UNIVERSITÉ D'ÉVRY VAL-D'ESSONNE

DOCUMENT DE SYNTHÈSE[1]
en vue d'obtenir
L'HABILITATION À DIRIGER DES RECHERCHES
Discipline : Mathématiques Appliquées
présenté et soutenu publiquement le 8 décembre 2008

par

Stéphane CRÉPEY

Titre :

MATHÉMATIQUES FINANCIÈRES ET FINANCE NUMÉRIQUE

FINANCIAL MODELING AND NUMERICAL FINANCE

Jury :

M. Bruno BOUCHARD	Examinateur
M. René CARMONA	Rapporteur
M. Mark DAVIS	Rapporteur
M. Laurent DENIS	Examinateur
Mme Nicole EL KAROUI	Rapportrice
Mme Monique JEANBLANC	Examinatrice
M. Huyen PHAM	Examinateur
M. Nizar TOUZI	Rapporteur

[1] Version finale.

Contents

Part I
Présentation générale

Introduction

La recherche présentée dans ce mémoire, correspondant à mes travaux depuis la thèse (soutenue en janvier 2001), se situe dans le domaine des mathématiques financières et de la finance numérique: problèmes d'évaluation et de couverture de produits dérivés en finance, et techniques probabilistes et d'équations aux dérivées partielles afférentes. Cette recherche se décompose suivant quatre thèmes:
• *(2001–04)* Modélisation de la volatilité (modèles à volatilité locale), dans le domaine des dérivés actions et indices actions (*equity and equity index derivatives*);
• *(2005–07)* Risque de défaut 'single name' et *equity to credit* (problématiques risque de crédit mettant en cause une unique contrepartie, et lien entre les sphères du credit et de l'equity; obligations convertibles en particulier);
• *(2008)* Risque de crédit 'multi name' (problème de la modélisation de la dépendance entre les défauts d'un portefeuille de crédit).
Dans chaque cas les approches privilégiées ont été celles les plus susceptibles de donner lieu à une utilisation effective en salle de marchés (avec les contraintes de calculabilité que cela comporte, calibrabilité de modèle notamment):
– étude des modèles à volatilité locale et de leur calibration dans le cas des dérivés actions,
– modèles à intensité (par opposition à structurels) pour le risque de crédit (single-name aussi bien que multi-name).
• Enfin les second et troisième points ci-dessus ont suscité des développements mathématiques dans le domaine des équations différentielles stochastiques rétrogrades ('équations backward') réfléchies avec sauts, ainsi que, dans le cas Markovien, des problèmes intégro-différentiels avec obstacles associés.

Ces différents aspects de ma recherche sont mis en perspective dans la suite de cette partie. Les ***principales contributions*** sont passées en revue en section 4. Divers aspects transverses: collaborations, expériences dans l'animation d'une recherche (encadrements d'étudiants en thèse en particulier), relations avec l'industrie, sont également évoqués au fil du texte. Travaux en cours et projets de recherche font l'objet de la section 5.

Les résultats obtenus font ensuite l'objet de présentations techniques plus poussées aux parties II à V. Pour les preuves d'une part et illustrations numériques d'autre part, on renvoie par défaut aux articles concernés. Occasionnellement cependant des éléments de démonstration ou d'illustration numérique sont donnés dans le texte du mémoire, quand jugés suffisamment simples et illustratifs.

1 Volatilité locale

Cette recherche a été menée en parallèle d'une activité de consulting à la société Artabel (dissoute depuis) entre 2001 et 2003, puis d'une activité de demi-ATER au département de mathématiques de l'université d'Évry en 2003-04. La société Artabel développait à l'époque des logiciels de pricing et calibration de modèle en finance, au sein d'une équipe de recherche-

développement en ingénierie financière dirigée par Rama Cont et Claude Martini. La notion de volatilité locale avait été introduite dans les années 90 à propos des marchés de dérivés actions et indice actions par Dupire [58] et Derman–Kani [56]. J'y avais déjà consacré une partie de ma thèse. C'est dans le cadre de mon activité chez Artabel que j'ai continué mon travail sur ce thème, sous plusieurs angles:
• Au plan mathématique, étendre les résultats théoriques de la thèse relativement au caractère bien posé d'une méthodologie de calibration par régularisation de type Tikhonov–Phillips [115, 109, 64] à des espaces fonctionnels suffisamment riches de nappes de volatilité (en s'affranchissant notamment de l'hypothèse technique de monotonie en temps de la fonction de volatilité dans la thèse [13]),
• Au plan pratique, élaborer un algorithme précis et rapide de calibration (par discrétisation du cadre précédent) adossé à de bonnes propriétés théoriques de stabilité, convergence et vitesse de convergence, plus performant en pratique que la méthode de Lagnado et Osher [94] implémentée dans la thèse,
• Analyser enfin, tant mathématiquement que numériquement sur données simulées et données réelles, les propriétés des modèles à volatilité locale en termes de qualité de couverture en delta, à concurrence de la méthode de couverture la plus communément utilisée par les traders (couverture en delta de Black–Scholes correspondant à la volatilité implicite de l'option couverte).

Les objectifs correspondant aux trois points ci-dessus ont été poursuivi pendant les années 2001–04, débouchant sur les publications [11, 12, 10], respectivement (voir aussi [9] pour un survey sur la méthode de régularisation de Tikhonov appliquée aux problèmes de calibration de modèle en finance).
L'algorithme de calibration proposé dans [10] a fait l'objet d'une **implémentation industrielle** chez Artabel dans les années 2001–03.

2 Equity to credit et dérivés de crédit single-name

2.1 Crédit Single-Name *versus* Multi-Name

Dans le domaine de la modélisation du risque de crédit (voir Bielecki–Rutkowski [29] pour une référence générale), les évènements qui doivent être pris en compte sont principalement le risque de défaut des émetteurs et les pertes associées à ces défauts, ainsi que l'évolution de la qualité de crédit des émetteurs. Il est donc important de travailler avec un modèle pertinent d'évènements de défaut, et de dynamique de la qualité des firmes concernées. On peut schématiquement distinguer deux types de problématiques risque de crédit, single-name *versus* multi-name:
• Le crédit *single-name* concerne l'évaluation et la couverture de payoffs defaultable du type $\mathbb{1}_{T<\theta}\phi(S_T)$ en T (ou $\mathbb{1}_{\tau<\theta}\phi(S_\tau)$ en cas d'exercice anticipé en τ, pour les produits à clauses américaines ou de types jeux comme les obligations convertibles), où:
– θ représente l'instant de défaut d'une entité de référence (firme),
– T est la date de maturité du produit, et
– ϕ est une *fonction de payoff* d'un sous-jacent (par exemple l'action de la firme) S.
Une question centrale est alors celle du lien entre Equity et Credit (*Equity to Credit*, ou *Credit to Equity*);
• Le crédit *multi-name* concerne pour sa part l'évaluation et la couverture de payoffs du

type $\phi(L_T)$ en T, où
– le *processus de perte* L est déterminé par les instants de défauts θ_i d'entités de référence (firmes), pour $i = 1, \ldots, n$ (typiquement $L_t = (1 - R) \sum_{i=1}^{n} \mathbb{1}_{\theta_i \leq t}$ pour un *taux de recouvrement* R), et
– ϕ est une *fonction de payoff* de la perte L en T.
La question centrale est alors celle de la *dépendance* entre les défauts des divers émetteurs. On renvoie le lecteur par exemple à Bielecki et al. [6, 30] pour une revue de synthèse des produits dérivés de crédit, typiquement le CDS, ou *Credit Default Swap*, pour le single-name, et le CDO, ou *Collaterized Debt Obligation*, pour le multi-name.

Suite à mon recrutement en septembre 2003 comme Maître de Conférences au département de mathématiques de l'Université d'Évry, au sein de l'équipe finance dirigée par Monique Jeanblanc, spécialiste du risque de crédit, mes centres d'intérêts se sont progressivement élargis du champ de la modélisation dérivés actions (voir section 1) à la modélisation des produits dérivés sujets au risque de défaut, ou *vulnérables* (notamment produits dérivés hybrides actions/crédit, dont l'exemple typique est fourni par les obligations convertibles), objet de la présente section 2, et plus récemment (ce sera l'objet de la section 3) aux produits dérivés de crédit multi-name.

2.2 Des obligations convertibles aux options de jeu vulnérables

Suscitée notamment par des échanges avec la société Ito33, société de développement de composants logiciels pour le front office spécialisée dans les obligations convertibles, en **partenariat de mécénat industriel** avec l'équipe finance d'Évry, cette recherche, menée en **collaboration** avec Monique Jeanblanc, Tomasz Bielecki et Marek Rutkowski [2, 3, 4, 5], concerne de manière plus générale les actifs contingents vulnérables, avec possiblement des clauses d'exercice anticipé (*put*) et/ou de rappel (*call*) de la part du détenteur et/ou de l'émetteur du produit. On a alors affaire à des produits vulnérables américains ou de type 'jeux' (*game options* [91]), dont les obligations convertibles constituent l'exemple de référence et la principale application.

Dans le domaine des méthodes dynamiques de pricing et de couverture en risque de crédit single-name, le courant dominant est celui des approches dites à forme *réduite* (par opposition aux modèles *structurels*, voir [29]).
L'idée, dans la lignée de Lando [95] ou Jarrow et Turnbull [86], est de ramener (*réduire*) l'étude d'un payoff defaultable à celle d'un payoff *default-free* (sans risque de défaut). Sous l'hypothèse, dite $(\mathcal{H})$, d'invariance des martingales par grossissement de filtration (voir par exemple [29]), il suffit en effet, pour tenir compte du défaut, d'utiliser un facteur d'actualisation *convenablement ajusté* par rapport au risque de défaut, au lieu du facteur d'actualisation au taux sans risque habituel, et d'introduire un terme de dividende fictif continument versé au taux γ, équivalent au recovery en cas de défaut. Le principal outil mathématique utilisé ici est celui de la *réduction de filtration* qui permet de se ramener à une économie virtuelle sans défaut.

Les travaux [3, 4] appliquent cette approche à forme réduite à des produits à clause(s) d'exercice et/ou rappel anticipé (options américaines et de type jeu), tant dans un cadre semimartingale abstrait [3] que dans un cadre markovien générique (pouvant comporter sauts, volatilité stochastique, changements de régimes..) [4].
Sur le plan pratique cette recherche apporte des réponses précises à des questions concrètes

comme la décomposition rigoureuse d'une obligation convertible en une composante obligataire pure et une composante optionnelle et la définition associée de spread et volatilité implicites d'une obligation convertible [2], ou encore la dérivation rigoureuse des équations du pricing et de la couverture d'une obligation convertible dans un modèle de diffusion avec saut au moment du défaut [5].

2.3 Équations backward et équations intégro-différentielles

Un des points importants mis en évidence par la recherche précédente est le rôle crucial joué dans ces problèmes par une certaine classe d'équations différentielles stochastiques rétrogrades (*équations backward*, voir El Karoui et al. [62] pour les liens généraux avec la finance) avec sauts, réfléchies ou doublement réfléchies dans les cas d'options américaines ou de type jeux. L'étude mathématique de ces équations backward, ainsi que des problèmes intégro-différentiels associés dans le cas Markovien, est traitée dans [15, 8].
Il est à noter que ces travaux, qui font l'objet de la partie IV de ce rapport, ont également des applications en crédit multi-name (voir partie V).

Dans le cadre d'une **collaboration** avec Anis Matoussi (Université du Maine), nous avons ainsi établi dans [15] des estimations a priori ainsi qu'un principe de comparaison qui faisaient défaut dans la littérature pour les équations différentielles stochastiques rétrogrades doublement réfléchies et à sauts, et qui sont la pierre angulaire de l'exploitation des résultats généraux de [2, 3] dans un cadre Markovien (voir [4, 5]).

Sur la base de ces estimations générales, on peut en effet montrer, dans le cas Markovien, la continuité de la solution de l'équation backward par rapport à la condition initiale du processus de facteurs sous-jacent (voir [8, Partie I]).

Il en découle ensuite (en utilisant également le principe de comparaison) la caractérisation analytique de la fonction de pricing (fonction valeur, dans un langage 'contrôle') comme unique solution du système d'équations intégro-différentielles avec obstacles associé (*équation de pricing* dans la suite, voir [8, Partie II]).
Ici il convient de préciser l'espace dans lequel on recherche une solution de l'équation de pricing. En effet, ne serait-ce qu'à cause des obstacles, on ne peut espérer l'existence d'une solution au sens classique (l'existence d'une solution classique faisant déjà défaut pour une simple option de put américain vanille dans le modèle de Black–Scholes). S'inscrivant dans la littérature (considérable) concernant les liens entre équations backward et équations aux dérivées partielles (voir par exemple Barles et al. [22] pour un cas de modèle avec sauts), on montre dans [8] que le processus d'état (première composante) de la solution de l'équation backward doublement réfléchie et à sauts est identifiable à l'unique *solution de viscosité* [46] (à croissance polynomiale dans les variables d'espace) de l'équation de pricing. L'équation est donc bien posée au sens des solutions de viscosité.
En marge de ce résultat, on établit également dans [8] la convergence vers la solution de tout schéma numérique satisfaisant certaines conditions de stabilité, consistence et monotonie (obtenues par adaptation à notre cadre des conditions de Barles et Souganidis [25]).

3 Modélisation de la dépendance et dérivés de crédit multi-name

Depuis une époque plus récente (2008), en lien en particulier avec la mise en place de la **Chaire Risque de crédit**, mes recherches se situent principalement dans le domaine du risque de crédit multi-name.

Par rapport à la situation du crédit single-name (voir section 2.1), les payoffs multi-name (multi- au sens parfois de quelques unités en ce qui concerne les *first-to-default swaps* par exemple, mais souvent de l'odre d'une centaine ou davantage, comme dans le cas des CDOs) posent un problème d'explosion combinatoire, à la fois au plan théorique et au plan pratique. De plus l'outil théorique utilisé dans le cas single-name pour se ramener au cas default-free, à savoir la réduction de filtration, devient très difficile à utiliser dans ce contexte. En effet l'hypothèse $(\mathcal{H})$ d'invariance des martingales par grossissement de filtration (voir section 2.1), peu coûteuse à un seul défaut, n'est typiquement pas vérifiée dans le cas multi-name. L'étude de ces problèmes de filtrations fait l'objet de la **thèse de Behnaz Zargari** (depuis septembre 2007; thèse en co-direction M. Jeanblanc–S. Crépey, en co-tutelle avec l'université de Sharif, Iran).

Face à ces difficultés on peut envisager trois types d'approches possibles.
Une première approche consiste à ne s'intéresser qu'à un aspect du modèle, suffisant pour traiter l'aspect *pricing* (pur, hors couverture). Ainsi, la connaissance de la loi jointe des instants de défaut vue de la date 0 détermine le prix en 0 de la plupart des *basket credit derivatives* (*CDOs*, k^{th}*-to-defaults*, etc.). Les *modèles à facteurs* (voir par exemple [29, 98]) consistent à postuler pour ces instants de défauts une loi jointe, généralement à base de copules, calibrable sur données de spreads de CDS et de tranches de CDO 'vanilles' observés sur les marchés. Mais ces modèles ne disent rien sur la *dynamique* de la loi des pertes, ni par conséquent sur le lien qui pourrait exister entre les prix correspondants et une quelconque notion de couverture.

Une seconde approche, dite *bottom-up*, consiste en des modèles dynamiques relatifs à une filtration incluant celles des indicatrices de défaut (approche directe sans réduction de filtration, puisque celle-ci n'est pas applicable en crédit multi-name). Dans le cas Markovien on peut alors dériver la loi des pertes du portefeuille par résolution du système d'équations de Kolmogorov associé — résolution par des méthodes déterministes si la dimension du problème le permet (problème à quelques noms), ou par des méthodes de simulation sinon (voir sections 3.2 et 5.2.2 plus bas).
C'est ainsi l'approche retenue dans [6] (voir aussi Frey et Backhaus [70], ou Herbertsson [82]), travail en **collaboration** avec Monique Jeanblanc, Tomasz Bielecki et Marek Rutkowski où on introduit un modèle dynamique de *migrations de ratings* d'un portefeuille de crédit (migrations possiblement influencées par des facteurs auxiliaires: variables de cycle..).

La troisième approche, plus récente (approches *top* et *top-down*, voir par Giesecke et Goldberg [74]) consiste à modéliser directement les sous-jacent des basket derivatives (typiquement: indices de CDS), plutôt que les constituants des indices dans les approches précédentes, quitte à redescendre aux CDS individuels par décomposition (dite *thinning*) de l'indice pour la couverture.

3.1 *Up and Down Credit Risk*

Les approches *top* et *top-down* sont à première vue séduisantes, dans la mesure où elles semblent ouvrir la voie à des modèles à la fois raisonnables en termes de coût de calcul (contrairement a priori aux approches bottom-up), et dynamiques, avec une notion de prix reliée à une notion de couverture.

L'objet du travail [1], en **collaboration** avec Monique Jeanblanc et Tomasz Bielecki, est d'approfondir les 'intuitions' précédentes concernant les approches top down. Ce travail fait apparaître de nombreux nombreuses difficultés, tant théoriques que pratiques: problèmes de filtrations (procédure de *thinning* difficilement soutenable en particulier dès lors que la filtration dans le modèle ne contient pas la filtration des indicatrices de défaut) ; problème de la détermination de dynamiques pertinentes pour les processus 'top', et risque de s'en tenir à des modèles simplistes, avec des conséquences dangereuses en termes de couverture.

3.2 Méthodes numériques

Au point de vue numérique le crédit multi-name pose un double enjeu de dimensionalité (au moins pour les modèles bottom-up) et d'évènements rares (rareté des évènements de défaut, même si ceci est à pondérer au vu des effets de contagion entre défauts, très sensibles sous la mesure de probabilité risque-neutre du pricing).

Si des formules analytiques sont parfois disponibles pour les produits vanilles (elles sont alors utilisées avec profit lors de la calibration du modèle), en revanche pour des produits plus exotiques les méthodes numériques sont incontournables. Les méthodes numériques déterministes étant exclues lorsque la dimension du modèle excède quelques unités (modèles bottom-up en particulier), on est alors conduit à utiliser des méthodes par simulation, avec les deux difficultés inhérentes dans le contexte du credit multi-name:
• problème des *évènements rares*: Dans [7] (travail en **collaboration** avec René Carmona) on étudie, notamment empiriquement sur la base d'études de cas, diverses techniques de type *importance sampling*, explicite quand le problème le permet (lorsqu'il peut donner lieu à des intuitions sur la mesure de probabilités à utiliser susceptible de privilégier les évènements rares considérés), ou implicite dans le cas contraire (méthodes de type *Interacting Particle Systems* à la Del Moral et al. [53, 54]);
• problème de la *dimension*: travail en cours avec Abdallah Rahal [16], voir section 5.2.2.

4 Principales contributions

Pour conclure cette revue des travaux effectués, on présente à présent une synthèse de leurs principales contributions à nos yeux. On peut ainsi retenir:

Au plan mathématique:

• Le Lemme 9.1 et le Théorème 9.2, relatifs aux propriétés de régularité (continuité et différentiabilité à l'ordre deux) de l'opérateur de pricing dans un modèle à volatilité locale 'irrégulière' ('mesurable') ;
• La Proposition 15.3, réduisant le problème du pricing d'une option de jeu vulnérable (*defaultable game option*) à celui d'une option de jeu standard (sans risque de crédit),

pour des valeurs ajustées (tenant compte du défaut) des taux d'intérêt et de dividendes ;
• La Proposition 16.4, qui donne une interprétation en termes de couverture du prix d'arbitrage d'une option de jeu vulnérable dans un modèle de marché 'général' (non Markovien, possiblement incomplet) ;
• Le Lemme 17.1, donnant la décomposition de la dynamique (de la partie martingale) du prix d'une option de jeu vulnérable, ainsi que le Lemme 17.2 analogue relatif aux prix des actifs primaires, et les implications de ces lemmes en termes de stratégies concrètes de couverture dans un cadre général (Propositions 17.3, 17.4 et 17.5) ou Markovien (Propositions 18.1 et 19.1) ;
• Les Théorèmes 22.2 et Propositions 23.1 établissant respectivement les estimations a priori et principe de comparaison pour une équation rétrograde doublement réfléchie dans un modèle 'assez général' (non Markovien, avec sauts) ;
• Le Théorème 26.3, qui établit (en utilisant les estimations et principe de comparaison abstraits précédents) le caractère bien posé de l'équation rétrograde doublement réfléchie correspondant au cas des applications typiques en finances (payoffs types options vanilles ou obligations convertibles dans des modèles Markoviens de diffusions à sauts et/ou changements de régimes) ;
• Le Théorème 27.1, qui établit le caractère bien posé du système d'équations intégro-différentielles avec double obstacle associé à l'équation rétrograde Markovienne précédente, ainsi que la convergence vers la solution de viscosité de ce système de tout schéma numérique stable, monotone et consistent ;

Au plan algorithmique:

• L'algorithme de calibration de volatilité locale de la section 11 (ou 12 pour le cas américain), basé sur la Proposition 10.2 de représentation du gradient de l'opérateur de pricing discrétisé par rapport à la fonction de volatilité (analogue discret du Théorème 9.3) ;
• Les algorithmes d'importance sampling et particulaires de la section 32 en crédit multi-name ;

Au plan du message 'ingénierie financière':

• La Proposition 13.1 établissant la supériorité d'un delta de volatilité locale quotidiennement recalibrée par rapport au delta de Black–Scholes implicite, pour la couverture d'une option par son sous-jacent dans un marché à skew persistent, positif ou négatif;
• Les analyses et conclusions de la partie 'crédit' (partie V):
– l'analyse de la section 33 relative à la comparaison entre un modèle de crédit *dynamique* calibré et un modèle statique en termes de couverture,
– l'importance d'utiliser un modèle *pertinent* pour cette dynamique, quitte au besoin à procéder par simulation (faute le cas échéant de formules fermées) pour pricer/calibrer,
– l'apport possible de techniques de *réduction de variance* de type importance sampling ou particulaires pour ces simulations.

5 Travaux en cours et projets de recherche

Dans la suite je souhaiterais donner aux travaux présentés ci-dessus les prolongements suivants.

5.1 *Pricing* versus *Greeking*

Les résultats de la section 2.3 relatifs aux équations du pricing (équations backward et intégro-différentielles) présentent la limite de ne concerner que la *fonction valeur* (fonction de pricing, dans la contexte de l'application financière) du problème de contrôle considéré, alors que la *stratégie* (correspondant formellement au *gradient* de la fonction valeur) est un enjeu au moins aussi important en pratique.

En effet, concernant l'application financière en salle de marchés, les prix de bons nombre de produits (dits 'vanilles') sont donnés de manière exogènes: ils s'imposent aux traders par l'effet de l'offre et de la demande, et sont cotés ouvertement sur les systèmes d'informations. Ces prix de marché sont utilisés comme input de la calibration des modèles (problème inverse du pricing consistant à déterminer un jeu de valeurs numériques des paramètres du modèle cohérent avec les observations de marché, cf. section 1.

L'objectif principal de la modélisation est alors la dérivation des paramètres de sensibilités, ou *Grecs*, les fonctions de pricing n'étant utilisées qu'au stade de la calibration. Ces Grecs, ou mesures des risques des produits financiers, sont ensuite utilisés pour déterminer la composition (dynamiquement dans le temps) du portefeuille de couverture.

Il apparaît donc important d'avoir des résultats, de convergence de schémas numériques notamment, concernant aussi bien la fonction valeur que son gradient. Concernant les schémas numériques par simulation, convergence de la fonction valeur et de son gradient (ou des processus de prix et de couverture, plus précisément) s'obtiennent typiquement conjointement (voir, par exemple, [36]). En ce qui concerne les schémas numériques déterministes, ce n'est pas nécessairement le cas, et notamment pas quand on entend les solutions des équations de pricing au sens des solutions de viscosité, pour lesquelles le gradient n'est a priori pas défini. Pour remédier à cette limitation des approches par solutions de viscosité, on se propose dans le cadre d'un **travail en cours** avec Anis Matoussi [14] de reformuler le problème de la relation backward/edp réfléchies et à sauts dans un contexte de solutions d'edp autre que celui des solutions de viscosité considéré dans [8], à savoir une certaine notion de solutions faibles (ou *solutions Sobolev*), introduite par Bally–Matoussi [21] et Barles–Lesigne [24].

État d'avancement Une première approche a été élaborée pour étendre à des modèles avec sauts, dans le cas de problèmes sans barrières, l'approche faible de Bally–Matoussi [21] et Barles–Lesigne [24]. Les aspects 'barrières' seront traités dans un second temps.

5.2 Méthodes numériques par simulation et *Curse of dimensionality*

Ce projet s'inscrit dans le cadre de la thèse de Abdallah Rahal, en co-direction M. Jeanblanc et S. Crépey, Université d'Évry, et Mustapha Jazar, Université libanaise. Une première approche est en cours [16]. Il s'agit d'étudier l'apport possible de méthodes de simulation/régression à la Longstaff–Schwartz [101] ou Gobet et al. [76] pour affronter les problèmes d'explosion combinatoire rencontrés aux sections précédentes, dans deux différents contextes.

5.2.1 *Path-dependence*

On a évoqué aux sections 2.3 et 5.1 les résultats théoriques de convergence de schémas numériques déterministes obtenus dans [8] concernant la fonction de pricing, et visés dans [14] pour son gradient. La mise en œuvre de ces schémas ne pose a priori pas de problème spécifique (en faisant tout de même attention aux difficultés liées aux sauts).
Cependant lorsqu'on considère les problèmes concrets de pricing liés aux obligations convertibles, il s'agit souvent de problèmes en grande dimension, du fait de clauses d'exercice anticipé hautement *path-dependent* (pouvant par exemple déboucher sur un modèle en dimension de l'ordre de trente, si l'on veut rendre compte finement des clauses du produit).

Au point de vue numérique les méthodes déterministes sont alors rendues inopérantes (*curse of dimensionality*). Il faut en ce cas avoir recours à des méthodes aléatoires basées sur des simulations de Monte Carlo — méthodes de type 'Monte Carlo américain' à la Longstaff–Schwartz [101] ou Gobet et al. [76] (voir aussi le Chapitre 8 de Glasserman [75] pour une référence générale), eu égard aux aspects 'arrêt optimal' du problème.

Autrement dit d'un point de vue plus mathématique, on privilégie alors l'approche probabiliste (approche équations backward plutôt qu'edp), et on s'intéresse à des schémas numériques de résolution des équations backward par simulation (voir, parmi d'autres, Touzi, Bouchard et al. [36, 35, 34], Gobet et al. [76], ou encore Chassagneux [41]). Mais la nature des clauses de path-dependence considérées sort de la littérature existante sur la discrétisation et la simulation d'équations rétrogrades réfléchies (avec dans les cas qui nous intéressent des barrières actives seulement sur des intervalles aléatoires en temps et non pour tout t), ce qui nécessite un travail spécifique d'adaptation et d'extension des résultats existants.
En relation avec la discussion de la section 2.3 concernant le calcul de la fonction de prix *versus* son gradient, il est intéressant de noter ici que les méthodes de résolution d'équations backward par simulation sont adossées à des propriétés de convergence concernant aussi bien le prix que la stratégie de couverture, ce qui les rend attractives également de ce point de vue (indépendamment des aspects dimensionalité).

5.2.2 *Bottom-Up Credit*

Une problématique de simulation/régression analogue à ci-dessus apparaît indépendamment dans un contexte, 'Poissonien' plutôt que 'Brownien', de chaînes de Markov en crédit multi-name. L'idée est alors de régresser dans la variable temporelle à état de la chaîne de Markov sous-jacente fixée, plutôt que dans les variables d'espace à date fixée dans Longstaff-Schwarz [101] ou Gobet at al. [76] (cf. section 5.2.1). On obtient ainsi une méthode de résolution par simulation du système d'équations différentielles de Kolmogorov associé à la chaîne, 'sur une région d'intérêt' de l'espace d'état (région définie par simulation de la chaîne de Markov à partir d'un point d'intérêt), en dimension potentiellement élevée (dans le cas de modèles *bottom-up*). Cette technique peut par exemple être mise à profit pour le calcul des Grecs sans resimulation.

5.3 Volatilité locale versus Intensité locale de défaut

Concernant le crédit multi-name (voir section 3 et partie V), un **projet de recherche** est envisagé avec Areski Cousin post-doctorant à Évry dans le cadre du **consortium industriel**

CRIS (projet d'élaboration d'une plate-forme indépendante de valorisation et gestion des produits dérivés de crédit, validé par le **pôle de compétitivité** mondial Finance Innovation, en collaboration avec Zeliade Systems, OTC-Conseils, JPLC, Dexia CL, Microsoft France et l'Université d'Évry).

L'idée est d'étendre aux marchés de corrélation de crédit l'analyse sur marchés de volatilité de [10]. Il s'agit de comparer en termes de couverture, sur une base empirique comme dans Cont et Kan [43], mais aussi sur la base d'une analyse mathématique esquissée à la section 33 du présent rapport, un modèle dynamique de crédit dynamique (en l'occurence le modèle *à intensité locale de défaut*, qui est au crédit multi-name ce que le modèle à volatilité locale de la section 1 est aux marchés de volatilité; voir par exemple Laurent et al. [97] ou Cont et Minca [44]), avec un modèle statique (modèle de copule Gaussien de Li [100], voir aussi Laurent [98]).

Part II

Local Volatility

Introduction

This part is a synthetic presentation of the papers [11, 12, 10, 9].

An important issue in quantitative finance is *model calibration.* The calibration problem is the *inverse* of the pricing problem. Instead of computing prices in a model with given values for its parameters, one wishes to compute the values of the model parameters that are consistent with observed prices. Now, it is well-known by physicists that such inverse problems are typically *ill-posed.* So, if one perturbs the data (e.g., if the observed prices move from some small amount between today and tomorrow), it is quite typical that a numerically determined best fit solution of the calibration problem switches from one 'basin of attraction' to the other, thus the numerically determined solution is*unstable.* To achieve robustness of model (re)calibration, we need to introduce some *regularization.* The most widely known and applicable regularization method is *Tikhonov(–Phillips)* [115, 109] regularization method. The paper [9] (see section 8 below) provides a survey on Tikhonov regularization applied to model calibration in finance.

Following an approach introduced by Lagnado and Osher [94], we study in [11, 12] the inverse problem, in finance, of calibrating a local volatility function from observed option prices, using Tikhonov regularization (see section 7 for a general presentation of this calibration problem). We consider this problem in two different settings: first, the generalized Black–Scholes model [11], and second, a trinomial tree discretization [12].
In [11] (see sections 6 and 9 below) we establish $W_p^{1,2}$-estimates for one-dimensional parabolic equations with measurable ingredients (Black–Scholes or Dupire equations with local volatility function used as a model diffusion coefficient). We deduce from these estimates that the method of Tikhonov regularization is applicable to the local volatility calibration problem. We thus prove the stability of the method, its convergence towards a minimum norm solution of the calibration problem, and discuss convergence rates issues.
In [12] we first establish analogous results in a trinomial tree discretization of the previous setting (section 10 below). Next (section 11) we present a parallel implementation of the method in the discrete setting, using a probabilistic interpretation to compute, at significantly reduced cost, the gradient of the cost criterion. Finally (section 12) we extend this methodology to the problem of calibration with American option prices.

In [10] (section 13 below) we compare the Profit and Loss arising from the delta-neutral dynamic hedging of options, using two possible values for the delta of the option. The first one is the Black–Scholes implied delta, while the second one is the *local* delta, namely the delta of the option in a generalized Black–Scholes model with a local volatility, recalibrated to the market smile every day (using a suitably regularized procedure for this calibration, like for instance the algorithm of [12]).
We explain why in negatively skewed markets the local delta should provide a better hedge than the implied delta during slow rallies or fast sell-offs, and a worse hedge, though to a lesser extent, during fast rallies or slow sell-offs. Since slow rallies and fast sell-offs are more likely to occur than fast rallies or slow sell-offs in negatively skewed markets (provided

we have physical as well as implied negative skewness), we conclude that on average the local delta provides a better hedge than the implied delta in negatively skewed markets. We obtain the same conclusion in the case of positively skewed markets.

6 Black–Scholes and Dupire Equations

Recall that a European *call* (respectively *put*) option with maturity date T and strike K, on an underlying asset S, means a right to buy (respectively sell), at price K, a unit of S at time T. We consider a theoretical financial market, with two traded assets: the savings account (riskless asset), with constant interest-rate r, and a risky stock, with price process S driven by a standard Brownian motion B under the historical probability $\mathbb{P}$. More precisely, the stock is assumed to obey the following dynamics:

$$dS_t = S_t(\mu_t dt + \sigma(t, S_t)dB_t) \ , \quad t > t_0 \ ; \quad S_{t_0} = S_0$$

for a (historical) *drift process* μ_t and a *volatility process* $\sigma(t, S_t)$. Moreover we postulate a continuously compounded dividend at constant rate q on S. Suppose finally the market to be liquid, non arbitrable and perfect. By standard arbitrage arguments and the Markov property of S, European vanilla calls/puts on S then have a theoretical fair price within the model that we will denote by $\Pi_{T,K}(t_0, S_0; a)$, where $a = \sigma^2/2$ and

$$\Pi_{T,K}(t_0, S_0; a) = e^{-r(T-t_0)}\mathbb{E}(S_T - K)^{+/-} \ . \tag{1}$$

Here $\mathbb{E}$ denotes the expectation with respect to the so-called *risk-neutral* probability $\mathbb{Q}$, under which

$$dS_t = S_t((r-q)dt + \sigma(t, S_t)dW_t) \ , \quad t > t_0 \ ; \quad S_{t_0} = S_0 \tag{2}$$

for a standard $\mathbb{Q}$ – Brownian motion W. Alternatively to their probabilistic representation (1)–(2), the prices Π can be expressed in terms of the solution to a differential equation. One can use either the *Black–Scholes*(–Merton) backward parabolic equation [33, 103] in the variables (t_0, S_0), which is

$$\begin{cases} -\partial_t\Pi - (r-q)S\partial_S\Pi - a(t,S)S^2\partial^2_{S^2}\Pi + r\Pi \ = \ 0, \quad t < T \\ \Pi|_T \ = \ (S-K)^{+/-} \ , \end{cases} \tag{3}$$

or the *Dupire* forward parabolic equation [58] in the variables (T, K), given by

$$\begin{cases} \partial_T\Pi - (q-r)K\partial_K\Pi - a(T,K)K^2\partial^2_{K^2}\Pi + q\Pi \ = \ 0, \quad T > t_0 \\ \Pi|_{t_0} \ = \ (S_0-K)^{+/-} \ . \end{cases} \tag{4}$$

Note that (1)–(2) can be viewed as the Feynman–Kac representation for the solution of (3). As for (4), it is but the Fokker–Planck equation for the transition probability density of S from (t_0, S_0) to (T, K), twice integrated with respect to K using the identity

$$\partial^2_{K^2}(S_0 - K)^{+/-} = \delta_{S_0}(K) \ ,$$

where δ_{S_0} denotes the Dirac mass at S_0.

