilde\theta_n)-\Psi_n^e(\theta_n)\to 0 \ P\text{-}ps$$

Si de plus

$$n(\Psi_n^e(\tilde\theta_n)-\Psi_n^e(\theta_n))\to 0 \ \textit{en probabilité}$$

alors, sous leurs hypothèses respectives, les théorèmes IV.2.1 et IV.2.2 restent valables à condition de remplacer θ_n par $\tilde\theta_n$ et ν_n par $\tilde\nu_n$.

Preuve : On montre le théorème central limite relatif à $\tilde\theta_n$ en prouvant que la suite $\sqrt{n}\|\tilde\theta_n-\theta_n\|$ tend vers 0 en probabilité.
En effet, ps à partir d'un certain rang, on a :

$$\Psi_n^e(\tilde\theta_n)-\Psi_n^e(\theta_n)\geq\sigma\|\tilde\theta_n-\theta_n\|^2$$

d'après la relation $(A1)$ établie au début de la preuve précédente, ce qui permet de conclure avec l'hypothèse du théorème.
Pour le théorème de la limite centrale concernant $\int\varphi d\tilde\nu_n$, il suffit de même de prouver que $\sqrt{n}\delta_n$ tend vers 0 en probabilité, en posant :

$$\delta_n=\frac{\Phi_n^e(\tilde\theta_n)}{\Psi_n^e(\tilde\theta_n)}-\frac{\Phi_n^e(\theta_n)}{\Psi_n^e(\theta_n)}$$

en posant pour tout u de E : :

$$\Phi_n^e(u)=\frac{\sum_{k=1}^n e^{u\cdot f(X_k)}\varphi(X_k)1_{\{f(X_k)\in E\}}}{n}$$

On réécrit δ_n sous la forme :

$$\sqrt{n}\delta_n = \frac{1}{\Psi_n^e(\tilde{\theta}_n)} \left(\sqrt{n}(\Phi_n^e(\tilde{\theta}_n) - \Phi_n^e(\theta_n)) - \frac{\Phi_n^e(\theta_n)}{\Psi_n^e(\theta_n)} \sqrt{n}(\Psi_n^e(\tilde{\theta}_n) - \Psi_n^e(\theta_n)) \right)$$

Sous les hypothèses du théorème IV.2.2, on sait que $\Psi_n^e(\theta_n)$ et $\Phi_n^e(\theta_n)$ convergent ps, ce qui constitue la loi forte des grands nombres pour ν_n, et qui ce retrouve, en prouvant selon la technique habituelle que Ψ_n^e et Φ_n^e convergent uniformément sur un voisinage compact de θ.
La première hypothèse de la proposition montre également que $\Psi_n^e(\tilde{\theta_n})$ converge.
Compte tenu de ceci, et en utilisant la seconde hypothèse, le deuxième terme du second membre tend vers 0 en probabilité.
Pour le premier terme, on établit :

$$\sqrt{n}(\Phi_n^e(\tilde{\theta}_n) - \Phi_n^e(\theta_n)) = \int_0^1 D\Phi_n^e(\theta_n + t(\tilde{\theta}_n - \theta_n))dt \cdot \sqrt{n}(\tilde{\theta}_n - \theta_n)$$

Par un argument de convergence uniforme habituel, l'intégrale du second membre converge ps, ce qui montre, avec la seconde hypothèse de la proposition, la convergence en probabilité du premier membre, donc de $\sqrt{n}\delta_n$, vers 0. ∎

Les résultats précédents ont prouvé que la méthode par minimisation de l'entropie relative est stable si le vecteur de contraintes C, ramené à $\mathbf{0}$, est parfaitement connu.
Compte tenu de l'imprécision la concernant, il est également nécessaire de regarder la stabilité par rapport aux variations des contraintes.
De ce point de vue, une condition nécessaire de stabilité est

$$\mathbf{0} \in Int(Conv(Supp(\mu^f)))$$

ce qui se traduit par $E = \mathbb{R}^d$.
Autrement, $\mathbf{0}$ serait sur la frontière de $(Conv(Supp(\mu^f)))$ et une variation de la contrainte $\mathbf{C}$ autour de $\mathbf{0}$ impliquerait l'absence de solution au problème, ou une solution non équivalente en temps que mesure (E serait différent) : le problème serait mal posé.
De plus, avec le choix $E = \mathbb{R}^d$, la mesure solution μ_E^θ est équivalente à μ, ce qui est désirable, la mesure a priori μ jouant essentiellement le rôle de déterminer les évènements néligeables et ceux qui ne le sont pas.
Ceci étant supposé, l'étude de la stabilité dans le cas limite (nombre de tirages Monte-Carlo infini) se fait par le calcul des sensibilités par rapport aux contraintes.
Les sensibilités s'interprètent elles-mêmes comme les coefficients de la régression linéaire de la fonction test φ sur f, sous la mesure solution μ_E^θ (voir par exemple [6]).
Il est intéressant de rappeler que le calcul de l'erreur associée au tcl fait également apparaître les coefficient de la même régression linéaire, sous ν_n cette fois, ce qui permet une estimation simultanée de l'erreur et des sensibilités.

Proposition V.0.5 *Supposons*

$$\mathbf{0} \in Int(Conv(Supp(\mu^f)))$$

et $\theta \in I \bigcap I_\varphi$.
Pour $\epsilon > 0$ *assez petit, presque sûrement à partir d'un certain rang, pour tout* $C \in B(\mathbf{0}, \epsilon)$ *, le problème de minimisation d'entropie sous la contrainte* C *admet une solution* ν_n^C *équivalente à* μ_n.
De plus, il existe $a > 0$ *tel que ps à partir d'un certain rang, pour tout* $C \in B(\mathbf{0}, \epsilon)$

$$\left| \int \varphi d\nu_n^C - \int \varphi d\nu_n \right| \leq a\|C\| \tag{V.0.1}$$

Remarque V.0.6 *Pour faire le lien avec ce qui vient d'être dit, on aurait un énoncé équivalent dans le cas limite.*
Pour ϵ assez petit, on peut alors associer un nombre a_∞ proche de la norme de la différentielle, c'est-à-dire de la sensibilité, du prix $\int \varphi d\nu^C$ par rapport à C, tel que :

$$\left| \int \varphi d\nu_\infty^C - \int \varphi d\mu^\theta \right| \leq a_\infty \|C\| \tag{V.0.2}$$

où ν_∞^C est la solution du problème de calibration limite sous les contraintes C.
Ainsi qu'il apparaitra au cours de la preuve, tout nombre a strictement supérieur à a_∞ convient dans la proposition, ce qui permet de se ramener ici également au calcul de la sensibilité.

Preuve : Choisissons $\epsilon_1 > 0$ tel que $B(\mathbf{0}, \epsilon_1) \subset Int(Conv(Supp(\mu^f)))$, ce qui est possible avec l'hypothèse de la proposition, et prenons $\epsilon < \epsilon_1$.
On montre que p.s. l'inclusion $B(\mathbf{0}, \epsilon) \subset Int(Conv(Supp(\mu_n^f)))$ est vraie à partir d'un rang N, ce qui permettra d'affirmer l'existence de la solution ν_n^C à partir du rang N pour tout $C \in B(\mathbf{0}, \epsilon)$.
Pour montrer l'inclusion souhaitée, on reprend l'argument habituel en remarquant d'abord qu'il existe un polyèdre contenant $B(\mathbf{0}, \epsilon_1)$ et dont les sommets appartiennent à $Supp(\mu^f)$.
Pour conclure, il suffit ensuite d'appliquer le Lemme I.2.3 ainsi que le lemme de Borel-Cantelli.
Après avoir établi cela, il s'agit d'examiner pour $n \geq N$ l'expression

$$\delta_n^C = \int \varphi d\nu_n^C - \int \varphi d\nu_n$$

Pour cela, posons pour tout u de $\mathbb{R}^d$

$$\delta_n(u) = \frac{\Phi_n^e(u)}{\Psi_n^e(u)} - \frac{\Phi_n^e(\theta_n)}{\Psi_n^e(\theta_n)} \tag{V.0.3}$$

On va lui appliquer la formule de Taylor en réarangeant l'expression de la dérivée :

$$\begin{aligned} \frac{d}{du}\left(\frac{\Phi_n^e(u)}{\Psi_n^e(u)}\right) &= \frac{\int \varphi f e^{u\cdot f} d\mu_n}{\int e^{u\cdot f} d\mu_n} - \frac{\int \varphi e^{u\cdot f} d\mu_n \int f e^{u\cdot f} d\mu_n}{(\int e^{u\cdot f} d\mu_n)^2} \\ &= \frac{\int \varphi(f - \Gamma_n(u)) e^{u\cdot f} d\mu_n}{\int e^{u\cdot f} d\mu_n} \end{aligned}$$

où l'on a posé

$$\forall u \in \mathbb{R}^d,\ \Gamma_n(u) = \frac{\int f e^{u\cdot f} d\mu_n}{\int e^{u\cdot f} d\mu_n}$$

La formule de Taylor appliquée à la différence (V.0.3) donne donc

$$\delta_n(u) = \int_0^1 \frac{\int \varphi(f - \Gamma_n(\theta_n + t(u - \theta_n)) e^{(\theta_n + t(u-\theta_n))\cdot f} d\mu_n}{\int e^{(\theta_n + t(u-\theta_n))\cdot f} d\mu_n} dt \cdot (u - \theta_n)$$

En prenant f à la place de φ, on établit pareillement, compte tenu aussi de $\Gamma_n(\theta_n) = \mathbf{0}$ (par définition de θ_n) :

$$\Gamma_n(u) = \int_0^1 \frac{\int (f - \Gamma_n(\theta_n + t(u - \theta_n)) \otimes (f - \Gamma_n(\theta_n + t(u - \theta_n)) e^{(\theta_n + t(u-\theta_n))\cdot f} d\mu_n}{\int e^{(\theta_n + t(u-\theta_n))\cdot f} d\mu_n} dt \cdot (u - \theta_n)$$

Considérons ensuite $\alpha > 0$ tel que la boule $\overline{B}(\theta, \alpha)$ soit incluse dans $I \bigcap I_\varphi$ ce qui est possible grâce à l'hypothèse de l'énoncé.
Alors, suivant une technique habituelle, on peut établir des résultats de convergence uniforme montrant qu'à partir d'un certain rang (tel qu'en particulier θ_n appartienne à $\overline{B}(\theta, \alpha)$), pour tout u de $\overline{B}(\theta, \alpha)$ on a :

$$|\delta_n(u)| \leq b \|u - \theta_n\|$$

et

$$\|u - \theta_n\| \leq c \|\Gamma_n(u)\|$$

avec

$$\begin{aligned} b &> \sup_{u \in \overline{B}(\theta,\alpha)} \left\| \int_0^1 \frac{\int \varphi(f - \Gamma(\theta + t(u-\theta))e^{(\theta+t(u-\theta))\cdot f} d\mu}{\int e^{(\theta+t(u-\theta))\cdot f} d\mu} dt \right\| \\ c &> \sup_{u \in \overline{B}(\theta,\alpha)} \left\| \left(\int_0^1 \frac{\int (f - \Gamma(\theta + t(u-\theta)) \otimes (f - \Gamma(\theta + t(u-\theta))e^{(\theta+t(u-\theta))\cdot f} d\mu}{\int e^{(\theta+t(u-\theta))\cdot f} d\mu} dt \right)^{-1} \right\| \end{aligned} \tag{V.0.4}$$

en posant également

$$\forall v \in \mathbb{R}^d, \ \Gamma(v) = \frac{\int f e^{v\cdot f} d\mu}{\int e^{v\cdot f} d\mu}$$

Le membre de droite de (V.0.4) minorant le choix de b est bien défini si α est choisi suffisamment petit car alors l'opérateur à inverser se trouve sur un voisinage assez petit de $\frac{\int (f-\Gamma(\theta)) \otimes (f-\Gamma(\theta)) e^{\theta\cdot f} d\mu}{\int e^{\theta \cdot f} d\mu}$ qui est inversible grâce à la condition géométrique (C).

Le théorème sera acquis avec $a = bc$ si l'on prouve que pour ϵ assez petit, on a, à partir d'un certain rang :

$$\forall C \in B(\mathbf{0}, \epsilon), \ \theta_n^C \in B(\theta, \alpha) \tag{V.0.5}$$

en posant :

$$\theta_n^C = arg\min \left\{ \int e^{u\cdot(f-C)} d\mu_n \right\}$$

On remarque que θ_n^C est caractérisée également par l'équation :

$$\Gamma_n(\theta_n^C) = C \tag{V.0.6}$$

Pour prouver l'existence de ϵ, en rappelant que Γ est un difféomorphisme sur I nul en $\mathbf{0}$, on va montrer qu'un candidat acceptable est $\epsilon_0 > 0$ choisi tel que

$$\overline{B}(\mathbf{0}, \epsilon_0) \subset \Gamma(B(\theta, \alpha)) \tag{V.0.7}$$

En utilisant la convergence uniforme de Γ_n sur le compact $\overline{B}(\theta, \alpha)$ de I, on en déduit que pour $\eta > 0$, à partir d'un certain rang, presque sûrement on a :

$$B(\mathbf{0}, \epsilon_0) \subset \Gamma_n(B(\theta, \alpha)) + B(\mathbf{0}, \eta) \tag{V.0.8}$$

Pour obtenir le résultat d'existence de ϵ, on montre les relations suivantes, à savoir qu'à partir d'un certain rang, presque-sûrement on a :

$$B(\mathbf{0}, \epsilon_0) \bigcap \Gamma_n(B(\theta, \alpha)) \neq \emptyset \tag{V.0.9}$$

$$B(\mathbf{0}, \epsilon_0) \bigcap Fr(\Gamma_n(B(\theta, \alpha))) = \emptyset \tag{V.0.10}$$

La relation (V.0.9) se déduit immédiatement de (V.0.8) en prenant $\eta < \epsilon_0$ (ou bien simplement en utilisant la convergence vers θ de θ_n où Γ_n s'annule).
Si l'on admet (V.0.10), il est alors aisé de conclure.
En effet, puisque $B(\mathbf{0}, \epsilon_0)$ est ouvert, l'égalité (V.0.10) s'écrit en terme de frontière relative dans $B(\mathbf{0}, \epsilon_0)$:

$$Fr_{B(\mathbf{0},\epsilon_0)}(\Gamma_n(B(\theta,\alpha)) \bigcap B(\mathbf{0},\epsilon_0)) = \emptyset \tag{V.0.11}$$

Ainsi, l'ensemble $\Gamma_n(B(\theta,\alpha)) \bigcap B(\mathbf{0},\epsilon_0)$ est ouvert et fermé dans $B(\mathbf{0},\epsilon_0)$ et, étant de plus non vide d'après (V.0.9), on conclut donc par la connexité de $B(\mathbf{0},\epsilon_0)$)

$$\Gamma_n(B(\theta,\alpha)) \bigcap B(\mathbf{0},\epsilon_0) = B(\mathbf{0},\epsilon_0)$$

ce qui signifie qu'à partir d'un certain rang, on a l'inclusion $B(\mathbf{0},\epsilon_0) \subset \Gamma_n(B(\theta,\alpha))$, et donc, avec l'équation caractéristique (V.0.6), la relation (V.0.5) est vérifiée avec $\epsilon = \epsilon_0$.
Pour prouver (V.0.10), il suffit d'établir le petit lemme qui suit. ∎

Lemme V.0.7 *Pour ϵ_0 satisfaisant la condition (V.0.7), on a presque sûrement à partir d'un certain rang*

$$d(\overline{B}(\mathbf{0},\epsilon_0), Fr(\Gamma_n(\overline{B}(\theta,\alpha)))) > 0 \tag{V.0.12}$$

Preuve : D'abord, on peut trouver $x_n \in \overline{B}(\theta,\alpha))$ et $y_n \in Fr(\Gamma_n(\overline{B}(\theta,\alpha)))$ réalisant

$$d(\overline{B}(\mathbf{0},\epsilon_0), Fr(\Gamma_n(\overline{B}(\theta,\alpha)))) = \|x_n - y_n\|$$

En introduisant $z_n \in Fr(\Gamma(\overline{B}(\theta,\alpha)))$, on obtient alors par l'inégalité triangulaire :

$$\|x_n - y_n\| \geq d(\overline{B}(\mathbf{0},\epsilon_0), Fr(\Gamma(\overline{B}(\theta,\alpha)))) - \|y_n - z_n\| \tag{V.0.13}$$

Avec les égalités $Fr(\Gamma_n(\overline{B}(\theta,\alpha))) = \Gamma_n(S(\theta,\alpha))$ et $Fr(\Gamma(\overline{B}(\theta,\alpha))) = \Gamma(S(\theta,\alpha))$ et du fait que Γ_n converge vers Γ uniformément sur $\overline{B}(\theta,\alpha)$, on en déduit que l'on peut choisir z_n tel que la suite $\|z_n - y_n\|$ tend vers 0.
D'autre part, avec l'inclusion supposée, le premier terme du second membre de l'inégalité (V.0.13) est non nul, si bien qu'il en est de même de $\|x_n - y_n\|$ à partir d'un certain rang, ce qui achève de prouver le lemme. ∎

Chapitre VI

Approximation de la probabilité martingale calibrée minimale

Ce chapitre est consacré à l'important problème de l'absence d'opportunité d'arbitrage, qui se traduit de façon courante par le fait que sous un changement de probabilité équivalente, le sous-jacent actualisé est une martingale.
Dans le cas où ce dernier est une suite finie de variables aléatoires, l'équivalence entre ces deux conditions est établie dans [63] si l'espace des états du hasard est de cardinal fini, puis elle a été énoncée sans cette restriction dans [40], ce qui constitue le Théorème de Dalang-Morton-Willinger (voir aussi [116] ainsi que [41] pour une présentation didactique).
Par ailleurs, dans le cas où le processus sous-jacent est à temps continu, le résultat n'est plus vrai et est toutefois examiné dans [120] qui montre que la condition de martingalité est nécessaire si l'on renforce l'absence d'opportunité d'arbitrage, dans le cas d'un processus borné dont les temps de sauts sont prévisibles et en nombre fini.
De manière générale, l'existence d'une probabilité martingale équivalente à la probabilité de départ permet d'assurer l'absence d'arbitrage et c'est sous cette condition, suffisante mais non nécessaire, qu'il est commode d'envisager le problème de calibration : un pricing en cohérence avec les prix d'options observés et dépourvu de possibilité d'arbitrage peut se faire sous une probabilité martingale calibrée.
Le problème du choix d'une probabilité martingale se pose dans le contexte des marchés incomplets et a par exemple été étudié dans [53], en retenant l'entropie relative comme critère de sélection.
On se propose ici d'étudier le problème de minimisation de l'entropie relative parmi les probabilités martingales calibrées.
Pour l'écrire précisément, on suppose que le sous-jacent $(S_t)_{t\in[0,T]}$ est un processus càdlàg à valeurs dans $\mathbb{R}^m$ avec $T > 0$ et on se place en fait dans la situation où c'est le processus canonique sur l'espace de Skorokhod des fonctions càdlàg m-dimensionnelles $\mathbb{D}_T$ et l'on note $(\mathcal{D}_t^0)_{0\leq t\leq T}$ la filtration canonique .
Pour compléter le cadre probabiliste, on se donne comme d'habitude une probabilité a priori μ, ce qui permet de considérer l'augmentation de $(\mathcal{D}_t^0)_{0\leq t\leq T}$, notée $(\mathcal{F}_t^0)_{0\leq t\leq T}$ en la complétant avec les ensembles μ-négligeables dont on notera $\mathcal{N}_\mu$ leur collection.
Finalement, on pose :

$$\forall 0 \leq t < T,\ \mathcal{F}_t = \bigcap_{t<u\leq T} \mathcal{F}_u^0$$

et l'on pose aussi

$$\mathcal{F}_T = \mathcal{F}_T^0$$

Quitte enfin à remplacer le sous-jacent par le processus actualisé, on considère que le taux sans risque est nul et l'on peut alors définir le problème de calibration suivant, en choisissant un critère convexe général J :

$$(PL^J_\infty)\left\{\begin{array}{l}\text{trouver } \nu \in \mathcal{P}(\mathbb{D}_T) \text{ qui minimise } H_J(\nu\|\mu)\\ \text{sous la contrainte que } (S_t) \text{ est une } \nu\text{-martingale et } \int f d\nu = \mathbf{0}\end{array}\right.$$

ce qui revient à trouver la probabilité minimale sur l'ensemble

$$\mathcal{M}_\infty = \left\{\nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \;\middle|\; (S_t) \text{ est une } \nu\text{-martingale et } \int f d\nu = \mathbf{0}\right\}$$

qu'il sera commode d'identifier à un sous-ensemble convexe de $L^1(\mathbb{D}_T, \mu)$.
Bien que l'on ne s'intéresse pas ici à la convergence Monte-carlo, on suppose que J possède les propriétés dont on l'avait muni dans cette étude.
Dans le cas de l'entropie relative, les résultats de Csiszar permettent d'obtenir l'existence de la solution du problème en un sens généralisé (I-projection) dés que $\exists \nu \in \mathcal{M}_\infty,\ H(\nu\|\mu) < +\infty$, la solution étant usuelle si $\mathcal{M}_\infty$ est fermé pour la norme de la variation (c'est-à-dire fortement dans $L^1(\mathbb{D}_T, \mu)$).
De manière plus générale, le Théorème 1 de [123] permet d'énoncer que

Théorème VI.0.8 *Si $\mathcal{A} \subset L^1(\mathbb{D}_T, \mu)$ est faiblement fermé et s'il existe $\varphi \in \mathcal{A}$ tel que $H_J(\varphi\mu\|\mu) < +\infty$, alors il existe $\varphi \in \mathcal{A}$ telle que $H_J(\varphi\mu\|\mu)$ soit minimale sur $\mathcal{A}$*

et ce résultat n'utilise des propriétés de J que le fait qu'elle est convexe, continue et vérifie $\lim_{u\to+\infty} \frac{J(u)}{u} = +\infty$.
En paraphrasant le théorème précédent, on énonce ainsi le résultat d'existence suivant :

Théorème VI.0.9 *Si l'ensemble des densités par rapport à μ des éléments de $\mathcal{M}_\infty$ est faiblement fermé dans $L^1(\mathbb{D}_T, \mu)$ et s'il existe $\nu \in \mathcal{M}_\infty,\ H_J(\nu\|\mu) < +\infty$ alors le problème de calibration (PL^J_∞) admet une solution ν_∞.*

Ici, l'unicité résulte bien sûr du fait qu'il s'agit de la recherche du minimum d'une fonction strictement convexe sur un ensemble convexe.
On supposera désormais que sous la probabilité a priori μ le processus sous-jacent S ainsi que le vecteur de payoff f sont bornés, ce qui permet d'assurer que l'ensemble $\mathcal{M}_\infty$ est faiblement fermé dans $L^1(\mathbb{D}_T, \mu)$.
Avec le théorème précédent, la solution du problème de calibration existe dès qu'il existe une probabilité martingale calibrée située à distance finie de la mesure a priori, l'objectif principal de l'étude qui suit étant alors d'approcher la solution du problème (PL^J_∞) en se donnant un nombre fini de contraintes.
Dans le cadre markovien, on formulera un nouveau problème de calibration et on l'approchera également avec des contraintes plus simples à condition de choisir le critère usuel de l'entropie relative, ce qui plaide de nouveau en faveur de cette dernière. Pour être complet, les aspects concernant la pénalisation seront également étudiés.
Cette étude rend ainsi compte de la possibilité de cohérence approchée de la calibration par Monte-Carlo avec le principe d'absence d'opportunité d'arbitrage.
En effet, puisque l'espace $\mathbb{D}_T$ est polonais, les résultats de Monte-Carlo exhibés antérieurement s'appliquent et on pourra obtenir une convergence Monte-Carlo vers la solution du problème martingale approché.
Bien entendu, dans le cadre de la simulation, on considère des trajectoires discrétisées en temps et l'on aborde également cet aspect pour obtenir le procédé de convergence.

VI.1 Convergence du problème de martingales approché

Dans le contexte d'un pricing par la méthode de Monte-Carlo, on n'envisage qu'un nombre fini de contraintes, ce qui fait que l'absence d'opportunité d'arbitrage ne sera qu'approximative.
Même si en pratique le nombre de contraintes retenues ne doit pas être trop important, une justification théorique est que les contraintes de martingalité du sous-jacent peuvent se réduire au cas dénombrable, ce qui permet ainsi d'envisager de les approcher par des contraintes en nombre fini.
On énonce :

Théorème VI.1.1 *On peut trouver une suite $(\varphi_n)_{n\in\mathbb{N}^*}$ de fonctions de $L^\infty(\mathbb{D}_T,\mu)$ telles que pour tout $\nu \in \mathcal{P}(\mathbb{D}_T)$ vérifiant $\nu \ll \mu$, le processus (S_t) est une ν-martingale si et seulement si :*

$$\forall n \in \mathbb{N}^* \int \varphi_n d\nu = 0 \tag{VI.1.1}$$

Preuve : Il s'agit d'établir qu'une suite de relations du type (VI.1.1) suffit à assurer que pour $0 \leq s < t \leq T$ et $B \in \mathcal{F}_s$ on a l'égalité

$$\int 1_B(S_t - S_s)d\nu = \mathbf{0} \tag{VI.1.2}$$

On peut identifier $\mathcal{D}_s^0$ à la tribu Borélienne de l'espace des trajectoires réelles càdlàg m-dimensionnelles sur [0,s] : comme c'est un espace Polonais donc séparable pour la topologie de Skorokhod, il possède une base dénombrable d'ouverts $(O_{s,i})_{i\in\mathbb{N}}$.
Supposons alors que pour tout couple (s,t) de rationnels tels que $0 \leq s < t \leq T$ et tout i entier, on ait :

$$\int 1_{O_{s,i}}(S_t - S_s)d\nu = \mathbf{0} \tag{VI.1.3}$$

et montrons que la relation (VI.1.2) est vérifiée pour tout $0 \leq s < t \leq T$.
En fixant d'abord s et t rationnels tels que $0 \leq s < t \leq T$ et en supposant que $O \in \mathcal{D}_s^0$ est ouvert, il est limite d'une suite dénombrable croissante d'ouverts s'écrivant comme réunion finie d'éléments de $(O_{s,i})_{i\in\mathbb{N}}$ et pour lesquels la relation (VI.1.2) est vraie en vertu de (VI.1.3) : on obtient alors par le théorème de la convergence dominée que (VI.1.2) est vraie pour $B = O$.
Si maintenant B est arbitraire dans $\mathcal{D}_s^0$, par régularité extérieure de la mesure Borélienne ν restreinte à $\mathcal{F}_s$, on peut trouver une suite décroissante d'ouverts $O_i \in \mathcal{F}_s$ convergeant presque sûrement vers B : on conclut en utilisant encore le théorème de la convergence dominée que (VI.1.2) est également vraie dans ce cas.
Après avoir ainsi établi le fait que la relation (VI.1.2) est vraie si s et t sont rationnels et pour $B \in \mathcal{D}_s^0$, montrons qu'elle est vraie dans le cas général.
Prenons donc $0 \leq s < t \leq T$, $B \in \mathcal{F}_s$ et considérons s' et t' rationnels tels que $s \leq s' < t \leq t' \leq T$.
Avec la définition de $\mathcal{F}_s$, on peut trouver $\tilde{B} \in \bigcap_{s<u\leq T} \mathcal{D}_u^0$ ne différant de B que d'un ensemble μ-négligeable.
Puisque $\tilde{B} \in \mathcal{D}_{s'}^0$, ce qui précède permet d'établir l'égalité :

$$\int 1_{\tilde{B}}(S_{t'} - S_{s'})d\nu = \mathbf{0}$$

Ensuite, en faisant décroître s' vers s et t' vers t, en utilisant la continuité à droite du processus sous-jacent $(S_u)_{u\in[0,T]}$ et en rappelant également que l'on a supposé que ce dernier est borné, on obtient finalement que la relation (VI.1.2) est vraie pour s et t quelconques et B dans $\bigcap_{s<u\leq T} \mathcal{D}_u^0$

donc dans $\mathcal{F}_s$ puisque les ensembles μ-négligeables sont ν-négligeables : cela traduit le fait que $(S_u)_{u\in[0,T]}$ est une martingale sous ν. Les relations (VI.1.3) écrites pour les couples de rationnels (s,t) tels que $0 \le s < t \le T$ et pour $i \in \mathbb{N}$ se réécrivent enfin comme (VI.1.1). ∎

Remarque VI.1.2 *Une façon pratique de réaliser la suite de contraintes $\varphi_n \in L^\infty(\mathbb{D}_T,\mu)$ du théorème précédent est de procéder à la discrétisation temporelle du sous-jacent.*
Pour cela, on se donne une suite croissante de sous-ensembles finis $(I_i)_{i\in\mathbb{N}^*}$ de [0,T] et on suppose que leur réunion $I_\infty = \bigcup_{i\in\mathbb{N}^*} I_i$ est dense dans [0,T].
Pour tout $j\in\mathbb{N}^*$, l'ensemble $\mathcal{C}((\mathbb{R}^m)^j)$ des fonctions continues à support compact sur $(\mathbb{R}^m)^j$ est séparable et possède donc une famille dense $(\gamma_{j,k})_{k\in\mathbb{N}^*}$.
Alors, pour $s\in I_\infty$ et $i\in\mathbb{N}^*$, on considère la suite finie $\{t_1,t_2,...,t_{n_{i,s}}\}$ d'éléments de I_i inférieurs ou égaux à s, si bien que l'on peut définir pour $t\in I_\infty$ avec t>s la famille de contraintes :

$$\int (S_t - S_s)\gamma_{n_{i,s},k}(S_{t_1},...,S_{t_{n_{i,s}}})d\nu = \mathbf{0}$$

L'étude de la discrétisation en temps sera approfondie dans le cadre markovien.

En se donnant une suite $\varphi_n \in L^\infty(\mathbb{D}_T,\mu)$ telle que définie dans le théorème précédent, on peut définir le problème approché pour $n\in\mathbb{N}^*$

$$(PL_n^J)\left\{\begin{array}{l}\text{trouver } \nu\in\mathcal{P}(\mathbb{D}_T) \text{ qui minimise } H_J(\nu\|\mu)\\ \text{sous les contraintes } \int f d\nu = \mathbf{0} \text{ et } \int \varphi_i d\nu = 0,\ 1\le i\le n\end{array}\right.$$

et l'on pose

$$\mathcal{M}_n = \left\{\nu\in\mathcal{P}(\mathbb{D}_T),\ \nu\ll\mu \;\middle|\; \int f d\nu = \mathbf{0} \text{ et } \int \varphi_i d\nu = 0,\ 1\le i\le n\right\}$$

Avec le Théorème VI.0.8, on énonce le théorème d'existence suivant, similaire au Théorème VI.0.9 :

Théorème VI.1.3 *S'il existe $\nu\in\mathcal{M}_n$, $H_J(\nu\|\mu) < +\infty$ alors le problème de calibration (PL_n^J) admet une solution (usuelle) unique ν_n.*

En effet, avec les hypothèses faites après le Théorème VI.0.9, on a supposé que f est bornée et de plus il en est de même des φ_i, ce qui fait que l'ensemble $\mathcal{M}_n$ est faiblement fermé dans $L^1(\mathbb{D}_T,\mu)$.
La convergence de la suite ν_n des solutions de (PL_n^J) est celle classique du problème des moments et est résolu grâce à des résultats de [123] généralisant ceux de [21] et [22] dans le cadre qui nous intéresse.
On considère que sont en force les hypothèses faites après le Théorème VI.0.9.

Théorème VI.1.4 *Supposons qu'il existe $\nu\in\mathcal{M}_\infty$, $H_J(\nu\|\mu) < +\infty$, si bien que la solution ν_∞ du problème (PL_∞^J) existe de même que celle ν_n du problème (PL_n^J).*
Alors la suite $H_J(\nu_n\|\mu)$ croît vers $H_J(\nu_\infty\|\mu)$ et ν_n tend vers ν_∞ en variation.

Preuve : Les ensembles $\mathcal{M}_n$ forment une suite décroissante d'ensembles convexes et de plus, avec le Théorème VI.1.1, on note l'égalité :

$$\mathcal{M}_\infty = \bigcap_{n=1}^{\infty} \mathcal{M}_n$$

La première convergence relève de l'exemple IV 2 de [123], reprenant le Théorème 3.1 de [22], tandis que la seconde est une conséquence du Lemme 2 de [123]. ∎

Remarque VI.1.5 *Il semble que la preuve des résultats de [123] soit vraie seulement dans le cas d'ensembles convexes alors qu'ils sont énoncés sans cette hypothèse.*
Le Théorème VI.0.8 est également une version corrigée de [123] qui affirmait l'unicité de la solution sans hypothèse de convexité.

En reprenant le cadre de la Remarque VI.1.2, on en déduit que l'on peut approcher le problème de calibration de la probabilité martingale en discrétisant temporellement le processus d'abord, puis en sélectionnant un nombre fini de contraintes.

VI.2 Cas markovien

Le choix des contraintes de martingalité φ_n est un problème important qui se prête à une simplification si l'on impose que le sous-jacent est un processus de Markov, à condition également d'adopter l'entropie relative comme critère de minimisation.
On rappelle que la propriété de Markov peut se voir comme la traduction mathématique du fait que l'information du marché est parfaite. Elle est vérifiée dans la plupart des modèles des mathématiques financières, avec pour conséquence pratique importante que le pricing des options ne fait pas intervenir l'historique des prix du sous-jacent.
Le problème modifié est

$$(PL_\infty^{\text{mark}}) \begin{cases} \text{trouver } \nu \in \mathcal{P}(\mathbb{D}_T) \text{ qui minimise } H(\nu \| \mu) \\ \text{sous la contrainte que } (S_t) \text{ est une } \nu\text{-martingale markovienne et } \int f d\nu = \mathbf{0} \end{cases}$$

ce qui revient à trouver la probabilité minimale sur l'ensemble

$$\mathcal{M}_\infty^{\text{mark}} = \left\{ \nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \,\middle|\, (S_t) \text{ est une } \nu\text{-martingale markovienne et } \int f d\nu = \mathbf{0} \right\}$$

Ce dernier ensemble n'est pas convexe, toutefois, on a :

Lemme VI.2.1 *L'ensemble $\mathcal{M}_\infty^{mark}$ est (fortement) fermé dans $L^1(\mathbb{D}_T, \mu)$.*

Preuve : L'ensemble $\mathcal{M}_\infty^{\text{mark}}$ est l'intersection de l'ensemble $\mathcal{M}_\infty$ qui est fortement fermé avec les hypothèses faites après le Théorème VI.0.9 et de l'ensemble

$$\mathcal{M}^{\text{mark}} = \left\{ \nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \,\middle|\, (S_t) \text{ est un processus markovien sous } \nu \right\}$$

et il suffit donc de montrer que ce dernier ensemble est fortement fermé.
Soit donc une suite α_n de $\mathcal{M}_\infty$ qui converge fortement vers α dans $L^1(\mathbb{D}_T, \mu)$.
Pour $0 \le s < t \le T$ et $\varphi \in \mathcal{C}_c(\mathbb{R}^m)$, c'est-à-dire une fonction continue à support compact, on calcule l'espérance conditionnelle de $\varphi(S_t)$ sachant $\mathcal{F}_s$ sous la probabilité α_n et l'on a par hypothèse l'existence d'une variable aléatoire Ψ_n mesurable par rapport à la tribu engendrée par S_s telle que :

$$\mathbb{E}^{\alpha_n}\left(\varphi(S_t) | \mathcal{F}_s\right) = \frac{\mathbb{E}\left(\frac{d\alpha_n}{d\mu} \varphi(S_t) \middle| \mathcal{F}_s \right)}{\mathbb{E}\left(\frac{d\alpha_n}{d\mu} \middle| \mathcal{F}_s \right)} = \Psi_n \ \mu\text{-ps} \tag{VI.2.1}$$

La convergence forte de $\frac{d\alpha_n}{d\mu}$ implique la convergence dans $L^1(\mathbb{D}_T, \mu)$ du numérateur et du dénominateur du second membre de la double égalité précédente, et, quitte à extraire une sous-suite, on peut suppposer que la convergence est μ-presque sûre et donc, en prenant la limite Ψ de Ψ_n, elle est mesurable par rapport à la tribu engendrée par S_s et l'on a :

$$\mathbb{E}^\alpha\left(\varphi(S_t)|\mathcal{F}_s\right) = \frac{\mathbb{E}\left(\frac{d\alpha}{d\mu}\varphi(S_t)\middle|\mathcal{F}_s\right)}{\mathbb{E}\left(\frac{d\alpha}{d\mu}\middle|\mathcal{F}_s\right)} = \Psi \ \mu\text{-ps}$$

ce qui donne :

$$\mathbb{E}^\alpha\left(\varphi(S_t)|\mathcal{F}_s\right) = \mathbb{E}^\alpha\left(\varphi(S_t)|\sigma(S_s)\right)$$

et assure la propriété de Markov. ∎

Du fait de la non convexité de l'ensemble $\mathcal{M}_\infty$, la fermeture forte de ce dernier ne suffit pas à assurer sa fermeture faible et on ne peut donc utiliser le Théorème VI.0.8 pour trouver la solution du problème $(PL_\infty^{\text{mark}})$, ce qui pourra cependant se faire grâce à un procédé d'approximation.

VI.2.1 Etude du problème discrétisé en temps

Pour passer au problème discret, on échantillonne le processus sous-jacent en considérant, suivant la Remarque VI.1.2 une suite croissante de sous-ensembles finis $(I_i)_{i\in\mathbb{N}^*}$ de [0,T] de réunion dense dans [0,T].
En se fixant $i_0 \in \mathbb{N}^*$, on pose $I_{i_0} = \{t_0, t_1, ..., t_M\}$ et l'on obtient la version discrétisée du processus $(S_{t_i})_{0\leq i\leq M}$.
On munit alors S de la filtration discrète $(\mathcal{F}_{t_i})_{0\leq i\leq M}$.
D'autre part, on se place dans la situation pratique où les options servant à la calibration sont du type "plain vanilla" : chacune verse un seul flux à une maturité donnée.
Sans perdre en généralité, on suppose que les maturités considérées sont les instants d'échantillonnage t_i, quitte à les rajouter à I_{i_0} et quitte à leur associer éventuellement un flux nul, ce qui permet d'écrire le payoff des options de calibration :

$$f(S) = \sum_{i=1}^{M} g_i(S_{t_i}) \tag{VI.2.2}$$

où les composantes non nulles du vecteur aléatoire d-dimensionnel $g_i(S_{t_i})$ sont les flux des options de calibration d'échéance t_i.
L'espace $\mathcal{C}_c(\mathbb{R}^m)$ des fonctions continues à support compact sur $\mathbb{R}^m$ étant séparable, on peut trouver une suite (γ_n) dense dans $\mathcal{C}_c(\mathbb{R}^m)$.
On définit le problème :

$$(PL_{n,i_0}^{\text{mark}}) \begin{cases} \text{trouver } \nu \in \mathcal{P}(\mathbb{D}_T) \text{ qui minimise } H(\nu\|\mu) \\ \text{sous les contraintes } \int f d\nu = \mathbf{0} \\ \text{et } \int \gamma_i(S_{t_j})(S_{t_{j+1}} - S_{t_j})d\nu = \mathbf{0},\ 1\leq i\leq n,\ 0\leq j\leq M-1 \end{cases}$$

et l'on pose

$$\mathcal{M}_{n,i_0}^{\text{mark}} = \left\{\nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \,\middle|\, \int f d\nu = \mathbf{0} \text{ et } \int \gamma_i(S_{t_j})(S_{t_{j+1}} - S_{t_j})d\nu = \mathbf{0},\ 1\leq i\leq n,\ 0\leq j\leq M-1\right\}$$

Il est également naturel d'introduire le problème limite discrétisé en temps, par analogie avec $(PL_{\infty}^{\text{mark}})$:

$$(PL_{\infty,i_0}^{\text{mark}}) \begin{cases} \text{trouver } \nu \in \mathcal{P}(\mathbb{D}_T) \text{ qui minimise } H(\nu\|\mu) \\ \text{sous la contrainte que } (S_{t_j})_{0\leq j\leq M} \text{ est une } \nu\text{-martingale markovienne et } \int f d\nu = \mathbf{0} \end{cases}$$

ce qui revient à trouver la probabilité minimale sur l'ensemble

$$\mathcal{M}_{\infty,i_0}^{\text{mark}} = \left\{ \nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \;\middle|\; (S_{t_j})_{0\leq j\leq M} \text{ est une } \nu\text{-martingale markovienne et } \int f d\nu = \mathbf{0} \right\}$$

L'étude de la convergence est résumée dans le résultat ci-dessus :

Théorème VI.2.2 *Supposons qu'il existe $\nu \in \mathcal{M}_{\infty,i_0}^{mark}$ équivalente à μ telle que $H(\nu\|\mu) < +\infty$ et supposons de plus que le processus sous-jacent $(S_{t_i})_{0\leq i\leq M}$ soit markovien sous μ.*
Alors le problème (PL_{∞,i_0}^{mark}) admet une solution unique ν_{∞,i_0}. Pour tout n le problème (PL_{n,i_0}^{mark}) admet une solution unique ν_n et de plus le processus $(S_{t_i})_{0\leq i\leq M}$ conserve la propriété de Markov sous ν_n.
D'autre part, la suite $H(\nu_n\|\mu)$ croît vers $H(\nu_{\infty,i_0}\|\mu)$ et ν_n tend vers ν_{∞,i_0} en variation.

En supposant que le sous-jacent soit markovien sous μ, on est cohérent avec l'hypothèse d'un marché parfait, ainsi qu'on l'a déjà évoqué, et l'on annonce ici que la calibration par minimisation de l'entropie relative permet de garder cette cohérence.
La difficulté empêchant d'appliquer les résultats classiques de convergence est que l'ensemble $\mathcal{M}_{\infty,i_0}^{\text{mark}}$ n'est ni faiblement fermé ni convexe, ce qui ne rend pas évident l'existence et l'unicité de la solution ν_{∞}.
Preuve : L'existence de la solution ν_n du problème de rang n se montre avec le Théorème VI.0.8 dont les conditions d'application sont vérifiées puisque $\mathcal{M}_{n,i_0}^{\text{mark}}$ est faiblement fermé dans $L^1(\mathbb{D}_T, \mu)$ et que de plus $\mathcal{M}_{\infty,i_0}^{\text{mark}} \subset \mathcal{M}_{n,i_0}^{\text{mark}}$, ce qui permet d'avoir l'existence de $\nu \in \mathcal{M}_n^{\text{mark}},\ H(\nu\|\mu) < +\infty$.
L'unicité est celle du minimum d'une fonction strictement convexe sur un ensemble convexe.
Pour passer à la suite, on explicite la solution ν_n comme une distribution de Boltzmann.
Avec l'hypothèse du théorème et l'inclusion ensembliste $\mathcal{M}_{\infty,i_0}^{\text{mark}} \subset \mathcal{M}_{n,i_0}^{\text{mark}}$, ce dernier ensemble contient une probabilité équivalente à μ, ce qui montre que la solution ν_n du problème (PL_n^{mark}) est également équivalente à μ (en se reportant à l'étude générale de la condition géométrique d'existence du problème de calibration avec contraintes linéaires en nombre fini).
En définitive, la probabilité ν_n s'écrit :

$$\nu_n = e^{\sum_{j=1}^{M} u_j \cdot g_j(S_{t_j}) + \sum_{j=1}^{M} \sum_{i=1}^{n} v_{i,j} \cdot \gamma_i(S_{t_{j-1}})(S_{t_j} - S_{t_{j-1}}) - w_0} \mu \tag{VI.2.3}$$

avec $u_j \in \mathbb{R}^d,\ 1 \leq j \leq M,\ v_{i,j} \in \mathbb{R}^m,\ 1 \leq j \leq M,\ 1 \leq i \leq n$ minimisant :

$$w_0 = \ln\left(\int e^{\sum_{j=1}^{M} u_j \cdot g_j(S_{t_j}) + \sum_{j=1}^{M} \sum_{i=1}^{n} v_{i,j} \cdot \gamma_i(S_{t_{j-1}})(S_{t_j} - S_{t_{j-1}})} d\mu \right)$$

Montrons à présent qu'en remplaçant la probabilité a priori μ par ν_n, la propriété de Markov du processus sous-jacent est préservée.
En prenant $1 \leq k \leq M$ et $\varphi \in L^{\infty}(\mathbb{R}^m)$, on calcule l'espérance conditionnelle de $\varphi(S_{t_k})$ sachant $\mathcal{F}_{t_{k-1}}$ sous la probabilité ν_n et, en reprenant la première égalité dans (VI.2.1) et compte tenu de

l'expression (VI.2.3), on a après simplification :

$$
\begin{aligned}
\mathbb{E}^{\nu_n}\left(\varphi(S_{t_k})|\mathcal{F}_{t_{k-1}}\right) &= \frac{\mathbb{E}\left(e^{\sum_{j=1}^{M} u_j\cdot g_j(S_{t_j})+\sum_{j=1}^{M}\sum_{i=1}^{n} v_{i,j}\cdot\gamma_i(S_{t_{j-1}})(S_{t_j}-S_{t_{j-1}})-w_0}\varphi(S_{t_k})\middle|\mathcal{F}_{t_{k-1}}\right)}{\mathbb{E}\left(e^{\sum_{j=1}^{M} u_j\cdot g_j(S_{t_j})+\sum_{j=1}^{M}\sum_{i=1}^{n} v_{i,j}\cdot\gamma_i(S_{t_{j-1}})(S_{t_j}-S_{t_{j-1}})-w_0}\middle|\mathcal{F}_{t_{k-1}}\right)} \\
&= \frac{\mathbb{E}\left(e^{\sum_{j=k}^{M} u_j\cdot g_j(S_{t_j})+\sum_{j=k}^{M}\sum_{i=1}^{n} v_{i,j}\cdot\gamma_i(S_{t_{j-1}})(S_{t_j}-S_{t_{j-1}})}\varphi(S_{t_k})\middle|\mathcal{F}_{t_{k-1}}\right)}{\mathbb{E}\left(e^{\sum_{j=k}^{M} u_j\cdot g_j(S_{t_j})+\sum_{j=k}^{M}\sum_{i=1}^{n} v_{i,j}\cdot\gamma_i(S_{t_{j-1}})(S_{t_j}-S_{t_{j-1}})}\middle|\mathcal{F}_{t_{k-1}}\right)}
\end{aligned}
\tag{VI.2.4}
$$

Au numérateur et au dénominateur du second membre de la dernière égalité, on prend sous la probabilité a priori l'espérance conditionellement à $\mathcal{F}_{t_{k-1}}$ d'une fonction de la trajectoire $(S_{t_j})_{k-1\leq j\leq M}$ et, grâce à la propriété de Markov, c'est une quantité mesurable par rapport à la tribu $\sigma(S_{t_{k-1}})$ engendrée par $S_{t_{k-1}}$ et il en est donc de même de $\mathbb{E}^{\nu_n}\left(\varphi(S_{t_k})|\mathcal{F}_{t_{k-1}}\right)$.
Puisque cela vaut pour toute fonction $\varphi \in L^{\infty}(\mathbb{R}^m)$, on en déduit que le processus discret $(S_{t_j})_{0\leq j\leq M}$ possède la propriété de Markov sous ν_n.
Pour l'étude de la convergence, on considère l'intersection

$$
\tilde{\mathcal{M}}^{\text{mark}}_{\infty,i_0} = \bigcap_{n=1}^{\infty} \mathcal{M}^{\text{mark}}_{n,i_0} \tag{VI.2.5}
$$

qui est faiblement fermé dans $L^1(\mathbb{D}_T,\mu)$ (et même convexe et fortement fermé) en temps qu'intersection d'ensembles ayant cette propriété.
Comme par ailleurs on a l'inclusion

$$
\mathcal{M}^{\text{mark}}_{\infty,i_0} \subset \tilde{\mathcal{M}}^{\text{mark}}_{\infty,i_0} \tag{VI.2.6}
$$

puisqu'elle est vraie en remplaçant le membre de droite par $\mathcal{M}^{\text{mark}}_{n,i_0}$ pour tout n, on en déduit avec l'hypothèse du présent théorème qu'il existe une probabilité $\nu \in \tilde{\mathcal{M}}^{\text{mark}}_{\infty}$ telle que $H(\nu\|\mu) < +\infty$. Avec le Théorème VI.0.8, on peut trouver une probabilité $\tilde{\nu}_\infty \in \tilde{\mathcal{M}}^{\text{mark}}_{\infty}$ telle que l'entropie relative $H(\tilde{\nu}_\infty\|\mu)$ soit minimale, l'unicité étant également claire avec des arguments de convexité. De plus, avec l'égalité ensembliste (VI.2.5) et exactement comme dans la preuve du Théorème VI.1.4, grâce aux résultats de [123], on obtient les résultats de convergence suivants :

$$
H(\nu_n\|\mu) \text{ croît vers } H(\tilde{\nu}_\infty\|\mu) \text{ et } \nu_n \text{ tend vers } \tilde{\nu}_\infty \text{ en variation} \tag{VI.2.7}
$$

La preuve du théorème sera achevée si l'on prouve que $\tilde{\nu}_\infty$ appartient à $\mathcal{M}^{\text{mark}}_{\infty,i_0}$ car l'inclusion (VI.2.6) montre alors que $\tilde{\nu}_\infty$ minimise l'entropie relative sur $\mathcal{M}^{\text{mark}}_{\infty,i_0}$: le problème $(PL^{\text{mark}}_{\infty,i_0})$ admet une solution unique qui vaut $\tilde{\nu}_\infty$ et qui vérifie les résultats de convergence (VI.2.7).
Montrons donc $\tilde{\nu}_\infty \in \mathcal{M}^{\text{mark}}_{\infty,i_0}$.

En effet, d'abord $(S_{t_i})_{0\leq i\leq M}$ est un processus de Markov sous ν_n qui est également une probabilité calibrée et ces propriétés se conservent par passage à la limite car f est bornée et puisque la propriété de Markov se préserve par convergence forte d'après la preuve du Lemme VI.2.1.
Il s'agit donc de montrer que le processus $(S_{t_i})_{0\leq i\leq M}$ est également une martingale sous $\tilde{\nu}_\infty$.
En utilisant l'appartenance de $\tilde{\nu}_\infty$ à $\tilde{\mathcal{M}}^{\text{mark}}_{\infty,i_0}$ défini par (VI.2.5), on a pour $0 \leq j \leq M-1$, et par densité de la suite γ_n dans $\mathcal{C}_c(\mathbb{R}^m)$, pour toute fonction φ de $\mathcal{C}_c(\mathbb{R}^m)$:

$$
\int \varphi(S_{t_j})(S_{t_{j+1}} - S_{t_j})\,d\tilde{\nu}_\infty = \mathbf{0} \tag{VI.2.8}
$$

Si l'on définit la mesure signée $A \in \mathcal{B}(\mathbb{R}^m) \mapsto \alpha(A) = \int 1_A(S_{t_j})(S_{t_{j+1}} - S_{t_j}) d\tilde{\nu}_\infty$, l'espace $\mathcal{C}_c(\mathbb{R}^m)$ est dense dans $L^1(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m), \alpha)$, car la variation $|\alpha|$ est une mesure bornée sur $\mathbb{R}^m$.
En particulier l'égalité (VI.2.8) s'étend aux fonctions $\varphi \in L^\infty(\mathbb{R}^m)$, ce qui se traduit par :

$$\mathbb{E}^{\tilde{\nu}_\infty}(S_{t_{j+1}} | \sigma(S_{t_j})) = S_{t_j}$$

pour $0 \leq j \leq M - 1$.
Avec la propriété de Markov, on réécrit le premier membre de cette dernière égalité comme $\mathbb{E}^{\tilde{\nu}_\infty}(S_{t_{j+1}} | \mathcal{F}_{t_j})$ et l'on en déduit donc que $(S_{t_j})_{0 \leq j \leq M}$ est une martingale sous la probabilité $\tilde{\nu}_\infty$, ce qu'il restait à montrer pour conclure que cette dernière appartient à $\mathcal{M}^{\text{mark}}_{\infty,i_0}$. ∎

Remarque VI.2.3 *Le choix de l'entropie relative n'intervient que parce qu'il préserve la nature markovienne du sous-jacent.*
Il est intéressant de remarquer cependant que les contraintes de calibration ou de martingalité ne doivent faire intervenir chacune la valeur du sous-jacent prise en un instant ou deux instants consécutifs, sinon on perd la propriété de Markov.

Remarque VI.2.4 *Avec le Théorème VI.0.8 et l'hypothèse faite au début du théorème précédent, on sait qu'il existe un élément unique ν_* de l'adhérence faible (ou forte puisqu'il s'agit de celle d'un convexe) $\overline{Conv(\mathcal{M}^{mark}_{\infty,i_0})}$ de l'enveloppe convexe de $\mathcal{M}^{mark}_{\infty,i_0}$ dans $L^1(\mathbb{D}_T, \mu)$, tel que $H(\nu_* \| \mu)$ soit minimale sur $\overline{Conv(\mathcal{M}^{mark}_{\infty,i_0})}$.*
Puisque le membre de droite de (VI.2.6) est un ensemble convexe faiblement fermé et que la preuve précédente montre que la solution $\nu_\infty = \tilde{\nu}_\infty$ du problème (PL^{mark}_{∞,i_0}) y minimise l'entropie relative par rapport à μ, on en déduit l'égalité : $\nu_* = \nu_\infty$.

VI.2.2 Convergence par raffinement de la discrétisation en temps

On va montrer à présent qu'en choisissant un nombre d'échantillons assez élevé (pour i_0 assez grand), on peut espérer obtenir une approximation satisfaisante du problème $(PL^{\text{mark}}_{\infty,i_0})$ par celle ν_{∞,i_0} du problème $(PL^{\text{mark}}_{\infty})$.
Avec l'étude précédente, pour un tel ordre i_0, on sait approcher ν_{∞,i_0} en se restreignant à un nombre fini de contraintes, ce qui autorise alors l'utilisation de la méthode de Monte-Carlo étudiée antérieurement.
Pour obtenir ces résultats, il faut renforcer la nature markovienne du sous-jacent sous la probabilité a priori et l'on énonce :

Théorème VI.2.5 *Supposons qu'il existe une probabilité $\nu \in \mathcal{M}^{mark}_\infty$ équivalente à μ telle que $H(\nu \| \mu) < +\infty$ et supposons de plus que le processus sous-jacent $(S_t)_{0 \leq t \leq T}$ possède la propriété de Markov forte sous μ.*
Alors le problème (PL^{mark}_∞) admet une solution ν_∞ et pour tout i le problème $(PL^{mark}_{\infty,i})$ admet une solution $\nu_{\infty,i}$.
De plus, la suite $H_J(\nu_{\infty,i} \| \mu)$ croît vers $H_J(\nu_\infty \| \mu)$ et $\nu_{\infty,i}$ tend vers ν_∞ en variation.