7 Calibration Problem

In (2) and (3)–(4), the yields r and q are assumed to be known constants (they could in fact be any deterministic known functions of time). The local volatility function σ, or $a = \sigma^2/2$, is an unknown function of time and stock. The *local volatility calibration problem* is the inverse problem that amounts to inferring the local volatility function a from market-quoted prices of liquid options, typically European vanilla calls and puts with various strikes and maturities. The local volatility function thus inferred is then used to price *exotic* (non vanilla) options, and value hedge ratios or derivative exposure, consistently with the market. This problem, known as *fitting the smile* by market practitioners, is hence the reconstruction of a local volatility function, supposed to be prevailing as the underlier's risk-neutral dynamics. It is indeed important for applications that the reconstruction of such a prevailing dynamics be as fair as possible. But this calibration problem is under-determined (since the set of observed prices is finite) and ill-posed, so that *ad hoc* stabilizing procedures must be used. A variant of the problem, also considered in our work (*American calibration problem*, as opposed to the previous *European calibration problem*), consists of the calibration of a local volatility function with *American* option prices.

These calibration problems have received intensive study in the late 90's to early 2000's. In order to recover a well-posed problem, one needs to introduce some regularization. Let us thus mention, among so many other references, Avellaneda et al. [19], which use entropic regularization, or Achdou et al. [17], which use as state variable the prices Π in the variables (T, K) (cf. the Dupire equation (4)), with H^2-regularization and a finite element discretization.

In our work, we focus on the approach introduced by Lagnado and Osher [94], based on H^1-Tikhonov(–Phillips) regularization [115, 109, 64]. This approach tackles the calibration problem as a minimization problem, of a cost criterion J_α defined by

$$2J_\alpha(a) = d\left(\Pi|_{obs}(a), \pi\right)^2 + \alpha\rho\left(a, a_0\right)^2 . \tag{5}$$

Here $d\left(\Pi|_{obs}(a), \pi\right)$ denotes the Euclidean distance between the model prices $\Pi(a)$ and the observed prices π, α is the so-called *regularization parameter*, and ρ is a penalty designed to keep a close to the *prior* a_0 (a priori guess for a), namely $\rho(a, a_0)^2 = \|a - a_0\|^2_{H^1(Q)}$, where

$$\|u\|^2_{H^1(Q)} = \int\int_Q \left(u(t, y)^2 + \|\nabla u(t, y)\|^2\right) dtdy , \tag{6}$$

the $H^1(Q)$-squared norm of u, with $Q = (t_0, \overline{T}) \times \mathbb{R}$ (in which $\overline{T}$ stands for the largest maturity of an option in the calibration data set).

8 Tikhonov regularization of nonlinear inverse problems

The purpose of this section is to give a crash course on Tikhonov regularization (cf. Crépey [9]). We thus consider:
- a closed convex non-void subset $\mathcal{A}$ of a Hilbert space $\mathcal{H}$,
- a direct operator ('pricing functional')

$$\mathcal{H} \supseteq \mathcal{A} \ni a \xrightarrow{\Pi} \Pi(a) \in \mathbb{R}^d$$

(in which a represents the set of model parameters),
• noisy data ('observed prices') π^{δ}, and
• a *prior* $a_0 \in \mathcal{H}$ (a priori guess for a).

The Tikhonov regularization method for *inverting* Π *at* π^{δ}, or estimating the model parameter a given the observation π^{δ}, consists in:
• Reformulating the inverse problem as the following *nonlinear least squares problem*:

$$\min_{a\in\mathcal{A}} \left\|\Pi\left(a\right)-\pi^{\delta}\right\|^{2} \tag{7}$$

to ensure *existence* of solutions,
• Selecting the solutions of the previous nonlinear least squares problem that minimize $\|a-a_0\|^2$ over the set of all solutions, and
• Introducing a trade-off between accuracy and regularity, parameterized by a level of regularization $\alpha > 0$, to ensure *stability*.

More precisely, we introduce the following *cost criterion*:

$$2J_{\alpha}^{\delta}\left(a\right)=\left\|\Pi\left(a\right)-\pi^{\delta}\right\|^{2}+\alpha\left\|a-a_0\right\|_{\mathcal{H}}^{2}\ . \tag{8}$$

Definition 8.1 Given α, δ and a further parameter η, where η represents an error tolerance on the minimization, a *regularized solution to the inverse problem for* Π *at* π^{δ}, is any model parameter $a_{\alpha}^{\delta,\eta} \in \mathcal{A}$ such that

$$J_{\alpha}^{\delta}\left(a_{\alpha}^{\delta,\eta}\right) \le J_{\alpha}^{\delta}\left(a\right)+\eta\,, \quad a\in\mathcal{A}\,.$$

Under suitable assumptions, one can show that the regularized inverse problem is well-posed, as follows. We first postulate that the direct operator Π satisfies the following continuity assumption.

Assumption 8.2 (Compactness) $\Pi\left(a_n\right)$ converges to $\Pi\left(a\right)$ in $\mathbb{R}^d$ if a_n weakly-converges to a in $\mathcal{H}$.

We then have the following *stability* result.

Theorem 8.1 (Stability) *Let* $\pi^{\delta_n} \to \pi^{\delta}$, $\eta_n \to 0$ *when* $n \to \infty$. *Then any sequence of regularized solutions* $a_{\alpha}^{\delta_n,\eta_n}$ *admits a subsequence which converges in* $\mathcal{H}$ *towards a regularized solution* $a_{\alpha}^{\delta,\eta=0}$.

Assuming further that the data lie in the range of the model leads to *convergence* properties of regularized solutions to (unregularized) solutions of the inverse problem as $\alpha \to 0$. Let us then make the following additional assumption on Π.

Assumption 8.3 (Range property) $\pi \in \Pi(\mathcal{A})$.

By an a_0 – *solution* to the inverse problem for Π at π, we mean any $a \in \operatorname{Argmin}_{\{\Pi(a)=\pi\}} \|a - a_0\|$. Note that the set of a_0-solutions is non-empty, under Assumption 8.3.

Theorem 8.2 (Convergence; see, for instance, Theorem 2.3 of Engl et al. [65]) *Let the* perturbed parameters $\alpha_n, \delta_n, \eta_n$ *and the* perturbed data $\pi_n \in \mathbb{R}^d$ *satisfy*

$$(n \in \mathbb{N}) \quad \|\pi - \pi_n\| \leq \delta_n,$$

$$(n \to \infty) \quad \alpha_n\,, \quad \delta_n^2/\alpha_n\,, \quad \eta_n/\alpha_n \quad \longrightarrow \quad 0.$$

Then any sequence of regularized solutions $a_{\alpha_n}^{\delta_n,\eta_n}$ *admits a subsequence which converges in* $\mathcal{H}$ *towards an* a_0*-solution* a *of the inverse problem for* Π *at* π*. In particular, in case when this problem admits a* unique a_0*-solution* a*, then* $a_{\alpha_n}^{\delta_n,\eta_n}$ *converges to* a*.*

Remark 8.4 In the special case where the direct operator Π is linear, Tikhonov regularization thus appears as an approximating scheme for the *pseudo-inverse* of Π.

Finally, assuming further regularity of Π, one can get *convergence rates* estimates, uniform over all data $\pi \in \Pi(\mathcal{A})$ sufficiently close and smooth with respect to the prior a_0 (so that the additional *source condition* (9) is satisfied below). Let us thus make the following additional assumption on Π.

Assumption 8.5 (Twice Gateaux differentiability) There exists linear and bilinear forms $d\Pi\,(a)$ on $\mathcal{H}$ and $d^2\Pi\,(a)$ on $\mathcal{H}^2$ such that

$$\begin{array}{c}\Pi\,(a+\varepsilon h) = \Pi\,(a) + \varepsilon d\Pi\,(a)\cdot h + \frac{\varepsilon^2}{2} d^2\Pi\,(a)\cdot(h,h) + \mathrm{o}\left(\varepsilon^2\right) \quad ; \quad a, a+h \in \mathcal{A} \\ \|d\Pi\,(a)\cdot h\| \leq C\,\|h\|\,, \quad \left\|d^2\Pi\,(a)\cdot(h,h')\right\| \leq C\,\|h\|\,\|h'\| \quad ; \quad a \in \mathcal{A}\,, \quad h, h' \in \mathcal{H}\end{array}$$

where C is a constant independent of $a \in \mathcal{A}$.

In the following theorem the operator

$$d\Pi\,(a)^* : \mathbb{R}^d \ni \lambda \mapsto d\Pi\,(a)^*\,\lambda \in \mathcal{H}$$

denotes the *adjoint* of

$$d\Pi\,(a) : \mathcal{H} \ni h \mapsto d\Pi\,(a)\,h \in \mathbb{R}^d\,,$$

in the sense that we have, for every $(h, \lambda) \in \mathcal{H} \times \mathbb{R}^d$ (see [64]):

$$\langle h, d\Pi(a)^\star\lambda\rangle_{\mathcal{H}} \;=\; \lambda^\mathsf{T} d\Pi(a).h\;.$$

Theorem 8.3 (Convergence Rates; see, for instance, Theorem 10.4 of Engl *et al* [64]) *Assume*

$$(n \in \mathbb{N}) \quad \|\pi - \pi_n\| \leq \delta_n,$$

$$(n \to \infty) \quad \alpha_n \longrightarrow 0\,, \quad \alpha_n \sim \delta_n\,, \quad \eta_n = \mathrm{O}\left(\delta_n^2\right)\;.$$

Then $\|a_{\alpha_n}^{\delta_n,\eta_n} - a\| = \mathrm{O}(\sqrt{\delta_n})$*, for any* a_0*-solution* a *of the inverse problem for* Π *at* π *such that*

$$a - a_0 = d\Pi\,(a)^*\,\lambda \tag{9}$$

for some λ *sufficiently small in* $\mathbb{R}^d$ *(in particular, there exists at most one such* a_0*-solution* a*).*

Remark 8.6 An interesting feature of Tikhonov regularization is that the data π does not need to belong to the range of the direct operator for applicability of the method — even if Assumption 8.3 is the simplest assumption for the previous results regarding convergence and convergence rates (in fact a minimal assumption for such results is the existence of a least squares solution to the inverse problem, see Proposition 3.2 of Binder *et al* [32]).

For implementation purposes, the problem of minimizing (8) is discretized, thus becoming effectively a *nonlinear minimization problem* on (some subset of) $\mathbb{R}^M$ (see, e.g., Nocedal and Wright [105]), where M is the number of model parameters to be estimated.

Of course an important issue in practice is the choice of the *regularization parameter* α, that determines the trade-off between accuracy and regularity in the method.

Sections 9 and 10 give two concrete settings in which assumptions 8.2 and 8.5 will be shown to be satisfied, so that all the results in this section are applicable (provided Assumption 8.3 holds, regarding Theorems 8.2 and 8.3).
Section 11 examines extension of the results of section 10 to the American calibration problem.

To alleviate the notation, we assume in sections 9 to 11 that there are only calls in the calibration data.

9 Continuous setting

The first setting, considered in Crépey [11], is that of the generalized Black–Scholes model introduced at section 6.

9.1 The setting

Given a plane strip $Q = (t_0, \overline{T}) \times \mathbb{R}$, constant bounds $0 < \underline{a} \leq \overline{a}$, and a real measurable function (the *prior*) a_0 on Q such that $\underline{a} \leq a_0 \leq \overline{a}$, let us denote by

$$\mathcal{A} = \left\{ a \in a_0 + H^1(Q) \quad ; \quad \underline{a} \leq a \leq \overline{a} \right\} ,$$

where $H^1(Q)$ is the usual Sobolev space of measurable functions on Q such that $\|u\|^2_{H^1(Q)} < +\infty$ (cf. (6)).

Let us also be given a finite subset $\mathcal{F} \subset \overline{Q} = [t_0, \overline{T}] \times \mathbb{R}$ with $|\mathcal{F}| = d \in \mathbb{N}^\star$. We define Π as the *pricing functional*

$$\mathcal{A} \ni a \stackrel{\Pi}{\longmapsto} \Pi_{\mathcal{F}}(t_0, y_0 = 0; a) \in \mathbb{R}^d ,$$

where

$$\Pi_{T,k}(t_0, y_0; a) = e^{-r(T-t_0)} \mathbb{E}(S_T - K)^+$$

denotes, for $(T, k) \in \mathcal{F}$, the price of the European call with maturity T and strike $K = e^k$, at the current date t_0 and underlying asset value $S_0 = e^{y_0}$, in the generalized (risk-neutral) Black–Scholes model (2) expressed with the logarithmic variable

$$y = \ln(\frac{S}{S_0}) - (r - q - \underline{a})(t - t_0) . \tag{10}$$

Note that Crépey [11] uses the logarithmic variable $y = \ln(S)$, instead of y as in (10). The present choice of variables gives rise to better stability conditions after discretization of the problem (cf. (16) below). With this choice, all the results in Crépey [11] are easily seen to be applicable. In particular, introducing the process $Y_t = \ln(\frac{S_t}{S_0}) - (r - q - \underline{a})(t - t_0)$, we have,

Lemma 9.1 (Crépey [11]) *There exists $\overline{p} = \overline{p}(\underline{a}, \overline{a}) \in]2, 3[$, such that if $p \in]2, \overline{p}[$, then, as (t, y) varies in $[0, T] \times \mathbb{R}$,*

$$\Theta(t, y) \; = \; \mathbb{E}\Big(\int_{s=t}^{T} e^{-r(s-t)} \Gamma(s, Y_s) \; ds \,|\, Y_t = y \Big) \; , \tag{11}$$

where Γ is given in $L_p(Q)$, defines the unique $W_p^{1,2}(Q)$-solution (solution almost everywhere with related partial derivatives in $L_p(Q)$) of the following Black–Scholes equation with source term Γ :

$$\begin{cases} -\partial_t \Theta - (\underline{a} - a)\, \partial_y \Theta - a \partial^2_{y^2} \Theta + r\Theta \; = \; \Gamma \quad \textit{on } Q \\ \Theta|_T \; = \; 0 \; . \end{cases}$$

Moreover,

$$\|\Theta\|_{W_p^{1,2}(Q)} \; \leq \; C \; \|\Gamma\|_{L_p(Q)} \; , \tag{12}$$

where $C = C_p(\underline{a}, \overline{a})$.

Here the difficulty lies in the absence of regularity of the function a (these results are in fact established under the mere hypothesis that the diffusion coefficient a is a measurable function of t and y such that $\underline{a} \leq a \leq \overline{a}$). An important challenge in particular is to show the estimate in (12) for a bound C *uniform with respect to a such that* $\underline{a} \leq a \leq \overline{a}$ (and also, locally uniform with respect to r, q and T).
In few words, the proof of Lemma 9.1 (see [11]) uses $W_p^{1,2}$-estimates for one-dimensional parabolic equations from Krylov [92], Fabes [67] and Stroock and Varadhan [114], in combination with (classical [46] and L_p - [47, 40]) viscosity solutions arguments.

One can then show by application of Lemma 9.1 (by making use of suitable Sobolev embeddings),

Theorem 9.2 *The pricing functional Π of this section satisfies the regularity assumptions 8.2 and 8.5.*

Again the main challenge here is to show the estimates in Assumption 8.5 for a bound C *independent of $a \in \mathcal{A}$.*

9.2 The gradient

Furthermore, we have the following gradient's representation,

Theorem 9.3 (Crépey [11]) *For $(T, k) \in \mathcal{F}$, the derivative of $\Pi_{T,k}(t_0, y_0; a)$ in the direction $h \in H^1(Q)$, admits the following Feynman–Kac representation:*

$$\begin{aligned} & d\Pi_{T,k}(t_0, y_0; a) \cdot h \\ & \quad = \int_{t=t_0}^{T} \int_{y=-\infty}^{\infty} h(t, y) \Gamma_{T,k}(t, y; a) \gamma_{t_0,y_0}(t, y; a) \; dy \; dt \; , \end{aligned} \tag{13}$$

where $\Gamma_{T,k}(t,y;a) = (\partial^2_{y^2} - \partial_y)\Pi_{T,k}(t,y;a)$, *and* $\gamma_{t_0,y_0}(t,y;a)$ *denotes, for almost every* $t > t_0$, *the transition probability density between* t_0 *and* t *discounted at rate* r, *that is,* $e^{-r(t-t_0)}\times$ *the density, of the process* $Y_t = \ln(\frac{S_t}{S_0}) - (r - q - \underline{a})(t - t_0)$.

Let Δ denote the Laplacian operator on $H^2(Q)$. In the set-up of this section, it is possible to give a more explicit formulation to the abstract source condition (9). So,

Theorem 9.4 (Crépey [11]) *Assuming moreover that* a *is uniformly continuous with respect to its space variable* y, *then, in the context of this section, condition (9) means that* $\Lambda = a - a_0$ *is the unique strong solution in* $H^2(Q)$ *of the following (nonlocal) problem:*

$$\left\{ \begin{array}{l} \Lambda - \Delta\Lambda = \sum_{(T,k)\in\mathcal{F};\ t\le T} \lambda_{T,k}\Gamma_{T,k}(t,y;a)\gamma_{t_0,y_0}(t,y;a), \quad Q\text{-a.e.} \\ \partial_n\Lambda = 0, \ \partial Q\text{-a.e.} \end{array} \right. \tag{14}$$

where the normal derivative $\partial_n\Lambda \in L_2(\partial Q)$ *is well defined, for* $\Lambda \in H^2(Q)$.

Finally, let ∇ denote the Gateaux derivative in $H^1(Q)$. The following result does not appear explicitly in Crépey [11], but it can be derived in the same manner as Theorem 9.4 above.

Theorem 9.5 *If* a *is uniformly continuous with respect to its space variable* y, *then* $u = \nabla J_\alpha(a) - \alpha(a - a_0)$ *belongs to* $H^2(Q)$, *and* u *is the unique strong solution in* $H^2(Q)$ *of the following problem:*

$$\left\{ \begin{array}{l} u - \Delta u = \sum_{(T,k)\in\mathcal{F};\ t\le T}(\Pi_{T,k}(t_0,y_0;a) - \pi_{T,k})\Gamma_{T,k}(t,y;a)\gamma_{t_0,y_0}(t,y;a), \quad Q\text{-a.e.} \\ \partial_n u = 0, \ \partial Q\text{-a.e.} \end{array} \right. \tag{15}$$

where the normal derivative $\partial_n u \in L_2(\partial Q)$ *is well defined, for* $u \in H^2(Q)$.

For simplicity we assume in the next two sections that all maturities with observed prices fall at steps of a constant time subdivision $(t_0, t_1, \ldots, t_p = \overline{T})$ with time step τ of $[t_0, \overline{T}]$, where $\overline{T} = \max_{(T,k)\in\mathcal{F}} T$.

10 Discrete Setting

It turns out that there is a natural discretization of the setting of section 9, which keeps all the required properties. The idea is to use a Markov chain algorithm to specify the same problem in a fully discrete setting. In order to handle the key point that the local volatility may vary within a range, we adopt a trinomial tree method where the mesh is fixed once for all, and the local volatility varies from node to node.

10.1 The setting

Define $\underline{\sigma}$ and $\overline{\sigma}$ such that $\underline{a}, \overline{a} = \underline{\sigma}^2/2, \overline{\sigma}^2/2$. We choose a Markov chain $(Y_n)_{0\le n\le N}$ with time step $\tau = (\overline{T} - t_0)/N$ and space step $\varepsilon = \beta\overline{\sigma}\sqrt{\tau}$, where β is some fixed parameter, the so-called *stretch factor*. Starting from $Y_0 = 0$, we look for a scheme with, at each node (t_n, y_m) of the tree, a local transition probability

$$p^+_{n,m} = P(Y_{n+1} = y_m + \varepsilon \mid Y_n = y_m)\ ,\ p^-_{n,m} = P(Y_{n+1} = y_m - \varepsilon \mid Y_n = y_m)\ ,$$

so that $P(Y_{n+1} = y_m \mid Y_n = y_m) = 1 - p^+_{n,m} - p^-_{n,m}$.

Then, given a local node volatility $a_{n,m} = a(t_{n+1}, y_m) \in [\underline{a}, \overline{a}]$, it is easy to show that the choice:

$$p^+_{n,m} = (\frac{a_{n,m}}{\varepsilon^2} - \frac{(a_{n,m} - \underline{a})}{2\varepsilon})\tau \ , \ p^-_{n,m} = (\frac{a_{n,m}}{\varepsilon^2} + \frac{(a_{n,m} - \underline{a})}{2\varepsilon})\tau$$

will yield nonnegative weights as long as the following stability conditions hold true:

$$\varepsilon \leq \frac{2\overline{a}}{(\overline{a} - \underline{a})} \ , \ 1 \leq \beta \ ; \tag{16}$$

and also, that the first and second moments of the Markov chain, after the change of variables

$$y \ \mapsto \ S = S_0 \exp(y + (r - q - \underline{a})(t - t_0)) \ ,$$

will match those of the continuous diffusion (2) with an $o(\tau)$ accuracy as $\tau \to 0$.

We denote by $\mathcal{T}$ the trinomial time-space tree supporting the evolution of the Markov chain (Y_n), and by $\mathcal{I}$ the sub-tree of $\mathcal{T}$ starting at (t_1, y_0). Now we consider the same setting as in section 9.1, except that:
• $\mathcal{A}$ now refers to the set of all possible functions $a = \sigma^2/2$ on $\mathcal{I}$ with $\underline{a} \leq a \leq \overline{a}$, and
• Π means the price, or discounted expectation of the payoff function, of a call in the tree with local volatility $\mathcal{T}$.
Obviously, $\mathcal{A}$ can be identified in a natural way with the product set $[\underline{a}, \overline{a}]^{N^2}$. Moreover, denoting by α_t and α_y auxiliary regularization parameters to be defined later, and by $\mathcal{B}$ the set of bottom nodes of $\mathcal{I}$ at the different time steps, we endow $\mathcal{A}$ with the inner product

$$\langle u, v \rangle_{h^1} = \alpha_t D_t(u, v) + \alpha_y D_y(u, v) \ , \tag{17}$$

where

$$D_t(u,v) \ = \ \tau^{-1}\varepsilon \sum_{(t_n,y_m)\in\mathcal{I}} \left(u\left(t_{n-1},y_m\right) - u\left(t_n,y_m\right)\right)\left(v\left(t_{n-1},y_m\right) - v\left(t_n,y_m\right)\right) \tag{18}$$

$$\begin{aligned} D_y(u,v) \ = \ \tau\varepsilon^{-1} \Bigg(& \sum_{(t_n,y_m)\in\mathcal{I}} \left(u\left(t_n,y_{m+1}\right) - u\left(t_n,y_m\right)\right)\left(v\left(t_n,y_{m+1}\right) - v\left(t_n,y_m\right)\right) \\ & + \sum_{(t_n,y_m)\in\mathcal{B}} \left(u\left(t_n,y_m\right) - u\left(t_n,y_{m-1}\right)\right)\left(v\left(t_n,y_m\right) - v\left(t_n,y_{m-1}\right)\right) \Bigg) \ , \end{aligned} \tag{19}$$

in which homogeneous boundary conditions are assumed, Neumann at the root node of $\mathcal{T}$ and Dirichlet at the other boundary nodes of $\mathcal{T}$.

In this setting, the pricing functional $a \mapsto \Pi\,(a)$ is multilinear, hence continuous. It is then easy to check the following result, which is the discrete analog of Theorem 10.1.

Proposition 10.1 *The pricing functional* Π *of this section satisfies the regularity assumptions 8.2 and 8.5.*

The abstract results of section 8 are thus applicable to the discretized problem as well.

10.2 The gradient

Furthermore, elementary computations yield the following,

Proposition 10.2 (Crépey [12]) *For $(T,k) \in \mathcal{F}$, the partial derivative of $\Pi_{T,k}(t_0, y_0; a)$ with respect to the value of the local volatility function a at node $(t_n, y_m) \in \mathcal{I}$ with $t_n \le T$, admits the following Feynman–Kac representation:*

$$\begin{aligned} & d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m) \\ & \quad = \Gamma_{T,k}(t_n, y_m; a)\gamma_{t_0,y_0}(t_{n-1}, y_m; a) \ \exp(-r\tau) \ \tau \ , \end{aligned} \tag{20}$$

where

$$\begin{aligned} & \Gamma_{T,k}(t_n, y_m; a) = \\ & \quad (\tfrac{1}{\varepsilon^2} + \tfrac{1}{2\varepsilon})\Pi_{T,k}(t_n, y_{m-1}; a) - \tfrac{2}{\varepsilon^2}\Pi_{T,k}(t_n, y_m; a) + (\tfrac{1}{\varepsilon^2} - \tfrac{1}{2\varepsilon})\Pi_{T,k}(t_n, y_{m+1}; a) \ , \end{aligned}$$

and γ_{t_0,y_0} denotes the map of Arrow–Debreu prices in $\mathcal{T}$ at the current state (t_0, y_0).

We now define the following discretized Laplacian operator on $\mathbb{R}^{N^2}$:

$$\begin{aligned} \Delta^\varepsilon_\tau u\,(t_n, y_m) &= \alpha_t \left(u\,(t_{n-1}, y_m) - 2u\,(t_n, y_m) + u\,(t_{n+1}, y_m)\right)/\tau^2 \\ & \quad + \alpha_y \left(u\,(t_n, y_{m-1}) - 2u\,(t_n, y_m) + u\,(t_n, y_{m+1})\right)/\varepsilon^2 \ , \ (t_n, y_m) \in \mathcal{I} \end{aligned}$$

with homogeneous boundary conditions, Neumann (respectively Dirichlet) at the origin and at an artificial $t_{N+1} = \overline{T} + \tau$ time step of the tree (respectively at the other boundary nodes of $\mathcal{T}$).

Proposition 10.3 (Crépey [12]) *In the context of this section, condition (9) means that $\Lambda = a - a_0$ is the unique solution in $\mathbb{R}^{N^2}$ of the following (nonlocal) problem:*

$$\begin{aligned} & -\varepsilon\Delta^\varepsilon_\tau \Lambda(t_n, y_m) = \\ & \sum_{(T,k)\in\mathcal{F};\ t_n \le T} \lambda_{T,k}\Gamma_{T,k}\,(t_n, y_m; a)\,\gamma_{t_0,y_0}\,(t_{n-1}, y_m; a)\exp(-r\tau) \ , \ (t_n, y_m) \in \mathcal{I} \ . \end{aligned} \tag{21}$$

Finally, let j_α denote the trinomial tree analog of the cost criterion J_α in (5), and let ∇ denote the gradient with respect to the h^1-inner product (17) in $\mathbb{R}^{N^2}$.

Proposition 10.4 (Crépey [12]) *$u = \nabla j_\alpha(a) - \alpha(a - a_0)$ is the unique solution in $\mathbb{R}^{N^2}$ of the following problem:*

$$\begin{aligned} -\varepsilon\Delta^\varepsilon_\tau u.(t_n, y_m) &= \sum_{(T,k)\in\mathcal{F};\ t_n \le T} (\Pi_{T,k}(t_0, y_0) - \pi_{T,k}) \\ & \quad \Gamma_{T,k}\,(t_n, y_m; a)\,\gamma_{t_0,y_0}\,(t_{n-1}, y_m; a)\exp(-r\tau) \ , \ (t_n, y_m) \in \mathcal{I} \ . \quad \square \end{aligned} \tag{22}$$

Remark 10.1 Observe the analogy between the identities (20), (21), (22) and (13), (14), (15), respectively.

11 A Trinomial Tree variant of the Lagnado and Osher Algorithm

Hence, a natural way to tackle numerically the Tikhonov regularized calibration problem in the tree, would consist in η-minimizing j_α with respect to $a \in [\underline{a}, \overline{a}]^{N^2}$, where $a = \sigma^2/2$ and N is the number of time steps in the tree, using the gradient with respect to the h^1-inner product (17) in $\mathbb{R}^{N^2}$. This gradient could be computed by solving (22) by standard methods.
In practice we use instead the *Euclidean* gradient, corresponding, as is seen immediately from above, to (cf. (22))

$$\widetilde{\nabla} j_\alpha(a) = \tau \sum_{(T,k)\in\mathcal{F};\ t_n \le T} (\Pi_{T,k}(t_0, y_0) - \pi_{T,k})$$
$$\Gamma_{T,k}(t_n, y_m; a)\, \gamma_{t_0,y_0}(t_{n-1}, y_m; a) \exp(-r\tau) - \alpha\varepsilon\Delta_\tau^\varepsilon(a - a_0)\ .$$

We thus spare the computational cost of solving (22). But our main point here is that sophisticated black-box bound-constrained gradient descent minimization routines, such as, for instance, `lbfgs` (see, e.g., Nocedal and Wright [105]), require the Euclidean gradient as their argument.

Moreover, in order to take care of normalization, we minimize a functional slightly different from j_α, namely $\widetilde{j}_\alpha$ defined by (cf. (17) for the definition of $\|\cdot\|_{h^1}$):

$$2\widetilde{j}_\alpha(a) = \frac{1}{d} \sum_{(T,k)\in\mathcal{F}} \left(\frac{\Pi_{T,k}(t_0, y_0; a) - \pi_{T,k}}{\omega_{T,k}} \right)^2 + \|a - a_0\|_{h^1}^2\ , \tag{23}$$

with, for every $(T, k) \in \mathcal{F}$:

$$\omega_{T,k} = \max(\Pi_{T,k}(t_0, y_0; \overline{a}) - \pi_{T,k}\ ,\ \pi_{T,k} - \Pi_{T,k}(t_0, y_0; \underline{a}))\ .$$

The regularization parameters α_t and α_y in $\|\cdot\|_{h^1}$ (cf. (17)) are chosen in a heuristic way (see [12] for the detail) devised to balance the contributions of the quadratic residual and penalty terms in the cost criterion at the minimum, so as to realize a fair compromise between accuracy and stability in the method.

With respect to Lagnado and Osher's algorithm [94], the main interest of this tree implementation follows from the probabilistic representation (20). Indeed this representation allows one to compute the gradient of the cost criterion by pricing the options and solving *one* Fokker–Planck equation in the tree, instead of pricing the options and solving one Black–Scholes equation with source term *by option and mesh node* in the Lagnado–Osher original presentation (see the detailed discussion in Crépey [12]). The accuracy of the Lagnado–Osher algorithm is preserved, but the computational time is drastically reduced. Typically the time can be reduced from about one hour to about one minute or less on a standard serial Pentium PC.
Moreover a parallel implementation allows one to gain a further factor. To do so, one shares between the available processors, for each maturity with observed prices, the computations relative to the options with various strikes. This can be done by using for instance the `MPI` library.

Another interest of this tree implementation is that explicit finite differences computations in the tree are less costly than implicit methods for computing option prices. Of course, explicit schemes are subject to the stability condition (16). But, to handle this condition, one only needs to take a space step $\varepsilon \leq \frac{2\overline{a}}{(\overline{a}-\underline{a})}$, and then a time step τ such that $\overline{\sigma}\sqrt{\tau} \leq \varepsilon$.

12 The American calibration problem

In the continuous setting of the generalized Black–Scholes model, it is an open question whether or not the theoretical results of Crépey [11] (cf. section 9) can be extended to the American calibration problem. Therefore we directly move to the discrete setting of section 10. Let us denote by π the call payoffs, namely

$$\pi_{T,k}\left(t_n, y_m\right) = (S_0 e^{(y_m + (r-q-\underline{a})(t_n - t_0))} - e^k)^+ \ , \ (T,k) \in \mathcal{F} \ .$$

We thus consider the pair

$$(\Pi, \widetilde{\Pi}) = (\Pi_{T,k}\left(t_n, y_m; a\right), \widetilde{\Pi}_{T,k}\left(t_n, y_m; a\right)) \ , \ (T,k) \in \mathcal{F} \ ,$$

jointly defined by the terminal condition $\Pi(T, \cdot; a) = \pi(T, \cdot)$, and for $n = N, \ldots, 1$:

i. $\widetilde{\Pi}(t_{n-1}, \cdot; a)$ defined from $\Pi(t_n, \cdot; a)$ in the same way as $\Pi(t_{n-1}, \cdot; a)$ from $\Pi(t_n, \cdot; a)$ in the European case;

ii. $\Pi(t_{n-1}, \cdot; a) = \max(\widetilde{\Pi}(t_{n-1}, \cdot; a), \pi(t_{n-1}, \cdot))$.

We denote by $\mathcal{C}^{n-1}_{T,k}(a)$, and call *Continuation Region* at time t_{n-1}, the set of all y_m such that $\widetilde{\Pi}_{T,k}\left(t_{n-1}, y_m; a\right) \geq \pi\left(t_{n-1}, y_m\right)$. We qualify the local volatility function a as *regular*, if this inequality is strict for every $y_m \in \mathcal{C}^{n-1}_{T,k}(a)$, for every $n = 1, \ldots, N$.

Remark 12.1 Since the *non regular* local volatility functions are to be sought among the solutions of one out of a finite number of equations, namely one for each node of $\mathcal{I}$, it is reasonable to expect that most local volatility functions are regular, or, more precisely, that the regular volatility functions form a full Lebesgue measure subset of $\mathcal{A}$.

An important difference between European and American call/put option prices is that in the exercise (non continuation) region of the American options, their prices are locally constant with respect to the local volatility function. Moreover, the following result regarding American option prices shows that these are less regular than the European prices with respect to the volatility function a.

Proposition 12.1 (Crépey [12]) *For every $(T,k) \in \mathcal{F}$, the American option price $\Pi = \Pi_{T,k}(t_0, y_0; a)$ is Lipschitz continuous and directionally differentiable with respect to the local volatility function $a \in \mathcal{A}$. Moreover, if a is regular, then Π is Gateaux differentiable at a, and the partial derivative of Π with respect to the value of the local volatility function a at node $(t_n, y_m) \in \mathcal{I}$ with $t_n \leq T$, admits the following Feynman–Kac representation:*

$$\begin{aligned} & d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m) \\ & \quad = \ \Gamma_{T,k}(t_n, y_m; a)\gamma^{T,k}_{t_0,y_0}(t_{n-1}, y_m; a) \ \exp(-r\tau) \ \tau \ , \end{aligned} \tag{24}$$

where

$$\Gamma_{T,k}(t_n, y_m; a) = \\ (\tfrac{1}{\varepsilon^2} + \tfrac{1}{2\varepsilon})\Pi_{T,k}(t_n, y_{m-1}; a) - \tfrac{2}{\varepsilon^2}\Pi_{T,k}(t_n, y_m; a) + (\tfrac{1}{\varepsilon^2} - \tfrac{1}{2\varepsilon})\Pi_{T,k}(t_n, y_{m+1}; a) \ ,$$

and $\gamma^{T,k}_{t_0,y_0}(t_{n-1}, y_m; a)$ *means the probability, discounted at rate* r*, that the Markov chain* y_n *goes from* (t_0, y_0) *to* (t_{n-1}, y_m) *through the continuation region* $\mathcal{C}_{T,k}(a)$*.*

So in particular $\gamma^{T,k}_{t_0,y_0}(t_{n-1}, y_m; a)$ is equal to 0 if $y_0 \notin \mathcal{C}^0_{T,k}(a)$ or $y_m \notin \mathcal{C}^{n-1}_{T,k}(a)$.

As in the European case, the *explicit* finite differences, or trinomial tree setting, is interesting because of the *probabilistic representation* (24) for the derivatives, in which the $\gamma^{T,k}$ can be computed forward in the tree, using Fokker–Planck discrete equations that express the composition of discounted probabilities in $\mathcal{C}_{T,k}(a)$. This results in a reduced cost computation procedure for these derivatives. Since γ now depends on T and k, one must solve one Fokker–Planck equation *by option*, instead of one Fokker–Planck equation as a whole in the European case. But this is still one or two orders of magnitude faster than solving one equation with source term *by option and mesh node* if one uses the Dynamic Programming characterization for the partial derivatives directly, without passing by the Fokker-Planck equation. Moreover, our computation for the $\gamma^{T,k}$ can be parallelized, in the same way as the one for the option prices (cf. section 11).

Since Π is continuous, the stability and convergence theorems 8.1 and (under assumption 8.3) 8.2 are applicable. But as Π is not everywhere differentiable, we cannot apply theorem 8.3 anymore.

12.1 American Calibration Algorithm

Now the calibration algorithm is the same as in the European case, except that we provide the minimization descent routines with the right-hand side of (24) instead of the Euclidean gradient of Π, knowing that both of them coincide, at the (presumably) full set of regular local volatility functions in $\mathcal{A}$. The overall computation cost of the calibration does not exceed twice the one required in the European case, both in the serial and in the MPI-parallel implementation.

For extensive reports on **numerical experiments** we refer the reader to [12]. Numerical experiments on real data sets involving several hundreds of input prices, support strong evidence that there is no trade-off between stability and accuracy when using this regularized calibration procedure, both in the European and in the American case. Except those corresponding to the shortest maturities, for which calibration is irrelevant, most prices are calibrated up to a few centimes of implied volatility, while the local and implied volatility functions thus calibrated exhibit satisfactory regularity and stability properties (see, e.g., Figure 1). The stability of the results contrasts with the instability that occurs if elementary reconstruction procedures are used for the local volatility function.
Finally the parallelization of the algorithm mentioned at the end of section 11 improves the speed of the algorithm by a further factor.

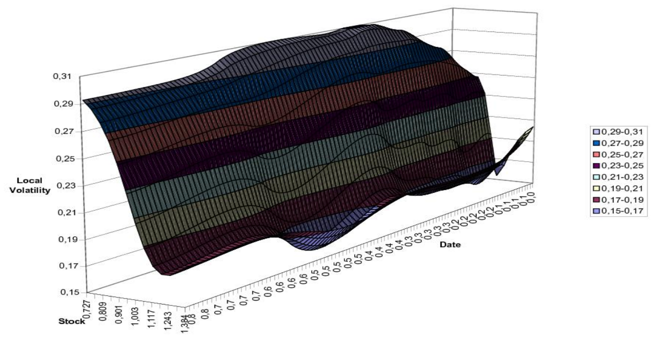

Figure 1: *Local Volatility calibrated on the DAX Index vanilla options, June 1 2001.*

13 Delta Hedging Vega Risk?

In this section we consider an agent who is short of one option with market price process Π on an underlyer S. Assume the riskless interest rate r in the economy and the dividend yield q on S to be zero, for notational simplicity. Let $T_1 \le T$ (the option's maturity). Practically, delta-hedging the option with the underlyer (and the riskless asset) on the time interval $[0, T_1]$ consists in rebalancing in a self-financed way, at every point in time of a subdivision (possibly random, though this is not the point here) $0 = t_0 \le t_1 \le \ldots \le t_p = T_1$ of $[0, T_1]$, a complementary position Δ in S, in order to minimize the overall exposure to 'small' moves of the underlying asset S.

The *tracking error*, or *profit-and-loss* (P&L for short) trajectory $e = (e_{t_k})_{0 \le k \le p}$, is obtained by adding up the following increments, starting with $e_0 = 0$, from $k = 0$ to $p - 1$:

$$\delta_k e = -\delta_k \Pi + \Delta_{t_k} \, \delta_k S \;, \tag{25}$$

where $\delta_k \Pi = \Pi_{t_{k+1}} - \Pi_{t_k}$, $\delta_k S = S_{t_{k+1}} - S_{t_k}$ and Δ_{t_k} is the number of units of S in the hedging portfolio on the time interval $(t_k, t_{k+1}]$.

In [10], our aim is to compare two delta-hedging strategies, with Δ_t respectively given by:
• The *Black–Scholes implied delta* of the option, that is

$$\Delta_t = \Delta_t^{bs} = \partial_S \Pi^{bs}(t, S_t, \Sigma_t)$$

where the function Π^{bs} and the number Σ_t stand for the Black–Scholes pricing function of the option and the market Black–Scholes implied volatility of the option at date t;
• Or, alternatively, the *local delta* of the option, that is

$$\Delta_t = \Delta_t^{lo} = \partial_S \Pi^{lo}(t, S_t, \sigma_t)$$

where Π^{lo} is the pricing function of the option in the local volatility model with volatility function σ_t calibrated to the full market vanilla Black–Scholes implied volatility surface observed at date t.

Remark 13.1 In the first bullet point we tacitly assume that the option's Black–Scholes implied volatility is well defined. This is for instance the case for European vanillas (European call and put options), provided their market prices lie within the arbitrage range.

We aim at determining which $\Delta = \Delta^{bs}$ or Δ^{lo} maintains the P&L trajectory e closest to 0 throughout the hedging period. The analysis of this section (see [10]) can be summarized as follows,

Proposition 13.1 (Crépey [10]) *When hedging an option with its underlyer (and the savings account), the local delta Δ^{lo} provides a better hedge than the Black–Scholes implied delta Δ^{bs}.*
At least this holds true in persistently skewed markets (positively or negatively), in terms of the risk-neutral variance of the delta-hedged P&L (and of the objective variance of the delta-hedged P&L as well, provided that the physical *as well as the risk-neutral market are persistently positively or negatively skewed).*

So, to make it short, in a persistently positively or negatively skewed market, the local delta works better than the implied delta for hedging an option, on average.

To develop our analysis we shall distinguish four stylized market regimes (see Crépey [10], Derman [55]): *fast rallies* (underlyer S quickly increasing), *slow sell-offs* (S slowly decreasing), *slow rallies* (S slowly increasing) and *fast sell-offs* (S quickly decreasing).
Though we were not able to formulate and establish this in a formalized and mathematical way, it is quite intuitive that, in a negatively skewed market, slow rallies and fast sell-offs are more likely to occur than fast rallies or slow sell-offs (provided we have *physical* as well as implied negative skewness). ***We take this for granted in the derivation of Proposition 13.1***, so there is still need for improvement here.

13.1 Analysis in a Local Volatility Model

In this section we operate in a theoretical market which would be given as a fixed *local volatility market model*, so

$$\Pi_t = \Pi^{lo}(t, S_t)\ ,\ \Sigma_t = \Sigma^{lo}(t, S_t)\ ,$$

for suitable *pricing functions* $\Pi^{lo}(t, S)$, $\Sigma^{lo}(t, S)$.

Note that since we are presently in the set-up of a local volatility model, the strategy Δ^{lo}, if applied in continuous time, would provide a perfect replication of the option by the underlyer (P&L trajectory e identically equal to zero). But we only consider *hedging in discrete time* here.

We first consider the problem of hedging a *vanilla* option in a *negatively skewed* local volatility market model. Since a vanilla option is Gamma *positive*, thus one can show that

$$\delta e^{lo} \text{ is negative at fast market regimes and positive at slow market regimes .} \tag{26}$$

Moreover, one has $\Pi^{lo}(t, S_t) = \Pi^{bs}(t, S_t, \Sigma^{lo}(t, S_t))$ by definition of the Black–Scholes implied volatility, so by chain differentiation:

$$\partial_S \Pi^{lo}(t, S_t) = \partial_S \Pi^{bs}(t, S_t, \Sigma_t) + \partial_\Sigma \Pi^{bs}(t, S_t, \Sigma^{lo}(t, S_t)) \partial_S \Sigma^{lo}(t, S_t)\ .$$

Now, for a vanilla option, one has $\partial_\Sigma \Pi^{bs} \geq 0$, and in the case of a negatively skewed local volatility market one can show that $\partial_S \Sigma^{lo} \leq 0$ (see [10]). Therefore,

$$\Delta_t^{lo} = \partial_S \Pi^{lo}(t, S_t) \leq \partial_S \Pi^{bs}(t, S_t, \Sigma_t) = \Delta_t^{bs} \ ,$$

and therefore, given (25),

$$\delta e^{bs} \leq \delta e^{lo} \text{ iff } \delta S \leq 0 \ . \tag{27}$$

Combining (26) and (27), we get the complete picture depicted in Table 1. As a consequence, Δ^{lo} provides a better hedge (δe closer to zero) than Δ^{bs} during fast sell-offs or slow rallies, and a worse hedge, though to a lesser extent, during slow sell-offs or fast rallies.

Market regime	**Rally**	**Sell-Off**
Slow	$0 \leq \delta e^{lo} \leq \delta e^{bs}$	$(\delta e^{bs})^+ \leq \delta e^{lo}$
Fast	$\delta e^{lo} \leq -(\delta e^{bs})^-$	$\delta e^{bs} \leq \delta e^{lo} \leq 0$

Table 1: *Vanilla option in a negatively skewed local volatility model.*

Since slow rallies and fast sell-offs are more likely to occur than fast rallies or slow sell-offs in negatively skewed markets (provided we have physical as well as implied negative skewness), the following conclusion follows,

Proposition 13.2 (Crépey [10]) *For a vanilla option in a negatively skewed local volatility market model, and provided we have* physical *as well as implied negative skewness, the local delta provides a better hedge than the implied delta* on average*, as well as* on average conditionally on the fact that the market is in a fast regime*, or* on average conditionally on the fact that the market is in a slow regime.

In the case of an *exotic option* with *negative Gamma/Vega exposure* (such as a reverse barrier option in the neighbourhood of the barrier), we simply apply a central symmetry in Table 1 (see Table 2), and the same conclusion follows.

Market regime	**Rally**	**Sell-Off**
Slow	$\delta e^{bs} \leq \delta e^{lo} \leq 0$	$\delta e^{lo} \leq -(\delta e^{bs})^-$
Fast	$(\delta e^{bs})^+ \leq \delta e^{lo}$	$0 \leq \delta e^{lo} \leq \delta e^{bs}$

Table 2: *Negative Gamma/Vega exposure in a negatively skewed local volatility model.*

Finally in a *positively* skewed local volatility model, we simply reverse the order of the columns in Table 1 (or 2), and the dominant market regimes are exchanged as well, hence the conclusions as in the case of positively skewed markets.

13.2 Analysis in a real market

In a real market, we can decompose the P&L increments in the following way:

$$\begin{aligned} \delta e^{lo} &= (-\delta\Pi^{lo} + \Delta^{lo}\delta S) + (\delta\Pi^{lo} - \delta\Pi) \\ \delta e^{bs} &= (-\delta\Pi^{lo} + \Delta^{bs}\delta S) + (\delta\Pi^{lo} - \delta\Pi) \end{aligned} \tag{28}$$

where $\delta\Pi$ denotes the increment of the market price of the option between the dates t_k and t_{k+1} while $\delta\Pi^{lo}$ represents the price increment predicted by the local volatility model calibrated at date t_k, given the new observations at date t_{k+1}. In the right-hand side of (28), the first terms behave as in the analysis of section 13.1, while the second terms are due to the misspecification at date t_{k+1} of the local volatility model calibrated at date t_k. This misspecification arises from the fact that the market-makers have revised their anticipations between date t_k and date t_{k+1}, according to the new market data observed at date t_{k+1} (and also, from time to time, according to more punctual economico-political macro news or events).
It seems reasonable to expect that:
(i) At fast market regimes with unexpectedly high levels of realized volatility, the market-makers will have a tendency to push the options' implied volatilities upwards compared to those predicted by the local volatility model calibrated at date t_k, whereas
(ii) At slow market regimes, the market-makers will have a tendency to push the options' implied volatilities downwards compared to those predicted by the local volatility model calibrated at date t_k.

In the case of a vanilla option, which is vega positive, this implies that:
(i) $\delta\Pi^{lo} \leq \delta\Pi$ at fast market regimes, and
(ii) $\delta\Pi \leq \delta\Pi^{lo}$ at slow market regimes.
By comparison with the situation in a negatively skewed local volatility model, δe^{bs} and δe^{lo} are pushed away from 0 by the same amount in Table 1, as an effect of the model misspecification terms in (28) (second terms in the right-hand side). So the situation depicted in Table 1 still holds true in the real market. The local delta thus remains better on average than the implied delta (but the performance of both deltas deteriorates, so the differential of performance between the two deltas is typically less in the real market than in a local volatility model).

Symmetrical analyses lead to the same conclusion in the case of an exotic option with negative Gamma/Vega exposure, and/or in the case of a persistently positively skewed market. Proposition 13.1 is thus established.

We refer the reader to [12] for reports on **numerical experiments** providing empirical and quantitative support to the previous conclusions, using both simulated and real time-series of equity-index data (note that real equity-index data have had a large negative implied skew since the stock market crash of October 1987).
Moreover we check numerically that the conclusions we draw are still true when *transaction costs* are taken into account.

Remark 13.2 By comparison (see in particular Proposition 13.2 and its analog which is still valid in a real, non local volatility market), the analysis of Derman [55] implied that, in negatively skewed markets, the implied delta should not be worse or could even be better *on average conditionally on the fact that the market is in a slow regime*, while the local delta should be better *on average conditionally on the fact that the market is in a fast regime*. In Derman's analysis one needs to know what is or will be the actual market regime, fast or slow, for making one's choice of a delta. The question of knowing which delta is better on average is left unanswered.

Part III

Defaultable Game Options

Introduction

In this part we provide a synthetic presentation of the main results of the papers [2, 3, 4, 5]. This series of papers, in collaboration with Monique Jeanblanc, Tomasz Bielecki and Marek Rutkowski, is motivated by applications to convertible bonds. Convertible bonds have two important and distinguishing features:

• early put (as for American options) and call clauses at the holder's and issuer's convenience, respectively;

• defaultability, since they are corporate bonds, and one of the main vehicles of the so called *equity to credit* and *credit to equity* strategies.

This led us to cast convertible bonds into a more general framework of *defaultable game options* (covering American and European options, defaultable or not, as special cases), adapting to credit risk the general definition of game options in Kifer [91].

Paper [2] (section 14 below) deals with the issue of *pricing*. Though we do not dwell upon this in this report, this simple study of arbitrage prices already has interesting practical applications. In particular (see Remark 14.4 and [2]) one can use it to establish a rigorous *decomposition* of a defaultable game option into a reference straight bond and an embedded game exchange option. This allows one to give a precise definition to commonly used terms of the *implied spread* and the *implied volatility* of a convertible bond. This decomposition also provides a *static replication strategy* of a defaultable game option in terms of the embedded straight bond and game exchange option.

The issue of *dynamic hedging* is dealt with in [3], in a rather general reduced-form model of credit risk (sections 15 to 17 below). Hence we postulate that the primary market filtration $\mathbb{G}$ admits the representation $\mathbb{G} = \mathbb{H} \vee \mathbb{F}$, where the filtration $\mathbb{H}$ is generated by the default indicator process $H_t = \mathbb{1}_{\{\theta \le t\}}$ and where $\mathbb{F}$ is some reference filtration.

The main result can be informally stated as follows: Under the assumption (thoroughly investigated in [15, 8], see part IV of this report) that a related doubly reflected *Backward Stochastic Differential Equation (BSDE)*, relative to the filtration $\mathbb{F}$ under some risk-neutral measure $\mathbb{Q}$, admits a solution (Y, F, K), then $\widetilde{\Pi} = Y$ is the minimal (pre-default) *superhedging price up to a* $(\mathbb{G}, \mathbb{Q})$ *– local martingale cost process*, the latter being equal to 0 in the case of complete markets. This notion of hedge *with local martingale cost* thus establishes a connection between arbitrage prices and hedging, in a rather general, possibly incomplete, market.

In [4] (section 18 below), we consider the specification of these results to the *Markovian set-up*. A complementary *variational inequality* approach may then be developed, and more *explicit and constructive hedging strategies* may be given.

Paper [5] (section 19 below) is an application and illustration of all the previous results in the case of convertible bonds in a primary market consisting of a savings account, a stock underlying the convertible bond, and an associated CDS contract.

14 Abstract Set-Up

14.1 Primary Market

Given a finite horizon date $T > 0$, we assume that the primary market is composed of the saving account and d risky assets with price processes defined on a filtered probability space $(\Omega, \mathbb{G}, \mathbb{P})$ (with $\mathcal{G}_0$ trivial, for simplicity), where $\mathbb{P}$ denotes the statistical probability measure. By default all random variables are real-valued and $\mathcal{G}_T$-measurable, all processes are real-valued and $\mathbb{G}$-adapted and all semimartingales are càdlàg, without loss of generality.

We postulate that:
• the *discount factor* process β, that is, the inverse of the savings account, is given as $\beta_t = \exp(-\int_0^t r_u du)$, for a bounded below short-term (presumably stochastic) interest rate process r;
• the primary risky assets, with $\mathbb{R}^d$-valued price process X, pay dividends, whose cumulative value process, denoted by $\mathcal{D}$, is modeled as an $\mathbb{R}^d$-valued process of finite variation. We define the *cumulative price* $\widehat{X}$ of the asset by

$$\beta_t \widehat{X}_t = \beta_t X_t + \int_{[0,t]} \beta_u \, d\mathcal{D}_u \, , \tag{29}$$

assumed to be a locally bounded (for simplicity of presentation in this report) semimartingale.

We assume that the primary market model is free of arbitrage opportunities (though presumably incomplete), in the sense that there exists a *risk-neutral measure* $\mathbb{Q}$, namely a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ for which $\beta\widehat{X}$ is a local martingale.

14.2 Defaultable Derivatives

Given a $[0, +\infty]$-valued stopping time θ representing the *default time* of a reference entity, we set

$$I_t = \mathbb{1}_{\{\theta \le t\}} \, , \; J_t = 1 - I_t \, .$$

Let Θ_t (or Θ, in case $t = 0$) denote the set of $[t, T]$-valued $\mathbb{G}$-stopping times, and let ν stand for $\sigma \wedge \tau \wedge \theta$, for any $\sigma, \tau \in \Theta_t$.

The next definition specifies to dividend paying defaultable derivatives the general notion of game option introduced by Kifer [91].

Definition 14.1 (Bielecki et al. [2]) A *defaultable game option* is a game option *(see Kifer [91])* with the *ex-dividend cumulative discounted cash flows* $\beta_t \pi(t; \sigma, \tau)$, where the $\mathcal{G}_\nu$-measurable random variable $\pi(t; \sigma, \tau)$ is given by the formula, for any *pricing time* $t \in [0, T]$, (holder) *call time* $\sigma \in \Theta_t$ and (issuer) *put time* $\tau \in \Theta_t$,

$$\beta_t \pi(t; \sigma, \tau) = \int_t^\nu \beta_u \, dD_u + \beta_\nu J_\nu \Big(\mathbb{1}_{\{\nu = \tau < T\}} L_\tau + \mathbb{1}_{\{\nu < \tau\}} U_\sigma + \mathbb{1}_{\{\nu = T\}} \xi \Big), \tag{30}$$

where:

• the *dividend process* $D = (D_t)_{t\in[0,T]}$ equals

$$D_t = \int_{[0,t]} J_u C_u du + R_u \, dI_u \ ,$$

for some *coupon rate process* $C = (C_t)_{t\in[0,T]}$, and some predictable locally bounded *recovery process* $R = (R_t)_{t\in[0,T]}$;
• the *put payment* $L = (L_t)_{t\in[0,T]}$ and the *call payment* $U = (U_t)_{t\in[0,T]}$ are càdlàg processes, and the *payment at maturity* ξ is a random variable such that

$$L \le U \text{ on } [0,T] \ , \ L_T \le \xi \le U_T \ .$$

Remark 14.2 **(i)** In [2, 3, 4, 5] one also copes with the case of *discrete coupons.*
(ii) Introducing *constrained sets of stopping policies* like '$\sigma \ge \bar{\sigma}$ for some $\bar{\sigma} \in \Theta$,' it is possible to consider defaultable American and European derivatives as special cases of defaultable Game Claims. More generally, the so called *call protection* $\bar{\sigma}$ is at the origin of a number of problems, theoretical as well as practical, not discussed in this report, which are dealt with at length in [2, 3, 4, 5, 16].

We further assume that R, L and ξ are bounded from below, so that the cumulative discounted payoff is bounded from below. Specifically, there exists a constant c such that

$$\int_{[0,t]} \beta_u \, dD_u + \beta_t J_t \Big(\mathbb{1}_{\{t<T\}} L_t + \mathbb{1}_{\{t=T\}} \xi \Big) \ge -c, \quad t \in [0,T] \ . \tag{31}$$

14.2.1 Convertible Bonds

The standing example of a defaultable game option is a (defaultable) *convertible bond.* To describe the covenants of a (stylized) convertible bond, we need to introduce some additional notation:

$\bar{N}$: the par (nominal) value,

S: the price process of the asset underlying the bond,

$\bar{R}$: the recovery process on the bond upon default of the issuer,

κ : the bond's conversion factor,

$\bar{P} \le \bar{C}$: the put and call nominal payments, respectively; by assumption $\bar{P} \le \bar{N} \le \bar{C}$.

Definition 14.3 A convertible bond is a (defaultable) game option with coupon process C, recovery process R^{cb} and payoffs L^{cb}, U^{cb}, ξ^{cb} such that

$$R_t^{cb} = (1-\eta)\kappa S_{t-} \vee \bar{R}_t \ , \ \xi^{cb} = \bar{N} \vee \kappa S_T \tag{32}$$

$$L_t^{cb} = \bar{P} \vee \kappa S_t \ , \ U_t^{cb} = \bar{C} \vee \kappa S_t \ . \tag{33}$$

See [2] for a more detailed description of covenants of convertible bonds, with further important real-life features like discrete coupons or call protection (cf. Remark 14.2).

14.3 Pricing and Hedging in the General Set-Up

14.3.1 Pricing

The notion of arbitrage price process of a game option referred to in the next result, is obtained by a suitable extension to game options of the No Free Lunch with Vanishing Risk condition of Delbaen and Schachermayer [51] (see Kallsen and Kühn [89], Bielecki et al. [2]). Now, it is so that this NFLVR condition on a 'candidate arbitrage price process' Π for a game option is essentially equivalent to Π being the value process of a related Dynkin Game [59] under a risk-neutral probability measure on the primary market. More precisely,

Proposition 14.1 (Bielecki et al. [2]) *Assume that a semimartingale* Π *is the value of the* Dynkin game *related to a game option under some risk-neutral measure* $\mathbb{Q}$ *on the primary market, that is, for* $t \in [0, T]$:

$$\begin{aligned} &\text{esssup}_{\tau\in\Theta_t}\text{essinf}_{\sigma\in\Theta_t}\mathbb{E}_{\mathbb{Q}}\big(\pi(t;\sigma,\tau)\,\big|\,\mathcal{G}_t\big) = \Pi_t \\ &\quad = \text{essinf}_{\sigma\in\Theta_t}\text{esssup}_{\tau\in\Theta_t}\mathbb{E}_{\mathbb{Q}}\big(\pi(t;\sigma,\tau)\,\big|\,\mathcal{G}_t\big)\ . \end{aligned} \tag{34}$$

Then Π *is an arbitrage price process for the game option. Moreover, a converse to this result holds under a suitable integrability assumption.*

Remark 14.4 Using this, one can establish a rigorous decomposition of an arbitrage price of a defaultable game option as the sum of the price of an embedded straight bond and of the price of an embedded game exchange option. This allows one to give a precise definition to commonly used terms of the *implied spread* and the *implied volatility* of a convertible bond (see [2, 5]).

14.3.2 Hedging

We adopt the definition of hedging game options stemming from successive developments, starting from the hedging of American options examined by Karatzas [90], and subsequently followed by El Karoui and Quenez [63], Kifer [91], Ma and Cvitanić [102] and Hamadène [79] (see also Schweizer [112]). This definition will be later shown to be consistent with the concept of arbitrage pricing of section 14.3.2 for a defaultable game option.

First, by a (self-financing) *primary trading strategy* (starting at time 0), we mean as usual a pair (w, ζ) such that:
• the constant w represents the *initial wealth*,
• ζ is a predictable locally bounded $\mathbb{R}^{1\otimes d}$-valued process (ζ *is an admissible strategy*, for short) representing holdings in primary risky assets.
The *wealth process* $\mathcal{W}^{w,\zeta} = \mathcal{W}$ of a primary trading strategy (w, ζ) is given by the formula, for $t \in [0, T]$,

$$\beta_t \mathcal{W}_t = w + \int_0^t \zeta_u \, d(\beta_u \widehat{X}_u). \tag{35}$$

We now introduce a (very large, to be specified later) class of hedges *with semimartingale cost process* ρ.

Definition 14.5 An *hedge with (semimartingale) cost process* ρ (*issuer hedge starting at time 0*) for the game option with ex-dividend cumulative cash flows π (cf. (30)) is represented by a triplet (w, ζ, σ) such that:
• (w, ζ) is a primary strategy with the wealth process $\mathcal{W}$ given by (35),
• the call time σ belongs to Θ, and the following inequality is valid, for every put time $\tau \in \Theta$,

$$\beta_\nu \mathcal{W}_\nu + \int_0^\nu \beta_u \, d\rho_u \geq \pi(0; \sigma, \tau), \quad \text{a.s.} \tag{36}$$

Remark 14.6 **(i)** The process ρ is to be interpreted as the (running) *financing cost*, that is, the amount of cash added to (if $d\rho_t \geq 0$) or withdrawn from (if $d\rho_t \leq 0$) the hedging portfolio in order to get a perfect, but no longer self-financing, hedge.
(ii) Hedges *at no cost* (that is, with $\rho = 0$) are thus in effect *superhedges.*
(iii) Analogous definitions and results hold for holder hedges.

This class of hedges with cost ρ is obviously too large for any practical purpose, so we will restrict our attention to hedges with a *local martingale* cost ρ under a particular risk-neutral measure $\mathbb{Q}$ (cf. the related notions of *risk-minimizing strategy* in Föllmer and Sondermann [69] and *mean self-financing hedge* in Schweizer [112]). ***In the sequel, we work under a fixed, but arbitrary, risk-neutral measure*** $\mathbb{Q}$***.*** In particular, we define $\mathbb{W}$ as the set of initial values w for which there exists an issuer hedge of the game option with the initial value w and with *local martingale* cost under $\mathbb{Q}$.

The following result gives some preliminary conclusions regarding the initial cost of a hedging strategy for the game option under the present, rather weak, assumptions. In Proposition 16.4, we shall see that, under stronger assumptions, the infimum is attained and thus we obtain equalities, rather than merely inequalities, in (37).

Lemma 14.2 *We have (with, by convention,* $\inf \emptyset = \infty$*)*

$$\inf_{\sigma \in \Theta} \sup_{\tau \in \Theta} \mathbb{E}_{\mathbb{Q}} \pi(0; \sigma, \tau) \leq \inf \mathbb{W} \, . \tag{37}$$

Proof. Assume that (w, ζ, σ^*) is an issuer hedge with local martingale cost ρ for the game option. Then (35) and (36) imply that, for any $t \in [0, T]$,

$$\begin{aligned} w = \beta_{t\wedge\sigma^*\wedge\theta} \mathcal{W}_{t\wedge\sigma^*\wedge\theta} - \int_0^{t\wedge\sigma^*\wedge\theta} \zeta_u \, d(\beta_u \widehat{X}_u) \geq \\ \pi(0; t, \sigma^*) - \int_0^{t\wedge\sigma^*\wedge\theta} \big(\zeta_u \, d(\beta_u \widehat{X}_u) + \beta_u \, d\rho_u\big) \, . \end{aligned} \tag{38}$$

The stochastic integral $\int_0^t \big(\zeta_u \, d(\beta_u \widehat{X}_u) + \beta_u \, d\rho_u\big)$ is a local martingale, as is then the stopped process $\int_0^{t\wedge\sigma^*\wedge\theta} \big(\zeta_u \, d(\beta_u \widehat{X}_u) + \beta_u \, d\rho_u\big)$. Moreover the latter process is bounded from below, by (38) and (31), so that it is a bounded from below local martingale and thus a supermartingale. Now, for any stopping time $\tau \in \Theta$, the inequality in formula (38) still holds with t replaced by τ. By taking expectations, we obtain (recall that σ^* is fixed)

$$w \geq \mathbb{E}_{\mathbb{Q}} \beta_0 \pi(0; \sigma^*, \tau) \, ,$$

for any τ in Θ, and thus,

$$w \geq \inf_{\sigma\in\Theta} \sup_{\tau\in\Theta} \mathbb{E}_{\mathbb{Q}}\pi(0;\sigma,\tau) .$$

The last inequality yields (37). □

15 Hazard Intensity Set-Up

We assume further that $\mathbb{G} = \mathbb{H} \vee \mathbb{F}$, where the filtration $\mathbb{H}$ is generated by the *default indicator process* $I_t = \mathbb{1}_{\{\theta \leq t\}}$ and $\mathbb{F}$ is some *reference filtration*. Moreover, we assume that the optional projection of J, defined by, for $t \in [0,T]$,

$$^{o}J_t = \mathbb{Q}(\theta > t \,|\, \mathcal{F}_t) =: Q_t$$

(*Azema's supermartingale*), is (strictly) positive, continuous and non-increasing.

Remark 15.1 **(i)** If Q is continuous, θ is a *totally inaccessible* $\mathbb{G}$-stopping time (see, e.g., [52]). Moreover, θ *avoids* $\mathbb{F}$*-stopping times*, in the sense that $\mathbb{Q}(\theta = \tau) = 0$, for any $\mathbb{F}$-stopping time τ (see Coculescu et al. [42]).
(ii) Assuming Q continuous, the further assumption that Q has a finite variation in fact implies that Q is non-increasing. This further assumption lies somewhere between assuming further the (stronger) $(\mathcal{H})$ (or *immersion*) Hypothesis and assuming further that θ is an $\mathbb{F}$-*pseudo-stopping time*. Recall that the $(\mathcal{H})$ Hypothesis means that all $\mathbb{F}$-local martingales are $\mathbb{G}$-local martingales; θ being an $\mathbb{F}$-*pseudo-stopping time* means that all $\mathbb{F}$-local martingales *stopped at* θ are $\mathbb{G}$-local martingales (see Nikeghbali and Yor [104]).

We assume for simplicity of presentation in this report that Q is time-differentiable, and we define the *default (hazard) intensity* γ, the *credit-risk adjusted interest rate* μ and the *credit-risk adjusted discount factor* α by, respectively

$$\gamma_t = -\frac{d\ln Q_t}{dt} \ , \ \mu_t = r_t + \gamma_t \ , \ \alpha_t = \beta_t \exp(-\int_0^t \gamma_u du) = \exp(-\int_0^t \mu_u du)$$

(note that the process α is time-differentiable and bounded, like β). Under our assumptions, the *compensated jump-to-default process* $H_t = I_t - \int_0^t J_u \gamma_u du$, $t \in [0,T]$, is known to be a $\mathbb{G}$-martingale.
The quantities $\widetilde{\tau}$ and $\widetilde{\Theta}$ introduced in the next lemma are called the *pre-default values* of τ and Θ, respectively.

Lemma 15.1 (see Bielecki et al. [3]) **(i)** *For any $\mathbb{G}$-adapted, resp. $\mathbb{G}$-predictable process Π over $[0,T]$ there exists an (unique) $\mathbb{F}$-adapted, resp. $\mathbb{F}$-predictable process $\widetilde{\Pi}$ over $[0,T]$ such that $J\Pi = J\widetilde{\Pi}$, resp. $J_{\cdot-}\Pi = J_{\cdot-}\widetilde{\Pi}$ over $[0,T]$.*
(ii) *For any $\tau \in \Theta$, there exists a $[0,T]$-valued $\mathbb{F}$-stopping time such that $\tau \wedge \theta = \widetilde{\tau} \wedge \theta$.*

In view of the structure of the payoffs π in (30), we thus may assume without loss of generality that the data C, R, L, U, ξ and the stopping policies σ, τ are defined relative to the filtration $\mathbb{F}$, rather than $\mathbb{G}$ above. More precisely, ***we assume in the sequel that C, L, U are $\mathbb{F}-$adapted, $\xi \in \mathcal{F}_T$, R is $\mathbb{F}$-predictable and σ, τ are $\mathbb{F}$-stopping times. For any $t \in [0,T]$, Θ_t (or Θ, in case $t = 0$) henceforth denotes the set of $[t,T]$-valued $\mathbb{F}$- (rather than $\mathbb{G}$- before) stopping times; ν denotes $\sigma \wedge \tau$ (rather than $\sigma \wedge \tau \wedge \theta$ before), for any $t \in [0,T]$ and $\sigma, \tau \in \Theta_t$.***

15.1 Reduction of Filtration

Under our assumptions, the next lemma (which is rather standard, if not for the presence of the stopping policies σ, τ therein) shows that the computation of conditional expectations of cash flows $\pi(t; \sigma, \tau)$ with respect to $\mathcal{G}_t$ can be reduced to the computation of conditional expectations of $\mathbb{F}$-equivalent cash flows $\widetilde{\pi}(t; \sigma, \tau)$ with respect to $\mathcal{F}_t$.

Lemma 15.2 *For any stopping times $\sigma, \tau \in \Theta_t$ we have that*

$$\mathbb{E}_{\mathbb{Q}}\big(\pi(t;\sigma,\tau)\,\big|\,\mathcal{G}_t\big) = J_t\,\mathbb{E}_{\mathbb{Q}}\big(\widetilde{\pi}(t;\sigma,\tau)\,\big|\,\mathcal{F}_t\big),$$

where $\widetilde{\pi}(t; \sigma, \tau)$ is given by, with $\nu = \tau \wedge \sigma$,

$$\alpha_t \widetilde{\pi}(t;\sigma,\tau) = \int_t^{\nu} \alpha_u f_u du + \alpha_\nu \big(\mathbb{1}_{\{\nu=\tau<T\}} L_\tau + \mathbb{1}_{\{\nu<\tau\}} U_\sigma + \mathbb{1}_{\{\nu=T\}}\xi\big) \tag{39}$$

where $f = C + \gamma R$.