Preuve : Si l'on note l'inclusion évidente $\mathcal{M}^{\text{mark}}_\infty \subset \mathcal{M}^{\text{mark}}_{\infty,i}$, on satisfait à l'hypothèse faite au début du Théorème VI.2.2 et on en déduit l'existence de $\nu_{\infty,i}$ pour tout $i \in \mathbb{N}^*$.
On pose alors :

$$\mathcal{A}^{\text{mark}}_\infty = \bigcap_{i \in \mathbb{N}^*} \overline{Conv(\mathcal{M}^{\text{mark}}_{\infty,i})} \tag{VI.2.9}$$

où le membre de droite est l'intersection de l'adhérence faible (ou forte) de l'enveloppe convexe des ensembles $\mathcal{M}_{\infty,i}^{\text{mark}}$.
Avec le Théorème VI.0.8 et l'hypothèse faite au début du présent théorème, on sait qu'il existe un unique élément ν_* de l'ensemble faiblement fermé $\mathcal{A}_{\infty}^{\text{mark}}$ tel que la quantité $H(\nu_*\|\mu)$ y soit minimale.
D'autre part, d'après la Remarque VI.2.4, la solution $\nu_{\infty,i}$ du problème $(PL_{\infty,i}^{\text{mark}})$ est telle que la quantité $H(\nu_{\infty,i}\|\mu)$ soit minimale sur $\overline{Conv(\mathcal{M}_{\infty,i}^{\text{mark}})}$.
Alors, avec l'égalité (VI.2.9) on obtient en suivant la preuve du Théorème VI.1.4 grâce aux résultats de [123] :

$$H(\nu_{\infty,i}\|\mu) \text{ croît vers } H(\nu_*\|\mu) \text{ et } \nu_{\infty,i} \text{ tend vers } \nu_* \text{ en variation} \tag{VI.2.10}$$

Alors, en notant l'inclusion $\mathcal{M}_{\infty}^{\text{mark}} \subset \mathcal{A}_{\infty}^{\text{mark}}$, si l'on montre que ν_* appartient à $\mathcal{M}_{\infty}^{\text{mark}}$, on l'identifiera à l'unique solution ν_∞ du problème $(PL_{\infty}^{\text{mark}})$ et les résultats de convergence (VI.2.10) achèveront la preuve.
Il s'agit donc de montrer $\nu_* \in \mathcal{M}_{\infty}^{\text{mark}}$.
D'abord pour tout $i \in \mathbb{N}^*$ le processus $(S_t)_{t\in I_i}$ est une martingale discrète possèdant la propriété de Markov sous $\tilde{\nu}_{\infty,i}$ qui est également une probabilité calibrée.
Comme ces propriétés se conservent par convergence forte, et que l'on dispose de (VI.2.10), alors ν_* est une probabilité calibrée sous laquelle le processus $(S_t)_{t\in I_\infty}$ est une martingale ainsi qu'un processus de Markov, la filtration étant bien sûr $(\mathcal{F}_t)_{t\in I_\infty}$ et en rappelant que l'on a posé $I_\infty = \bigcup_{i\in\mathbb{N}^*} I_i$.
Il suffit donc d'étendre sous ν_* la propriété de Markov au processus $(S_t)_{0\leq t\leq T}$ et de montrer également que c'est une martingale.
Réglons d'abord le problème de la martingalité : pour $s,t \in I_\infty = \bigcup_{i\in\mathbb{N}^*} I_i$, avec s<t, on a d'après ce qui vient d'être dit :

$$\mathbb{E}^{\nu_*}(S_t|\mathcal{F}_s) = S_s \tag{VI.2.11}$$

pour généraliser cette relation à $s < t$ quelconques dans $[0,T]$, on prend $t' \geq t > s' \geq s$ avec $t', s' \in I_\infty$ et l'on fait tendre t' vers t, ce qui donne grâce à (VI.2.11) et compte tenu de la continuité à droite de $(S_u)_{0\leq u\leq T}$ et de sa bornitude :

$$\mathbb{E}^{\nu_*}(S_t|\mathcal{F}_{s'}) = S_{s'} \tag{VI.2.12}$$

En appliquant le Théorème I 3.13 de [76], puisque la filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ répond aux conditions usuelles (en étant en particulier continue à droite), on peut trouver une version continue à droite de la martingale $(\mathbb{E}^{\nu_*}(S_t|\mathcal{F}_u))_{0\leq u\leq T}$ et donc, comme le sous-jacent $(S_u)_{0\leq u\leq T}$ est également continu à droite, on peut faire décroître s' vers s dans (VI.2.12) pour obtenir l'égalité (VI.2.11) : $(S_t)_{0\leq t\leq T}$ est donc une martingale sous ν_*.
On passe ensuite à la propriété de Markov et pour cela, en prenant une fonction $\varphi \in \mathcal{C}_c(\mathbb{R}^m)$, c'est-à-dire continue et à support compact (ou tout simplement continue car le sous-jacent est borné) et pour $s,t \in I_\infty$ avec $s < t$, on a l'égalité :

$$\mathbb{E}^{\nu_*}(\varphi(S_t)|\mathcal{F}_s) = \mathbb{E}^{\nu_*}(\varphi(S_t)|\sigma(S_s)) \tag{VI.2.13}$$

On prend de nouveau $t' \geq t > s' \geq s$ avec $t', s' \in I_\infty$ et l'on fait tendre t' vers t, ce qui donne grâce à la relation précédente et compte tenu de la continuité à droite de $(S_u)_{0\leq u\leq T}$ et de la bornitude de φ :

$$\mathbb{E}^{\nu_*}(\varphi(S_t)|\mathcal{F}_{s'}) = \mathbb{E}^{\nu_*}(\varphi(S_t)|\sigma(S_{s'})) \tag{VI.2.14}$$

En reprenant le Théorème I 3.13 de [76], le membre de gauche possède une version continue à droite $(\mathbb{E}^{\nu_*}(\varphi(S_t)|\mathcal{F}_u))_{0\leq u\leq T}$ et l'on écrit ainsi :

$$\mathbb{E}^{\nu_*}(\varphi(S_t)|\mathcal{F}_s) = \lim_{\substack{s'\in I_\infty \\ s'\downarrow s}} \mathbb{E}^{\nu_*}(\varphi(S_t)|\sigma(S_{s'})) \tag{VI.2.15}$$

On remarque que le membre de droite est mesurable par rapport à la tribu $\bigcap_{s'>s} \sigma(S_u,\ s \le u \le s')$.
Cependant, en utilisant le Théorème II 7.7 de [76] et la propriété de Markov forte du sous-jacent sous μ, on obtient la continuité à droite de la filtration augmentée (à partir de l'instant s) $\sigma(S_u,\ s \le u \le s') \vee \mathcal{N}_\mu$, et, avec ce qui précède, le membre de droite de (VI.2.15) est mesurable par rapport à $\sigma(S_s) \vee \mathcal{N}_\mu$ et, finalement, avec (VI.2.15) on en déduit bien (VI.2.13) c'est-à-dire la propriété de Markov. ∎

VI.3 Approche pénalisée

VI.3.1 Problème de base de la probabilité martingale

Dans les différents problème de calibration précédents, on s'impose la satisfaction exacte des contraintes de prix d'options aussi bien que celles permettant d'approcher la martingalité.
D'une certaine manière, les contraintes de martingalité, excluant les arbitrages, sont plus impératives que celles des prix d'options, qui ne sont en fait connus qu'à l'intérieur d'une fourchette ainsi qu'on l'a évoqué ailleurs.
Il est alors naturel de poser le problème de calibration en pénalisant les contraintes de prix d'options mais en imposant l'exactitude par rapport aux contraintes de martingalité, et de manière plus précise, on cherche à résoudre le problème :

$$(PL_\infty^{J,\chi}) \left\{ \text{ trouver } \nu \in \mathcal{M}_\infty^0 \text{ qui minimise } H_J(\nu\|\mu) + \chi\left(\int f d\nu \right) \right.$$

où l'on a défini l'ensemble

$$\mathcal{M}_\infty^0 = \{\nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \mid (S_t) \text{ est une } \nu\text{-martingale}\}$$

et où la fonction $\chi : \mathbb{R}^d \longrightarrow \mathbb{R}$ est une fonction strictement convexe sci, minimale en $\mathbf{0}$ et possédant les autres propriétés données dans l'étude antérieure de la pénalisation, avec comme choix usuel :

$$\forall u \in \mathbb{R}^d,\ \chi(u) = \frac{1}{2} \sum_{i=1}^{d} \frac{1}{\gamma_i} (u_i)^2$$

où les $\gamma^i,\ 1 \le i \le d$ sont d nombres strictement positifs.
Pour définir ensuite le problème discret, on reprend la suite de fonctions $(\varphi_n)_{n \in \mathbb{N}^*}$ introduites dans le Théorème VI.1.1, et l'on note l'ensemble :

$$\mathcal{M}_n^0 = \left\{ \nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \ \middle|\ \int \varphi_i d\nu = 0,\ 1 \le i \le n \right\}$$

On approche la résolution du problème $(PL_\infty^{J,\chi})$ par celle du problème suivant :

$$(PL_n^{J,\chi}) \left\{ \text{ trouver } \nu \in \mathcal{M}_n^0 \text{ qui minimise } H_J(\nu\|\mu) + \chi\left(\int f d\nu \right) \right.$$

On énonce le résultat suivant, en rappelant encore ici que la fonction vectorielle f et le processus $(S_t)_{0 \le t \le T}$ sont supposés bornés sous μ :

Théorème VI.3.1 *Supposons qu'il existe $\nu \in \mathcal{M}_\infty^0$, $H_J(\nu\|\mu) < +\infty$.*
Alors le problème de calibration $(PL_\infty^{J,\chi})$ admet une solution usuelle unique ν_∞ et de même pour tout $n \in \mathbb{N}^*$ le problème $(PL_n^{J,\chi})$ admet une solution usuelle unique ν_n .
De plus, la suite $H_J(\nu_n\|\mu) + \chi\left(\int f d\nu_n\right)$ croît vers $H_J(\nu_\infty\|\mu) + \chi\left(\int f d\nu_\infty\right)$ et $\frac{d\nu_\infty}{d\mu}$ tend vers $\frac{d\nu_\infty}{d\mu}$ faiblement dans $L^1(\mathbb{D}_T, \mu)$.
Cette dernière convergence est forte si l'on fait le choix $H_J(.\|\mu) = H(.\|\mu)$ l'entropie relative par rapport à μ.

Remarquons que l'hypothèse du théorème est vérifiée dès que l'on suppose que le processus sous-jacent est martingale sous la probabilité a priori μ, ce qui peut être considéré comme un choix pratique.
Ceci constitue donc un avantage par rapport à la calibration exacte pour lequel l'existence d'une probabilité martingale calibrée à distance finie de la mesure a priori est plus contraignant.
Remarquons aussi que le problème discret peut lui même s'approcher par Monte-Carlo, ainsi qu'on l'a étudié antérieurement.
Preuve : Avec le Théorème 2.7 A de [21] ou le Lemme 1 de [123] et si l'on étend la définition de $H_J(.\|\mu)$ à $L^1(\mu)$ en posant $H_J(.\|\mu) : \varphi \in L^1(\mu) \mapsto \int J(\varphi) d\mu$ (et en posant $J(u) = +\infty$ pour $u < 0$) alors cette fonctionnelle a ses ensembles de niveau faiblement compacts, en définissant pour $\alpha \in \mathbb{R}$ l'ensemble de niveau α :

$$B_\alpha = \left\{ \varphi \in L^1(\mu) \parallel \int J(\varphi) d\mu \leq \alpha \right\}$$

Pour que cela soit vrai, la propriété $\lim_{u \to +\infty} \frac{J(u)}{u} = +\infty$ est essentielle, ainsi qu'elle l'était pour énoncer le Théorème VI.0.8. Montrons qu'il en est de même pour la fonctionnelle pénalisée $\varphi \in L^1(\mu) \mapsto H_J(\varphi\mu\|\mu) + \chi\left(\int f\varphi d\mu\right)$.
D'abord, la fonction $H_J(.\|\mu)$ est sci faiblement, ce qui traduit la fermeture faible de ses ensembles de niveau, et l'on en déduit donc que la fonction $\varphi \in L^1(\mu) \mapsto H_J(\varphi\mu\|\mu) + \chi\left(\int f\varphi d\mu\right)$ est sci faiblement car on a la continuité faible du second terme (f étant bornée).
En particulier, l'ensemble de niveau de cette fonctionnelle pénalisée

$$\left\{ \varphi \in L^1(\mu) \parallel H(\varphi\mu\|\mu) + \chi\left(\int f\varphi\mu\right) \leq \alpha \right\}$$

est faiblement fermé donc compact car inclus dans $B_{\alpha - \chi(\mathbf{0})}$.
Alors, avec la compacité faible des ensembles de niveau de la fonctionnelle pénalisée et le fait que l'ensemble $\mathcal{M}_\infty^0$ est faiblement fermé, le problème $(PL_\infty^{J,\chi})$ se ramène à la minimisation d'une fonction strictement convexe sci faiblement sur un ensemble convexe faiblement compact où elle n'est pas identiquement infinie, par hypothèse du présent théorème : la solution ν_∞ existe et est unique pour le problème $(PL_\infty^{J,\chi})$ et on obtient de même l'unique solution ν_n du problème $(PL_n^{J,\chi})$.
Le problème d'existence étant résolu, les résultats de convergence reposent sur l'égalité :

$$\mathcal{M}_\infty^0 = \bigcap_{n \in \mathbb{N}^*} \mathcal{M}_n^0$$

où le membre de droite est la limite de la suite décroissante des ensembles $\mathcal{M}_n^0$.
Alors, avec la Proposition 1.4 de [21] et la compacité faible des ensembles de niveau de la fonctionnelle pénalisée, on a la convergence :

$$H_J(\nu_n\|\mu) + \chi\left(\int f d\nu_n\right) \to_{n\to\infty} H_J(\nu_\infty\|\mu) + \chi\left(\int f d\nu_\infty\right) \tag{VI.3.1}$$

tandis qu'avec la Proposition 1.5 de [21] ainsi que l'unicité de la solution ν_∞, on obtient la convergence faible de ν_n vers celle-ci.
Pour achever la preuve du théorème, il s'agit de montrer que cette dernière convergence est forte dans le cas où la pénalisation s'applique au choix usuel de l'entropie relative pénalisée.
D'après le Théorème 2.7 de [22] ainsi que la Définition 2.4 de cette même référence, l'entropie relative est Kadec, ce qui signifie que :

$$(Kadec)\left\{\begin{array}{l}\text{si } \psi_n \to \psi \text{ faiblement dans } L^1(\mu)\\ \text{et si } H(\psi_n\mu\|\mu) \to H(\psi\mu\|\mu) < +\infty\\ \text{alors } \psi_n \to \psi \text{ fortement}\end{array}\right.$$

Puisque le terme de pénalisation $\chi\left(\int f d\nu_n\right)$ converge, à cause de la convergence faible de la suite ν_n, la relation de convergence (VI.3.1) implique celle de $H(\nu_n\mu\|\mu)$ vers $H(\psi\mu\|\mu) < +\infty$ et on conclut par la propriété de Kadec de l'entropie relative. ∎

VI.3.2 Cadre markovien

Dans cette situation, on se restreint au choix de l'entropie relative. On considère d'abord une discrétisation en temps à un rang i_0 fixé si bien que l'on cherche à résoudre :

$$(PL_{\infty,i_0}^{\text{mark},\chi})\left\{\text{ trouver } \nu \in \mathcal{M}_{\infty,i_0}^{\text{mark},*} \text{ qui minimise } H(\nu\|\mu) + \chi\left(\int f d\nu\right)\right.$$

en posant :

$$\mathcal{M}_{\infty,i_0}^{\text{mark},*} = \left\{\nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \;\middle|\; (S_{t_i})_{0\le i\le M} \text{ est une } \nu\text{-martingale markovienne}\right\}$$

et on l'approche par le problème sujet à un nombre fini de contraintes :

$$(PL_{n,i_0}^{\text{mark},\chi})\left\{\text{ trouver } \nu \in \mathcal{M}_{n,i_0}^{\text{mark},*} \text{ qui minimise } H(\nu\|\mu) + \chi\left(\int f d\nu\right)\right.$$

où l'on pose

$$\mathcal{M}_{n,i_0}^{\text{mark},*} = \left\{\nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \;\middle|\; \int \gamma_i(S_{t_j})(S_{t_{j+1}} - S_{t_j})d\nu = \mathbf{0},\ 1 \le i \le n,\ 0 \le j \le M-1\right\}$$

en reprenant la suite (γ_n) dense dans $\mathcal{C}_c(\mathbb{R}^m)$.
On peut alors répéter le Théorème VI.2.2 pratiquement à l'identique :

Théorème VI.3.2 *Supposons qu'il existe $\nu \in \mathcal{M}_{\infty,i_0}^{mark,0}$ équivalente à μ et telle que $H(\nu\|\mu) < +\infty$ et supposons de plus que le processus sous-jacent $(S_{t_i})_{0\le i\le M}$ soit markovien sous μ.*
Alors le problème $(PL_{\infty,i_0}^{mark,\chi})$ admet une solution unique ν_{∞,i_0}, pour tout n le problème $(PL_{n,i_0}^{mark,\chi})$ admet une solution unique ν_n et de plus le processus $(S_{t_i})_{0\le i\le M}$ conserve la propriété de Markov sous ν_n.
Enfin, la suite $H(\nu_n\|\mu) + \chi\left(\int f d\nu_n\right)$ croît vers $H(\nu_{\infty,i_0}\|\mu) + \chi\left(\int f d\nu_{\infty,i_0}\right)$ et ν_n tend vers ν_{∞,i_0} en variation.

Preuve : La solution ν_n existe et est unique puisque l'on a vu dans la preuve du Théorème VI.3.1 que l'entropie relative pénalisée est une fonction strictement convexe faiblement sci et d'ensembles de niveau faiblement compact ce qui fait qu'elle atteint son minimum en un unique point sur l'ensemble convexe faiblement fermé $\mathcal{M}_{n,i_0}^{\text{mark},*}$, où elle n'est pas identiquement infinie. La propriété de Markov de ν_n est alors acquise puisque son expression est du type (VI.2.3).
Pour l'étude de la convergence, on introduit l'ensemble :

$$\tilde{\mathcal{M}}_{\infty,i_0}^{\text{mark},*} = \bigcap_{n\in\mathbb{N}^*} \mathcal{M}_{n,i_0}^{\text{mark},*} \tag{VI.3.2}$$

qui est convexe et faiblement fermé et où l'entropie relative pénalisée atteint son minimum en un unique élément $\tilde{\nu}_\infty$ pour les mâmes raisons qui ont conclut à l'existence de ν_n.
Alors, exactement comme dans la preuve du Théorème VI.3.1, on prouve :

$$\begin{cases} H(\nu_n\|\mu) + \chi\left(\int f d\nu_n\right) \to H(\tilde{\nu}_\infty\|\mu) + \chi\left(\int f d\tilde{\nu}_\infty\right) < +\infty \\ \text{et } \nu_n \to \tilde{\nu}_\infty \text{ fortement} \end{cases}$$

Puis on termine comme dans la preuve du Théorème VI.2.2 en prouvant que $\tilde{\nu}_\infty$ est l'unique solution ν_{∞,i_0} du problème $(PL_{\infty,i_0}^{\text{mark},\chi})$. ∎

On s'intéresse ensuite à la convergence du problème discrétisé en temps afin d'approcher :

$$(PL_\infty^{\text{mark},\chi}) \left\{ \text{ trouver } \nu \in \mathcal{M}_\infty^{\text{mark},*} \text{ qui minimise } H(\nu\|\mu) + \chi\left(\int f d\nu\right) \right.$$

où l'on a posé

$$\mathcal{M}_\infty^{\text{mark},*} = \left\{ \nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \;\middle|\; (S_t) \text{ est une } \nu\text{-martingale markovienne} \right\}$$

qui est un ensemble non convexe mais fortement fermé dans $L^1(\mu)$ puisque c'est ce qui a été établi dans la preuve du Lemme VI.2.1.
On peut alors énoncer :

Théorème VI.3.3 *Supposons qu'il existe une probabilité $\nu \in \mathcal{M}_\infty^{mark,*}$ équivalente à μ et telle que $H(\nu\|\mu) < +\infty$, et supposons de plus que le processus sous-jacent $(S_t)_{0\le t\le T}$ possède la propriété de Markov forte sous μ.*
Alors le problème $(PL_\infty^{mark,\chi})$ admet une solution unique ν_∞ et pour tout i le problème $(PL_{\infty,i}^{mark})$ admet une solution unique $\nu_{\infty,i}$.
De plus la suite $H(\nu_{\infty,i}\|\mu) + \chi\left(\int f d\nu_{\infty,i}\right)$ croît vers $H(\nu_\infty\|\mu) + \chi\left(\int f d\nu_\infty\right)$ et $\nu_{\infty,i}$ tend vers ν_∞ en variation.

Preuve : L'existence et l'unicité de $\nu_{\infty,i}$ vient du théorème précédent.
On pose ensuite :

$$\mathcal{A}_\infty^{\text{mark},*} = \bigcap_{i\in\mathbb{N}^*} \overline{Conv\left(\mathcal{M}_{\infty,i}^{\text{mark},*}\right)} \tag{VI.3.3}$$

D'une part, pour des raisons semblables à celles de la Remarque VI.2.4, la solution $\nu_{\infty,i}$ minimise l'entropie relative pénalisée sur $\overline{Conv\left(\mathcal{M}_{\infty,i}^{\text{mark},*}\right)}$, d'autre part, en répétant les arguments employés au début de la preuve du Théorème VI.3.2, on en déduit l'existence d'une probabilité unique ν_* minimisant l'entropie pénalisée sur l'ensemble convexe faiblement fermé (car fortement) $\mathcal{A}_\infty^{\text{mark},*}$ et de plus, exactement comme dans la preuve du Théorème VI.3.1, on prouve :

$$\begin{cases} H(\nu_{\infty,i}\|\mu) + \chi\Big(\int f d\nu_{\infty,i}\Big) \to H(\nu_*\|\mu) + \chi\Big(\int f d\nu_*\Big) < +\infty \\ \text{et } \nu_{\infty,i} \to \nu_* \text{ fortement} \end{cases}$$

Alors on achève la preuve comme dans celle du Théorème VI.2.5 en montrant que ν_* est la solution ν_∞ du problème $(PL_\infty^{\text{mark},\chi})$. ∎

VI.4 Caractérisation de la probabilité martingale calibrée minimale

Dans cette dernière partie, on souhaite caractériser davantage la solution limite des problèmes approchés dont la convergence a été établie précédemment.
On se contente ici d'examiner le cas de l'entropie relative en se restreignant au problème non pénalisé, c'est à dire en considérant le problème :

$$(PL_\infty)\left\{ \text{ trouver } \nu \in \mathcal{M}_\infty \text{ qui minimise } H(\nu\|\mu)\right.$$

où l'on a défini l'ensemble

$$\mathcal{M}_\infty = \left\{\nu \in \mathcal{P}(\mathbb{D}_T),\ \nu \ll \mu \mid (S_t)_{0\le t\le T} \text{ est une martingale sous } \nu \text{ et } \int f d\nu = \mathbf{0}\right\}$$

En dehors des aspects de calibration, la caractérisation de la probabilité martingale minimale est donnée par le Théorème 2.3 de [53] et, dans le cas des processus à temps discret, aboutit à une formulation plus naturelle avec le Théorème 2.5 de [53].
La version naturelle en temps continu est par ailleurs explicitée grâce à la Proposition 3.2 de [58] et c'est donc ce résultat que l'on peut adapter sans difficulté :

Théorème VI.4.1 *Supposons qu'il existe une probabilité $\nu_0 \in \mathcal{M}_\infty$ équivalente à μ et telle que $H(\nu_0\|\mu) < +\infty$.*
Une probabilité ν est la solution du problème (PL_∞) si et seulement si :

1. $\nu \in \mathcal{M}_\infty$
2. $$\frac{d\nu}{d\mu} = ce^{\int_{[0,T]} \eta_t \cdot dS_t + \theta\cdot f} \quad \mu\text{-ps} \tag{VI.4.1}$$
   avec $\theta \in \mathbb{R}^d$ et η processus prévisible tel que $\int_{[0,T]} \eta_t \cdot dS_t \in L^1(\nu)$ et $c > 0$
3. $\mathbb{E}^\nu\Big(\int \eta\cdot dS + \theta\cdot f\Big) = 0$

Preuve : On reprend la preuve de la Proposition 3.2 de [58] en commençant par remarquer que l'existence de ν_0 supposée dans le théorème ainsi que des considérations déjà faites montrent que la solution du problème (PL_∞) existe bien.

En admettant dans un premier temps que l'ensemble des conditions du théorème est nécessaire, on en vérifie la suffisance en écrivant d'abord :

$$\begin{aligned} H(\nu\|\mu) &= \ln(c) + \mathbb{E}^{\nu}\left(\int \eta \cdot dS + \theta \cdot f\right) \\ &= \ln(c) \end{aligned} \tag{VI.4.2}$$

où l'on a utilisé 2 puis 3.
D'autre part, si ν_∞ est la solution du problème (PL_∞), on a la suite de relations, puisque l'on a admis que les conditions du théorème sont nécessaires :

$$\begin{aligned} H(\nu_\infty\|\mu) &= H(\nu_\infty\|\nu) + \int \ln\left(\frac{d\nu}{d\mu}\right) d\nu_\infty \\ &\geq \ln(c) + \mathbb{E}^{\nu_\infty}\left(\int \eta \cdot dS + \theta \cdot f\right) \\ &\geq \ln(c) \end{aligned} \tag{VI.4.3}$$

où la première inégalité est due à la positivité de l'entropie relative tandis que la seconde vient de 3.
Avec la condition 3, les relations (VI.4.2) et (VI.4.3) montrent alors l'égalité $\nu = \nu_\infty$ par unicité de la solution du problème de calibration.
Finalement, pour la nécessité, on peut adapter le Théorème 2.2 de [53] et montrer que ν_∞ est équivalente à μ.
Alors, en appliquant le Théorème 3.1 de [36], on sait que ν_∞ est de la forme :

$$\nu_\infty = c e^{\varphi_0} \mu$$

avec $c > 0$ et

$$\varphi_0 \in \overline{K + Vect\{f^i,\ 1 \leq i \leq M\}}^{\nu_\infty}$$

l'adhérence dans $L^1(\nu_\infty)$ de $K + Vect\{f^i,\ 1 \leq i \leq M\}$ en posant :

$$K = Vect\{H(S_t^i - S_s^i) \mid H \in L^\infty(\mathbb{D}_T, \mathcal{F}_s),\ 0 \leq s \leq t \leq T,\ 1 \leq i \leq m\}$$

Par ailleurs, par application du Théorème I.42 de [110], l'ensemble $\overline{K}^{\nu_\infty} + Vect\{f^i,\ 1 \leq i \leq M\}$ est fermé en tant que somme d'un sous-espace fermé et d'un sous-espace de dimension finie et l'on en déduit :

$$\varphi_0 \in \overline{K}^{\nu_\infty} + Vect\{f^i,\ 1 \leq i \leq M\}$$

Or avec la Remarque III.2 de [120], c'est-à-dire en utilisant le Corollaire 2.5.2 de [128] qui s'applique puisque S est une martingale sous ν_∞, l'adhérence $\overline{K}^{\nu_\infty}$ de K dans $L^1(\nu_\infty)$ est incluse dans l'espace des intégrales stochastiques

$$\left\{ \int_{[0,T]} \eta_t \cdot dS_t \in L^1(\nu_\infty) \;\middle|\; \eta \text{ prévisible} \right\}$$

et la preuve est achevée puisque si $\int_{[0,T]} \eta_t \cdot dS_t$ appartient à $\overline{K}^{\nu_\infty}$, alors $\mathbb{E}^{\nu_\infty}\left(\int_{[0,T]} \eta_t \cdot dS_t\right) = \mathbf{0}$
■

Chapitre VII

Prix d'options dans un modèle où la volatilité saute avec le sous-jacent

VII.1 Une modélisation du krach

On se propose d'étudier un modèle d'actif qui intègre la présence de chute du prix sous la forme d'un saut négatif du sous-jacent accompagné d'un saut positif de la volatilité, et qui, en dehors des instants de sauts, suit le modèle de Black-Scholes.
Le processus de prix (S_t), supposé unidimensionnel, suit l'équation différentielle stochastique (EDS) suivante :

$$dS_t = S_{t-}(\mu dt + \sigma_{t-} dW_t + \alpha dN_t) \tag{VII.1.1}$$

où (W_t) est un mouvement brownien et (N_t) est un processus de Poisson indépendant de (W_t) et d'intensité λ et où l'on s'est de plus fixé un saut de sous-jacent $\alpha > -1$ et un rendement μ constants.
Le processus de volatilité subit quant à lui un saut à la hausse, tout en restant inférieur à un certain niveau. Dans une première partie on étudiera le cas d'un retour progressif de la volatilité. Dans ce contexte, les discontinuités aléatoires du processus de volatilité n'interviennent qu'aux instants de sauts du processus de Poisson (N_t) et l'on peut alors facilement adapter les développements de [83] relatifs aux aspects de couverture, ce qui est le premier objectif de ce chapitre. Dans une seconde partie, on envisagera un retour brutal de la volatilité au niveau normal, ajoutant des temps de discontinuité aléatoires, ce qui a essentiellement pour effet de modifier la couverture.

Dans cette partie, pour évaluer le prix des options ainsi que leur couverture, on suppose que la dynamique (VII.1.1) décrit l'évolution du sous-jacent sous une probabilité martingale choisie.
Cela revient à imposer la condition :

$$\mu + \alpha\lambda = r \tag{VII.1.2}$$

où r est le taux sans risque considéré constant.
En effet, l'EDS (VII.1.1) se réécrit pour l'actif actualisé $\hat{S}_t = e^{-rt}S_t$

$$d\hat{S}_t = \hat{S}_{t-}((-r + \mu + \alpha\lambda)dt + \sigma_{t-} dW_t + \alpha d\tilde{N}_t) \tag{VII.1.3}$$

où l'on note le processus de Poisson compensé $\tilde{N}_t = N_t - \lambda t$.
Sous la condition $(VII.1.2)$, cette EDS définit bien une martingale (locale) puisque (W_t) et (M_t)

le sont (ainsi donc que les intégrales stochastiques par rapport à ces deux processus).

Le choix de cette dynamique suppose une procédure de calibration associée, ce qui peut se faire en ajustant les paramètres du modèle de manière à ce que le smile simulé reflète celui que l'on observe.
Une manière de simplifier cette opération est de pouvoir disposer de formules approchées qui sont également établies dans ce chapitre.

VII.2 Cas d'un retour progressif de la volatilité

VII.2.1 Présentation du modèle

On va d'abord préciser le modèle défini par l'EDS (VII.1.1) en construisant le processus de volatilité.
A cet effet, on note $(\tilde{T}_i)_{i\in\mathbb{N}^*}$ la suite des instants de sauts associés au processus de poisson (N_t) :

$$\forall i \geq 1,\ \tilde{T}_i = inf\{t > 0,\ N_t \geq i\}$$

et l'on se donne aussi une constante

$$\tilde{T}_0 \leq 0$$

correspondant au dernier saut dans le passé.
On définit alors σ_t :

$$\left\{ \begin{array}{lcl} \sigma_t^2 & = & \min\left\{\sigma_{max}^2(t), \sigma_{min}^2(t) + \Delta\tilde{\sigma}_t^2\right\} \\ d\Delta\tilde{\sigma}_t^2 & = & \Delta\sigma^2 dN_t + \Delta\tilde{\sigma}_{t-}^2 \psi(t - \tilde{T}_{N_{t-}})dt \end{array} \right. \tag{VII.2.1}$$

où $\sigma_{min}(.)$ est une fonction déterministe représentant le niveau bas de volatilité (régime hors crash) et où $\sigma_{max}(.)$ est également une fonction déterministe désignant le niveau maximal de volatilité toléré.
Afin de borner en définitive la volatilité, on supposera :

$$\underline{\sigma} \leq \sigma_{min}(.), \sigma_{max}(.) \leq \overline{\sigma} \tag{VII.2.2}$$

où $\underline{\sigma}$ et $\overline{\sigma}$ sont des constantes strictement positives.
Dans la pratique, on fixera en fait $\sigma_{max}(.)$ à $\overline{\sigma}$ et l'on règlera essentiellement le niveau $\sigma_{min}(.)$.
L'incrément du carré de la volatilité $\Delta\tilde{\sigma}_t^2$ subit un saut valant $\Delta\sigma^2$ à l'instant $\tilde{T}_i$ tandis que son accroissement relatif instantané est donné par $\psi(t - \tilde{T}_i)$ pour $\tilde{T}_i < t \leq \tilde{T}_{i+1}$.
On choisira $\psi = \frac{\chi'}{\chi}$ où $\chi : \mathbb{R}+ \to \mathbb{R}_+^*$ est une fonction strictement positive et dérivable, valant 1 en 0 et décroissant vers 0 : typiquement, on peut exiger qu'elle atteigne un niveau $\beta \ll 1$ au bout d'un temps caractéristique τ.
L'intérêt de définir la volatilité de cette manière est que l'on obtient un retour approximatif au niveau normal $\sigma_{min}(t)$ si le temps τ s'est écoulé depuis le dernier saut et que l'on préserve le fait que la partie aléatoire $\Delta\tilde{\sigma}_t^2$ ne subit de discontinuités qu'aux instants $\tilde{T}_i$: les sauts interviennent donc suivant le processus de Poisson (N_t).
De façon implicite, on doit se donner un niveau de volatilité initial, ou plutôt de l'incrément de son carré $\Delta\tilde{\sigma}_0^2$ pour résoudre la dernière EDS (VII.2.1).

Avec $\tilde{T}_0$ et S_0, on obtient les conditions initiales du modèle.
Dans la pratique, on prendra :

$$\Delta\tilde{\sigma}_0^2 = 0 \text{ et } \tilde{T}_0 = 0$$

On autorise la volatilité minimale à varier car elle est destinée à jouer en pratique le role de la volatilité implicite du modèle de Black-Scholes.
Enfin, le fait de passer par le carré de la volatilité simplifie les développements calculatoires, ainsi que le suggère la formule de Black-Scholes.
Pour résumer, le modèle étudié est défini par les équations (VII.1.1) et (VII.2.1), que l'on complète par la condition (VII.1.2) pour se placer sous une probabilité risque neutre.
A titre de précision, avec le Corollaire I 3.16 de [68] par exemple, la condition suffisante (VII.1.2) est aussi nécessaire puisque le premier terme du second membre de (VII.1.3) doit définir une martingale (locale) et doit donc être nul puisqu'il est à variation bornée et est continu (donc prévisible).
Terminons cette présentation en remarquant que l'on peut résoudre explicitement l'EDS (VII.1.3) ce qui fait intervenir dans un cas simple la formule classique de l'exponentielle de Doléans-Dade : cela s'écrit (avec le Théorème I 4.61 de [68] par exemple) :

$$S_t = S_0 \exp\left(\mu t - \frac{1}{2}\int_0^t \sigma_{s-}^2 ds + \int_0^t \sigma_{s-} dW_s\right)(1+\alpha)^{N_t} \tag{VII.2.3}$$

VII.2.2 Valorisation d'une option européenne

Dans le modèle précédemment exposé, on souhaite donner un prix p_t à une option délivrant en T le pay-off $\phi(S_T)$, où ϕ est mesurable.
On choisit de définir le prix à la date 0 comme étant la valeur initiale d'une stratégie permettant de couvrir au mieux l'option au sens du critère quadratique.
On se donne un portefeuille autofinançant constitué à chaque instant t d'une quantité H_t^0 d'actif non risqué $S_t^0 = e^{rt}$ et H_t du sous-jacent S_t où H_t^0 et H_t sont supposés prévisibles.
Sa valeur V_t est donc :

$$V_t = H_t^0 S_t^0 + H_t S_t$$

et la condition d'autofinancement s'écrit (ici formellemnent) :

$$dV_t = H_t^0 r e^{rt} dt + H_t dS_t \tag{VII.2.4}$$

Pour que cette équation ait un sens, il suffit de se limiter aux stratégies admissibles selon la définition suivante :

Définition VII.2.1 *Une stratégie admissible est un couple de processus (H_t, H_t^0) prévisibles tels que :*

$$\int_0^T |H_s^0| ds < +\infty \ P\text{-}ps \tag{VII.2.5}$$

et

$$\mathbb{E}\left(\int_0^T H_s^2 S_s^2 ds\right) < +\infty \tag{VII.2.6}$$

et tels que l'EDS (VII.2.4) soit vérifiée.

La valeur actualisée du portefeuille suit donc l'EDS :

$$d\hat{V}_t = H_t d\hat{S}_t$$

ce qui montre que cette valeur actualisée vaut :

$$\hat{V}_t = V_0 + \int_0^t H_s d\hat{S}_s \qquad \text{(VII.2.7)}$$

On commence ainsi par montrer que le problème de la couverture fait essentiellement intervenir l'actif risqué :

Lemme VII.2.2 *Soit (H_t) un processus adapté et continu à gauche tel que :*

$$\mathbb{E}\left(\int_0^T H_s^2 S_s^2 ds\right) < +\infty$$

et soit $V_0 \in \mathbb{R}$.
Il existe un seul processus (H_t^0) tel que le couple (H_t^0, H_t) définisse une stratégie admissible.

Preuve : On a forcément :

$$H_t^0 = V_0 + \int_0^t H_s d\hat{S}_s - H_t\hat{S}_t$$

et il est alors facile de vérifier la condition (VII.2.5).
De plus, H_t^0 est un processus adapté et on montre aisément qu'il est continu à gauche, donc prévisible. ∎

Suivant les développements de la section VII 3.2 de [83], en appliquant l'identité $\mathbb{E}(Z^2) = (\mathbb{E}(Z))^2 + \mathbb{E}(Z - \mathbb{E}(Z))^2$, le risque quadratique

$$R_0^T = \mathbb{E}\left(e^{-rT}(V_T - \phi(S_T))\right)^2$$

se réécrit :

$$R_0^T = \left(V_0 - \mathbb{E}(e^{-rT}\phi(S_T))\right)^2 + \mathbb{E}\left((\hat{V}_T - V_0) - e^{-rT}(\phi(S_T) - \mathbb{E}(\phi(S_T)))\right)^2$$

Seul le premier terme du second membre dépend de la quantité initiale V_0 d'après (VII.2.7).
Cela prouve que la stratégie optimale, minimisant le risque quadratique a pour valeur initiale le prix $p_0 = \mathbb{E}(e^{-rT}\phi(S_T))$.
De manière générale, à un instant t, si l'on cherche à minimiser le risque quadratique

$$R_t^T = \mathbb{E}\left(\left(e^{-rT}(V_T - \phi(S_T))\right)^2 \middle| \mathcal{F}_t\right)$$

on obtient la valeur en t de la stratégie optimale

$$\forall 0 \leq t < T,\ p_t = e^{-r(T-t)}\mathbb{E}(\phi(S_T)|\mathcal{F}_t)$$

que l'on adopte ainsi comme prix en t de l'option.
On va alors exprimer p_t en faisant intervenir la fonction $(t, x, \sigma^2) \in [0, T[\times\mathbb{R}_+^* \times \mathbb{R}_+^* \mapsto \mathcal{H}(t, x, \sigma^2)$ donnant le prix en t de la même option dans le modèle de Black-Scholes avec le taux sans risque r, la volatilité σ et un niveau de sous-jacent égal à x à l'instant t.
En effet, compte tenu de la relation (VII.1.2), l'expression (VII.2.3) de S_t permet d'écrire :

$$S_T = S_t(1+\alpha)^{N_T - N_t}e^{-\alpha\lambda(T-t)}X_{T-t} \qquad \text{(VII.2.8)}$$

où (X_s) est la solution de

$$\begin{cases} X_0 = 1 \\ dX_s = X_t(rds + \sigma_{s-}dW_{t+s}) \end{cases}$$

En raison de l'indépendance du mouvement Brownien (W_t) et du processus de Poisson (N_t) et du fait que la volatilité ne dépend que de ce dernier, on en déduit en conditionnant par rapport à $(N_s)_{s\geq t}$:

$$p_t = \mathbb{E}\left(\mathcal{H}\left(t, S_t(1+\alpha)^{N_T-N_t}e^{-\alpha\lambda(T-t)}, \frac{\int_t^T \sigma_{s-}^2 ds}{T-t}\right)\Big|\mathcal{F}_t\right)$$

Enfin, l'espérance conditionnelle ci-dessus ne dépend de $\mathcal{F}_t$ qu'à travers S_t, $\Delta\tilde{\sigma}_t^2$ et $\tilde{T}_{N_t}$ et cela conduit à écrire :

$$p_t = \mathbb{E}\left(\mathcal{H}\left(t, S_t(1+\alpha)^{N_T-N_t}e^{-\alpha\lambda(T-t)}, \frac{\int_t^T \sigma_{s-}^2 ds}{T-t}\right)\Big|S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t}\right) \tag{VII.2.9}$$

Dans ce modèle, l'état présent est donc résumé par le triplet $(S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})$ qui est ainsi le processus markovien à considérer à la place du sous-jacent seul.
Le choix des conditions initiales tient d'ailleurs compte de ce fait.
Remarquons en particulier que dans ce modèle, on s'attend à un retour éventuel de la volatilité à la normale que dans le cas où n_t est non nul. En pratique, on dispose d'une formule semi-explicite pour $\mathcal{H}$ et le prix p_t, pour un triplet $(S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})$ donné, s'obtient sous la forme d'une série :

$$\begin{aligned} p_t \;&=\; e^{-\lambda(T-t)}\sum_{k\geq 0}\frac{(\lambda(T-t))^k}{k!} \\ &\mathbb{E}\left(\mathcal{H}\left(t, x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \frac{\int_t^T \sigma_k^2(s-, v, \tilde{T}, (\tilde{U}_i^k)_{1\leq i\leq k})ds}{T-t}\right)\right)_{(x,v,\tilde{T})=(S_t,\Delta\tilde{\sigma}_t^2,\tilde{T}_{N_t})} \end{aligned}$$

où pour tout $k \geq 1$ $(\tilde{U}_i^k)_{1\leq i\leq k}$ est une famille obtenue par le réordonnement de k variables indépendantes de loi uniforme sur $[t, T]$: c'est la loi des temps de sauts de (N_s) dans $[t, T]$ conditionellement au nombre de sauts $N_T - N_t$ fixé à k.
On définit également pour tout entier $k \geq 0$ et pour tout s de $[t, T]$:

$$\sigma_k^2(s-, v, \tilde{T}, (\tilde{U}_i^k)_{1\leq i\leq k}) = \min\left\{\sigma_{max}^2(s), \sigma_{min}^2(s) + \tilde{\xi}_{s-}\right\} \tag{VII.2.10}$$

en posant

$$\tilde{\xi}_{s-} = \begin{cases} \chi(s-\tilde{T})v \text{ si } t < s \leq \tilde{U}_1^k \\ \chi(s-\tilde{U}_i^k)\left(\tilde{\xi}_{\tilde{U}_i^k-} + \Delta\sigma^2\right), \; 1 \leq i \leq k \text{ si } \tilde{U}_i^k < s \leq \tilde{U}_{i+1}^k \end{cases} \tag{VII.2.11}$$

avec par convention $\tilde{U}_{k+1}^k = T$.

VII.2.3 Couverture : calcul du delta

Cet aspect est exposé dans [83] dans le cas d'une volatilité constante et l'on peut parfaitement en adapter les développements dans le cas présent où la volatilité est une fonction déterministe du processus de sauts, indépendant du Brownien.
En reprenant le cadre de la section précédente, on se donne un portefeuille autofinançant constitué à chaque instant t d'une quantité H_t^0 d'actif non risqué $S_t^0 = e^{rt}$ et H_t du sous-jacent S_t où H_t^0 et H_t sont supposés prévisibles.
On rappelle que la stratégie optimale, minimisant le risque quadratique

$$R_0^T = \mathbb{E}\left(\left(e^{-rT}(V_T - \phi(S_T)\right)^2\right)$$

a pour valeur initiale le prix $p_0 = \mathbb{E}(e^{-rT}\phi(S_T))$.
Afin de déterminer cette stratégie optimale, on réécrit d'abord le prix p_t sous la forme explicite :

$$p_t = F(t, S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})$$

en posant :

$$F(t,x,v,\tilde{T}) = \\ e^{-\lambda(T-t)} \sum_{k\geq 0} \frac{(\lambda(T-t))^k}{k!} \mathbb{E}\left(\mathcal{H}\left(t, x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \frac{\int_t^T \sigma_k^2(s-, v, \tilde{T}, (\tilde{U}_i^k)_{1\leq i\leq k})ds}{T-t}\right)\right) \tag{VII.2.12}$$

On fera les hypothèses suivantes sur la fonction $\mathcal{H}$ donnant le prix Black-Scholes de l'option :

$$\begin{cases} (t,x,\sigma^2) \mapsto \mathcal{H}(t,x,\sigma^2) \text{ est } \mathcal{C}^{1,2,1} \text{ sur }]0,T[\times\mathbb{R}_+^* \times \mathbb{R}_+^* \\ \left|\frac{\partial\mathcal{H}}{\partial x}\right|(t,x,\sigma^2) + \left|\frac{\partial^2\mathcal{H}}{\partial x^2}\right|(t,x,\sigma^2) \\ \qquad\qquad + \left|\frac{\partial\mathcal{H}}{\partial\sigma^2}\right|(t,x,\sigma^2) \leq A(1+Bx^a) \text{ sur }]0,T[\times\mathbb{R}_+^* \times [\underline{\sigma^2}, \overline{\sigma^2}] \end{cases} \tag{VII.2.13}$$

avec $a, A, B \leq 0$.
Cette hypothèse ne porte en fait que sur la fonction de payoff ϕ. Par exemple, dans le cas d'un Call, cette condition est vérifiée puisque la condition de régularité est classiquement vraie tandis que l'on a

$$\frac{\partial\mathcal{H}}{\partial x}(t,x,\sigma^2) = N(d_1)$$

et

$$\frac{\partial\mathcal{H}}{\partial\sigma^2}(t,x,\sigma^2) = \frac{1}{2\sigma}x\sqrt{T-t}\, n(d_1)$$

où N et n désignent respectivement les fonctions de répartition et de densité de la loi gaussienne centrée réduite tandis que d_1 est une fonction de t, de x et de σ.

Lemme VII.2.3 *Sous la condition (VII.2.13), la fonction* $(t,x,v,\tilde{T}) \mapsto F(t,x,v,\tilde{T})$ *est de classe* $\mathcal{C}^{1,2,1,1}$.

Preuve : Pour la dérivée par rapport à x, par exemple, c'est immédiat avec la condition (VII.2.13) puisque l'on a

$$e^{-\lambda(T-t)} \sum_{k\geq 0} \frac{(\lambda(T-t))^k}{k!} \left(1 + Bx^a(1+\alpha)^{ak} e^{-a\alpha\lambda(T-t)}\right)(1+\alpha)^k e^{-\alpha\lambda(T-t)} < +\infty$$

ce qui permet d'appliquer en particulier le théorème de dérivation sous l'intégrale et d'obtenir :

$$\begin{aligned} \frac{\partial F}{\partial x}(t,x,v,\tilde{T}) &= e^{-\lambda(T-t)} \sum_{k\geq 0} \frac{(\lambda(T-t))^k}{k!}(1+\alpha)^k e^{-\alpha\lambda(T-t)} \\ &\quad \mathbb{E}\left(\frac{\partial\mathcal{H}}{\partial x}\left(t, x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \frac{\int_t^T \sigma_k^2(s-, v, \tilde{T}, (\tilde{U}_i^k)_{1\leq i\leq k})ds}{T-t}\right)\right) \end{aligned} \tag{VII.2.14}$$

Les autres dérivées se traitent de façon identique en notant en particulier que la fonction :

$$(t, v, \tilde{T}) \mapsto \frac{\int_t^T \sigma_k^2(s-, v, \tilde{T}, (\tilde{U}_i^k)_{1\leq i\leq k})ds}{T-t}$$

est de classe $\mathcal{C}^{1,1,1}$. ∎

On énonce :

Lemme VII.2.4 *Si la valeur initiale d'une stratégie admissible (au sens de la définition VII.2.1) est donnée par :*

$$V_0 = \mathbb{E}(e^{-rT}\phi(S_T)) = F(0, S_0, \Delta\tilde{\sigma}_0^2, \tilde{T}_0)$$

alors le risque quadratique a pour expression :

$$\begin{aligned} R_0^T &= \mathbb{E}\Bigg(\int_0^T \left(\frac{\partial F}{\partial x}(s, S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t}) - H_t\right)^2 e^{-2rt} S_t^2 \sigma_{t-}^2 dt \\ &+ \int_0^T \lambda e^{-2rt}\left(F(t, S_t(1+\alpha), \Delta\tilde{\sigma}_t^2 + \Delta\sigma^2, t) - F(t, S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t}) - H_t\alpha S_t\right)^2 dt\Bigg) \end{aligned} \tag{VII.2.15}$$

Preuve : On suit la preuve de la Proposition VII 3.2 de [83].
Elle repose sur la formule d'Ito appliquée au processus $\hat{F}(t, \hat{S}_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})$ où l'on a posé :

$$\hat{F}(t, x, v, z) = e^{-rt}F(t, e^{rt}x, v, z)$$

et où l'on rappelle que $\hat{S}_t = e^{-rt}S_t$ est l'actif actualisé.
Puisque F est de classe $\mathcal{C}^{1,2,1,1}$, que $(\Delta\tilde{\sigma}_t^2)$ est un processus à variation bornée et $\tilde{T}_{N_t}$ est un processus de saut pur on a :

$$\begin{aligned} \hat{F}(t, \hat{S}_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t}) - \hat{F}(0, S_0, \Delta\tilde{\sigma}_0^2, \tilde{T}_0) &= \int_0^t \frac{\partial \hat{F}}{\partial s}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})ds \\ &+ \int_0^t \frac{\partial \hat{F}}{\partial x}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})\sigma_{s-}\hat{S}_{s-}dW_s \\ &-\lambda\alpha \int_0^t \frac{\partial \hat{F}}{\partial x}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})ds \\ &+\frac{1}{2}\int_0^t \frac{\partial^2 \hat{F}}{\partial x^2}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})\sigma_{s-}^2\hat{S}_{s-}^2 ds \\ &+ \int_0^t \frac{\partial \hat{F}}{\partial v}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})\Delta\tilde{\sigma}_{s-}^2\psi(s - \tilde{T}_{N_{s-}})ds \\ &+ \sum_{0\leq s\leq t}\left(\hat{F}(s, \hat{S}_s, \Delta\tilde{\sigma}_s^2, \tilde{T}_{N_s}) - \hat{F}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})\right) \end{aligned}$$

Le dernier terme du second membre regroupe les sauts de $\hat{F}(t, \hat{S}_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})$ et se réécrit :

$$\int_0^t \left(\hat{F}(s, \hat{S}_{s-}(1+\alpha), \Delta\tilde{\sigma}_{s-}^2 + \Delta\sigma^2, s) - \hat{F}(s, \hat{S}_{s-}, \Delta\tilde{\sigma}_{s-}^2, \tilde{T}_{N_{s-}})\right) dN_s$$

En compensant cette intégrale, on voit alors que $\hat{F}(t, \hat{S}_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t}) - \hat{F}(0, S_0, \Delta\tilde{\sigma}_0^2, \tilde{T}_0)$ est une semi-martingale spéciale dont la composante prévisible à variation bornée est nulle puisque cette

semi-martingale est aussi une martingale en tant que prix actualisé de l'option obtenu en prenant l'espérance conditionnelle de son payoff.
Ainsi, on a :

$$\hat{F}(t,\hat{S}_t,\Delta\tilde{\sigma}_t^2,\tilde{T}_{N_t}) - \hat{F}(0,S_0,\Delta\tilde{\sigma}_0^2,\tilde{T}_0) = \int_0^t \frac{\partial\hat{F}}{\partial x}(s,\hat{S}_{s-},\Delta\tilde{\sigma}_{s-}^2,\tilde{T}_{N_{s-}})\sigma_{s-}\hat{S}_{s-}dW_s$$
$$+\int_0^t \left(\hat{F}(s,\hat{S}_{s-}(1+\alpha),\Delta\tilde{\sigma}_{s-}^2+\Delta\sigma^2,s) - \hat{F}(s,\hat{S}_{s-},\Delta\tilde{\sigma}_{s-}^2,\tilde{T}_{N_{s-}})\right) d(N_s-\lambda s)$$

D'autre part la valeur actualisée du portefeuille $\hat{V}_t = e^{-rt}V_t$ s'écrit, compte tenu de (VII.1.3) et (VII.1.2)

$$\begin{aligned}\hat{V}_t &= V_0 + \int_0^t H_s d\hat{S}_s \\ &= V_0 + \int_0^t H_s\hat{S}_{s-}\left(\sigma_{s-}dW_s + \alpha d(N_s-\lambda s)\right)\end{aligned}$$

En tenant compte de la valeur initiale du portefeuille, on obtient :

$$\hat{V}_t - \hat{F}(t,\hat{S}_t,\Delta\tilde{\sigma}_t^2,\tilde{T}_{N_t}) = M_t^{(1)} + M_t^{(2)}$$

où l'on a défini les deux martingales locales

$$\begin{aligned}M_t^{(1)} &= \int_0^t \left(H_s - \frac{\partial\hat{F}}{\partial x}(s,\hat{S}_{s-},\Delta\tilde{\sigma}_{s-}^2,\tilde{T}_{N_{s-}})\right)\sigma_{s-}\hat{S}_{s-}dW_s \\ M_t^{(2)} &= -\int_0^t \left(\hat{F}(s,\hat{S}_{s-}(1+\alpha),\Delta\tilde{\sigma}_{s-}^2+\Delta\sigma^2,s) - \hat{F}(s,\hat{S}_{s-},\Delta\tilde{\sigma}_{s-}^2,\tilde{T}_{N_{s-}}) - \alpha H_s\hat{S}_{s-}\right) d(N_s-\lambda s)\end{aligned}$$

On peut prouver qu'il s'agit en fait de martingales de carré intégrable.
Pour ce faire, on introduit les crochets

$$< M^{(1)}, M^{(1)} >_t = \int_0^t (Y_s^{(1)})^2 ds$$
$$< M^{(2)}, M^{(2)} >_t = \int_0^t (Y_s^{(2)})^2 \lambda ds$$

en posant :

$$\begin{aligned}Y_s^{(1)} &= \left(H_s - \frac{\partial\hat{F}}{\partial x}(s,\hat{S}_s,\Delta\tilde{\sigma}_s^2,\tilde{T}_{N_s})\right)\sigma_s\hat{S}_s \\ Y_s^{(2)} &= -\left(\hat{F}(s,\hat{S}_s(1+\alpha),\Delta\tilde{\sigma}_s^2+\Delta\sigma^2,s) - \hat{F}(s,\hat{S}_s,\Delta\tilde{\sigma}_s^2,\tilde{T}_{N_s}) - \alpha H_s\hat{S}_s\right)\end{aligned}$$

Montrons que :

$$\mathbb{E}\left(\int_0^T (Y_s^{(1)})^2 ds\right) < +\infty \tag{VII.2.16}$$

En effet, avec l'expression (VII.2.14), on a

$$\begin{aligned}\frac{\partial\hat{F}(t,x,v,\tilde{T})}{\partial x} &= e^{-\lambda(T-t)}\sum_{k\geq 0}\frac{(\lambda(T-t))^k}{k!}(1+\alpha)^k e^{-\alpha\lambda(T-t)} \\ &\quad \mathbb{E}\left(\frac{\partial\mathcal{H}}{\partial x}\left(t, e^{rt}x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \frac{\int_t^T \sigma_k^2(s-,v,\tilde{T},(\tilde{U}_i^k)_{1\leq i\leq k})ds}{T-t}\right)\right)\end{aligned} \tag{VII.2.17}$$

Compte tenu de la condition (VII.2.13), on obtient donc la majoration :

$$\left|\frac{\partial \hat{F}(t,x,v,\tilde{T})}{\partial x}\right| \leq$$
$$e^{-\lambda(T-t)}\sum_{k\geq 0}\frac{(\lambda(T-t))^k}{k!}(1+\alpha)^k e^{-\alpha\lambda(T-t)} A\left(1+Be^{art}x^a(1+\alpha)^{ak}e^{-a\alpha\lambda(T-t)}\right) \tag{VII.2.18}$$

D'autre part, en utilisant l'expression (VII.2.3), ainsi qu'un développement en série du type (VII.2.12) on a pour tout $b>0$:

$$\begin{aligned}\mathbb{E}\left(\hat{S}_t^b\right) &\leq S_0^b e^{-\lambda t}\sum_{k\geq 0}\frac{(\lambda t)^k}{k!}e^{-\frac{1}{2}\underline{\sigma}^2 tb}e^{\frac{1}{2}\overline{\sigma}^2 tb^2}(1+\alpha)^{bk} \\ &= S_0^b e^{t\left(\lambda((1+\alpha)^b-1)+\frac{b}{2}(\overline{\sigma}^2 b-\underline{\sigma}^2)\right)}\end{aligned} \tag{VII.2.19}$$

et donc

$$\mathbb{E}\left(\int_0^T \hat{S}_t^b dt\right) < +\infty$$

Alors, avec (VII.2.19), on obtient bien l'intégrabilité souhaitée (VII.2.16).
De manière tout à fait similaire, en utilisant l'inégalité des accroissements finis pour majorer la valeur absolue du terme $\hat{F}(s,\hat{S}_s(1+\alpha),\Delta\tilde{\sigma}_s^2+\Delta\sigma^2,s)-\hat{F}(s,\hat{S}_s,\Delta\tilde{\sigma}_s^2,\tilde{T}_{N_s})$ de $Y_s^{(2)}$, on aboutit à

$$\mathbb{E}\left(\int_0^T (Y_s^{(2)})^2\lambda ds\right) < +\infty \tag{VII.2.20}$$

Alors, classiquement (b/ du Théorème I 4.40 de [68] par exemple), on en déduit que $M^{(1)}$ et $M^{(2)}$ sont deux martingales de carré intégrable.