As a corollary to the previous results, we have,

Proposition 15.3 (Bielecki et al. [3]) *If an $\mathbb{F}$-semimartingale $\widetilde{\Pi}$ solves the $\mathbb{F}$-Dynkin Game with payoff $\widetilde{\pi}$, in the sense that, for any $t \in [0, T]$,*

$$\begin{aligned} \operatorname{esssup}_{\tau\in\Theta_t} \operatorname{essinf}_{\sigma\in\Theta_t} \mathbb{E}_{\mathbb{Q}}\big(\widetilde{\pi}(t;\sigma,\tau)\,\big|\,\mathcal{F}_t\big) &= \widetilde{\Pi}_t \\ &= \operatorname{essinf}_{\sigma\in\Theta_t} \operatorname{esssup}_{\tau\in\Theta_t} \mathbb{E}_{\mathbb{Q}}\big(\widetilde{\pi}(t;\sigma,\tau)\big|\,\mathcal{F}_t\big)\ , \end{aligned}$$

then $\Pi := J\widetilde{\Pi}$ is a $\mathbb{G}$-semimartingale solving the $\mathbb{G}$-Dynkin Game with payoff π.

Hence by Proposition 14.1 Π is an arbitrage price for the option, with related pre-default price process $\widetilde{\Pi}$. A converse to this result may be established under a suitable integrability assumption.

We thus effectively moved our considerations from the original market subject to the default risk, in which cash flows are discounted according to the discount factor β, to the fictitious default-free market, in which cash flows are discounted according to the credit risk adjusted discount factor α.

16 Pre-default Model

16.1 Doubly Reflected BSDE

The next step consists in modeling $\widetilde{\Pi}$ as the state-process Y of a solution (Y, F, K), assumed to exist, to the following *doubly reflected BSDE* with data α, C, R, L, U, ξ (see Cvitanić and Karatzas [48], Hamadène and Hassani [80], Crépey et al. [15, 8]):

$$\begin{aligned} \alpha_t Y_t = \alpha_T \xi + \textstyle\int_t^T \alpha_u \left(f_u du + dK_u - dF_u\right)\ ,\ t \in [0,T], \\ L_t \le Y_t \le U_t, \quad t \in [0,T], \\ \textstyle\int_0^T (Y_u - L_u)\, dK_u^+ = \int_0^T (U_u - Y_u)\, dK_u^- = 0\ . \end{aligned} \tag{40}$$

Definition 16.1 By a *solution* to (40), we mean a triplet (Y, F, K) such that:
• the *state process* Y is an $\mathbb{F}$-adapted, càdlàg process,
• $\int_0^{\cdot} \alpha \, dF$ is an $\mathbb{F}$-martingale vanishing at time 0,
• K is an $\mathbb{F}$-adapted continuous finite variation process vanishing at time 0,
• all conditions in (40) are satisfied, where in the third line K^+ and K^- denote the Jordan components of K.
Here by *Jordan decomposition* we mean the decomposition $K = K^+ - K^-$, where the non-decreasing continuous processes K^+ and K^- vanish at time 0 and define mutually singular measures.

Note that the first line of (40) may be rewritten as, for $t \in [0, T]$ (recall $\mu = r + \gamma$):

$$Y_t = \xi + \int_t^T (f_u - \mu_u Y_u) du + dK_u - dF_u \; . \tag{41}$$

As investigated in Hamadène and Hassani [80] or Crépey and Matoussi [15] and Crépey [8], existence and uniqueness of a solution to (40) (under suitable L_2-integrability conditions on the data and the solution) is essentially equivalent to the so-called *Mokobodski condition*, namely, the existence of a *quasimartingale* Θ (see section 22) such that $L \le \Theta \le U$ on $[0, T]$. It is thus satisfied when one of the barriers is a quasimartingale and, in particular, when one of the barriers is given as $S \vee c$ where S is a square-integrable Itô process and c is a constant in $\mathbb{R} \cup \{-\infty\}$ (see Theorem 26.3 below). This covers, for instance, the lower payoff L of a *convertible bond* (see Definition 14.3 and Bielecki et al. [2, 5]).

To support our modeling standing assumption that $\widetilde{\Pi} = Y$, we have the following (standard) *verification principle.*

Proposition 16.1 *The $\mathbb{F}$-semimartingale Y solves the $\mathbb{F}$-Dynkin Game with payoff $\widetilde{\pi}$. More precisely, for any $t \in [0, T]$, the pair of stopping times $\sigma^*, \tau^* \in \Theta_t$ given by*

$$\sigma^* = \inf \Big\{ u \in [t, T] \, ; \, Y_u \ge U_u \Big\} \wedge T, \quad \tau^* = \inf \Big\{ u \in [t, T] \, ; \, Y_u \le L_u \Big\} \wedge T, \tag{42}$$

is a saddle-point of this game, in the sense that we have, for any $\sigma, \tau \in \Theta_t$:

$$\mathbb{E}_{\mathbb{Q}}\big(\widetilde{\pi}(t; \sigma^*, \tau) \,\big|\, \mathcal{F}_t\big) \le Y_t \le \mathbb{E}_{\mathbb{Q}}\big(\widetilde{\pi}(t; , \sigma, \tau^*) \,\big|\, \mathcal{F}_t\big).$$

Hence, by Proposition 15.3, $\Pi := J\widetilde{\Pi}$ (where we set $\widetilde{\Pi} = Y$) is an arbitrage price for the option, with related pre-default price process $\widetilde{\Pi}$.

16.2 Connection with Hedging

Let us set further, for $t \in [0, T]$,

$$\Pi_t = \mathbb{1}_{\{t<\theta\}} \widetilde{\Pi}_t \; , \; \beta_t \widehat{\Pi}_t = \beta_t \Pi_t + \int_{[0,t]} \beta_u \, dD_u \tag{43}$$

where we recall that $D_t = \int_{[0,t]} J_u C_u du + R_u \, dI_u$. We define G by $G_0 = 0$ and, for $t \in [0, T]$,

$$\int_{[0,t]} \beta_u dG_u = \beta_t \widehat{\Pi}_t + \int_0^t \beta_u J_u \, dK_u. \tag{44}$$

The following lemma is key in what follows. It allows one in particular to interpret (44) as the canonical decomposition of the $\mathbb{G}$-semimartingale $\beta\widehat{\Pi}$. In particular G is the canonical $\mathbb{G}$-local martingale component of $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t \widehat{\Pi}_t)$.

Lemma 16.2 *The process G defined by (44) is a $\mathbb{G}$-local martingale (stopped at θ).*

Proof. We have by (40), for every $t \in [0,T]$,

$$\int_0^t \alpha_u \, dF_u = \alpha_t \widetilde{\Pi}_t - \widetilde{\Pi}_0 + \int_0^t \alpha_u \, dK_u + \int_0^t \alpha_u \, (C_u + \gamma_u R_u) du$$

So by standard computations (cf. Lemma 15.2), for any $0 \le t \le u \le T$,

$$\mathbb{E}_{\mathbb{Q}}\Big(\beta_t^{-1} \int_t^u \beta_v \, dG_v \,\Big|\, \mathcal{G}_t\Big) = J_t \, \mathbb{E}_{\mathbb{Q}}\Big(\alpha_t^{-1} \int_t^u \alpha_v \, dF_v \,\Big|\, \mathcal{F}_t\Big) = 0 \, .$$

□

Some of the arguments underlying the following result are classical, and already present for instance in Lepeltier and Maingueneau [99]. Proposition 16.3 can thus be seen as an extension of their results to the defaultable case, in which two filtrations are involved. Note that our assumptions are made relative to the filtration $\mathbb{F}$ (the one with respect to which the BSDE (40) is defined), whereas conclusions are drawn relative to the filtration $\mathbb{G}$.

Proposition 16.3 (Bielecki et al. [3]) *For any admissible strategy ζ, the triplet (Π_0, ζ, σ^*) (where Π_0 is defined by (43) and σ^* by (42) with $t = 0$ therein) is an hedge with* $\mathbb{G}$-local martingale *cost process $\rho(\zeta) = \rho$ given by $\rho_0 = 0$ and*

$$d\rho_t = dG_t - \zeta_t \, \beta_t^{-1} d(\beta_t \widehat{X}_t). \tag{45}$$

Proof. By Lemma 16.2, the process ρ is a $\mathbb{G}$-local martingale. Let $\mathcal{W}$ denote the wealth process of the primary strategy (Π_0, ζ). So $\mathcal{W}_0 = \Pi_0$ and for $t \in [0,T]$:

$$d(\beta_t \mathcal{W}_t) = \zeta_t \, d(\beta_t \widehat{X}_t) = \beta_t (dG_t - \, d\rho_t) \, .$$

Thus

$$\beta_t \mathcal{W}_t + \int_0^t \beta_u \, d\rho_u = \int_0^t \beta_u \, dG_u + \Pi_0 = \beta_t \widehat{\Pi}_t + (\Pi_0 - \widehat{\Pi}_0) + \int_0^t \beta_u J_u \, dK_u \, , \tag{46}$$

by (44). For any $\tau \in \Theta$, set $\nu = \tau \wedge \sigma^*$. From the definition of σ^*, the minimality conditions in (40) and the continuity of K^-, it follows that $K^- = 0$ and thus $K \ge 0$ on $[0, \sigma^*]$. Since $\tau \le \sigma^*$, (46) thus yields

$$\beta_\nu \mathcal{W}_\nu + \int_0^\nu \beta_u \, d\rho_u \ge \int_0^\nu \beta_u \, dD_u + \beta_\nu J_\nu Y_\nu \ge \pi(0; \sigma^*, \tau)$$

where the last inequality holds because $Y_\nu = Y_{\sigma^*} \ge U_{\sigma^*}$ on the event $\sigma^* < \tau$. □

Remark 16.2 **(i)** The situation where ρ can be made equal to zero by the choice of a suitable strategy ζ in Proposition 16.3 corresponds to a particular form of hedgeability of a game option in which an issuer (or an holder) is able to hedge all risks embedded in a defaultable game option. The case where $\rho \neq 0$ corresponds either to non-hedgeability of a game option or to the situation in which an issuer (or a holder) is able to hedge, but she *prefers* not to hedge all the risks embedded in the option, for instance, she may be *willing* to take some bets regarding specific risk directions. That is why we do not postulate a priori that ρ should be minimized in some sense as, for instance, in Schweizer [112].
(ii) It is possible to introduce the issuer *trivial hedge* $(\Pi_0, 0, \sigma^*)$ with the $\mathbb{G}$-local martingale cost

$$\rho_t^0 = G_t, \quad t \in [0, T].$$

Obviously, this hedge is of a minor practical interest, since it implicitly assumes one is not interested in hedging. The trivial hedge or, more precisely, the existence of any hedge is used in the proof of Proposition 16.4, however.

Let us now draw some conclusions from Lemma 14.2 and Proposition 16.3. In the context of specific (Cox–Ross–Rubinstein or Black–Scholes) models, analogous results can be found in Kifer [91]. Our contribution here is an extension of these results to the present set-up involving a reduction of filtration, as well as to more general models.

Proposition 16.4 (Bielecki et al. [3]) *Under the assumptions of Proposition 16.3, we have that* $\Pi_0 = \min \mathbb{W}$*, so* Π_0 *is the minimum of initial wealths of an issuer hedge with a* $\mathbb{G}$*-local martingale cost.*

Given our definition of hedging with a cost and the definition of Π_0, the fact that there exists a hedge with initial wealth Π_0 and $\mathbb{G}$-local martingale cost is by no means surprising. The minimality statement establishes a connection between arbitrage prices and hedging in a general, incomplete market.

It is also easy to see that one could state analogous definitions and results regarding hedging a defaultable game option, starting at any date $t \in [0, T]$. Note that in the special case of *European options*, the results can be further specified in terms of payoff's *replication at* T (up to a $\mathbb{G}$-local martingale cost), rather than (super-)hedging (up to a $\mathbb{G}$-local martingale cost) in the above sense.

17 Analysis of Hedging Strategies

17.1 Discounted Cumulative Value Dynamics

Let $H_t = I_t - \int_0^t J_u \gamma_u du$ stand for the compensated jump-to-default $\mathbb{G}$-martingale. Our analysis of hedging strategies will rely on the following lemma, which yields the dynamics of the price process $\widehat{\Pi}$ of a game option or, more precisely, of the martingale component G of $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t \widehat{\Pi}_t)$.

Lemma 17.1 *The* $\mathbb{G}$*-local martingale* G *defined by (44) satisfies, for* $t \in [0, T \wedge \theta]$:

$$dG_t = dF_t + \Delta\Pi_t \, dH_t \tag{47}$$

with $\Delta\Pi_t = R_t - \widetilde{\Pi}_{t-}$.

Sketch of Proof. This follows by computations similar to those of the proof of Kusuoka's Theorem 2.3 in [93] (where the $(\mathcal{H})$ hypothesis and a more specific Brownian filtration $\mathbb{F} = \mathbb{F}^W$ are assumed therein), using the avoidance property that $\mathbb{Q}(\theta = \tau) = 0$ for any $\mathbb{F}$-stopping time τ. □

In analogy with the previous developments regarding the option to be hedged, we assume henceforth that the dividend (vector) process $\mathcal{D}$ of the *primary market* price process X is given as

$$\mathcal{D}_t = \int_{[0,t]} J_u\, \mathcal{C}_u du + \mathcal{R}_u\, dH_u$$

for suitable coupon rate and recovery processes $\mathcal{C}$ and $\mathcal{R}$. We assume that $X = J\widetilde{X}$, without loss of generality with respect to the hedging application at hand (in particular any value of the primary market at θ is embedded in the recovery part of the dividend process $\mathcal{D}$ for X). We further define, along with the cumulative price $\widehat{X}$ as usual, the *pre-default cumulative price*, given by, for $t \in [0, T]$:

$$\bar{X}_t = \widetilde{X}_t + \alpha_t^{-1} \int_0^t \alpha_u\, g_u du\ .$$

where we set $g = \mathcal{C} + \gamma\mathcal{R}$.

The following result is the analog, relative to the primary market, of identity (47) for a game option.

Lemma 17.2 *One has, for* $t \in [0, T \wedge \theta]$:

$$\beta_t^{-1} d(\beta_t \widehat{X}_t) = \alpha_t^{-1} d(\alpha_t \bar{X}_t) + \Delta X_t\, dH_t \tag{48}$$

with $\Delta X_t = \mathcal{R}_t - \widetilde{X}_{t-}$. *Moreover,* $\alpha\bar{X}$ *is an* $\mathbb{F}$*-local martingale.*

Plugging (48) and (47) into (45), we get the following ***fundamental decomposition of the hedging cost*** ρ of the strategy (Π_0, ζ, σ^*). This decomposition will be exploited in various ways in the remaining sections of this part.

Proposition 17.3 *Under the previous assumptions, for any admissible strategy* ζ*, the related cost* $\rho = \rho(\zeta)$ *in Proposition 16.3 satisfies, for every* $t \in [0, T \wedge \theta]$,

$$d\rho_t = dG_t - \zeta_t\, \beta_t^{-1} d(\beta_t \widehat{X}_t) = \Big[dF_t - \zeta_t \alpha_t^{-1} d(\alpha_t \bar{X}_t)\Big] + \Big[\Delta\Pi_t - \zeta_t \Delta X_t\Big] dH_t\ . \tag{49}$$

17.2 Hedging via Orthogonal Decompositions

Let us further be given a reference vector-valued $\mathbb{F}$-square integrable martingale M. In any particular application, the choice of this process will depend on the problem at hand (see [4]).

We assume that F and $\int_0^{\cdot} \alpha^{-1} d(\alpha\bar{X})$ are $\mathbb{F}$-square integrable martingales. They thus admit the following Galtchouk-Kunita-Watanabe decompositions (relative to $\mathbb{F}$), for $t \in [0,T]$:

$$\begin{aligned} dF_t &= V_t\, dM_t + dm_t \\ \alpha_t^{-1} d(\alpha_t \bar{X}_t) &= \bar{V}_t\, dM_t + d\bar{m}_t \end{aligned}$$

for $\mathbb{F}$-square integrable martingales m and $\bar{m}$ orthogonal (in $\mathbb{F}$) to M. In this situation, (49) yields,

Proposition 17.4 *For* $t \in [0, T \wedge \theta]$,

$$d\rho_t = \Big[(V_t, \Delta\Pi_t) - \zeta_t(\bar{V}_t, \Delta X_t)\Big] d \begin{pmatrix} M_t \\ H_t \end{pmatrix} + \Big[dm_t - \zeta_t d\bar{m}_t\Big] . \tag{50}$$

The following result justifies the informal statement that the strategy $\widehat{\zeta}$ (resp. $\widetilde{\zeta}$) therein hedges the risk sources M and H (resp. M). In this result, the symbol $[\cdot,\cdot]$ denotes the square bracket with respect to $\mathbb{G}$. Recall that an $\mathbb{F}$-martingale stopped at θ is a $\mathbb{G}$-local martingale.

Proposition 17.5 (Bielecki et al. [3]) (i) *Assume, in addition, that the system*

$$(V_t, \Delta\Pi_t) = \zeta_t(\bar{V}_t, \Delta X_t)$$

has an (admissible) solution $\widehat{\zeta}$ *on* $[0, T\wedge\theta]$. *Then the cost* $\widehat{\rho} = \rho(\widehat{\zeta})$ *satisfies, for* $t \in [0, T\wedge\theta]$,

$$d\widehat{\rho}_t = dm_t - \widehat{\zeta}_t\, d\widehat{m}_t.$$

Moreover the processes $\widehat{\rho}$ *and* $M_{\cdot\wedge\theta}$ *and* $\widehat{\rho}$ *and* $H_{\cdot\wedge\theta}$ *are* orthogonal in $\mathbb{G}$, *in the sense that* $[\widehat{\rho}, M_{\cdot\wedge\theta}]$ *and* $[\widehat{\rho}, H_{\cdot\wedge\theta}]$ *are* $\mathbb{G}$*-local martingales.*
(ii) *Alternatively to (i), assume that the system*

$$V_t = \zeta_t \bar{V}_t$$

has an (admissible) solution $\widetilde{\zeta}$ *on* $[0, T\wedge\theta]$. *Then the cost* $\widetilde{\rho} = \rho(\widetilde{\zeta})$ *satisfies, for* $t \in [0, T\wedge\theta]$,

$$d\widetilde{\rho}_t = (\Delta\Pi_t - \widetilde{\zeta}_t \Delta X_t)\, dH_t + (dm_t - \widetilde{\zeta}_t\, d\widehat{m}_t).$$

Moreover the processes $\widetilde{\rho}$ *and* $M_{\cdot\wedge\theta}$ *are* orthogonal in $\mathbb{G}$, *in the sense that* $[\widetilde{\rho}, M_{\cdot\wedge\theta}]$ *is a* $\mathbb{G}$*-local martingale.*

In relation with Remark 16.2(i), note that the situation of Proposition 17.5(i) corresponds to the hedgeable case, where the cost $\widehat{\rho}$ vanishes for a strategy $\widehat{\zeta}$. The situation of Proposition 17.5(ii) corresponds to the case of unhedged default risk.

18 Hedging in the Markovian Set-Up

In the Markovian case, an alternative to the previous Galtchouk-Kunita-Watanabe decompositions consists in applying the Itô formula to the pre-default price process $Y_t = u(t, Z_t)$, where Z denotes a relevant Markovian *factor process*, and $u = u(t,z)$ a related *pricing function*.

Let us thus assume that the BSDE (40) is *Markovian*, in the sense that the input data $\mu = r+\gamma, f = C+\gamma R$, L, U and ξ of (40) (with the first line of (40) represented by (41)) are given by Borel-measurable functions of an $(\Omega, \mathbb{F}, \mathbb{Q})$-Markov process Z, so

$$\mu_t = \mu(t,Z_t)\,,\ f_t = f(t,Z_t)\,,\ \xi = \phi(Z_T)\,,\ L_t = \ell(t,Z_t)\,,\ U_t = h(t,Z_t)\,.$$

We assume more specifically that the factor process Z is the solution to the following Markovian (forward) SDE in $\mathbb{R}^k$ (with time-homogeneous coefficients for notational simplicity):

$$dZ_t = b(Z_t)\,dt + \sigma(Z_t)\,dW_t + \delta(Z_{t-})dN_t\ , \tag{51}$$

where:
• W is a k-dimensional Brownian motion, and
• N is the compensated counting measure of a market point process with related intensity $\lambda(Z_t)$ and conditional jump probability measure $p(Z_t, de)$ (so $\delta(Z_{t-})dN_t$ in (51) is a shorthand for $\int_E \delta(Z_{t-},e)N(dt,de)$, where E is the mark space).
Note that Z in (51) is a well-defined Markov process under suitable (Lipschitz and growth) assumptions on the model coefficients (see section 25 for a concrete example), with related generator

$$\mathcal{Z}u(z) = \partial u(z)b(z) + \tfrac{1}{2}\mathrm{Tr}[a(z)\mathcal{H}u(z)] + \lambda(z)\big(\overline{\delta}u(z) - \partial u(z)\widehat{\delta}(z)\big)$$

where $a(z) = \sigma(z)\sigma(z)^{\mathsf{T}}$, ∂u and $\mathcal{H}u$ denote the *row-gradient* and the *Hessian* of u with respect to z, and where we denote

$$\overline{\delta}u(z) = \int_E (u(z+\delta(z,e)) - u(z))p(z,de)\ ,\ \widehat{\delta}(z) = \overline{\delta}\mathrm{Id}_{\mathbb{R}^k}(z) = \int_E \delta(z,e)p(z,de)\ .$$

In the present Markovian set-up the valuation PIDE formally related to a game option writes, with $f = C + \gamma R$ (cf. equation (84) in section 27 below; see also [4]):

$$\begin{aligned}\min\Big(\max\Big((\partial_t + \mathcal{Z})u(t,z) + f(t,z) - \mu(t,z)u(t,z),\\ \ell(t,z) - u(t,z)\Big), h(t,z) - u(t,z)\Big) = 0, \quad t<T,\ z\in\mathbb{R}^k,\end{aligned} \tag{52}$$

with terminal condition $u(T,z) = \phi(z)$. Under mild conditions, the PIDE (52) is well-posed (in a viscosity [8] or weak [27, 28, 21, 14] sense), and its solution $u(t,z)$ is related to the solution (Y,F,K) of (40) as follows, for $t\in[0,T]$:

$$\begin{aligned} Y_t &= u(t,Z_t)\\ dF_t &= \partial u\sigma(t,Z_t)dW_t + \delta u(t,Z_t)dN_t\end{aligned} \tag{53}$$

(and it is also possible to give a PDE interpretation to the obstacle term K in the solution of (40) [20, 14]). Accordingly, (41) takes the following form:

$$-\,du(t,Z_t) = (f_t - \mu_t u(t,Z_t))dt + dK_t - \partial u\sigma(t,Z_t)dW_t - \delta u(t,Z_t)dN_t\ . \tag{54}$$

Let us assume the same structure (without barriers) on the primary market price process X, so $\widetilde{X}_t = v(t,Z_t)$, where, setting $g(t,z) = \mathcal{C}(t,z) + \gamma(t,z)\mathcal{R}(t,z)$,

$$-\,dv(t,Z_t) = (g_t - \mu_t v(t,Z_t))dt - \partial v\sigma(t,Z_t)dW_t - \delta v(t,Z_t)dN_t\ . \tag{55}$$

Exploiting (54) and (55) in (49), we get,

Proposition 18.1 *For $t \in [0, T \wedge \theta]$,*

$$d\rho_t = \Big[(\partial u \sigma(t, Z_t), \delta u(t, Z_t), \Delta u(t, Z_t)) - \zeta_t(\partial v \sigma(t, Z_t), \delta v(t, Z_t), \Delta v(t, Z_t))\Big] d \begin{pmatrix} W_t \\ N_t \\ H_t \end{pmatrix} . \quad (56)$$

This decomposition of the hedging cost ρ can then be used for devising practical hedging schemes of a defaultable game option, like superhedging ($\rho = 0$), hedging only the market (spread) risk W, hedging only the default risk H, or min-variance hedging: see section 19 and Bielecki et al. [3, 4, 5].

19 A Simple Example

In this section we specify the factor process Z of (51) as the following scalar diffusion on $\mathbb{R}_+$ relative to a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ with $\mathbb{F} = \mathbb{F}^W$, for a scalar $(\mathbb{F}, \mathbb{Q})$-Wiener process W:

$$dZ_t = Z_t\Big(\big(r(t) - q(t) + \eta\gamma(Z_t)\big)\, dt + \sigma\, dW_t\Big), \quad Z_0 = z \in \mathbb{R} \quad (57)$$

where:
• $r(t)$ and $q(t)$ represent deterministic riskless interest-rates and dividend yields on the stock of a reference entity (firm),
• the function $\gamma \geq 0$ will be interpreted later as a local default intensity,
• $\eta \leq 1$ is a real constant, to be interpreted later as the *fractional loss upon default* on the stock price of the firm, and
• the *volatility* σ is taken as constant in this report, for notational simplicity.
In particular Z is a Markov process with generator

$$\mathcal{Z} = (r - q + \eta\gamma)z\partial_z + \frac{\sigma^2 z^2}{2}\partial^2_{z^2}. \quad (58)$$

We refer the reader to Remark 19.1 below regarding the specification of the drift of Z in (57)–(58).

We assume that θ is defined in terms of Z via the so-called canonical construction, so (with, by convention, $\inf \emptyset = \infty$),

$$\theta = \inf\Big\{\, t \in [0, \infty];\ \int_0^t \gamma(Z_u)\, du \geq \varepsilon \Big\},$$

where ε is a unit exponential random variable on $(\Omega, \mathcal{G}, \mathbb{Q})$ independent of W, with $\mathcal{F} \subseteq \mathcal{G}$. Thus

$$Q_t = \mathbb{Q}(\theta > t \,|\, \mathcal{F}_t) = \exp\big(- \int_0^t \gamma(Z_u)\, du\big) = \mathbb{Q}(\theta > t \,|\, \mathcal{F}) ,$$

which is (strictly) positive, continuous and non-increasing on $\mathbb{R}_+$. In particular immersion holds between $\mathbb{F}$ and $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$, where $\mathbb{H}$ is the natural filtration of the indicator process I of θ. So the general assumptions of section 15 are satisfied.

Now we define the *pre-default stock price process* of the firm by $\widetilde{S} = Z$, and the related *discounted cumulative stock price process* as (with as usual $J = 1 - I$):

$$\beta_t \widehat{S}_t = \beta_t J_t \widetilde{S}_t + \int_{[0,t]} \beta_u \big(q(u) J_u \widetilde{S}_u du + (1-\eta) \widetilde{S}_u dI_u \big) \ .$$

Let us further examine the valuation in the present model of a (protection payer, say) CDS written on the reference entity. Note that the following analysis can be extended to a *rolling CDS* more realistically used as an hedging instrument [5].

Consistently with arbitrage requirements (cf. [4, 5]), we assume that the *pre-default CDS price* $\widetilde{B}_t$, $t \in [0, \mathcal{T}]$ (for some constant CDS maturity $\mathcal{T} > T$) is given as $\widetilde{B}_t = \widetilde{B}(t, \widetilde{S}_t)$, where the *pre-default CDS pricing function* $\widetilde{B}(t, Z)$ is the unique (classical) solution to the following PDE:

$$(\partial_t + \mathcal{Z}) \widetilde{B}(t,z) + \delta(t,z) - \mu(t,z) \widetilde{B}(t,z) = 0 \,, t < \mathcal{T} \ , \ \widetilde{B}(\mathcal{T}, z) = 0,$$

where:
- the operator $\mathcal{Z}$ is given by (58),
- $\delta(t,z) = \bar{\nu} - \nu(t)\gamma(z)$ is the *pre-default dividend function* of the CDS (see [5]),
- $\mu(t,z) = r(t) + \gamma(z)$ is the *credit-risk adjusted interest rate.*

The *discounted cumulative CDS price* $\beta \widehat{B}$ equals, for every $t \in [0, T]$,

$$\beta_t \widehat{B}_t = \beta_t J_t \widetilde{B}_t + \int_{[0,t]} \beta_u \big(\bar{\nu}\, J_u du - \nu(u)\, dI_u \big) \ .$$

Remark 19.1 Given the specification of the drift of Z in (57), an easy computation shows that $\beta \widehat{S}$ and $\beta \widehat{B}$ are $\mathbb{G}$-local martingales (see [5]). The arbitrage assumption of section 14.1 is thus satisfied, as can also be seen by application of the general *arbitrage pre-default drift condition* of [4].

Denoting $\widetilde{X}_t = \begin{pmatrix} \widetilde{S}_t \\ \widetilde{B}_t \end{pmatrix} = v(t, Z_t)$, we have in this set-up (∂ standing for the partial derivative with respect to Z):

$$\big(\sigma Z_t \partial v(t, Z_t), \Delta v(t, Z_t)\big) = \begin{pmatrix} \sigma Z_t & -\eta Z_t \\ \sigma Z_t \partial \widetilde{B}(t, Z_t) & \nu(t) - \widetilde{B}_t \end{pmatrix} \ . \qquad (59)$$

Let us additionally be given a *(game) option* on this market with related pre-default price process $Y_t = u(t, Z_t)$ over $[0, T]$, and pricing equation (52) with $\mathcal{Z}$ given as (58) therein.

By application of (56), we have the following decomposition of the cost ρ of the strategy ζ for hedging the option, where $H = I - \int_0^{\cdot} J_t \gamma(Z_t) dt$ stands as usual for the $\mathbb{G}$-compensated jump-to-default process.

Proposition 19.1 *For every* $t \in [0, T \wedge \theta]$,

$$d\rho_t = \Big[\big(\sigma Z_t \partial u(t, Z_t), \Delta u(t, Z_t)\big) - \zeta_t \big(\sigma Z_t \partial v(t, Z_t), \Delta v(t, Z_t)\big) \Big] d \begin{pmatrix} W_t \\ H_t \end{pmatrix} \ .$$

Consequently, if the matrix $\big(\sigma Z_t \partial v(t, Z_t), \Delta v(t, Z_t)\big)$ (cf. (59)) is invertible on $[0, \theta \wedge T]$, one can superhedge the option (market and default risk, so $\rho = 0$) by setting, for $t \in [0, T \wedge \theta]$,

$$\zeta_t = \widehat{\zeta}_t := \big(\sigma Z_t \partial u(t, Z_t), \Delta u(t, Z_t)\big)\big(\sigma Z_t \partial v(t, Z_t), \Delta v(t, Z_t)\big)^{-1}$$

(and σ^* defined as usual by (42)).

Otherwise it is still possible to hedge the market risk (represented by W) by setting $\zeta = \widetilde{\zeta}$ with, for $t \in [0, T \wedge \theta]$,

$$\widetilde{\zeta}^1_t = \tfrac{\partial u}{\partial v}(t, Z_t) \ , \ \widetilde{\zeta}^2_t = 0$$

(and σ^* as before), whence in this case

$$d\rho_t = \Big[\Delta u(t, Z_t) + \eta Z_t \widetilde{\zeta}^1_t\Big] dH_t = \Big[R_t - u(t, Z_t) + \eta Z_t \tfrac{\partial u}{\partial v}(t, Z_t)\Big] dH_t \ .$$

Part IV

BSDE and PDE Results

Introduction

In part III we essentially reduced the problem of pricing and hedging defaultable game options to that of solving related doubly reflected BSDEs (or the associated integro-differential variational inequalities, in the Markovian case).

In this part, which is a synthesis of the main results of Crépey and Matoussi [15] and Crépey [8], we tackle the resulting BSDE and PIDE problems. Note that the results of this part are also used in Part V, where analogous equations arise in the context of portfolio credit risk.

In [15] (sections 20 to 24 below), a priori estimates and comparison principles are derived for reflected or doubly reflected BSDEs in the rather general set-up of a model driven by a continuous local martingale and an integer-valued random measure.

In [8] (sections 25 to 27 below), we use these results to establish the well-posedness of a *Markovian doubly reflected BSDE*, and of the associated *system of partial integro-differential double obstacle problem*, in a rather flexible Markovian set-up made of a Jump–Diffusion model with Regimes. As an aside we prove the convergence of any *stable, monotone and consistent* approximation scheme to the above pide system.

Section 28 presents the mapping between the mathematical set-up of this part and the financial problems of parts III and V.

Note that the papers [15, 8] also deal with BSDEs *with random terminal time*, and the related Cauchy–Dirichlet problems in the Markovian case. This is actually an important issue for applications, in which the random terminal time may for instance represent a call protection (cf. Remark 14.2(ii)). ***To keep it simple here, we do not deal with these issues in the present report, referring the interested reader to [15, 8] for the related developments.***

20 Abstract Set-Up

Let us be given a finite time horizon $T > 0$, a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ with $\mathcal{F}_T = \mathcal{F}$, satisfying the usual conditions of right-continuity and completeness. By default we declare that a *random variable* is $\mathcal{F}$-measurable, and that a *process* is defined on the time interval $[0, T]$ and $\mathbb{F}$-adapted. All semimartingales are assumed to be càdlàg, without restriction.

Let $B = (B_t)_{t\in[0,T]}$ be a d-dimensional standard Brownian motion. Given an auxiliary measured space $(E, \mathcal{B}_E, \rho)$, where ρ is a non-negative σ-finite measure on $(E, \mathcal{B}_E)$, let $\mu = (\mu(dt, de))_{t\in[0,T],e\in E}$ be an *integer valued random measure* on $\big([0, T] \times E, \mathcal{B}([0, T]) \otimes \mathcal{B}_E\big)$ (see, e.g., Jacod–Shiryaev [84, Definition II.1.13 page 68]).

We assume that the compensator of μ is defined by $\zeta_t(\omega, e)\rho(de)dt$, for some $\mathcal{P} \otimes \mathcal{B}_E$-measurable non-negative bounded (random) function ζ, where $\mathcal{P}$ is the predictable σ-field on $\Omega \times [0, T]$. The introduction of the (not really standard) random density ζ is motivated by applications (for a concrete example see section 25.2 below, equation (76)). We denote by $\widetilde{\mu}(dt, de) = \mu(dt, de) - \zeta_t(e)\rho(de)dt$ the compensatrix of μ.

By default in the sequel, all (in)equalities between random quantities are to be understood $d\mathbb{Q}$-almost surely, $d\mathbb{Q} \otimes dt$-almost everywhere or $d\mathbb{Q} \otimes dt \otimes \zeta_t(e)\rho(de)$-almost everywhere, as suitable in the situation at hand. Moreover we omit all dependences in ω of any process or random function in the notation, for simplicity.