On a ensuite le développement suivant

$$\left(\hat{V}_t-\hat{F}(t,\hat{S}_t,\Delta\tilde{\sigma}_t^2,\tilde{T}_{N_t})\right)^2 = (M_t^{(1)})^2+(M_t^{(2)})^2+2M_t^{(1)}M_t^{(2)}$$

La définition de la variation quadratique et des calculs classiques donnent :

$$\begin{aligned}(M_t^{(1)})^2 &= 2\int_0^t M_{s-}^{(1)}Y_{s-}^{(1)}dW_s+\int_0^t (Y_{s-}^{(1)})^2 ds \\ (M_t^{(2)})^2 &= 2\int_0^t M_{s-}^{(2)}Y_{s-}^{(2)}d(N_s-\lambda s)+\int_0^t (Y_{s-}^{(2)})^2 dN_s \\ M_t^{(1)}M_t^{(2)} &= \int_0^t M_{s-}^{(2)}Y_{s-}^{(1)}dW_s+\int_0^t M_{s-}^{(1)}Y_{s-}^{(2)}d(N_s-\lambda s)\end{aligned}$$

La condition (VII.2.16) ainsi que le numéro b/ du Théorème I 4.50 de [68] montrent que

$$\int_0^t M_{s-}^{(1)}Y_{s-}^{(1)}dW_s = (M_t^{(1)})^2-[M^{(1)},M^{(1)}]_t$$

est une martingale puisque $M^{(1)}$ est une martingale de carré intégrable .
De même, puisque $M^{(1)}$ et $M^{(2)}$ sont deux martingales de carré intégrables et de crochet nul, le numéro b/ du Théorème I 4.50 de [68] montre que leur produit $M^{(1)}M^{(2)}$ est une martingale

(On pourrait aussi utiliser le Lemme VII 2.3 de [83]).

De la même façon, le résultat de [68] précédemment cité montre que $(M_t^{(2)})^2 - \int_0^t (Y_{s-}^{(2)})^2 dN_s$ est une martingale et d'autre part, on a, par définition du compensateur (voir le i/ du Théorème II 1.8 de [68]) :

$$\mathbb{E}\left(\int_0^t (Y_{s-}^{(2)})^2 dN_s\right) = \mathbb{E}\left(\int_0^t (Y_{s-}^{(2)})^2 \lambda ds\right) < +\infty$$

la dernière inégalité provenant de (VII.2.20).
Cela prouve que l'intégrale compensée $\int_0^t (Y_{s-}^{(2)})^2 d(N-\lambda s)_s$ est de moyenne nulle et, avec un raisonnement similaire, on montre que c'est même une martingale, ce qui achève de traiter le cas de $(M_t^{(2)})^2$ (On aurait également pu utiliser le Lemme VII 2.2 de [83] pour aboutir à ce résultat.)
En définitive, on peut écrire :

$$\left(\hat{V}_t - \hat{F}(t, \hat{S}_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})\right)^2 = M_t + \int_0^t (Y_{s-}^{(1)})^2 ds + \int_0^t (Y_{s-}^{(2)})^2 \lambda ds$$

où (M_t) est une martingale.
On obtient le résultat annoncé en prenant l'espérance. ∎

Ensuite, pour déterminer le delta (H_t) de la couverture optimale, il suffit d'annuler la dérivée de l'expression du risque R_0^T par rapport à (H_t) :

Proposition VII.2.5 *Le delta du portefeuille optimal est donné par :*

$$H_t = \Delta(t, S_{t-}, \Delta\tilde{\sigma}_{t-}^2, \tilde{T}_{N_{t-}})$$

avec

$$\Delta(t, S_{t-}, \Delta\tilde{\sigma}_{t-}^2, \tilde{T}_{N_{t-}}) \equiv$$
$$\frac{1}{1+\frac{\lambda\alpha^2}{\sigma_{t-}^2}}\left(\frac{\partial F}{\partial x}(t, S_{t-}, \Delta\tilde{\sigma}_{t-}^2, \tilde{T}_{N_{t-}}) + \frac{\lambda\alpha^2}{\sigma_{t-}^2}\frac{F(t, S_{t-}(1+\alpha), \Delta\tilde{\sigma}_{t-}^2 + \Delta\sigma^2, t) - F(t, S_{t-}, \Delta\tilde{\sigma}_{t-}^2, \tilde{T}_{N_{t-}})}{\alpha S_{t-}}\right) \tag{VII.2.21}$$

et le risque a alors pour expression :

$$R_0^T =$$
$$\mathbb{E}\left(\int_0^T \frac{\lambda\alpha^2}{1+\frac{\lambda\alpha^2}{\sigma_t^2}} e^{-2rt} S_t^2 \left(\frac{\partial F}{\partial x}(t, S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t}) - \frac{F(t, S_t(1+\alpha), \Delta\tilde{\sigma}_t^2 + \Delta\sigma^2, t) - F(t, S_t, \Delta\tilde{\sigma}_t^2, \tilde{T}_{N_t})}{\alpha S_t}\right)^2 dt\right)$$

Preuve : Le processus H_t ainsi défini est adapté et continu à gauche.
De plus, il est facile de voir que la condition (VII.2.6) est vérifiée en reprenant ce qui a été fait pour établir (VII.2.16) et (VII.2.20).
Le Lemme VII.2.2 montre que l'on a peut définir une unique stratégie admissible de valeur initiale $V_0 = \mathbb{E}(e^{-rT}\phi(S_T)) = F(0, S_0, \Delta\tilde{\sigma}_0^2, \tilde{T}_0)$ et telle que le processus (H_t) détermine la quantité d'actif S_t.
Finalement, avec le Lemme VII.2.15, on voit qu'il s'agit de la stratégie optimale et l'expression du risque résiduel minimal s'en déduit par calcul. ∎

Remarque VII.2.6 *Pour voir le sens de la couverture donnée par le delta modifié (VII.2.21), on peut discuter suivant l'ordre de grandeur du rapport* $\frac{\lambda\alpha^2}{\sigma_{t-}^2}$ *en distinguant deux régimes limite, à* $\Delta\sigma^2$ *fixé :*

1. *Si* $\frac{\lambda\alpha^2}{\sigma_{t-}^2} \ll 1$*, on se couvre essentiellement grâce à la dérivée continue, la contribution discrète devenant négligeable.*
   L'interprétation est claire puisque les sauts deviennent négligeables devant la volatilité avec pour conséquence le fait que la partie Brownienne devient prépondérante et la volatilité est approximativement déterministe : on est donc ramené au cadre de Black-Scholes et le delta est forcément positif dans le cas d'un Call.
2. *Si* $\frac{\lambda\alpha^2}{\sigma_{t-}^2} \gg 1$*, on se trouve dans le cas d'un modèle de sauts pur.*
   Dans ce cas, le delta est encore positif pour un Call.
3. *Il existe une situation intermédiaire intéressante où le delta continu (Black-Scholes) est faible devant le delta discret ("Poissonnien") et où ce dernier est négatif, contrairement au cas limite précédent.*
   En effet, si la contribution Brownienne n'est plus négligeable, le saut de volatilité peut compenser le saut négatif du sous-jacent, impliquant une augmentation du prix : le delta Poissonnien s'écrit alors comme un rapport négatif qui impose que le delta global est également négatif.

VII.3 Cas d'un retour brutal de la volatilité

VII.3.1 Le modèle

Dans cette partie, on envisage le cas où le retour de la volatilité à la normale s'effectue de façon instantanée au bout d'un temps aléatoire.
On reprend donc la dynamique du prix (S_t) définie par (VII.1.1) :

$$dS_t = S_{t-}(\mu dt + \sigma_{t-} dW_t + \alpha dN_t) \tag{VII.3.1}$$

où σ_t est donnée par l'expression suivante :

$$\sigma_t^2 = \min\left\{\sigma_{max}^2(t), \sigma_{min}^2(t) + n_t \Delta\sigma^2\right\} \tag{VII.3.2}$$

où $\sigma_{min}(.)$ est la fonction déterministe représentant le niveau bas de volatilité (régime hors crash) tandis qu'est constants l'incrément du carré de la volatilité $\Delta\sigma^2$ ainsi que la durée τ du saut avant retour à la normale.
Le nombre n_t d'incréments du carré de la volatilité à l'instant t est défini par récurrence : on se donne d'abord un nombre entier n_0 afin de fixer le niveau de volatilité initiale. Avec le dernier temps du sous-jacent $\tilde{T}_0$ ainsi qu'un temps aléatoire τ_0, on peut initier le procédé récursif suivant :

$$n_t = \begin{cases} n_0 1_{\{t-\tilde{T}_0<\tau_0\}} & \text{si } 0 \le t < \tilde{T}_1 \\ \left(n_{\tilde{T}_{N_t}-} + 1\right) 1_{\{t-\tilde{T}_{N_t}<\tau_{N_t}\}} & \text{si } t \ge \tilde{T}_1 \end{cases} \tag{VII.3.3}$$

Après constatation du ième saut le niveau de volatilité retombe ainsi à la normale si un nouveau saut n'a pas été observé au bout d'une durée τ_i.
Les temps de retour de la volatilité à la normale sont modélisés par une suite de variables aléatoires positives $(\tau_i)_{i\in\mathbb{N}}$ supposées indépendantes entre elles ainsi que de (N_t) et (W_t).

Pour travailler dans un cadre markovien, on fera l'hypothèse que pour $i \geq 1$, les variables τ_i suivent la même loi $F(dt)$ et, pour être cohérent, on suppose aussi que τ_0 suit la loi $F_0(dt)$ définie par

$$F_0(dt) = \begin{cases} F_0(dt) = 1_{]-\tilde{T}_0,+\infty[}(t)\frac{F(dt)}{F(]-\tilde{T}_0,+\infty[)} & \text{si } n_0 > 0 \\ \delta_{\tau_0} \text{ avec } \tau_0 \leq -\tilde{T}_0 \text{ et déterministe} & \text{sinon} \end{cases} \tag{VII.3.4}$$

Quand il s'agit de variables exponentielles, on dispose d'un modèle assez proche de celui étudié dans [97]. Par ailleurs, on examinera le cas particulier où les temps de retour sont déterministes. Avec cette modification, la condition de martingalité du sous-jacent (VII.1.2) est valable ainsi que l'expression (VII.2.3) de S_t.

VII.3.2 Valorisation

Afin de donner des prix d'options, on va définir la filtration adoptée. Avec les temps de retour aléatoires, on doit tenir compte d'un nouveau risque. On introduit le processus $\hat{N}$ du nombre de retours à la normale de la volatilité :

$$\hat{N}_t = \sum_{i=0}^{\infty} 1_{\{t \geq \tilde{T}_i + \tau_i > 0 \text{ et } \tilde{T}_{i+1} > \tilde{T}_i + \tau_i\}}$$

C'est également le nombre de retours à 0 du processus (n_t) (abusivement si n_0 est nul).
On se donne alors comme filtration $(\hat{\mathcal{F}}_t)$ la filtration augmentée de celle $(\hat{\mathcal{F}}_t^0)$ engendrée par les processus (W_t), (N_t) et $(\hat{N}_t)$, en rappelant :

$$\hat{\mathcal{F}}_t = \mathcal{N} \vee \hat{\mathcal{F}}_t^0$$

où on a introduit l'ensemble $\mathcal{N}$ des parties négligeables.
Cette filtration vérifie en particulier les conditions usuelles. Ce résultat n'est pas essentiel pour la suite et on ne donne ici que les principaux arguments de la preuve. Celle-ci repose sur la Proposition II 7.7 de [76] par exemple, et il suffirait de montrer que le processus $(W_t, N_t, \tilde{T}_{N_t}, \hat{T}_{\hat{N}_t})$, continu à droite et markovien possède également la propriété de Markov forte. On le vérifierait en adaptant la Proposition VIII 5 de [28] par exemple, relative à la propriété de Feller.
Comme le processus $(S_t, n_t, \tilde{T}_{N_t}))$ est markovien, on peut suivre des considérations faites dans la section précédente. Il est donc possible de développer le prix d'une option européenne selon l'expression suivante :

$$\begin{aligned} p_t \quad = \quad & e^{-\lambda(T-t)} \sum_{k \geq 0} \frac{(\lambda(T-t))^k}{k!} \mathbb{E}\Bigg(\mathcal{H}\Bigg(t, x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \\ & \frac{\int_t^T \sigma_k^2(s-, n, \tilde{T}, (\tilde{U}_i^k)_{1 \leq i \leq k}, (\tilde{\tau}_i^k)_{0 \leq i \leq k}) ds}{T-t} \Bigg)\Bigg)_{(x,n,\tilde{T})=(S_t,n_t,\tilde{T}_{N_t})} \end{aligned} \tag{VII.3.5}$$

où pour tout $k \geq 1$ $(\tilde{U}_i^k)_{1 \leq i \leq k}$ est une famille obtenue par le réordonnement de k variables indépendantes de loi uniforme sur [t,T] et $(\tilde{\tau}_i^k)_{0 \leq i \leq k})$ est une famille de variables indépendantes des variables uniformes précédentes et ayant pour loi celle de $(\tau_i)_{0 \leq i \leq k}$.
On a posé pour $k \geq 0$

$$\sigma_k^2(s-, n, \tilde{T}, (\tilde{U}_i^k), (\tilde{\tau}_i^k)_{0 \leq i \leq k}) = \min\left\{\sigma_{max}^2(s), \sigma_{min}^2(s) + \tilde{\nu}_{s-}\Delta\sigma^2\right\} \tag{VII.3.6}$$

avec

$$\tilde{\nu}_{s-} = \begin{cases} n 1_{\{s-\tilde{T}<\tilde{\tau}_0^k\}} \text{ si } t < s \leq \tilde{U}_1^k \\ \left(\tilde{\nu}_{\tilde{U}_i^k-} + 1\right) 1_{\{s-\tilde{U}_i^k<\tilde{\tau}_i^k\}}, \ 1 \leq i \leq k \text{ si } \tilde{U}_i^k < s \leq \tilde{U}_{i+1}^k \end{cases} \tag{VII.3.7}$$

en posant comme auparavant $\tilde{U}_{k+1}^k = T$.

Remarque VII.3.1 *Avec les hypothèses d'indépendance, toute l'information est résumée par le processus markovien* $(S_t, n_t, \tilde{T}_{N_t})$.
Si l'on suppose de plus que les durées d'attente $(\tau_i)_{i\in\mathbb{N}}$ *suivent des lois exponentielles, l'absence de mémoire fait que le prix ne dépend en fait que du seul processus* (S_t, n_t)*, ce qui est un avantage théorique par rapport au modèle plus simple où l'on suppose les retours déterministes.*

VII.3.3 Un calcul de compensateur

Afin de couvrir le risque, il s'agit d'abord de compenser le processus $\hat{N}$.
Introduisons à cet effet les temps d'arrêt :

$$T_0^* = \tilde{T}_0$$

et

$$\forall i \geq 1, \ T_i^* = \min\{\tilde{T}_i, \tilde{T}_{i-1} + \tau_{i-1}\}$$

Alors :

Lemme VII.3.2 *Le compensateur de* $(\hat{N}_t)$ *est le processus* $(\hat{A}_t)$ *défini par*

$$d\hat{A}_t = \sum_{n\in\mathbb{N}} 1_{]\tilde{T}_n, T_{n+1}^*]}(t) \frac{\delta_{\tilde{T}_n} * F_n(dt)}{F_n([t-\tilde{T}_n, +\infty])}$$

où $F_n(dt)$ *est la loi de* τ_n*, égale à F(dt) pour* $n \geq 1$ *et définie par l'expression (VII.3.4) pour* $n = 0$.

Preuve : Il s'agit essentiellement d'adapter la preuve du Théorème III 1.33 de [68], en commençant par introduire les tribus suivantes, pour tout $n \in \mathbb{N}$:

$$\begin{aligned} \mathcal{G}_n &= \mathcal{N} \vee \sigma(\tilde{T}_i, T_i^*, \ 0 \leq i \leq n) \\ \mathcal{G}_{n+1}^* &= \mathcal{G}_n \vee \sigma(T_{n+1}^*) \end{aligned}$$

et en notant $\mathcal{P}^W$ la tribu prévisible pour la filtration Brownienne.
On énonce le résultat suivant, en en reportant la démonstration après celle du Lemme VII.3.2 :

Lemme VII.3.3 *Un processus H est prévisible pour la filtration* $(\hat{\mathcal{F}}_t)$ *si et seulement si* H_0 *est* $\hat{\mathcal{F}}_0$*-mesurable et si pour chaque* $n \in \mathbb{N}$ *il existe un processus* $H(n)$ *qui est* $\mathcal{P}^W \vee (\mathcal{G}_n \otimes \mathbb{R}_+)$*-mesurable et un processus* $H_*(n+1)$ *qui est* $\mathcal{P}^W V(\hat{\mathcal{G}}_{n+1} \otimes \mathbb{R}_+)$*-mesurable tels que*

$$H = H_0 + \sum_{n\in\mathbb{N}} \left(H(n) 1_{]\tilde{T}_n, T_{n+1}^*]} + H_*(n+1) 1_{]T_{n+1}^*, \tilde{T}_{n+1}]} \right) \tag{VII.3.8}$$

Tout d'abord, le processus $(\hat{A}_t)$ est prévisible en vertu du lemme énoncé ci-dessus.
Considérons à présent un processus prévisible positif H et montrons l'égalité :

$$\mathbb{E}\left(\int H_s d\hat{N}_s \right) = \mathbb{E}\left(\int H_s d\hat{A}_s \right)$$

On a d'une part

$$\int H_s 1_{]T^*_{n+1},\tilde{T}_{n+1}]}(s)d\hat{N}_s = \int H_s 1_{]T^*_{n+1},\tilde{T}_{n+1}]}(s)d\hat{A}_s = 0$$

et il suffit alors de montrer, puisque le processus H est de la forme (VII.3.8) :

$$\mathbb{E}\left(\int H(n)_s 1_{]\tilde{T}_n,T^*_{n+1}]}(s)d\hat{N}_s\right) = \mathbb{E}\left(\int H(n)_s 1_{]\tilde{T}_n,T^*_{n+1}]}(s)d\hat{A}_s\right)$$

On calcule par définition de $\hat{A}_t$, en rappelant que la loi des temps τ_n est $F_n(dt)$ tandis que celle des temps $\tilde{T}_{n+1} - \tilde{T}_n$ est $\lambda e^{-\lambda t}dt$:

$$\begin{aligned}
&\mathbb{E}\left(\int H(n)_s 1_{]\tilde{T}_n,T^*_{n+1}]}(s)d\hat{A}_s\right)\\
&= \mathbb{E}\left(\int H(n)_s 1_{]\tilde{T}_n,T^*_{n+1}]}(s)\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n([s-\tilde{T}_n,+\infty])}\right)\\
&= \mathbb{E}\Bigg(\int 1_{\{u<v\}}F_n(du))\lambda e^{-\lambda v}dv \int H(n)_s 1_{\{\tilde{T}_n<s\leq\tilde{T}_n+u\}}\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n([s-\tilde{T}_n,+\infty])}\Bigg)\\
&\quad +\mathbb{E}\left(\int 1_{\{u\geq v\}}F_n(du)\lambda e^{-\lambda v}dv \int H(n)_s 1_{\{\tilde{T}_n<s\leq\tilde{T}_n+v\}}\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n([s-\tilde{T}_n,+\infty])}\right)
\end{aligned}$$

En appliquant ensuite le Théorème de Fubini entre autres, on a :

$$\begin{aligned}
&\mathbb{E}\left(\int H(n)_s 1_{]\tilde{T}_n,T^*_{n+1}]}(s)d\hat{A}_s\right)\\
&= \mathbb{E}\left(\int H(n)_s 1_{\{\tilde{T}_n<s\}}\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n([s-\tilde{T}_n,+\infty])}\int 1_{\{s-\tilde{T}_n\leq u<v\}}F_n(du)\lambda e^{-\lambda v}dv\right)\\
&\quad +\mathbb{E}\left(\int H(n)_s 1_{\{\tilde{T}_n<s\}}\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n[s-\tilde{T}_n,+\infty])}\int 1_{\{s-\tilde{T}_n\leq v\leq u\}}F_n(du)\lambda e^{-\lambda v}dv\right)\\
&= \mathbb{E}\Bigg(\int H(n)_s 1_{\{\tilde{T}_n<s\}}\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n([s-\tilde{T}_n,+\infty])}\int 1_{\{s-\tilde{T}_n\leq u\}}1_{\{s-\tilde{T}_n\leq v\}}F_n(du)\lambda e^{-\lambda v}dv\Bigg)\\
&= \mathbb{E}\left(\int H(n)_s 1_{\{\tilde{T}_n<s\}}\frac{\delta_{\tilde{T}_n} * F_n(ds)}{F_n([s-\tilde{T}_n,+\infty])}F_n([s-\tilde{T}_n,+\infty])1_{\{s-\tilde{T}_n<v\}}\lambda e^{-\lambda v}dv\right)\\
&= \mathbb{E}\left(\int H(n)_s 1_{\{\tilde{T}_n<s\}}\delta_{\tilde{T}_n} * F_n(ds)1_{\{s-\tilde{T}_n<v\}}\lambda e^{-\lambda v}dv\right)\\
&= \mathbb{E}\left(H(n)_{\tilde{T}_n+\tau_n}1_{\{\tilde{T}_{n+1}>\tilde{T}_n+\tau_n\}}\right)\\
&= \mathbb{E}\left(\int H(n)_s 1_{]\tilde{T}_n,T^*_{n+1}]}(s)d\hat{N}_s\right)
\end{aligned}$$

ce qui établit l'égalité voulue et donc le lemme. ∎

Preuve du Lemme VII.3.3 : On suit la preuve du Lemme III 1.29 de [68] en montrant dans un premier temps que $\hat{\mathcal{F}}_t$ coincide avec la tribu :

$$\begin{aligned}\mathcal{K}_t \;\; = \;\; & \Big\{ B \in \mathcal{P}(\Omega) \,\Big|\, \forall n \in \mathbb{N},\ \exists B_n \in \mathcal{F}_t^W \vee \mathcal{G}_n, B_{n+1}^* \in \mathcal{F}_t^W \vee \mathcal{G}_{n+1}^*, \\ & B \bigcap \{t < T_{n+1}^*\} = B_n \bigcap \{t < T_{n+1}^*\} \text{ et } B \bigcap \{t < \tilde{T}_{n+1}\} = B_{n+1}^* \bigcap \{t < \tilde{T}_{n+1}\} \Big\}\end{aligned} \tag{VII.3.9}$$

D'abord, on a l'inclusion $\mathcal{K}_t \subset \hat{\mathcal{F}}_t$.
En effet, si $B \in \mathcal{K}_t$, on a alors l'égalité :

$$B = \bigcup_{n \in \mathbb{N}} \left[B_n \bigcap \{\tilde{T}_n \leq t < T_{n+1}^*\} \right] \bigcup \left[B_{n+1}^* \bigcap \{T_{n+1}^* \leq t < \tilde{T}_{n+1}\} \right]$$

où les ensembles B_n et B_{n+1}^* sont associée à B comme dans (VII.3.9).
Puisque l'on a $\forall 1 \leq i \leq n,\ \tilde{T}_i \leq \tilde{T}_n$ et $T_i^* \leq \tilde{T}_n$, alors, en utilisant la définition de $\mathcal{G}_n$, on a :

$$\mathcal{G}_n \subset \hat{\mathcal{F}}_{\tilde{T}_n} \tag{VII.3.10}$$

On en déduit alors facilement :

$$B_n \bigcap \{\tilde{T}_n \leq t\} \in \hat{\mathcal{F}}_t$$

De même, on note :

$$\mathcal{G}_n^* \subset \hat{\mathcal{F}}_{T_{n+1}^*} \tag{VII.3.11}$$

et l'on obtient

$$B_{n+1}^* \bigcap \{T_{n+1}^* \leq t\} \in \hat{\mathcal{F}}_t$$

L'inclusion cherchée en résulte.
Montrons l'inclusion inverse :

$$\hat{\mathcal{F}}_t \subset \mathcal{K}_t$$

Soit en effet $s, u \leq t$, $p, q \in \mathbb{N}$, $A \in \mathcal{F}_t^W$ et prenons $B = A \bigcap \{N_s = p,\ \hat{N}_u = q\}$.
Posons pour tout n :

$$B_n = A \bigcap \left\{ \sum_{i=1}^{n} 1_{[0,s]}(\tilde{T}_i) = p \text{ et } \sum_{i=1}^{n} 1_{[0,u]}(T_i^*) 1_{\{T_i^* \neq \tilde{T}_i\}} = q \right\}$$

et

$$B_{n+1}^* = A \bigcap \left\{ \sum_{i=1}^{n} 1_{[0,s]}(\tilde{T}_i) = p \text{ et } \sum_{i=1}^{n+1} 1_{[0,u]}(T_i^*) 1_{\{T_i^* \neq \tilde{T}_i\}} = q \right\}$$

Alors, on a

$$B \bigcap \{t < T_{n+1}^*\} = B_n \bigcap \{t < T_{n+1}^*\}$$

et

$$B \bigcap \{t < \tilde{T}_{n+1}\} = B_{n+1}^* \bigcap \{t < \tilde{T}_{n+1}\}$$

ce qui prouve que B appartient à $\mathcal{K}_t$.
Comme $\mathcal{N} \subset \mathcal{K}_t$, on obtient l'inclusion souhaitée.
On a ainsi l'égalité

$$\hat{\mathcal{F}}_t = \mathcal{K}_t \tag{VII.3.12}$$

Passons maintenant à la preuve de l'énoncé du lemme et prouvons le caractère suffisant de la condition en posant pour $n \in \mathbb{N}$:

$$H(n) = Y(n) Z(n)$$

où $Y(n)$ est un processus $\mathcal{F}^W$-prévisible, $Z(n)$ est une variable $\mathcal{G}_n$-mesurable.
D'après l'inclusion (VII.3.10), $Z(n)$ est $\hat{\mathcal{F}}_{\tilde{T}_n}$-mesurable et donc le processus $Z(n)1_{]\tilde{T}_n,T^*_{n+1}]}$ est prévisible (Proposition I 2.5 de [68]).
Puisque $Y(n)$ est prévisible, on en déduit donc qu'il en est de même de $H(n)1_{]\tilde{T}_n,T^*_{n+1}]}$.
On montre de manière identique la prévisibilité du processus

$$H_*(n+1)1_{]T^*_{n+1},\tilde{T}_n]}$$

avec

$$H_*(n+1) = Y_*(n+1)Z_*(n+1)$$

où $Y_*(n)$ est un processus $\mathcal{F}^W$-prévisible, $Z_*(n+1)$ est une variable $\mathcal{G}^*_{n+1}$-mesurable.
En utilisant un argument de linéarité et de classe monotone, on en déduit la prévisibilité de tout processus de la forme énoncée dans le lemme.
Pour la réciproque, par un argument de classe monotone, il suffit d'examiner le cas évident où $H = 1_{B\times\{0\}}$ avec $B \in \hat{\mathcal{F}}_0$, auquel cas c'est évident, ainsi que $H = 1_{B\times]t,s]}$ avec $B \in \hat{\mathcal{F}}_t$.
Dans ce dernier cas, avec l'égalité (VII.3.12) et la définition (VII.3.9), on peut associer à B les ensembles B_n et B^*_{n+1} pour $n \in \mathbb{N}$ et il suffit de poser $H(n) = 1_{B_n\times]t,s]}$ et $H_*(n+1) = 1_{B^*_{n+1}\times]t,s]}$.
∎

Dans le cas particulier d'un retour déterministe on obtient le résultat suivant :

Lemme VII.3.4 *Si les temps de retour τ_n sont déterministes, égaux à une constante τ , alors le processus $\hat{N}_t$ est un processus prévisible.*

Preuve : On peut déduire ce résultat du Lemme VII.3.2, en montrant que $\hat{N}_t$ est égal à son compensateur $\hat{A}_t$.
On préfère cependant une preuve directe en réécrivant le nombre de retours sous la forme :

$$\hat{N}_t = 1_{\{\tilde{T}_1 > \tilde{T}_0+\tau>0\}}1_{[\tilde{T}_0+\tau,+\infty[}(t) + \sum_{i=1}^{\infty} 1_{\{\tilde{T}_{i+1} > \tilde{T}_i+\tau\}}1_{[\tilde{T}_i+\tau,+\infty[}(t)$$

Or la loi des temps $\tilde{T}_n - \tilde{T}_{n-1}$ entre les instants de krach est à densité et donc ne charge pas $\{\tau\}$, si bien que le processus $\hat{N}_t$ est indistinguable du processus

$$\hat{N}^*_t = 1_{\{\tilde{T}_1 \geq \tilde{T}_0+\tau>0\}}1_{[\tilde{T}_0+\tau,+\infty[}(t) + \sum_{i=1}^{\infty} 1_{\{\tilde{T}_{i+1} \geq \tilde{T}_i+\tau\}}1_{[\tilde{T}_i+\tau,+\infty[}(t)$$

Pour i fixé, $\tilde{T}_i+\tau$ est un temps prévisible (I 2.8 de [68] par exemple) et $1_{\{\tilde{T}_{i+1}\geq\tilde{T}_i+\tau\}}$ est $\mathcal{F}_{(\tilde{T}_i+\tau)-}$-mesurable puisque :

$$\{\tilde{T}_{i+1} \geq \tilde{T}_i + \tau\}^c = \bigcup_{q\in\mathbb{Q}}\{\tilde{T}_{i+1} \leq q\}\bigcap\{q < \tilde{T}_i + \tau\}$$

Avec la Proposition I 2.12 de [68], on en déduit que $1_{\{\tilde{T}_{i+1}\geq\tilde{T}_i+\tau\}}1_{[\tilde{T}_i+\tau,+\infty[}$ est prévisible, ce qui prouve le résultat. ∎

VII.3.4 Couverture

On rappelle que les instants de krach sont ceux d'un processus de Poisson homogène et l'on suppose de plus que les lois des temps τ_n pour $n \geq 1$ sont à densité $F(dt) = \phi(t)dt$ telle que :

$$\left\{ \begin{array}{l} (\hat{A}_t)_{0\leq t\leq T} \text{ est absolument continue} \\ \frac{d\hat{A}_t}{dt} \leq \hat{\lambda}, \quad 0 \leq t \leq T \end{array} \right. \tag{VII.3.13}$$

où $\hat{\lambda}$ est une constante positive.
Par exemple, cette condition est vérifiée dès que les temps de retour suivent une loi exponentielle.
De manière plus générale, pour que la condition (VII.3.13) soit satisfaite, il suffit d'imposer pour la loi $F_n(dt)$ des temps de retour τ_n, $n \geq 0$ en posant $F_n = F$ pour $n \geq 1$:

$$\left\{ \begin{array}{l} F_n \text{ est à densité par rapport à la mesure de Lebesgue} \\ \frac{1}{F_n([t,+\infty])}\frac{dF_n}{dt} \leq \hat{\lambda}, \quad 0 \leq t \leq T \end{array} \right. \tag{VII.3.14}$$

La résolution est immédiate et, compte tenu de l'expression (VII.3.4) de F_0, on est conduit au choix :

$$\left\{ \begin{array}{l} F(dt) = \phi(t)e^{-\int_0^t \phi(s)ds} \\ \text{où } \phi \text{ mesurable telle que } 0 \leq \phi(t) \leq \overline{\lambda}, \quad \forall 0 \leq t \leq T \text{ avec } \int_0^{+\infty} \phi(t)dt = +\infty \end{array} \right.$$

Pour le calcul de la couverture, l'expression (VII.3.5) permet d'écrire le prix de l'option sous la forme

$$p_t = G(t, S_t, n_t, \tilde{T}_{N_t})$$

en supposant vérifiée la condition (VII.2.13).
Comme pour le Lemme VII.2.3, il est alors loisible de prouver que la fonction G est asez régulière (de classe $\mathcal{C}^{1,2,1,1}$ par exemple).
Le Lemme fondamental VII.2.4 devient

Lemme VII.3.5 *Si la valeur initiale de la stratégie est donnée par :*

$$V_0 = \mathbb{E}(e^{-rT}\phi(S_T)) = G(0, S_0, n_0, \tilde{T}_0)$$

alors le risque quadratique a pour expression :

$$\begin{aligned} R_0^T &= \mathbb{E}\Big(\int_0^T \left(\frac{\partial G}{\partial x}(t, S_t, n_t, \tilde{T}_{N_t}) - H_t\right)^2 e^{-2rt} S_t^2 \sigma_{t-}^2 dt \\ &\quad + \int_0^T \lambda e^{-2rt} \Big(G(t, S_t(1+\alpha), n_t + 1, t) - G(t, S_t, n_t, \tilde{T}_{N_t}) - H_t \alpha S_t\Big)^2 dt \\ &\quad + \int_0^T e^{-2rt} \Big(G(t, \hat{S}_t, 0, \tilde{T}_{N_t}) - G(t, \hat{S}_t, n_t, \tilde{T}_{N_t})\Big)^2 d\hat{A}_t \Big) \end{aligned} \tag{VII.3.15}$$

Preuve : Suivant les lignes de la preuve du Lemme VII.2.15, on introduit la fonction

$$\hat{G}(t, x, n, z) = e^{-rt} G(t, e^{rt}x, n, z)$$

et l'on applique la formule d'Itô

$$\begin{aligned}
\hat{G}(t,\hat{S}_t,n_t,\tilde{T}_{N_t}) - \hat{G}(0,S_0,n_0,\tilde{T}_0) &= \int_0^t \frac{\partial \hat{G}}{\partial s}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})ds \\
&+ \int_0^t \frac{\partial \hat{G}}{\partial x}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\sigma_{s-}\hat{S}_{s-}dW_s \\
&-\lambda\alpha \int_0^t \frac{\partial \hat{G}}{\partial x}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})ds \\
&+\frac{1}{2}\int_0^t \frac{\partial^2 \hat{G}}{\partial x^2}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\sigma_{s-}^2\hat{S}_{s-}^2 ds \\
&+ \sum_{0\leq s\leq t}\left(\hat{G}(s,\hat{S}_s,n_s,\tilde{T}_{N_s}) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right)
\end{aligned}$$

Le dernier terme se réécrit :

$$\begin{aligned}
&\textstyle\int_0^t \left(\hat{G}(s,\hat{S}_{s-}(1+\alpha),n_{s-}+1,s) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right) dN_s \\
&+ \textstyle\int_0^t \left(\hat{G}(s,\hat{S}_{s-},0,\tilde{T}_{N_{s-}}) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right) d\hat{N}_s
\end{aligned}$$

en tenant compte en particulier du fait que les instants de saut des processus (N_t) et $(\hat{N}_t)$ forment deux ensembles disjoints (presque sûrement).
En compensant ces deux intégrales, et en tenant compte en particulier de l'expression du compensateur $\hat{A}_t$ de $\hat{N}_t$ donnée par le Lemme VII.3.2, on obtient :

$$\begin{aligned}
&\hat{G}(t,\hat{S}_t,n_t,\tilde{T}_{N_t}) - \hat{G}(0,S_0,n_0,\tilde{T}_0) = \textstyle\int_0^t \frac{\partial \hat{G}}{\partial x}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\sigma_{s-}\hat{S}_{s-}dW_s \\
&+ \textstyle\int_0^t \left(\hat{G}(s,\hat{S}_{s-}(1+\alpha),n_{s-}+1,s) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right) d(N_s-\lambda s) \\
&+ \textstyle\int_0^t \left(\hat{G}(s,\hat{S}_{s-},0,\tilde{T}_{N_{s-}}) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right) d(\hat{N}_s-\hat{A}_s)
\end{aligned}$$

Ensuite, en considérant la valeur actualisée du portefeuille $\hat{V}_t$ et en continuant comme dans la preuve du Lemme VII.2.15, on écrit :

$$\hat{V}_t - \hat{G}(t,\hat{S}_t,n_t,\tilde{T}_{N_t}) = M_t^{(1)} + M_t^{(2)} + M_t^{(3)}$$

où l'on a défini les trois martingales locales

$$\begin{aligned}
M_t^{(1)} &= \int_0^t \left(H_s - \frac{\partial \hat{G}}{\partial x}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right)\sigma_{s-}\hat{S}_{s-}dW_s \\
M_t^{(2)} &= -\int_0^t \left(\hat{G}(s,\hat{S}_{s-}(1+\alpha),n_{s-}+1,s) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}}) - \alpha H_s\hat{S}_{s-}\right) d(N_s-\lambda s) \\
M_t^{(3)} &= -\int_0^t \left(\hat{G}(s,\hat{S}_{s-},0,\tilde{T}_{N_{s-}}) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right) d(\hat{N}_s-\hat{A}_s)
\end{aligned}$$

Il s'agit en fait de martingales de carré intégrable : en effet, pour les deux premières, on reprend à l'identique ce qui a été fait dans la preuve du Lemme VII.2.15 tandis que pour la troisième, il s'agit encore de prouver (b/ du Théorème I 4.40 de [68]) que l'on a :

$$< M^{(1)},M^{(1)} >_t = \int_0^t \left(\hat{G}(s,\hat{S}_{s-},0,\tilde{T}_{N_{s-}}) - \hat{G}(s,\hat{S}_{s-},n_{s-},\tilde{T}_{N_{s-}})\right)^2 d < \hat{N}-\hat{A},\hat{N}-\hat{A} >_t < +\infty \tag{VII.3.16}$$

Pour calculer le crochet $< \hat{N} - \hat{A}, \hat{N} - \hat{A} >_t$, on utilise le fait qu'il s'agit du compensateur de la variation quadratique $[\hat{N} - \hat{A}, \hat{N} - \hat{A}]_t$ (b/ de la Proposition I 4.50 de [68]).
En utilisant le fait que $\hat{N}_t$ et $\hat{A}_t$ sont à variation finie et que de plus $\hat{A}_t$ est supposé continu, on a (avec le a/ de La Proposition I 4.49 de [68]) :

$$\begin{aligned}
[\hat{N} - \hat{A}, \hat{N} - \hat{A}]_t &= \sum_{0 \leq s \leq t} \left(\Delta \hat{N}_s \right)^2 \\
&= \hat{N}_t
\end{aligned}$$

On en déduit donc

$$< \hat{N} - \hat{A}, \hat{N} - \hat{A} >_t \quad = \quad \hat{A}_t$$

En utilisant l'hypothèse

$$\frac{d\hat{A}_t}{dt} \leq \hat{\lambda}$$

on peut adapter ce qui a été fait pour les deux premières martingales et obtenir la condition d'intégrabilité (VII.3.16) souhaitée, ce qui montre que $M^{(3)}$ est également une martingale.
On a ensuite le développement suivant

$$\left(\hat{V}_t - \hat{G}(t, \hat{S}_t, n_t, \tilde{T}_{N_t}) \right)^2 = \sum_{1 \leq i \leq 3} (M_t^{(i)})^2 + +2 \sum_{1 \leq i < j \leq 3} M_t^{(i)} M_t^{(j)}$$

Une application de la formule d'Ito et des calculs classiques de variations quadratiques, similaires à celui que l'on vient d'établir, donnent, en tenant compte en particulier que $\hat{A}_t$ est absolument continu et que N_t et $\hat{N}_t$ n'ont pas de saut en commun :

$$\begin{aligned}
(M_t^{(1)})^2 &= \int_0^t M_{s-}^{(1)} Y_{s-}^{(1)} dW_s + \int_0^t (Y_{s-}^{(1)})^2 ds \\
(M_t^{(2)})^2 &= \int_0^t M_{s-}^{(2)} Y_{s-}^{(2)} d(N_s - \lambda s) + \int_0^t (Y_{s-}^{(2)})^2 dN_s \\
(M_t^{(3)})^2 &= \int_0^t M_{s-}^{(3)} Y_{s-}^{(3)} d(\hat{N}_s - \hat{A}_s) + \int_0^t (Y_{s-}^{(3)})^2 d\hat{N}_s \\
M_t^{(1)} M_t^{(2)} &= \int_0^t M_{s-}^{(2)} Y_{s-}^{(1)} dW_s + \int_0^t M_{s-}^{(1)} Y_{s-}^{(2)} d(N_s - \lambda s) \\
M_t^{(1)} M_t^{(3)} &= \int_0^t M_{s-}^{(3)} Y_{s-}^{(1)} dW_s + \int_0^t M_{s-}^{(1)} Y_{s-}^{(3)} d(\hat{N}_s - \hat{N}_s) \\
M_t^{(2)} M_t^{(3)} &= \int_0^t M_{s-}^{(3)} Y_{s-}^{(2)} d(N_s - \lambda s) + \int_0^t M_{s-}^{(2)} Y_{s-}^{(3)} d(\hat{N}_s - \hat{A}_s)
\end{aligned}$$

où l'on a posé :

$$\begin{aligned}
Y_s^{(1)} &= \left(H_s - \frac{\partial \hat{G}}{\partial x}(s, \hat{S}_s, n_s, \tilde{T}_{N_s}) \right) \sigma_s \hat{S}_s \\
Y_s^{(2)} &= -\left(\hat{G}(s, \hat{S}_s(1+\alpha), n_{s-} + 1, s) - \hat{G}(s, \hat{S}_s, n_{s-}, \tilde{T}_{N_s}) - \alpha H_s \hat{S}_s \right) \\
Y_s^{(3)} &= -\left(\hat{G}(s, \hat{S}_{s-}, 0, \tilde{T}_{N_{s-}}) - \hat{G}(s, \hat{S}_{s-}, n_{s-}, \tilde{T}_{N_{s-}}) \right)
\end{aligned}$$

Encore une fois, avec des développements similaires à ceux de la preuve du Lemme VII.2.15, reposant sur le b/ de la Proposition I 4.50 de [68] et le fait que les $M^{(i)}$ sont des martingales de

carré intégrable, on montre que les produits croisés sont des martingales (à cause de la nullité de la covariation quadratique) et sont en particulier intégrables, et qu'il en est de même des deux premiers carrés compensés $(M_t^{(1)})^2 - \int_0^t (Y_{s-}^{(1)})^2 ds$ et $(M_t^{(2)})^2 - \int_0^t (Y_{s-}^{(2)})^2 \lambda ds$.
Pour traiter le troisième carré $(M_t^{(3)})^2$, on reprend ce qui a été détaillé dans la preuve du Lemme VII.2.15 pour le deuxième carré : tout d'abord, avec le résultat de [68] précédemment cité, on obtient le fait que $(M_t^{(3)})^2 - \int_0^t (Y_{s-}^{(3)})^2 d\hat{N}_s$ est une martingale.
D'autre part, on a, par définition du compensateur :

$$\begin{aligned} \mathbb{E}\left(\int_0^t (Y_{s-}^{(3)})^2 d\hat{N}_s\right) &= \mathbb{E}\left(\int_0^t (Y_{s-}^{(3)})^2 d\hat{A}_s\right) \\ &\leq \mathbb{E}\left(\int_0^t (Y_{s-}^{(3)})^2 \hat{\lambda} ds\right) < +\infty \end{aligned}$$

la dernière inégalité n'étant autre chose que (VII.3.16).
Cela prouve que l'intégrale compensée $\int_0^t (Y_{s-}^{(3)})^2 d(\hat{N} - \hat{A})_s$ est de moyenne nulle et, avec un raisonnement similaire, on montre que c'est même une martingale, ce qui achève de traiter le cas de $(M_t^{(3)})^2$.

Finalement, on obtient :

$$\left(\hat{V}_t - \hat{G}(t, \hat{S}_t, n_t, \tilde{T}_{N_t})\right)^2 = M_t + \int_0^t (Y_{s-}^{(1)})^2 ds + \int_0^t (Y_{s-}^{(2)})^2 \lambda ds + \int_0^t (Y_{s-}^{(3)})^2 d\hat{A}_s$$

où M_t est une martingale et cela donne le résultat annoncé en prenant l'espérance. ∎

L'expression du risque quadratique comporte un terme supplémentaire indépendant du portefeuille, ce qui montre que le delta de la couverture reste inchangé par rapport au cas déterministe. On énonce :

Proposition VII.3.6 *Le delta du portefeuille optimal est donné par :*

$$H_t = \hat{\Delta}(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}})$$

avec

$$\begin{aligned} \hat{\Delta}(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}}) \equiv \; & \frac{1}{1 + \frac{\lambda\alpha^2}{\sigma_{t-}^2}} \Bigg(\frac{\partial G}{\partial x}(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}}) \\ & + \frac{\lambda\alpha^2}{\sigma_{t-}^2} \frac{G(t, S_{t-}(1+\alpha), n_{t-}+1, t) - G(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}})}{\alpha S_{t-}} \Bigg) \end{aligned} \tag{VII.3.17}$$

et le risque a alors pour expression :

$$\begin{aligned} R_0^T = \; & \mathbb{E}\Bigg(\int_0^T \frac{\lambda\alpha^2}{1 + \frac{\lambda\alpha^2}{\sigma_{t-}^2}} e^{-2rt} S_t^2 \left(\frac{\partial G}{\partial x}(t, S_t, n_t, \tilde{T}_{N_t}) - \frac{G(t, S_t(1+\alpha), n_t+1, t) - G(t, S_t, n_t, \tilde{T}_{N_t})}{\alpha S_t} \right)^2 dt \\ & + \int_0^T e^{-2rt} \left(G(t, \hat{S}_t, 0, \tilde{T}_{N_t}) - G(t, \hat{S}_t, n_t, \tilde{T}_{N_t}) \right)^2 d\hat{A}_t \Bigg) \end{aligned}$$

Preuve : C'est immédiat en suivant la preuve de la Proposition VII.2.5. ∎

Il est important de noter que le cas où les durées de retour τ_i sont déterministes n'entre pas strictement dans ce qui précède puisqu'alors le compensateur $\hat{A}_t$ n'est pas absolument continu. Cependant, en reprenant la preuve du Lemme VII.3.15, la martingale locale $M^{(3)}$ est en fait nulle du fait de l'égalité de $\hat{N}_t$ et de son compensateur $\hat{A}_t$.
On est donc ramené au Lemme VII.2.15 de la section précédente et l'on peut réécrire le Lemme VII.3.7 :

Proposition VII.3.7 *Dans le cas où les temps τ_i sont égaux à une constante $\tau > 0$, le delta du portefeuille optimal est donné par :*

$$H_t = \hat{\Delta}(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}})$$

avec

$$\begin{aligned}\hat{\Delta}(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}}) \quad &\equiv \quad \frac{1}{1+\frac{\lambda\alpha^2}{\sigma_{t-}^2}}\Bigg(\frac{\partial G}{\partial x}(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}}) \\ &+ \frac{\lambda\alpha^2}{\sigma_{t-}^2}\frac{G(t, S_{t-}(1+\alpha), n_{t-}+1, t) - G(t, S_{t-}, n_{t-}, \tilde{T}_{N_{t-}})}{\alpha S_{t-}}\Bigg)\end{aligned} \tag{VII.3.18}$$

et le risque a alors pour expression :

$$R_0^T \quad = \quad \mathbb{E}\left(\int_0^T \frac{\lambda\alpha^2}{1+\frac{\lambda\alpha^2}{\sigma_t^2}} e^{-2rt} S_t^2 \left(\frac{\partial G}{\partial x}(t, S_t, n_t, \tilde{T}_{N_t}) - \frac{G(t, S_t(1+\alpha), n_t+1, t) - G(t, S_t, n_t, \tilde{T}_{N_t})}{\alpha S_t}\right)^2 dt\right)$$

Remarque VII.3.8 *Le cas de temps de retour déterministes se traite en fait comme la limite du modèle de la section précédente en faisant tendre la fonction χ vers l'indicatrice $1_{[0,\tau[}$, ce qui montre bien intuitivement que les résultats ne sont pas modifiés.*

VII.3.5 Expression développée du smile

Dans cette partie, on se propose d'approcher la volatilité implicite du modèle avec saut de volatilité et retour instantané.
On part de l'expression du prix (VII.3.5) que l'on réécrit :

$$G(t, x, n, \tilde{T}) = e^{-\lambda(T-t)} \sum_{k \geq 0} \frac{(\lambda(T-t))^k}{k!} \mathbb{E}\left(\mathcal{K}\left(t, x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \tilde{\sigma}_k\right)\right) \tag{VII.3.19}$$

en définissant les variables aléatoires :

$$\tilde{\sigma}_k^2 = \frac{\int_t^T \sigma_k^2(s-, n, \tilde{T}, (\tilde{U}_i^k)_{1 \leq i \leq k}, (\tilde{\tau}_i^k)_{0 \leq i \leq k}) ds}{T-t} \tag{VII.3.20}$$

et en adoptant la fonction de prix Black-Scholes usuelle :

$$\mathcal{K}(t, x, \sigma) = \mathcal{H}(t, x, \sigma^2)$$

Ici, on se place dans le cas d'un Call européen de strike K.
La volatilité implicite $\sigma^{(imp)}$ vérifie alors par définition :

$$\mathcal{K}(t,x,\sigma^{(imp)}) = G(t,x,n,\tilde{T}) \tag{VII.3.21}$$

Alors, si l'on se fixe une volatilité de référence $\sigma^{(0)}$, un développement au premier ordre du membre de gauche par rapport à la volatilité donne :

$$\sigma^{(imp)} - \sigma^{(0)} \simeq \frac{G(t,x,n,\tilde{T}) - \mathcal{K}(t,x,\sigma^{(0)})}{\frac{\partial \mathcal{K}}{\partial \sigma}(t,x,\sigma^{(0)})} \tag{VII.3.22}$$

L'expression de la volatilité implicite s'obtiendra en choisissant un développement approprié du prix Black-Scholes dans l'expression (VII.3.19) du prix injectée dans le membre de droite de (VII.3.22).
On supposera bien sûr que les sauts du sous-jacent et de la volatilité sont assez faibles pour que l'on puisse admettre de tels développements limités.

Afin d'examiner en particulier l'influence du saut de volatilité, on commencera à considérer le cas où ce dernier est absent (ou tout au moins négligeable).

Cas où le saut de volatilité est négligeable

Dans ce cas, les volatilités effectives définies par (VII.3.20) sont constantes (à t fixé) et égales à

$$\sigma^{(0)} = \left(\frac{\int_t^T \sigma_{min}^2(s)ds}{T-t} \right)^{\frac{1}{2}} \tag{VII.3.23}$$

On développe donc le prix Black-Scholes dans (VII.3.19) par rapport au sous-jacent seul et de plus, ainsi qu'il apparaitra dans la suite, pour obtenir un effet de smile (dépendance par rapport au strike), on doit pousser le développement jusqu'à l'ordre trois.
Cela donne, si on effectue le développement à l'ordre quatre :

$$\begin{aligned}\mathcal{K}\left(t, x(1+\alpha)^k e^{-\alpha\lambda(T-t)}, \tilde{\sigma}_k\right) \quad \simeq \quad & \mathcal{K}(t,x,\sigma^{(0)}) + \frac{\partial \mathcal{K}}{\partial x}(t,x,\sigma^{(0)})\Delta_k x + \frac{1}{2}\frac{\partial^2 \mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)})\Delta_k^2 x^2 \\ & + \frac{1}{6}\frac{\partial^3 \mathcal{K}}{\partial x^3}(t,x,\sigma^{(0)})\Delta_k^3 x^3 + \frac{1}{24}\frac{\partial^4 \mathcal{K}}{\partial x^4}(t,x,\sigma^{(0)})\Delta_k^4 x^4 \end{aligned} \tag{VII.3.24}$$

où l'on a posé :

$$\Delta_k = (1+\alpha)^k e^{-\alpha\lambda(T-t)} - 1 \tag{VII.3.25}$$

En injectant cela dans l'expression (VII.3.19), on obtient une approximation du prix :

$$\begin{aligned} G(t,x,n,\tilde{T}) \quad \simeq \quad & \mathcal{K}(t,x,\sigma^{(0)}) + \frac{1}{2}\frac{\partial^2 \mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)}) < \Delta^2 > x^2 + \frac{1}{6}\frac{\partial^3 \mathcal{K}}{\partial x^3}(t,x,\sigma^{(0)}) < \Delta^3 > x^3 \\ & + \frac{1}{24}\frac{\partial^4 \mathcal{K}}{\partial x^4}(t,x,\sigma^{(0)}) < \Delta^4 > x^4 \end{aligned} \tag{VII.3.26}$$

où l'on a posé pour $i \in \mathbb{N}$:

$$< \Delta^i > = e^{-\lambda(T-t)} \sum_{k \geq 0} \frac{(\lambda(T-t))^k}{k!} \Delta_k^i \tag{VII.3.27}$$

et où l'on a tenu compte de :

$$< \Delta^1 > = 0$$

D'autre part, des calculs aisés donnent, puisque l'on est dans le cas d'un Call de strike K :

$$\begin{aligned}
\frac{\partial \mathcal{K}}{\partial \sigma}(t,x,\sigma^{(0)}) &= x\sqrt{T-t}\frac{1}{\sqrt{2\pi}}e^{-\frac{d_1^2}{2}} \\
\frac{\partial^2 \mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)}) &= \frac{1}{x\sigma^{(0)}\sqrt{T-t}}\frac{1}{\sqrt{2\pi}}e^{-\frac{d_1^2}{2}} \\
\frac{\partial^2 \mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)}) &= \frac{1}{x^2\sigma^{(0)}(T-t)}\frac{\partial \mathcal{K}}{\partial \sigma}(t,x,\sigma^{(0)}) \\
\frac{\partial^3 \mathcal{K}}{\partial x^3}(t,x,\sigma^{(0)}) &= -\frac{\partial^2 \mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)})\frac{1}{x}\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right) \\
&= -\frac{1}{x^3\sigma^{(0)}(T-t)}\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)\frac{\partial \mathcal{K}}{\partial \sigma}(t,x,\sigma^{(0)}) \\
\frac{\partial^4 \mathcal{K}}{\partial x^4}(t,x,\sigma^{(0)}) &= \frac{1}{x^2}\left[\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)^2+\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)-\frac{1}{(\sigma^{(0)})^2(T-t)}\right] \\
&\quad \frac{\partial^2 \mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)}) \\
&= \frac{1}{x^4\sigma^{(0)}(T-t)}\left[\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)^2+\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)-\frac{1}{(\sigma^{(0)})^2(T-t)}\right] \\
&\quad \frac{\partial \mathcal{K}}{\partial \sigma}(t,x,\sigma^{(0)})
\end{aligned} \tag{VII.3.28}$$

où l'on a posé

$$d_1 = \frac{\ln(\frac{x}{K}) + \left(r+\frac{(\sigma^{(0)})^2}{2}\right)(T-t)}{\sigma^{(0)}\sqrt{T-t}} \tag{VII.3.29}$$

Finalement, en injectant (VII.3.26) dans (VII.3.22), on obtient le développement :

$$\begin{aligned}
\sigma^{(imp)} - \sigma^{(0)} \simeq\ & \frac{1}{2\sigma^{(0)}(T-t)} < \Delta^2 > - \frac{1}{6\sigma^{(0)}(T-t)}\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right) < \Delta^3 > \\
&+\frac{1}{24\sigma^{(0)}(T-t)} < \Delta^4 > \\
&\left[\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)^2+\left(1+\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}\right)-\frac{1}{(\sigma^{(0)})^2(T-t)}\right]
\end{aligned} \tag{VII.3.30}$$

Si l'on se limite à l'ordre trois, ce développement se réécrit :

$$\sigma^{(imp)} - \sigma^{(AM)} \simeq -\frac{1}{6(\sigma^{(0)})^3(T-t)^2} < \Delta^3 > \ln\frac{x}{K} \tag{VII.3.31}$$

où $\sigma^{(AM)}$ désigne la volatilité implicite à la monnaie.
Typiquement, si la maturité est faible et le saut négatif, on aura :

$$< \Delta^3 >< 0$$

et l'on observe la décroissance de la volatilité implicite en fonction du strike.

Pour le voir, on écrit :

$$
\begin{aligned}
<\Delta^3> &= e^{-\lambda(T-t)}\sum_{k\geq 0}\frac{(\lambda(T-t))^k}{k!}\Big((1+\alpha)^{3k}e^{-3\alpha\lambda(T-t)}-3(1+\alpha)^{2k}e^{-2\alpha\lambda(T-t)} \\
&\quad +3(1+\alpha)^k e^{-\alpha\lambda(T-t)}-1\Big) \\
&= e^{-\lambda(T-t)}\Big(e^{\lambda(1+\alpha)^3(T-t)-3\alpha\lambda(T-t)}-3e^{\lambda(1+\alpha)^2(T-t)-2\alpha\lambda(T-t)} \\
&\quad +3e^{\lambda(1+\alpha)(T-t)-\alpha\lambda(T-t)}-e^{\lambda(T-t)}\Big) \\
&= \Big(2-3e^{\alpha^2\lambda(T-t)}+e^{3\alpha^2\lambda(T-t)+\alpha^3\lambda(T-t)}\Big) \\
&\simeq \alpha^3\lambda(T-t)<0
\end{aligned}
$$

Cas où le saut de volatilité est appréciable

Dans le cas général, on propose de remplacer le développement (VII.3.24) par un développement à l'ordre deux par rapport au sous-jacent et à la volatilité :

$$
\begin{aligned}
\mathcal{K}\left(t,x(1+\alpha)^k e^{-\alpha\lambda(T-t)},\tilde{\sigma}_k\right) &\simeq \mathcal{K}(t,x,\sigma^{(0)})+\frac{\partial\mathcal{K}}{\partial x}(t,x,\sigma^{(0)})\Delta_k x+\frac{1}{2}\frac{\partial^2\mathcal{K}}{\partial x^2}(t,x,\sigma^{(0)})\Delta_k^2 x^2 \\
&\quad +\frac{\partial\mathcal{K}}{\partial\sigma}(t,x,\sigma^{(0)})(\tilde{\sigma}_k-\sigma^{(0)})+\frac{1}{2}\frac{\partial^2\mathcal{K}}{\partial\sigma^2}(t,x,\sigma^{(0)})(\tilde{\sigma}_k-\sigma^{(0)})^2 \\
&\quad +\frac{\partial^2\mathcal{K}}{\partial\sigma\partial x}(t,x,\sigma^{(0)})(\tilde{\sigma}_k-\sigma^{(0)})\Delta_k x
\end{aligned}
\tag{VII.3.32}
$$

où $\sigma^{(0)}$ est la volatilité effective en absence de saut donnée par l'expression (VII.3.23).
Un calcul fastidieux permet de compléter les relations (VII.3.28) :

$$
\begin{aligned}
\frac{\partial^2\mathcal{K}}{\partial x\partial\sigma}(t,x,\sigma^{(0)}) &= -\frac{1}{x}\left(\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}-1\right)\frac{\partial\mathcal{K}}{\partial\sigma}(t,x,\sigma^{(0)}) \\
\frac{\partial^2\mathcal{K}}{\partial\sigma^2}(t,x,\sigma^{(0)}) &= \sqrt{T-t}\,d_1\left(\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}-1\right)\frac{\partial\mathcal{K}}{\partial\sigma}(t,x,\sigma^{(0)})
\end{aligned}
\tag{VII.3.33}
$$

Alors, en injectant (VII.3.32) dans (VII.3.22), on obtient le nouveau développement :

$$
\begin{aligned}
\sigma^{(imp)}-\sigma^{(0)} &\simeq <\tilde{\sigma}-\sigma^{(0)}>+\frac{1}{2\sigma^{(0)}(T-t)}<\Delta^2>-\left(\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}-1\right)<(\tilde{\sigma}-\sigma^{(0)})\Delta> \\
&\quad +\frac{1}{2}\sqrt{T-t}\,d_1\left(\frac{d_1}{\sigma^{(0)}\sqrt{T-t}}-1\right)<(\tilde{\sigma}-\sigma^{(0)})^2>
\end{aligned}
\tag{VII.3.34}
$$

où l'on a posé pour $i,j\in\mathbb{N}$:

$$
<(\tilde{\sigma}-\sigma^{(0)})^j\Delta^i>=e^{-\lambda(T-t)}\sum_{k\geq 0}\frac{(\lambda(T-t))^k}{k!}\Delta_k^i\mathbb{E}\Big((\tilde{\sigma}_k-\sigma^{(0)})^j\Big) \tag{VII.3.35}
$$

Compte tenu de (VII.3.29), et puisque $<\Delta>=0$, on obtient :

$$\begin{aligned}\sigma^{(imp)}-\sigma^{(0)} &\simeq Q+\left(\frac{r}{(\sigma^{(0)})^3}<(\tilde{\sigma}-\sigma^{(0)})^2>-\frac{1}{(\sigma^{(0)})^2(T-t)}<\tilde{\sigma}\Delta>\right)\ln\frac{x}{K}\\ &\quad+\frac{1}{2(\sigma^{(0)})^3(T-t)}<(\tilde{\sigma}-\sigma^{(0)})^2>\left(\ln\frac{x}{K}\right)^2\end{aligned}$$

où Q est un terme qui est indépendant du strike K.
Dans la pratique, deux types de développement, peuvent être adoptés en faisant intervenir la volatilité à la monnaie $\sigma^{(AM)}$ comme volatilité de référence.
On propose :

1. En dehors de la monnaie $(K \geq x)$

$$\begin{aligned}\sigma^{(imp)}-\sigma^{(AM)} &\simeq \left(\frac{r}{(\sigma^{(0)})^3}<(\tilde{\sigma}-\sigma^{(0)})^2>-\frac{1}{(\sigma^{(0)})^2(T-t)}<\tilde{\sigma}\Delta>\right)\ln\frac{x}{K}\\ &\quad+\frac{1}{2(\sigma^{(0)})^3(T-t)}<(\tilde{\sigma}-\sigma^{(0)})^2>\left(\ln\frac{x}{K}\right)^2\end{aligned}$$

2. Dans la monnaie $(K \leq x)$

$$\begin{aligned}\sigma^{(imp)}-\sigma^{(AM)} &\simeq \left(\frac{r}{(\sigma^{(0)})^3}<(\tilde{\sigma}-\sigma^{(0)})^2>-\frac{1}{(\sigma^{(0)})^2(T-t)}<\tilde{\sigma}\Delta>\right)\\ &\quad\left[\frac{x-K}{x}+\frac{1}{2}\left(\frac{x-K}{x}\right)^2\right]\\ &\quad+\frac{1}{2(\sigma^{(0)})^3(T-t)}<(\tilde{\sigma}-\sigma^{(0)})^2>\left(\frac{x-K}{x}\right)^2\end{aligned}$$

La deuxième expression s'obtient à partir de la deuxième avec un développement de $\ln\frac{x}{K} = -\ln\left(1-\frac{x-K}{x}\right)$.

Chapitre VIII

Calibration d'une diffusion avec sauts

VIII.1 Position du problème

L'objectif de cette partie est de calibrer un modèle d'actif de manière à pouvoir donner des prix d'options en cohérence avec ceux des Call et des Puts liquides traités sur le marché ainsi qu'avec la condition d'absence d'opportunité d'arbitrage.
Une façon bien connue pour résoudre le problème du smile consiste à choisir un modèle à volatilité stochastique. En particulier, le modèle dit à volatilité locale, où la volatilité ne dépend que du temps et du niveau du sous-jacent (ce qui restreint considérablement sa nature stochastique), est largement répandu depuis qu'il a été établi dans [43] une formule explicite donnant la volatilité locale en fonction des prix des Call.
Comme elle exige de disposer d'une nappe régulière de ces prix, fonction dérivable de la maturité et fonction deux fois dérivable du strike, cette formule n'est pas directement utilisable, ce qui a suscité de nombreux travaux dont [5], [82], [67], [17], [27] avec pour objectif une calibration stable.
De manière générale, les modèles à volatilité stochastique permettent de rendre compte de la structure de smile à maturité élevée tandis qu'il est également important d'ajouter une composante de sauts afin de modéliser les phénomènes observés à maturité courte pour lesquels la volatilité ne joue pas beaucoup.
La tentative naturelle d'intégrer le phénomène de smile constaté empiriquement en permettant au sous-jacent de subir des discontinuités est proposée historiquement dans [93] et est généralisée sous de nombreux aspects, ainsi qu'il est discuté dans [111] par exemple.

Dans l'étude qui suit, le cadre de départ est un espace de probabilité $(\Omega, \mathcal{F}, P)$ sur lequel le prix (S_t), de dimension m, subit des sauts introduits par une mesure de Poisson homogène $\Pi(dt, dz)$ sur l'espace $\mathbb{R} = \mathbb{R}^m$ muni de sa tribu Borélienne $\mathcal{R}$ et on note $\pi(dt, dz)$ la mesure intensité associée. Pour simplifier, on considérera par la suite que le sous-jacent est unidimensionnel ($m = 1$).
D'autre part, la composante martingale continue du prix est obtenue à partir d'un processus de Wiener (W_t) supposé indépendant de $\Pi(dt, dz)$ et l'on se donnera comme filtration $(\mathcal{F}_t)$ la filtration augmentée de celle engendrée par W et Π : $(\mathcal{F}_t)$ vérifie les conditions usuelles.
Dans ce contexte, sous la probabilité P, on suppose que S_t est une diffusion avec sauts, solution de l'EDS suivante, en se plaçant à un horizon fini T :

$$dS_t = S_{t-}[b(t,S_{t-})dt + \sigma(t,S_{t-})dW_t + \int_{z\in\mathbf{R}} \Phi(t,S_{t-},z)(\Pi(dt,dz) - \pi(dt,dz))], \quad 0 \le t \le T \tag{VIII.1.1}$$

où l'on conviendra que :

$$S_{0-} = S_0 > 0$$

et où l'on fait les hypothèses suivantes :

1. On suppose que la mesure d'intensité $\pi(dt,dz)$ admet la factorisation $\pi(dt,dz) = dt \otimes \rho(dz)$ où $\rho(dz)$ est une mesure positive *sigma*-finie sur $\mathbf{R}$.
2. Le taux d'intérêt r_t est supposé déterministe et borné et les fonctions $(t,x) \mapsto b(t,e^x)$, $(t,x) \mapsto \sigma(t,e^x)$ et $(t,x) \mapsto \Phi(t,e^x,z)$ sur $[0,T]\times\mathbb{R}$ sont mesurables et localement lipschitziennes en x : pour tout $n \in \mathbb{N}$ il existe une constante α_n ainsi qu'une fonction $\beta_n : \mathbf{R} \mapsto \mathbb{R}_+$ vérifiant $\int \beta_n(z)^2\rho(dz) < +\infty$ telles que pour $t \le n,\ |x| \le n,\ |x'| \le n$:

$$\begin{aligned} |b(t,e^x) - b(t,e^{x'})| &\le \alpha_n|x-x'| \\ |\sigma(t,e^x) - \sigma(t,e^{x'})| &\le \alpha_n|x-x'| \\ |\Phi(t,e^x,z)1_{\{|\Phi(t,e^x,z)|\le 1\}|} - \Phi(t,e^{x'},z)|1_{\{|\Phi(t,e^{x'},z)|\le 1\}|} &\le \beta_n(z)|x-x'| \\ |\Phi(t,e^x,z)1_{\{|\Phi(t,e^x,z)|> 1\}|} - \Phi(t,e^{x'},z)|1_{\{|\Phi(t,e^{x'},z)|> 1\}|} &\le \beta_n^2(z)|x-x'| \end{aligned} \tag{VIII.1.2}$$

3. La fonction $(t,x) \mapsto b(t,x)$ est bornée sur $[0,T]\times\mathbb{R}$ et l'on a aussi :

$$\forall (t,x) \in [0,T]\times\mathbb{R}, \quad 0 < \underline{\sigma} \le \sigma(t,x) \le \overline{\sigma} \tag{VIII.1.3}$$

   où $\underline{\sigma}$ et $\overline{\sigma}$ sont des constantes.
4. On suppose de plus que Φ vérifie

$$\forall (t,x,z) \in [0,T]\times\mathbb{R}\times\mathbb{R},\ |\Phi(t,x,z)| \le \Phi_0(z) \tag{VIII.1.4}$$

   où $\Phi_0 \ge 0$ est une fonction mesurable bornée vérifiant :

$$\int_{\mathbf{R}} \Phi_0^2(z)\rho(dz) < +\infty \tag{VIII.1.5}$$

   et qu'enfin

$$\forall (t,x,z) \in [0,T]\times\mathbb{R}\times\mathbf{R},\ \Phi(t,x,z) \ge \Phi_1 > -1 \tag{VIII.1.6}$$

   où Φ_1 est un réel fixé.

L'introduction d'un processus de Poisson ponctuel pour représenter les sauts du sous-jacent est justifié dans le chapitre II de [86] (et plus précisémment dans [45]) et l'on discutera plus loin les hypothèses présentées ci-dessus.

Ainsi qu'on l'a dit au début de cette introduction, on va proposer une procédure de calibration du modèle et, pour situer l'approche retenue, il faut signaler que les modèles de diffusion avec sauts ont suscité différentes méthodes de calibration dès le papier [93].
Dans [2], le modèle considéré est proche de (VIII.1.1) et l'effort de calibration est concentré sur la fonction de volatilité : exactement comme dans [43], cela suppose que l'on a obtenu une nappe régulière de volatilité implicite.
On peut également remarquer qu'en choisissant un modèle avec sauts, on se place délibérément

en situation de marché incomplet et donc, suivant le point de vue financier adopté au Chapitre VI, la calibration peut aussi être envisagée comme le choix d'une probabilité martingale parmi toutes celles qui lui sont équivalentes, un critère valable pour ce choix étant celui de l'entropie relative.
Dans le cas du modèle régi par (VIII.1.1), il s'agit précisément de contourner le problème posé par l'incomplétude du marché et le nombre insuffisant de contraintes en cherchant la probabilité martingale Q sur Ω minimisant la distance en entropie par rapport à P sous des contraintes de prix d'options liquides.
En dehors de la problématique liée à la calibration, l'existence et la détermination de la probabilité martingale minimisant l'entropie relative par rapport à une loi a priori dans le cadre d'un marché incomplet général est discuté dans [53] et [58] tandis que [94] et [34] par exemple s'intéressent au cas de diffusions avec sauts.
En adaptant les résultats des deux premières références, on a donné brièvement au Chapitre VI une caractérisation générale de la probabilité martingale calibrée minimale qui s'applique au modèle (VIII.1.1) car ce dernier génère un processus localement borné.
Cette caractérisation est cependant trop théorique car elle ne permet pas de construire la solution.
Récemment dans [35], grâce à une technique de descente du gradient, la méthode de la minimisation de l'entropie relative a permis de calibrer la classe des modèles de diffusion avec sauts lorsque le logarithme du sous-jacent est un processus de Lévy, ce qui revient à prendre constants les coefficient b, σ et Φ du présent modèle, la restriction que l'on ajoute ici étant de toutefois de borner les sauts à cause de la condition (VIII.1.4).
Dans la perspective de résoudre le problème de minimisation d'entropie dans le cas d'une diffusion avec sauts du type de celle proposée ici, on va suivre une méthode proche de la technique de contrôle optimal proposée dans [5] pour calibrer la surface de volatilité locale et reposant également sur la minimisation de l'entropie relative et pour cette étude, la référence principale sera [113].
En anticipant sur la suite, on insiste sur le fait que tous les paramètres du modèle a priori sont supposés fixés, et que la fonction à déterminer pour le modèle calibré n'est plus la volatilité comme dans le cas de [2], ce qui impliquerait la perte de l'absolue continuité entre mesures, mais le compensateur du processus de sauts.

Enfin, on reprécise le fait que le problème ne présente ici d'intérêt que si les coefficients du modèle a priori ne sont pas tous constants, auquel cas on est ramené à [35].

Tout d'abord, on introduit de nouveau la distance en entropie relative

$$H(Q\|P) = \begin{cases} \int \ln(\frac{dQ}{dP})dQ & \text{si } Q \text{ est absolument continue par rapport à } P \\ +\infty & \text{sinon} \end{cases}$$

et l'on rappelle le problème de calibration

$$(P^{\mathcal{M},\mathbf{C}}) \quad \text{trouver } Q \in \mathcal{M} \text{ qui minimise } H(Q\|P) \text{ sous la contrainte } \int f(S)dQ = \mathbf{C}$$

où l'on note $\mathcal{M}$ l'ensemble des probabilités martingales sur $(\Omega, \mathcal{F})$ et où $f = (f^1, ..., f^d)$ désigne le vecteur des fonctions de payoff actualisées de d options liquides dont on observe les prix $\mathbf{C} = (C^1, ..., C^d)$.
Ce problème n'admet alors de solution que sous la condition géométrique détaillée antérieurement :

Lemme VIII.1.1 *Pour que le problème* $(P^{\mathcal{M},\mathbf{C}})$ *admette une solution il est nécessaire qu'il existe un sous-espace vectoriel* E *de* $\mathbb{R}^d$ *tel que si* $P^{f\circ S}_{|E}$ *désigne la restriction à* E *de l'image de* P *par* $f \circ S$,

$$(C^{\mathbf{C}}) \begin{cases} \mathbf{C} \in Int_E(Conv(Supp(P^{f\circ S}_{|E}))) \text{ l'intérieur dans } E \text{ de l'enveloppe convexe du support de } P^{f\circ S}_{|E} \\ \forall \eta \in \mathcal{P}(\Omega) \text{ absolument continue par rapport à } P \text{ t.q. } \int_\Omega f(S)d\eta = \mathbf{C},\ \eta((f\circ S)^{-1}(E)) = 1\ . \end{cases}$$

Afin d'utiliser des techniques classiques de contrôle, on va en fait se restreindre à un sous-ensemble $\mathcal{M}'$ de $\mathcal{M}$ permettant de définir un espace de contrôles convenables, et en particulier, pour des raisons également financières, il est préférable de rechercher une solution équivalente à la probabilité a priori P.
Après avoir fait ce choix, on définira la version pénalisée du problème de calibration qui a l'avantage d'assurer l'unicité de la solution du problème dual et c'est en définitive avec elle que l'étude sera menée. L'objectif est de montrer l'existence d'une solution stable.

VIII.2 Choix d'un ensemble de probabilités martingales

L'objet de cette partie est d'abord de restreindre le choix de la solution du problème à un ensemble convenable de probabilités martingales équivalentes puis de donner un critère suffisant pour résoudre le problème (Lemme VIII.2.7).
Commençons par discuter les hypothèses faites sur le modèle régi par l'EDS (VIII.1.1) définissant le sous-jacent comme un processus de diffusion avec sauts.
L'existence et l'unicité de la solution de ce type d'EDS et le lien avec le problème des martingales, ont été étudiés dans les articles [121], [79] et plus généralement dans [86] et, ici, on se réfère essentiellement aux conditions données par [68].
Ici, par commodité, les conditions sont adaptées à l'EDS transformée de (VIII.1.1) par le changement de variable logarithmique :

$$X_t = \ln(S_t)$$

Afin de préciser cela, sous réserve d'existence de la solution S_t de l'EDS (VIII.1.1) (l'unicité n'étant pas forcément acquise), cette dernière s'écrit selon la formule de l'exponentielle de Doléans-Dade

$$\begin{aligned} S_t &= S_0 e^{\int_0^t b(s,S_{s-})ds + \int_0^t \sigma(s,S_{s-})dW_s - \frac{1}{2}\int_0^t \sigma(s,S_{s-})^2 ds} e^{\int_{[0,t]\times\mathbf{R}} \Phi(s,S_{s-},z)(\Pi-\pi)(ds,dz)} \\ &\quad \prod_{0<\tau_n\le t} (1+\Phi(\tau_n, S_{\tau_n -}, \zeta_{\tau_n}))e^{-\Phi(\tau_n,S_{\tau_n -},\zeta_{\tau_n})} \end{aligned} \tag{VIII.2.1}$$

où, avec la Proposition II 1.14 de [68], on a introduit un processus ζ à valeurs dans $\mathbf{R}$ ainsi qu'une suite (τ_n) de temps d'arrêt tels que la mesure Π s'écrive :

$$\Pi(dt,dz) = \sum_{n\ge 0} \delta_{(\tau_n,\zeta_{\tau_n})} \tag{VIII.2.2}$$

Avec le Théorème I 4.61 de [68], le produit infini du membre de droite de (VIII.2.1) est absolument convergent p.s. et alors ne s'annule pas avec la condition (VIII.1.6) : la solution S_t de l'EDS (VIII.1.1) est strictement positive et on peut la réécrire plus simplement :

$$\begin{aligned} S_t &= S_0 e^{\int_0^t b(s,S_{s-})ds + \int_0^t \sigma(s,S_{s-})dW_s - \frac{1}{2}\int_0^t \sigma(s,S_{s-})^2 ds} e^{\int_{[0,t]\times\mathbf{R}} \Phi(s,S_{s-},z)(\Pi-\pi)(ds,dz)} \\ &\quad e^{-\int_{[0,t]\times\mathbf{R}}(\Phi-\ln(1+\Phi))(s,S_{s-},z)\Pi(ds,dz)} \end{aligned} \tag{VIII.2.3}$$

Ainsi, le changement de variable $X_t = \ln(S_t)$ est valide et l'EDS vérifiée par X est :

$$\begin{aligned} dX_t &= (b_*(t, X_{t-}) - \frac{1}{2}\sigma_*^2(t, X_{t-}))dt + \sigma_*(t, X_{t-})dW_t + \int_{z\in\mathbf{R}} \ln(1 + \Phi_*(t, X_{t-}, z))(\Pi - \pi)(dt, dz) \\ &+ \int_{z\in\mathbf{R}} (\ln(1 + \Phi_*(t, X_{t-}, z)) - \Phi_*(t, X_{t-}, z))\pi(dt, dz), \quad 0 \le t \le T \qquad \text{(VIII.2.4)} \end{aligned}$$

où l'on a posé :

$$\begin{aligned} &\forall(t, x) \in [0, T] \times \mathbb{R}, \ b_*(t, x) = b(t, e^x) \\ &\forall(t, x) \in [0, T] \times \mathbb{R}, \ \sigma_*(t, x) = \sigma(t, e^x) \\ &\forall(t, x) \in [0, T] \times \mathbb{R}, \ \forall z \in \mathbf{R}, \Phi_*(t, x, z) = \Phi(t, e^x, z) \end{aligned}$$

et où l'on a utilisé les conditions (VIII.1.4) et (VIII.1.5) pour assurer la définition de la dernière intégrale du membre de droite de (VIII.2.4).
Remarquons par ailleurs que ces mêmes majorations (VIII.1.4) et (VIII.1.5) assurent l'existence de la deuxième intégrale (voir [68] Théorème II 1.33 b/ et Définition II 1.27).
Si l'on reprend alors les hypothèses données dans l'introduction, l'ensemble de conditions (VIII.1.2) exprime que les coefficients de l'EDS (VIII.1.1) sont localement Lipschitziens et on vérifie qu'ils ont une croissance linéaire car b et σ sont bornés tandis que l'on a la majoration (VIII.1.4) pour Φ.
Le Théorème III 2.32 de [68] assure alors l'unicité et l'existence (au sens fort) de la solution de l'EDS (VIII.2.4), puis, en revenant à la variable originale, celles de la solution de l'EDS (VIII.1.1).

Remarque VIII.2.1 *On a supposé que le coefficient de sauts Φ est borné (avec Φ_0), ce qui implique l'hypothèse pratique que les sauts de X_t sont bornés.*
Cette hypothèse servira surtout à assurer la validité de la majoration exponentielle (VIII.4.19) dont l'intérêt sera discuté plus loin.
En anticipant sur la suite, dans le cas où les options servant à la calibration sont bornées, on peut relâcher cette hypothèse et affaiblir également la condition d'intégrabilité (VIII.1.5).
Le problème de calibration $(P^{\mathcal{M}'})$ a donc bien un sens et l'on se donne à présent l'ensemble $\mathcal{M}'$ de probabilités où l'on souhaite en rechercher la solution.
On considère la tribu prévisible $\mathcal{P}$ pour $(\Omega, (\mathcal{F}_t))$, la tribu $\mathcal{R}$ associée à l´espace $\mathbf{R}$ et l'on note $\tilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{R}$.
Classiquement, dit ici de manière informelle, avec les Théorèmes III 3.24 et III 5.19 de [68] résumés dans le Théorème 3.2 de [34] par exemple, un changement de probabilité équivalente à P détermine un processus prévisible $(G_t(\omega))$ et une fonction $\tilde{\mathcal{P}}$-mesurable $(H_t(\omega, z))$ ajoutant un "drift" au Brownien en changeant la mesure d'intensité de la mesure de Poisson.
On va exprimer cela de façon précise en se limitant au cas où $(G_t(\omega))$ et $(H_t(\omega, z))$ sont bornées, ce qui suffit dans la pratique.
Lemme VIII.2.2 *Soit $(H_t(\omega, z))$ une fonction $\tilde{\mathcal{P}}$-mesurable strictement positive et soit $\psi \in L^2(\mathbf{R}, \rho) \bigcap L^\infty(\mathbf{R}, \rho)$ telle que*

$$\forall(\omega, t, z) \in \Omega \times [0, T] \times \mathbb{R}, \quad |\ln(H_t)|(\omega, z) \le \psi \qquad \text{(VIII.2.5)}$$

Soit G_t le processus prévisible défini par l'égalité

$$G_t = \frac{1}{\sigma(t, S_{t-})}\left[r_t - b(t, S_{t-}) - \int_{\mathbf{R}} \Phi(t, S_{t-}, z)(H_t(., z) - 1)d\rho(z)\right] \qquad \text{(VIII.2.6)}$$

Alors le processus

$$Z_t^H = e^{\int_0^t G_s dW_s - \frac{1}{2}\int_0^t G_s^2 ds} e^{\int_{]0,t]\times\mathbf{R}} (H_s(.,z)-1)(\Pi-\pi)(ds,dz)} \prod_{0<\tau_n\le t} H_{\tau_n}(., \zeta_{\tau_n}) e^{1-H_{\tau_n}(.,\zeta_{\tau_n})} \qquad \text{(VIII.2.7)}$$

est une martingale positive, où les temps d'arrêt τ_n et le processus ζ sont associés à l'écriture (VIII.2.2) de la mesure de Poisson Π.
On définit de plus une probabilité martingale Q^H équivalente à P sur $(\Omega, \mathcal{F}_T, P)$ par :

$$\left.\frac{dQ^H}{dP}\right|_{\mathcal{F}_T} = Z_T^H$$

Sous Q^H, $(W_t - \int_0^t G_s ds)_{0\leq t\leq T}$ est un processus de Wiener et le compensateur de la mesure ponctuelle $\Pi(dt,dz)$ est $H(t,z)\pi(dt,dz)$ et le sous-jacent actualisé $\tilde{S}_t \equiv e^{-\int_0^t r_s ds} S_t$ est une martingale.

Preuve : Sans perte de généralité, on peut supposer

$$\forall(\omega,z) \in \Omega\times\mathbf{R},\ H_0(\omega,z) = 1 \tag{VIII.2.8}$$

puisque cette valeur n'intervient pas dans l'expression (VIII.2.7) de la densité et que π ne charge pas $\{0\}\times\mathbf{R}$.
Ensuite, le processus prévisible G_t est défini et borné ainsi qu'on peut le montrer de façon précise : avec l'inégalité des accroissements finis appliquée à $x \mapsto e^x$ en $\ln H_t$ ainsi que la majoration (VIII.2.5), on obtient en effet :

$$|H_t - 1| \leq e^{\|\psi\|_\infty} |\ln(H_t)|$$

et en utilisant encore l'hypothèse (VIII.2.5), on a simplement :

$$|H_t - 1| \leq e^{\|\psi\|_\infty} \psi \tag{VIII.2.9}$$

On peut alors écrire :

$$\left|\int_{\mathbf{R}} \Phi(t,S_{t-},z)(H_t(.,z)-1)d\rho(z)\right| \leq e^{\|\psi\|_\infty}\left(\int_{\mathbf{R}} \Phi_0^2(z)d\rho(z)\right)^{\frac{1}{2}}\left(\int_{\mathbf{R}}\psi^2(z)d\rho(z)\right)^{\frac{1}{2}} \tag{VIII.2.10}$$

et le membre de droite de (VIII.2.10) est fini en vertu de (VIII.1.5) ainsi que de l'hypothèse sur ψ.
En injectant cette majoration dans la relation (VIII.2.6) et en tenant compte de la relation (VIII.1.3) ainsi que du fait que les coefficients r_t et b sont supposés bornés, on conclut que G_t est borné.
Alors, le fait que Z_t^H est une martingale est une conséquence de la Proposition I 1.11 de [87] mais on en redonne les éléments de preuve.
En utilisant la formule d'Itô, ou grâce à l'expression de l'exponentielle de Doléans-Dade et la relation (VIII.2.2), on voit que Z_t^H vérifie l'EDS

$$dZ_t^H = Z_{t-}^H G_t dW_t + Z_{t-}^H \int_{\mathbf{R}} (H_t(.,z)-1)(\Pi-\pi)(dt,dz) \tag{VIII.2.11}$$

avec

$$Z_{0-}^H = Z_0^H = 1$$

c'est-à-dire que l'on a :

$$Z_t^H = 1 + \int_0^t Z_{s-}^H G_s dW_s + \int_{]0,t]} Z_{s-}^H \int_{\mathbf{R}} (H_s(.,z)-1)(\Pi-\pi)(ds,dz) \tag{VIII.2.12}$$

Avec (VIII.2.7), on a $Z_t^H = Z_{t-}^H H_{\tau_n}(.,\zeta_{\tau_n})$ et donc c'est une martingale locale qui est localement de carré intégrable et dont on peut calculer le crochet oblique.
Avec le Théorème I 4.52 de [68], on calcule d'abord la variation quadratique :

$$[Z^H, Z^H]_t = \int_0^t (Z_s^H G_s)^2 ds + \int_{]0,t]\times\mathbf{R}} (Z_{s-}^H (H_s(.,z)-1))^2 \Pi(ds,dz)$$

En compensant la dernière expression, il vient :

$$< Z^H >_t = \int_0^t (Z_s^H G_s)^2 ds + \int_{]0,t]\times\mathbf{R}} (Z_s^H (H_s(.,z)-1))^2 \rho(dz)ds$$

en remarquant que le crochet oblique est bien défini car la martingale locale $Z_t^H = Z_{t-}^H H_{\tau_n}(.,\zeta_{\tau_n})$ est localement de carré intégrable.
Montrons alors $\mathbb{E}^P < Z^H >_T < +\infty$.
Or on peut écrire

$$Z_t^H = e^{M_t - \frac{1}{2}\int_0^t G_s^2 ds - \int_{]0,t]\times\mathbf{R}} (H_s(z)-1-\ln(H_s(z)))\pi(ds,dz)} \qquad \text{(VIII.2.13)}$$

où M_t est définie par

$$M_t = \int_0^t G_s dW_s + \int_{]0,t]\times\mathbf{R}} \ln(H_s(.,z))(\Pi-\pi)(ds,dz) \qquad \text{(VIII.2.14)}$$

et donc, compte tenu de $x-1-\ln x \geq 0$, il vient $Z_t^H \leq e^{M_t}$.
Mais M_t est une martingale vérifiant les conditions d'application du Lemme VIII.2.4 énoncé après la présente preuve : elle est quasi-continue à gauche (avec le Corollaire II 1.19 de [68] puisque $\pi(\{t\}\times\mathbf{R}) = 0\ \forall t$), ses sauts sont bornés et son crochet oblique vérifie : $< M >_t = \int_0^t G_s^2 ds + \int_{]0,t]\times\mathbf{R}} (\ln(H_s(.,z)))^2 \rho(dz)ds \leq Kt$ où K est une constante, à cause en particulier de la majoration (VIII.2.5) de $|\ln(H)|$) et de la condition d'intégrabilité de la mesure de Lévy ρ.
Le lemme précédemment cité permet alors d'obtenir :

$$\mathbb{E}^P(e^{2\sup_{0\leq s\leq T}|M_s|}) < +\infty \qquad \text{(VIII.2.15)}$$

Avec le fait que G est bornée et en obtenant aisément $\sup_{0\leq t\leq T}\int_{\mathbf{R}}(H_t(z)-1)^2\rho(dz) < +\infty$, on en déduit bien que $\mathbb{E}^P[Z^H,Z^H]_T = \mathbb{E}^P < Z^H >_T < +\infty$ et donc que Z_t^H est une martingale de carré intégrable avec le Corollaire 3 du Théorème II 6.27 de [101] et en particulier Q^H est une probabilité.
Pour terminer la preuve, on sait que sous Q^H, la mesure Π a pour compensateur $\tilde{\pi}$ donné par le Théorème III 3.17 de [68] par :

$$\tilde{\pi}(dt,dz) = Y\pi(dt,dz) \qquad \text{(VIII.2.16)}$$

où Y est une fonction positive $\tilde{\mathcal{P}}$-mesurable caractérisée par :

$$\mathbb{E}^P\left[\int_{[0,T]\times\mathbf{R}} Y_s(.,z) Z_{s-}^H U_s(.,z)\Pi(ds,dz)\right] = \mathbb{E}^P\left[\int_{[0,T]\times\mathbf{R}} Z_s^H U_s(.,z)\Pi(ds,dz)\right] \qquad \text{(VIII.2.17)}$$

pour tout U $\tilde{\mathcal{P}}$-mesurable et positive.
Mais, avec l'écriture (VIII.2.2) de la mesure de Poisson ainsi que l'expression (VIII.2.7) de la densité et la condition initiale (VIII.2.8) , on a :

$$\begin{aligned}\int_{[0,T]\times\mathbf{R}} Z_s^H U_s(.,z)\Pi(ds,dz) &= \sum_{0\leq\tau_n\leq T} Z_{\tau_n}^H U_{\tau_n}(.,\zeta_{\tau_n}) \\ &= \sum_{0\leq\tau_n\leq T} Z_{\tau_n-}^H H_{\tau_n}(.,\zeta_{\tau_n}) U_{\tau_n}(.,\zeta_{\tau_n}) \\ &= \int_{[0,T]\times\mathbf{R}} Z_{s-}^H H_s(.,z) U_s(.,z)\Pi(ds,dz)\end{aligned}$$

et la relation caractéristique (VIII.2.17) tient avec $Y = H$, ce qui règle le cas de la mesure de sauts.
D'autre part, on sait que sous Q^H, le processus (W_t) est une martingale (Théorème II 2 de [101]).
De plus, toujours sous Q^H, (W_t) reste continu et il est alors clair que son triplet de caractéristiques (au sens de la Définition II 2.6 de [68]) est du type $(B_t, C_t, 0)$ où (B_t) et (C_t) sont des processus réels à variation bornée, (C_t) étant également continu.
Avec le Théorème III 3.24 de [68], on obtient :

$$\begin{aligned} C_t &= t && \text{(VIII.2.18)} \\ B_t &= \int_0^t \beta_s ds && \text{(VIII.2.19)} \\ & && \text{(VIII.2.20)} \end{aligned}$$

où (β_t) est un processus prévisible.
Par définition des processus caractéristiques, et puisque (W_t) est continu, (B_t) est le compensateur prévisible de (W_t) et le processus $(W_t - B_t)$ est donc une martingale locale.
Cette martingale locale est de plus continue avec l'identité (VIII.2.19) (et cela s'obtient également avec le Lemme I 4.24 de [68]) et, de plus, son crochet oblique est donnée par la première égalité.
En appliquant la caractérisation classique de Lévy donnée par exemple dans le Théorème II 4.4 de [68] ou dans le Théorème III 3.16 de [76], on en déduit que le processus $(W_t - B_t)$ est un processus de Wiener.
Pour identifier le processus β, la relation (3.28) du Théorème III 3.24 de [68] le caractérise par la relation :

$$\langle (Z^H)^c_t, W_t \rangle = \int_0^t \beta_s Z^H_{s-} ds \qquad \text{(VIII.2.21)}$$

où $(Z^H)^c$ est la composante martingale continue de Z^H relativement à la probabilité P.
Mais l'EDS (VIII.2.11) implique :

$$(Z^H)^c_t = \int_0^t Z^H_{s-} G_s dW_s$$

et la relation caractéristique tient avec $\beta = G$.
On a donc obtenu le fait que $(W_t - \int_0^t G_s ds)$ est un processus de Wiener sous Q^H et que la mesure ponctuelle Π a pour compensateur $H\pi$.
Pour finir, avec l'égalité (VIII.2.6), le sous-jacent actualisé $\tilde{S}_t$ obéit à l'EDS suivante sous Q^H :

$$d\tilde{S}_t = \tilde{S}_{t-}\sigma(t, S_{t-})(dW_t - G_t dt) + \tilde{S}_{t-} \int_{\mathbf{R}} \Phi(t, S_{t-}, z)(\Pi - H_t(., z)\pi)(dt, dz) \qquad \text{(VIII.2.22)}$$

Il est alors facile de vérifier que cette martingale locale sous Q^H est une vraie martingale en suivant les lignes de la preuve concernant le processus de densité Z_t^H, martingale vraie sous P. ∎

Remarque VIII.2.3 *On a utilisé l'égalité $Z^H_{t-} = Z^H_t$ ps pour $0 \leq t \leq T$.*
Cela résulte du fait que la mesure aléatoire $\Pi(dt, dz)$ définit un processus à accroissement indépendants et stationnaires, d'après le Théorème II 4.15 de [68] et n'admet donc pas de temps fixe de discontinuité d'après ce même théorème, ou en invoquant le n° II 4.3 de [68].

On a également fait usage de la majoration exponentielle suivante, valable pour les martingales à sauts bornés, et qui généralise la Proposition 15 de [86] en se déduisant exactement de la même manière du Théorème 13 de [86] :

Lemme VIII.2.4 *Pour $k \geq 0$ on peut déterminer un nombre $c < +\infty$ tel que si $(M_t)_{0 \leq t \leq T}$ est une martingale locale réelle, nulle en 0, quasi-continue à gauche, telle que $\forall t \geq 0$, $< M, M >_t \leq kt$ et que tous ses sauts soient bornés par k, alors on a la majoration exponentielle :*

$$\mathbb{E}^P(e^{\sup_{0\leq t\leq T}|M_t|}) \leq c \tag{VIII.2.23}$$

Dans la suite, il est commode de travailler en variable logarithmique, c'est-à-dire avec la solution $X_t = \ln(S_t)$ de l'EDS (VIII.2.4).
Le lemme précédent permet de considérer un ensemble de probabilités martingales paramétré par la fonction H déterminant l'intensité de la mesure de Poisson Π. Comme il est souhaitable de la choisir de manière à conserver le caractère markovien du sous-jacent, les changements d'intensité considérés seront du type $K(t, X_{t-}, z)$ où K est choisi dans $\mathcal{K}^{\overline{H}}$ avec $\overline{H} > 1$, en définissant

$$\mathcal{K}^{\overline{H}} = \left\{ K(t,x,z) \text{ fonctions Boréliennes sur } [0,T]\times\mathbb{R}\times\mathbf{R} \text{ telles que } |\ln(K)| \leq \ln(\overline{H})\psi_0 \right\}$$

où ψ_0 est une fonction positive fixée de $L^2(\mathbf{R}, \rho)$ telle que $\|\psi_0\|_\infty \leq 1$.

Il est alors naturel d'introduire la solution $(X_u^{t,x})_{t\leq u\leq T}$ de l'EDS (VIII.2.4) avec la condition initiale :

$$X_t^{t,x} = x$$

Alors, avec le Lemme VIII.2.2, on peut définir la famille de processus densité, pour $K \in \mathcal{K}^{\overline{H}}$, $(t,x) \in [0,T]\times\mathbb{R}$, en arrangeant l'expression (VIII.2.7) :

$$\begin{aligned}\forall u \in [t,T],\ Z_u^{K,t,x} &= e^{\int_t^u G_s^{K,t,x} dW_s - \frac{1}{2}\int_t^u (G_s^{K,t,x})^2 ds} \\ &\quad e^{\int_{]t,u]\times\mathbf{R}} \ln(K(s,X_{s-}^{t,x},z))(\Pi-\pi)(ds,dz) - \int_{]t,u]\times\mathbf{R}}(K-1-\ln(K))(s,X_{s-}^{t,x},z)\pi(ds,dz)}\end{aligned} \tag{VIII.2.24}$$

en posant :

$$\forall u \in [t,T],\ G_u^{K,t,x} = \frac{1}{\sigma_*(u, X_{u-}^{t,x})}\left[r_u - b_*(u, X_{u-}^{t,x}) - \int_{\mathbf{R}} \Phi_*(u, X_{u-}^{t,x}, z)(K(u, X_{u-}^{t,x}, z) - 1)d\rho(z)\right] \tag{VIII.2.25}$$

La famille $Z^{K,t,x}$ permet enfin de définir une famille de probabilités sur $\mathcal{F}_T$ en posant

$$\left.\frac{dQ^{K,t,x}}{dP}\right|_{\mathcal{F}_T} = Z_T^{K,t,x} \tag{VIII.2.26}$$

Pour le moment, on considère seulement $(t,x) = (0, X_0)$.
Si l'on note enfin

$$\mathcal{K}_0^{\overline{H}} = \left\{ K \in \mathcal{K}^{\overline{H}} \text{ et } \int \|f_*(X)\| dQ^{K,0,X_0} < +\infty \right\}$$

où l'on a naturellement posé :

$$\forall x \in \mathbb{R},\ f_*(x) = f(e^x)$$

on se propose de remplacer le problème général $(P^{\mathcal{M},\mathbf{C}})$ par :

$(P^{\overline{H},\mathbf{C}})$ trouver $K \in \mathcal{K}_0^{\overline{H}}$ qui minimise $H(Q^{K,0,X_0}\|P)$ sous la contrainte $\int f_*(X)dQ^{K,0,X_0} = \mathbf{C}$

Avec le Lemme I.2.2 de [87], on obtient le résultat suivant

Lemme VIII.2.5 *L'ensemble*

$$\left\{ Q^{K,0,X_0} \;\middle|\; K \in \mathcal{K}^{\overline{H}} \right\}$$

est convexe.

et donc le problème de calibration relève donc théoriquement de l'analyse convexe classique.
Par rapport au problème général $(P^{\mathcal{M},\mathbf{C}})$, puisque l'intensité $K > 0$ détermine un changement de probabilité $Q^{K,0,X_0}$ équivalente à la probabilité d'origine P, et pas seulement absolument continue par rapport à cette dernière, la condition d'existence est renforcée :

Lemme VIII.2.6 *Pour que le problème $(P^{\overline{H},\mathbf{C}})$ admette une solution, il est nécessaire que la condition suivante soit vérifiée*

$$(C_e^{\mathbf{C}}) \begin{cases} \mathbf{C} \in ri(Conv(Supp(P^{f_*(X)}))) \\ \textit{l'intérieur relatif de l'enveloppe convexe du support de } P^{f_*(X)} \end{cases}$$

C'est en fait une condition nécessaire de non vacuité de l'ensemble

$$\mathcal{K}_c^{\overline{H},\mathbf{C}} = \left\{ K \in \mathcal{K}_0^{\overline{H}} \text{ telle que } \int f_*(X) dQ^{K,0,X_0} = \mathbf{C} \right\}$$

On attaque alors le problème $(P^{\overline{H},\mathbf{C}})$ en introduisant la fonction valeur (initiale) :

$$\forall \theta \in \mathbb{R}^d,\ V(\theta) = \sup_{K \in \mathcal{K}_0^{\overline{H}}} \left[\theta \cdot \int f_*(X) dQ^{K,0,X_0} - H(Q^{K,0,X_0} \| P) \right] \tag{VIII.2.27}$$

ce qui permet d'énoncer un résultat fondamental

Proposition VIII.2.7 *Si la fonction $\theta \mapsto V(\theta) - \mathbf{C} \cdot \theta$ admet un minimum θ^* où elle est différentiable et s'il existe une fonction $K^* \in \mathcal{K}_0^{\overline{H}}$ réalisant*

$$V(\theta^*) = \theta^* \cdot \int f_*(X) dQ^{K^*,0,X_0} - H(Q^{K^*,0,X_0} \| P) \tag{VIII.2.28}$$

alors K^* est solution du problème $(P^{\overline{H},\mathbf{C}})$.

On n'en fait pas la preuve, similaire à celle de la Proposition VIII.3.1 à venir, et reposant sur une technique de dualité classique.
On formulera dans la section prochaine le problème de calibration sous forme pénalisée en fixant les objectifs de la présente étude. Signalons également que cette dernière mériterait d'être complétée par celle de la condition géométrique $(C_e^{\mathbf{C}})$.
On peut en effet considérer que c'est une condition nécessaire pour que le choix de la mesure a priori μ soit cohérent avec les observations $\mathbf{C}$.
Cela revient à déterminer l'image de l'ensemble $\left\{ Q^{K,0,X_0} \;\middle|\; K \in \mathcal{K}^{\overline{H}} \right\}$ par le prix des options de calibration.
Dans le cas d'un seul prix, on peut se référer à [9] ainsi qu'à [44] et dans le cas général, il s'agit d'un ensemble convexe en tant qu'image d'un convexe par une application linéaire.
Il serait bien sûr intéressant d'obtenir le fait qu'il est d'intérieur non vide (avec un bon choix des options de calibration) et d'avoir une caractérisation plus fine, ce qui ne sera pas plus approfondie dans ce chapitre.

VIII.3 Pénalisation du problème

VIII.3.1 Formulation générale

Le programme défini dans le lemme précédent a l'inconvénient de ne pas assurer l'existence et l'unicité du minimum si la fonction valeur V n'est pas strictement convexe.
Dans ce dernier cas, potentiellement, on n'a pas un comportement régulier de la solution face aux variations du prix $\mathbf{C}$ des options servant de contraintes.
L'idée est alors de pénaliser la fonction à minimiser en introduisant une fonction $\chi : \mathbb{R}^d \mapsto \mathbb{R}$ convexe, minimale en $\mathbf{0}$, deux fois différentiable et de hessien défini positif, et telle que :

$$\chi'(\mathbb{R}^d) = \mathbb{R}^d \tag{VIII.3.1}$$

On considère alors la fonction convexe conjuguée $\chi^* : \mathbb{R}^d \longrightarrow]-\infty, +\infty]$ définie par :

$$\forall u \in \mathbb{R}^d,\ \chi^*(u) = \sup_{\theta\in\mathbb{R}^d} (u \cdot \theta - \chi(\theta)) \tag{VIII.3.2}$$

et on sait que χ^* est finie et est strictement convexe sur $\mathbb{R}^d$ puisque χ est fermée (car finie et continue) et de type de Legendre et qu'elle vérifie (VIII.3.1).
On sait aussi qu'elle est minimale en $\mathbf{0}$ et que sa conjuguée est χ.
Un choix usuel est :

$$\chi : \theta \in \mathbb{R}^d \longmapsto \frac{1}{2}\sum_{i=1}^{d} \omega_i(\theta_i)^2$$

avec

$$\forall 1 \leq i \leq d,\ \omega_i > 0$$

On se propose donc de remplacer le problème de calibration $(P^{\overline{H},\mathbf{C}})$ par :

$$(P^{\overline{H},\mathbf{C}}_{\chi^*}) \quad \text{trouver } K \in \mathcal{K}_0^{\overline{H}} \text{ qui minimise } H(Q^{K,0,X_0}\|P) + \chi^*(\textstyle\int f_*(X)dQ^{K,0,X_0} - \mathbf{C})$$

La fonction $\theta \mapsto V(\theta)+\chi(-\theta)-\mathbf{C}\cdot\theta$ est semi-continue inférieurement (puisque V est en particulier la borne supérieure d'une famille de fonctions continues), et, avec le terme de pénalisation, elle est strictement convexe et tend vers l'infini en l'infini : elle admet donc un minimum unique.
L'énoncé qui suit donne le programme de l'étude :

Proposition VIII.3.1 *La fonction $\theta \mapsto V(\theta) + \chi(-\theta) - \mathbf{C} \cdot \theta$ admet un minimum unique θ^*. De plus, si elle est différentiable en θ^* et s'il existe une fonction $K^* \in \mathcal{K}_0^{\overline{H}}$ réalisant*

$$V(\theta^*) = \theta^* \cdot \int f_*(X)dQ^{K^*,0,X_0} - H(Q^{K^*,0,X_0}\|P) \tag{VIII.3.3}$$

alors K^* est solution du problème $(P^{\overline{H},\mathbf{C}}_{\chi^*})$.