21 Reflected and Doubly Reflected BSDEs

Definition 21.1 A *solution* to the doubly reflected backward stochastic differential equation (R2BSDE for short) with data (g, ξ, L, U), is a quadruple (Y, Z, V, K), such that:

$$\left.\begin{array}{l} \textbf{(i) } Y_t = \xi + \displaystyle\int_t^T g_s(Y_s, Z_s, V_s)ds + K_T - K_t \\ \qquad\qquad - \displaystyle\int_t^T Z_s dB_s - \int_t^T \int_E V_s(e)\widetilde{\mu}(ds, de)\ ,\ t \in [0, T] \\ \textbf{(ii) } L_t \le Y_t \le U_t,\ t \in [0, T] \\ \text{ and } \displaystyle\int_0^T (Y_t - L_t)dK_t^+ = \int_0^T (U_t - Y_t)dK_t^- = 0\ , \end{array}\right\} (\mathcal{E})$$

with:

• $Y \in \mathcal{S}^2$, the space of real valued càdlàg processes such that

$$\|Y\|_{\mathcal{S}^2}^2 := \mathbb{E}\big[\sup_{t\in[0,T]} |Y_t|^2\big] \ <\ +\infty\,;$$

• $Z \in \mathcal{H}_d^2$ (or $\mathcal{H}^2$, in case $d = 1$), the space of $\mathbb{R}^{1\otimes d}$-valued predictable processes Z such that

$$\|Z\|_{\mathcal{H}_d^2} := \Big(\mathbb{E}\Big[\int_0^T |Z_t|^2\, dt\Big]\Big)^{\frac{1}{2}} \ <\ +\infty\ ;$$

• $V \in \mathcal{H}_\mu^2$, the space of $\widetilde{\mathcal{P}}$-measurable functions $V : \Omega \times [0, T] \times E \to \mathbb{R}$ such that

$$\|V\|_{\mathcal{H}_\mu^2} := \Big(\mathbb{E}\Big[\int_0^T |V_t|_t^2 dt\Big]\Big)^{\frac{1}{2}} \ <\ +\infty$$

where $|v|_t^2 = \displaystyle\int_E v(e)^2\zeta_t(e)\rho(de)$;

• $K \in \mathcal{V}^2$, the space of finite variation processes with continuous Jordan components $K^\pm \in \mathcal{S}^2$ null at time 0;

In particular:

• $\displaystyle\int_0^\cdot Z_t dB_t$ and $\displaystyle\int_0^\cdot\int_E V_t(e)\widetilde{\mu}(dt, de)$ are martingales, for any $Z \in \mathcal{H}_d^2$ and $V \in \mathcal{H}_\mu^2$;
• $K = K^+ - K^-$ where $K^\pm$ define mutually singular measures on $\mathbb{R}^+$, for any $K \in \mathcal{V}^2$.

In the case of a progressive process X we shall abusively use the notation $\|X\|_{\mathcal{H}_d^2}$ for $\Big(\mathbb{E}\Big[\displaystyle\int_0^T |X_t|^2\, dt\Big]\Big)^{\frac{1}{2}}$ whenever this is a well defined quantity in $\mathbb{R}\cup\{\infty\}$, whether the process X belongs to $\mathcal{H}_d^2$ or not (whether X is predictable or not).

Let us now consider the case when there is only one barrier, say, for instance, a lower barrier L. A *solution* to the reflected BSDE (RBSDE, for short) with data (g, ξ, L), is a quadruple $(Y, Z, V, K) \in (\mathcal{S}^2, \mathcal{H}^2_d, \mathcal{H}^2_\mu, \mathcal{V}^2)$ with $K^- = 0$ (so $K = K^+$ is a continuous non-decreasing process vanishing at time 0) such that:

$$\left.\begin{aligned} &\textbf{(i) } Y_t = \xi + \int_t^T g_s(Y_s, Z_s, V_s)ds + K_T - K_t \\ &\qquad - \int_t^T Z_s dB_s - \int_t^T \int_E V_s(e)\widetilde{\mu}(ds, de)\ ,\ t \in [0, T] \\ &\textbf{(ii) } L_t \le Y_t,\ t \in [0, T] \text{ and } \int_0^T (Y_t - L_t)dK_t = 0\ . \end{aligned}\right\} (\mathcal{E}')$$

When there is no barrier, we define likewise solutions to BSDEs with data (g, ξ).

We assume, denoting $\mathcal{M}_\rho = \mathcal{M}(E, \mathcal{B}_E, \rho; \mathbb{R})$ in (H.1.iii):
(**H.0**) $\xi \in \mathcal{L}^2$, the space of square integrable real valued ($\mathcal{F}_T$-measurable) random variables such that

$$\|\xi\|^2_{\mathcal{L}^2} := \mathbb{E}\big[|\xi|^2\big] \ < \ +\infty\ ;$$

(**H.1.i**) $g_\cdot(y, z, v)$ is a progressively measurable process, for any $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\otimes d}$, $v \in \mathcal{M}_\rho$;
(**H.1.ii**) $\|g_\cdot(0, 0, 0)\|_{\mathcal{H}^2} < \infty$;
(**H.1.iii**) g is uniformly Lipschitz with respect to (y, z, v), in the sense that there exists a constant $\Lambda \ge 0$ such that for every $(\omega, t) \in \Omega \times [0, T], y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^{1\otimes d}$, $v, v' \in \mathcal{M}_\rho$:

$$|g_t(y, z, v) - g_t(y', z', v')| \le \Lambda(|y - y'| + |z - z'| + |v - v'|_t)\ ;$$

(**H.2.i**) L and U are càdlàg processes in $\mathcal{S}^2$;
(**H.2.ii**) $L_t \le U_t$, $t \in [0, T)$ and $L_T \le \xi \le U_T$.

22 A Priori Bound and Error Estimates

This section extends to R2BSDEs with jumps the results of El Karoui et al. [61] regarding RBSDEs in a continuous set-up (see also [60] for a survey).

Recall that a *quasimartingale* X can be defined as a difference of two non-negative supermartingales, the minimal such decomposition being called the *Rao decomposition* of X (see sections VI.38 to VI.42 and Appendix 2 of Dellacherie and Meyer [52]; see also Protter [110, Chapter III, section 4]). In particular:
• Any process given as $S \vee c$ where S stands for an Itô process in $\mathcal{S}^2$ and c is a constant in $\mathbb{R} \cup \{-\infty\}$, is a quasimartingale in $\mathcal{S}^2$ (cf. Proposition 24.3 below);
• Any quasimartingale in $\mathcal{S}^2$ is a special semimartingale with canonical decomposition

$$L_t = L_0 + M_t + A_t\ ,\ t \in [0,\, T] \tag{60}$$

for a *uniformly integrable* martingale component M and a predictable process *of integrable variation* A.

Assuming that L (resp. U) is a quasimartingale in $\mathcal{S}^2$, Lemma 22.1 below provides an explicit representation for the process K^+ (resp. K^-) of a solution, assumed to exist, to $(\mathcal{E})$. Since the roles of L and U are entirely symmetrical in this regard (considering the problem with data $(-g, -\xi, -L, -U)$), we only state and prove the results regarding L.

Lemma 22.1 *Let (Y, Z, V, K) be a solution to ($\mathcal{E}$), in case L is a quasimartingale in $\mathcal{S}^2$ with canonical decomposition (60). Then*

$$dK_t^+ \leq \mathbb{1}_{\{Y_t = L_t\}}\big(\, g_t^-(Y_t, Z_t, V_t)dt \, + dA_t^-\big) \; , \tag{61}$$

where $A = A^+ - A^-$ is the Jordan decomposition of A.
In particular, if $dA_t^- \leq \alpha_t dt$ for some progressively measurable time-integrable process α, then K^+ is a time-differentiable process with derivative k^+ such that

$$k_t^+ \leq \mathbb{1}_{\{Y_t = L_t\}}\big(g_t^-(Y_t, Z_t, V_t) + \alpha_t\big) \; , \; t \in [0, T] \; . \tag{62}$$

Sketch of Proof. This follows by identification of the expressions for $d(Y - L)$ and $d(Y - L)^+$ respectively obtained by using ($\mathcal{E}$) and the Itô-Tanaka formula (see [15]). □

Using the direct control over K^+ (or K^-) provided by Lemma 22.1, and controlling the remaining terms of the solution to ($\mathcal{E}$) by the equation, one can then derive the following a priori bound and error estimates.

Theorem 22.2 (Crépey and Matoussi [15]) *Let us consider a sequence of R2BSDE problems as in the first part of Lemma 22.1, with data and solutions indexed by n, the data being bounded in the sense that the driver coefficients g^n are Λ-equilipschitz, and for some constant c_1 :*

$$\|\xi^n\|^2_{\mathcal{L}^2} + \|g^n_\cdot(0,0,0)\|^2_{\mathcal{H}^2} + \|L^n\|^2_{\mathcal{S}^2} + \|U^n\|^2_{\mathcal{S}^2} + \|A^{n,-}\|^2_{\mathcal{S}^2} \;\leq\; c_1 \; .$$

Then we have for some constant $c(\Lambda)$:

$$\|Y^n\|^2_{\mathcal{S}^2} + \|Z^n\|^2_{\mathcal{H}^2_d} + \|V^n\|^2_{\mathcal{H}^2_\mu} + \|K^{n,+}\|^2_{\mathcal{S}^2} + \|K^{n,-}\|^2_{\mathcal{S}^2} \leq c(\Lambda)c_1 \; . \tag{63}$$

Indexing by ^{n,p} the differences $X^n - X^p$ for any sequence (X^n), we also have:

$$\|Y^{n,p}\|^2_{\mathcal{S}^2} + \|Z^{n,p}\|^2_{\mathcal{H}^2_d} + \|V^{n,p}\|^2_{\mathcal{H}^2_\mu} + \|K^{n,p}\|^2_{\mathcal{S}^2} \leq \tag{64}$$
$$c(\Lambda)c_1\Big(\|\xi^{n,p}\|^2_{\mathcal{L}^2} + \|g^{n,p}_\cdot(Y^n_\cdot, Z^n_\cdot, V^n_\cdot)\|^2_{\mathcal{H}^2} + \|L^{n,p}\|_{\mathcal{S}^2} + \|U^{n,p}\|_{\mathcal{S}^2}\Big) \; .$$

Remark 22.1 By symmetry the same results are valid in case the U^n are quasimartingales (with $dA^{n,+} \leq \alpha_t^n dt$ for some progressively measurable processes α^n such that $\|\alpha^n\|_{\mathcal{H}^2}$, for the last part of the theorem).

In the case of RBSDE problems like ($\mathcal{E}'$), we have likewise the following result (without specific structure assumptions on the barrier, here).

Theorem 22.3 *Let us consider a sequence of RBSDE problems, the data being bounded in the sense that the driver coefficients g^n are Λ-equilipschitz, and for some constant c_1 :*

$$\|\xi^n\|^2_{\mathcal{L}^2} + \|g^n_\cdot(0,0,0)\|^2_{\mathcal{H}^2} + \|L^n\|^2_{\mathcal{S}^2} \;\leq\; c_1 \; .$$

Then we have for some constant $c(\Lambda)$:

$$\|Y^n\|^2_{\mathcal{S}^2} + \|Z^n\|^2_{\mathcal{H}^2_d} + \|V^n\|^2_{\mathcal{H}^2_\mu} + \|K^n\|^2_{\mathcal{S}^2} \leq c(\Lambda)c_1 \; . \tag{65}$$

Indexing by ^{n,p} the differences $X^n - X^p$, for any sequence (X^n), we also have:

$$\|Y^{n,p}\|^2_{\mathcal{S}^2} + \|Z^{n,p}\|^2_{\mathcal{H}^2_d} + \|V^{n,p}\|^2_{\mathcal{H}^2_\mu} + \|K^{n,p}\|^2_{\mathcal{S}^2} \leq \tag{66}$$
$$c(\Lambda)c_1\Big(\|\xi^{n,p}\|^2_{\mathcal{L}^2} + \|g^{n,p}_\cdot(Y^n_\cdot, Z^n_\cdot, V^n_\cdot)\|^2_{\mathcal{H}^2} + \|L^{n,p}\|_{\mathcal{S}^2}\Big) \; .$$

23 Comparison Principle

In this section we specialize (H.1) to the case where

$$g_t(y,z,v) = h_t\Big(y, z, \int_E v(e)\eta_t(e)\zeta_t(e)\,\rho(de)\Big)\;, \tag{67}$$

for a $\mathcal{P}\otimes\mathcal{B}_E$-measurable non-negative function $\eta_t(e)$ with $|\eta_t|_t$ bounded, and a $\mathcal{P}\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}^{1\otimes d})\otimes\mathcal{B}(\mathbb{R})$-measurable function $h:\Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^{1\otimes d}\times\mathbb{R}\to\mathbb{R}$ such that:
(H.1.i)' $h_.(y,z,r)$ is a progressively measurable process, for any $y\in\mathbb{R}$, $z\in\mathbb{R}^{1\otimes d}$, $r\in\mathbb{R}$;
(H.1.ii)' $h_.(0,0,0)$ is a square integrable process;
(H.1.iii)' $|h_t(y,z,r)-h_t(y',z',r')|\le\Lambda\big(|y-y'|+|z-z'|+|r-r'|\big)$, for any $(\omega,t)\in\Omega\times[0,T]$, $y,y'\in\mathbb{R}$, $z,z'\in\mathbb{R}^{1\otimes d}$ and $r,r'\in\mathbb{R}$;
(H.1.iv)' $r\mapsto h_t(y,z,r)$ is non-decreasing, for any $(\omega,t,y,z)\in\Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^{1\otimes d}$.

Using in particular the fact that

$$|\int_E (v(e)-v'(e))\eta_t(e)\zeta_t(e)\,\rho(de)\;|\le|v-v'|_t|\eta_t|_t\;,$$

with $|\eta_t|_t$ bounded, then it is immediate that g defined by (67) satisfies (H.1).

Our next result is a comparison result for $(\mathcal{E})$ in this case, extending to doubly reflected BSDEs the comparison principle of Proposition 2.6 in Barles et al. [22] for BSDEs without barriers (see [22, Remark 2.7 page 64] for a counter-example in the general case, not assuming (H.1.iv)').

Proposition 23.1 (Crépey and Matoussi [15]) *Let (Y,Z,V,K) and (Y',Z',V',K') be solutions to the R2BSDEs with data (g,ξ,L,U) and (g',ξ',L',U') satisfying assumptions (H.0)–(H.1)–(H.2). We assume further that g satisfies (H.1)'. Then $Y_t\le Y'_t$, $t\in[0,T]$, whenever:*
(i) $\xi\le\xi'$,
(ii) $g_t(Y'_t,Z'_t,V'_t)\;\le\;g'_t(Y'_t,Z'_t,V'_t)\,,t\in[0,T]$
(iii) $L_t\le L'_t$ *and* $U_t\le U'_t$, $t\in[0,T]$.

Sketch of Proof. The proof consists in extending to reflected BSDEs the classical proof by linearization for establishing comparison in the case without barriers (see, e.g., El Karoui, Peng and Quenez [62]). □

Note that this comparison principle admits obvious specifications to RBSDEs and BSDEs. In the latter case, we recover the comparison principle of Barles et al. [22].

24 Existence and Uniqueness Issues

We now deal with the issues of existence and uniqueness of solutions to $(\mathcal{E})$ and $(\mathcal{E}')$.

As for uniqueness, an application of the error estimates of Theorems 22.2 and 22.3 yields,

Proposition 24.1 *Uniqueness holds for $(\mathcal{E})$ and $(\mathcal{E}')$.*

In order to get existence results, we need to make the following *square integrable martingale predictable representation* assumption:
(**R**) Every square integrable martingale M admits a representation

$$M_t = M_0 + \int_0^t Z_s \, dB_s + \int_0^t \int_E V_s(e)\widetilde{\mu}(ds, de) \,, \quad t \in [0, T] \tag{68}$$

for some $Z \in \mathcal{H}^2_d$ and $V \in \mathcal{H}^2_\mu$.

We also strengthen Assumption (H.2.i) into:
(**H.2.i**)$'$ L and U are càdlàg *quasi-left continuous* processes in $\mathcal{S}^2$,
where we recall for the reader's convenience that for a càdlàg process X, quasi-left continuity is equivalent to the existence of sequence of totally inaccessible stopping times which exhausts the jumps of X, implying that ${}^pX = X_{\cdot -}$ (Jacod–Shiryaev [84, Propositions I.2.26 page 22 and I.2.35 page 25]).

We thus work under assumptions (R)–(H.0)–(H.1)–(H.2)', where (H.2)' denotes (H.2) with (H.2.i) strengthened into (H.2.i)'.

The following existence result is essentially contained in earlier results by Hamadène and Ouknine [81] and Hamadène [80]. By *the Mokobodski condition* in this proposition, we mean the existence of a quasimartingale X with Rao components in $\mathcal{S}^2$ such that $L \le X \le U$ over $[0, T]$. In view of the properties of quasimartingales recalled at the first paragraph of section 22, this is tantamount to the existence of non-negative supermartingales X^1, X^2 belonging to $\mathcal{S}^2$ and such that $L \le X^1 - X^2 \le U$ over $[0, T]$.

Proposition 24.2 *Assuming (R)–(H.0)–(H.1)–(H.2)':*
(i) *Existence holds for ($\mathcal{E}'$);*
(ii) *Existence of a solution to ($\mathcal{E}$) is equivalent to the Mokobodski condition. In particular, existence holds for ($\mathcal{E}$) when L or U is a quasimartingale with Rao components in $\mathcal{S}^2$.*

Remark 24.1 In the situation of Proposition 24.2(ii), L or U is obviously a quasimartingale in $\mathcal{S}^2$ as postulated in Lemma 22.1, and the estimates of Theorem 22.2 are thus applicable.

We conclude this section by a proposition motivated by convertible bonds related R2BSDEs in finance, in which the lower barrier L is typically given by a call payoff functional of the underlying stock price process S (cf. Definition 14.3), the latter being typically modeled as a jump-diffusion with (possibly) random coefficients. We thus have the following result, whose proof is similar to that of Lemma 22.1.

Proposition 24.3 (Crépey and Matoussi [15, 8]) *Let S be given as an Itô process with square integrable special semimartingale components, so*

$$S_t = S_0 + \int_0^t a_s ds + \int_0^t z_s dB_s + \int_0^t \int_E v_s(e)\widetilde{\mu}(ds, de) \,, \; t \in [0, \, T] \tag{69}$$

for some $z \in \mathcal{H}^2_d$, $v \in \mathcal{H}^2_\mu$, and a progressively measurable time-integrable process a such that $\|a\|_{\mathcal{H}^2} < +\infty$. Let in turn L be given as $L = S \vee c$, for some constant $c \in \mathbb{R} \cup \{-\infty\}$.
Then L is a càdlàg quasi-left continuous quasimartingale with Rao components in $\mathcal{S}^2$. Moreover L satisfies all the conditions in Lemma 22.1 (including the hypotheses on L in (H.2)), with in particular α in (62) given by a^-, the negative part of a in (69).

25 Jump–Diffusion Setting with Regimes

We now present a rather generic specification for a Markovian factor process F underlying a BSDE, and we show how it fits into the abstract set-up of the previous sections.

Given integers d and k, we define the following linear operator $\mathcal{G}$ acting on regular functions $u = u^i(t,x)$, $(t,x,i) \in E = [0,T] \times \mathbb{R}^d \times I$, with $I = \{1,\dots,k\}$:

$$\begin{aligned}\mathcal{G}u^i(t,x) &= \partial_t u^i(t,x) + \tfrac{1}{2}\sum_{l,q=1}^{d} a^i_{l,q}(t,x)\partial^2_{x_l x_q} u^i(t,x) \\ &+ \sum_{l=1}^{d}\Big(b^i_l(t,x) - \int_{\mathbb{R}^d}\delta^i_l(t,x,y) f^i(t,x,y) m(dy)\Big)\partial_{x_l} u^i(t,x) \\ &+ \int_{\mathbb{R}^d}\big(u^i(t,x+\delta^i(t,x,y)) - u^i(t,x)\big) f^i(t,x,y) m(dy) \\ &+ \sum_{j\in I} n^{i,j}(t,x)(u^j(t,x) - u^i(t,x)) \ .\end{aligned} \tag{70}$$

In (70) $m(dy)$ is a *finite jump measure* on $\mathbb{R}^d$ (not charging $\{0_d\}$ where 0_d stands for the null in $\mathbb{R}^d$), and all the coefficients are Borelian functions such that:
• the $a^i(t,x)$ are d-dimensional *covariance* matrices, with $a^i(t,x) = \sigma^i(t,x)\sigma^i(t,x)^{\mathsf{T}}$ for some d-dimensional *dispersion* matrices $\sigma^i(t,x)$:
• the $b^i(t,x)$ are d-dimensional *drift* vector coefficients;
• the *jump intensity functions* $f^i(t,x,y)$ are bounded, and the *jump size functions* $\delta^i(t,x,y)$ are absolutely integrable with respect to $m(dy)$;
• the *regime switching intensity functions* $n^{i,j}(t,x)$ are such that the $n^{i,j}(t,x)$ are non-negative and bounded whenever $i \neq j$, and $n^{i,i}(t,x) = 0$ for every i.
We shall often find convenient to denote $v(t,x,i,\dots)$ rather than $v^i(t,x,\dots)$ for a function v of $(t,x,i,\dots)$, and $n(t,x,i,j)$, for $n^{i,j}(t,x)$.

Proposition 25.1 (Crépey [8]) *Under suitable conditions (see [8]), there exists a stochastic basis* $(\Omega,\mathbb{F},\mathbb{Q})$ *on* $[0,T]$ *endowed with a d-dimensional Brownian motion* B, *an* integer-valued random measure χ *and a càdlàg process* $F = (X,N)$ *on* $[0,T]$ *with initial condition* (x,i) *at time 0, such that:*
• *The* $\mathbb{R}^d$*-valued process* X *satisfies, for* $t \in [0,T]$:

$$dX_t = b(t,F_t)\,dt + \sigma(t,F_t)\,dB_t + \int_{\mathbb{R}^d}\delta(t,F_{t-},y)\,\widetilde{\chi}(dt,dy) \ , \tag{71}$$

where the $\mathbb{Q}$*-compensatrix* $\widetilde{\chi}$ *of* χ *is given by*

$$\widetilde{\chi}(dt,dy) = \chi(dt,dy) - f(t,F_t,y)m(dy)dt \ ;$$

• *The* $\mathbb{Q}$*-compensatrix* $\widetilde{\nu}$ *of the integer-valued random measure* ν *on* I *counting the number of transitions* $\nu_t(j)$ *of* N *to state* j *between time* 0 *and time* t, *is given by*

$$d\widetilde{\nu}_t(j) = d\nu_t(j) - n(t,F_t,j)\,dt \ . \tag{72}$$

Note that the construction of such a model, with mutual dependence between X and N, is a non-trivial issue. It is treated in detail in Crépey [8] by a *Markovian change of probability* approach (see also [6]), under suitable conditions on the coefficients.

Remark 25.1 **(i)** If we suppose that the coefficients b, σ, δ and f do not depend on i, then X is a jump-diffusion process. Alternatively, if n does not depend on x, then N is an (inhomogeneous) continuous-time Markov chain. In general N defines the so-called *regime* of the coefficients b, σ, δ and f, whence the name of *Jump–Diffusion Setting with Regimes* for this model.
For simplicity we do not consider the "infinite activity" case, that is, the case when the jump measure m is unbounded. Note however that our results could be extended to *Lévy measures* m without major changes if wished (see e.g. Barles et al. [22], recently complemented by Barles and Imbert [23]).
(ii) More specific sub-cases or related models were frequently considered in the literature. So:
• Barles et al. [22] consider jumps in X but no regimes N, and Lévy jump measures $m(dy)$ (cf. (i));
• Pardoux et al. [107] consider a diffusion model with regimes, which corresponds to the special case of our model in which f is equal to 0, and the regimes are driven by a standard Poisson process with constant intensity;
• Becherer and Schweizer consider in [26] a diffusion model with regimes, including the model of Pardoux et al. [107] as a special case, and which corresponds to the special case of our model in which f is equal to 0.

25.1 Itô formula and Martingale Representation

In this model we have the following Itô formula:

$$\begin{aligned} du(t, F_t) &= \mathcal{G}u(t, F_t)dt + \partial u(t, F_t)\sigma(t, F_t)dB_t \\ &+ \int_{\mathbb{R}^d} \big(u(t, X_{t-} + \delta(t, F_{t-}, y), N_{t-}) - u(t, F_{t-})\big)\widetilde{\chi}(dt, dy) \\ &+ \sum_{j\in I} (u(t, X_{t-}, j) - u(t, F_{t-}))d\widetilde{\nu}_t(j) \ , \ t \geq 0 \end{aligned} \tag{73}$$

for any system of $\mathcal{C}^{1,2}$-functions $u = (u^j)_{j\in I}$.
In particular $(\Omega, \mathbb{F}, \mathbb{Q}, F)$ is a solution to the *time-dependent local martingale problem* with generator $\mathcal{G}$ and initial condition (t, x, i) (see Ethier–Kurtz [66, sections 7.A and 7.B]).

Moreover, still under the above mentioned conditions (see [8]), χ and ν cannot jump together, and every $(\Omega, \mathbb{F}, \mathbb{Q})$-square integrable martingale M admits a representation

$$M_t = M_0 + \int_0^t Z_s \, dB_s + \int_0^t \int_{\mathbb{R}^d} \widetilde{V}_s(y)\widetilde{\chi}(ds, dy) + \sum_{j\in I} \int_0^t \widehat{V}_s(j)d\widetilde{\nu}_s(j) \, , \ \ t \in [0, T] \tag{74}$$

for some $Z \in \mathcal{H}^2_d$, $\widetilde{V} \in \mathcal{H}^2_\chi$ and $\widehat{V} \in \mathcal{H}^2_\nu$. Finally the following estimates are available, for any $p \in [2, +\infty)$:

$$\|X\|^p_{\mathcal{S}^p_d} \leq C_p\big(1 + |x|^p\big) \ . \tag{75}$$

25.2 Mapping with the Abstract Set-Up

Let 0_d stand for the null in $\mathbb{R}^d$. It is easy to check that the model $F = (X, N)$ is a (rather generic) Markovian specification of the abstract set-up of section 20, with (cf. section 20 and [15]):

• E, the subset $\big(\mathbb{R}^d \times \{0\}\big) \cup \big(\{0_d\} \times I\big)$ of $\mathbb{R}^{d+1}$;

• $\mathcal{B}_E$, the sigma field generated by $\mathcal{B}(\mathbb{R}^d) \times \{0\}$ and $\{0_d\} \times \mathcal{I}$ on E, where $\mathcal{B}(\mathbb{R}^d)$ and $\mathcal{I}$ stand for the Borel sigma field on $\mathbb{R}^d$ and the sigma field of all parts of I, respectively;

• $\rho(de)$ and $\zeta_t(e)$ respectively given by, for any $e = (y, j) \in E$:

$$\rho(de) = \begin{cases} m(dy) & \text{if} \quad j = 0 \\ 1 & \text{if} \quad y = 0_d \end{cases} , \quad \zeta_t(e) = \begin{cases} f(t, F_t, y) & \text{if} \quad j = 0 \\ n(t, F_t, j) & \text{if} \quad y = 0_d \ ; \end{cases} \tag{76}$$

• μ, the integer-valued random measure on $\big([0, T] \times E, \mathcal{B}([0, T]) \otimes \mathcal{B}_E\big)$ counting the jumps of X of size $y \in A$ and the jumps of N to state j between 0 and t, for any $t \geq 0$, $A \in \mathcal{B}(\mathbb{R}^d), j \in I$.

We denote for short:

$$(E, \mathcal{B}_E, \rho) = (\mathbb{R}^d \oplus I, \mathcal{B}(\mathbb{R}^d) \oplus \mathcal{I}, m(dy) \oplus \mathbf{1}) \ .$$

So, in the present context, the abstract set $\mathcal{M}_\rho$ of section 20 can be identified with the product space

$$\mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R}) \times \mathbb{R}^k \ , \tag{77}$$

and the compensator of μ is given by, for any $t \geq 0, A \in \mathcal{B}(\mathbb{R}^d), j \in I$:

$$\int_0^t \int_{A \oplus \{j\}} \zeta_s(e) \rho(de) ds = \int_0^t \int_A f(s, F_s, y) m(dy) ds + \int_0^t n(s, F_s, j)\, ds \ ,$$

where $A \oplus \{j\}$ is a notation for $\big(A \times \{0\}\big) \cup \big(\{0_d\} \times \{j\}\big)$.

Note finally that (74) is a martingale representation of the form (68), with for $e = (y, j)$:

$$V_s(de) = \begin{cases} \widetilde{V}_s(y) \text{ if } j = 0 \\ \widehat{V}_s(j) \text{ if } y = 0_d \ . \end{cases}$$

Hence the model F has the martingale representation property (R) of section 24.

26 Markovian BSDEs

We consider, in the Jump–Diffusion Setting with Regimes of section 25, the BSDE naturally connected with the Itô formula (73), namely for $t \geq 0$:

$$-\, dY_t = g(t, F_t, Y_t, Z_t, V_t) dt - Z_t dB_t - \int_{\mathbb{R}^d} \widetilde{V}_t(y) \widetilde{\chi}(dt, dy) - \sum_{j \in I} \widehat{V}_t(j) d\widetilde{\nu}_t(j)$$

with $V = (\widetilde{V}, \widehat{V})$, possibly supplemented by suitable barrier and minimality conditions, and for a suitable driver coefficient $g(t, F_t, y, z, v)$, where $v = (\widetilde{v}, \widehat{v})$ denotes a generic element of the product space $\mathcal{M}_\rho$ specified under the form of (77).

Let $\mathcal{P}$ denote the class of functions u on $[0, T] \times \mathbb{R}^d \times I$ such that u^i is Borel-measurable with polynomial growth in x for any $i \in I$. Let us further be given real-valued continuous *running cost functions* $\widetilde{g}^i(t, x, u, z, r)$ (where $(u, z, r) \in \mathbb{R}^k \times \mathbb{R}^{1 \otimes d} \times \mathbb{R}$), *terminal cost functions* $\Psi^i(x)$, and *lower and upper obstacle functions* $\ell^i(t, x)$ and $h^i(t, x)$, such that:

(M.0) Ψ lies in $\mathcal{P}$;
(M.1.i) $(t,x,i) \mapsto \widetilde{g}^i(t,x,0,0,0)$ lies in $\mathcal{P}$;
(M.1.ii) $\widetilde{g}$ is uniformly Λ – Lipschitz continuous with respect to (u,z,r), in the sense that Λ is a constant such that for every $(t,x,i) \in [0,T]\times\mathbb{R}^d\times I$ and $(u,z,r),(u',z',r') \in \mathbb{R}^k\times\mathbb{R}^{1\otimes d}\times\mathbb{R}$:

$$\left|\widetilde{g}^i(t,x,u,z,r) - \widetilde{g}^i(t,x,u',z',r')\right| \le \Lambda\left(\left|u-u'\right| + \left|z-z'\right| + \left|r-r'\right|\right) ;$$

(M.1.iii) $\widetilde{g}$ is non-decreasing with respect to r ;
(M.2.i) ℓ and h lie in $\mathcal{P}$;
(M.2.ii) $\ell \le h$, $\ell(T,\cdot) \le \Psi \le h(T,\cdot)$;

We define further for any $(t,y,z,v) \in [t,T]\times\mathbb{R}\times\mathbb{R}^{1\otimes d}\times\mathcal{M}_\rho$, with $v = (\widetilde{v},\widehat{v}) \in \mathcal{M}_\rho$:

$$g(t,F_t,y,z,v) = \widetilde{g}(t,F_t,\widetilde{u}_t,z,\widetilde{r}_t) - \sum_{j\in I}\widehat{v}_j n(t,F_t,j) , \tag{78}$$

where $\widetilde{u}_t = \widetilde{u}_t(y,\widehat{v})$ and $\widetilde{r}_t = \widetilde{r}_t(\widetilde{v})$ are defined as

$$(\widetilde{u}_t)^j = \begin{cases} y, & j = N_t \\ y + \widehat{v}_j, & j \neq N_t \end{cases} , \ \widetilde{r}_t = \int_{\mathbb{R}^d}\widetilde{v}(y)f(t,F_t,y)m(dy) . \tag{79}$$

We then consider the R2BSDE data set given in terms of the factor process F as

$$g_t(\omega,y,z,v) = g(t,F_t,y,z,v) , \ \xi = \Psi(F_T) , \ L_t = \ell(t,F_t) , \ U_t = h(t,F_t) . \tag{80}$$

We refer the reader to section 28.2 for simple examples in finance.

Proposition 26.1 *The data (80) satisfy assumptions (H.0)–(H.1)–(H.2)'.*

Proof. Given (M.0)–(M.1)–(M.2) and the estimate (75) on X, the verification of (H.0)–(H.1)–(H.2)' is straightforward (see [8] for the detail). □

The next step consists in specifying, in the model F, a concrete class of processes S which satisfy the abstract conditions of Proposition 24.3.

Proposition 26.2 (Crépey and Matoussi [15]) *Let $\phi = (\phi^i)_{i\in I}$ be a system of real-valued functions $\phi^i = \phi^i(t,x)$ of class $\mathcal{C}^{1,2}$ on $[0,T]\times\mathbb{R}^d$ such that*

$$\phi,\ \mathcal{G}\phi,\ \partial\phi\sigma,\ (t,x,i) \mapsto \int_{\mathbb{R}^d}|\phi^i(t,x+\delta^i(t,x,y))|m(dy) \ \in \mathcal{P} . \tag{81}$$

Then the process S defined by, for $t \in [0,T]$:

$$S_t = \phi(t,F_t) ,$$

is an Itô-Lévy process with square integrable special semimartingale decomposition components as postulated in Proposition 24.3, with related process a in (69) given as $a_t = \mathcal{G}\phi(t,F_t)$, for $t \in [0,T]$.

Proof. Under our polynomial growth assumptions and given the estimates (75) on X, the result easily follows by application of the Itô formula (73) to $\phi(t, F_t)$. □

Example 26.1 The standing example we have in mind for S in Proposition 24.3 is $S = X^1$, the first component of X of our model $F = (X, N)$ (assuming $d \geq 1$ therein). This corresponds to the case where $\phi^i(t, x) = x_1$ in Proposition 26.2. Note that in this case:

$$\mathcal{G}\phi = b_1,\ \partial\phi\sigma = \sigma_1,\ \int_{\mathbb{R}^d} |\phi^i(t, x + \delta^i(t, x, y))| m(dy) = \int_{\mathbb{R}^d} |x_1 + \delta_1^i(t, x, y)| m(dy)\ ,$$

so that (81) reduces to

$$b_1,\ \sigma_1,\ (t, x, i) \mapsto \int_{\mathbb{R}^d} |\delta_1^i(t, x, y)| m(dy) \in \mathcal{P}\ . \tag{82}$$

Putting everything together, we get,

Theorem 26.3 (Crépey and Matoussi [15]) *Given the data (80) with ℓ specified as $\phi \vee c$ where ϕ satisfies (81) (e.g., $\phi = x_1$, assuming (82)) and for some constant $c \in \mathbb{R} \cup \{-\infty\}$, then the related R2BSDE ($\mathcal{E}$) admits a unique solution (Y, Z, V, K). Moreover K^+ is a time-differentiable process with time-derivative k^+ satisfying (62). The RBSDE ($\mathcal{E}'$) also admits a unique solution. All the estimates and comparison principle derived earlier in this part are applicable.*

The results of this part are thus applicable to convertible bonds (cf. section 14.2.1), in rather general jump-diffusion models (see [4, 5]).

27 Variational Inequality Approach

We now work in the set-up and under the assumptions of sections 25–26, with ℓ given as in Theorem 26.3. We denote by (Y, Z, V, K), with $V = (\widetilde{V}, \widehat{V})$, the unique solution to ($\mathcal{E}$) (cf. Theorem 26.3). Our next goal is to establish the connection between the Markovian R2BSDE ($\mathcal{E}$) with data (80), and a related system of obstacles problems. We shall consider this issue from the point of view of *viscosity solutions* to this system. We refer the reader to the books by Bensoussan and Lions [27, 28] for results in suitable spaces of weak Sobolev solutions. An alternative weak Sobolev solutions approach will be dealt with in Crépey and Matoussi [14].

For any (real-valued, vector-valued or matrix-valued) functions u on E, let $\delta u^i(t, x, \cdot)$ and $\Delta u^{i,\cdot}(t, x)$ (or $\delta u^i(t, x)$ and $\Delta u^i(t, x)$, for short) denote the functions

$$\mathbb{R}^d \ni y \overset{\delta u^i(t,x)}{\longrightarrow} u^i(t, x + \delta^i(t, x, y)) - u^i(t, x)$$
$$I \ni j \overset{\Delta u^i(t,x)}{\longrightarrow} u^j(t, x) - u^i(t, x)$$

and let

$$\overline{\delta} u^i(t, x) = \textstyle\int_{\mathbb{R}^d} \delta u^i(t, x, y) f^i(t, x, y) m(dy)$$
$$\overline{\Delta} u^i(t, x) = \textstyle\sum_{j \in I} \Delta u^{i,j}(t, x) n^{i,j}(t, x)\ .$$

Denoting (cf. 70)

$$\widetilde{\mathcal{G}}u^i(t,x) = \mathcal{G}u^i(t,x) - \overline{\Delta}u^i(t,x) \ , \tag{83}$$

we introduce the following *PIDE obstacle problem* (system of k coupled semi-linear PIDEs with obstacles in space dimension d):

$$\begin{aligned} \min\Big(\max\Big(\widetilde{\mathcal{G}}u^i(t,x) + \widetilde{g}^i(t,x,u(t,x),(\partial u\sigma)^i(t,x),\overline{\delta}u^i(t,x)), \\ \ell^i(t,x) - u^i(t,x)\Big), h^i(t,x) - u^i(t,x)\Big) = 0 \end{aligned} \tag{84}$$

on $[0,T) \times \mathbb{R}^d \times I$, supplemented by the terminal condition Ψ (our terminal cost function) at T.

Note that as opposed to the set-up of Becherer and Schweizer [26] where *linear* reaction-diffusion systems of parabolic equations are considered in a diffusion model with regimes (*without jumps* in X), here, due to the nonlinearities (presence of the *obstacles* and of the nonlinear term $\widetilde{g}$) in (84) and of the jumps in X, equation (84) typically does not admit a classical solution. Also note in this regard that we make no non-degeneracy assumption on the diffusion coefficient σ of X.

By a *solution* to (84), we thus mean a *viscosity solution* with polynomial growth in x to (84), adapting the general definitions of viscosity solutions for nonlinear PDEs (see [46, 68]) to (finite activity) *jumps* and *systems* of PIDEs as in Alvarez and Tourin [18], Pardoux et al. [107, 22], Pham [108], or Briani et al. [39].
Again (see Remark 25.1(i)), we exclude "small jumps" in the model (restricting our attention to finite jump measures m) for simplicity, yet our approach can be extended to more general Lévy jump measures without major changes if wished. But the viscosity solution techniques involved become much more complicated, as demonstrated in Barles and Imbert [23].
We also adapt to systems of PIDEs the notions of *stable, monotone and consistent approximation schemes* originally introduced for non linear PDEs by Barles and Souganidis [25]; see also Briani, La Chioma and Natalini [39], Cont and Voltchkova [45] or Jakobsen et al. [85] for various extensions of these results to PIDEs.

The following results thus extend to models with regimes (and therefore *systems* of PIDEs) the results of [25, 39], among others.

Theorem 27.1 (Crépey [8]) *Under suitable technical conditions (see [8]):*
(i) *Equation (84) admits a unique solution u. Moreover, we have for every $t \in [0,T]$:*

$$Y_t = u(t,F_t) \tag{85}$$

$$\widehat{V}_t(j) = u^j(t,X_{t-}) - u(t,F_{t-}) \ , \ j \in I \tag{86}$$

$$\int_0^t g(s,F_s,Y_s,Z_s,V_s)ds = \int_0^t \Big[\widetilde{g}(s,F_s,u(s,X_s),Z_s,\widetilde{r}_s) - \overline{\Delta}u(s,F_s)\Big]ds \tag{87}$$

with in (87) (cf. (79));

$$u(s,X_s) := (u^j(s,X_s))_{j\in I} \ , \ \widetilde{r}_s = \int_{\mathbb{R}^d} \widetilde{V}_s(y)f(s,F_s,y)m(dy) \ ;$$

(ii) *Any stable, monotone and consistent approximation scheme (u_h) for u, converges locally uniformly to u as $h \to 0$.*

Sketch of Proof. In part (i), existence and representation results are obtained by BSDE techniques, using the results of section 26; uniqueness follows from typical viscosity arguments (cf. in particular Barles et al. [22]).
Part (ii) extends to models with regimes (whence *systems* of PIDEs) the classical convergence arguments of Barles and Souganidis [25]. □

28 Mapping with Financial Applications

28.1 Model Dynamics

The Jump–Diffusion Setting with Regimes of section 25 admits versatile applications in financial modeling.

In Bielecki et al. [3] (see part III), this model is presented as a flexible risk-neutral pricing model in finance, for *equity and equity-to-credit (single-name credit) derivatives.* In this case the main component of the model (the one in which the *payoffs* of the product under consideration are expressed) is X, while N represents *implied pricing regimes* which may be viewed as a simple, whence robust, way, to implement *stochastic volatility* (whereas more standard, diffusive, forms of stochastic volatility, may be accounted for in the diffusive component of X).
In the context of single-name credit derivatives, N may also represent the credit rating of the reference obligor. So, in the area of *structural arbitrage*, *credit–to–equity* models and/or *equity–to–credit* models are studied. Our market model nests both types of interactions. For example, if one of the factors is the price process of the equity issued by a credit name, and if credit migration intensities depend on this factor, then we have an equity–to–credit type interaction. On the other hand, if the credit rating of the obligor impacts the equity dynamics, then we deal with a credit–to–equity type interaction.

In Bielecki et al. [6] (see section 30 below), this model is used in the context of *multi-name credit risk* for the valuation and hedging of basket credit derivatives. The main component in the model is then the 'Markov chain like' component N, representing the vector of (implied) credit ratings of the reference obligors, which is modulated by the 'jump-diffusion like' component X, representing the evolution of economic variables which impact the likelihood of credit rating migrations. *Frailty* and *default contagion* are accounted for in the model by the coupled interaction between N and X.

28.2 Cost Functionals

In the context of typical (risk-neutral) pricing problems in finance:
• the function $\widetilde{g}$ is of the form

$$\widetilde{g}^i(t,x,u,z,r) = c^i(t,x) - \mu^i(t,x)u^i + \sum_{j\in I} n^{i,j}(t,x)(u^j - u^i) \ , \tag{88}$$

thus (cf. (85)–(87))

$$g(t,F_t,Y_t,Z_t,V_t) = c(t,F_t) - \mu(t,F_t)Y_t \tag{89}$$

in (87), for *dividend-yield and riskless interest-rate* functions c and μ;
• $\Psi(F_T)$ corresponds to a *terminal payoff* that is paid by the issuer to the holder at time T

if the contract was not exercised before T;
• $\ell(F_t)$, resp. $h(F_t)$, corresponds to a *lower*, resp. *upper payoff* that is paid by the issuer to the holder of the claim in the event of early termination of the contract at the initiative of the holder, resp. issuer.

The contingent claims under consideration are thus general *Game Contingent Claims* (see [2, 3, 4, 5]), covering American Claims (and European Claims) as special cases. From the point of view of the financial interpretation, the components of F are observable *factors* (see section 28.1).

Note that $\widetilde{g}$ in (88) does not depend on z nor r, so $\widetilde{g}^i(t, x, u, z, r) = \widetilde{g}^i(t, x, u)$ therein. However, modeling the pricing problem under the historical probability (as opposed to the risk-neutral probability in (88)), would lead to a '*z-dependent*' driver coefficient function $\widetilde{g}$. Moreover we tacitly assumed above a perfect, frictionless financial market. Accounting for market imperfections would lead to *nonlinear* coefficients $\widetilde{g}$.

Also note that in a context of *vulnerable claims* (*single-name credit risk*, cf. Part III), it is enough, to account for credit-risk, to work with suitably *credit-spread adjusted interest-rates* μ and *recovery-adjusted dividend-yields* c in (88) (see Part III and [3, 4]).

Part V
Portfolio Credit Risk

Introduction

The goal of this part is to present some works related to the valuation and hedging of portfolio credit risk (valuation and hedging of basket credit derivatives in particular) [6, 1, 7]. Thus, we are concerned with modeling dependent defaults, and more generally, in the context of several (possibly implied) credit ratings of underlying credit instruments, dependent credit migrations. On the mathematical level, we are concerned with modeling dependence between random times (or processes) and with evaluation of functionals of (dependent) random times (or processes).

In [6] (section 30 below), we propose a fairly general Markovian model of portfolio credit risk, nesting several models that were previously studied in the literature.

In [1] (section 31), we use (a specific sub-case of) the previous model for illustrating the fact that the choice of a credit model with simplistic dynamics may have dangerous consequences in terms of hedging.

Of course complex models are computationally intensive. In case of basket credit derivatives, they typically do not give access to closed-form pricing formulae, so that pricing and calibration in the model needs to be done by simulation. An important issue in this regard is *variance reduction*. Importance sampling is often regarded as the method of choice when it comes to variance reduction. Importance sampling and related particle methods for portfolio credit risk are dealt with in [7] (section 32).

Finally section 33, which is the concluding section of this report, presents a ***research project*** inspired by the equity derivatives analysis of [10] (see section 13 above). The objective is to compare a dynamic model with the static Gaussian copula model which has long been the industry standard, in terms of hedging of a credit portfolio derivative with the credit index.

This part of the report is less mathematically-oriented than the previous ones. Here our motivation is rather to discuss important ***financial engineering issues*** regarding portfolio credit derivatives modeling:
(i) Is *dynamic better than static?* (section 33),
(ii) *Which dynamics* really matters (section 31),
(iii) Admittedly, computations (by simulation, typically) are intensive in models with pertinent, possibly high-dimensional and complex dynamics, however suitable forms of *importance sampling* may help in this regard (section 32).

29 Definitions and Preliminaries

Considering a pool (portfolio) of n credit names, we denote by τ_i the default time corresponding to the i-th name, by $H_t^i = \mathbb{1}_{\tau_i \le t}$ the related default indicator processes, and by R an homogeneous and constant recovery at default. We define the *cumulative default process* N and the *cumulative loss process* L by $N_t = \sum_{i=1}^n H_t^i$ and $L_t = (1-R)N_t$, respectively.

Suppose that $\xi = \pi(L_T)$ represents a (bounded, say) payoff at the maturity time T, representing a specific credit portfolio derivative claim. We assume zero interest-rates, for simplicity. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ a filtration that represents flow of information we use for pricing, and by $\mathbb{E}$ expectation relative to a risk-neutral pricing measure $\mathbb{Q}$ on a primary market of hedging instruments. The pricing task for the derivative with payoff ξ at time T thus amounts to computation of the conditional expectation $\mathbb{E}(\xi|\mathcal{F}_t)$.

29.1 CDS Index and CDO Tranches

We refer the reader to, for instance, Bielecki et al. [6, 30], for a review of real-life credit derivatives products, typically the *Credit Default Swap* (CDS) for single-name credit, and the *Collaterized Debt Obligation* (cash CDO or synthetic CDO tranches) for multi-name credit.

In few words, these are all swapped products with two legs, a default protection leg and a fee leg, and a notion of fair spread Σ_t at time t defined much as in the case of interest-rate swaps, so that the two legs of the contract would have equal values at time t, if the contractual spread was equal to Σ_t. Of course the contractual spread of the contract is fixed once for all at time 0 (starting time of the swap), at Σ_0, and at time t the contract has therefore a value which can be positive or negative depending on Σ_t.

29.1.1 Stylized Tranches

From the modeling point of view, the default protection leg is the more challenging, since it crucially depends on the model of dependence used for default times. In order to simplify the analysis, we shall consider stylized (protection legs of) CDS index contracts and CDO tranches, with payoff of the form

$$\xi = (\frac{L_T}{n} - K)^+ \wedge (\mathcal{K} - K)$$

at T, where the *attachment and detachment points* K and $\mathcal{K}$ are such that $0 \leq K \leq \mathcal{K} \leq 100\%$. In particular, we shall consider *equity tranches* (i.e. attachment $K = 0$), resp. *(super)senior tranches* (i.e. detachment $\mathcal{K} = 100\%$), with payoffs (recall $L_T = (1 - R)N_T$):

$$\pi^+(N_T) = \frac{L_T}{n} \wedge k \text{ , resp. } \pi^-(N_T) = (\frac{L_T}{n} - k)^+$$

at T, where the 'strike' (detachment, resp. attachment point) k belongs to $[0, 1]$. In this formalism the stylized *credit index* corresponds to the stylized equity tranche with $k = 100\%$ (or senior tranche with $k = 0$), namely to the payoff

$$p(N_T) = \frac{L_T}{n} = (\pi^+ + \pi^-)(N_T) \ .$$

With a slight abuse of terminology, we shall refer to our stylized loss derivatives as index and (equity and senior) tranches, assuming henceforth in the case of the equity and senior tranches that $0 < k < 1 - R$.

Given a pricing filtered probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, we also introduce the (cum-dividend) *equity/senior tranche and index price processes* $\Pi^\pm_t$ and P_t, and the (stylized) *bp spreads* $\Sigma^\pm_t$

and S_t, respectively defined by, for $t \in [0, T]$:

$$\Pi_t^+ = \mathbb{E}\Big(\pi^+(N_T)|\mathcal{F}_t\Big) \; , \; \Pi_t^- = \; \mathbb{E}\Big(\pi^-(N_T)|\mathcal{F}_t\Big) \; , \; P_t = \mathbb{E}\Big(p(N_T)|\mathcal{F}_t\Big)$$

$$\Sigma_t^+ = \frac{10^4\mathbb{E}\Big(\pi^+(N_T)-\pi^+(N_t)|\mathcal{F}_t\Big)}{(\pi^+(n)-\pi^+(N_t))(T-t)} \; , \; \Sigma_t^- = \frac{10^4\mathbb{E}\Big(\pi^-(N_T)-\pi^-(N_t)|\mathcal{F}_t\Big)}{(\pi^-(n)-\pi^-(N_t))(T-t)} \; , \; S_t = \frac{10^4\mathbb{E}\Big(p(N_T)-p(N_t)|\mathcal{F}_t\Big)}{(p(n)-p(N_t))(T-t)} \quad (90)$$

29.2 Li Model and implied Correlations

The one factor Gaussian copula model, *or Li model* (see Li [100] or Laurent [98]), is the financial industry quotation standard for multi-name credit derivatives. Without entering into details, let us only mention that at the current time t, the Li model parameters are t, a correlation parameter $\rho \in [0, 1]$, and a family $F = (F^i)_{1 \le i \le n}$ of marginal time-to-default cumulative distribution functions over $[t, +\infty)$.

As the Black(–Scholes) formula on volatility markets, the Li model is usually used in the reverse-engineering mode for quoting CDO tranches in terms of their *Li implied correlations*, given an instrumental instance F_t of F (typically inferred in some way from the related marginal CDS spread curves at time t).
More precisely, at time t, denoting by $\Sigma^{li}(T, K, \mathcal{K}; t, F_t, \rho)$ the fair spread of the $(T, K, \mathcal{K})$-tranche in the Li model with parameters F_t and ρ, and by $\Sigma_t^{ma}(T, K, \mathcal{K})$ the market spread of the tranche:
• The *Li compound implied correlation* of the tranche is the value of the correlation $\widetilde{\rho}_t$ in a Li model such that

$$\Sigma^{li}(T, K, \mathcal{K}; t, F_t, \widetilde{\rho}_t) = \Sigma_t^{ma}(T, K, \mathcal{K}) \, ; \qquad (91)$$

• The *Li base implied correlation* of the tranche is the value of the correlation ρ_t in a Li model such that

$$\Sigma^{li}(T, 0, \mathcal{K}; t, F_t, \rho_t) = \Sigma_t^{ma}(T, 0, \mathcal{K}) \, , \qquad (92)$$

where $\Sigma_t^{ma}(T, 0, \mathcal{K})$ denotes a synthetic market spread computed from the observed market spreads for the tranches with detachment point $\le \mathcal{K}$ (see, e.g., [106]).
Base implied correlation is more stable numerically than compound implied correlation, because $\Sigma_t^{li}(T, K, \mathcal{K}; t, F_t, \rho)$ is monotone (decreasing) with respect to ρ for $K = 0$, but not for $K > 0$ [106].

Much like the Black–Scholes implied volatility surface on volatility derivatives markets, the market *Li implied compound correlation surface* is then defined as the surface of market Li compound implied correlations obtained as $(K, \mathcal{K})$ varies over a standardized set of successive intervals (like $(0\%, 3\%)$, $(3\%, 6\%)$, $(6\%, 9\%)$, $(9\%, 12\%)$ and $(12\%, 22\%)$ on the DJ iTraxx market, a family of CDS indices for Europe and Asia) and T varies over the set of maturities with quoted CDO tranches (e.g., $1yr, 2yr, 3yr, 5yr, 7yr, 10yr$).
Likewise, the market *Li implied base correlation surface* is defined as the surface of market Li base implied correlations obtained as $\mathcal{K}$ varies over a set of standardized tranches detachment points, and T varies over the set of maturities with quoted CDO tranches.

At fixed T, the market Li implied correlation is typically *convex* with respect to the 'strike' variable in the case of the *compound* correlation (yielding the so-called *compound correlation smile*), and increasing with respect to the 'strike' variable in the case of the *base* correlation

(yielding the so called *base correlation skew*).
Credit models are then assessed on their ability to reproduce the market (base or compound) implied correlation surface, for suitably calibrated values of their parameters: See, e.g., Figure 2 in section 30.1 below, in which a market implied compound correlation smile is compared to a calibrated implied compound correlation smile in the credit migrations model of [6] (cf. section 30).

29.2.1 Stylized Li Model

In the context of stylized tranches and index as defined in section 29.1.1, we shall consider in section 33 a suitably specified version of the Li model in which the F_t – Li model parameter at time t is determined by assuming that all the marginal CDS spread curves at time t are constant and equal to the index spread S_t. So

$$\Pi^{li}(T, K, \mathcal{K}; t, F_t, \rho) = \widetilde{\Pi}^{li}(T, K, \mathcal{K}; t, S_t, \rho) \,, \tag{93}$$

for a *reduced Li pricing function* $\widetilde{\Pi}^{li}$.

For stylized tranches and with this specification of the Li model, the equation defining the market Li implied (base, say) correlation tranche at time t writes, equivalently to (92):

$$\widetilde{\Pi}^{li}(T, 0, \mathcal{K}; t, S_t, \rho_t) = \Pi^{ma}_t(T, 0, \mathcal{K}) \,. \tag{94}$$

Note that in the case of the index, the Li price does not depend on the correlation parameter, so

$$P^{li}(T, 0, 100\%; t, F_t, \rho) = \widetilde{P}^{li}(T, 0, 100\%; t, S_t) \,, \tag{95}$$

and there is thus no well defined index Li implied correlation parameter.

30 Markovian Market Model

Modeling of dependent defaults and credit migrations was considered by many authors, who proposed several alternative approaches to this important issue. The detailed analysis of these methods is beyond the scope of this text. Let us simply mention a few of them:
• Modeling correlated defaults in a static framework using copulae (see, e.g., Laurent [98]),
• Factor approach (Duffie and Garleanu [57], Davis and Lo [49], Jarrow and Yu [87], Yu [116], Frey and Backhaus [70, 72]),
• Modeling the portfolio loss distribution in a top-down approach (Giesecke and Goldberg [74], Schönbucher [111]).

In [6], we propose a fairly general Markovian model that nests several of the above models. In particular, this model covers jump-diffusion dynamics and continuous time Markov chains. It allows for incorporating several credit names, and thus it is suitable when dealing with valuation and hedging of basket credit products (such as, basket credit default swaps or collateralized debt obligations), possibly in a multiple credit ratings environment.
This model also comprises as a special case the Homogeneous Groups Model already considered by several authors, which is exposed in some detail in section 30.2 in view of later use in this report.

For simplicity we shall limit ourself here to the description of a slightly alleviated version of the general model of [6]. Let as usual the pricing probability space be denoted by $(\Omega, \mathbb{F}, \mathbb{Q})$. We consider d obligors (or credit names) and we assume that the current credit quality of each reference entity can be classified into ν rating categories, numbered from 0 to $\nu - 1$. By convention, the category 0 corresponds to default. Let N^l, $l = 1, 2, \ldots, d$ be processes on $(\Omega, \mathbb{F}, \mathbb{Q})$ taking values in the finite state space $I = \{0, 1, 2, \ldots, \nu - 1\}$. The process N^l represents the evolution of credit ratings of the l^{th} reference entity. We define the *default time* τ_l of the l^{th} reference entity by setting

$$\tau_l = \inf\{\, t > 0 \,:\, N_t^l = 0\} \,.$$

We assume that the default state 0 is absorbing, so that for each name the default event can only occur once. We denote by $\mathcal{N} = (N^1, N^2, \ldots, N^d)$ the joint credit rating process of the portfolio. The state space of N is $\mathcal{I} := I^d$. We also consider a jump-diffusion vector-process, X, representing the evolution of relevant economic variables, like short rate or equity price processes. Finally we assume that, given $X_t = x$, the intensity matrix of $\mathcal{N}$ is given by $(\lambda^{\imath,\jmath}(x))_{(\imath,\jmath)\in\mathcal{I}^2}$, so that the process $F = (X, \mathcal{N})$ is jointly Markov under $\mathbb{Q}$.

We recognize in the model F a special case of the jump-diffusion setting with regimes (with $\mathcal{N}$ and λ here in the role of N and n therein) of section 25 (see also the comments of section 28.1). ***Therefore all the results derived in sections 25 to 27 are applicable here.***

Moreover we impose a further structure on the intensity λ, namely, $\lambda^{\imath,\jmath}(x) = 0$ if the vectors of ratings $\imath$ and $\jmath$ differ by more than one component. In other words, the ratings of different credit names may not change simultaneously. The advantage of this restriction is that, for the purpose of simulating the next jump time and state of $\mathcal{N}$ conditional on $(X_t, \mathcal{N}_t) = (x, \imath)$, it is enough, rather than dealing with the ν^d-dimensional intensity measure $(\lambda^{\imath,\jmath}(x))_{\jmath\in\mathcal{I}}$, to deal with d intensity measures ('which obligor jumps'), each of dimension ν ('where it jumps').

Within the present set-up, the current credit rating of one credit name directly impacts the intensity of transition of the rating of another credit name. This property, known as *frailty*, may contribute to default contagion.

30.1 Computational issues

Since, in case of basket credit derivatives, we typically do not have access to closed-form pricing formulae in such model, pricing and calibration in this model need to be done by simulation.

Figure 2 shows that the model performs very well in regard to the calibration issue. Since the calibration is done by simulation, it may take some time however, like up to a few minutes for calibrating a whole (one maturity-)set of individual CDSs and CDO tranches, using $m = 10^5$ simulated model trajectories and a well-chosen specification of the intensities involving seven parameters.

For more details about the simulation, calibration and implementation issues and extensive reports on numerical experiments, we refer the reader to Bielecki et al. [6, 30]. See also section 32 and [7] for related variance reduction issues.

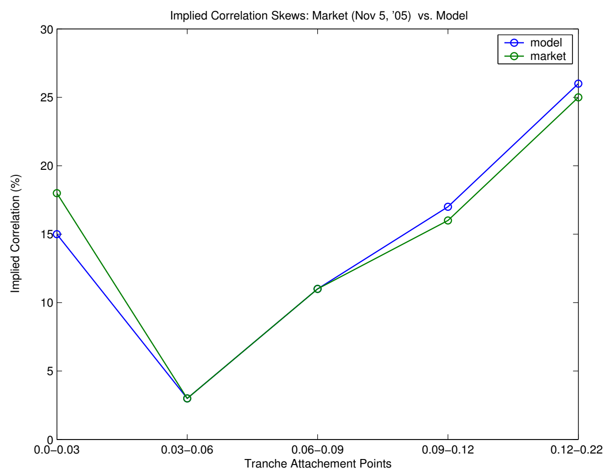

Figure 2: *Market and Model Li compound implied correlation smiles for CDO tranches of a fixed maturity.*

30.2 Homogeneous Groups Model

For later use, we now describe in some detail a more specific version of the previous model, considered for different purposes by various authors in [1, 7, 71, 83], among others. In this specification of the model, there is no factor process X involved. We thus deal with a continuous-time Markov Chain $\mathcal{N}$ and $\mathbb{F} = \mathbb{F}^{\mathcal{N}}$.

More precisely, a pool of n credit names is organized in d homogeneous groups of $(\nu - 1)$ obligors (so $n = (\nu - 1)d$, assuming $\frac{n}{d}$ integer), and N^l now represents the number of defaulted obligors in the l^{th} group (instead of representing the credit rating of obligor l previously; so the interpretation of l and N^l have changed, but the mathematical structure of the model is preserved). Moreover we assume that the N^l's can only jump one at a time *and by one*, so that we in fact deal with a d-variate Markov point process $\mathcal{N} = (N^1, \ldots, N^d)$. For each l, the ($\mathbb{F}$-)intensity of N^l is assumed to be of the form

$$\lambda^l(\mathcal{N}_t) = (\nu - 1 - N^l_t)\widetilde{\lambda}^l(\mathcal{N}_t) \ , \tag{96}$$

for an *aggregated intensity function* $\lambda^l = \lambda^l(\imath)$, and *pre-default individual* intensity function $\widetilde{\lambda}^l(\imath)$, where $\imath = (i_1, \ldots, i_d) \in \mathcal{I} = I^d$ (recall $I = \{0, 1, \ldots, \nu - 1\}$).
Since we assume that there are no common jumps between processes N^l, so the jump intensities λ^l are in one-to-one correspondence with the generator Λ of $\mathcal{N}$, which consists of a $\nu^d \otimes \nu^d$ matrix Λ (a very sparse matrix, since the components of $\mathcal{N}$ may only jump by one and only one at a time).

For $d = 1$, we recover the so called *Local Intensity Model* (pure birth process stopped at level n) of Laurent, Cousin and Fermanian [97], Cont and Minca [14] or Herbertsson [83], for the portfolio cumulative default process N. This model is, in a sense, the analog for credit derivatives of the local volatility model for equity and equity index derivatives, that we considered at length in part II: see sections 13 and 33 for the details and limits of this analogy.

At the other end of the spectrum, for $d = n$ (i.e. when each group has only a single element), we are in effect modeling the vector of default indicator processes $H = (H^i)_{1 \le i \le n}$ of the pool.

As d varies between 1 and n, we thus get a variety of models of credit risk, ranging from pure 'top-down' models for $d = 1$, to pure 'bottom-up' models for $d = n$ (see section 31).

30.2.1 Pricing in the Homogeneous Groups Model

Since $\mathcal{N}$ is a Markov process and N is a function of $\mathcal{N}$, the model price process of the stylized tranche with payoff $\pi(N_T)$ as of section 29.1.1 writes, for $t \in [0, T]$:

$$\Pi_t = \mathbb{E}(\pi(N_T)|\mathcal{F}_t) = u(t, \mathcal{N}_t) , \tag{97}$$

where $u(t, \imath)$ or $u_\imath(t)$ for $t \in [0, T]$ and $\imath \in \mathcal{I} = I^d$, is the *pricing function* (system of time-functionals $u_\imath$), solution to the following *pricing equation* (system of ODEs) with generator Λ:

$$(\partial_t + \Lambda)u = 0 \text{ on } [0, T) , \tag{98}$$

with terminal condition $u_\imath(T) = \pi(\imath)$, for $\imath \in \mathcal{I}$.
Likewise, the groups losses distribution at time t, that is, $q_\imath(t) = \mathbb{Q}(\mathcal{N}_t = \imath)$ for $t \in [0, T]$ and $\imath \in \mathcal{I}$, can be characterized in terms of the associated forward Kolmogorov equations (see, e.g., [7]).
These pricing and transition probability backward and forward Kolmogorov equations can then be solved by various means, like numerical matrix exponentiation (since the model is time-homogeneous).

However, even if the matrix Λ is very sparse, its size is prohibitive in most cases as far as deterministic numerical methods are concerned. For instance, in the case of $d = 5$ groups of $\nu - 1 = 25$ names, one gets $\nu^{2d} = 26^{10}$. So for high values of d, Monte Carlo methods appear to be the only viable computational alternative.
As will be seen in section 32 (cf. Carmona and Crépey [7]), these Monte Carlo methods can be made quite efficient by application of suitable importance sampling techniques.

Remark 30.1 Observe that in the *fully homogeneous case* where $\widetilde{\lambda}^l(\imath) = \widehat{\lambda}(\sum_{1 \le \ell \le d} i_\ell)$ for some function $\widehat{\lambda} = \widehat{\lambda}(i)$ (independent of l), the model (whatever the nominal value of d / structure of the matrix generator used for encoding the model) effectively reduces to a local intensity model (with $d = 1$ and pre-default individual intensity $\widehat{\lambda}(i)$ therein).
The case of a constant $\widehat{\lambda}$ corresponds to the situation of homogeneous and independent obligors.
In general, introducing parsimonious parameterizations of the intensities allows one to account for inhomogeneity between groups, and/or for defaults contagion.

31 Up and Down Credit Risk

Various approaches to valuation and hedging of derivatives written on credit portfolios differ between themselves depending on what is the content of the model filtration $\mathbb{F}$. Thus, loosely speaking, these approaches differ between themselves depending on what they take to be sufficient information so to price (and consequently to hedge) credit portfolio derivatives.

The approach that we dub the *top* approach takes as $\mathbb{F}$ the filtration generated by the loss process and, possibly also, by an additional relevant (low dimensional) factor process, say Y. Thus, in this case, $\mathbb{F} = \mathbb{F}^L \vee \mathbb{F}^Y$. Examples are papers by Laurent, Cousin and Fermanian [97], Cont and Minca [14] (these are 'pure top' papers, $\mathbb{F} = \mathbb{F}^L$) or Schönbucher [111] and Sidenius, Piterbarg and Andersen [113].

The so-called *top-down* approach starts from *top*, that is, it starts with modeling of evolution of the portfolio loss process subject to information structure $\mathbb{F}$. Then, it attempts to "decompose" the dynamics of the portfolio loss process *down* on the individual constituent names of the portfolio. This is done by a method of random thinning formalized in Giesecke and Goldberg [74].

The approach that we dub the *bottom-up* approach takes as $\mathbb{F}$ the filtration generated by process H and, possibly also, by a further factor process Z. Thus, in this case, $\mathbb{F} = \mathbb{F}^H \vee \mathbb{F}^Z$. Examples are Bielecki, Crépey, Jeanblanc and Rutkowski [6] (section 30 above), Bielecki, Vidozzi and Vidozzi [31], Frey and Backhaus [70, 71] or Duffie and Garleanu [57].

In [1] we provide some insights to the fact that information (namely, the choice of a relevant model filtration) is the major issue for handling portfolio credit derivatives. This is done by the means of mathematical demonstration and of numerical simulations.

Let thus τ denote an $\widehat{\mathbb{F}}$-stopping time where $\widehat{\mathbb{F}} \subseteq \mathbb{F}$. Let Λ and $\widehat{\Lambda}$ denote the $\mathbb{F}$-compensator and the $\widehat{\mathbb{F}}$-compensator of τ, respectively. We have the following result, which establishes the relation between Λ and $\widehat{\Lambda}$ (and the related $\mathbb{F}$- and $\widehat{\mathbb{F}}$- intensity processes λ and $\widehat{\lambda}$, whenever they exist). We denote by ${}^o.$ and $.^p$ the optional and dual predictable projections on the sub-filtration $\widehat{\mathbb{F}}$ (see, e.g., Dellacherie and Meyer [52]).

Proposition 31.1 (See [1]) *$\widehat{\Lambda}$ is the dual predictable projection of Λ on $\widehat{\mathbb{F}}$, so*

$$\widehat{\Lambda} = \Lambda^p \, . \tag{99}$$

Moreover, in case $\widehat{\Lambda}$ and Λ are time-differentiable with related $\widehat{\mathbb{F}}$- and $\mathbb{F}$- intensity processes $\widehat{\lambda}$ and λ, then $\widehat{\lambda}$ is the optional projection of λ on $\widehat{\mathbb{F}}$, so

$$\widehat{\lambda} = {}^o\lambda \, . \tag{100}$$

We refer the reader to [1] for other mathematical aspects of the paper.

In the sequel of this section we shall rather focus on a rather striking numerical illustration, developed in [1], of the fact that, even for basket credit derivatives which can be considered as derivatives on the (non-traded) loss process L in the sense that their payoff processes are given as functions of L, this loss process L is not a sufficient statistic for pricing and hedging them. This negative result regarding the top approaches is therefore an argument in favor of bottom-up approaches.

31.1 Hedging Issues in the Homogeneous Groups Model

We consider the benchmark problem of hedging CDO tranches with the related CDS index in the Homogeneous Groups Model of section 30.2. For simplicity we work with *stylized* CDS index and equity and senior CDO tranches as defined in section 29.1.1.

Note that in this model, it is possible to replicate, dynamically in continuous time, any terminal payoff at T, provided d non-redundant hedging instruments are available (see Bielecki, Vidozzi and Vidozzi [31] or Frey and Backhaus [70]; see also Laurent, Cousin and Fermanian [97] for results in the special case where $d = 1$). From the mathematical side this corresponds to the fact that the model is of (Davis-Varaiya) *multiplicity* d [50], in general. So, in general, it is not possible to replicate a payoff, such as tranche, *by the index alone* in this model, unless the model dimension d is equal to one (or reducible to one, cf. Remark 30.1). Now our point is that this potential lack of replicability is not purely speculative, but can be very significant in practice.