Preuve : Suivant la technique classique, on fait apparaitre la minimisation de l'entropie pénalisée comme la réalisation d'un programme "dual" associé au problème primal, consistant à minimiser parmi les éléments $K \in \mathcal{K}_0^{\overline{H}}$ la fonctionnelle

$$\begin{aligned}
& H(Q^{K,0,X_0}\|P) + \chi^*(\int f_*(X)dQ^{K,0,X_0} - \mathbf{C}) \\
= & \sup_{\theta\in\mathbb{R}^d} \Big[H(Q^{K,0,X_0}\|P) - \theta\cdot\Big(\int f_*(X)dQ^{K,0,X_0} - \mathbf{C}\Big) - \chi(-\theta)\Big] \\
= & -\inf_{\theta\in\mathbb{R}^d} \Big\{\Big[\theta\cdot\int f_*(X)dQ^{K,0,X_0} - H(Q^{K,0,X_0}\|P)\Big] + \chi(-\theta) + \mathbf{C}\cdot\theta\Big\} \\
\geq & -\inf_{\theta\in\mathbb{R}^d} \Big(V(\theta) + \chi(-\theta) - \mathbf{C}\cdot\theta\Big) \\
= & -\Big(V(\theta^*) + \chi(-\theta^*) - \mathbf{C}\cdot\theta^*\Big)
\end{aligned} \tag{VIII.3.4}$$

D'autre part, en reprenant la fonction K^* introduite dans la proposition, on obtient la majoration, valable pour tout $\theta \in \mathbb{R}^d$:

$$V(\theta) \geq \theta\cdot\int f_*(X)dQ^{K^*,0,X_0} - H(Q^{K^*,0,X_0}\|P)$$

avec égalité pour $\theta = \theta^*$.
On en déduit :

$$V(\theta) - V(\theta^*) \geq (\theta-\theta^*)\cdot\int f_*(X)dQ^{K^*,0,X_0}$$

et donc $\int f_*(X)dQ^{K^*,0,X_0}$ appartient au sous-différentiel en θ^* de la fonction convexe V, c'est-à-dire, par hypothèse de différentiabilité :

$$\int f_*(X)dQ^{K^*,0,X_0} = V'(\theta^*) \tag{VIII.3.5}$$

Comme aussi $\theta \mapsto V(\theta) + \chi(-\theta) - \mathbf{C}\cdot\theta$ atteint son minimum en θ^*, en différentiant et en tenant compte de (VIII.3.5), on obtient l'égalité :

$$\chi'(-\theta^*) = \int f_*(X)dQ^{K^*,0,X_0} - \mathbf{C}$$

ce qui donne

$$\chi^*\Big(\int f_*(X)dQ^{K^*,0,X_0} - \mathbf{C}\Big) = -\theta^*\cdot\Big(\int f_*(X)dQ^{K^*,0,X_0} - \mathbf{C}\Big) - \chi(-\theta^*)$$

Avec cette relation ainsi que l'égalité (VIII.3.3), on obtient :

$$H(Q^{K^*,0,X_0}\|P) + \chi^*(\int f_*(X)dQ^{K^*,0,X_0} - \mathbf{C}) = -\Big(V(\theta^*) + \chi(-\theta^*) - \mathbf{C}\cdot\theta^*\Big)$$

et cela prouve que K^* est optimal d'après (VIII.3.4). ∎

Ainsi, à condition de vérifier les hypothèses de la proposition précédente, on voit que pour résoudre le problème $(P_{\chi^*}^{\overline{H},\mathbf{C}})$, il suffit d'attaquer le problème de minimisation duale :

$$(P_{\chi}^{*,\mathbf{C}}) \quad \text{trouver } \theta(\mathbf{C}) \in \mathbb{R}^d \text{ qui minimise } V(\theta) + \chi(-\theta) - \mathbf{C}\cdot\theta$$

L'énoncé suivant constitue alors un résultat de stabilité essentiel :

Lemme VIII.3.2 *Si $\mathbf{C}_0$ est un vecteur de $\mathbb{R}^d$ tel que la fonction valeur V est deux fois continuement différentiable au voisinage du point $\theta(\mathbf{C}_0)$, alors la solution $\theta(\mathbf{C})$ du problème de minimisation duale $(P_\chi^{*,\mathbf{C}})$ est continuement différentiable d'un voisinage ouvert de $\mathbf{C}_0$ vers un voisinage ouvert de $\theta(\mathbf{C}_0)$.*

Preuve : Par un argument de convexité, la condition de minimum de $\theta(\mathbf{C})$ où V est supposé différentiable s'exprime par la condition du premier ordre :

$$V'(\theta(\mathbf{C})) - \chi'(-\theta(\mathbf{C})) = \mathbf{C}$$

On en déduit le lemme par le théorème des fonctions implicites puisque la différentielle seconde $V''(\theta(\mathbf{C}_0)) + \chi''(-\theta(\mathbf{C}_0))$ définit un opérateur inversible car associé à une forme bilinéaire définie positive. ∎

L'étude de la stabilité sera poursuivie ultérieurement lorsque l'on examinera la variation des prix d'options à “pricer” avec celle du prix $\mathbf{C}$ des options de calibration.
La réalisation des hypothèses de la Proposition VIII.3.1 et du Lemme VIII.3.2 définit ainsi l'objectif du reste de ce chapitre.

VIII.4 Formulation HJB du problème

On se fixe désormais pour programme de résoudre le problème dual pénalisé de minimisation $(P_\chi^{*,\mathbf{C}})$ établi dans la section précédente.
Pour que cette approche soit valide, il s'agit d'abord d'étudier la fonction valeur afin de vérifier les hypothèses de la Proposition VIII.3.1 et du Lemme VIII.3.2 : cela permettra d'une part de revenir du problème dual au problème primal pénalisé $(P_{\chi^*}^{\overline{H},\mathbf{C}})$ et d'autre part d'obtenir la stabilité espérée.
Par une méthode de contrôle optimal, on commence à obtenir dans cette partie la fonction valeur ainsi que des conditions assurant l'existence de l'intensité optimale K^* , sous réserve que l'équation de la programmation dynamique associée admette une solution régulière : c'est le Théorème VIII.4.3.
Il restera ensuite à vérifier que la fonction valeur est de classe C^2 et que les conditions d'existence précédentes sont bien vérifiées, ce qui fait l'objet d'une autre partie.
Tout d'abord, avec (VIII.2.27), la fonction valeur réalise la maximisation de l'espérance du coût terminal :

$$\mathbb{E}^P\left(\theta \cdot f_*(X) Z_T^{K,0,X_0} - \ln(Z_T^{K,0,X_0}) Z_T^{K,0,X_0}\right)$$

où le processus de densité $Z_u^{K,t,x}$ est défini par la relation (VIII.2.24).
On se trouve ainsi dans le cadre d'un programme de contrôle optimal du processus markovien $(X_t, Z_t^H)_{0\leq t\leq T}$ et on se propose ici de l'aborder avec l'équation de la programmation dynamique associée, réduite au seul processus non contrôlé (X_t).
On fera par ailleurs l'hypothèse que le payoff des options de calibration est du type :

$$f_*(X) = \int_0^T \bar{l}(t, X_t)dt + h(X_T) \tag{VIII.4.1}$$

où la fonction de coût intégral $\bar{l}$ et celle de coût terminal h sont mesurables, à valeurs dans $\mathbb{R}^d$ et vérifient la condition de croissance exponentielle suivante :

$$(RS)\left\{\begin{array}{l}\text{Les fonctions } h \text{ et } \bar{l} \text{ vérifient la condition de croissance exponentielle :}\\ \|h(x)\| + \|\bar{l}(t,x)\| \leq \kappa(1+|x|+e^x)\\ \text{ avec } \kappa > 0\end{array}\right.$$

Le terme de croissance linéaire est en pratique inutile, mais il ne sera guère couteux à traiter.
Dans le cas réaliste où l'on cherche à retrouver le "smile" du marché, on ne peut adopter la forme (VIII.4.1) de $f_*(X)$ qu'au prix d'une régularisation permettant de remplacer par le coût intégral $\int_0^T \bar{l}(t,X_t)dt$ une suite de flux discrets $h_i(X_{T_i})$, $1 \leq i \leq p$ versés à p maturités $0 < T_1 < ... < T_p \leq T$ par les véritables options (Call ou Put de maturité T_i) : comme dans [113], pour le problème de calibration de la volatilité, il suffit de remplacer chaque option de calibration par la moyenne temporelle sur le payoff étendu temporellement autour de la maturité T_i.
Ainsi qu'on le verra ultérieurement, il n'est pas nécessaire d'adopter la forme délocalisée du payoff (VIII.4.1) pour pouvoir résoudre le problème de calibration.
La délocalisation présente essentiellement comme intérêt de permettre un contrôle sans les discontinuités qui apparaîtraient inévitablement aux maturités intermédiaires T_i en absence de délocalisation et qui n'auraient pas de signification financière.

VIII.4.1 Enoncé des résultats

Commençons par écrire l'EDS vérifiée par le processus X sous $Q^{K,0,X_0}$, avec $K \in \mathcal{K}^{\overline{H}}$:

$$\begin{aligned} dX_t &= (r_t - \frac{1}{2}\sigma_*^2(t,X_{t-}))dt + \sigma_*(t,X_{t-})dW_t^K \\ &+ \int_{z\in\mathbf{R}} \ln(1+\Phi_*(t,X_{t-},z))(\Pi - K(t,X_{t-},z)\pi)(dt,dz) \\ &- \int_{z\in\mathbf{R}} (\Phi_* - \ln(1+\Phi_*))(t,X_{t-},z)K(t,X_{t-},z)\pi(dt,dz), \quad 0 \leq t \leq T \end{aligned} \tag{VIII.4.2}$$

où W^K est un mouvement Brownien sous $Q^{K,0,S_0}$ le mouvement Brownien sous $Q^{K,0,S_0}$ obtenu par le changement de drift (VIII.2.6) dans le Lemme VIII.2.2, si bien que le générateur de la diffusion avec sauts (X_t), $\mathcal{G}^K$, est défini pour toute fonction W de classe C^2 sur $\mathbb{R}$ par :

$$\begin{aligned} (\mathcal{G}^K W)(t,x) &= (r_t - \frac{1}{2}\sigma_*^2(t,x))\frac{\partial W}{\partial x} + \frac{1}{2}\sigma_*^2(t,x)\frac{\partial^2 W}{\partial x^2} \\ &+ \int_{\mathbf{R}} \left(W(x + \ln(1+\Phi_*(t,x,z))) - W(x) - \Phi_*(t,x,z)\frac{\partial W}{\partial x}\right) K(t,x,z)\rho(dz) \end{aligned} \tag{VIII.4.3}$$

La définition du terme intégral ne pose pas de problème lorsque ρ est une mesure bornée tandis que l'on peut voir qu'elle est valide dans le cas général en écrivant

$$\begin{aligned} &\left| W(x+\ln(1+\Phi_*(t,x,z))) - W(x) - \Phi_*(t,x,z)\frac{\partial W}{\partial x}(x)\right| \leq \\ &(\ln(1+\Phi_*(t,x,z)))^2 \int_0^1 \left|\frac{\partial^2 W}{\partial x^2}\right|(x + s\ln(1+\Phi_*(t,x,z)))(1-s)ds \\ &+ |(\ln(1+\Phi_*(t,x,z)) - \Phi_*(t,x,z))| \left|\frac{\partial W}{\partial x}\right|(x) \end{aligned} \tag{VIII.4.4}$$

Puisque $\ln(1+\Phi_*)$ est borné avec en particulier la minoration (VIII.1.6), l'intégrale du membre de droite est une fonction bornée de z, pour (t,x) fixé.
En appliquant ensuite la formule de Taylor au premier et deuxième ordre en 0 à la fonction $x \mapsto \ln(1+x)$ et en tenant encore compte du fait que $\ln(1+\Phi_*)$ est borné, on peut finalement écrire :

$$\begin{aligned}&\left|W(x+\ln(1+\Phi_*(t,x,z))) - W(x) - \Phi_*(t,x,z)\frac{\partial W}{\partial x}(x)\right| \leq \\ &B\left[\left|\frac{\partial W}{\partial x}\right|(t,x) + \sup_{\ln(1+\Phi_1)\leq y\leq \ln(1+\|\Phi_0\|_\infty)}\left|\frac{\partial^2 W}{\partial x^2}\right|(t,x+y)\right]\Phi_0^2(z)\end{aligned} \tag{VIII.4.5}$$

où B est une constante : le membre de gauche de (VIII.4.5) est donc ρ-intégrable et (VIII.4.3) a un sens car $K()$ est bornée.
On peut énoncer en reprenant la solution $(X_u^{t,x})_{t\leq u\leq T}$ de l'EDS (VIII.2.4) avec la condition initiale $X_t^{t,x} = x$:

Lemme VIII.4.1 *Sous l'hypothèse (RS), on a la condition d'intégrabilité suivante*

$$\forall K \in \mathcal{K}^{\overline{H}},\ \mathbb{E}^{Q^{K,t,x}}\left[\|h\|(X_T^{t,x}) + \int_t^T \|\bar{l}\|(s, X_s^{t,x})ds\right] < +\infty \tag{VIII.4.6}$$

qui fournit en particulier l'égalité ensembliste

$$\mathcal{K}_0^{\overline{H}} = \mathcal{K}^{\overline{H}} \tag{VIII.4.7}$$

Puisque l'on a besoin de résultats intermédiaires, on reporte la preuve du lemme.
Comme conséquence du résultat précédent, la fonction valeur donnée par la relation (VIII.2.27) s'écrit, toujours sous l'hypothèse (RS) :

$$V(\theta) = \sup_{K\in\mathcal{K}^{\overline{H}}}\left\{\theta\cdot\mathbb{E}^{Q^{K,0,S_0}}\left[\int_0^T \bar{l}(s,X_{s-})ds + h(X_T)\right] - H(Q^{K,0,X_0}\|P)\right\} \tag{VIII.4.8}$$

et, par extension, on peut définir la fonction valeur $V^\theta(t,x)$ sur $[0,T]\times\mathbb{R}_+$ en posant

$$\forall (t,x)\in[0,T]\times\mathbb{R},\ V^\theta(t,x) = \sup_{K\in\mathcal{K}^{\overline{H}}}\left\{\theta\cdot\mathbb{E}^{Q^{K,t,x}}\left[\int_t^T \bar{l}(s,X_s^{t,x})ds + h(X_T^{t,x})\right] - H(Q^{K,t,x}\|P)\right\} \tag{VIII.4.9}$$

D'autre part, on peut également faire apparaître l'entropie relative comme une fonction de coût intégral :

Lemme VIII.4.2 *Avec les notations précédentes, on a :*

$$H(Q^{K,t,x}\|P) = \mathbb{E}^{Q^{K,t,x}}\left[\int_t^T \eta(K)(s,X_{s-}^{t,x})ds\right] \tag{VIII.4.10}$$

où l'on a posé pour tout $(s,y)\in[0,T]\times\mathbb{R}$

$$\begin{aligned}\eta(K)(s,y) &= \int_{\mathbf{R}}[K(s,y,z)(\ln(K(s,y,z))-1)+1]\rho(dz)\\ &+\frac{1}{2\sigma_*^2(s,y)}\left[b(s,y) - r_s + \int \Phi_*(s,y,z)(K(s,y,z)-1)\rho(dz)\right]^2\end{aligned} \tag{VIII.4.11}$$

Preuve : L'expression (VIII.2.24) du processus de densité $Z^{K,t,x}$ se réécrit sous la probabilité $Q^{K,t,x}$:

$$Z_u^{K,t,x} = e^{\int_t^u G_s^{K,t,x} d\tilde{W}_s + \frac{1}{2}\int_t^u (G_s^{K,t,x})^2 ds} e^{\int_{]t,u]\times\mathbf{R}} \ln(K(s,X_{s-}^{t,x},z))(\Pi - K(s,X_{s-}^{t,x},z)\pi)(ds,dz) + \int_{]t,u]\times\mathbf{R}} (K\ln(K)-K+1)(s,X_{s-}^{t,x},z)\pi(ds,dz)} \tag{VIII.4.12}$$

en posant également : $\tilde{W}_u = W_t - \int_t^u G_s^{K,t,x} ds$.
En compensant sous $Q^{K,t,x}$, on a :

$$\ln Z_u^{K,t,x} = \int_t^u \eta(K)(s, X_{s-}^{t,x}) ds + L_u$$

en définissant la $Q^{K,t,x}$-martingale locale

$$L_u = \int_t^u G_s^{K,t,x} d\tilde{W}_s + \int_{]t,u]\times\mathbf{R}} \ln K(s, X_{s-}^{t,x}, z)(\Pi(ds,dz) - K(s, X_{s-}^{t,x}, z)\pi(ds,dz))$$

qui est en fait une $Q^{K,t,x}$-martingale puisque $G^{K,t,x}$ est bornée et avec l'hypothèse sur $\ln(K)$. ∎

En reprenant le générateur infinitésimal de la diffusion défini par (VIII.4.3), on recherche à présent la fonction valeur comme solution continue sur $[0,T]\times\mathbb{R}$ et $C^{1,2}$ sur $[0,T[\times\mathbb{R}$ du système :

$$(HJB^\theta) \begin{cases} \frac{\partial W}{\partial t} + \sup_{K\in\mathcal{K}^{\overline{H}}} \Big(\mathcal{G}^K W - \eta(K)\Big) + \theta\cdot\bar{l} = 0 \text{ sur }]0,T[\times\mathbb{R} \\ W(T,x) = \theta\cdot h(x) \text{ sur } \mathbb{R} \end{cases}$$

Comme on l'a déjà annoncé, cette équation ne provient pas d'un résultat classique de contrôle, mais on va s'y ramener : suivant l'intuition, tout se passe comme si le processus contrôlé est le sous-jacent $(X_t)_{0\le t\le T}$.
La condition d'existence de la densité optimale nécessite quelques notations, apparaissant naturellement dans la preuve du prochain théorème : en considérant une solution W^θ fixée du système (HJB^θ), on pose

$$U^\theta(t,x,z) = W^\theta(t, x + \ln(1+\Phi_*(t,x,z))) - W^\theta(t,x) - \Phi_*(t,x,z)\frac{\partial W^\theta}{\partial x} \tag{VIII.4.13}$$

et l'on montre alors que l'optimum peut se rechercher implicitement sous la forme

$$K^\theta(t,x,z) - 1 = (J^*)'\left(U^\theta(t,x,z) - \lambda^{(\theta,t,x)}\frac{\Phi_*(t,x,z)}{\sigma_*(t,x)}\right) \tag{VIII.4.14}$$

avec

$$\lambda^{(\theta,t,x)} = \frac{1}{\sigma_*(t,x)}\left(\int (K^\theta(t,x,z)-1)\Phi_*(t,x,z)\rho(dz) + b_*(t,x) - r_t\right) \tag{VIII.4.15}$$

et

$$(J^*)'(u) = \begin{cases} e^u - 1 & \text{si } |u| \le \ln(\overline{H}) \\ \overline{H}^{-1} - 1 & \text{si } u < -\ln(\overline{H}) \\ \overline{H} - 1 & \text{si } \ln(\overline{H}) < u \end{cases} \tag{VIII.4.16}$$

Théorème VIII.4.3 *Faisons l'hypothèse (RS) sur les fonctions de payoff et supposons aussi que le système (HJB^θ) admette une solution W^θ continue sur $[0,T] \times \mathbb{R}$, de classe $C^{1,2}$ sur $[0,T[\times\mathbb{R}$, satisfaisant de plus la condition de croissance exponentielle*

$$|W^\theta(t,x)| \leq \gamma(1+|x|+e^x) \tag{VIII.4.17}$$

où γ est une constante positive, indépendante de $t \in [0,T]$.
Alors le système d'équations (VIII.4.14) et (VIII.4.15) admet une solution unique K^θ et c'est une fonction Borélienne sur $[0,T] \times \mathbb{R} \times \mathbf{R}$ telle que $|\ln(K)| \leq \ln(\overline{H})$.
Dès que l'on suppose en plus que $K^\theta \in \mathcal{K}^{\overline{H}}$, c'est-à-dire $|\ln(K)| \leq \ln(\overline{H})\psi_0$, alors

$$\forall(t,x) \in [0,T]\times\mathbb{R},\ V^\theta(t,x) = W^\theta(t,x) = \theta\cdot\mathbb{E}^{Q^{K^\theta,t,x}}\left[\int_t^T \bar{l}(s,X_s^{t,x})ds+h(X_T^{t,x})\right]-H(Q^{K^\theta,t,x}\|P)$$

En particulier, à l'origine, on a

$$V(\theta) = W^\theta(0,X_0) = \theta\cdot\mathbb{E}^{Q^{K^\theta,0,X_0}}\left[\int_0^T \bar{l}(s,X_s)ds + h(X_T)\right] - H(Q^{K^\theta,0,X_0}\|P)$$

où $V(\theta)$ est défini par (VIII.2.27) et K^θ est l'optimum.

Notons que dans le cas où ρ est bornée, on peut prendre $\forall z, \psi_0(z) = 1$ et donc $K^\theta \in \mathcal{K}^{\overline{H}}$ dès que $|\ln(K)| \leq \ln(\overline{H})$. On y reviendra ultérieurement.
La section suivante est consacrée à la preuve du résultat précédent.

VIII.4.2 Preuve des résultats de l'étude précédente

On va établir ici la preuve du Théorème VIII.4.3 ainsi que du Lemme VIII.4.1.
On va d'abord justifier l'hypothèse du théorème concernant le comportement de la solution en énonçant deux majorations utiles dans la suite, obtenues par une méthode classique :

Lemme VIII.4.4 *Il existe une constante C telle que*

$$\forall(t,x) \in [0,T] \times \mathbb{R},\ \mathbb{E}^P\left[\sup_{t\leq u\leq T}|X_u^{t,x}|^2\right] \leq C(1+x^2) \tag{VIII.4.18}$$

et

$$\forall(t,x) \in [0,T] \times \mathbb{R},\ \mathbb{E}^P\left[\sup_{t\leq u\leq T} e^{2X_u^{t,x}}\right] \leq Ce^{2x} \tag{VIII.4.19}$$

Preuve : Commençons à prouver la deuxième inégalité en posant, pour revenir exceptionnellement à la diffusion d'origine :

$$S_u^{x,t} = e^{X_u^{x,t}}$$

qui est la solution de l'EDS (VIII.1.1) avec pour condition initiale :

$$S_t = e^x$$

On a donc :

$$S_u^{x,t} = e^x+\int_t^u S_{s-}^{t,x}b(s,S_{s-}^{t,x})ds+\int_t^u S_{s-}^{t,x}\sigma(s,S_{s-}^{t,x})dW_s+\int_{]t,u]\times\mathbf{R}} S_{s-}^{t,x}\Phi(s,S_{s-}^{t,x},z)(\Pi(ds,dz)-\pi(ds,dz)) \tag{VIII.4.20}$$

On peut adapter le Problème V 3.15 de [76] par exemple en utilisant l'inégalité :

$$(|a_1| + \ldots + |a_n|)^p \leq n^{p-1}(|a_1|^p + \ldots + |a_n|^p)$$

ce qui donne :

$$\begin{aligned}
\sup_{t\leq u\leq v} (S_u^{x,t})^2 &\leq 4\Bigg\{e^{2x} + \sup_{t\leq u\leq v}\left[\int_t^u S_{s-}^{t,x} b(s, S_{s-}^{t,x})ds\right]^2 + \sup_{t\leq u\leq v}\left[\int_t^u S_{s-}^{t,x}\sigma(s, S_{s-}^{t,x})dW_s\right]^2 \\
&\quad + \sup_{t\leq u\leq v}\left[\int_{]t,u]\times\mathbf{R}} S_{s-}^{t,x}\Phi(s, S_{s-}^{t,x}, z)(\Pi(ds, dz) - \pi(ds, dz))\right]^2\Bigg\} \\
&\leq 4\Bigg\{e^{2x} + (u-t)\int_t^u (S_{s-}^{t,x})^2 b^2(s, S_{s-}^{t,x})ds + \sup_{t\leq u\leq v}\left[\int_t^u S_{s-}^{t,x}\sigma(s, S_{s-}^{t,x})dW_s\right]^2 \\
&\quad + \sup_{t\leq u\leq v}\left[\int_{]t,u]\times\mathbf{R}} S_{s-}^{t,x}\Phi(s, S_{s-}^{t,x}, z)(\Pi(ds, dz) - \pi(ds, dz))\right]^2\Bigg\}
\end{aligned}$$

On localise ensuite en introduisant les temps d'arrêt

$$\sigma_n = \inf\{s \geq t \mid |S_s^{t,x}| \geq n\}$$

Alors, en prenant l'espérance et en utilisant l'inégalité de Doob , on obtient :

$$\begin{aligned}
\mathbb{E}^P\left[\sup_{t\leq u\leq\sigma_n\wedge v} (S_u^{x,t})^2\right] &\leq C_0\Bigg\{e^{2x} + (v-t)\mathbb{E}^P\left[\int_t^{\sigma_n\wedge v} b^2(s, S_{s-}^{t,x})(S_{s-}^{t,x})^2 ds\right] \\
&\quad + \mathbb{E}^P\left[\int_t^{\sigma_n\wedge v} \sigma^2(s, S_{s-}^{t,x})(S_{s-}^{t,x})^2 ds\right] \\
&\quad + \mathbb{E}^P\left[\int_{]t,\sigma_n\wedge v]\times\mathbf{R}} \Phi^2(s, S_{s-}^{t,x}, z)(S_{s-}^{t,x})^2\pi(ds, dz)\right]\Bigg\}
\end{aligned}$$

où C_0 est une constante.
Puisque les coefficients sont bornés et avec les hypothèses sur la mesure ρ et sur Φ, on en déduit

$$\forall v \in [t, T],\ \mathbb{E}^P\left[\sup_{t\leq u\leq\sigma_n\wedge v} (S_u^{x,t})^2\right] \leq C(T)\left\{e^{2x} + \int_t^v \mathbb{E}^P\left[\sup_{t\leq u\leq\sigma_n\wedge s} (S_u^{t,x})^2\right]ds\right\} \tag{VIII.4.21}$$

où C(T) est une nouvelle constante (qui dépend de T).
Mais, par définition de σ_n, le terme sous l'espérance du premier membre de l'inégalité précédente est majoré par $n^2(1+M)^2$, où M est un majorant de $|\Phi|$. Le premier membre de (VIII.4.21) est donc fini. Alors, le Lemme de Gronwall donne

$$\mathbb{E}^P\left[\sup_{t\leq u\leq\sigma_n\wedge v} (S_u^{x,t})^2\right] \leq C(T)e^{2x}$$

En prenant $v = T$, on conclut avec le lemme de Fatou en faisant tendre n vers l'infini.
L'inégalité (VIII.4.18) s'obtient suivant les mêmes lignes. ∎

On énonce alors deux majorations fondamentales :

Corollaire VIII.4.5 *Pour $K \in \mathcal{K}^{\overline{H}}$ fixé, on a*

$$\mathbb{E}^{Q^{K,t,x}}\left[\sup_{t\leq s\leq T} |X_s^{t,x}|\right] \leq C(1+|x|) \tag{VIII.4.22}$$

$$\mathbb{E}^{Q^{K,t,x}}\left[\sup_{t\leq s\leq T} e^{X_s^{t,x}}\right] \leq Ce^x \tag{VIII.4.23}$$

où C est une constante indépendante de K, t, x.

Preuve : Les deux inégalités se traitent de manière identique, en écrivant l'inégalité de Cauchy-Schwarz pour établir (VIII.4.22) par exemple :

$$\mathbb{E}^{Q^{K,t,x}}\left[\sup_{t\leq s\leq T}|X_s^{t,x}|\right] \leq \left[\mathbb{E}^P(Z_T^{K,t,x})^2\right]^{\frac{1}{2}}\left\{\mathbb{E}^P\left[\sup_{t\leq s\leq T}(X_s^{t,x})^2\right]\right\}^{\frac{1}{2}} \tag{VIII.4.24}$$

Avec le lemme précédent, on obtient le résultat à condition de majorer $\mathbb{E}^P(Z_T^{K,t,x})^2$ uniformément par rapport à $K \in \mathcal{K}^{\overline{H}}, t, x$, ce qui est un résultat du type du Lemme I 2.1 de [87].
On peut en fait le montrer si l'on reprend la relation (VIII.2.15) en la précisant comme une majoration uniforme, ce qu'exprime exactement le Lemme VIII.2.4 puisque la martingale M_t définie par (VIII.2.14) est quasi-continue à gauche, que les sauts ainsi que la quantité $\frac{<M>_t}{t}$ sont bornés uniformément. ∎

Donnons alors la preuve d'un lemme jusque là admis :
Preuve du Lemme VIII.4.1 : Avec l'hypothèse (RS) de croissance exponentielle des fonctions de payoff, le résultat précédent permet d'obtenir la condition d'intégrabilité (VIII.4.6). ∎

Le comportement de la fonction valeur s'obtient de même avec le résultat suivant
Lemme VIII.4.6 *Sous l'hypothèse (RS) sur les fonctions de payoff, la fonction valeur, pour θ fixé, est à croissance exponentielle :*

$$|V^\theta(t,x)| \leq \gamma^\theta(1+|x|+e^x) \tag{VIII.4.25}$$

où γ^θ est une constante positive.

Preuve : On reprend l'expression (VIII.4.9) de la fonction valeur :

$$\forall (t,x)\in[0,T]\times\mathbb{R},\ V^\theta(t,x) = \sup_{K\in\mathcal{K}^{\overline{H}}}\left\{\theta\cdot\mathbb{E}^{Q^{K,t,x}}\left[\int_t^T \bar{l}(s,X_s^{t,x})ds + h(X_T^{t,x})\right] - H(Q^{K,t,x}\|P)\right\} \tag{VIII.4.26}$$

et le lemme est alors immédiat avec le Corollaire VIII.4.5 ainsi que la condition (RS) de croissance des payoff et puisqu'il est facile de majorer $H(Q^{K,t,x}\|P)$ uniformément par rapport à $K \in \mathcal{K}^{\overline{H}}, t, x$ (car il en est de même de $\eta(K)(t,x)$ donné par (VIII.4.11)). ∎

Passons à présent à la preuve du théorème principal :
Preuve du Théorème VIII.4.3 : Il s'agit essentiellement d'adapter le Théorème de vérification du controle optimal en horizon fini (Théorème III 8.1 de [50] par exemple).
En se fixant $K \in \mathcal{K}^{\overline{H}}$ et compte tenu de la condition de croissance exponentielle de la solution, on a :

$$\mathbb{E}^{Q^{K,t,x}}(W^\theta(T,X_T^{t,x})) - W^\theta(t,x) + \theta\cdot\mathbb{E}^{Q^{K,t,x}}\left[\int_t^T \bar{l}(s,X_s^{t,x})ds\right] - H(Q^{K,t,x}\|P) \leq 0 \tag{VIII.4.27}$$

Ce résultat est établi précisémment dans le Lemme VIII.4.8.
Avec la condition terminale, cela se réécrit :

$$W^\theta(t,x) \geq \theta\cdot\mathbb{E}^{Q^{K,t,x}}\left[\int_t^T \bar{l}(s,X_s^{t,x})ds + h(X_T^{t,x})\right] - H(Q^{K,t,x}\|P)$$

c'est-à-dire

$$W^\theta(t,x) \geq V^\theta(t,x)$$

Admettons ensuite qu'il existe $K^\theta \in \mathcal{K}^{\overline{H}}$ tel que pour tout $(t,x) \in [0,T] \times \mathbb{R}$:

$$K^\theta(t,x,.) = \arg\max_{K \in \mathcal{K}^{\overline{H}}} \Big(\mathcal{G}^K W^\theta - \eta(K)\Big)(t,x)$$

Dans ce cas, en prenant $K = K^\theta$ dans les relations précédentes, celles-ci sont vérifiées avec égalité, ce qui prouve la conclusion du théorème.
Il reste donc à prouver l'existence de K^θ, ce qui revient à résoudre le problème suivant :

$$(\mathcal{E}) \left\{ \begin{array}{l} \text{trouver une fonction } K^\theta \in \mathcal{K}^{\overline{H}} \text{ telle que} \\ \forall (t,x) \in [0,T] \times \mathbb{R},\ K^\theta \in \arg\max_{K \in \mathcal{K}^{\overline{H}}} \Big(\int_{\mathbf{R}} U^\theta(t,x,z) K(t,x,z) \rho(dz) - \eta(K)(t,x) \Big) \end{array} \right.$$

en rappelant la notation (VIII.4.13) :

$$U^\theta(t,x,z) \quad = \quad W^\theta(t, x + \ln(1+\Phi_*(t,x,z))) - W^\theta(t,x) - \Phi_*(t,x,z)\frac{\partial W^\theta}{\partial x} \tag{VIII.4.28}$$

Le problème serait presque résolu si l'on obtenait pour tout $(t,x) \in [0,T] \times \mathbb{R}$ l'existence, en tant que classe de fonctions, de $z \mapsto K^\theta(t,x,z) \in \mathcal{B}_0^{\overline{H}} = \Big\{ k(.)$ fonction positive ρ-mesurable telle que $\forall z \in \mathbf{R},\ |\ln(k(z))| \leq \ln(\overline{H})\psi_0(z) \Big\}$ réalisant :

$$K^\theta(t,x,.) = \arg\max_{k \in \mathcal{B}_0^{\overline{H}}} \left(\int_{\mathbf{R}} U^\theta(t,x,z) k(z) \rho(dz) - \eta(k)(t,x) \right) \tag{VIII.4.29}$$

l'unicité de la solution étant une conséquence de la stricte convexité de la fonctionnelle η.
En réalité, on ne résoudra pas directement ce problème car la majoration dans la définition de $\mathcal{B}_0^{\overline{H}}$ n'est pas gratuite. En notant que $k - 1 \in L^2(\mathbf{R},\rho)$ si $k \in \mathcal{B}_0^{\overline{H}}$ (avec une majoration du type (VIII.2.9)), on est conduit à faire le changement de variable correspondant et à résoudre le problème "élargi" :

$$K^\theta(t,x,.) - 1 = \arg\max_{k \in L^2(\mathbf{R},\rho)} \left(\int_{\mathbf{R}} U^\theta(t,x,z) k(z) \rho(dz) - \eta_0(k)(t,x) \right) \tag{VIII.4.30}$$

avec $\forall k \in L^2(\mathbf{R},\rho),\ \eta_0(k) = \eta(k+1)$.
Pour que (VIII.4.29) soit réalisée, il suffira alors de vérifier que $K^\theta(t,x,.)$ appartient à $\mathcal{B}_0^{\overline{H}}$.
On vérifie d'abord que $U^\theta \in L^2(\mathbf{R},\rho)$, en utilisant une majoration du type (VIII.4.5) (avec également le fait que Φ_0 est bornée), si bien que la première intégrale du second membre de (VIII.4.30) est bien définie, tandis que le Lemme VIII.4.2 permet d'écrire η_0 sous la forme :

$$\eta_0(k)(t,x) = \int J(k(z))\rho(dz) + \frac{1}{2}\left[\alpha(t,x) + \int \beta(t,x,z)k(z)\rho(dz)\right]^2 \Bigg\} \tag{VIII.4.31}$$

où l'on a défini la fonction J sur $\mathbb{R}$ par :

$$J(u) = \left\{ \begin{array}{ll} (u+1)\ln(u+1) - (u+1) + 1 & \text{si } \overline{H}^{-1} \leq u+1 \leq \overline{H} \\ +\infty & \text{sinon} \end{array} \right. \tag{VIII.4.32}$$

et l'on a également noté :

$$\alpha(t,x) = \frac{b_*(t,x) - r_t}{\sigma_*(t,x)} \tag{VIII.4.33}$$

et

$$\beta(t,x,z) = \frac{\Phi_*(t,x,z)}{\sigma_*(t,x)} \tag{VIII.4.34}$$

En tenant compte du fait que $\beta \in L^2(\mathbf{R},\rho)$ grâce à (VIII.1.4), et puisque J est positive, on voit donc que $\eta_0(.)(t,x)$ est définie pour $k \in L^2(\mathbf{R},\rho)$, ce qui achève de donner un sens au problème d'optimisation associé à (VIII.4.30).
En fixant toujours $(t,x) \in [0,T] \times \mathbb{R}$, on considère le produit scalaire naturel de l'espace $L^2(\mathbf{R},\rho)$, et le problème fait apparaître la fonction convexe conjuguée $\eta_0^*(.)(t,x)$ de $\eta_0(.)(t,x)$, ce qui a un sens puisque l'examen de l'expression (VIII.4.31) montre que $\eta_0(.)(t,x)$ est une fonctionnelle strictement convexe en tant que somme d'une intégrale d'une fonction strictement convexe et du carré d'une forme affine.
Elle a la forme d'une entropie pénalisée, ce qui permet d'appliquer les techniques de dualité utilisées dans ce cas et sa conjuguée prise en $z \in \mathbf{R} \mapsto U^\theta(t,x,z)$, se définit alors :

$$\begin{aligned}
\eta_0^*(U^\theta)(t,x) &= \sup_{k \in L^2(\mathbf{R},\rho)} \left[\int U^\theta(t,x,z)k(z)\rho(dz) - \eta_0(k)(t,x) \right] \\
&= \sup_{k \in L^2(\mathbf{R},\rho)} \left\{ \int U^\theta(t,x,z)k(z)\rho(dz) \right. \\
&\quad \left. - \int J(k(z))\rho(dz) - \frac{1}{2}\left[\alpha(t,x) + \int \beta(t,x,z)k(z)\rho(dz)\right]^2 \right\}
\end{aligned} \tag{VIII.4.35}$$

Notons pour $k \in L^2(\mathbf{R},\rho)$

$$L(k,\lambda) = \int U^\theta(t,x,z)k(z)\rho(dz) - \int J(k(z))\rho(dz) - \lambda\left[\alpha(t,x) + \int \beta(t,x,z)k(z)\rho(dz)\right] + \frac{1}{2}\lambda^2$$

Alors, en utilisant la relation :

$$\forall \lambda \in \mathbb{R},\ \frac{1}{2}a^2 \geq \lambda a - \frac{1}{2}\lambda^2 \tag{VIII.4.36}$$

avec égalité si et seulement si

$$\lambda = a$$

on a, avec (VIII.4.35) :

$$\begin{aligned}
\eta_0^*(U^\theta)(t,x) &= \sup_{k \in L^2(\mathbf{R},\rho)} \inf_{\lambda \in \mathbb{R}} L(k,\lambda) \\
&\leq \inf_{\lambda \in \mathbb{R}} \sup_{k \in L^2(\mathbf{R},\rho)} L(k,\lambda)
\end{aligned} \tag{VIII.4.37}$$

et l'on examine pour $\lambda \in \mathbb{R}$

$$\begin{aligned}
\sup_{k \in L^2(\mathbf{R},\rho)} L(k,\lambda) &= \sup_{k \in L^2(\mathbf{R},\rho)} \left[\int (U^\theta(t,x,z) - \lambda\beta(t,x,z))k(z)\rho(dz) - \int J(k(z))\rho(dz) \right] \\
&\quad -\lambda\alpha(t,x) + \frac{1}{2}\lambda^2
\end{aligned} \tag{VIII.4.38}$$

Avec le Théorème 2 ainsi que le Lemme 1 de [104], on exprime la fonction conjuguée apparaissant comme premier terme du second membre :

$$\begin{aligned}\sup_{k\in L^2(\mathbf{R},\rho)}\Bigg[\int(U^\theta(t,x,z)-\lambda\beta(t,x,z))k(z)\rho(dz)&\\ -\int J(k(z))\rho(dz)\Bigg] &= \int J^*(U^\theta(t,x,z)-\lambda\beta(t,x,z))\rho(dz)\end{aligned}\tag{VIII.4.39}$$

où J^* est la conjuguée de J définie par :

$$\forall u\in\mathbb{R},\ \ J^*(u)=\sup_{v\in\mathbb{R}}(uv-J(v))\tag{VIII.4.40}$$

Les hypothèses du corollaire cité sont en effet satisfaites puisque J est une fonction convexe propre semi continue inférieurement et que l'espace $L^2(\mathbf{R},\rho)$ est "décomposable" (selon le sens donné dans la section 3 b de [104]).
Sans faire appel à ces résultats, on pourrait cependant se contenter de majorer le membre de gauche de (VIII.4.39) par celui de droite.
On détermine à partir de (VIII.4.32) :

$$J^*(u)=\begin{cases} e^u-1-u & \text{si } |u|\leq \ln(\overline{H})\\ (\overline{H}^{-1}-1)u-J(\overline{H}^{-1}-1) & \text{si } u<-\ln(\overline{H})\\ (\overline{H}-1)u-J(\overline{H}-1) & \text{si } \ln(\overline{H})<u\end{cases}\tag{VIII.4.41}$$

ce qui montre que J^* est continuement dérivable :

$$(J^*)'(u)=\begin{cases} e^u-1 & \text{si } |u|\leq \ln(\overline{H})\\ \overline{H}^{-1}-1 & \text{si } u<-\ln(\overline{H})\\ \overline{H}-1 & \text{si } \ln(\overline{H})<u\end{cases}\tag{VIII.4.42}$$

Il est alors aisé d'établir avec la définition (VIII.4.40) :

$$\forall u,v\in\mathbb{R},\ \ J(v)\geq uv-J^*(u)\tag{VIII.4.43}$$

avec égalité si et seulement si :

$$v=(J^*)'(u)$$

En particulier, on vérifie la dérivabilité de la fonction de $\lambda\in\mathbb{R}$

$$\begin{aligned}\sup_{k\in L^2(\mathbf{R},\rho)} L(k,\lambda) &= \int J^*(U^\theta(t,x,z)-\lambda\beta(t,x,z))\rho(dz)\\ &\quad -\lambda\alpha(t,x)+\frac{1}{2}\lambda^2\end{aligned}\tag{VIII.4.44}$$

Montrons en effet que les hypothèses du théorème de dérivation sont remplies : si l'on considère un intervalle ouvert borné Λ de $\mathbb{R}$, alors en reprenant une majoration du type (VIII.4.4) et compte tenu de (VIII.1.4) et du fait que Λ est un ensemble borné, on obtient la majoration :

$$|U^\theta(t,x,z)-\lambda\beta(t,x,z)|\leq M(t,x)(\Phi_0^2(z)+\Phi_0(z))\leq M(t,x)(1+\|\Phi_0\|_\infty)\Phi_0(z)\tag{VIII.4.45}$$

ou $M(t,x)$ ne dépend ni de $z\in\mathbf{R}$ ni de $\lambda\in\Lambda$.
D'autre part, l'inégalité des accroissements finis appliquée à $u\mapsto e^u$ dans (VIII.4.42) donne pour $|u|\leq\ln(\overline{H})$:

$$|(J^*)'|(u)\leq\overline{H}|u|\tag{VIII.4.46}$$

Cette inégalité, vraie également pour $|u| > \ln(\overline{H})$, permet d'écrire en vertu de (VIII.4.45)

$$|(J^*)'(U^\theta(t,x,z) - \lambda\beta(t,x,z))|\beta(t,x,z) \leq M_0(t,x)\Phi_0(z)\beta(t,x,z) \tag{VIII.4.47}$$

ou $M_0(t,x)$ ne dépend ni de $z \in \mathbf{R}$ ni de $\lambda \in \Lambda$: le membre de droite étant ρ-intégrable, il est donc loisible de dériver l'intégrale dans (VIII.4.44).
Cette dernière expression définit donc une fonction dérivable et de plus strictement convexe en tant que somme de trois fonctions convexes dont la dernière l'est strictement.
Comme le premier terme est du signe positif de J^* (car cette dernière est positive sur $[-\ln(\overline{H}, \ln(\overline{H}]$ et que ses dérivées en $-\ln(\overline{H})$ et $\ln(\overline{H})$ sont respectivement négative et positive car on a choisi $\overline{H} > 1$), on obtient donc le fait que $\sup_{k\in L^2(\mathbf{R},\rho)} L(k,\lambda)$ est une fonction qui tend vers $+\infty$ à l'infini : ainsi, en tant que fonction strictement convexe et dérivable, elle admet un minimum unique $\lambda^{(\theta,t,x)}$ où sa dérivée s'annule :

$$\lambda^{(\theta,t,x)} = \int (K^\theta(t,x,z) - 1)\beta(t,x,z)\rho(dz) + \alpha(t,x) \tag{VIII.4.48}$$

où l'on a posé :

$$K^\theta(t,x,z) - 1 = (J^*)'(U^\theta(t,x,z) - \lambda^{(\theta,t,x)}\beta(t,x,z)) \tag{VIII.4.49}$$

Avec cette expression, on peut appliquer le cas d'égalité de (VIII.4.43) dans le deuxième membre de la relation (VIII.4.44) évaluée en $\lambda = \lambda^{(\theta,t,x)}$, en notant également que $K^\theta(t,x,.) - 1 \in L^2(\mathbf{R},\rho)$.
Ainsi, avec (VIII.4.38), on trouve :

$$\begin{aligned}
\inf_{\lambda\in\mathbb{R}} \sup_{k\in L^2(\mathbf{R},\rho)} L(k,\lambda) &= \sup_{k\in L^2(\mathbf{R},\rho)} L(k,\lambda^{(\theta,t,x)}) \\
&= \int J^*(U^\theta(t,x,z) - \lambda^{(\theta,t,x)}\beta(t,x,z))\rho(dz) \\
&\quad -\lambda^{(\theta,t,x)}\alpha(t,x) + \frac{1}{2}(\lambda^{(\theta,t,x)})^2 \qquad \text{(VIII.4.50)} \\
&= \int (U^\theta(t,x,z) - \lambda^{(\theta,t,x)}\beta(t,x,z))(K^\theta(t,x,z) - 1)\rho(dz) \\
&\quad - \int J(K^\theta(t,x,z) - 1)\rho(dz) - \lambda^{(\theta,t,x)}\alpha(t,x) + \frac{1}{2}(\lambda^{(\theta,t,x)})^2 \\
&= \int U^\theta(t,x,z)(K^\theta(t,x,z) - 1)\rho(dz) - \int J(K^\theta(t,x,z) - 1)\rho(dz) \\
&\quad -\lambda^{(\theta,t,x)}\left[\alpha(t,x) + \int \beta(t,x,z)(K^\theta(t,x,z) - 1)\rho(dz)\right] + \frac{1}{2}(\lambda^{(\theta,t,x)})^2 \\
&= \int U^\theta(t,x,z)(K^\theta(t,x,z) - 1)\rho(dz) - \int J(K^\theta(t,x,z) - 1)\rho(dz) \\
&\quad -\frac{1}{2}\left[\alpha(t,x) + \int \beta(t,x,z)(K^\theta(t,x,z) - 1)\rho(dz)\right]^2
\end{aligned}$$

puisque la relation (VIII.4.48) permet de réaliser le cas d'égalité de (VIII.4.36).
Avec la dernière égalité, on obtient

$$\begin{aligned}\inf_{\lambda\in\mathbb{R}}\sup_{k\in L^2(\mathbf{R},\rho)} L(k,\lambda) &= \int U^\theta(t,x,z)(K^\theta(t,x,z)-1)\rho(dz)\\ &\quad -\eta_0(K^\theta-1)(t,x) \qquad \text{(VIII.4.51)}\\ &\leq \sup_{k\in L^2(\mathbf{R},\rho)}\left(\int U^\theta(t,x,z)k(z)\rho(dz)-\eta_0(k)\right) \qquad \text{(VIII.4.52)}\\ &= \eta_0^*(U^\theta)(t,x)\\ &\leq \inf_{\lambda\in\mathbb{R}}\sup_{k\in L^2(\mathbf{R},\rho)} L(k,\lambda)\end{aligned}$$

en réécrivant en dernier l'inégalité (VIII.4.37).
Les inégalités précédentes sont ainsi des égalités, ce qui prouve en particulier, avec (VIII.4.51) et (VIII.4.52) :

$$\begin{aligned}\sup_{k\in L^2(\mathbf{R},\rho)}\left(\int U^\theta(t,x,z)k(z)\rho(dz)-\eta_0(k)(t,x)\right) &= \int U^\theta(t,x,z)(K^\theta(t,x,z)-1)\rho(dz)\\ &\quad -\eta_0(K^\theta-1)(t,x) \qquad \text{(VIII.4.53)}\end{aligned}$$

Si l'on prouve que $\lambda^{(\theta,t,x)}$ est une fonction mesurable de (t,x), la relation (VIII.4.49) montre que la fonction $(t,x,z)\mapsto K^\theta(t,x,z)$ est mesurable.
La preuve de la mesurabilité de $(t,x)\mapsto\lambda^{(\theta,t,x)}$ est reportée au lemme VIII.4.9.
De plus, avec (VIII.4.42), on a $|\ln(K^\theta(t,x,z))|\leq\ln(\overline{H})$ et l'on en déduit la première partie du théorème.
Enfin, à condition d'avoir $K^\theta\in\mathcal{K}^{\overline{H}}$, on obtient l'optimum cherché et c'est la deuxième assertion du théorème. ∎

Remarque VIII.4.7 *Si l'on suppose que la mesure de Lévy ρ est bornée, l'existence de la solution au problème d'optimisation (VIII.4.30) s'obtient en utilisant des principes généraux.*
En fixant $(t,x)\in[0,T]\times\mathbb{R}$, il s'agit en effet d'optimiser sur $L^2(\mathbf{R},\rho)$ la fonctionnelle ξ définie par :

$$\xi(k) = \int_{\mathbf{R}} U^\theta(t,x,z)k(z)\rho(dz)-\eta_0(k)(t,x)$$

Il suffit en fait de résoudre le problème précédent sur

$$\mathcal{B}^{\overline{H}} = \left\{k\in L^2(\mathbf{R},\rho)\ \textit{telle que}\ \overline{H}^{-1}-1\leq k\leq\overline{H}-1\right\}$$

C'est un ensemble fortement borné de $L^2(\mathbf{R},\rho)$ puisque l'on a ici supposé que ρ est finie et c'est également un ensemble fortement fermé puisque de toute suite fortement convergente on peut extraire une sous-suite qui converge ρ-presque sûrement.
Puisque l'on peut montrer que la fonctionnelle strictement concave ξ est semi-continue supérieurement (et même continue) pour la topologie forte de $L^2(\mathbf{R},\rho)$ et puisque $L^2(\mathbf{R},\rho)$ est réflexif, on déduit du Corollaire III 20 de [29] que ξ admet bien un maximum $K^\theta(t,x,.)-1$.
La continuité forte de ξ est immédiate en raison du fait que $U^\theta(t,x,.)$ et $\beta(t,x,.)$ appartiennent à $L^2(\mathbf{R},\rho)$ et en raison également du fait que J est dérivable, à dérivée bornée sur $[\overline{H}^{-1}-1,\overline{H}-1]$.

Pour achever la preuve du Théorème VIII.4.3, il faut d'abord prouver la relation (VIII.4.27).

Lemme VIII.4.8 *Avec l'hypothèse (VIII.4.17) de croissance exponentielle de W^θ, la relation (VIII.4.27) est vérifiée.*

Preuve : On considère $0 \leq t \leq \tilde{T} < T$ et l'on définit pour tout entier positif n le temps d'arrêt

$$\sigma_n = \inf\{t \leq s \leq \tilde{T} \ ; \ |X_s^{t,x}| \geq n\}$$

pour obtenir par application de la formule d'Itô, puisque W^θ est de classe $C^{1,2}$ sur $[0,T[\times\mathbb{R}$:

$$\begin{aligned}
\mathbb{E}^{Q^{K,t,x}}\left(W^\theta(\tilde{T}\wedge\sigma_n, X^{t,x}_{\tilde{T}\wedge\sigma_n})\right) - W^\theta(t,x) &= \mathbb{E}^{Q^{K,t,x}}\left[\int_t^{\tilde{T}\wedge\sigma_n}\left(\frac{\partial W^\theta}{\partial s} + \mathcal{G}^K W^\theta(s,X_s^{t,x})\right)ds\right] \\
&\leq \mathbb{E}^{Q^{K,t,x}}\left[\int_t^{\tilde{T}\wedge\sigma_n}\left(\eta(K)(s,X_s^{t,x}) - \theta\cdot\bar{l}(s,X_s^{t,x})\right)ds\right]
\end{aligned}$$

où l'on a utilisé :

$$\frac{\partial W^\theta}{\partial t} + \mathcal{G}^K W^\theta - \eta(K) + \theta\cdot\bar{l} \leq 0$$

On a donc

$$W^\theta(t,x) \geq \mathbb{E}^{Q^{K,t,x}}\left(W^\theta(\tilde{T}\wedge\sigma_n, X^{t,x}_{\tilde{T}\wedge\sigma_n})\right) - \mathbb{E}^{Q^{K,t,x}}\left[\int_t^{\tilde{T}\wedge\sigma_n}\left(\eta(K)(s,X_s^{t,x}) - \theta\cdot\bar{l}(s,X_s^{t,x})\right)ds\right] \tag{VIII.4.54}$$

Pour traiter le premier terme du membre de droite de (VIII.4.54), on utilise la croissance exponentielle de W^θ :

$$\mathbb{E}^{Q^{K,t,x}}\left[\sup_{t\leq s\leq T}|W^\theta|(s,X_s^{t,x})\right] \leq \gamma\mathbb{E}^{Q^{K,t,x}}\left[1 + \sup_{t\leq s\leq T}|X_s^{t,x}| + \sup_{t\leq s\leq T} e^{X_s^{t,x}}\right] < +\infty \tag{VIII.4.55}$$

d'après le Corollaire VIII.4.5, ce qui justifie que par convergence dominée on a :

$$\mathbb{E}^{Q^{K,t,x}}\left(W^\theta(\tilde{T}\wedge\sigma_n, X^{t,x}_{\tilde{T}\wedge\sigma_n})\right) \longrightarrow_{n\to\infty} \mathbb{E}^{Q^{K,t,x}}\left(W^\theta(\tilde{T}, X^{t,x}_{\tilde{T}})\right) \longrightarrow_{\tilde{T}\to T} \mathbb{E}^{Q^{K,t,x}}\left(W^\theta(T, X^{t,x}_{T})\right) \tag{VIII.4.56}$$

Pour la convergence de droite, en plus de la majoration (VIII.4.55), on a utilisé la continuité de W^θ sur $[0,T]\times\mathbb{R}$ ainsi que le fait que le processus $X^{t,x}$ n'admet pas de temps fixe de discontinuité, et en particulier en T, d'après la Remarque VIII.2.3.
Le sort du deuxième terme peut se régler par convergence dominée (en tenant compte du fait que $\eta(K)$ est bornée et que l'on a la condition d'intégrabilité (VIII.4.6) :

$$\begin{aligned}
&\mathbb{E}^{Q^{K,t,x}}\left[\int_t^{\tilde{T}\wedge\sigma_n}\left(\eta(K)(s,X_s^{t,x}) - \bar{l}(s,X_s^{t,x})\right)ds\right] \\
&\to_{n\to\infty} \mathbb{E}^{Q^{K,t,x}}\left[\int_t^{\tilde{T}}\left(\eta(K)(s,X_s^{t,x}) - \bar{l}(s,X_s^{t,x})\right)ds\right] \\
&\to_{\tilde{T}\to T} H(Q^{K,t,x}\|P) - \mathbb{E}^{Q^{K,t,x}}\left[\int_t^T \bar{l}(s,X_s^{t,x})ds\right]
\end{aligned} \tag{VIII.4.57}$$

Avec (VIII.4.56) et (VIII.4.57), la relation (VIII.4.54) donne exactement l'inégalité (VIII.4.27) souhaitée.

∎

Une question de mesurabilité restait à régler pour conclure la prévisibilité de l'intensité optimale :

Lemme VIII.4.9 *La fonction* $(t,x) \mapsto \lambda^{(\theta,t,x)}$ *est mesurable.*

Preuve : On peut regrouper les relations (VIII.4.48) et (VIII.4.49) sous la forme :

$$\chi(t,x,\lambda^{(\theta,t,x)}) = 0 \tag{VIII.4.58}$$

avec

$$\chi(t,x,\lambda) = \lambda - \int (J^*)'(U^\theta(t,x,z) - \lambda\beta(t,x,z))\rho(dz) - \alpha(t,x) + \int \beta(t,x,z)\rho(dz) \tag{VIII.4.59}$$

Pour montrer que cette relation définit $\lambda^{(\theta,t,x)}$ comme une fonction mesurable de (t,x), puisque les conditions du Théorème des fonctions implicites ne sont pas forcément vérifiées, on peut utiliser un procédé de définition dichotomique de $\lambda^{(\theta,t,x)}$.
On considère la suite croissante de parties finies de $\mathbb{R}$:

$$\left\{ \frac{j}{2^N} \;\middle|\; -N2^N \leq j \leq N2^N \right\}$$

et l'on pose :

$$\lambda_N^{(\theta,t,x)} = \begin{cases} -N & \text{si } \chi(t,x,-N) \geq 0 \\ \frac{j}{2^N} & \text{si } \chi(t,x,\frac{j-1}{2^N}) < 0 \leq \chi(t,x,\frac{j}{2^N}),\ -N2^N+1 \leq j \leq N2^N \\ N & \text{si } \chi(t,x,N) < 0 \end{cases}$$

Puisque la fonction $\lambda \mapsto \chi(t,x,\lambda)$ est strictement croissante en tant que dérivée de la fonction strictement convexe sur $\mathbb{R}$ définie par le second membre de (VIII.4.44), ce procédé définit bien $\lambda_N^{(\theta,t,x)}$ et de plus cette suite converge avec N vers l'unique racine de $\lambda \mapsto \chi(t,x,\lambda)$: la mesurabilité de la limite $(t,x) \mapsto \lambda^{(\theta,t,x)}$ résulte donc de celle de $(t,x) \mapsto \lambda_N^{(\theta,t,x)}$ pour tout N. ∎

VIII.5 Régularisation du problème pénalisé et solution dans les espaces Hölderiens

Dans cette section, on se propose de résoudre le problème primal de calibration primal pénalisé $(P_{\chi^*}^{\overline{H},\mathbf{C}})$, en rappelant que cela passe par l'étude de la régularité de la fonction valeur $\theta \mapsto V(\theta)$ d'après la Proposition VIII.3.1.
En adoptant toujours la forme délocalisée (VIII.4.1) de la fonction de payoff f_*, les résultats de l'étude précédente sont valable et la régularité sera acquise avec celle de la fonction $\theta \mapsto W^\theta$, solution du système (HJB^θ) à condition de renforcer l'hypothèse (RS) par des conditions de nature Hölderienne.
On résoudra alors le problème grâce au théorème des fonctions implicites dans le cadre des

espaces Hölderiens, ainsi que cela a été fait dans [113] pour la calibration du modèle de volatilité locale.
Par ailleurs, afin d'assurer que l'on peut trouver un optimum dans la classe de contrôles $\mathcal{K}^{\overline{H}}$, il convient également de choisir

$$\psi_0 = \min\{\Phi_0, 1\} \tag{VIII.5.1}$$

ainsi qu'il apparaitra ultérieurement.
Avec le Théorème VIII.4.3 et la la Proposition VIII.3.1, l'objectif essentiel de cette partie est alors d'établir l'existence d'une solution $C^{1,2}$ du système (HJB^θ) et de montrer que cette solution est également régulière par rapport à θ, dans un voisinage du minimum θ^* de $\theta \mapsto V(\theta)+\chi(-\theta)-\mathbf{C}\cdot\theta$.
On consacre le début de cette partie à définir les espaces de Banach qui permettront l'emploi du théorème des fonctions implicites.

VIII.5.1 Espaces fonctionnels utilisés

En reprenant les notations de [55], on va utiliser des résultats classiques en introduisant d'abord pour $0 < \alpha < 1$ l'espace $C^{\frac{\alpha}{2},\alpha}$ des fonctions $\varphi(t,x)$ Höldériennes sur $[0,T] \times \mathbb{R}$ d'exposant $\frac{\alpha}{2}$ en t et α en x, c'est-à-dire telles que soit finie la semi-norme

$$< \varphi >^{(\alpha)} \equiv < \varphi >_t^{(\frac{\alpha}{2})} + < \varphi >_x^{(\alpha)} \tag{VIII.5.2}$$

avec

$$< \varphi >_t^{(\frac{\alpha}{2})} = \inf\{c \geq 0 \mid |\varphi(t,x) - \varphi(t',x)| \leq c|t-t'|^{\frac{\alpha}{2}},\ \forall t, t', x\} \tag{VIII.5.3}$$

et

$$< \varphi >_x^{(\alpha)} = \inf\{c \geq 0 \mid |\varphi(t,x) - \varphi(t,x')| \leq c|x-x'|^{\alpha},\ \forall t, x, x'\} \tag{VIII.5.4}$$

L'espace $C^{\frac{\alpha}{2},\alpha}$ est de Banach si on le munit de la norme

$$\|.\|_\alpha \equiv \|.\|_0 + < \varphi >^{(\alpha)} \tag{VIII.5.5}$$

où $\|.\|_0$ désigne la norme sup

$$\|\varphi\|_0 = \sup\{|\varphi(t,x)|,\ (t,x) \in [0,T] \times \mathbb{R}\} \tag{VIII.5.6}$$

On définit de même l'espace $C^{\frac{1+\alpha}{2},1+\alpha}$ des fonctions φ de classe $C^{0,1}$ sur $[0,T] \times \mathbb{R}$ telles que soit finie la quantité

$$< \varphi >^{(1+\alpha)} \equiv < \varphi >_t^{(\frac{1+\alpha}{2})} + \left\langle \frac{\partial \varphi}{\partial x} \right\rangle^{(\alpha)} \tag{VIII.5.7}$$

et $C^{\frac{1+\alpha}{2},1+\alpha}$ est un espace de Banach muni de la norme

$$\|\varphi\|_{1+\alpha} \equiv \|\varphi\|_0 + \left\| \frac{\partial \varphi}{\partial x} \right\|_0 + < \varphi >^{(1+\alpha)} \tag{VIII.5.8}$$

On définit enfin l'espace $C^{\frac{2+\alpha}{2},2+\alpha}$ des fonctions φ de classe $C^{1,2}$ sur $[0,T] \times \mathbb{R}$ telles que soit finie la quantité

$$< \varphi >^{(2+\alpha)} \equiv \left\langle \frac{\partial \varphi}{\partial t} \right\rangle^{(\alpha)} + \left\langle \frac{\partial^2 \varphi}{\partial x^2} \right\rangle^{(\alpha)} + \left\langle \frac{\partial \varphi}{\partial x} \right\rangle_t^{(\frac{1+\alpha}{2})} \tag{VIII.5.9}$$

et $C^{\frac{2+\alpha}{2},2+\alpha}$ est aussi un espace de Banach muni de la norme

$$\|\varphi\|_{2+\alpha} \equiv \sum_{0 \leq 2i+j \leq 2} \left\| \frac{\partial^{i+j} \varphi}{\partial t^i \partial x^j} \right\|_0 + < \varphi >^{(2+\alpha)} \tag{VIII.5.10}$$

VIII.5.2 Solution de l'équation de la programmation dynamique dans les espaces Hölderiens

Afin de pouvoir utiliser le cadre Hölderien, on est amené à faire des hypothèses sur les coefficients du modèle :

$$(RH) \begin{cases} \sigma_*^2, b_*, r_t \in C^{\frac{\alpha}{2},\alpha} \\ \bar{l} \in (C^{\frac{\alpha}{2},\alpha})^d \\ h \in (C^{2+\alpha})^d \end{cases}$$

où $0 < \alpha < 1$ est fixé.
Avec les deux dernières conditions de (RH), les fonctions de payoff sont bornées et donc l'hypothèse (RS) est en force.
On va compléter (RH) par une condition portant sur le coefficient du processus de sauts en supposant l'existence de Φ_α $(\mathbf{R}, \rho)$-mesurable et de $d_{\Phi_*} > 0$ telles que $\forall (t,x), (t',x') \in [0,T] \times \mathbb{R}$, $\forall z \in \mathbf{R}$, $\forall 0 \leq s \leq 1$

$$(RH') \begin{cases} |\Phi_*(t,x,z) - \Phi_*(t',x',z)| \leq \Phi_\alpha(z)(|t-t'|^{\frac{\alpha}{2}} + |x-x'|^\alpha) \\ |[x + s\ln(\Phi_*(t,x,z))] - [x' + s\ln(\Phi_*(t',x',z))]| \leq d_{\Phi_*}(|t-t'|^{\frac{1}{2}} + |x-x'|) \\ \int_{\mathbf{R}} \min\{\Phi_\alpha^2(z), \Phi_\alpha(z)\}\rho(dz) \leq d_{\Phi_*} \\ \Phi_\alpha \geq \Phi_0 \end{cases}$$

Remarque VIII.5.1 *Quitte à remplacer Φ_α par $\Phi_0 + \Phi_\alpha$, on peut toujours se ramener au cas où la dernière hypothèse $\Phi_\alpha \geq \Phi_0$ est satisfaite.*
On a en effet la double inégalité

$$\frac{1}{2}[\min\{\Phi_0, \Phi_0^2\} + \min\{\Phi_\alpha, \Phi_\alpha^2\}] \leq \min\{\Phi_0 + \Phi_\alpha, (\Phi_0 + \Phi_\alpha)^2\} \leq 4[\min\{\Phi_0, \Phi_0^2\} + \min\{\Phi_\alpha, \Phi_\alpha^2\}] \tag{VIII.5.11}$$

ce qui fait que le remplacement suggéré n'affecte en rien la condition d'intégrabilité dans (RH').

L'essentiel de l'étude repose sur le résultat suivant :

Proposition VIII.5.2 *On fait les hypothèses (RH) et (RH') et on considère une fonction vectorielle $\tilde{h} \in (C^{2+\alpha})^d$.*
On peut alors trouver un voisinage ouvert $\mathbf{U}$ de $(\mathbf{0}_{\mathbb{R}^d}, r, \tilde{h})$ dans $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d$ tel que pour $(\theta, b_*, h) \in \mathbf{U}$, le système (HJB^θ) admet une solution $W^\theta \in C^{\frac{2+\alpha}{2},2+\alpha}$ et telle que la solution K^θ du système d'équations (VIII.4.14) et (VIII.4.15) est un élément de $\mathcal{K}^{\overline{H}}$.
De plus, quitte à réduire $\mathbf{U}$, la fonction $(\theta, b_*, h) \in \mathbf{U} \mapsto W^\theta \in C^{\frac{2+\alpha}{2},2+\alpha}$ est indéfiniment différentiable.

La régularité par rapport à h est énoncée pour l'étude ultérieure des payoff non délocalisés (Théorème VIII.7.2).
Il serait également aisé d'étendre la régularité énoncée en incluant les autres paramètres du modèle σ_* et Φ_* ainsi que le coût intégral $\bar{l}$.

Remarque VIII.5.3 *La preuve de la Proposition VIII.5.2, donnera l'existence et l'unicité de la solution W^θ du système (HJB^θ) dans un voisinage de $\mathbf{0}_{\mathbb{R}[0,T]}$ dans $C^{\frac{2+\alpha}{2},2+\alpha}$.*
Le Théorème VIII.4.3 montre alors que l'unicité est globale et identifie W^θ à la fonction valeur V^θ définie par (VIII.4.9). La solution W^θ satisfait en effet aux conditions du Théorème VIII.4.3 puisque ce dernier est valide si l'on remplace la condition (RS) par la condition (RH) et puisque la fonction W^θ est bornée en tant qu'élément $C^{\frac{2+\alpha}{2},2+\alpha}$ et qu'elle vérifie donc la condition (VIII.4.17) du Théorème VIII.4.3.

La preuve de la proposition précédente est reportée à la section prochaine.

Solution du problème primal régularisé, stabilité de la solution

On suppose désormais fixé le payoff terminal h en s'autorisant à faire varier le multiplicateur de Lagrange θ ainsi que le coefficient b_*.
Sous les conditions de régularité (RH) et (RH'), la Proposition VIII.5.2 fournit l'existence d'une solution W^θ du système $(\overline{HJB}^\theta)$ dans $C^{\frac{2+\alpha}{2},2+\alpha}$.
Puisque l'on peut alors appliquer le Théorème VIII.4.3 en suivant les considérations faites au cours de la Remarque VIII.5.3, on en déduit

Corollaire VIII.5.4 *Sous les conditions (RH) et (RH') et pour $(\theta, b_*) \in \mathbf{U}$, voisinage ouvert assez petit de $(\mathbf{0}_{\mathbb{R}^d}, r)$ dans $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha}$, il existe un élément $K^\theta \in \mathcal{K}^{\overline{H}}$ tel que l'on ait $\forall (t,x) \in [0,T] \times \mathbb{R}$:*

$$V^\theta(t,x) = W^\theta(t,x) = \theta \cdot \mathbb{E}^{Q^{K^\theta},t,x}\left[\int_t^T \bar{l}(s, X_s^{t,x})ds + h(X_T^{t,x})\right] - H(Q^{K^\theta,t,x}\|P) \qquad \text{(VIII.5.12)}$$

et tel que $(\theta, b_*) \in \mathbf{U} \mapsto V^\theta \in C^{\frac{2+\alpha}{2},2+\alpha}$ est indéfiniment différentiable.

Le prochain résultat donne la solution du problème de calibration pénalisé $(P_\chi^{\overline{H},\mathbf{C}})$ associé au système (HJB^θ).
On commence par définir le vecteur de $\mathbb{R}^d$:

$$\mathbf{C}_r = \mathbb{E}^P(f_*(X^r)) \qquad \text{(VIII.5.13)}$$

où X^r désigne la solution de l'EDS (VIII.2.4) avec $b_* = r$ et où l'on rappelle que la fonction de payoff f_* est donnée par (VIII.4.1).
On énonce :

Théorème VIII.5.5 *Sous les conditions (RH) et (RH'), il existe un voisinage $\tilde{\mathbf{U}}$ de $(r, \mathbf{C}_r)$ dans $C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d$ tel que pour tout $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}}$, le problème $(P_{\chi^*}^{\overline{H},\mathbf{C}})$ admet une solution et c'est également la solution du problème $(P_{\chi^*}^{\overline{H}',\mathbf{C}})$ pour tout $\overline{H}' \geq \overline{H}$.*

Preuve : Notons de façon explicite V^{θ,b_*} pour faire apparaître la dépendance en b_* de la solution du système (HJB^θ).
Alors, compte tenu du Corollaire VIII.5.4, on peut trouver un voisinage ouvert $\mathbf{U}$ de $(\mathbf{0}_{\mathbb{R}^d}, r)$ dans $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha}$ tel que la fonction

$$\Gamma : (\theta, b_*, \mathbf{C}) \in \mathbf{U} \times \mathbb{R}^d \mapsto V^{\theta,b_*} + \chi(-\theta) - \mathbf{C}\cdot\theta \in C^{\frac{2+\alpha}{2},2+\alpha} \qquad \text{(VIII.5.14)}$$

est indéfiniment différentiable et il en est de même de la fonction :

$$\overline{\Gamma} : (\theta, b_*, \mathbf{C}) \in \mathbf{U} \times \mathbb{R}^d \mapsto V^{\theta,b_*}(0, X_0) + \chi(-\theta) - \mathbf{C}\cdot\theta \in \mathbb{R} \qquad \text{(VIII.5.15)}$$

ce qui permet de considérer en particulier sa différentielle suivant θ :

$$D_\theta\overline{\Gamma} : (\theta, b_*, \mathbf{C}) \in \mathbf{U} \times \mathbb{R}^d \mapsto D_\theta V^{\theta,b_*}(0, X_0) - D_\theta\chi(-\theta) - \mathbf{C} \in \mathbb{R}^d \qquad \text{(VIII.5.16)}$$

ainsi que sa différentielle seconde

$$D_\theta^2\overline{\Gamma} : (\theta, b_*, \mathbf{C}) \in \mathbf{U} \times \mathbb{R}^d \mapsto D_\theta^2 V^{\theta,b_*}(0, X_0) + D_\theta^2\chi(-\theta) \qquad \text{(VIII.5.17)}$$

qui est inversible de $\mathbb{R}^d$ sur $\mathbb{R}^d$ en tant que forme bilinéaire strictement positive car on rappelle qu'il en est de même de $D_\theta^2\chi$ par hypothèse et que V^{θ,b_*} est convexe en θ.
Si on peut montrer que $D_\theta\overline{\Gamma}(\mathbf{0}_{\mathbb{R}^d}, r, \mathbf{C}_r) = \mathbf{0}_{\mathbb{R}^d}$ le théorème des fonctions implicites permet

d'affirmer l'existence d'un voisinage ouvert $\mathbf{U}_1$ de $\mathbf{0}_{\mathbb{R}^d}$ ainsi que d'un second voisinage ouvert $\mathbf{U}_2$ de $(r, \mathbf{C}_r)$ dans $C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d$ tels que $\mathbf{U}_1 \times \mathbf{U}_2 \subset \mathbf{U} \times \mathbb{R}$ et que l'équation

$$D_\theta \overline{\Gamma}(\theta, b_*, \mathbf{C}) = \mathbf{0}_{\mathbb{R}^d} \tag{VIII.5.18}$$

définisse de manière unique dans $\mathbf{U}_1$ le multiplicateur de Lagrange θ comme une fonction $\theta^*(b_*, \mathbf{C})$ indéfiniment différentiable de $(b_*, \mathbf{C}) \in \mathbf{U}_2$.
Puisque l'équation (VIII.5.18) est la condition du premier ordre de minimum de $\theta \mapsto \overline{\Gamma}(\theta, b_*, \mathbf{C})$, pour b_* et $\mathbf{C}$ fixés et que c'est également une condition suffisante avec la convexité de cette dernière fonction, on en déduit, avec le Corollaire VIII.5.4 que la Proposition VIII.3.1 est vérifié avec $\theta^* = \theta^*(b_*, \mathbf{C})$ et $K^* = K^{\theta^*(b_*,\mathbf{C})}$ qui est donc solution du problème $(P_{\chi^*}^{\overline{H},\mathbf{C}})$.
Le théorème est donc prouvé avec $\tilde{\mathbf{U}} = \mathbf{U}_2$ à condition de satisfaire (VIII.5.18) pour $(\theta, r, \mathbf{C}) = (\mathbf{0}_{\mathbb{R}^d}, r, \mathbf{C}_r)$.
En reprenant les arguments précédents, cela équivaut à dire que $\theta \mapsto \overline{\Gamma}(\theta, r, \mathbf{C}_r)$ est minimale en $\mathbf{0}_{\mathbb{R}^d}$.
Ce point résulte de la relation de dualité établie au début de la preuve de la Proposition VIII.3.1 : en choisissant $b_* = r$ et $K = 1$ (ce qui est possible avec la définition de $\mathcal{K}^{\overline{H}}$, la probabilité $Q^{K,0,X_0}$ est égale à la mesure a priori puisque cette dernière rend alors martingale l'actif actualisé (on peut bien sûr vérifier directement l'égalité $Q^{K,0,X_0} = P$ grâce aux relations (VIII.2.26) et (VIII.2.24) avec $G^{K,t,x} = 0$ d'après (VIII.2.25)).
La relation (VIII.3.4) donne alors, en utilisant la définition (VIII.5.13) de $\mathbf{C}_r$:

$$0 = H(P\|P) + \chi^*\bigg(\int f_*(X^r)dP - \mathbf{C}_r\bigg) \geq -\inf_{\theta \in \mathbb{R}^d}\bigg(V^{\theta,r}(0, X_0) + \chi(-\theta) - \mathbf{C}_r \cdot \theta\bigg) \tag{VIII.5.19}$$

D'autre part, avec l'expression (VIII.4.9) de V^θ, on en déduit également

$$V^{\mathbf{0}_{\mathbb{R}^d},r}(0, X_0) = \sup_{K \in \mathcal{K}^{\overline{H}}} -H(Q^{K,t,x}\|P) \geq 0$$

en prenant encore une fois $K = 1$ et donc $Q^{K,t,x} = P$.
Comme l'entropie relative est positive, la dernière inégalité est en fait une égalité, ce qui permet d'avoir :

$$\overline{\Gamma}(\mathbf{0}_{\mathbb{R}}, r, \mathbf{C}_r) = 0$$

Avec la relation (VIII.5.19), il vient alors :

$$\overline{\Gamma}(\mathbf{0}_{\mathbb{R}}, r, \mathbf{C}_r) = \inf_{\theta \in \mathbb{R}^d} \overline{\Gamma}(\theta, r, \mathbf{C}_r) \tag{VIII.5.20}$$

et c'est la condition de minimum cherché.
Enfin, on verra que le choix du voisinage $\tilde{\mathbf{U}}$ de $(r, \mathbf{C}_r)$ peut être fait tel que l'on puisse remplacer $(J^*)'$ par une régularisée dans l'expression (VIII.4.14) de l'optimum K^θ : en particulier, la valeur de $(J^*)'(u)$ pour $|u| \geq \ln(\overline{H})$ ne joue pas, ce qui prouve aussi que les équations (VIII.4.14) et (VIII.4.15) reste valables si l'on augmente $\overline{H}$, et W^θ demeure donc solution. ∎

Remarque VIII.5.6 *Quitte à le restreindre, le voisinage $\tilde{\mathbf{U}}$ peut être pris du type :*

$$\{(b_*, \mathbf{C}) \in C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d \mid \|\mathbf{C} - \mathbf{C}_r\| + \|b_* - r\|_\alpha < \epsilon\}$$

avec $\epsilon > 0$.
La quantité $\|\mathbf{C} - \mathbf{C}_r\| + \|b_* - r\|_\alpha$ a l'interprétation naturelle de mesurer la proximité entre la loi a priori et les contraintes de martingalité et de prix observés.
Le Théorème VIII.5.5 affirme donc que si P est assez proche des contraintes au sens de ce critère, on obtient la solution du problème de calibration pénalisée.

Pour terminer cette partie, on montre que le prix calibré des options européennes est stable de manière générale.

Théorème VIII.5.7 *Si ψ est une variable aléatoire dans $L^q(\Omega, P)$, $1 < q < \infty$ fixé , alors, sous les conditions (RH) et (RH'), le prix calibré $(b_*, \mathbf{C}) \in C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d \mapsto \mathbb{E}^{Q^{K^{\theta^*(b_*,\mathbf{C})},0,X_0}}(\psi)$ définit une fonction indéfiniment différentiable sur le voisinage $\tilde{\mathbf{U}}$ de $(r, \mathbf{C}_r)$.*

La preuve est renvoyée à la prochaine section.
Toutefois, en adaptant le Théorème 3.3 de [113], on peut remarquer que le résultat précédent peut être démontré de manière directe dans le cas d'une option à payoff régularisé du type

$$g_*(X) = \int_0^T \tilde{g}(t, X_t)dt + \tilde{g}_T(X_T) \tag{VIII.5.21}$$

En effet, sous les conditions (RH) et (RH') et en supposant de plus $\tilde{g} \in C^{\frac{\alpha}{2},\alpha}$, $\tilde{g}_T \in C^{\alpha}$, on peut établir que le prix calibré de l'option $\int g_*(X^{0,X_0})dQ^{K^{\theta^*(b_*,\mathbf{C})},0,X_0}$, s'écrit $V_{g_*}^{\theta^*(b_*,\mathbf{C}),b_*}$ où $V_{g_*}^{\theta,b_*}$ est la solution du système :

$$\begin{cases} \frac{\partial W}{\partial t} + \mathcal{G}^{K^\theta} W + \tilde{g} = 0 \text{ sur } [0,T] \times \mathbb{R} \\ W(T,x) = \tilde{g}_T(x) \text{ sur } \mathbb{R} \end{cases}$$

où l'on rappelle que K^θ est l'optimum défini au Corollaire VIII.5.4.
Comme on pourra voir d'autre part dans la preuve de la Proposition VIII.5.2 que K^θ est donné de manière explicite par (VIII.5.32), la régularité de $(\theta, b_*) \in \mathbf{U} \mapsto V_{g_*}^{\theta,b_*}$ s'obtient comme celle de V^θ, établie à la Proposition VIII.5.2, et l'on en déduit donc la régularité du prix de l'option par composition avec θ^*.

VIII.5.3 Preuve des résultats de l'étude précédente

Dans cette partie, on donne la preuve de la Proposition VIII.5.2 ainsi que du Théorème VIII.5.7.
Preuve de la Proposition VIII.5.2 : Introduisons pour toute fonction W de classe $C^{1,2}$ l'opérateur linéaire du second ordre

$$A(t,x)W = (r_t - \frac{1}{2}\sigma_*^2(t,x))\frac{\partial W}{\partial x} + \frac{1}{2}\sigma_*^2(t,x)\frac{\partial^2 W}{\partial x^2} \tag{VIII.5.22}$$

ainsi que l'opérateur non linéaire

$$\begin{aligned} H(t,x)W &= \sup_{k-1 \in L^2(\mathbf{R},\rho)} \Bigg[\int_{\mathbf{R}} \Bigg(W(t, x + \ln(1 + \Phi_*(t,x,z))) - W(t,x) \\ &\quad -\Phi_*(t,x,z)\frac{\partial W}{\partial x}\Bigg) k(z)\rho(dz) - \eta(k) \Bigg] \\ &= \int_{\mathbf{R}} \Bigg(W(t, x + \ln(1 + \Phi_*(t,x,z))) - W(t,x) - \Phi_*(t,x,z)\frac{\partial W}{\partial x}\Bigg)\rho(dz) \\ &\quad + \sup_{k \in L^2(\mathbf{R},\rho)} \Bigg[\int_{\mathbf{R}} \Bigg(W(t, x + \ln(1 + \Phi_*(t,x,z))) - W(t,x) \\ &\quad -\Phi_*(t,x,z)\frac{\partial W}{\partial x}\Bigg) k(z)\rho(dz) - \eta_0(k) \Bigg] \end{aligned} \tag{VIII.5.23}$$

Avec l'hypothèse (RH), il est enfin licite de faire le changement de fonction $u = W - \theta \cdot h$ dans (HJB^θ), et, avec les notations précèdentes, cela conduit à résoudre ce système sous la forme

$$(HJB_0^\theta) \left\{ \begin{array}{l} \frac{\partial u}{\partial t} + A(t,x)u + H(t,x)[u + \theta \cdot h] + \theta \cdot [A(t,x)h + \bar{l}] = 0 \text{ sur } [0,T] \times \mathbb{R} \\ u \in \mathcal{H}_0^\alpha \end{array} \right.$$

où l'on a également posé

$$\mathcal{H}_0^\alpha \equiv \{u \in C^{\frac{2+\alpha}{2}, 2+\alpha} \mid u(T,x) = 0 \text{ sur } \mathbb{R}\} \qquad \text{(VIII.5.24)}$$

Le membre de gauche de l'EDP ci dessus définit la fonction

$$\begin{aligned} \Psi : (u, \theta, b_*, h) \in \mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d \longmapsto \quad & \frac{\partial u}{\partial t} + A(t,x)u + H(t,x)[u + \theta \cdot h] \\ & + \theta \cdot [A(t,x)h + \bar{l}] \in C^{\frac{\alpha}{2},\alpha} \end{aligned} \qquad \text{(VIII.5.25)}$$

On note que la dépendance de Ψ en b_*, la dérive de la diffusion, se fait à travers le terme d'entropie η qui figure dans l'opérateur intégro-différentiel $H(t,x)$.
D'autre part, il n'est pas tout à fait évident que Ψ prend effectivement ses valeurs dans $C^{\frac{\alpha}{2},\alpha}$, mais ce fait recevra confirmation dans la suite.
On vérifie alors que $\Psi(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}) = 0$ et la proposition sera acquise si l'on peut appliquer le théorème des fonctions implicites au point $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$.
On va démontrer la C^∞-différentiabilité de $(u, \theta, h) \mapsto \Psi(u, \theta, b_*, h)$ au voisinage du point $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$ et on va prouver que la différentielle $D_u(\Psi)(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}) : \mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \to C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d$ est continuement inversible.
En regroupant les parties linéaire et bilinéaire de Ψ, on pose

$$L : (u, \theta, h) \in \mathcal{H}_0^\alpha \times \mathbb{R}^d \times (C^{2+\alpha})^d \mapsto \frac{\partial u}{\partial t} + A(t,x)u + \theta \cdot [A(t,x)h + \bar{l}] \in C^{\frac{\alpha}{2},\alpha} \qquad \text{(VIII.5.26)}$$

On prouve alors que l'opérateur L est bien à valeurs dans $C^{\frac{\alpha}{2},\alpha}$ et qu'il est continu.
Cela se voit en considérant les opérateurs $u \in C^{\frac{1+\alpha}{2},1+\alpha} \mapsto \frac{\partial u}{\partial x} \in C^{\frac{\alpha}{2},\alpha}$, $u \in C^{\frac{2+\alpha}{2},2+\alpha} \mapsto \frac{\partial^2 u}{\partial x^2} \in C^{\frac{\alpha}{2},\alpha}$, $(u,v) \in C^{\frac{\alpha}{2},\alpha} \times C^{\frac{\alpha}{2},\alpha} \mapsto uv \in C^{\frac{\alpha}{2},\alpha}$ et $(\theta, u) \in \mathbb{R}^d \times (C^{\frac{\alpha}{2},\alpha})^d \mapsto \theta \cdot u \in C^{\frac{\alpha}{2},\alpha}$.
la continuité de ces opérateurs s'obtient en utilisant les définitions (VIII.5.5), (VIII.5.8), (VIII.5.10) des normes $\|.\|_\alpha$, $\|.\|_{1+\alpha}$ et $\|.\|_{2+\alpha}$ et la continuité de L en résulte.
En conclusion, en tant que somme de deux opérateurs continus respectivement linéaire et bilinéaire, L est C^∞-différentiable sur $\mathcal{H}_0^\alpha \times \mathbb{R}^d \times (C^{2+\alpha})^d$.

Passons donc à la partie intégro-différentielle de Ψ, c'est-à-dire à l'étude de l'opérateur

$$H : (u, \theta, b_*, h) \in \mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d \mapsto H(u + \theta \cdot h) \in C^{\frac{\alpha}{2},\alpha} \qquad \text{(VIII.5.27)}$$

Afin d'expliciter H et de vérifier en particulier qu'il est à valeurs dans $C^{\frac{\alpha}{2},\alpha}$, on reprend l'expression (VIII.5.23) de H, pour trouver

$$K^\theta(t,x,.) - 1 = \arg \max_{k \in L^2(\mathbf{R},\rho)} \left(\int_{\mathbf{R}} U^{\theta,h}(t,x,z) k(z) \rho(dz) - \eta_0(k)(t,x) \right) \qquad \text{(VIII.5.28)}$$

en posant

$$\begin{aligned} U^{\theta,h}(t,x,z) \quad = \quad & u(t, x + \ln(1 + \Phi(t,x,z))) - u(t,x) - \Phi(t,x,z) \frac{\partial u}{\partial x}(t,x) \\ & + \theta \cdot \left[h(x + \ln(1 + \Phi(t,x,z))) - h(x) - \Phi(t,x,z) \frac{\partial h}{\partial x}(x) \right] \end{aligned} \qquad \text{(VIII.5.29)}$$

On suit ce qui a été fait pour résoudre le problème (VIII.4.30) dans la preuve du Théorème VIII.4.3, en remplaçant simplement U^θ par $U^{\theta,h}$.
Compte tenu du fait que l'inégalité (VIII.4.52) est réalisée avec égalité et compte tenu aussi de l'égalité (VIII.4.50), ainsi que des définitions (VIII.5.27) et (VIII.5.23) des opérateurs intégraux, on peut écrire

$$\begin{aligned} H(u,\theta,b_*,h) &= \int [U^{\theta,h}(t,x,z) + J^*(U^{\theta,h}(t,x,z) - \lambda^{(u,\theta,b_*,h)}(t,x)\beta(t,x,z))]\rho(dz) \\ &\quad -\lambda^{(u,\theta,b_*,h)}(t,x)\frac{b_*(t,x)-r_t}{\sigma_*(t,x)} + \frac{1}{2}(\lambda^{(u,\theta,b_*,h)}(t,x))^2 \end{aligned} \tag{VIII.5.30}$$

où, suivant les relations (VIII.4.48) et (VIII.4.49), la fonction $\lambda^{(u,\theta,b_*,h)}(t,x)$ est définie implicitement, de manière unique, par l'équation

$$\lambda^{(u,\theta,b_*,h)}(t,x) - \int (J^*)'(U^{\theta,h}(t,x,z) - \lambda^{(u,\theta,b_*,h)}(t,x)\beta(t,x,z))\beta(t,x,z)\rho(dz) - \frac{b_*(t,x)-r_t}{\sigma_*(t,x)} = 0 \tag{VIII.5.31}$$

Dans ce qui précède, on a repris la notation (VIII.4.34)

$$\beta(t,x,z) = \frac{\Phi_*(t,x,z)}{\sigma_*(t,x)}$$

et on a fait apparaitre explicitement la dépendance par rapport à b_*.
Accessoirement, on trouve l'optimum

$$K^\theta(t,x,z) - 1 = (J^*)'(U^{\theta,h}(t,x,z) - \lambda^{(u,\theta,b_*,h)}(t,x)\beta(t,x,z)) \tag{VIII.5.32}$$

Ainsi qu'on l'a déjà remarqué lors de la preuve du Lemme VIII.4.9, la non dérivabilité de $J^{*'}$ en $\ln(\overline{H})$ et $-\ln(\overline{H})$ peut faire obstacle à l'application du théorème des fonctions implicites à l'équation (VIII.5.31) et l'on considère pour cela une fonction $\tilde{J}^*$ de classe C^∞, à support compact et égale à J^* sur un voisinage de l'origine.
Cela permet de définir une version régularisée du premier membre de (VIII.5.31) :

$$\begin{aligned} \tilde{\zeta} \quad &: \quad (\lambda,u,\theta,b_*,h) \in C^{\frac{\alpha}{2},\alpha} \times \mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d \\ &\mapsto \lambda - \int \tilde{J}^{*'}(U^{\theta,h}(.,.,z) - \lambda\beta(.,.,z))\beta(.,.,z)\rho(dz) - \frac{b_* - r_.}{\sigma_*} \in C^{\frac{\alpha}{2},\alpha} \end{aligned} \tag{VIII.5.33}$$

On peut alors montrer que $\tilde{\zeta}$ est effectivement à valeurs dans $C^{\frac{\alpha}{2},\alpha}$, qu'elle est de classe C^∞ et que sa différentielle par rapport à la première variable λ est

$$D_\lambda\tilde{\zeta}(\lambda,u,\theta,b_*,h) = \left[1 + \int \tilde{J}^{*''}(U^{\theta,h}(.,.,z) - \lambda\beta(.,.,z))\beta^2(.,.,z)\rho(dz)\right] id_{C^{\frac{\alpha}{2},\alpha}} \tag{VIII.5.34}$$

où le membre de droite désigne l'endomorphisme identité sur $C^{\frac{\alpha}{2},\alpha}$.
Cela se voit en faisant subir à (VIII.5.33) une différentiation formelle par rapport à λ qui ne nécessite d'être justifiée que pour le deuxième terme : celle-ci sera acquise avec le Lemme VIII.5.10.
On voit alors que $D_\lambda\tilde{\zeta}(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}) = [1 + \int_{\mathbf{R}} \beta^2(.,.,z)\rho(dz)] id_{C^{\frac{\alpha}{2},\alpha}}$ est inversible sur $C^{\frac{\alpha}{2},\alpha}$et que son inverse $[1 + \int_{\mathbf{R}} \beta^2(.,.,z)\rho(dz)]^{-1} id_{C^{\frac{\alpha}{2},\alpha}}$ est continue.
Comme d'autre part $\tilde{\zeta}(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}) = \mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}$, le théorème des fonctions implicites énonce que l'on peut trouver un voisinage ouvert B_0 de $\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}$ dans $C^{\frac{\alpha}{2},\alpha}$ ainsi qu'un voisinage ouvert B_1 de $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$ dans $\mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d$ tels que l'équation

$$\tilde{\zeta}(\lambda,u,\theta,b_*,h) = \mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}} \tag{VIII.5.35}$$

définit dans $B_0 \times B_1$ l'élément λ comme une fonction C^∞

$$(u,\theta,b_*,h) \in B_1 \mapsto \tilde{\lambda}^{(u,\theta,b_*,h)} \in B_0 \tag{VIII.5.36}$$

On note pour la suite l'égalité

$$\tilde{\lambda}^{(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}},\mathbf{0}_{\mathbb{R}^d},(r_t),\tilde{h})} = \mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}} \tag{VIII.5.37}$$

Avec (VIII.5.30), la version régularisée de H s'écrit alors, en se restreignant à l'ouvert B_1 :

$$\begin{aligned}\tilde{H}(u,\theta,b_*,h) &= \int [U^{\theta,h}(.,.,z) + \tilde{J}^*(U^{\theta,h}(.,.,z) - \tilde{\lambda}^{(u,\theta,b_*,h)}\beta(.,.,z))]\rho(dz)\\ &\quad -\tilde{\lambda}^{(u,\theta,b_*,h)}\frac{b_* - r.}{\sigma_*} + \frac{1}{2}(\tilde{\lambda}^{(u,\theta,b_*,h)})^2\end{aligned} \tag{VIII.5.38}$$

Compte tenu en particulier du fait que $(u,\theta,b_*,h) \in B_1 \mapsto \tilde{\lambda}^{(u,\theta,b_*,h)} \in C^{\frac{\alpha}{2},\alpha}$ est C^∞, on peut montrer avec le Lemme VIII.5.10 qu'il en est de même de $(u,\theta,b_*,h) \in B_1 \mapsto \tilde{H}(u,\theta,b_*,h) \in C^{\frac{\alpha}{2},\alpha}$.
Finalement, $\tilde{\lambda}^{(u,\theta,b_*,h)}$ et $\lambda^{(u,\theta,b_*,h)}$ coincident sur B_1, quitte à restreindre ce dernier en supposant :

$$\forall(\lambda,u,\theta,b_*,h) \in B_0 \times B_1,\ -\ln\overline{H} < U^{\theta,h} - \lambda\beta < \ln\overline{H} \tag{VIII.5.39}$$

puisqu'alors la régularisation de J^* en $\tilde{J}^*$ est inutile dans l'équation (VIII.5.31), tant que (λ,u,θ,b_*,h) demeure dans $B_0 \times B_1$.
Par conséquent, en reprenant, la condition (VIII.5.39) et en rappelant que $\tilde{\lambda}^{(u,\theta,b_*,h)}$ est sur B_1 à valeurs dans B_0, il apparaît donc que sur B_1 l'expression (VIII.5.38) de l'opérateur $\tilde{H}$ est identique à celle (VIII.5.30) de H.
En rassemblant tout ce qui précède, on a donc prouvé que $\Psi = L + H$ est à valeurs dans $C^{\frac{\alpha}{2},\alpha}$, de classe C^∞ sur le voisinage ouvert B_1 de $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}},\mathbf{0}_{\mathbb{R}^d},(r_t),\tilde{h})$.
Poursuivant le programme fixé, montrons que $D_u(\Psi)(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}},\mathbf{0}_{\mathbb{R}^d},(r_t),\tilde{h}) : \mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d \to C^{\frac{\alpha}{2},\alpha}$ est continuement inversible.
On peut calculer de manière générale la différentielle $D_u(\Psi)(u,\theta,b_*,h)$ en $(u,\theta,b_*,h) \in B_1$, et l'on a vu que cela revenait à utiliser la version régularisée $\tilde{J}^*$ de J^*, ce que l'on fait par commodité.
On a donc pour $(u,\theta,b_*,h) \in B_1$, compte tenu de (VIII.5.26) :

$$D_u\Psi(u,\theta,b_*,h) = \frac{\partial}{\partial t} + A + D_u\tilde{H}(u,\theta,b_*,h) \tag{VIII.5.40}$$

Seule la partie non linéaire nécessite d'être traitée, mais ce qui précède en a préparé l'étude.
D'abord, on sait que $\tilde{\lambda}^{(u,\theta,b_*,h)}$ est C^∞ sur B_1, et donc en particulier différentiable par rapport à u.
En utilisant encore une fois le Lemme VIII.5.10, on peut ensuite différentier par rapport à u l'expression (VIII.5.38) de $\tilde{H}$:

$$\begin{aligned}\forall v \in \mathcal{H}_0^\alpha,\ D_u\tilde{H}(u,\theta,b_*,h).v &= \int [1 + \tilde{J}^{*\prime}(U^{\theta,h}(.,.,z) - \tilde{\lambda}^{(u,\theta,b_*,h)}\beta(.,.,z))]\Delta_* v(.,.,z)\rho(dz)\\ &\quad + \Big\{ -\int \tilde{J}^{*\prime}(U^{\theta,h}(.,.,z) - \tilde{\lambda}^{(u,\theta,b_*,h)}\beta(.,.,z))\beta(.,.,z)\rho(dz)\\ &\quad - \frac{b_* - r.}{\sigma_*} + \tilde{\lambda}^{(u,\theta,b_*,h)}\Big\} D_u\tilde{\lambda}^{(u,\theta,b_*,h)}.v\end{aligned} \tag{VIII.5.41}$$

où l'on a posé

$$\Delta_* v(t,x,z) = v(t, x+\ln(1+\Phi_*(t,x,z))) - v(t,x) - \Phi_*(t,x,z)\frac{\partial v}{\partial x}(t,x) \tag{VIII.5.42}$$

et en particulier, puisque l'on dispose également de l'égalité (VIII.5.37) :

$$\begin{aligned}\forall v \in \mathcal{H}_0^\alpha,\ D_u\tilde{H}(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}).v &= \int [1+\tilde{J^*}'(0)]\Delta_* v(.,.,z)\rho(dz) \\ &\quad - \int \tilde{J^*}'(0))\beta(.,.,z)\rho(dz) D_u\tilde{\lambda}^{(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})}.v \\ &= \int \Delta_* v(.,.,z)\rho(dz) \end{aligned} \tag{VIII.5.43}$$

où l'on a pris en compte l'expression (VIII.4.41) de J^*.
Alors, avec (VIII.5.40), on obtient la différentielle de Ψ en $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$

$$\begin{aligned} D_u\Psi(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}) : v \in \mathcal{H}_0^\alpha \mapsto \quad & Lv + \textstyle\int \Delta_* v(.,.,z)\rho(dz) \\ & \in C^{\frac{\alpha}{2},\alpha} \end{aligned} \tag{VIII.5.44}$$

Il s'agit pour conclure de montrer que la différentielle $D_u\Psi(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$ calculée ci-dessus possède un inverse continu : l'équation $\Psi(u,\theta,b_*,h)$ définit ainsi implicitement u comme une fonction prenant ses valeurs dans $C^{\frac{2+\alpha}{2},2+\alpha}$ et de classe C^∞ sur un voisinage de $(\mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$ dans $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d$, et il en est de même de $W = u + \theta\cdot h$.
Le résultat d'inversibilité qu'il faut produire repose sur des techniques classiques concernant les équations intégro-différentielles aux dérivées partielles et est traité à part dans le lemme qui suit la preuve de la proposition.
Pour achever cette dernière, on montre que K^θ appartient à $\mathcal{K}^{\overline{H}}$: en effet, l'inégalité (VIII.4.4) permet d'établir, compte tenu du fait que Φ_0 est bornée :

$$|U^{\theta,h}(t,x,z) - \tilde{\lambda}^{(u,\theta,b_*,h)}(t,x)\beta(t,x,z)| \le M_0(\|u\|_{2+\alpha} + \|\theta\|\|h\|_{2+\alpha} + \|\tilde{\lambda}^{(u,\theta,b_*,h)}\|_0)\Phi_0(z) \tag{VIII.5.45}$$

En utilisant ensuite le fait que $\tilde{\lambda}^{(u,\theta,b_*,h)}$ et $\lambda^{(u,\theta,b_*,h)}$ coincident sur B_1, ouvert assez petit de $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h})$ dans $\mathcal{H}_0^\alpha \times \mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d$, l'égalité (VIII.5.32) donne en injectant la majoration (VIII.5.45) dans (VIII.4.46) :

$$|K^\theta(t,x,z) - 1| \le M_0\overline{H}(\|u\|_{2+\alpha} + \|\theta\|\|h\|_{2+\alpha} + \|\tilde{\lambda}^{(u,\theta,b_*,h)}\|_0)\Phi_0(z) \tag{VIII.5.46}$$

Comme l'expression (VIII.4.42) de $(J^*)'$ permet d'encadrer celle (VIII.5.32) de K^θ :

$$\frac{1}{\overline{H}} \le K^\theta \le \overline{H}$$

l'inégalité des accroissements finis appliquée à $x \mapsto \ln(1+x)$ donne :

$$|\ln(K^\theta)| \le \overline{H}|K^\theta - 1|$$

Alors, si l'ouvert B_1 est assez petit pour que l'on ait :

$$\|u\|_{2+\alpha} + \|\theta\|\|h\|_{2+\alpha} + \|\tilde{\lambda}^{(u,\theta_{b_*},h)}\|_0 \le \frac{\ln(\overline{H})}{M_0\overline{H}^2}$$

on en déduit avec (VIII.5.46) :

$$|\ln(K^\theta)| \le \ln(\overline{H})\Phi_0$$

ce qui signifie bien l'appartenance de K^θ à $\mathcal{K}^{\overline{H}}$ avec le choix $\psi_0 = \Phi_0$. ∎

Pour la suite, il est commode d'introduire les quantités suivantes pour toute fonction $\varphi \in L^0([0,T] \times \mathbb{R} \times \mathbf{R})$, l'ensemble des fonctions mesurables sur $[0,T] \times \mathbb{R} \times \mathbf{R}$:

$$\|\varphi\|_{\mathcal{E}_\alpha} \equiv \inf \Big\{ c > 0,\ \forall z \in \mathbf{R},\ \|\varphi(.,.,z)\|_0 \le c \text{ et } \|\varphi(.,.,z)\|_\alpha \le c\Phi_\alpha(z) \Big\} \qquad \text{(VIII.5.47)}$$

et

$$\|\varphi\|_{\mathcal{E}_\alpha} \equiv \inf \Big\{ c > 0,\ \forall z \in \mathbf{R},\ \|\varphi(.,.,z)\|_0 \le c \text{ et } < \varphi(.,.,z) >_\alpha \le c\Phi_\alpha(z) \Big\} \qquad \text{(VIII.5.48)}$$

et on définit les espaces :

$$\mathcal{E}_\alpha = \Big\{ \varphi \in L^0([0,T] \times \mathbb{R} \times \mathbf{R}) \mid \|\varphi\|_{\mathcal{E}_\alpha} < +\infty \Big\}$$

et

$$\mathcal{E}^0_\alpha = \Big\{ \varphi \in L^0([0,T] \times \mathbb{R} \times \mathbf{R}) \mid \|\varphi\|_{\mathcal{E}^0_\alpha} < +\infty \Big\}$$

On donne alors deux majorations utiles dont on reporte la preuve après celle du Lemme VIII.5.11 :

Lemme VIII.5.8 *On peut trouver une constante M telle que pour $\psi_0, \psi_1, \psi_2 \in L^0([0,T]\times\mathbb{R}\times\mathbf{R})$ on a :*

$$\forall z \in \mathbf{R},\ \|\psi_0(.,.,z))\psi_1(.,.,z)\|_\alpha \le M\|\psi_0\|_{\mathcal{E}^0_\alpha}\|\psi_1\|_{\mathcal{E}_\alpha}\Phi_\alpha(z) \qquad \text{(VIII.5.49)}$$

et

$$\forall z \in \mathbf{R},\ \|\psi_0(.,.,z))\psi_1(.,.,z)\psi_2(.,.,z)\|_\alpha \le M\|\psi_0\|_{\mathcal{E}^0_\alpha}\|\psi_1\|_{\mathcal{E}_\alpha}\|\psi_2\|_{\mathcal{E}_\alpha}\min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\} \qquad \text{(VIII.5.50)}$$

Pour compléter la preuve de la Proposition VIII.5.2, on énonce ensuite le résultat suivant :

Lemme VIII.5.9

$$\forall \varphi \in C^{\frac{\alpha}{2},\alpha} \exists ! v \in \mathcal{H}^\alpha_0,\ D_u\Psi(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}}, \mathbf{0}_{\mathbb{R}^d}, (r_t), \tilde{h}).v = \varphi \qquad \text{(VIII.5.51)}$$

De plus, on vérifie :

$$\|v\|_{2+\alpha} \le c\|\varphi\|_\alpha \qquad \text{(VIII.5.52)}$$

où c est une constante.

Preuve : Afin d'appliquer le Théorème II.3.1. de [55] sur les équations intégro-différentielles, on utilise la majoration (VIII.4.5) avec $W(x) = x$, en la renforçant puisqu'alors les dérivées de W sont bornées, ce qui donne

$$|\ln(1 + \Phi_*(t,x,z)) - \Phi_*(t,x,z)| \le B\Phi^2_0(z)$$

et même, puisque Φ_0 est bornée

$$|\ln(1 + \Phi_*(t,x,z)) - \Phi_*(t,x,z)| \le B\min\{\Phi_0(z), \Phi^2_0(z)\}$$

quitte à augmenter B.
On s'autorise alors la réécriture de l'équation (VIII.5.51) sous la forme

$$\begin{aligned} L(t,x)v &+ \left[\int_{\mathbf{R}} [\Phi_*(t,x,z) - \ln(1+\Phi_*(t,x,z))]\rho(dz)\right] \frac{\partial v}{\partial x}(t,x) + \\ &\int_{\mathbf{R}} \left[v(t, x+\ln(1+\Phi_*(t,x,z))) - v(t,x) - \ln(1+\Phi_*(t,x,z)) \frac{\partial v}{\partial x}(t,x) \right] \rho(dz) = \varphi \end{aligned} \tag{VIII.5.53}$$

Le premier terme L de (VIII.5.53) constitue un opérateur linéaire uniformément parabolique (car σ_* est uniformément minoré), à coefficient dans $C^{\frac{\alpha}{2},\alpha}$ tandis que le troisième terme a la forme spécifiée dans le Théorème II.3.1. de [55], avec entre autre la première condition de (RH').
Le lemme résulte alors du théorème cité à condition de prouver que le second terme linéaire de (VIII.5.53) a son coefficient dans $C^{\frac{\alpha}{2},\alpha}$, ce qui revient à montrer que

$$\left\| \int_{\mathbf{R}} (\Phi_*(.,.,z) - \ln(1+\Phi_*(.,.,z)))\rho(dz) \right\|_\alpha < +\infty \tag{VIII.5.54}$$

La formule de Taylor avec reste intégral appliquée en 0 à la fonction $u \mapsto \ln(1+u)$ donne :

$$\ln(1+\Phi_*) - \Phi_* = -\Phi_*^2 \int_0^1 \frac{1-s}{(1+s\Phi_*)^2} ds \tag{VIII.5.55}$$

Alors, la majoration (VIII.5.50) du Lemme VIII.5.8 s'écrit avec $\psi_0 = \frac{1-s}{(1+s\Phi_*)^2}$ et $\psi_1 = \psi_2 = \Phi_*$:

$$\forall z \in \mathbf{R},\ \|(\ln(1+\Phi_*) - \Phi_*)(.,.,z)\|_\alpha \leq M \sup_{0\leq s\leq 1} \left\| \frac{1-s}{(1+s\Phi_*)^2} \right\|_{\mathcal{E}_\alpha^0} \|\Phi_*\|_{\mathcal{E}_\alpha}^2 \min\{\Phi_\alpha(z), \Phi_\alpha^2(z)\} \tag{VIII.5.56}$$

La condition (RH') montre que :

$$\|\Phi_*\|_{\mathcal{E}_\alpha} < +\infty$$

et également, en utilisant aussi la minoration (VIII.1.6) :

$$\sup_{0\leq s\leq 1} \left\| \frac{1-s}{(1+s\Phi_*)^2} \right\|_{\mathcal{E}_\alpha^0} < +\infty$$

En intégrant ensuite la relation (VIII.5.56), et compte tenu de la condition d'intégrabilité sur Φ_α dans (RH'), il s'ensuit la relation voulue (VIII.5.54).

∎

On donne également un résultat de différentiabilité concernant une fonctionnelle intégrale utilisée à plusieurs reprises dans la preuve de la Proposition VIII.5.2.