Since delta-hedging in continuous time is expensive in terms of transaction costs, and because main changes occur at default times in this model (in fact, default times are the only events in this model, if not for time flow and the induced time-decay effects), we shall focus on *semi-static hedging* in what follows, only updating at default times the composition of the hedging portfolio. More specifically, denoting by t_1 the first default time of a reference obligor, we shall examine the result at t_1 of a static hedging strategy on the random time interval $[0, t_1]$.
Let Π and P denote the tranche and index model price processes, respectively. Using a constant hedge ratio $\widehat{\delta}_0$ over the time interval $[0, t_1]$, the *tracking error* or *profit-and-loss* of the delta-hedged tranche at t_1 writes:

$$e_{t_1} = (\Pi_{t_1} - \Pi_0) - \widehat{\delta}_0 (P_{t_1} - P_0) \ . \tag{101}$$

The question we want to consider is whether it is possible to make this quantity 'small', in terms, say, of (risk-neutral) variance, relative to the variance of $\Pi_{t_1} - \Pi_0$ (which corresponds to the risk without hedging), by a suitable choice of $\widehat{\delta}_0$. It is expected that this should depend:
• First, on the characteristics of the tranche, and in particular on the value of the strike k: A high strike equity tranche or low strike senior tranche (*in-the-money* tranche) is quite close to the index in terms of cashflows, and should therefore exhibit a higher degree of correlation and be easier to hedge with the index, than a low strike equity tranche or high strike senior tranche (*out-of-the money* tranche);
• Second, on the 'degree of Markovianity' of the loss process L, which in the case of the homogeneous groups model depends both on the model nominal dimension d and on the specification of the intensities (see, e.g., Remark 30.1).

Moreover, it is intuitively clear that for too large values of t_1 time-decay effects matter and the hedge should be rebalanced at some intermediate points of the time interval $[0, t_1]$ (even though no default occurred yet). To keep it as simple as possible we shall merely apply a cutoff and restrict our attention to the random set $\{\omega : t_1(\omega) < T_1\}$ for some fixed $T_1 \in [0, T]$.

31.2 Numerical Results

We work with the above model for $d = 2$ and $\nu = 5$. We thus consider a two-dimensional model of a stylized credit portfolio of $n = 8$ obligors. The model generator is a $\nu^d \otimes \nu^d$ – (sparse) matrix with $\nu^{2d} = 5^4 = 625$. Recall that the computation time for exact pricing (using matrix exponentiation) in such model grows as ν^{2d}, which motivated the previous modest choices for d and ν.

Moreover we take the $\widetilde{\lambda}^l$'s given by, (cf. (96)):

$$\widetilde{\lambda}^1(t,\imath) = \frac{2(1+i_l)}{9n} \ , \ \widetilde{\lambda}^2(t,\imath) = \frac{16(1+i_l)}{9n} \ . \quad (102)$$

So in this case (which is an admittedly extreme case of inhomogeneity between two independent groups of obligors), the individual intensities of the obligors of group 1 and 2 are given as $\frac{1+i_1}{36}$ and $\frac{8(1+i_2)}{36}$, where i_1 and i_2 represent the number of currently defaulted obligors in groups 1 and 2, respectively.
For instance, at time 0 with $\mathcal{N}_0 = (0,0)$, the individual intensities of obligors of group 1 and 2 are equal to $1/36$ and $8/36$, respectively; the average individual intensity at time 0 is thus equal to $1/8 = 0.125 = 1/n$.

Model Simulation In this toy model the simulation takes the following very simple form (see [70] or [6] for more details in more general set-ups):
Compute Π_0 and P_0 by numerical matrix exponentiation (cf. section 30.2.1), and then for every $j = 1,\dots,m$:
• Draw a pair $(\widetilde{t}^j_1, \widehat{t}^j_1)$ of independent exponential random variables with parameter (cf. (96)–(102))

$$(\lambda^1_0, \lambda^2_0) = 4 \times (\frac{1}{36}, \frac{8}{36}) = (\frac{1}{9}, \frac{8}{9}) \ ; \quad (103)$$

• Set $t^j_1 = \min(\widetilde{t}^j_1, \widehat{t}^j_1)$ and $\mathcal{N}_{t^j_1} = (1,0)$ or $(0,1)$ depending on whether $t^j_1 = \widetilde{t}^j_1$ or $\widehat{t}^j_1$;
• Compute $\Pi_{t^j_1}$ and $P_{t^j_1}$ by numerical matrix exponentiation.

Doing this for $m = 10^4$, we got 9930 draws with $t_1 < T = 5yr$, among which 6299 ones with $t_1 < T_1 = 1yr$, subdividing themselves into 699 defaults in the first group of obligors and 5600 defaults in the second one.

31.2.1 Pricing

We consider two $T = 5yr$-tranches in the above model: an equity tranche with $k = 30\%$, corresponding to a payoff $\frac{(1-R)N_T}{n} \wedge k = (\frac{60N_T}{8} \wedge 30)\%$, and a senior tranche defined as the complement of the equity tranche to the index, thus with payoff $(\frac{(1-R)N_T}{n} - k)^+ = (\frac{60N_T}{8} - 30)^+\%$.

The portfolio loss distribution (computed by numerical matrix exponentiation) and the result of the pricing of the tranches and of the index at times 0 and t_1 (on the subset of the draws for which $t_1 < T$) are displayed in Figures 3, 4 and 5.

The left pane of Figure 3 represents the histogram of the loss distribution at the time horizon T; we indicate by a vertical line the loss level x beyond which the equity tranche is wiped out, and the senior tranche starts being hit (so $\frac{(1-R)x}{n} = k$, e.g. $x = 4$).
The right pane of Figure 3 displays the equity (labeled by $+$), senior ($\times$) and index ($\circ$) tranche prices at t_1 (in ordinate) versus t_1 (in abscissa), for all the points in the simulated data with $t_1 < 5$ (9930 points). Blue and red points correspond to defaults in the first ($\mathcal{N}_{t_1} = (1,0)$) and in the second ($\mathcal{N}_{t_1} = (0,1)$) group of obligors, respectively. We also represented in black the points $(0,\Pi_0)$ (for the tranches) and $(0,P_0)$ (for the index).

Note that there is virtually no error involved in the previous computations, in the sense that our simulation is exact (without simulation bias), and the prices and loss probabilities are computed by (quasi-exact) matrix exponentiation.

The left pane of Figure 3 represents the histogram of the loss distribution at the time horizon T; we indicate by a vertical line the loss level x beyond which the equity tranche is wiped out, and the senior tranche starts being hit (so $\frac{(1-R)x}{n} = k$, e.g. $x = 4$).
The right pane of Figure 3 displays the equity (labeled by $+$), senior ($\times$) and index ($\circ$) tranche prices at t_1 (in ordinate) versus t_1 (in abscissa), for all the points in the simulated data with $t_1 < 5$ (9930 points). Blue and red points correspond to defaults in the first ($\mathcal{N}_{t_1} = (1,0)$) and in the second ($\mathcal{N}_{t_1} = (0,1)$) group of obligors, respectively. We also represented in black the points $(0, \Pi_0)$ (for the tranches) and $(0, P_0)$ (for the index).

Note that in the case of the senior tranche and of the index, there is a clear difference between prices at t_1 depending on whether t_1 corresponds to a default in the first or in the second group of obligors, whereas in the case of the equity tranche there seems to be little difference in this regard.
In view of the portfolio loss distribution in the left pane, this can be explained by the fact that in the case of the equity tranche, the probability conditional on t_1 that the tranche will be wiped out at maturity is important unless t_1 is rather large. Therefore the equity tranche price at t_1 is close to $k = 30\%$ for t_1 close to 0. Moreover for t_1 close to T the intrinsic value of the tranche at t_1 constitutes the major part of the equity tranche price at t_1 (since the tranche has low time-value close to maturity). In conclusion the state of $\mathcal{N}$ at t_1 has a low impact on Π_{t_1}, unless t_1 is in the middle of the time-domain.
On the other hand, in the case of the senior tranche or in case of the index, the state of $\mathcal{N}$ at t_1 has a high impact on the corresponding price, unless t_1 is close to T (in which case intrinsic value effects are dominant). This explains the 'two-track' pictures seen for the senior tranche and for the index on the right pane of Figure 3, except close to T (whereas the two-tracks are superimposed close to 0 and T in the case of the equity tranche).

Looking at these results in terms of price changes $\Pi_0 - \Pi_{t_1}$ of a tranche versus the corresponding index price changes $P_0 - P_{t_1}$, we obtain the graphs of Figure 4 for the equity tranche and 5 for the senior tranche. We consider all points with $t_1 < T$ on the left panes and focus on the points with $t_1 < T_1$ on the right ones. We use the same blue/red color code as above, and we further highlight in green on the left panes the points with $t_1 < 1$, which are focused upon on the right panes.
Figure 4 gives a further graphical illustration of the low level of correlation between price changes of the equity tranche and of the index. Indeed the cloud of points on the right pane is obviously "far from a straight line", due to the partitioning of points between blue points / defaults in group one on one segment versus red points / defaults in group two on a different segment.
On the opposite (Figure 5), at least for t_1 not too far from 0 (right pane), there is an evidence of linear correlation between price changes of the senior tranche and of the index, since in this case the blue and the red segments are not far from being on a common line.

31.2.2 Hedging

We then computed the (empirical, risk-neutral) variance of $\Pi_{t_1} - \Pi_0$ and of the profit-and-loss e_{t_1} in (101) (restricting attention to the subset $t_1 < T_1 = 1$), using for $\widehat{\delta}_0$ the empirical

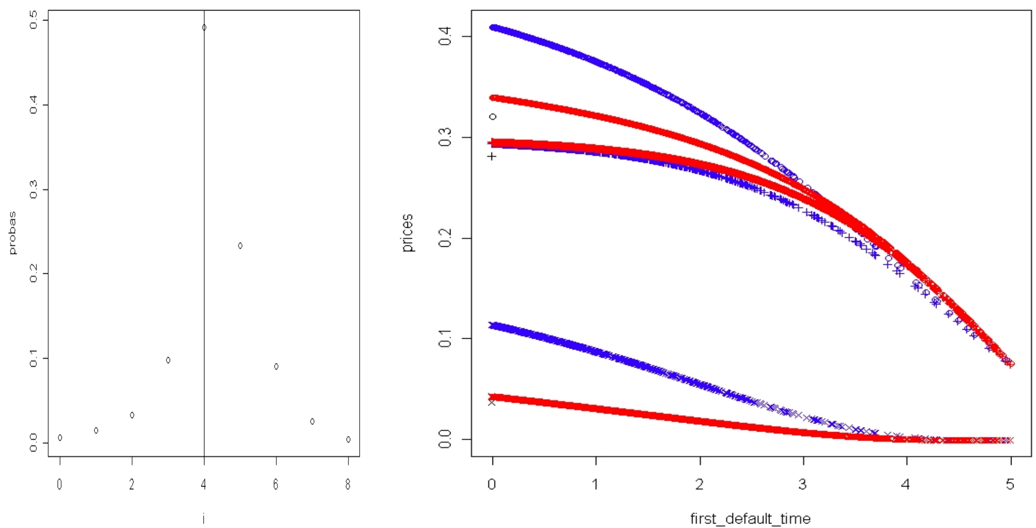

Figure 3: *(Left) Portfolio loss distribution at maturity $T = 5yr$; (Right) Tranche Prices at t_1 for $t_1 < T = 5$ (equity tranche (+), senior tranche ($\times$) and index ($\circ$)). On this and the following Figures, blue and red points correspond to defaults in the first and in the second group of obligors, respectively.*

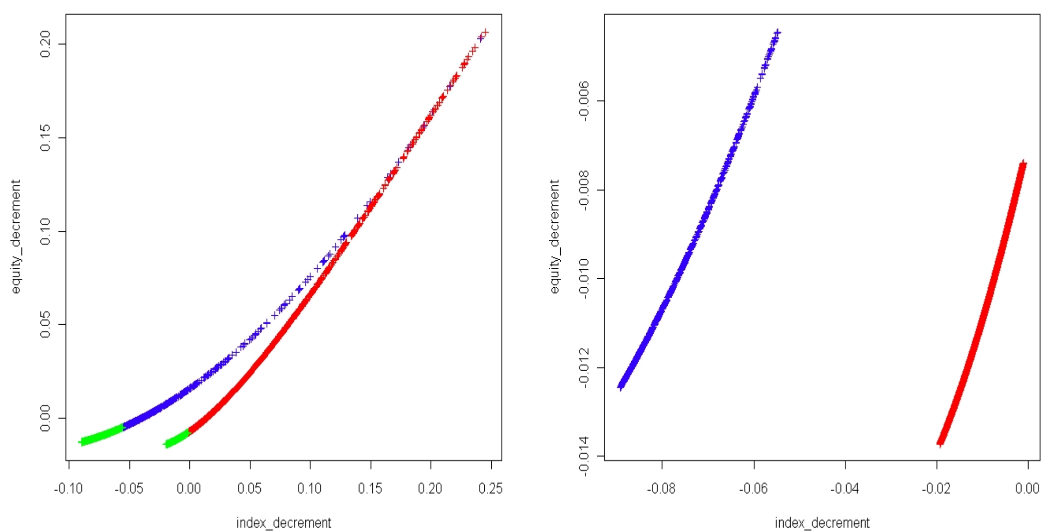

Figure 4: *Equity vs Index Price Changes between 0 and t_1 ($t_1 < T = 5$, left pane; zoom on $t_1 < T_1 = 1$, right pane).*

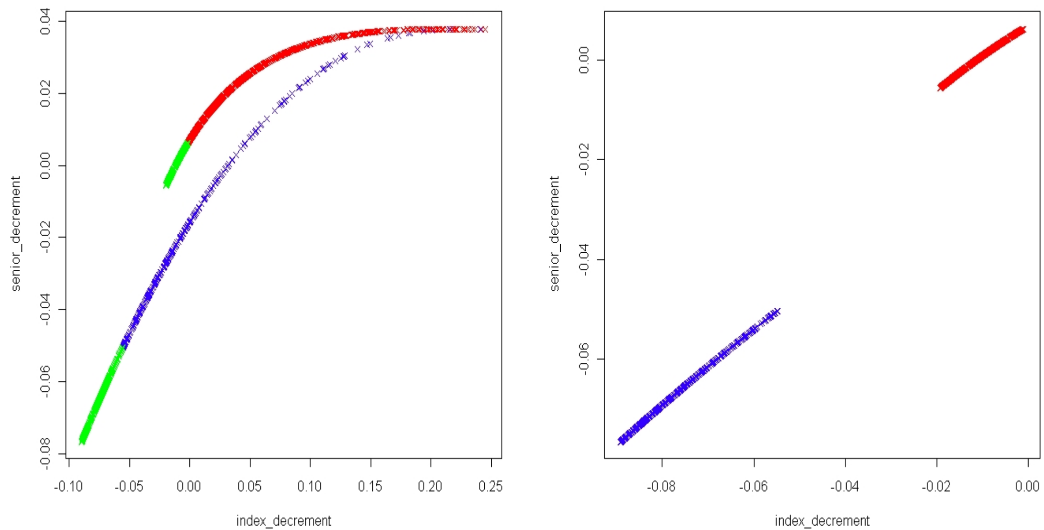

Figure 5: *Senior vs Index Price Changes between 0 and* t_1 *(*$t_1 < T = 5$*, last pane; zoom on* $t_1 < T_1 = 1$*, right pane).*

regression delta of the tranche with respect to the index at time 0, so

$$\widehat{\delta}_0 = \frac{\widehat{\mathbb{C}\mathrm{ov}}(\Pi_{t_1} - \Pi_0, P_{t_1} - P_0)}{\widehat{\mathbb{V}\mathrm{ar}}(P_{t_1} - P_0)} \,. \tag{104}$$

The results are displayed in Tables 3 and 4. Note that the prices and deltas of the equity and senior tranche of same strike k respectively sum up to P and to one, by construction. So the results for the senior tranche could in a sense be deduced from those for the equity tranche and conversely. However we present detailed results for the equity and senior tranche, for the reader's convenience.

In Table 3:
• $\Sigma_0 = \frac{10^4}{kT}\Pi_0$ or $\frac{10^4}{(1-R-k)T}\Pi_0$ (for the equity or senior tranche) or $S_0 = \frac{10^4}{(1-R)T}P_0$ (for the index), the stylized bp spreads defined in terms of the related prices by (90) for $t = 0$,
• δ_0^1, δ_0^2 and δ_0, the functions $\frac{\delta^1 u}{\delta^1 v}$, $\frac{\delta^2 u}{\delta^2 v}$ and the *continuous time min-variance delta function* (as is easily shown)

$$\frac{\lambda^1(\delta^1 u)(\delta^1 v) + \lambda^2(\delta^2 u)(\delta^2 v)}{\lambda^1(\delta^1 v)^2 + \lambda^2(\delta^2 v)^2} = \frac{\lambda^1(\delta^1 v)^2}{\lambda^1(\delta^1 v)^2 + \lambda^2(\delta^2 v)^2}(\frac{\delta^1 u}{\delta^1 v}) + \frac{\lambda^2(\delta^2 v)^2}{\lambda^1(\delta^1 v)^2 + \lambda^2(\delta^2 v)^2}(\frac{\delta^2 u}{\delta^2 v})$$

evaluated at $t = 0$ and $\imath = \mathcal{N}_{0-} = (0, 0)$, so

$$\delta_0^1 = \tfrac{u_{1,0} - u_{0,0}}{v_{1,0} - v_{0,0}}(0) \ , \ \delta_0^2 = \tfrac{u_{0,1} - u_{0,0}}{v_{0,1} - v_{0,0}}(0) \tag{105}$$

$$\delta_0 = \tfrac{\lambda_0^1(u_{1,0} - u_{0,0})(v_{1,0} - v_{0,0}) + \lambda_0^2(u_{0,1} - u_{0,0})(v_{0,1} - v_{0,0})}{\lambda_0^1(v_{1,0} - v_{0,0})^2 + \lambda_0^2(v_{0,1} - v_{0,0})^2} \tag{106}$$

where we recall from (103) that $(\lambda_0^1, \lambda_0^2) = (\frac{1}{9}, \frac{8}{9})$.
The three deltas δ_0^1, δ_0^2 and δ_0 were thus computed by matrix exponentiation for the various terms $u, v_\imath(0)$ involved in formulas (105), (106).

Remark 31.1 The instantaneous min-variance delta δ_0 (which is a suitably weighted average of δ_0^1 and δ_0^2) can be considered as a measure of the *moneyness* of a tranche: *out-of-the-money* low strike equity tranche or high strike senior tranche with δ_0 less than 0.5, versus *in-the-money* tranche high strike equity tranche or low strike senior tranche with δ_0 greater than 0.5. The further out-of-the money a tranche and/or 'the less Markovian' a porfolio loss process L, the poorer the hedge by the index (cf. end of section 31.2.2).

	Π_0 or P_0	Σ_0 or S_0	δ_0^1	δ_0^2	δ_0
Eq	0.2821814	1881.209	0.1396623	0.7157741	0.2951399
Sen	0.03817907	254.5271	0.8603377	0.2842259	0.7048601

Table 3: *Time $t = 0$ – Prices, Spreads and Instantaneous Deltas in the Semi-Homogeneous Model.*

In Table 4 (cf. also (104)):
• ρ in column two is the empirical correlation of the tranche price increments $\Pi_{t_1} - \Pi_0$ versus the index price increments $P_{t_1} - P_0$,
• $R2 = \rho^2$ in column three is the *coefficient of determination* of the regression,
• Dev in column 4 stands for $\widehat{\mathrm{Stdev}}(\Pi_{t_1} - \Pi_0)/\Pi_0$,
• The *hedging variance reduction factor* RedVar $= \frac{\widehat{\mathbb{V}\mathrm{ar}}(\Pi_{t_1}-\Pi_0)}{\widehat{\mathbb{V}\mathrm{ar}}(e_{t_1})}$ in the last column is equal to $\frac{1}{1-\rho^2}$.

Remark 31.2 It is expected that $\widehat{\delta}_0$ should converge to δ_0 in the limit where the cutoff T_1 would tend to zero, provided the number of simulations m jointly goes to infinity. For $T_1 = 1yr$ and $m = 10^4$ simulations however, we shall see below that there is a clear discrepancy between δ_0 and $\widehat{\delta}_0$, and all the more so that we are in a non-homogeneous model with low correlation between the tranche and index price changes between times 0 and t_1. The reason is that the coefficient of determination ($R2$) of the linear regression with slope $\widehat{\delta}_0$ is given by $R2 = \rho^2$. In case ρ is small, $R2$ is even smaller, and the significance of the estimator (for low T_1's) $\widehat{\delta}_0$ of δ_0 is low too. In other words, in case ρ is small, we recover mainly noise through $\widehat{\delta}_0$; this however does not weaken our statements below regarding the ability or not to hedge the tranche by the index, since the variance reduction factor RedVar $= \frac{\widehat{\mathbb{V}\mathrm{ar}}(\Pi_{t_1}-\Pi_0)}{\widehat{\mathbb{V}\mathrm{ar}}(e_{t_1})}$ is equal to $\frac{1}{1-\rho^2}$, which only depends on ρ and not on $\widehat{\delta}_0$.

	$\widehat{\delta}_0$	ρ	$R2$	Dev	RedVar
Eq	-0.00275974	-0.03099014	0.0009603885	0.006612626	**1.000961**
Sen	1.002760	0.9960836	0.9921825	0.07475331	**127.9176**

Table 4: *Hedging Tranches by the Index in the Semi-Homogeneous Model.*

Recall that qualitatively the senior tranche's dynamics is rather close to that of the index (at least for t_1 close to 0, see Section 31.2.1, right pane of Figure 5). Accordingly, we find that hedging the senior tranche with the index is possible (variance reduction factor of about 128 in bold blue in the last column). This case thus seems to be supportive of the claim according to which one could use the index for hedging a loss derivative, even in a non Markovian model of portfolio loss process L.

But in the case of the equity tranche we get the opposite message: the index is useless for hedging the equity tranche (variance reduction factor essentially equal to 1 in bold red in the table, so *no variance reduction* in this case).

Incidentally this also means that hedging the senior tranche by the equity tranche, or vice versa, is not possible either.

We conclude that in general, at least for certain ranges of the model parameters and tranche characteristics (strong to strong non-Markovianity of L and out-of-the money to far out-of-the money tranche), hedging tranches with the index may not be possible in a non Markovian model of portfolio loss process L.

31.2.3 Fully Homogeneous Case

For confirmation of the previous analysis and interpretation of the results, we redid the computations using the same values as before for all the model, products and simulation parameters, except for the fact that the following pre-default individual intensities were used, for $l = 1, 2$:

$$\widetilde{\lambda}^l(\imath) = \frac{1}{n} + \frac{\sum_{1\leq\ell\leq d} i_\ell}{nd} =: \widehat{\lambda}(\sum_{1\leq\ell\leq d} i_\ell) \ . \tag{107}$$

For instance, at time 0 with $\mathcal{N}_0 = 0$, the individual intensities of the obligors are all equal to $1/8 = 0.125 = 1/n$.

We are thus in a case of homogeneous obligors, reducible to a local intensity model (with $d = 1$ and pre-default individual intensity $\widehat{\lambda}(i)$ therein, see Remark 30.1). So in this case we expect that hedging tranches by the index should work, including in the case of the out-of-the-money equity tranche.

This is what happens numerically. This time the red and blue curves are superimposed on the analogs of Figures 3, 4 and 5 (not reproduced here, see [1]). This is consistent with the fact that the identity of a defaulted name has no bearing in this case, given the present specification of the intensities.

Looking at Table 6, we find as before that hedging the senior tranche with the index works very well (still better than before, variance reduction factor of 11645 in bold blue in the last column; yet this may partly due to a moneyness effect: the senior tranche is further in-the-money than before, with an senior tranche δ_0 of about 0.7 in Table 3 versus 0.8 in Table5). But as opposed to the situation in the semi-homogeneous case, hedging the equity tranche with the index also works very well (variance reduction factor of about 123 in bold purple in the last column), and this holds even though the equity tranche is further out-of-the-money now than it was before, with an equity tranche δ_0 of about 0.3 in Table 3 versus 0.2 in Table 5 (cf. Remark 31.1). This also means that hedging the equity tranche by the senior tranche, or vice versa, is quite effective in this case.

These results support our previous analysis that the impossibility of hedging the equity tranche by the index in the semi-hompogeneous model was due to the non-Markovianity of the loss process L.

Note incidentally that $\widehat{\delta}_0$ and δ_0 are closer now (in Tables 5–6) than they were previously (in Tables 3–4). This is consistent with the fact that $R2$ is now larger than before ($\widehat{\delta}_0$ and δ_0 would be even closer if the cutoff T_1 was less than $1yr$, provided of course the number of simulations m is large enough; see Remark 31.2).

	Π_0 or P_0	Σ_0 or S_0	δ_0^1	δ_0^2	δ_0
Eq	0.2850154	1900.103	0.2011043	0.2011043	0.2011043
Sen	0.1587075	1058.050	0.7988957	0.7988957	0.7988957

Table 5: *Time $t = 0$ – Prices, Spreads and Instantaneous Deltas in the Fully-Homogeneous Model.*

	$\widehat{\delta}_0$	ρ	$R2$	Dev	RedVar
Eq	0.0929529	0.9959361	0.9918887	0.004754811	**123.2852**
Sen	0.9070471	0.999957	0.9999141	0.04621152	**11645.15**

Table 6: *Hedging Tranches by the Index in the Fully-Homogeneous Model.*

Conclusions

For credit derivatives with (stylized) payoff given as $\pi(L_T)$ at maturity time T, it is tempting to adopt a Black–Scholes like approach, modeling L as a Markov point process and performing factor hedging of one derivative by another, balancing the related sensitivities computed by a suitable Itô-Markov formula (like, for instance, (73)). However, since the loss process L is far from being Markovian in the market (unless maybe additional factors are considered to form a Markovian vector state-process), this loss process is not a sufficient statistics for the purpose of valuation and hedging of portfolio credit risk. In other words, ignoring the potentially non-Markovian dynamics of L for pricing and/or hedging may cause huge model risk, even though the payoffs of the products at hand are given as functions of L.

This effectively means that the prices of credit derivatives depend on factors others than L, like the identity (and not only the number) of the defaulted names, ratings or implied ratings (and not only identities) of survivors, etc. This makes of course perfect sense since it is rather clear that the default of a major name in the index does not bear the same informational content as that of an arbitrary firm, and, moreover, pricing is done by agents with regard to the quality of the remaining names in the portfolio rather than with regard to the defaulted names.

Our conclusion is that the bottom-up approach is the best-suited for an adequate risk management of portfolio credit derivatives.
At this point one may raise the issue of the so called *curse of dimensionality* that is commonly associated with the bottom-up approaches. However, recent developments in the bottom up modeling enable one to efficiently cope with this curse of dimensionality. It is thus possible to specify high-dimensional ('bottom-up') dynamic Markovian models of portfolio credit risk *with automatically calibrated model marginals* (to the individual CDS curves, say), see Bielecki, Vidozzi and Vidozzi [31].
Much like in the standard static copula framework, this effectively reduces the main compu-

tational cost issue, that relative to model calibration, to calibration of the few *dependence parameters* in the model at hand. This calibration can thus be achieved in a very reasonable time, including by pure simulation procedures if need be (without using any closed pricing formulae, if there aren't any in the model under consideration). Appropriate reduction variance methods (see the next section and [7]) may help in this regard.

32 Importance Sampling for Markovian Credit Portfolios

In Markovian models of credit risk, the portfolio loss distribution can be computed by numerical resolution of the related forward Kolmogorov equations (see section 30.2.1). But a practical implementation of deterministic numerical schemes is precluded by the curse of dimensionality for models of dimension greater than a few units. Simulation approaches are then the only reasonable alternative. Yet simulation methods are typically slow, which makes it important to find efficient variance reduction techniques (see, e.g., Glasserman [75]).

Importance sampling (IS for short) is regarded as the method of choice when it comes to variance reduction. However, IS does not always make sense. Often times, it is not clear which change of measure will significantly reduce the variance, and moreover, producing Monte Carlo random samples after the change of measure can be impractical if at all possible. In this section, based on the paper [7] (joint work with René Carmona), we compare the results of IS to a more sophisticated method of Del Moral and Garnier [54] (see also Del Moral [53]). This method is based on the properties of twisted Feynman-Kac expectations and the approximation of their values by interacting particle systems (IPS for short, also called henceforth *implicit importance sampling*, as opposed to the previous *explicit importance sampling*).

32.1 Importance Sampling in a Nutshell

The problem at hand is the computation of small probabilities of events and related expectations of the form $\mathbb{E}f(X_n)$, relative to a (possibly time inhomogeneous) Markov chain $(X_i)_{1\leq i\leq n}$ with transition kernel $K(X_{i-1},\cdot)$ with respect to a filtered (in discrete time) probability space $(\Omega,\mathbb{F},\mathbb{Q})$, X_i being a random element taking values in a general measurable space $(E,\mathcal{E})$.

32.1.1 Explicit Importance Sampling for Makov Chains

Given weight functions $w_i = w_i(x_1,\ldots,x_i)$ such that $\mathbb{E}\left(w_i(X_1,\ldots,X_i)\,|\,\mathcal{F}_{i-1}\right) = 1$, let $\widetilde{\mathbb{Q}}$ stand for the *twisted probability measure* defined by the w_i's, in the sense that we have, for every $i\in\mathbb{N}$:

$$\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{Q}}\,|_{\mathcal{F}_i} = \prod_{1\leq l\leq i} w_l(X_1,\ldots,X_l)\,. \tag{108}$$

The basic importance sampling algorithm based on $\widetilde{\mathbb{Q}}$ runs as follows. Given the initial condition $\widetilde{\xi}_0 = 0$, draw for $i=1,\ldots,n$, for $j=1,\ldots,m$:

$$\widetilde{\xi}_i^j \rightsquigarrow \widetilde{K}_i(\widetilde{\xi}_{i-1}^j,\cdot) \tag{109}$$

where $\widetilde{K}_i$ stands for the transition kernel of X at time $i-1$ under the twisted probability measure $\widetilde{\mathbb{Q}}$. We then have the following unbiased (at fixed m), asymptotically convergent (as $m \to \infty$) estimate (see, e.g., Glasserman [75]):

$$\mathbb{E}f(X_n) = \widetilde{\mathbb{E}}\left(\frac{f(X_n)}{\prod_{1\le i\le n} w_i(X_1,\ldots,X_i)}\right) \approx \frac{1}{m}\sum_{j=1}^{m} \frac{f(\widetilde{\xi}^j_n)}{\prod_{1\le i\le n} w_i(\widetilde{\xi}^j_1,\ldots,\widetilde{\xi}^j_i)}$$

Moreover, in order to minimize the variance, one should use weight functions w_n such that $\prod_{1\le i\le n} w_i(X_1,\ldots,X_i)$ is proportional to $|f(X_n)|$, and for obvious practical reasons the choice of w should give rise to an algorithm that can be easily implemented. This is typically done by resorting to a suitable version of the Girsanov theorem.

32.1.2 Interacting Particles System Approach

Roughly speaking, the method proposed by Del Moral and Garnier [54] is based on the deformation of the Markov chain successive transitions by way of mutations and selections in order to force the chain into the rare events of interest. This strategy is reminiscent of classical importance sampling. However, the main difference is that while the Monte Carlo sample of an importance sampling computation are generated from the *twisted* distribution, the Monte Carlo samples used in an IPS Monte Carlo computation are generated under the original distribution of the chain. In other words, the knowledge of the distribution of the underlying Markov chain is not really necessary. All we need to have in order to implement the IPS Monte Carlo computations is a *black box* capable of generating Monte Carlo samples from the distribution of the chain.

The basic simulation algorithm runs as follows. We choose an integer m which we shall interpret as the number of particles. A particle at time i is an element

$$\xi^j_i = (\xi^j_{0,i}, \xi^j_{1,i}, \ldots, \xi^j_{i,i}) \in E^{i+1}$$

where the superscript j of the particle ranges from 1 to m. The particles system (indexed by j) starts from an initial condition $\xi^j_0 = x_0$, for every j. Given weight functions w_i, the system evolves between times $i-1$ and i for $i = 1,\ldots,n$ according to the following *selection/mutation dynamics*, for $j = 1,\ldots,m$ (see Table 7):

$$(\xi^j_{0,i}, \xi^j_{1,i}, \ldots, \xi^j_{i-1,i}) \rightsquigarrow \textstyle\sum_{l=1}^m w_{i-1}(\xi^l_{i-1})\delta_{\xi^l_{i-1}} \quad (110)$$

$$\xi^j_{i,i} \rightsquigarrow K(\xi^j_{i-1,i}, \cdot) \quad (111)$$

where $\delta_{\xi^l_{i-1}}$ denotes a Dirac mass at ξ^l_{i-1}.

$\xi^1_{0,0}$	$\xi^1_{0,1}, \xi^1_{1,1}$	$\cdots$	$\xi^1_{0,n}, \xi^1_{1,n}, \ldots, \xi^1_{n,n}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\xi^m_{0,0}$	$\xi^m_{0,1}, \xi^m_{1,1}$	$\cdots$	$\xi^m_{0,n}, \xi^m_{1,n}, \ldots, \xi^m_{n,n}$

Table 7: *Selection/Mutation Dynamics.*

We then have the following unbiased (at fixed m), asymptotically convergent (as $m \to \infty$) estimate (see Del Moral and Garnier [54], Del Moral [53]):

$$\mathbb{E}f(X_n) \approx \left(\frac{1}{m} \sum_{j=1}^{m} \frac{f(\xi^j_{n,n})}{\prod_{1 \le i < n} w_i(\xi^j_{1,n}, \ldots, \xi^j_{i,n})} \right) \left(\prod_{1 \le i < n} \frac{1}{m} \sum_{j=1}^{m} w_i(\xi^j_i) \right) \tag{112}$$

Moreover, in order to minimize the variance, one should use weight functions w_i favoring the occurrence of the rare event of interest (without involving too large normalizing constants). Finally, the choice of w should give rise to an algorithm that can be easily implemented.

32.2 Armageddon

In presence of strong contagion between obligors (as often priced by portfolio credit derivatives markets), the loss distribution has a very different structure than in the independent case.

To address this issue we consider an homogeneous Local Intensity Model of portfolio credit risk (see section 30.2) with $n = 125$ credit names and individual pre-default instantaneous intensity at time t equal to $\widetilde{\lambda}_t = a \exp(bN_t/n)$, for non-negative parameters a and b. For $a = 1/n$ and $b = 0$, we recover a model of independent obligors. Positive values of b allow one to account for defaults contagion.