Lemme VIII.5.10 *Sous les hypothèses (RH) et (RH'), la fonctionnelle*

$$I_{\tilde{J}^*} : (\lambda, u, \theta, h) \in C^{\frac{\alpha}{2},\alpha} \times C^{\frac{1+\alpha}{2},1+\alpha} \times \mathbb{R}^d \times (C^{2+\alpha})^d \mapsto \int \tilde{J}^*(U^{\theta,h}(.,.,z) - \lambda\beta(.,.,z))\rho(dz)$$

où $U^{\theta,h}$ est défini par (VIII.5.29), est indéfiniment différentiable et à valeurs dans $C^{\frac{\alpha}{2},\alpha}$.
Sa différentielle partielle d'ordre k par rapport aux trois premières variables en $(\lambda,u,\theta,h) \in C^{\frac{\alpha}{2},\alpha} \times C^{\frac{1+\alpha}{2},1+\alpha} \times \mathbb{R}^d \times (C^{2+\alpha})^d$ est donnée par :

$$\begin{aligned} D^{(k)}_{\lambda,u,\theta} I_{\tilde{J}^*}(\lambda,u,\theta) \quad &: \quad ((\lambda_1,u_1,\theta_1),\ldots,(\lambda_k,u_k,\theta_k)) \in (C^{\frac{\alpha}{2},\alpha} \times C^{\frac{1+\alpha}{2},1+\alpha} \times \mathbb{R}^d)^k \mapsto \\ &\int \tilde{J}^{*(k)}(U^{\theta,h}(.,.,z) - \lambda\beta(.,.,z)) \prod_{i=1}^k [\Delta_*(u_i + \theta_i \cdot h) - \lambda_i\beta(.,.,z)]\rho(dz) \end{aligned} \tag{VIII.5.58}$$

Preuve : On rappelle que l'opérateur Δ_* a été défini par la relation (VIII.5.42) que l'on réécrit ici pour $v \in C^{\frac{2+\alpha}{2},2+\alpha}$:

$$\Delta_* v(t,x,z) = v(t, x + \ln(1+\Phi_*(t,x,z))) - v(t,x) - \Phi_*(t,x,z)\frac{\partial v}{\partial x}(t,x)$$

et l'on a :

$$U^{\theta,h}(t,x,z) = \Delta_*(u+\theta\cdot h)(t,x,z)$$

On peut montrer l'existence d'une constante C_0 :

$$\forall v \in C^{\frac{2+\alpha}{2},2+\alpha},\ \|\Delta_* v(.,.,z)\|_\alpha \leq C_0\|v\|_{2+\alpha}\min\{\Phi_\alpha(z), \Phi_\alpha^2(z)\} \tag{VIII.5.59}$$

ainsi que

$$\forall v \in C^{\frac{2+\alpha}{2},2+\alpha},\ \|\Delta_* v(.,.,z)\|_0 \leq C_0\|v\|_{2+\alpha}\Phi_0^2(z) \tag{VIII.5.60}$$

ce qui fait l'objet du Lemme VIII.5.11, et il est également facile de prouver, quitte à choisir C_0 plus grand :

$$\forall v \in C^{\frac{2+\alpha}{2},2+\alpha},\ \|\beta(.,.,z)\|_\alpha \leq C_0\Phi_\alpha(z) \tag{VIII.5.61}$$

ainsi que :

$$\forall v \in C^{\frac{2+\alpha}{2},2+\alpha},\ \|\beta(.,.,z)\|_0 \leq C_0\Phi_0(z) \tag{VIII.5.62}$$

Les quatre majorations précédentes ainsi que le fait que la fonction Φ_0 est bornée, montrent, que l'application

$$(\lambda, u, \theta, h) \in C^{\frac{\alpha}{2},\alpha} \times C^{\frac{1+\alpha}{2},1+\alpha} \times \mathbb{R}^d \times (C^{2+\alpha})^d \mapsto \Delta_*(u+\theta\cdot h) - \lambda\beta \tag{VIII.5.63}$$

est à valeurs dans $\mathcal{E}_\alpha$, l'un des espaces introduits avant le Lemme (VIII.5.8), et est continue à l'origine : elle est donc partout indéfiniment différentiable puisqu'elle ne comprend que des termes linéaires et bilinéaires continus.
De plus, c'est une application linéaire si l'on fige la dernière variable h et alors elle est égale à sa différentielle partielle d'ordre 1, les dérivées supérieures s'annulant.
Par conséquent, le lemme résultera de la composition des différentielles si l'on montre la différentiabilité à tout ordre de la fonctionnelle intégrale :

$$\tilde{I} : \varphi \in \mathcal{E}_\alpha \mapsto \int \tilde{J}^*(\varphi(.,.,z))\rho(dz) \tag{VIII.5.64}$$

en prenant également le soin de vérifier qu'elle est à valeurs dans $C^{\frac{\alpha}{2},\alpha}$.
Le résultat s'obtient par récurrence et l'on commence par se fixer $k \in \mathbb{N}$ et à poser :

$$\kappa = \varphi_1 ... \varphi_k \tag{VIII.5.65}$$

avec $\varphi_i \in \mathcal{E}_\alpha,\ 1 \leq i \leq k$.
En se donnant aussi un accroissement $\psi \in \mathcal{E}_\alpha$ de φ, on prépare l'étude de la différentiabilité en formant la quantité

$$\begin{aligned} E_k(t,x) &= \int \tilde{J^*}^{(k)}\Big((\varphi+\psi)(t,x,z)\Big)\kappa(t,x,z)\rho(dz) - \int \tilde{J^*}^{(k)}(\varphi(t,x,z))\kappa(t,x,z)\rho(dz) \\ &\quad - \int_{\mathbf{R}} \tilde{J^*}^{(k+1)}(\varphi(t,x,z))\psi(t,x,z)\kappa(t,x,z)\rho(dz) \\ &= \int_{\mathbf{R}}\int_0^1 \tilde{J^*}^{(k+2)}\Big(\varphi(t,x,z) + s\psi(t,x,z)\Big)(1-s)ds\psi^2(t,x,z)\kappa(t,x,z)\rho(dz) \end{aligned} \tag{VIII.5.66}$$

et toute la suite de l'étude se ramène à celle de l'expression définie pour $(t,x) \in [0,T] \times \mathbb{R}$, et pour $\psi_0, \psi_1, \psi_2 \in \mathcal{E}_\alpha$ par :

$$\int_{\mathbf{R}} \tilde{J^*}^{(k)}\Big(\psi_0(t,x,z)\Big)\psi_1(t,x,z)\psi_2(t,x,z)\rho(dz) \tag{VIII.5.67}$$

Comme il est facile de montrer que l'on a $\tilde{J^*}^{(k)}(\psi_0) \in \mathcal{E}^0_\alpha$ avec $\|\tilde{J^*}^{(k)}(\psi_0)\|_{\mathcal{E}^0_\alpha} \leq \sup M(1+\|\psi_0\|_{\mathcal{E}_\alpha})$ où M est une constante, on en déduit, quitte à augmenter M :

$$\forall z \in \mathbf{R},\ \|\tilde{J^*}^{(k)}(\psi_0(.,.,z))\psi_1(.,.,z)\psi_2(.,.,z)\|_\alpha \leq M(1+\|\psi_0\|_{\mathcal{E}_\alpha})\|\psi_1\|_{\mathcal{E}_\alpha}\|\psi_2\|_{\mathcal{E}_\alpha} \min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\} \tag{VIII.5.68}$$

Ceci fait l'objet du Lemme VIII.5.8 énoncé plus loin.
On en déduit aisément :

$$\left\| \int_{\mathbf{R}} \tilde{J^*}^{(k)}(\psi_0(.,.,z))\psi_1(.,.,z)\psi_2(.,.,z)\rho(dz) \right\|_\alpha \leq M_k(1+\|\psi_0\|_{\mathcal{E}_\alpha})\|\psi_1\|_{\mathcal{E}_\alpha}\|\psi_2\|_{\mathcal{E}_\alpha} \tag{VIII.5.69}$$

avec $M_k = M \int_{\mathbf{R}} \min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\}$ qui est finie par hypothèse.

Ceci étant établi pour toutes fonctions $\psi_0, \psi_1, \psi_2 \in \mathcal{E}_\alpha$, on reprend la fonction $\varphi \in \mathcal{E}_\alpha$ ainsi que les fonctions $\varphi_1, ..., \varphi_k \in \mathcal{E}_\alpha$ pour $k \geq 0$.
Par ailleurs, comme $\tilde{J^*}(0) = \tilde{J^*}^{(1)}(0) = 0$, la formule de Taylor avec reste intégral permet de ramener l'étude des cas $k=0$ et $k=1$ au cas $k=2$ et d'amorcer la récurrence.
On écrit en effet :

$$\int \tilde{J^*}(\varphi(.,.,z))\rho(dz) = \int_0^1 \int \tilde{J^*}^{(2)}(s\varphi(.,.,z))(1-s)\varphi(.,.,z)^2\rho(dz)ds \tag{VIII.5.70}$$

et

$$\int \tilde{J^*}^{(1)}(\varphi(.,.,z))\varphi_1(.,.,z)\rho(dz) = \int_0^1 \int \tilde{J^*}^{(2)}(s\varphi(.,.,z))\varphi(.,.,z)\varphi_1(.,.,z)\rho(dz)ds \tag{VIII.5.71}$$

Les deux dernières égalités donnent en vertu de (VIII.5.69), si l'on choisit $\psi_0 = s\varphi$, $\psi_1 = \varphi$ et $\psi_2 = \varphi$ ou $\psi_2 = \varphi_1$:

$$\left\| \int \tilde{J^*}(\varphi(.,.,z))\rho(dz) \right\|_\alpha \leq M_2(1+\|\varphi\|_{\mathcal{E}\alpha})\|\varphi\|^2_{\mathcal{E}\alpha} \tag{VIII.5.72}$$

et

$$\left\| \int \tilde{J^*}^{(1)}(\varphi(.,.,z))\varphi_1(.,.,z)\rho(dz) \right\|_\alpha \leq M_2(1+\|\varphi\|_{\mathcal{E}\alpha})\|\varphi\|_{\mathcal{E}\alpha}\|\varphi_1\|_{\mathcal{E}\alpha} \tag{VIII.5.73}$$

où M_2 est la valeur de la constante M_k pour $k=2$.
En particulier, le membre de droite de l'inégalité (VIII.5.72) est fini, ce qui signifie que la fonctionnelle $\tilde{I}$ définie par (VIII.5.64) est bien à valeurs dans $C^{\frac{\alpha}{2},\alpha}$.
De plus, avec la majoration (VIII.5.73), l'intégrale du membre de gauche définit une application linéaire continue $\tilde{I}_1(\varphi) : \varphi_1 \in \mathcal{E}_\alpha \mapsto \tilde{I}_1(\varphi)(\varphi_1) \in C^{\frac{\alpha}{2},\alpha}$.
Avec la relation (VIII.5.66) pour $k=0$ et $\kappa=1$ et en prenant $\psi_0 = \varphi + s\psi$ et $\psi_1 = \psi_2 = \psi$ dans la majoration (VIII.5.69), on obtient :

$$\|\tilde{I}(\varphi+\psi) - \tilde{I}(\varphi) - \tilde{I}_1(\varphi)(\psi)\|_\alpha \leq (1+\|\psi\|_{\mathcal{E}_\alpha}\|)\|\psi\|^2_{\mathcal{E}_\alpha} \tag{VIII.5.74}$$

Avec cette relation ainsi que la continuité de $\tilde{I}_1$, on obtient que cette dernière est la différentielle de la fonctionnelle $\tilde{I}$ à valeurs dans $C^{\frac{\alpha}{2},\alpha}$.

Pour examiner les différentielles d'ordre supérieur, on choisit pour $k \geq 2,\ \psi_0 = \varphi,\ \psi_1 = \varphi_1, \psi_2 = \varphi_2 ... \varphi_k$ dans la relation (VIII.5.67) : la majoration (VIII.5.69) donne alors la relation :

$$\left\| \int \tilde{J^*}^{(k)}(\varphi(.,.,z)) \prod_{i=1}^{k} \varphi_i(.,.,z)\rho(dz) \right\|_\alpha \leq M_k(1 + \|\psi_0\|_{\mathcal{E}\alpha}) \prod_{i=1}^{k} \|\varphi_i\|_{\mathcal{E}\alpha} \qquad \text{(VIII.5.75)}$$

où l'on a utilisé la relation $\left\| \prod_{i=2}^{k} \varphi_i \right\|_{\mathcal{E}\alpha} \leq \prod_{i=2}^{k} \|\varphi_i\|_{\mathcal{E}\alpha}$ et cela prouve que l'intégrale du membre de gauche de (VIII.5.75) définit une application multilinéaire symétrique continue $\tilde{I}_k(\varphi)$: $(\varphi_i)_{1 \leq i \leq k} \in \mathcal{E}_\alpha^k \mapsto \tilde{I}_k(\varphi)(\varphi_1, ..., \varphi_k) \in C^{\frac{\alpha}{2},\alpha}$ pour $k \geq 2$.
De manière similaire à (VIII.5.74) mais en prenant κ donné par (VIII.5.65) dans l'inégalité (VIII.5.66), on peut établir pour $k \geq 1$

$$|||\tilde{I}_k(\varphi + \psi) - \tilde{I}_k(\varphi) - \tilde{I}_{k+1}(\varphi)(\psi)|||_\alpha \leq (1 + \|\psi\|_{\mathcal{E}_\alpha})\|\psi\|^2_{\mathcal{E}_\alpha} \qquad \text{(VIII.5.76)}$$

où la norme triple $|||.|||_\alpha$ désigne la norme naturelle des applications multilinéaires continues de $\mathcal{E}_\alpha$ vers $C^{\frac{\alpha}{2},\alpha}$.
Ainsi, pour $k \geq 1$, avec (VIII.5.76) et la continuité de $\tilde{I}_{k+1}$, on obtient que cette dernière est la différentielle de $\tilde{I}_{k+1}$: c'est aussi la différentielle d'ordre k de $\tilde{I}$, ce qui permet d'achever la preuve du lemme. ∎

Lemme VIII.5.11 *On a l'existence d'une constante C_0 telle que*

$$\forall v \in C^{\frac{2+\alpha}{2},2+\alpha},\ \|\Delta_* v(.,.,z)\|_\alpha \leq C_0 \|v\|_{2+\alpha} \min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\} \qquad \text{(VIII.5.77)}$$

et

$$\forall v \in C^{\frac{2+\alpha}{2},2+\alpha},\ \|\Delta_* v(.,.,z)\|_0 \leq C_0 \|v\|_{2+\alpha} \Phi^2_0(z) \qquad \text{(VIII.5.78)}$$

Preuve : La dernière inégalité a déjà été vue sous une forme plus précise avec (VIII.4.5) et on démontre la première suivant le même principe en utilisant la formule de Taylor avec reste intégral pour $v \in C^{2+\alpha}$:

$$\begin{aligned} \Delta_* v(t,x,z) &= v(t, x + \ln(1 + \Phi_*(t,x,z))) - v(t,x) - \Phi_*(t,x,z)\frac{\partial v}{\partial x}(t,x) \\ &= \int_0^1 \frac{\partial^2 v}{\partial x^2}(t, x + s\ln(1 + \Phi_*(t,x,z)))(1-s)ds(\ln(1 + \Phi_*(t,x,z)))^2 \\ &\quad + B(t,x,z)\frac{\partial v}{\partial x}(t,x) \end{aligned} \qquad \text{(VIII.5.79)}$$

où l'on a noté $B(t,x,z) = \Phi_*(t,x,z) - \ln(1 + \Phi_*(t,x,z))$.
Pour poursuivre, compte tenu de la minoration (VIII.1.6), on note l'inégalité :

$$|\ln(1 + \Phi_*)| \leq M\Phi_* \qquad \text{(VIII.5.80)}$$

où M est une constante et, en appliquant l'inégalité des accroissements finis à $u \mapsto \ln(1+u)$, on remarque aussi que l'hypothèse (RH') implique :

$$(RH'') \begin{cases} |\ln(1 + \Phi_*(t,x,z)) - \ln(1 + \Phi_*(t',x',z))| \leq \frac{1}{1+\Phi_1}\Phi_\alpha(z)(|t-t'|^{\frac{\alpha}{2}} + |x - x'|^\alpha) \\ \forall (t,x),(t',x') \in [0,T] \times \mathbb{R},\ \forall z \in \mathbf{R} \end{cases}$$

Cela prouve que l'on a $\ln(1+\Phi_*) \in \mathcal{E}_\alpha$, et puisque l'on a également avec en particulier la deuxième condition de (RH') $\frac{\partial^2 v}{\partial x^2}(t, x + s\ln(1+\Phi_*(t,x,z))) \in \mathcal{E}^0_\alpha$, la relation (VIII.5.50) donne avec (VIII.5.79) :

$$
\begin{aligned}
\|\Delta_* v(.,.,z)\|_\alpha &\leq M \sup_{0\leq s\leq 1} \left\| \frac{\partial^2 v}{\partial x^2}(.,. + s\ln(1+\Phi_*)) \right\|_{\mathcal{E}^0_\alpha} \|\ln(1+\Phi_*)\|^2_{\mathcal{E}_\alpha} \min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\} \\
&\quad + \left\| B(.,.,z)\frac{\partial v}{\partial x} \right\|_\alpha && \text{(VIII.5.81)} \\
&\leq M_1 \left\| \frac{\partial^2 v}{\partial x^2} \right\|_\alpha \min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\} + \left\| \frac{\partial v}{\partial x} \right\|_\alpha \|B(.,.,z)\|_\alpha && \text{(VIII.5.82)} \\
& && \text{(VIII.5.83)}
\end{aligned}
$$

et la majoration (VIII.5.77) en résulte compte tenu de (VIII.5.56).

∎

Preuve du Lemme VIII.5.8 : On ne montre que la relation (VIII.5.50), la preuve de (VIII.5.49) se faisant suivant les mêmes lignes.
D'abord on a en fixant $z \in \mathbf{R}$

$$
\|\psi_0\psi_1\psi_2(.,.,z)\|_0 \leq \|\psi_0(.,.,z)\|_{\mathcal{E}^0_\alpha} \|\psi_1(.,.,z)\|_{\mathcal{E}_\alpha} \|\psi_2(.,.,z)\|_{\mathcal{E}_\alpha} \min\{\Phi_\alpha(z), \Phi^2_\alpha(z)\} \qquad \text{(VIII.5.84)}
$$

Pour établir cela, on a utilisé :

$$
\forall \varphi \in \mathcal{E}_\alpha,\ \forall z \in \mathbf{R},\ \|\varphi(.,.,z)\|_0 \leq \|\varphi\|_{\mathcal{E}_\alpha} \min\{1, \Phi_\alpha(z)\} \qquad \text{(VIII.5.85)}
$$

Ensuite, pour évaluer $<\psi_0\psi_1\psi_2(.,.,z)>^{(\alpha)} = <\psi_0\psi_1\psi_2(.,.,z)>^{(\frac{\alpha}{2})}_t + <\psi_0\psi_1\psi_2(.,.,z)>^{(\alpha)}_x$, il est commode d'utiliser la double inégalité, valable pour tout $v \in C^{\frac{\alpha}{2},\alpha}$

$$
\frac{1}{2} < v >^{(\alpha)} \leq \sup_{t,t'\in[0,T], x,x'\in\mathbb{R}} \frac{|v(t,x) - v(t',x')|}{|t-t'|^{\frac{\alpha}{2}} + |x-x'|^\alpha} \leq < v >^{(\alpha)} \qquad \text{(VIII.5.86)}
$$

et, en se donnant alors (t,x) et $(t',x') \in [0,T]\times\mathbb{R}$, on établit la majoration

$$
|\psi_0\psi_1\psi_2(t,x,z) - \psi_0\psi_1\psi_2(t',x',z)| \leq A'_1 + A'_2 + A'_3 \qquad \text{(VIII.5.87)}
$$

avec

$$
\begin{aligned}
A'_1 &= |\psi_0(t,x,z) - \psi_0(t',x',z)||\psi_1(t,x,z)||\psi_2(t,x,z)| \\
A'_2 &= |\psi_0(t',x',z)||\psi_1(t,x,z) - \psi_1(t',x',z)||\psi_2(t,x,z)| \\
A'_3 &= |\psi_0(t',x',z)||\psi_1(t',x',z)||\psi_2(t,x,z) - \psi_2(t',x',z)|\rho(dz)
\end{aligned}
$$

En reprenant en particulier l'inégalité de droite dans (VIII.5.86), on a :

$$\begin{aligned}
A'_1 &\leq \sup_{z'\in\mathbf{R}}\{<\psi_0(.,.,z')>^{(\alpha)}\Phi_\alpha(z')^{-1})\}(|t-t'|^{\frac{\alpha}{2}}+|x-x'|^\alpha) \\
&\quad \sup_{z'\in\mathbf{R}}\{\|\psi_1(.,.,z')\|_0\min\{1,\Phi_\alpha(z')\}^{-1}\}\sup_{z'\in\mathbf{R}}\|\psi_2(.,.,z')\|_0\min\{\Phi_\alpha(z),\Phi_\alpha^2(z)\} \\
A'_2 &\leq \sup_{z'\in\mathbf{R}}\|\psi_0(.,.,z')\|_0\sup_{z'\in\mathbf{R}}\{\|\psi_2(.,.,z')\|_0\min\{1,\Phi_\alpha(z')\}^{-1}\}\sup_{z'\in\mathbf{R}}\{<\psi_1(.,.,z')>^{(\alpha)}\Phi_\alpha(z')^{-1}\} \\
&\quad (|t-t'|^{\frac{\alpha}{2}}+|x-x'|^\alpha)\min\{\Phi_\alpha(z),\Phi_\alpha^2(z)\} \\
A'_3 &\leq \sup_{z'\in\mathbf{R}}\|\psi_0(.,.,z')\|_0\sup_{z'\in\mathbf{R}}\{\|\psi_1(.,.,z')\|_0\min\{1,\Phi_\alpha(z')\}^{-1}\}\sup_{z'\in\mathbf{R}}\{<\psi_2(.,.,z')>^{(\alpha)}\Phi_\alpha(z')^{-1}\} \\
&\quad (|t-t'|^{\frac{\alpha}{2}}+|x-x'|^\alpha)\min\{\Phi_\alpha(z),\Phi_\alpha^2(z)\}
\end{aligned}$$

Compte tenu de (VIII.5.86) ainsi que de la définition de la norme $\|.\|_{\mathcal{E}_\alpha}$ on établit l'existence d'une constante M_0 telle que :

$$\begin{aligned}
&|\psi_0\psi_1\psi_2(t,x,z)-\psi_0\psi_1\psi_2(t',x',z)| \leq \\
&M|(1+\|\psi_0\|_{\mathcal{E}_\alpha})\|\psi_1\|_{\mathcal{E}_\alpha}\|\psi_2\|_{\mathcal{E}_\alpha}(|t-t'|^{\frac{\alpha}{2}}+|x-x'|^\alpha)\min\{\Phi_\alpha(z),\Phi_\alpha^2(z)\}
\end{aligned} \quad \text{(VIII.5.88)}$$

Alors, avec (VIII.5.84) et compte tenu de l'inégalité de gauche dans (VIII.5.86), on en déduit la majoration (VIII.5.50). ∎

Preuve du Théorème VIII.5.7 : Avec l'expression (VIII.5.32), le processus optimal calibré s'écrit

$$K^{\theta^*(b_*,\mathbf{C})}(t,x,z) = (J^*)'\left(\Delta_* V^{\theta^*(b_*,\mathbf{C}),b_*}(t,x,z) - \overline{\lambda}^{(V^{\theta^*(b_*,\mathbf{C}),b_*},\theta^*(b_*,\mathbf{C}),b_*)}(t,x)\beta(t,x,z)\right)+1 \quad \text{(VIII.5.89)}$$

où l'on a noté $\overline{\lambda}^{(W,\theta,b_*)} = \tilde{\lambda}^{(W+\theta\cdot h,\theta,b_*,h)}$, en considérant que le payoff h est fixé, et où on rappelle que $\theta^*(b_*,\mathbf{C})$ est le multiplicateur de Lagrange optimal déterminé implicitement par l'équation (VIII.5.18) dans la preuve du Théorème VIII.5.5.
Comme dans cette dernière, on fait apparaître explicitement la dépendance en b_* de la solution V^{θ,b_*} du système (HJB^θ).
Les applications $(b_*,\mathbf{C}) \in \tilde{\mathbf{U}} \mapsto \theta^*(b_*,\mathbf{C}) \in \mathbb{R}^d$ et $(b_*,\mathbf{C}) \in \tilde{\mathbf{U}} \mapsto V^{\theta,b_*} \in C^{\frac{2+\alpha}{2},2+\alpha}$ sont indéfiniment différentiables (ce que l'on a établi dans la preuve du théorème précédemment cité et dans le Corollaire VIII.5.4) et de plus, avec la relation (VIII.5.36) et la définition de $\overline{\lambda}$, il existe un voisinage $\tilde{\mathbf{V}}$ de $(\mathbf{0}_{\mathbb{R}[0,T]\times\mathbb{R}},\mathbf{0}_{\mathbb{R}^d},(r_t))$ dans $C^{\frac{2+\alpha}{2},2+\alpha}\times\mathbb{R}^d\times C^{\frac{\alpha}{2},\alpha}$ tel que l'application $(W,\theta,b_*) \in \tilde{\mathbf{V}} \mapsto \overline{\lambda}^{(W,\theta,b_*)} \in C^{\frac{\alpha}{2},\alpha}$ est de classe C^∞.
En utilisant en particulier le Lemme VIII.5.11, on en déduit alors que l'application

$$(b_*,\mathbf{C}) \in \tilde{\mathbf{U}} \mapsto \overline{D}^{b_*,\mathbf{C}} = \Delta_*\overline{V}^{\theta^*(b_*,\mathbf{C}),b_*} - \overline{\lambda}^{(\overline{V}^{\theta^*(b_*,\mathbf{C}),b_*},\theta^*(b_*,\mathbf{C}),b_*)}\beta \in \mathcal{E}_\alpha \quad \text{(VIII.5.90)}$$

est indéfiniment différentiable où l'on rappelle que l'ensemble $\mathcal{E}_\alpha$ a été défini au début de la preuve du Lemme VIII.5.10.
D'autre part dans l'expression (VIII.5.89), $\tilde{\mathbf{U}}$ peut être choisi assez petit pour que l'on puisse remplacer J^* par une régularisée $\tilde{J}^*$ indéfiniment différentiable et bornée ainsi que ses dérivées et l'on peut montrer que l'application $\varphi \in \mathcal{E}_\alpha \mapsto (\tilde{J}^*)'(\varphi)$ est à valeurs dans $\mathcal{E}_\alpha$ et est infiniment

différentiable.
Pour le voir, il suffit de remarquer que $(\tilde{J}^*)'(0) = 0$ si bien que l'on a :

$$\tilde{J}^{*\prime}(\varphi) = \int_0^1 \tilde{J}^{*(2)}(s\varphi)\varphi ds \tag{VIII.5.91}$$

tandis qu'en utilisant la relation (VIII.5.49), on établit pour $\psi_0, \psi_1 \in \mathcal{E}_\alpha$ et $k \geq 0$, en notant aussi la majoration $\|\tilde{J}^{*(k)}(\psi_0)\|_{\mathcal{E}^0_\alpha} \leq M(1 + \|\psi_0\|_{\mathcal{E}_\alpha})$ où M est une constante :

$$\|\tilde{J}^{*(k)}(\psi_0)\psi_1\|_{\mathcal{E}_\alpha} \leq M_k(1 + \|\psi_0\|_{\mathcal{E}_\alpha})\|\psi_1\|_{\mathcal{E}_\alpha} \tag{VIII.5.92}$$

Avec (VIII.5.91), en choisissant $\psi_0 = s\varphi,\ 0 \leq s \leq 1,\ \psi_1 = \varphi$, on en déduit bien $\tilde{J}^{*\prime}(\varphi) \in \mathcal{E}_\alpha$ tandis que la différentiabilité s'obtient par récurrence sur $k \geq 0$ appliquée à la quantité

$$\begin{aligned} D_k &= \tilde{J}^{*(k+1)}(\varphi+\psi)\kappa - \tilde{J}^{*(k+1)}(\varphi)\kappa - \tilde{J}^{*(k+1)}(\varphi)\psi\kappa \\ &= \int_0^1 \tilde{J}^{*(k+2)}(\varphi + s\psi)(1-s)ds\psi^2\kappa \end{aligned} \tag{VIII.5.93}$$

avec $\kappa = 1$ si $k = 0$ et sinon $\kappa = \prod_{i=1}^k \varphi_i$ avec $\varphi_1, ..., \varphi_k \in \mathcal{E}_\alpha$.
Grâce une nouvelle fois à la relation (VIII.5.92), on en déduit alors que $\varphi \in \mathcal{E}_\alpha \mapsto (\tilde{J}^*)'(\varphi) \in \mathcal{E}_\alpha$ est indéfiniment différentiable, de différentielle d'ordre k en $\varphi \in \mathcal{E}_\alpha$ l'application multilinéaire $(\varphi_i)_{1\leq i\leq k} \in \mathcal{E}^k_\alpha \mapsto \tilde{J}^{*(k+1)}(\varphi)(\varphi_1, ..., \varphi_k) \in \mathcal{E}_\alpha$.
Ceci étant, par composition des applications, on en déduit que l'application :

$$(b_*, \mathbf{C}) \in \tilde{\mathbf{U}} \mapsto K^{\theta^*(b_*,\mathbf{C})} - 1 = (\tilde{J}^*)'(\overline{D}^{b_*,\mathbf{C}}) \tag{VIII.5.94}$$

est à valeurs dans $\mathcal{E}_\alpha$ et est infiniment différentiable et on voit également facilement qu'il en est de même de :

$$(b_*, \mathbf{C}) \in \tilde{\mathbf{U}} \mapsto \ln(K^{\theta^*(b_*,\mathbf{C})}) \tag{VIII.5.95}$$

Alors, avec (VIII.2.24), et (VIII.2.25), la densité de la probabilité calibrée $Q^{K^{\theta^*(b_*,\mathbf{C})},0,X_0}$ s'écrit sous la forme :

$$Z_T^{K^{\theta^*(b_*,\mathbf{C})},0,X_0} = e^{\overline{R}((\Lambda_0,\Lambda_1,\Lambda_2)(K^{\theta^*(b_*,\mathbf{C})}-1,\ln(K^{\theta^*(b_*,\mathbf{C})})))} \tag{VIII.5.96}$$

où l'on a introduit la fonctionnelle linéaire $\overline{R}$:

$$\begin{aligned} \overline{R} : (\varphi_0, \varphi_1, \varphi_2) \in (C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha \quad \mapsto \quad & \int_0^T \varphi_0(t, X^{0,X_0}_{t-})dt + \int_0^T \varphi_1(t, X^{0,X_0}_{t-})dW_t \\ & + \int_{[0,T]\times\mathbf{R}} \varphi_2(t, X^{0,X_0}_{t-}, z)(\Pi(dt,dz) - \rho(dz)dt) \end{aligned} \tag{VIII.5.97}$$

et où l'on a posé pour tout $\psi_1, \psi_2 \in \mathcal{E}_\alpha$

$$\Lambda_2(\psi_1,\psi_2) = \psi_2 \tag{VIII.5.98}$$
$$\Lambda_1(\psi_1,\psi_2) = \frac{1}{\sigma_*(.,.)}\left[r(.) - b_*(.,.) - \int_{\mathbf{R}} \Phi_*(.,.,z)\psi_1(.,.,z)d\rho(z)\right] \tag{VIII.5.99}$$
$$\Lambda_0(\psi_1,\psi_2) = \frac{1}{2}\Lambda_1^2(\psi_1,\psi_2) + \int_{\mathbf{R}} (e^{\psi_2(.,.,z)} - 1 - \psi_2(.,.,z))\rho(dz) \tag{VIII.5.100}$$

La fonction $\Lambda_2 : \mathcal{E}_\alpha^2 \to \mathcal{E}_\alpha$ est évidemment de classe C^∞ et il en est de même des fonctions $\Lambda_i : \mathcal{E}_\alpha^2 \to C^{\frac{\alpha}{2},\alpha}$, $i = 1, 2$.
Pour le voir pour Λ_1, il s'agit essentiellement de le faire pour la fonctionnelle linéaire $\psi_1 \in \mathcal{E}_\alpha \mapsto \int_{\mathbf{R}} \Phi_*(.,.,z)\psi_1(.,.,z)d\rho(z) \in C^{\frac{\alpha}{2},\alpha}$ qui est continue donc indéfiniment différentiable en raison de la majoration (VIII.5.50) qui donne

$$\forall z \in \mathbf{R},\ \|\Phi_*(.,.,z))\psi_1(.,.,z)\|_\alpha \leq M\|1\|_{\mathcal{E}_\alpha^0}\|\Phi_*\|_{\mathcal{E}_\alpha}\|\psi_1\|_{\mathcal{E}_\alpha}\min\{\Phi_\alpha(z), \Phi_\alpha^2(z)\} \qquad \text{(VIII.5.101)}$$

et la quantité à droite est finie et intégrable.
Ensuite, compte tenu de la différentiabilité à tout ordre de Λ_1, celle de Λ_0 se résume à celle de l'intégrale du membre de droite de (VIII.5.100) qui s'écrit en appliquant la formule de Taylor avec reste intégral à la fonction $u \mapsto e^u - 1 - u$ nulle en 0 ainsi que sa dérivée première $u \mapsto e^u - 1$, comme l'intégrale par rapport à ρ :

$$e^{\psi_2(.,.,z)} - 1 - \psi_2(.,.,z) = \int_0^1 e^{s\psi_2(.,.,z)}\psi_2^2(.,.,z)(1-s)ds \qquad \text{(VIII.5.102)}$$

On montre alors qu'il s'agit d'une fonction indéfiniment différentiable de $\mathcal{E}_\alpha$ dans $\mathcal{E}_\alpha$ en utilisant comme ci-dessus la majoration (VIII.5.50) et en suivant ce qui a été fait plus haut pour la fonction $\varphi \in \mathcal{E}_\alpha \mapsto (\tilde{J}^*)'(\varphi)$ puis on conclut en intégrant.
Ainsi, avec la régularité des fonctions Λ_i, $1 \leq i \leq 3$, ainsi que celle de $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}} \mapsto K^{\theta*(b_*,\mathbf{C})} - 1$ et de $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}} \mapsto \ln(K^{\theta^*(b_*,\mathbf{C})})$, le théorème sera acquis si l'on montre que la fonctionnelle $e^{\overline{R}}$ est de classe C^∞ de $(C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha$ dans $L^{q^*}(\Omega, P)$ où q^* est l'exposant conjugué de $q : q^{-1} + q_*^{-1} = 1$.
En effet, par composition avec l'application $\varphi \in L^{q^*} \mapsto \mathbb{E}^P(\varphi\psi)$ qui est de classe C^∞ car continue linéaire, on en déduira que $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}} \mapsto \mathbb{E}^P[Z_T^{K_n^{\theta_n(b_*,\mathbf{C})},0,X_0}\psi]$ est de classe C^∞.
On va prouver de manière générale que la fonctionnelle $e^{\overline{R}}$ mise en jeu dans l'expression (VIII.5.96) est à valeurs dans $L^p(\Omega, P)$ pour tout $1 \leq p < +\infty$ et de classe C^∞ (pour la norme de $L^p(\Omega, P)$) et que sa différentielle d'ordre k en $(\varphi_0, \varphi_1, \varphi_2) \in (C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha$ est l'application multilinéaire

$$\begin{aligned} \overline{R}_k^{(\varphi_0,\varphi_1,\varphi_2)} \quad &: \quad (\psi_0^1, \psi_1^1, \psi_2^1), ..., (\psi_0^k, \psi_1^k, \psi_2^k) \in (C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha \mapsto \\ &\qquad e^{\overline{R}(\varphi_0,\varphi_1,\varphi_2)} \prod_{i=1}^k \overline{R}(\psi_0^i, \psi_1^i, \psi_2^i) \end{aligned} \qquad \text{(VIII.5.103)}$$

Pour examiner ensuite la différentiabilité, suivant la preuve du Lemme VIII.5.10 et fixant $(\psi_0^i, \psi_1^1, \psi_2^i)$, $1 \leq i \leq k$, on pose $\kappa = \prod_{i=1}^k \overline{R}(\psi_0^i, \psi_1^i, \psi_2^i)$ et l'on va prouver que la fonctionnelle $e^{\overline{R}}\kappa$ prend ses valeurs dans $L^p(\Omega, P)$ et qu'elle différentiable sur $(C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha$.
On forme la quantité pour $(\psi_0, \psi_1, \psi_2) \in (C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha$:

$$\begin{aligned} Q_\kappa &= e^{\overline{R}(\varphi_0+\psi_0,\varphi_1+\psi_1,\varphi_2+\psi_2)}\kappa - e^{\overline{R}(\varphi_0,\varphi_1,\varphi_2)}\kappa - e^{\overline{R}(\varphi_0,\varphi_1,\varphi_2)}\overline{R}(\psi_0,\psi_1,\psi_2)\kappa \\ &= \int_0^1 e^{\overline{R}(\varphi_0+s\psi_0,\varphi_1+s\psi_1,\varphi_2+s\psi_2)}(\overline{R}(\psi_0,\psi_1,\psi_2))^2\kappa(1-s)ds \end{aligned} \qquad \text{(VIII.5.104)}$$

En appliquant l'inégalité de Hölder, on obtient la majoration :

$$\mathbf{E}^P(Q_\kappa^p) \leq \sup_{0\leq s\leq 1}\left[\mathbf{E}^P(e^{3p\overline{R}(\varphi_0+s\psi_0,\varphi_1+s\psi_1,\varphi_2+s\psi_2)})\right]^{\frac{1}{3}}\left[\mathbf{E}^P(\overline{R}(\psi_0,\psi_1,\psi_2))^{6p})\right]^{\frac{1}{3}}\left[\mathbf{E}^P(\kappa^{3p})\right]^{\frac{1}{3}} \qquad \text{(VIII.5.105)}$$

On écrit alors pour le premier facteur du second membre

$$\mathbf{E}^P(e^{3p\overline{R}(\varphi_0+s\psi_0,\varphi_1+s\psi_1,\varphi_2+s\psi_2)}) \leq e^{3pT(\|\varphi_0\|_0+\|\psi_0\|_0)}\mathbf{E}^P\left(e^{3pM_T^{(s)}}\right) \qquad \text{(VIII.5.106)}$$

en définissant la martingale

$$M_t^{(s)} = \int_0^T \int_{\mathbf{R}} (\varphi_1+s\psi_1)(t, X_{t^-}^{0,X_0}, z)\rho(dz)dW_t + \int_{[0,T]\times\mathbf{R}} (\varphi_2+s\psi_2)(t, X_{t^-}^{0,X_0}, z)(\Pi(dt,dz)-\rho(dz)dt)$$

et le deuxième facteur du membre de droite de (VIII.5.106) est uniformément majoré par rapport à s en vertu du Lemme VIII.2.4 puisque la martingale quasi-continue à gauche $M_t^{(s)}$ a ses sauts bornés par $\|\varphi_2\|_{0,\mathcal{E}_\alpha} + \|\psi_2\|_{0,\mathcal{E}_\alpha}$ et que l'on a :

$$< M^{(s)}, M^{(s)} >_t \leq 2[\|\varphi_1\|_0^2 + \|\psi_1\|_0^2 + (\|\varphi_2\|_{0,\mathcal{E}_\alpha}^2 + \|\psi_2\|_{0,\mathcal{E}_\alpha}^2)\int_{\mathbf{R}} \min\{\Phi_\alpha(z), \Phi_\alpha^2(z)\}\rho(dz)]t$$

En fait, on peut même écrire d'après ce lemme, les fonctions $\varphi_0,\ \varphi_1,\ \varphi_2$ étant fixées :

$$\sup_{0\leq s\leq 1}\left[\mathbf{E}^P(e^{3p\overline{R}(\varphi_0+s\psi_0,\varphi_1+s\psi_1,\varphi_2+s\psi_2)})\right] = \zeta((\psi_i)_{0\leq i\leq 2}) \qquad \text{(VIII.5.107)}$$

où $\zeta((\psi_i)_{0\leq i\leq 2})$ reste borné lorsque $\|\psi_0\|_0 + \|\psi_1\|_0 + \|\psi_2\|_{0,\mathcal{E}_\alpha}$ est borné.
Pour traiter le deuxième facteur dans la relation (VIII.5.105) et obtenir une majoration du moment d'ordre p', $1 \leq p' < \infty$ de $\overline{R}(\psi^0, \psi_1, \psi_2)$, on reprend la définition (VIII.5.97) : avec l'inégalité $(|A_1| + |A_2| + |A_3|)^{p'} \leq 3^{p'-1}(|A_1|^{p'} + |A_2|^{p'} + |A_3|^{p'})$ où $A_i,\ 1 \leq i \leq 3$ sont les trois termes de $\overline{R}(\psi^0, \psi_1, \psi_2)$, et en appliquant l'inégalité de Jensen pour A_1, on obtient :

$$\mathbb{E}^P(\overline{R}(\psi_0,\psi_1,\psi_2)^{p'}) \leq 3^{p'-1}(T^{p'-1}\|\psi_0\|_0^{p'} + \|\psi_1\|_0^{p'}|M^{(1)}|^{p'} + \|\psi_2\|_{0,\mathcal{E}_\alpha}^{p'}|M^{(2)}|^{p'} \qquad \text{(VIII.5.108)}$$

où $M^{(1)}$ est une martingale continue telle que $< M^{(1)}, M^{(1)} >_t \leq t$ et $M^{(2)}$ est une martingale quasi-continue à gauche, dont les sauts sont bornés par 1 et telle que $< M^{(2)}, M^{(2)} >_t \leq t$.
En appliquant encore une fois la majoration exponentielle du Lemme VIII.2.4, on peut donc trouver une constante C_0 telle que

$$\mathbb{E}^P(\overline{R}(\psi_0,\psi_1,\psi_2)^{p'}) \leq C_0(\|\psi_0\|_0 + \|\psi_1\|_0 + \|\psi_2\|_{0,\mathcal{E}_\alpha}))^{p'} \qquad \text{(VIII.5.109)}$$

Cette relation fournit la majoration du second facteur du membre de droite de (VIII.5.105) avec p'=6p, tandis qu'avec l'inégalité de Hölder, elle permet de majorer également le dernier facteur.
Tout compte fait, la majoration (VIII.5.105) donne :

$$[\mathbf{E}^P(Q_\kappa^p)]^{\frac{1}{p}} \leq (\|\psi_0\|_0+\|\psi_1\|_0+\|\psi_2\|_{0,\mathcal{E}^\alpha})^2\prod_{i=1}^k(\|\psi_0^i\|_0+\|\psi_1^i\|_0+\|\psi_2^i\|_{0,\mathcal{E}_\alpha})\zeta((\psi_i)_{0\leq i\leq 2}) \quad \text{(VIII.5.110)}$$

où $\zeta((\psi_i)_{0\leq i\leq 2})$ reste borné lorsque $\|\psi_0\|_0 + \|\psi_1\|_0 + \|\psi_2\|_{0,\mathcal{E}_\alpha}$ est borné.
D'autre part, selon les mêmes lignes, avec l'inégalité de Hölder, la relation (VIII.5.109) ainsi que la relation (VIII.5.107) avec $\psi_i = 0,\ 0 \leq i \leq 2$, on établit :

$$[\mathbf{E}^P(e^{\overline{R}(\varphi_0,\varphi_1,\varphi_2)}\kappa)^p]^{\frac{1}{p}} \leq C_k\prod_{i=1}^k(\|\psi_0^i\|_0 + \|\psi_1^i\|_0 + \|\psi_2^i\|_{0,\mathcal{E}_\alpha}) \qquad \text{(VIII.5.111)}$$

où C_k est une constante.
En prenant k=0 dans (VIII.5.111) , ou en utilisant simplement la relation (VIII.5.107), on montre que l'application $e^{\overline{R}}$ est à valeur dans $L^p(\Omega, P)$ tandis que pour $k \geq 1$, on établit la continuité de l'application k-linéaire $\overline{R}_k^{(\varphi_0,\varphi_1,\varphi_2)}$ définie par (VIII.5.103).
Grâce à (VIII.5.110), on voit donc par récurrence que la fonctionnelle $e^{\overline{R}}$ est de classe C^∞ de $(C^{\frac{\alpha}{2},\alpha})^2\times\mathcal{E}_\alpha$ dans $L^p(\Omega, P)$ pour $1 \leq p < +\infty$ et en particulier pour $p = q^*$, et que sa différentielle d'ordre k en $(\varphi_0, \varphi_1, \varphi_2) \in (C^{\frac{\alpha}{2},\alpha})^2 \times \mathcal{E}_\alpha$ est l'application multilinéaire $\overline{R}_k^{(\varphi_0,\varphi_1,\varphi_2)}$, ce qu'il fallait démontrer pour terminer la preuve. ∎

VIII.6 Cas où la mesure de Lévy est bornée

VIII.6.1 Contrôle dépendant explicitement de la taille des sauts

Dans cette partie, on examine brièvement le cas où l'on suppose que la mesure de Lévy ρ vérifie $\rho(\mathbf{R}) < +\infty$, ce qui permet de prendre prendre $\psi_0 \equiv 1$ dans la définition de $\mathcal{K}^{\overline{H}}$. On recherchera donc l'intensité optimale dans

$$\mathcal{K}_b^{\overline{H}} \quad = \quad \left\{ K(t,x,z) \text{ fonctions Boréliennes sur } [0,T] \times \mathbb{R} \times \mathbf{R} \text{ telles que } |\ln(K)| \leq \ln(\overline{H}) \right\}$$

Envisagé sous forme pénalisée et toujours en supposant que le payoff f_* est du type intégral (VIII.4.1), le problème est donc :

$$(P_{b\chi^*}^{\overline{H},\mathbf{C}}) \quad \text{trouver } K \in \mathcal{K}_b^{\overline{H}} \text{ qui minimise } H(Q^{K,0,X_0} \| P) + \chi^*(\textstyle\int f_*(X) dQ^{K,0,X_0} - \mathbf{C})$$

Les résultats généraux s'appliquent bien entendu et l'on se contente de redonner le Théorème VIII.5.5 :

Théorème VIII.6.1 *Sous les conditions (RH) et (RH'), il existe un voisinage $\tilde{\mathbf{U}}$ de $(r, \mathbf{C}_r)$ dans $C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d$ tel que pour tout $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}}$, le problème $(P_{b\chi^*}^{\overline{H},\mathbf{C}})$ admet une solution.*

VIII.6.2 Contrôle ne dépendant pas explicitement de la taille des sauts

A présent, on envisage la situation où l'on restreint l'ensemble des contrôles à $\mathcal{K}_u^{\overline{H}}$ avec :

$$\mathcal{K}_u^{\overline{H}} \quad = \quad \left\{ K(t,x) \text{ fonctions Boréliennes sur } [0,T] \times \mathbb{R} \text{ telles que } |\ln(K)| \leq \ln(\overline{H}) \right\}$$

Le problème est donc

$$(P_{u\chi^*}^{\overline{H},\mathbf{C}}) \quad \text{trouver } K \in \mathcal{K}_u^{\overline{H}} \text{ qui minimise } H(Q^{K,0,X_0} \| P) + \chi^*(\textstyle\int f_*(X) dQ^{K,0,X_0} - \mathbf{C})$$

et, toujours sous l'hypothèse restrictive (RS), on doit résoudre le système

$$(HJB_u^\theta) \begin{cases} \frac{\partial W}{\partial t} + \sup_{K \in \mathcal{K}_u^{\overline{H}}} \left(\mathcal{G}^K(W) - \eta(K) \right) + \theta \cdot \bar{l} = 0 \text{ sur }]0,T[\times \mathbb{R} \\ W(T,x) = \theta \cdot h(x) \text{ sur } \mathbb{R} \end{cases}$$

On peut établir un résultat semblable au Théorème VIII.4.3 à condition de vérifier l'existence de l'optimum de l'Hamiltonien associé au problème de contrôle.
Il est facile de s'en assurer, en partant de la preuve faite pour ce théorème, car l'optimum (VIII.4.30) doit alors être modifié en

$$\begin{aligned} K^\theta(t,x) &= \arg \max_{|\ln(k)| \leq \ln(\overline{H})} \left(\int_{\mathbf{R}} U^\theta(t,x,z) \rho(dz) k - \eta(k)(t,x) \right) \\ &= \arg \max_{|\ln(k+1)| \leq \ln(\overline{H})} \left(\int_{\mathbf{R}} U^\theta(t,x,z) \rho(dz) k - \eta_0(k)(t,x) \right) \end{aligned} \tag{VIII.6.1}$$

cette dernière réécriture n'étant appliquée que pour retrouver des notations précédentes, et l'on rappelle que l'on a posé $\eta_0(u) = \eta(u+1)$.

Le membre de droite définit la conjuguée en $\int_{\mathbf{R}} U^\theta(t,x,z)\rho(dz) \in \mathbb{R}$ de l'entropie $k \in [\overline{H}^{-1} - 1, \overline{H} - 1] \mapsto \eta_0^*(k)(t,x)$, ce qui s'écrit en adaptant (VIII.4.35) et avec une notation abusive :

$$\begin{aligned}
\eta_0^*\Big(\int_{\mathbf{R}} U^\theta(t,x,z)\rho(dz)\Big)(t,x) &= \sup_{\overline{H}^{-1}-1\le k\le \overline{H}-1} \Big[\int kU^\theta(t,x,z)\rho(dz) - \eta_0(k)(t,x)\Big] \\
&= \sup_{k\in\mathbb{R}} \Big\{ k\int U^\theta(t,x,z)\rho(dz) \\
&\quad -J(k)\int \rho(dz) - \frac{1}{2}\Big[\alpha(t,x) + k\int \beta(t,x,z)\rho(dz)\Big]^2\Big\}
\end{aligned} \tag{VIII.6.2}$$

où la fonction J est définie par (VIII.4.32).
On en déduit que $K^\theta(t,x)$ existe bien comme maximum d'une fonction continue sur un compact, l'unicité étant également acquise par la concavité stricte de cette fonction.
La mesurabilité peut également se vérifier suivant le procédé dichotomique employé dans la preuve du Lemme VIII.4.9 car, de façon plus précise, le problème d'optimisation posé par (VIII.4.37) et (VIII.4.38) conduit à examiner

$$\begin{aligned}
\inf_{\lambda\in\mathbb{R}}\sup_{k\in\mathbb{R}} L(k,\lambda) &= \inf_{\lambda\in\mathbb{R}}\Big\{\sup_{k\in\mathbb{R}}\Big[k\int (U^\theta(t,x,z) - \lambda\beta(t,x,z))\rho(dz) - \rho(\mathbf{R})J(k)\Big] \\
&\quad -\lambda\alpha(t,x) + \frac{1}{2}\lambda^2\Big\} \\
&= \inf_{\lambda\in\mathbb{R}}\Big\{\rho(\mathbf{R})J^*\Big(\int (U^\theta(t,x,z) - \lambda\beta(t,x,z))\rho(dz)(\rho(\mathbf{R}))^{-1}\Big) \\
&\quad -\lambda\alpha(t,x) + \frac{1}{2}\lambda^2\Big\}
\end{aligned} \tag{VIII.6.3}$$

De manière similaire aux relations (VIII.4.48) et (VIII.4.49), on obtient le λ optimal sous forme implicite :

$$\lambda^{(\theta,t,x)} = (K^\theta(t,x) - 1)\int \beta(t,x,z)\rho(dz) + \alpha(t,x) \tag{VIII.6.4}$$

avec :

$$K^\theta(t,x) - 1 = (J^*)'\Big(\int (U^\theta(t,x,z) - \lambda^{(\theta,t,x)}\beta(t,x,z))\rho(dz)(\rho(\mathbf{R}))^{-1}\Big) \tag{VIII.6.5}$$

qui est bien l'optimum cherché puisque l'on peut montrer facilement l'égalité (VIII.4.53) sous la forme

$$\sup_{\ln(k)\le\ln(\overline{H})}\Big(\int U^\theta(t,x,z)\rho(dz)k - \eta_0(k)(t,x)\Big) = \int U^\theta(t,x,z)(K^\theta(t,x)-1)\rho(dz) - \eta_0(K^\theta - 1)(t,x) \tag{VIII.6.6}$$

La méthode du Lemme VIII.4.9 peut alors être utilisée pour obtenir la mesurabilité de K^θ qui est donc bien un élément de $\mathcal{K}_u^{\overline{H}}$.
On peut alors énoncer

Théorème VIII.6.2 *Faisons l'hypothèse* (RS) *sur les fonctions de payoff et supposons que* (HJB_u^θ) *admette une solution* W^θ *continue sur* $[0,T]\times\mathbb{R}$*, de classe* $C^{1,2}$ *sur* $]0,T[\times\mathbb{R}$ *et satisfaisant de plus la condition de croissance exponentielle*

$$|W^\theta(t,x)| \le \gamma(1 + |x| + e^x) \tag{VIII.6.7}$$

où γ est une constante positive, indépendante de $t \in [0,T]$.
Alors le système d'équations (VIII.6.5) et (VIII.6.4) admet une solution unique $K^\theta \in \mathcal{K}_u^{\overline{H}}$ telle que l'on a :

$$\begin{aligned}\forall (t,x) \in [0,T] \times \mathbb{R},\ W^\theta(t,x) &= \sup_{K \in \mathcal{K}_u^{\overline{H}}} \left\{ \theta \cdot \mathbb{E}^{Q^{K,t,x}} \left[\int_t^T \bar{l}(X_s^{t,x}) ds + h(X_T^{t,x}) \right] - H(Q^{K,t,x} \| P) \right\} \\ &= \theta \cdot \mathbb{E}^{Q^{K^\theta,t,x}} \left[\int_t^T \bar{l}(X_s^{t,x}) ds + h(X_T^{t,x}) \right] - H(Q^{K^\theta,t,x} \| P)\end{aligned}$$

En particulier, on obtient

$$V(\theta) \ = \ \theta \cdot \mathbb{E}^{Q^{K^\theta,t,x}} \left[\int_t^T (X_s^{t,x}) ds + h(X_T^{t,x}) \right] - H(Q^{K^\theta,t,x} \| P)$$

L'étude de la régularité s'adapte ensuite sous les hypothèses Höldériennes d'usage, compte tenu du fait que l'intensité optimale est obtenue en résolvant le système d'équations (VIII.6.4) et (VIII.6.5).
On obtient alors une version adaptée de la Proposition VIII.5.2 et du Théorème VIII.5.5 :

Proposition VIII.6.3 *On fait les hypothèses (RH) et (RH') et on considère une fonction vectorielle $\tilde{h} \in (C^{2+\alpha})^d$.*
On peut alors trouver un voisinage ouvert $\mathbf{U}$ de $(\mathbf{0}_{\mathbb{R}^d}, r, \tilde{h})$ dans $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha} \times (C^{2+\alpha})^d$ tel que pour $(\theta, b_*, h) \in \mathbf{U}$, le système (HJB_u^θ) admet une solution $W^\theta \in C^{\frac{2+\alpha}{2},2+\alpha}$ et telle que la solution K^θ du système d'équations (VIII.6.5) et (VIII.6.4) est un élément de $\mathcal{K}_u^{\overline{H}}$.
De plus, quitte à réduire $\mathbf{U}$, la fonction $(\theta, b_*, h) \in \mathbf{U} \mapsto W^\theta \in C^{\frac{2+\alpha}{2},2+\alpha}$ est indéfiniment différentiable.

et

Théorème VIII.6.4 *Sous les conditions (RH) et (RH'), il existe un voisinage $\tilde{\mathbf{U}}$ de $(r, \mathbf{C}_r)$ dans $C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d$ tel que pour tout $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}}$, le problème $(P_{u\chi^*}^{\overline{H},\mathbf{C}})$ admet une solution.*

VIII.7 Résolution du problème de calibration pour des fonctions de payoff non délocalisées

Dans cette partie, on se propose de résoudre le problème dans le cas où la fonction de payoff des options de calibration n'est pas du type délocalisé (VIII.4.1) mais est constituée d'une suite de flux discrets, ce qui correspond à la situation réelle où il s'agit de Put et de Call européens.
Dans ce cas, les options de calibration ne versent chacune qu'un seul flux à une seule maturité (options "plain vanilla") si bien que le payoff total (supposé actualisé) se décompose en un flux (vectoriel) $h(X_T)$ versé à la maturité T, ainsi qu'une suite $h_i(X_{T_i})$ de flux intermédiaires versés à p maturités $0 < T_1 < ... < T_p < T$.
De façon précise, on découpe l'intervalle d'indices entiers $\{1 \leq j \leq d\}$ en une partition $(I_i, 1 \leq i \leq p) \bigcup \{I\}$ telle que si $(e_j, 1 \leq j \leq d)$ désigne la base canonique de $\mathbb{R}^d$, alors on pose pour tout $x \in \mathbb{R}$

$$\forall 1 \leq i \leq p,\ h_i(x) = \sum_{j \in I_i} h_i^j(x) e_j$$

et

$$h(x) = \sum_{j \in I} h^j(x) e_j$$

ce qui permet d'écrire, pour le processus $(X_t, 0 \leq t \leq T)$:

$$f_*(X) = \sum_{i=1}^{p} h_i(X_{T_i}) + h(X_T) \tag{VIII.7.1}$$

qui est la version non délocalisée de (VIII.4.1).
En considérant le problème de calibration pénalisée $(P^{\overline{H},\mathbf{C}})$ dans le cas d'une mesure de Lévy ρ générale, on l'attaque via l'équation de la programmation dynamique.

VIII.7.1 Formulation HJB pour les fonctions de payoff non délocalisées

Pour adapter le Théorème VIII.4.3, on commence par faire les hypothèses de régularité suivantes :

$$(\widehat{RS}) \begin{cases} \text{Les fonctions de payoff } h_i,\ 1 \leq i \leq p \text{ et } h \text{ vérifient la condition de croissance exponentielle :} \\ \|h(x)\| + \sum_{i=1}^{p} \|h_i(x)\| \leq \kappa(1 + |x| + e^x) \\ \text{avec } \kappa > 0 \end{cases}$$

En effet, la fonction valeur naturellement associée au problème de calibration peut se définir comme :

$$\forall (t,x) \in [0,T] \times \mathbb{R},\ \hat{V}^\theta(t,x) = \sup_{K \in \mathcal{K}^{\overline{H}}} \left\{ \theta \cdot \mathbb{E}^{Q^{K,t,x}} \left[\sum_{i=1}^{p} 1_{\{t \leq T_i\}} h_i(X_{T_i}^{t,x}) + h(X_T^{t,x}) \right] - H(Q^{K,t,x} \| P) \right\} \tag{VIII.7.2}$$

et l'on a bien sûr

$$V(\theta) = \hat{V}^\theta(0, X_0)$$

où l'on rappelle que $V(\theta)$ est définie par (VIII.2.27).
La définition (VIII.7.2) a un sens puisque l'on a le résultat suivant, prouvé comme dans le Lemme VIII.4.1 :

Lemme VIII.7.1 *Sous l'hypothèse $(\widehat{RS})$, on a la condition d'intégrabilité suivante*

$$\forall K \in \mathcal{K}^{\overline{H}},\ \mathbb{E}^{Q^{K,t,x}} \left[\sum_{i=1}^{p} \|h_i\|(X_{T_i}^{t,x}) 1_{\{t \leq T_i\}} + \|h\|(X_T^{t,x}) \right] < +\infty \tag{VIII.7.3}$$

qui fournit en particulier l'égalité ensembliste

$$\mathcal{K}_0^{\overline{H}} = \mathcal{K}^{\overline{H}} \tag{VIII.7.4}$$

La fonction valeur $\hat{V}^\theta$ est alors recherchée comme solution du système :

$$(\widehat{HJB}^\theta) \begin{cases} \frac{\partial W}{\partial t} + \sup_{K \in \mathcal{K}^{\overline{H}}} \Big(\mathcal{G}^K(W) - \eta(K) \Big) = 0 \text{ sur }]T_{i-1}, T_i[\times \mathbb{R},\ 1 \leq i \leq p \\ W(T_i, x) = \begin{cases} W(T_i+, x) + \theta \cdot h_i(x) \text{ sur } \mathbb{R} \text{ si } 1 \leq i \leq p \\ \theta \cdot h(x) \text{ sur } \mathbb{R} \text{ si } i = p+1 \end{cases} \end{cases}$$

en posant $T_{p+1} = T$ et en imposant $T_p < T$.
On peut écrire formellement

$$(\widehat{HJB}^\theta) \Big\{ \frac{\partial W}{\partial t} + \sup_{K \in \mathcal{K}^{\overline{H}}} \Big(\mathcal{G}^K(W) - \eta(K) \Big) = \theta \cdot \Big(\sum_{i=1}^{p} h_i(x)\delta(t - T_i) + h(x)\delta(t - T) \Big) \text{ sur } [0,T]$$

On peut montrer que sous certaines hypothèses de régularité, et sous réserve d'existence, la solution $\widehat{W}^\theta$ de $(\widehat{HJB}^\theta)$ est effectivement la fonction valeur $\hat{V}^\theta$ en commençant par noter

$$U_0^\theta(t,x,z) \quad = \quad \widehat{W}^\theta(t, x + \ln(1+\Phi_*(t,x,z))) - W^\theta(t,x) - \Phi_*(t,x,z)\frac{\partial W^\theta}{\partial x} \tag{VIII.7.5}$$

et par définir implicitement

$$K^\theta(t,x,z) - 1 = (J^*)'\Big(U_0^\theta(t,x,z) - \lambda^{(\theta,t,x)}\frac{\Phi_*(t,x,z)}{\sigma_*(t,x)}\Big) \tag{VIII.7.6}$$

avec

$$\lambda^{(\theta,t,x)} = \frac{1}{\sigma_*(t,x)}\Big(\int (K^\theta(t,x,z)-1)\Phi_*(t,x,z)\rho(dz) + b_*(t,x) - r_t\Big) \tag{VIII.7.7}$$

où $(J^*)'$ a pour expression (VIII.4.16).
Pour identifier la solution $\widehat{W}^\theta$ à la fonction valeur $\hat{V}^\theta$ et afin de ramener cette solution (non unique a priori) à la régularité requise au Théorème VIII.4.3, on suppose qu'elle est de classe $C^{1,2}$ dans chaque pavé $]T_{i-1},T_i[\times\mathbb{R},\ 1\le i\le p+1$, à croissance exponentielle et qu'elle est prolongeable en une fonction continue sur $[T_{i-1},T_i]\times\mathbb{R},\ 1\le i\le p+1$.
En faisant également l'hypothèse que la solution K^θ du système d'équations (VIII.7.6) et (VIII.7.7) est un élément de $\mathcal{K}^{\overline{H}}$, on vérifie alors que $\widehat{W}^\theta$ résout les p+1 problèmes de contrôle pour $1\le i\le p+1$

$$(\widehat{HJB}^\theta_{(i)})\begin{cases} \frac{\partial W}{\partial t} + \sup_{K\in\mathcal{K}^{\overline{H}}}\Big(\mathcal{G}^K(W) - \eta(K)\Big) = 0 \text{ sur }]T_{i-1},T_i[\times\mathbb{R} \\ W(T_i,x) = \hat{V}^\theta(T_i,x) \end{cases}$$

On commence par examiner le problème $(\widehat{HJB}^\theta_{(p+1)})$ pour lequel c'est acquis car la condition terminale $\hat{V}^\theta(T,x)$ s'écrit $\theta\cdot h(x)$: la fonction $\widehat{W}^\theta$ est donc solution de $(\widehat{HJB}^\theta_{(p+1)})$ sur $]T_p,T_{p+1}[$.
En appliquant le Théorème VIII.4.3, avec la condition $K^\theta\in\mathcal{K}_b^{\overline{H}}$ ainsi que les hypothèses portant sur la régularité du payoff h par $(\widehat{RS})$ et celle supposée de $\widehat{W}^\theta$, cette dernière est égale à la fonction valeur $\hat{V}^\theta$ dans l'intervalle semi-ouvert $]T_p,T]$.
Si l'on passe au problème $(\widehat{HJB}^\theta_{(p)})$, la condition terminale $\hat{V}^\theta(T_p,x) = \hat{V}^\theta(T_p+,x)+\theta\cdot h_p(x)$ se réécrit donc $\widehat{W}^\theta(T_p+,x)+\theta\cdot h_p(x)$, ce qui montre que $\widehat{W}^\theta$ est également solution de $(\widehat{HJB}^\theta_{(p)})$ sur $]T_{p-1},T_p[$.
En poursuivant la récurrence, on prouve donc que $\widehat{W}^\theta$ est solution des $p+1$ systèmes $(\widehat{HJB}^\theta_{(i)})$, $1\le i\le p+1$ et est égale à la fonction valeur $\hat{V}^\theta$.
On énonce

Théorème VIII.7.2 *Faisons l'hypothèse* $(\widehat{RS})$ *sur les fonctions de payoff et supposons que* $(\widehat{HJB}^\theta)$ *admette une solution* $\widehat{W}^\theta$ *de classe* $C^{1,2}$ *sur chacun des pavés ouverts* $]T_{i-1},T_i[\times\mathbb{R},\ 1\le i\le p+1$ *qui soit également prolongeable en une fonction continue sur* $[T_{i-1},T_i]\times\mathbb{R},\ 1\le i\le p+1$, *satisfaisant de plus la condition de croissance exponentielle*

$$|\widehat{W}^\theta(t,x)| \le \gamma(1+|x|+e^x) \tag{VIII.7.8}$$

où γ *est une constante positive, indépendante de* $t\in[0,T]$.
Alors le système d'équations (VIII.4.14) et (VIII.4.15) admet une solution unique $K^\theta\in\mathcal{K}^{\overline{H}}$

et c'est un élément de $\mathcal{K}_b^{\overline{H}}$*, c'est-à-dire une fonction Borélienne sur* $[0,T]\times\mathbb{R}\times\mathbf{R}$ *telle que* $|\ln(K)|\leq\ln(\overline{H})$*.*
Si de plus $K^\theta\in\mathcal{K}^{\overline{H}}$*, alors*

$$\forall(t,x)\in[0,T]\times\mathbb{R},\ \hat{V}^\theta(t,x)=\widehat{W}^\theta(t,x)=\theta\cdot\mathbb{E}^{Q^{K^\theta},t,x}\left[\sum_{i=1}^p 1_{\{t\leq T_i\}}h_i(X_{T_i}^{t,x})+h(X_T^{t,x})\right]-H(Q^{K^\theta,t,x}\|P)$$

où $\hat{V}^\theta$ *est donnée par (VIII.7.2).*
En particulier, à l'origine, on a

$$V(\theta)=\widehat{W}^\theta(0,X_0)=\theta\cdot\mathbb{E}^{Q^{K^\theta},0,X_0}\left[\sum_{i=1}^p h_i(X_{T_i}^{t,x})+h(X_T^{t,x})\right]-H(Q^{K^\theta,t,x}\|P)$$

où $V(\theta)$ *est défini par (VIII.2.27).*

VIII.7.2 Régularité de la solution du système HJB

Pour que le système $(\widehat{HJB}^\theta)$ possède une solution régulière, on est conduit à adopter des hypothèses de régularité Hölderienne mais, bien entendu, la solution sera sujette à être discontinue aux différentes maturités, ce qui avait motivé la technique de délocalisation.
On va d'abord revenir au système $(\widehat{HJB}^\theta)$ en le réécrivant comme la succession de $p+1$ problèmes $1\leq i\leq p+1$

$$(\widehat{HJB}^\theta_{(i)*})\begin{cases}\frac{\partial W}{\partial t}+\sup_{K\in\mathcal{K}^{\overline{H}}}\Big(\mathcal{G}^K(W)-\eta(K)\Big)=0 \text{ sur }]T_{i-1},T_i[\times\mathbb{R}\\ W(T_i,x)=\begin{cases}W(T_i+,x)+\theta\cdot h_i(x)\text{ sur }\mathbb{R}\text{ si }1\leq i\leq p\\ \theta\cdot h(x)\text{ sur }\mathbb{R}\text{ si }i=p+1\end{cases}\end{cases}$$

avec $T_{p+1}=T$ et en imposant $T_p<T$.
Chacun de ces systèmes est du type (HJB^θ) ce qui permet de lui appliquer la Proposition VIII.5.2 sous réserve de remplacer la condition (RH) par :

$$(\widehat{RH})\begin{cases}\sigma_*^2,b_*,r_t\in C^{\frac{\alpha}{2},\alpha}\\ h_i\in(C^{2+\alpha})^d,\ 1\leq i\leq p\text{ et }h\in(C^{2+\alpha})^d\end{cases}$$

et de conserver l'hypothèse (RH') : il existe Φ_α $(\mathbf{R},\rho)$-mesurable et $d_{\Phi_*}>0$ telles que $\forall(t,x),(t',x')\in[0,T]\times\mathbb{R},\ \forall z\in\mathbf{R},\ \forall 0\leq s\leq 1$

$$(RH')\begin{cases}|\Phi_*(t,x,z)-\Phi_*(t',x',z)|\leq\Phi_\alpha(z)(|t-t'|^{\frac{\alpha}{2}}+|x-x'|^\alpha)\\ |[x+s\ln(\Phi_*(t,x,z))]-[x'+s\ln(\Phi_*(t',x',z))]|\leq d_{\Phi_*}(|t-t'|^{\frac{1}{2}}+|x-x'|)\\ \int_{\mathbf{R}}\min\{\Phi_\alpha^2(z),\Phi_\alpha(z)\}\rho(dz)\leq d_{\Phi_*}\\ \Phi_\alpha\geq\Phi_0\end{cases}$$

Remarque VIII.7.3 *On pourrait affaiblir la condition* $(\widehat{RH})$ *en n'imposant de régularité que dans chaque pavé* $]T_{i-1},T_i[\times\mathbb{R},\ 1\leq i\leq p+1$ *mais cela serait artificiel et finalement non désirable pour les mêmes raisons qui ont conduit à la délocalisation.*

En examinant le système $(\widehat{HJB}^\theta_{(p+1)})$, on est dans les conditions d'application de la proposition précédemment citée qui montre alors que l'on dispose d'une solution $\widehat{W}^\theta_{(p+1)}$ qui définit une application indéfiniment différentiable de $(\theta,b_*)\in\mathbf{U}\mapsto\widehat{W}^\theta_{(p+1)}\in C^{\frac{2+\alpha}{2},2+\alpha}(]T_p,T[\times\mathbb{R})$ où $\mathbf{U}$ est

un un voisinage ouvert de $(\mathbf{0}_{\mathbb{R}^d}, r)$ dans $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha}$.
En passant ensuite au problème $(\widehat{HJB}^{\theta}_{(p)})$, la condition terminale se réécrit $\widehat{W}^{\theta}_{(p+1)}(T_p+, x) + \theta \cdot h_p(x)$ qui est donc, d'après ce qui précède une application indéfiniment différentiable de $(\theta, b_*) \in \mathbf{U}$ dans $C^{2+\alpha}$.
Alors, avec la Proposition VIII.5.2 et par composition des différentielles, on en déduit que le système $(\widehat{HJB}^{\theta}_{(p)})$ a une solution $\widehat{W}^{\theta}_{(p)}$ qui définit une application indéfiniment différentiable de $(\theta, b_*) \in \mathbf{U} \mapsto \widehat{W}^{\theta}_{(p)} \in C^{\frac{2+\alpha}{2}, 2+\alpha}(]T_{p-1}, T_p[\times\mathbb{R})$, en réduisant au besoin $\mathbf{U}$.
En poursuivant la récurrence on montre ainsi le résultat suivant :
Proposition VIII.7.4 *On fait les hypothèses* $(\widehat{RH})$ *et* (RH').
On peut alors trouver un voisinage ouvert $\mathbf{U}$ *de* $(\mathbf{0}_{\mathbb{R}^d}, r)$ *dans* $\mathbb{R}^d \times C^{\frac{\alpha}{2},\alpha}$ *tel que pour* $(\theta, b_*) \in \mathbf{U}$, *le système* $(\widehat{HJB}^{\theta})$ *admet une solution* $\widehat{W}^{\theta}$ *telle que pour* $1 \leq i \leq p$ *elle définisse par restriction à* $]T_{i-1}, T_i[\times\mathbb{R}$ *une fonction* $\widehat{W}^{\theta}_{(i)} \in C^{\frac{2+\alpha}{2}, 2+\alpha}(]T_{i-1}, T_i[\times\mathbb{R})$.
De plus, quitte à réduire $\mathbf{U}$, *la fonction* $(\theta, b_*) \in \mathbf{U} \mapsto \widehat{W}^{\theta}_{(i)} \in C^{\frac{2+\alpha}{2}, 2+\alpha}(]T_{i-1}, T_i[\times\mathbb{R})$ *est indéfiniment différentiable pour* $1 \leq i \leq p$.
Dans, ce cas, l'ensemble des résultats établis dans le cas de la fonction de payoff délocalisée $\overline{f}_{n*}$ s'adapte immédiatement.
Notamment, si l'on pose

$$\mathbf{C}_r = \mathbb{E}^P(f_*(X^r)) \tag{VIII.7.9}$$

où l'on rappelle que X^r désigne la solution de l'EDS (VIII.2.4) avec $b_* = r$, on énonce le résultat d'existence locale :
Théorème VIII.7.5 *Sous les conditions* $(\widehat{RH})$ *et* (RH'), *il existe un voisinage* $\tilde{\mathbf{U}}$ *de* $(r, \mathbf{C}_r)$ *dans* $C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d$ *tel que pour tout* $(b_*, \mathbf{C}) \in \tilde{\mathbf{U}}$, *le problème* $(P^{\overline{H},\mathbf{C}}_{\chi^*})$ *admet une solution* $K^{b_*,\mathbf{C}}$.
et on le complète en donnant la propriété de stabilité des prix d'options :
Théorème VIII.7.6 *Si* Φ *est une variable aléatoire dans* $L^q(\Omega, P)$, $1 < q < \infty$ *fixé , alors, sous les conditions* $(\widehat{RH})$ *et* (RH'), *le prix calibré* $(b_*, \mathbf{C}) \in C^{\frac{\alpha}{2},\alpha} \times \mathbb{R}^d \mapsto \mathbb{E}^{Q^{K^{b_*,\mathbf{C}}},0,X_0}(\Phi)$ *définit une fonction indéfiniment différentiable sur le voisinage* $\tilde{\mathbf{U}}$ *de* $(r, \mathbf{C}_r)$.
Remarquons que les hypothèses de régularité $(\widehat{RH})$ faites sur les options de calibration ne sont pas satisfaites dans le cas des Call et des Put.

VIII.7.3 Etude de la délocalisation des fonctions de payoff

Pour finir, on considère encore le cas où la fonction de payoff des options de calibration est de la forme (VIII.7.1) et l'on va procéder à une régularisation du système $(\widehat{HJB}^{\theta})$. On va montrer que l'on est conduit à délocaliser le payoff sous la forme (VIII.4.1).
L'hypothèse $(\widehat{RH})$ que l'on a été amené à faire pour résoudre le problème de calibration renforce en particulier les conditions $(\widehat{RS})$ et l'on étudie alors la délocalisation sous l'hypothèse

$$(\widehat{RS}_0) \begin{cases} \text{Les fonctions } h_i,\ 1 \leq i \leq p \text{ sont deux fois dérivables et} \\ \text{sont bornées ainsi que leurs dérivées et le payoff terminal h} \\ \|h(x)\| + \sum_{i=1}^p \|h_i(x)\| + \left\|\frac{\partial h_i}{\partial x}(x)\right\| + \left\|\frac{\partial^2 h_i}{\partial x^2}(x)\right\| \leq \kappa \\ \text{ avec } \kappa > 0 \end{cases}$$

et on définit la fonction :

$$\forall (t,x) \in [0,T] \times \mathbb{R},\ l(t,x) = \sum_{i=1}^{p} h_i(x) 1_{\{t \leq T_i\}} \tag{VIII.7.10}$$

On pourrait en fait faire l'étude en remplaçant la bornitude par la croissance exponentielle :

$$\|h(x)\| + \sum_{i=1}^{p} \|h_i(x)\| + \left\| \frac{\partial h_i}{\partial x}(x) \right\| + \left\| \frac{\partial^2 h_i}{\partial x^2}(x) \right\| \leq \kappa(1 + |x| + e^x)$$

Afin de faire disparaître les discontinuités aux maturités T_i que présente la fonction valeur $\hat{V}^\theta$, on effectue un changement de fonction en posant :

$$\hat{V}_0^\theta(t,x) = \hat{V}^\theta(t,x) + \theta \cdot l(0, X_0) - \theta \cdot l(t,x) \tag{VIII.7.11}$$

et l'on note la préservation du caractère $C^{1,2}$ dans les intervalles $]T_{i-1}, T_i[\times\mathbb{R}$ ainsi que celle de la valeur initiale $V(\theta)$.
Ce changement de fonction dans le système $(\widehat{HJB})$ montre que $\hat{V}_0^\theta$ est solution du système :

$$(\widehat{HJB}_0^\theta) \begin{cases} \frac{\partial W}{\partial t} + \sup_{K \in \mathcal{K}^{\overline{H}}} \Big(\mathcal{G}^K(W + \theta \cdot l) - \eta(K) \Big) = 0 \text{ sur }]0,T[\times\mathbb{R} \\ W(T,x) = \theta \cdot (h(x) + l(0, X_0)) \text{ sur } \mathbb{R} \end{cases}$$

On se propose ensuite de régulariser ce système en se donnant une fonction $\phi : \mathbb{R} \mapsto \mathbb{R}_+$ continuement dérivable, à support compact et d'intégrale unité :

$$\int_{\mathbb{R}} \phi(t)dt = 1$$

et l'on pose pour $n \in \mathbb{N}^*$ et $t \in \mathbb{R}$

$$\phi_n(t) = n\phi(nt)$$

La régularisation de l'équation $(\widehat{HJB}_0^\theta)$ consiste alors à modifier la fonction $l(t,x)$ définie par (VIII.7.10) en posant :

$$\forall (t,x) \in [0,T] \times \mathbb{R},\ \tilde{l}(t,x) = \sum_{i=1}^{p} h_i(x)(\phi_n * 1_{]-\infty,T_i]})(t) \tag{VIII.7.12}$$

et en notant d'autre part :

$$\begin{aligned} \forall (t,x) \in [0,T] \times \mathbb{R},\ \bar{l}(t,x) &\equiv -\frac{\partial \tilde{l}}{\partial t}(t,x) && \text{(VIII.7.13)} \\ &= \sum_{i=1}^{p} h_i(x)\phi_n(t - T_i) && \text{(VIII.7.14)} \end{aligned}$$

La fonction $\bar{l}$ ainsi définie apparaît comme une fonction de payoff régularisée en temps et, avec la modification introduite par la définition (VIII.7.12), on remplace d'abord le système $(\widehat{HJB}_0^\theta)$ par :

$$(\widehat{HJB}_{\text{rég}}^\theta) \begin{cases} \frac{\partial W}{\partial t} + \sup_{K \in \mathcal{K}^{\overline{H}}} \Big(\mathcal{G}^K(W + \theta \cdot \tilde{l}) - \eta(K) \Big) = 0 \text{ sur } [0,T] \times \mathbb{R} \\ W(T,x) = \theta \cdot (h(x) + l(0, X_0)) \text{ sur } \mathbb{R} \end{cases}$$

En tenant compte des conditions de régularité spatiale de $\tilde{l}$ grâce à $(\widehat{RS}_0)$ et temporelle par convolution dans (VIII.7.12), il est ensuite licite d'effectuer le nouveau changement de fonction :

$$\overline{W} = W + \theta \cdot (\tilde{l} - l(0, X_0)) \tag{VIII.7.15}$$

qui préserve le caractère $C^{1,2}$ de la solution tandis que la valeur initiale n'est pas modifiée car $\theta \cdot (\tilde{l}(0, X_0) - l(0, X_0))$ est nul si la régularisation n'est pas trop poussée (avec un support de Φ assez faible).
Avec la définition (VIII.7.13), le système $(\widehat{HJB}^{\theta}_{\text{rég}})$ se réécrit :

$$(\overline{HJB}^{\theta}_{\text{rég}}) \begin{cases} \frac{\partial \overline{W}}{\partial t} + \sup_{K \in \mathcal{K}^{\overline{H}}} \left(\mathcal{G}^K \overline{W} - \eta(K) \right) + \theta \cdot \bar{l} = 0 \text{ sur } [0,T] \times \mathbb{R} \\ \overline{W}(T,x) = \theta \cdot (h(x) + \tilde{l}(T,x)) \text{ sur } \mathbb{R} \end{cases}$$

Lorsque la régularisation n'est pas trop poussée, $\tilde{l}(T,.)$ est identiquement nul et l'on retombe exactement sur le système (HJB^{θ}) avec $\bar{l}$ donné par (VIII.7.14), ce qui permet de l'interpréter ce dernier comme la version régulière du système $(\widehat{HJB}^{\theta}_0)$ ou $(\widehat{HJB}^{\theta})$.

VIII.8 Perspective de solutions avec des hypothèses de régularité faibles

En adaptant des résultats de [49], il est possible d'introduire la notion de solution généralisée du système (HJB^{θ}) sous réserve de vérifier une condition d'existence et d'intégrabilité de la densité des distributions de transition du processus sous-jacent.

$$(HD) \begin{cases} \text{Sous la probabilité initiale, les distributions de transition de } X^{t,x}, \\ \text{solution de l'EDS (VIII.2.4), ont une densité } f_* \text{ telle que :} \\ \exists p_* > 1,\ \forall 0 \le t < s < T,\ \int_{[s,T]\times\mathbb{R}} f_*(t,x,t',x')^{p_*} dt' dx' < \infty \end{cases}$$

On peut alors envisager de résoudre (HJB^{θ}) dans le sens suivant :

Définition VIII.8.1 *On dit que $W : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ est une solution généralisée de (HJB^{θ}) si W vérifie (HJB^{θ}) pour presque tout $(t,x) \in]0,T[\times\mathbb{R}$ avec la condition terminale en T et avec les conditions de régularité suivantes*

1. *W est continue sur $[0,T] \times \mathbb{R}$ et satisfait la condition de croissance exponentielle :*

$$\forall (t,x) \in]0,T[\times\mathbb{R},\ |W(t,x)| \le \gamma(1 + |x| + e^x) \tag{VIII.8.1}$$

   où γ est une constante positive.

2. *Les dérivées partielles $\frac{\partial W}{\partial t}$, $\frac{\partial W}{\partial x}$, $\frac{\partial W}{\partial t}$ existent au sens généralisé et sont des éléments de $L^q(B)$ pour tout ensemble mesurable borné $B \subset]0,T[\times\mathbb{R}$.*

où q >1 vérifie :

$$q^{-1} + p_*^{-1} < 1 \tag{VIII.8.2}$$

Alors, au moins dans le cas où la mesure de Lévy est bornée, on peut montrer la validité du théorème de vérification VIII.4.3. Si on reprend la preuve du Théorème VIII.4.3, cela revient à montrer que le Lemme VIII.4.8 reste vrai, ce qui se fait en s'inspirant des techniques utilisées dans le Lemme V 11.2 et le Théorème V 11.1 de [49]. Il s'agit essentiellement d'étendre la formule d'Itô aux fonctions généralisées et, dans le cas d'une diffusion continue, cette généralisation se trouve dans [80] par exemple.
L'existence même d'une solution généralisée à l'équation (HJB^{θ}) peut s'obtenir dans les espaces de Sobolev avec poids introduits dans [13] dont on reprend les notations en posant pour $\mu \ge 0$:

$\forall x \in \mathbb{R},\ \beta_\mu(x) = e^{-\mu\sqrt{1+x^2}}$.
On note pour $p \geq 0$, $L^{p,\mu}(\mathbb{R})$ l'espace des fonctions u telles que $u\beta_\mu \in L^p(\mathbb{R})$ et l'on introduit l'espace $W^{1,p,\mu}(\mathbb{R})$ des fonctions u telles que $u\beta_\mu,\ u'\beta_\mu \in L^{p,\mu}(\mathbb{R})$, ainsi que $W^{2,p,\mu}(\mathbb{R})$ des fonctions u telles que $u\beta_\mu,\ u'\beta_\mu,\ u''\beta_\mu \in L^{p,\mu}(\mathbb{R})$.
On note enfin $\mathcal{W}^{1,2,p,\mu}(]0,T[\times\mathbb{R})$ l'espace des fonctions $u \in L^p(]0,T[;W^{2,p,\mu}(\mathbb{R}))$ telles que $\frac{\partial u}{\partial t} \in L^p(]0,T[;L^{p,\mu}(\mathbb{R}))$.
Faisons alors l'hypothèse suivante :

$$(RS^{p,\mu})\ :\ h \in W^{2,p,\mu}(\mathbb{R}) \text{ et } \bar{l} \in L^p(]0,T[\times\mathbb{R})$$

En adaptant le Théorème III.3.1 de [13], on prouve :

Proposition VIII.8.2 *Supposons* $2 \leq p < \infty$ *et* $\mu > 0$ *et supposons aussi que la dérivée ordinaire* $\frac{\partial \sigma_*}{\partial x}$ *existe et est bornée.*
Alors, sous l'hypothèse $(RS^{p,\mu})$*, le système* (HJB^θ) *admet dans* $\mathcal{W}^{1,2,p,\mu}(]0,T[\times\mathbb{R})$ *une solution unique.*

Si l'on renforce alors les hypothèses en supposant $(RS^{\infty,0})$, on montre la continuité de la solution W^θ qui est donc une solution généralisée du système (HJB^θ) si l'on montre qu'elle est à croissance exponentielle et même bornée.
Pour pouvoir résoudre le problème de calibration, en dehors de cette dernière vérification ainsi que celle de la condition (HD), il resterait à prouver la régularité de la solution généralisée par rapport à θ. Il n'est cependant pas évident d'être en situation d'utiliser le théorème des fonctions implicites, ce qui est un obstacle à cette perspective. Eventuellement en vue d'une recherche future, il serait cependant intéressant d'examiner sous quelles conditions de régularité supplémentaires on peut lever cet obstacle.

Conclusion et perspectives

La connaissance de la loi des marchés financiers à partir de l'information révélée par les prix suscite d'intenses recherches. Il s'agit d'un problème inverse pour lequel on sait utiliser des techniques classiques (détermination d'une densité de probabilité à partir de ses moments). Il reste cependant à approfondir les aspects spécifiques aux modèles de la finance concernant le choix de ces derniers ainsi que leur calibration. On s'est intéressé à certains de ces aspects.

Dans une première partie, on a fait l'étude d'une technique de calibration par la méthode de Monte-Carlo [6]. A la lumière de résultats d'analyse convexe, on a prouvé la convergence d'estimateurs des prix d'options calibrés.
On rappelle qu'on cherchait à pondérer n tirages indépendants X_1,...,X_n obtenus suivant une loi a priori du sous-jacent. De façon précise, il s'agissait de déterminer un vecteur $(p_i^n)_{1\leq i\leq n}$ de $[0,1]^n$ tel que $\sum_{i=1}^n p_i^n = 1$ et $\sum_{i=1}^n f(X_i)p_i^n = \mathbf{0}$, minimisant

$$\sum_{i=1}^{n} J(np_i)$$

où J est une fonction convexe.
Sous une condition géométrique, on a montré que ce problème admettait une solution à partir d'un certain rang n. Lorsque le critère à minimiser est l'entropie relative, on a prouvé la convergence étroite de $\nu_n = \sum_{i=1}^n p_i^n \delta_{X_i}$. La limite ν s'obtient par minimisation du critère dual associé à un problème de calibration limite. Ce n'est la solution de ce dernier problème que si ν vérifie les contraintes. Dans le cas contraire, on fait apparaître une I-projection généralisée qui est naturellement normalisée en une probabilité.
Cependant, dans de nombreux cas on a signalé l'intérêt à choisir un critère différent de l'entropie relative. Classiquement, il peut s'agir de critères de Burg ou de critères bâtis sur une fonction puissance. Dans un premier temps, on a considéré une extension à des critères proches de l'entropie relative. Les résultats de convergence de la méthode de Monte-Carlo pondérée subsistent alors. En particulier, il est remarquable que cette convergence assure que la solution "généralisée" ν du problème limite est une probabilité.
C'est avec cette extension de l'entropie relative que l'on a ensuite envisagé de relâcher les contraintes de calibration. Cela a été fait de deux manières : en pénalisant le critère de minimisation ou bien en considérant des contraintes de type ensemble. On a alors énoncé des résultats de convergence semblables à ceux du cas contraint élémentaire.
En adoptant un critère convexe proche de l'entropie relative, on a pu traiter le cas où la fonction de payoff des options de calibration n'est pas bornée, au prix que la loi limite ne soit pas toujours calibrée. Cependant, en se restreignant à une fonction de payoff bornée, on a étendu les résultats de convergence à des critères plus généraux. Cette extension englobe notamment les critères construits avec une fonction puissance d'exposant $p > 1$. Il n'a cependant pas été possible

de faire la généralisation au cas $p < 1$ ni à celui de l'entropie de Burg. Dans une perspective future, il serait raisonnable d'approfondir l'étude de ces deux types de critères afin de les utiliser éventuellement pour la calibration par la méthode de Monte-Carlo.
Les extensions obtenues précédemment présentent un intérêt financier (avec le choix d'une fonction d'utilité) et, de manière plus générale, sont susceptibles d'application aux problèmes inverses où l'on ne souhaite pas forcément utiliser l'entropie relative. Cette dernière jouit néanmoins de propriétés spécifiques remarquables. Sa version pénalisée, largement utilisée, méritait un éclairage particulier. En examinant le problème de calibration limite, on a effectivement montré que ce dernier admettait une solution généralisée dans le sens où elle est la limite pour la norme de la variation de toute suite minimisante : c'est donc une transposition naturelle de la notion de I-projection généralisée. Avec les résultats précédents, on l'a également identifiée avec la limite du problème de calibration par la méthode de Monte-Carlo.
Après avoir prouvé la convergence de cette dernière, on a établi un résultat de type théorème central limite auquel on était en droit de s'attendre dans le cadre d'utilisation de M-estimateurs. Dans le cas de l'entropie relative, on a pu faire apparaître la possibilité d'une réduction de variance. La stabilité de la méthode a également été démontrée.
L'évaluation des prix par la méthode de Monte-Carlo calibrée en minimisant l'entropie ou un critère plus général présente l'inconvénient de ne pas garantir la martingalité de la probabilité limite. Cet aspect du sujet a donc été examiné pour compléter l'étude. En fait, on a montré qu'un nombre suffisant de contraintes supplémentaires permet d'approcher la condition de martingalité du sous-jacent actualisé. On a essentiellement utilisé des résultats de convergence pour les problèmes de moments. L'entropie relative a été systématiquement adoptée lorsque la propriété de Markov est en jeu. Pour terminer, on a donné une caractérisation de la probabilité martingale minimale calibrée. Dans une perspective de futures applications, il conviendrait de choisir les contraintes de martingalité approchée de manière appropriée.
On a donc pu valider un ensemble de résultats concernant l'utilisation de la méthode de Monte-Carlo pour la calibration des options financières. Ces résultats sont cependant d'une portée plus générale et s'étendent pour l'essentiel à la résolution de problèmes du type de la programmation partiellement finie. Les aspects spécifiques de la finance concernent la propriété de martingale et celle de Markov au sujet desquelles il est raisonnable d'adopter l'entropie relative.

Dans cette partie assez générale, on n'a traité que du problème de calibration, le choix de la mesure a priori n'étant astreinte qu'à satisfaire une condition géométrique, sauf dans le cas de la pénalisation. Il est cependant essentiel de bien déterminer la loi a priori. On s'est donc intéressé par la suite à des modèles particuliers, susceptibles de créer le smile attendu.

La deuxième partie proposait le phénomène de krach comme explication du smile. En incorporant des sauts dans la volatilité lorsque le sous-jacent subit une baisse brutale et en imposant à la volatilité un retour au régime normal, on a étudié les aspects de valorisation et de couverture d'options européennes. Par un calcul du compensateur des sauts, on a évalué le risque que comporte un portefeuille optimal constitué avec l'option et son sous-jacent. La calibration a été abordée avec une formule approchée du smile. Cette étude doit se considérer comme préparatoire à celle d'un modèle plus réaliste.

Dans la dernière partie, on a étudié un modèle de sauts incorporant la volatilité locale. L'entropie relative a été le critère retenu pour calibrer l'intensité des sauts du modèle et l'on s'est principalement attaché à démontrer la stabilité de la méthode grâce à des techniques de contrôle optimal. Le choix des espaces Höldériens a permis d'utiliser le théorème des fonctions implicites et d'obtenir la robustesse de la calibration voulue. Il n'est pas apparu évident de relâcher les conditions de régularité Höldériennes imposées sur les coefficients du modèles ainsi que sur les fonctions de payoff des options de calibration. Cela peut constituer cependant un objectif raisonnable. En vue des applications, il resterait à produire une méthode numérique efficace.

Bibliographie

[1] Amemiya T. (1985). Advanced Econometrics. *Harvard University Press, cambridge, MA*

[2] Andersen L., Andreasen J. (2000). Jump Diffusion Models : Volatility Smile Fitting and Numerical Methods for Pricing. *Review of Derivatives Research*, 4, 231-262

[3] Avellaneda, M., Levy A., Paras A. (1995). Pricing and Hedging Derivative Securities in Markets with Uncertain Volatilities. *Applied Mathematical Finance*, 2, 73-88

[4] Avellaneda, Paras A. (1996). Managing the Volatility Risk of Portfolios of Derivative Securities : The Lagrangian Uncertain Volatility Model. *Applied Mathematical Finance*, 3, 21-52

[5] Avellaneda, M., Friedman C., Holmes R., Samperi D (1997). Calibrating Volatility Surfaces via Relative Entropy Minimization. *Applied Mathematical Finance*, 4(1), 37-64

[6] Avellaneda M., Buff R., Friedman C., Grandchamp N., Kruk L. et Newman J. (2000). Weighted Monte Carlo : A new Technique for Calibrating Asset-Pricing Models. *Intern. J. of Theor. and Appl. Finance*, 4(1), 91-119

[7] Bachelier L. (1900). Théorie de la Spéculation. Thèse de Doctorat, Faculté des Sciences de Paris *Annales Scientifiques de l'Ecole Normale Supérieure*, Suppl. 3, 1017, 21

[8] Bahadur R.R., Zabell S.L. (1979). Large Deviations of the Sample Mean in General Vector Spaces. *The Annals of Probability*, 7(4), 587-621

[9] Bellamy N., Jeanblanc M. (2000). Incompleteness of Markets Driven by a Mixed Diffusion. *Finance and Stochastics*, 4, 209-222

[10] Bellini F., Frittelli M. (2002). On the Existence of Minimax Martingale Measures. *Mathematical Finance*, 12(1), 1-21

[11] Ben-Bassat B. (1978). f-Entropies, Probability of Error, and Feature Selection *Information and Control*, 39, 227-242

[12] Ben-Tal A., Borwein J.M., Teboulle M. (1988). A Dual Approach to Multidimensional L_p Spectral Estimation Problems. *SIAM J. Control and Optimization*, 26(4), 985-996

[13] Bensoussan A., Lions J.L. (1982). Contrôle Impulsionnel et Inéquations Quasi Variationnelles. *Dunod*

[14] Berthon J., Gallais-Hamonno G. (1994). Les Options négociables. *Presses Universitaires de france*

[15] Bismut J.M. (1975). Contrôle des Processus de Sauts. *C.R.Acad.Sc.Paris, Série A*, 281, 767-770

[16] Black F., Scholes M. (1973). The Pricing of Options and Corporate Liabilities. *J. Political Econom.*, 72, 637

[17] Bodurtha J.N., Jermakyan M. (1999). Nonparametric Estimation of an Implied Volatility Surface. *Journal of Computational Finance*, 2(4), 29-60

[18] Bonnans J.F., Gilbert J.C., Lemaréchal C., Sagastizabal C.A. (1997). Numerical Optimization *Springer*

[19] Borovkov A.A. (1998). Mathematical Statistics *Gordon and Breach Science Publishers*

[20] Borwein J.M. (1981). A Lagrange Multiplier Theorem and a Sandwich Theorem for Convex Relations. *Math. Scand.*, 48, 189-204

[21] Borwein J.M., Lewis A.S. (1991). On the Convergence of Moment Problems. *Trans. Amer. Math. Soc.*, 325, 249-271

[22] Borwein J.M., Lewis A.S. (1991). Convergence of Best Entropy Estimates. *SIAM J. Optimizat.*, 1, 191-205

[23] Borwein J.M., Lewis S. (1991). Duality Relationships for Entropy-Like Minimization Problems. *SIAM J. Control and Optimization*, 29(2), 325-338

[24] Borwein J.M., Lewis S. (1992). Partially-Finite Convex Programming, Part I : Quasi Relative Interiors and Duality Theory. *Mathematical Programming*, 57, 15-48

[25] Borwein J.M., Lewis S. (1992). Partially-Finite Convex Programming, Part II : Explicit Lattice Models. *Mathematical Programming*, 57, 49-83

[26] Borwein J.M., Lewis S. (1993). Partially-Finite Programming in L_1 and the Existence of Maximum Entropy Estimates. *SIAM J. Optimization*, 3(2), 248-267

[27] Bouchouev I., Isakov V. (1999). Uniqueness, Stability and Numerical Methods for the Inverse Problem that Arises in Financial Markets. *Inverse Problems*, 15, R95-R116

[28] Bouleau N. (1988). Processus Stochastiques et Applications. *Hermann*

[29] Brézis H. (1993). Analyse fonctionnelle, Théorie et applications. *Masson*

[30] Brillouin L. (1958). La Science et la Théorie de l'Information. *Gabay*

[31] Buchen P.W., Kelly M. (1996). The Maximum Entropy Distribution of an Asset Inferred from Option Prices *Journal of Financial and Quantitative Analysis*, 31(1), 143-159

[32] Bühlmann H., Delbaen F., Embrechts P., Shiryaev A.N. (1996). No-Arbitrage, change of Measure and Conditional Esscher Transforms. *CWI Quaterly*, 9(4), 291-317

[33] Cattiaux P., Leonard C (1994). Minimization of the Kullback Information of Diffusion Processes. *Ann. Inst. Henri Poincaré*, 30(1), 83-132

[34] Chan T. (1999). Pricing Contingent Claims on Stocks Driven by Lévy Processes. *The Annals of Applied Probability*, 9(2), 504-528

[35] Cont R., Tankov P. (2002). Calibration of Jump-Diffusion Models : a Robust Non-Parametric Approach. *working paper, CMAP*

[36] Csiszar I. (1975). I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1), 146-158

[37] Csiszar I. (1981). Why Least Squares and Maximum Entropy ? An Axiomatics Approach to Inference for Linear Inverse Problems. *The Annals of Statistics*, 19(4), 2032-2066

[38] Csiszar I. (1984). Sanov property, generalized I-projection and a conditional limit theorem. *The Annals of Probability*, 12(3),p.768-793

[39] Dacunha-Castelle D., Gamboa F. (1990). Maximum d'Entropie et Problème de Moments. *Ann. Inst. Henri Poincaré*, 26(4), 567-596

[40] Dalang R.C., Morton A., Willinger W. (1990). Equivalent Martingale Measures and No-Arbitrage in Stochastic Securities Market Models. *Stochastics Stochastics Rep.*, 29(2), 185-201

[41] Delbaen F. (1999). The Dalang-Morton-Willinger Theorem. *http ://www.math.ethz.ch/ ˜delbaen*

[42] Delbaen F., Grandits P., Rheinländer T., Samperi D., Schweizer M., Stricker C. (2000). Exponential Hedging and Entropic Penalties. *working paper*

[43] Dupire B. (1994). Pricing with a Smile. *Risk Magazine*, 7(1), 18-20

[44] Eberlein E., Jacod J. (1997). On the Range of Option Pricing. *Finance and Stochast.*, 1, 131-140

[45] El Karoui N., Lepeltier J.-P. (1977). Représentation des Processus Ponctuels Multivariés à l'Aide d'un Processus de Poisson. *Z. Wahrscheinlichkeitstheorie Vew. Gebiete*, 39, 111-133

[46] El Karoui N., Rouge R. (2000). Pricing via Utility Maximization and Entropy. *Math. Finance*, 10(2), 259-276

[47] Engl H.W., Landl G. (1993). Convergence Rates for Maximum Entropy Regularization. *SIAM J.Numer.Anal.*, 30(5), 1509-1536

[48] Esche F., Schweizer (2003). Minimal Entropy Preserves the Lévy Property : How and Why. *preprint, University of Munich*

[49] Fleming W. H., Rishel R. W. (1975). Deterministic and Stochastic Optimal Control. *Springer Verlag*

[50] Fleming W. H., Soner H. M. (1993). Controlled Markov Processes and Viscosity Solutions. *Springer Verlag*

[51] Föllmer H., Schweizer M. (1991). Hedging of Contingent Claims under Incomplete Information. in M. H. A. davis and R.J. Elliot ed. *Applied Stochastic Analysis*, Gordon and breach, New York, 5, 389-414

[52] Foster F.D., Whiteman C.H. (2000). Bayesian Prediction, Entropy, and Option Pricing in the U.S. Soybean Market, 1993-1997. *working paper, University of Iowa*

[53] Frittelli M. (2000). The Minimal Entropy Martingale Measure and the Valuation Problem in Incomplete Markets. *Mathematical Finance*, 10(1), 39-52

[54] Gamboa F., Gassiat E (1997). Bayesian Methods and Maximum Entropy for Ill-Posed Inverse Problems. *The Annals of Statistics*, 25(1), 328-350

[55] Garroni M.G., Menaldi J.L. (1992). Green Functions for Second Order Parabolic Integro-Differential Problems. *Longman Scientific & Technical*

[56] Glasserman P., Jin Y. (1999). Comparing Stochastic Discount Factors Through their Implied Measures. *preprint*, 25 p.

[57] Gopikrishnan P., Plerou V., Nunes Amaral L.A. (1999). Scaling of the Distribution of Fluctuations of Financial Market Indices. Preprint disponible à *http ://xxx.lanl.gov/cond-mat/9905305*

[58] Grandits P., Rheinländer T (2002). On the Minimal Entropy Martingale Measure. *The Annals of Probability*, 30(3), 1003-1038

[59] Groeneboom P., Oosterhoff J., Ruymgaart F.H. (1979). Large Deviation Theorems for Empirical Probability Measures *The Annals of Probability*, 7(4), 553-586

[60] Gzyl H (1995). The Method of Maximum Entropy *World Scientific*

[61] Hammersley J.M., Handscomb D.C (1964). Monte Carlo Methods *Chapman&Hall*

[62] Harrison J.M., Kreps D. (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, 20, 381-408

[63] Harrison J.M., Pliska S.R. (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications*, 11, 215-260

[64] Henderson V, Hobson D., Howison S., Kluge T. (2003). A Comparison of q-Optimal Option Prices in a Stochastic Volatility Model with Correlation . *Working Paper*

[65] Heston S. L., (1993). A Closed-Form Solution For Options With Stochastic Volatilities *Review of Financial Studies*, 6(2), 327-343

[66] Hull J. C. (1997). Options, Futures, and Other Derivatives. *Prentice Hall*

[67] Jackson N., Süli E., Howison S. (1998). Computation of Deterministic Volatility Surfaces. *Report, Oxford University Comp. Laboratory*

[68] Jacod J., Shiryaev A.N. (1987). Limit Theorems for Stochastic Processes. *Springer-Verlag*

[69] Jaynes E.T. (1957). Information Theory and Statistical Mechanics. *The Physical Review*, 106(4), 620-630

[70] Jaynes E.T. (1963). Information Theory and Statistical Mechanics. In : *Brandeis University Summer Institute Lectures in Theoretical Physics, K.W.Ford editor*

[71] Jourdain B. et Nguyen L. (2001). Minimisation de l'entropie relative par méthode de Monte-Carlo. *Comptes Rendus Acad. Sc., Série 1*, 332(4), 345-3500

[72] Jupp P.E. et Mardia K.V. (1983). A note on the maximum entropy principle. *Scandin. J. Statist.*, 10, p.45-47

[73] Kallsen J. (2002). Utility-Based Derivative Pricing in Incomplete Markets. *Mathematical Finance-Bachelier Congress 2000, Springer*, 313-338

[74] Kallsen J. (2002). Derivative Pricing Based on Local Utility Maximization. *Finance and Stochastics*, 6, 115-140

[75] Karatzas I., Lehoczky J.P., Shreve S., Xu G.L. (1991). Martingale and Duality Methods for Utility Maximization in an Incomplete Market. *SIAM J. Control and Optimization*, 29(3), 702-730

[76] Karatzas I., Shreve S. (1997). Brownian Motion and Stochastic Calculus. *Springer Verlag*

[77] Karatzas I., Shreve S. E. (1999). Methods of Mathematical Finance *Springer Verlag*

[78] Kloeden P.E., Platen E. (1995). Numerical Solution of Stochastic Differential Equations *Springer Verlag*

[79] Komatsu T. (1973). Markov Processus Associated with Certain Integro-Differential Operators. *Osaka Journal of Mathematics*, 10, 271-303

[80] Krylov N.V. (1980). Controlled Diffusion Processes. *Springer Verlag*

[81] Kullback S., Leibler R.A. (1951). On Information and Sufficiency. *Ann. Math. Statist.*, 22, 79-86

[82] Lagnado R., Osher S. (1997). A Technique for Calibrating Derivative Security Pricing Models : Numerical Solution of an Inverse Problem. *Journal of Computational Finance*, 1(1), 13-25

[83] Lamberton D., Lapeyre B. (1997). Introduction au Calcul Stochastique Appliqué à la Finance. *Ellipse*

[84] Lapeyre B., Pardoux E., Sentis R. (1998). Méthodes de Monte-Carlo pour les Equations de Transport et de Diffusion. *Springer-Verlag*

[85] Lehrer E., Smorodinsky R. (2000). Relative Entropy in Sequential Decision Problems. *Journal of Math. Econ.*, 33, 425-439

[86] Lepeltier J.P., Marchal B. (1976). Problème des Martingales et Equations Différentielles Stochastiques Associées à un Opérateur Intégro-Différentiel. *Annales de l'Institut Henri Poincaré-Section B*, 12(1), 43-103

[87] Lepeltier J.P., Marchal B. (1977). Sur l'Existence de Politiques Optimales dans le Contrôle Intégro-Différentiel. *Annales de l'Institut Henri Poincaré-Section B*, 13(1), 45-97

[88] Lin J. (1991). Divergence Measures based on the Shannon Entropy. *IEEE Trans. on Information Th.*, 37(1), 145-151

[89] Mandelbrot B.B. (1963). The Variation of Certain Speculative Prices. *J. Business*, 36, 394

[90] Mandelbrot B.B. (1997). Fractals and Scaling in Finance. *Springer*

[91] Mantegna R.N., Stanley H.E. (1995). Scaling Behaviour in the Dynamics of an Economic Index *Nature*, 376, 46

[92] Merton R. (1973). Theory of Rational Option Pricing *Bell J. Econom. Managem. Sci*, 4, 141

[93] Merton, R.C. (1976). Option Pricing When Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*, 5, 125-144

[94] Miyahara Y. (1999). Minimal Relative Entropy Martingale Measures and Their Applications to Option Pricing Theory. In : *Proceedings of JIC99, The 5-th JAFEE International Conference*, 316-323

[95] Miyahara Y., Novikov A. (2003). Geometric Lévy Process Pricing Model. *working paper*

[96] Musiela M., Rutkowski M. (1998). Martingale Methods in Financial Modelling. *Springer Verlag*

[97] Naik V. (1993). Option Valuation and Hedging Strategies with Jumps in the Volatility of Asset Returns. *The Journal of Finance*, 48(5), 1969-1984

[98] Paul W., Baschnagel J. (1999). Stochastic Processes. From Physics to Finance. *Springer Verlag*

[99] Pinsker M.S. (1964). Information and Information Stability of Random Variables and Processes *Holden-Day, INC*

[100] Press H.P., Teukolsky S.A., Vetterling W.T, Flannery B.P., (1994). Numerical Recipes in C, The Art of Scientific Computing *Cambridge University Press*

[101] Protter P. (1990). Stochastic Integration and Differential Equations. *Springer-Verlag*

[102] Reza F.M. (1994). An Introduction to Information Theory. *Dover*

[103] Robertson J.C., Tallman E.W., Whiteman C.H. (2002). Forecasting Using Relative Entropy. *working paper, Federal Reserve Bank of Atlanta*

[104] Rockafellar R.T. (1968). Integrals which are Convex Functionals. *Pacific Journal of Mathematics*, 24(3), 525-539

[105] Rockafellar R.T. (1971). Integrals which are Convex Functionals, II. *Pacific Journal of Mathematics*, 39(2), 439-469

[106] Rockafellar R.T. (1972). Convex analysis. *Princeton Mathematical Series, 28*

[107] Rockafellar R.T. (1974). Conjugate Duality and Optimization. *CBMS-NSF Regional Conference Series in Applied Mathematics 16, Society for Industrial and Applied Mathematics, Philadelphia, PA*

[108] Rouge R., El Karoui N. (2000). Pricing Via Utility Maximization and Entropy. *Mathematical Finance*, 10(2), 259-276

[109] Rudin W. (1995). Analyse réelle et complexe. *Masson*

[110] Rudin W. (1995). Analyse Fonctionnelle. *Ediscience International*

[111] Runggaldier W.J. (2003). Jump-Diffusion Models. In : *Handbook of Heavy Tailed Distributions in Finance(S.T. Rachev, ed.), Handbooks in Finance, Book 1(W.Ziemba Series Ed.),Elsevier/North-Holland*, 169-209

[112] Samperi D. (2000). Model Calibration Using Entropy and Geometry. *working paper, Decision Synergy Inc.*

[113] Samperi D. (2002). Calibrating a Diffusion Pricing Model with Uncertain Volatility : Regularization and Stability *Mathematical Finance*, 12(1), 71-87

[114] Samuelson P. A. (1965). Rational Theory of Warrant Pricing. *Industrial Management Review*, 6, 13 Suppl. 3, 1017, 21

[115] Sanov I.N. (1961). On the Probability of large Deviations of Random Variables *Sel. Transl. Math. Statist. Proba.*, 1, 213-244

[116] Schachermayer W. (1992). A Hilbert Space Proof of the Fundamental Thorem of Asset Pricing in Finite Discrete Time. *Insurance Math. Econom.*, 11(4), 249-257

[117] Shannon C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, 379-423

[118] Slomczynski W., Zastawniak T. (2002). Utility Maximizing Entropy and Second law of Thermodynamics *preprint* 25 p.

[119] Stone M. (1974). Large Deviations of Empirical Probability Measures. *The Annals of Statistics*, 2(2), 362-366

[120] Stricker C. (1990). Arbitrage et Lois de Martingale. *Ann. Inst. Henri Poincaré*, 26(3), 451-460

[121] Strook D. (1975). Processus Diffusion Associated with Lévy generators. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, 32, 209-244

[122] Stutzer M. J. (2000). Simple Entropic Derivation of a Generalized Black-Scholes Option Pricing Model *Entropy*, 2, 70-77

[123] Teboulle M., Vajda I. (1993). Convergence of Best ϕ-Entropy Estimates. *IEEE Transactions on Information Theory*, 39(1), 297-301

[124] Topsoe F. (1979). Information Theoretical Optimization Techniques *Kybernetyka*, 15(1), 8-27

[125] Van Campenhout J.M., Cover T.M. (1981). Maximum Entropy and Conditional Probability. *IEEE Trans. on Information Theory*, 27(4)

[126] Vasicek O.A. (1980). A Conditional Law of Large Numbers. *The Annals of Probability*, 8(1), 142-147

[127] Viala P, Briys E. (1995). Eléments de Théorie Financière. *Nathan*

[128] Yor M. (1976-1977). Sous-Espaces Denses dans L^1 ou H^1 et Représentation des Martingales. *Université de Strasbourg, Séminaire de Probabilités*, 12, 265-309