Figure 6 gives plots of the loss distribution for $T = 5yr$ in two different sets of parameters. For the left pane we used the values $a = 0.01$ and $b = 0$. This corresponds to the case of independent obligors. For the right pane, we used the values $a = 0.01$ and $b = 13$ which correspond to a case of extreme contagion. These distributions were computed by matrix exponentiation of the one-dimensional model generator Λ ($126 \otimes 126$ matrix, see section 30.2.1). Note the different vertical scales.

In the case of independent obligors (left pane), the structure of the loss distribution is basically that of a Poisson distribution (truncated at the level n). The right-tail of the distribution goes exponentially fast to zero, which makes high levels of the loss extremely rare. The probability of the Armageddon (everyone defaulted in the portfolio) by the time of maturity $T = 5yr$, is equal to $1.044507e - 164$. Importance sampling (IS or IPS) methods are thus in this case a complete necessity for computing by simulation high-losses related quantities like the price of a super-senior CDO tranche (see section 29.1).

In the case of extreme contagion (right pane), we observe the so-called Armageddon effect: the default of all the obligors within a finite time horizon becomes an event with significant probability, $7.106e - 03$, so the order of one percent. Moreover there are no extremely rare levels of the loss any more. The less likely loss level is the level $i = 115$, with a loss probability of $1.108e - 06$. For such a model, importance sampling methods are not strictly needed, since a standard Monte Carlo method with 10^6 samples will basically do a good job at estimating the $5yr$ loss distribution with a reasonable accuracy over the whole range of the loss levels. Importance sampling methods may be used however for the purpose of *variance reduction*: estimating by simulation the $5yr$ loss distribution (or part of it) using, say, 10^4 samples instead of 10^6.

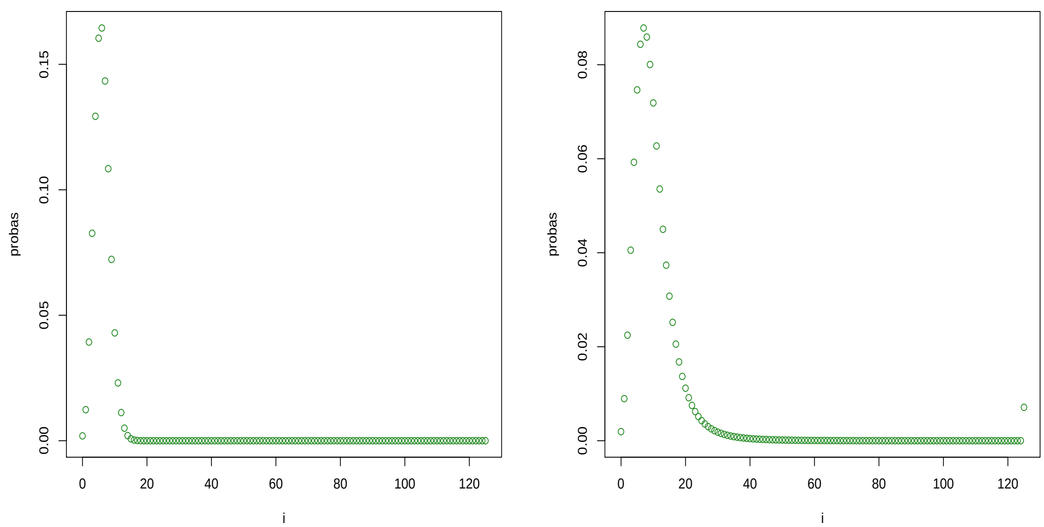

Figure 6: *$T = 5yr$ loss distributions. Left pane: Independent obligors ($a = 0.01, b = 0$); Right pane: Extreme contagion ($a = 0.01, b = 13$).*

32.2.1 Results for Importance Sampling

In the case of independent obligors ($a = 0.01, b = 0$), an (appropriate, see [7] for the detail) IS method succeeds very well in estimating the 5yr loss distribution on the whole range of loss levels (left pane of Figure 7). But, oddly enough, this is not true any more in the case of extreme contagion. We can see on the right pane of Figure 7 that the IS method is completely inefficient for estimating the probabilities of loss levels with probabilities less than, say, 10^{-3}, though $m = 10^4$ simulations were used in this experiment. The explanation of this negative result is that the weights involved in the change of measures become extreme, creating huge fluctuations and a large variance in this case, making the method essentially useless in practice.

32.2.2 Results for Interacting Particle Systems

As opposed to the IS method in section 32.2.1, an (appropriate [7]) IPS method does show some ability in capturing the events of probability 10^{-5} to 10^{-6} in the case of extreme contagion (right pane of Figure 8), though only $m = 10^4$ trajectories were used in this experiment.

In the independent case (left pane of Figure 8) the performances of the IPS algorithm are good for computing not too small probabilities (until the loss level $i = 36$ with exact probability $2.126281e - 18$, on this specific example). For higher levels of the loss, the related probabilities are too small and the generic IPS methodology is not sufficient to provide reasonable estimates, more specifically problem-dependent methodologies should be considered instead (see, e.g., Johansen, Del Moral and Doucet [88]).

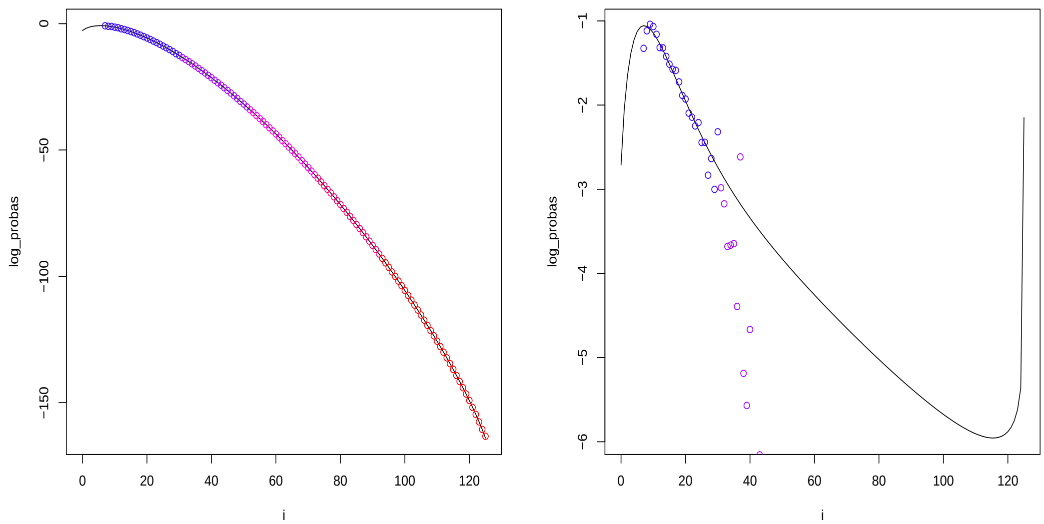

Figure 7: *Log-probabilities (exact versus simulated by IS with $m = 10^4$ draws) for independent obligors ($a = 0.01, b = 0$; left) or in the case of extreme contagion ($a = 0.01, b = 13$; right).*

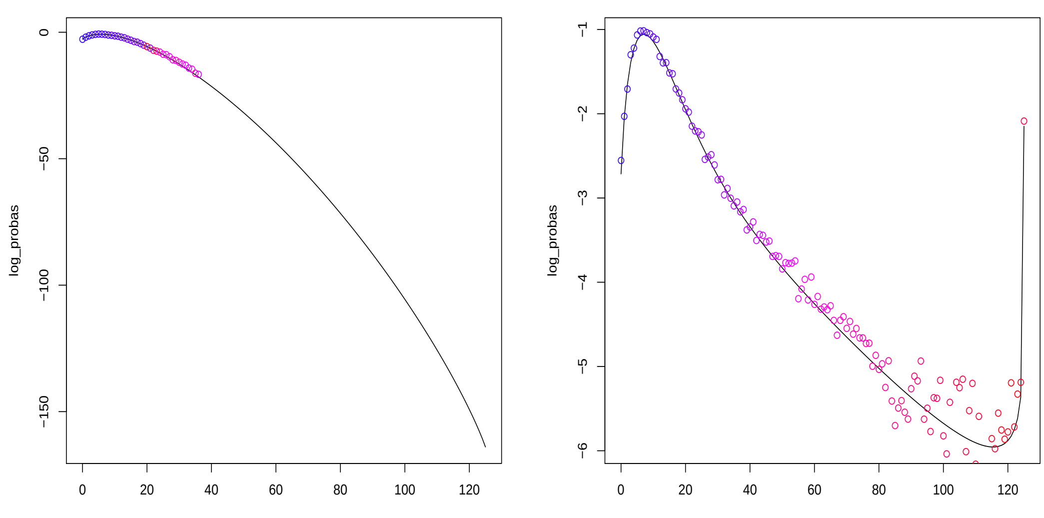

Figure 8: *Log-probabilities (exact versus simulated by IPS with $m = 10^4$ draws) for independent obligors ($a = 0.01, b = 0$; left) or in the case of extreme contagion ($a = 0.01, b = 13$; right).*

Conclusions

In summary, explicit importance sampling methods *can do wonders* when models for the loss distribution are simple enough for a Girsanov like transformation to be identified, a random generator for the distorted probability structure be available at the cost of a low overhead, and the corresponding densities be easily computed along the samples.

However, rare event probability estimation based on interacting particle systems can be a very useful substitute when no obvious Girsanov change of measure is available, or when the Monte Carlo simulations are based on a computer implementation in the form of a black box which cannot be open and modified for the purpose of importance sampling.

33 Delta-hedging Correlation Risk?

In conclusion of this report, the ***preliminary analysis*** of this section, which at this stage corresponds to no more but a ***research project***, aims at comparing two deltas with respect to the task of delta-hedging dynamically in discrete time a CDO tranche with the related credit index and the savings account, in a positively skewed base correlation market:
(i) The local delta Δ_t^{lo}, namely the delta of the tranche in a local intensity model calibrated to the market implied correlation smile at time t,
(ii) The Li implied delta Δ_t^{li}, namely the delta of the tranche in a Li model with the implied base correlation of the tranche at time t.
In these definitions:
• The *local intensity model* is the pure birth process described in section 30.2 for modeling a credit portfolio cumulative default process N,
• The *Li model* is the one factor Gaussian copula model of section 29.2, which has long been the industry standard for dealing with credit correlation products.

The question we want to address is which delta performs better, in the sense of maintaining, on average, P&L trajectory of a delta-hedged CDO tranche closer to 0 throughout the hedging period. Our motivation in this section is thus analogous to that of section 13 (see [10]), but relatively to credit correlation derivatives here, instead of volatility derivatives there.

For simplicity we take zero interest-rates, we consider stylized equity and senior CDO tranches and CDS index contracts of maturity T as defined in section 29.1.1, and we accordingly consider the version of the Li model as of section 29.2.1.

Let $T_1 \leq T$ denote the hedging horizon. Delta-hedging in discrete time the tranche (sold, say) with the index and the riskless asset over the time interval $[0, T_1]$, consists in rebalancing in a self-financed way, at every point in time of a subdivision (possibly random, though this is not the point here) $0 = t_0 \leq t_1 \leq \ldots \leq t_p = T_1$ of $[0, T_1]$, a complementary position Δ in the index, in order to minimize the overall exposure to 'small' moves of the index.

The *tracking error*, or *profit-and-loss* (P&L for short) trajectory $e = (e_{t_k})_{0 \leq k \leq p}$, is obtained by adding up the following P&L increments, starting with $e_0 = 0$, from $k = 0$ to $p - 1$:

$$\delta_k e = -\delta_k \Pi + \Delta_{t_k} \, \delta_k P \; , \tag{113}$$

where (cf. section 29.1.1):

• $\delta_k \Pi$ and $\delta_k P$ are the increments of the tranche and index values between times t_k and t_{k+1},
• Δ_{t_k} is the index delta (number of units of index contract in the hedging portfolio over the time interval $(t_k, t_{k+1}]$).

We want to compare the P&L trajectories e obtained using two strategies, with Δ_t and Δ_t given by:
• The *Li implied delta* of the option, that is

$$\Delta_t = \Delta_t^{li} = \frac{\partial_S \Pi^{li}(t, S_t, \rho_t)}{\partial_S P^{li}(t, S_t)}$$

where the functions $\Pi^{li}(t, S, \rho)$ and $P^{li}(t, S)$ stand for the Li pricing function of the tranche and of the index (the *reduced Li pricing functions* $\widetilde{\Pi}(t, S, \rho)$ and $\widetilde{P}^{li}(t, S)$, in the terminology and notation of section 29.2.1), and where the number ρ_t stands for the (stylized) Li implied correlation of the tranche at time t as of (94);
• Or, alternatively, the *local delta* of the tranche, that is

$$\Delta_t = \Delta_t^{lo} = \Delta_i^{lo}(t) = \frac{\Pi_{i+1}^{lo}(t) - \Pi_i^{lo}(t)}{P_{i+1}^{lo}(t) - P_i^{lo}(t)}$$

evaluated at $i = N_{t-}$, where $\Pi_i^{lo}(t)$ and $P_i^{lo}(t)$ with $t \in [0, T]$ and $i \in \{0, \ldots, n\}$ refer to the tranche and index pricing functions in the sense of section 30.2.1 (case where $d = 1$ therein), in a local intensity model calibrated to the full market Li implied base correlation surface at time t (see section 29.2).

We assume a steeply positively skewed base correlation market ('steep market' for short henceforth, or 'systemic market'; see section 29.2 for the definition of the base correlation), as typically observed in the market in recent years, corresponding to a high level of contagion and of positive correlation between the level of index spreads and that of (implied or realized) correlations. As will be explained in Remark 33.3, an analogous analysis can be performed in the case of 'even markets' with a flat or close to flat base correlation smile, or 'idiosyncratic markets', as was for instance the case during the correlation breakdown crisis of Spring 2005.

To develop our analysis we shall distinguish two stylized market regimes: *widening* and *tightening*, corresponding to values of the (stylized) index spread and index contract increasing and decreasing, respectively.

Remark 33.1 In the volatility analysis of section 13 we first distinguished four market regimes: *fast sell-offs*, *slow rallies*, *slow sell-offs* and *fast rallies*, and then noticed than dominant regimes are fast sell-offs and slow rallies in the case of a negatively skewed volatility market (respectively slow sell-offs and rallies in the case of a negatively skewed volatility market). Here we directly focus on the dominant market regimes, namely (quickly) widening (sometimes) and (gently) tightening (most of the time).

33.1 Analysis in a Local Intensity Model

In this section we operate in a theoretical market given as a fixed local intensity model, with an assumingly steeply positively skewed implied base correlation smile ('steep local intensity model').

Note that in the set-up of a local intensity model, the strategy Δ^{lo}, if applied in continuous time, would provide a perfect replication of the tranche by the index (P&L trajectory e identically equal to zero). But we consider *hedging in discrete time*.

In the set-up of a local intensity model, one has (using also (93), regarding the last identity):

$$\Pi_t = \Pi^{lo}(t, N_t)\ ,\ P_t = P^{lo}(t, N_t)\ ,\ S_t = S^{lo}(t, N_t)\ ,\ \rho_t = \rho^{lo}(t, N_t)\ ,$$

for suitable functions $\Pi^{lo}(t,i)$, $P^{lo}(t,i)$, $S^{lo}(t,i)$, $\rho^{lo}(t,i)$ (also denoted $\Pi^{lo}_i(t)$, $P^{lo}_i(t)$, $S^{lo}_i(t)$, $\rho^{lo}_i(t)$).

33.1.1 Equity Tranche

Let us first consider the case of an equity tranche.

Convexity with respect to Realized Correlation One then has,

$$\delta e^{lo} \text{ is positive in tightening regimes and negative in widening regimes .} \tag{114}$$

To support (114), note that one has,

$$\delta_k e^{lo} = -\delta_k \Pi + \Delta^{lo}_{t_k}\,\delta_k P = \int_{t_k}^{t_{k+1}} (\Delta^{lo}_{t_k} - \Delta^{lo}_t)\,dP_t\ , \tag{115}$$

where it is well known that for an equity tranche $\Delta^{lo}_i(t)$ is non-decreasing with respect to time t. Therefore:
• On a (small) time interval $[t_k, t_{k+1}]$ with no default, one has $\delta_k P \le 0$ (so this corresponds to a tightening regime), and $\delta_k e^{lo} \ge 0$,
• On a (small) time interval $[t_k, t_{k+1}]$ with one default, one has $\delta_k P \ge 0$ (so this corresponds to a widening regime), and $\delta_k e^{lo} \le 0$.

Remark 33.2 One might be surprised by the fact that in a local intensity model $\delta_k e^{lo}$ is, at first sight, directional, in the sense that the sign of $\delta_k e^{lo}$ is driven by that of $\delta_k P$. One possible interpretation is that this is an effect of coincidence in this model of spreads tightening (resp. widening) with a stylized form of realized correlation (resp. decorrelation), namely nothing happening to anybody in the portfolio but for time-decay (resp. default of one name whilst the others stay alive). Now it is well-known that, consistently with (114), an equity tranche is Gamma negative with respect to the realized correlation (note in this regard that an equity tranche is essentially equivalent to a short put position on the portfolio loss; see, e.g., Gallo et al. [73]).

Ordering between the two deltas One has by definition of the Li base implied correlation, for every $t \in [0,T]$ and $i \in \{0,\ldots,n\}$:

$$\begin{aligned}&\Pi^{lo}_{i+1}(t) - \Pi^{lo}_i(t) = \Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_{i+1}(t)) - \Pi^{li}(t, S^{lo}_i(t), \rho^{lo}_i(t))\\ &= \big(\Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_{i+1}(t)) - \Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_i(t))\big)\\ &+ \big(\Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_i(t)) - \Pi^{li}(t, S^{lo}_i(t), \rho^{lo}_i(t))\big)\ .\end{aligned} \tag{116}$$

Moreover it is well known that for an equity tranche, one has,

$$\partial_\rho \Pi^{li}(t, S, \rho) \le 0 \; .$$

Finally, in the case of a steep local intensity model, one has,

$$\rho^{lo}_{i+1}(t) \ge \rho^{lo}_i(t) \; .$$

So

$$\Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_{i+1}(t)) \le \Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_i(t)) \; .$$

Therefore, by (116) used at $i = N_{t-}$:

$$\begin{aligned} \Delta^{lo}_t = \Delta^{lo}_i(t) &= \frac{\Pi^{lo}_{i+1}(t) - \Pi^{lo}_i(t)}{P^{lo}_{i+1}(t) - P^{lo}_i(t)} \\ &\le \frac{\Pi^{li}(t, S^{lo}_{i+1}(t), \rho^{lo}_i(t)) - \Pi^{li}(t, S^{lo}_i(t), \rho^{lo}_i(t))}{P^{lo}_{i+1}(t) - P^{lo}_i(t)} \approx \Delta^{li}_t \; . \end{aligned}$$

In view of (113), it follows that

$$\delta e^{li} \le \delta e^{lo} \text{ iff } \delta P \le 0 \; . \tag{117}$$

Tightening	**Widening**
$(\delta e^{li})^+ \le \delta e^{lo}$	$\delta e^{lo} \le -(\delta e^{li})^-$

Table 8: *Equity tranche in a steep local intensity model.*

Synthesis Combining (114) and (117), we get the picture depicted in Table 8. It might thus be so that in some cases the Li delta provides a better hedge than the local delta. But recall that we are in a local intensity model, in which the strategy Δ^{lo}, if applied in continuous time, would provide a perfect replication of the tranche by the index. This means that for hedge rebalancing frequencies large enough (like one week or less) δe^{lo} is very close to 0, and Table 8 reduces to Table 9:

Tightening	**Widening**
$\delta e^{li} \le 0 \simeq \delta e^{lo}$	$\delta e^{lo} \simeq 0 \le \delta e^{li}$

Table 9: *Case of a moderate to high rebalancing frequency in Table 8.*

33.1.2 Senior Tranche

In the case of a senior tranche, which is *positively* sensitive to correlation and Gamma *positive* with respect to the realized correlation, we simply exchange the contents of the cells in each of the Tables 8 and 9, yielding Tables 10 and 11.

Tightening	Widening
$\delta e^{lo} \leq -(\delta e^{li})^-$	$(\delta e^{li})^+ \leq \delta e^{lo}$

Table 10: *Senior tranche in a steep local intensity model.*

Tightening	Widening
$\delta e^{lo} \simeq 0 \leq \delta e^{li}$	$\delta e^{li} \leq 0 \simeq \delta e^{lo}$

Table 11: *Case of a moderate to high rebalancing frequency in Table 10.*

33.2 Analysis in a real market

In a steep market, not necessarily given by a local intensity model anymore (but still considering our stylized tranches, index and related definitions of spread and implied correlation), we can decompose the P&L increments in the following way:

$$\begin{aligned} \delta e^{lo} &= \Big(- \delta\Pi^{lo} + \Delta^{lo}\delta P^{lo}\Big) + \Big(\delta\Pi^{lo} - \delta\Pi - \Delta^{lo}(\delta P^{lo} - \delta P)\Big) \\ \delta e^{li} &= (-\delta\Pi^{lo} + \Delta^{li}\delta P^{lo}) + \Big(\delta\Pi^{lo} - \delta\Pi - \Delta^{li}(\delta P^{lo} - \delta P)\Big) \ , \end{aligned} \tag{118}$$

where $\delta\Pi$ and δP denote the increments of the market price of the tranche and of the index between the dates t_k and t_{k+1}, while $\delta\Pi^{lo}$ and δP^{lo} stand for the price increment predicted by the local intensity model calibrated at date t_k, given the new observations at date t_{k+1}.

In the right-hand side of (118):
• the first terms behave as in the analysis of section 33.1, whereas
• the second terms are due to the misspecification at date t_{k+1} of the local intensity model calibrated at date t_k.
This misspecification arises from the fact that the market-makers have revised their anticipations between date t_k and date t_{k+1}, according to the new market data observed at date t_{k+1} (and also, from time to time, according to more punctual economico-political macro news or events). Since there is more model risk on the tranches than on the index (unless the tranche is very close to the index: case of an equity tranche with k close to 1, or of a senior tranche with k close to 0), so typically:

$$\frac{|\delta\Pi - \delta\Pi^{lo}|}{|\delta\Pi^{lo}|} \geq \frac{|\delta P - \delta P^{lo}|}{|\delta P^{lo}|} \ ,$$

and even in the case of at-the-money tranches, for which the differential between the two deltas is the most significant (since $\Delta^{lo} \simeq \Delta^{li} \simeq 0$, resp. 1 in the case of far out-of-the money, resp. in-the-money tranches):

$$|\delta\Pi - \delta\Pi^{lo}| \gg \frac{|\delta\Pi^{lo}|}{|\delta P^{lo}|}|\delta P - \delta P^{lo}| \ .$$

As for assessing the impact of the misspecification terms in (118), is it thus probably acceptable to consider, 'at first order', that:

$$\delta\Pi^{lo} - \delta\Pi - \Delta^{lo}(\delta P^{lo} - \delta P) \simeq \delta\Pi^{lo} - \delta\Pi - \Delta^{li}(\delta P^{lo} - \delta P) \simeq \delta\Pi^{lo} - \delta\Pi \ ,$$

and thus, by (118),

$$\begin{aligned} \delta e^{lo} &\simeq \Big(-\delta\Pi^{lo} + \Delta^{lo}\delta P^{lo}\Big) + \Big(\delta\Pi^{lo} - \delta\Pi\Big) \\ \delta e^{li} &\simeq (-\delta\Pi^{lo} + \Delta^{li}\delta P^{lo}) + \Big(\delta\Pi^{lo} - \delta\Pi\Big) . \end{aligned} \tag{119}$$

Moreover, it seems reasonable to expect that:
(i) At market regimes with defaults and widening credit spreads, and thus, since we are in a steep market, with increasing levels of realized correlations, the market-makers will have a tendency to push the tranches' Li implied base correlations upwards compared to those predicted by the local intensity model calibrated at date t_k, whereas
(ii) At slow market regimes with tightening credit spreads, the market-makers will have a tendency to push the tranches' Li implied base correlations downwards compared to those predicted by the model calibrated at date t_k.

In the case of an equity tranche, which is negatively sensitive to correlation, this implies that:
(i) $\delta\Pi \leq \delta\Pi^{lo}$ in widening regimes, and
(ii) $\delta\Pi \geq \delta\Pi^{lo}$ in tightening regimes.
By comparison with the situation of a steep local intensity model, δe^{li} and δe^{lo} are pushed away from 0 by the same amount in Table 9 (assuming a moderate to high rebalancing frequency), as an effect of the misspecification terms $\delta\Pi^{lo} - \delta\Pi$ in (119).

Likewise, in the case of a senior tranche, one can check that by comparison with the situation in a local intensity model that δe^{li} and δe^{lo} are pushed away from 0 by the same amount in Table 11 (assuming a moderate to high rebalancing frequency), as an effect of the misspecification terms $\delta\Pi^{lo} - \delta\Pi$ in (119).

The final results of the previous findings are summed-up in Table 12.

Market regime	**Tightening**	**Widening**
Equity tranche	$\delta e^{li} \leq \delta e^{lo} \leq 0$	$0 \leq \delta e^{lo} \leq \delta e^{li}$
Senior tranche	$0 \leq \delta e^{lo} \leq \delta e^{li}$	$\delta e^{li} \leq \delta e^{lo} \leq 0$

Table 12: *Synthesis of the results in the case of steep market (for a moderate to high hedging rebalancing frequency).*

Consistently with the fact the a local intensity model fits the market over the full set of CDO tranches at every point in time, whereas the Li model only provides a per tranche fit, we thus find that the local delta provides a better hedge (in the sense of maintaining the P&L increments closer to 0) than the Li delta (under the market risk-neutral measure, and, provided we have physical as well as implied positive skewness, under the objective measure as well).

Remark 33.3 In the situation of an even (as opposed to steep) local intensity model (, with a flat or close to flat base correlation smile, the ordering between the two deltas changes. The results analogous to those of Tables 8 and 10 are displayed in Table 13.
Moving further to the situation of a *real* even market, not necessarily given by a local intensity model anymore, but with a negative correlation between index spreads and correlations of credit spread, we expect that for an equity tranche, which is negatively sensitive to correlation:

(i) $\delta\Pi \geq \delta\Pi^{lo}$ in widening regimes,
(ii) $\delta\Pi \leq \delta\Pi^{lo}$ in tightening regimes,
and the opposite results for a senior tranche. So, finally, the situation of Table 13 is still valid in a real even market.
Again, consistently with the fact the a local intensity model fits the market over the full set of CDO tranches at every point in time, whereas the Li model only provides a per tranche fit, we find that the local delta provides a better hedge than the Li delta in an even market. Note that the conclusion is in fact clearer in an even market than in a steep market, since for an even market the conclusion holds irrespectively of the frequency of the hedge rebalancing.

Market regime	**Tightening**	**Widening**
Equity tranche	$0 \leq \delta e^{lo} \leq \delta e^{li}$	$\delta e^{li} \leq \delta e^{lo} \leq 0$
Senior tranche	$\delta e^{li} \leq \delta e^{lo} \leq 0$	$0 \leq \delta e^{lo} \leq \delta e^{li}$

Table 13: *Equity and senior tranches in an even local intensity model.*

Needless to say, the rough argumentation of this section is only a first attempt to establish an analogy between the mechanisms of equity and credit portfolio derivatives markets. In particular it would be important to check whether it can be assessed numerically.

References as Author or Co-Author *(in Bold: articles published or accepted in Refereed Journals or Books)*

[1] Bielecki, T.R., Crépey, S. and Jeanblanc, M.: Up and Down Credit Risk. *Submitted*.

[2] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Arbitrage pricing of defaultable game options with applications to convertible bonds. Forthcoming in ***Quantitative Finance*** (available online at `www.defaultrisk.com`).

[3] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Valuation and hedging of defaultable game options in a hazard process model. *Submitted*, 2008 (available online at `www.defaultrisk.com`).

[4] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Defaultable options in a Markovian intensity model of credit risk. ***Mathematical Finance***, vol. 18, pp. 493-518, 2008.

[5] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Convertible bonds in a defaultable diffusion model. *Submitted*, 2007 (available online at `www.defaultrisk.com`).

[6] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Valuation of basket credit derivatives in the credit migrations environment. ***Handbook of Financial Engineering***, J. Birge and V. Linetsky eds., Elsevier, 2007.

[7] Carmona, R. and Crépey, S.: Importance Sampling and Interacting Particle Systems for the Estimation of Markovian Credit Portfolios Loss Distribution, *Submitted*, 2008.

[8] Crépey, S.: About the pricing equations in finance. *Submitted*, 2008.

[9] Crépey, S.: Tikhonov Regularization. Forthcoming in ***Encyclopedia of Quantitative Finance***, editor Rama Cont.

[10] Crépey, S.: Delta-hedging Vega Risk? ***Quantitative Finance***, 2004, 4, 559–579.

[11] Crépey, S.: Calibration of the local volatility in a generalized Black–Scholes model using Tikhonov regularization. ***SIAM Journal on Mathematical Analysis***, Vol. 34 No 5 (2003), pp. 1183–1206.

[12] Crépey, S.: Calibration of the local volatility in a trinomial tree using Tikhonov regularization. ***Inverse Problems***, 19 (2003), pp. 91–127.

[13] S. Crépey. *Contribution à des méthodes numériques appliquées à la Finance et aux Jeux Différentiels*, Thèse de Doctorat, Ecole Polytechnique–INRIA Sophia Antipolis, France, Janvier 2001.

[14] Crépey, S. and Matoussi, A.: About the Greeking equations in finance. *In Preparation.*

[15] Crépey, S. and Matoussi, A.: Reflected and doubly reflected BSDEs with jumps: A priori estimates and comparison principle. ***Annals of Applied Probability***, Vol 18, Issue 5 (October 2008), pp. 2041–69.

[16] Crépey, S. and Rahal, A.: Computing Pricing Functions by Simulation. *In Preparation.*

Other References

[17] Achdou, Y., Indragoby, G. and Pironneau, O.: Volatility Calibration with American Options. *Methods and Applications of Analysis*, International Press Vol. 11, No. 3, pp. 001-024, September 2004.

[18] Alvarez O., Tourin, A.: Viscosity solutions of nonlinear integro-differential equations. *Annales de l'institut Henri Poincaré (C) Analyse non linéaire*, 13 no. 3 (1996), p. 293-317.

[19] Avellaneda, M., Friedman, C., Holmes, R. and Samperi, D.: Calibrating volatility surfaces via relative-entropy minimization, *Applied Math. Finance*, 41 (1997), pp. 37–64.

[20] Bally V., Caballero E., Fernandez, B. and El-Karoui N.: Reflected BSDE's PDE's and Variational Inequalities, to appear in *Bernoulli.*

[21] Bally, V. and Matoussi, A.: Weak solutions for SPDEs and Backward doubly stochastic differential equations, *Journal of Theoretical Probability,* Vol. **14,** No. 1, 125-164 (2001).

[22] Barles, G., Buckdahn, R. and Pardoux, E.: Backward Stochastic Differential Equations and Integral-Partial Differential Equations. *Stochastics and Stochastics Reports*, Vol. 60, pp. 57-83 (1997).

[23] Barles, G., Imbert, C.: Second-Order Elliptic Integro-Differential Equations: Viscosity Solutions' Theory Revisited. *Annales de l'IHP*, To Appear.

[24] Barles, G. and L. Lesigne: SDE, BSDE and PDE. *Backward Stochastic differential Equations. Pitman Research Notes in Mathematics Series* 364, Eds El Karoui, N. and Mazliak, L., pp. 47-80 (1997).

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

[26] Becherer, D. and Schweizer, M.: Classical Solutions to Reaction-Diffusion Systems for Hedging with Interacting Itô and Point Processes. *Annals of Applied Probability*, 15, 1111-1144, 2005.

[27] Bensoussan, A. and Lions, J.L.: *Applications of Variational Inequalities in Stochastic Control*, North-Holland, Amsterdam, 1982.

[28] Bensoussan, A. and Lions, J.L.: *Impulse Control and Quasi-variational Inequalities*, Gauthier-Villars, Paris, 1984.

[29] Bielecki, T.R. and Rutkowski, M.: *Credit Risk: Modeling, Valuation and Hedging.* Springer, 2002.

[30] Bielecki, T.R., Vidozzi, A. and Vidozzi, L.: An efficient approach to valuation of credit basket products and ratings triggered step-up bonds. *Working Paper*, 2006.

[31] Bielecki, T.R., Vidozzi, A. and Vidozzi, L.: A Markov Copulae Approach to Pricing and Hedging of Credit Index Derivatives and Ratings Triggered Step–Up Bonds, *Working Paper.*

[32] Binder, A., Engl, H. W., Groetsch, C. W., Neubauer, A. and Scherzer, O.: Weakly closed nonlinear operators and parameter identification in parabolic equations by Tikhonov regularization, *Appl. Anal.* 55 (1994), pp. 13–25.

[33] Black, F. and Scholes, M.: The pricing of options and corporate liabilities, *J. Pol. Econ.*, 81 (1973), pp. 637–659.

[34] Bouchard, B. and Chassagneux, J.-F.: Discrete time approximation for continuously and discretely reflected BSDE's. Forthcoming in *Stochastic Processes and their Applications.*

[35] Bouchard B. and Elie R: Discrete time approximation of decoupled Forward-Backward SDE with jumps, Forthcoming in *Stochastic Processes and their Applications.*

[36] Bouchard, B. and Touzi, N.: Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications* 111 (2004), 175–206.

[37] Brémaud, P.: *Point processes and queues, martingale dynamics*, Springer, 1981.

[38] Brémaud, P. and Yor, M.: Changes of filtrations and of probability measures. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 45, 269–295, 1978.

[39] Briani, M., La Chioma, C. and Natalini, R.: Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory, *Numer. Math.*, 2004, vol. 98, no4, pp. 607-646.

[40] Caffarelli, L., Crandall, M.G., Kocan, M. and Swiech, A.: On viscosity solutions of fully non linear equations with measurable ingredients, *CPAM*, XLIX (1996), pp. 365–397.

[41] Chassagneux, J.-F.: Discrete-time approximation of doublly reflected BSDEs. *Working Paper*, 2007.

[42] Coculescu, D. and Nikeghbali, A.: Hazard processes and martingale hazard processes. *Working Paper*, 2008.

[43] Cont, R. and Kan, Y.H.: Dynamic hedging of portfolio credit derivatives, *Working Paper*, 2008.

[44] Cont, R. and Minca, A.: Recovering Portfolio Default Intensities Implied by CDO Quotes, *Working Paper*, 2008.

[45] Cont, R., Voltchkova, K. Finite difference methods for option pricing methods in exponential Levy models, *SIAM Journal on Numerical Analysis*, Vol 43, 2005.

[46] Crandall, M., Ishii, H. and Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc.*, 1992.

[47] Crandall, M.G., Kocan, M. and Swiech, A.: L^p-theory for fully nonlinear parabolic equations, *CPDE*, 25 (2000), 11 and 12, pp. 1997–2053.

[48] Cvitanic, J. and Karatzas, I.: Backward Stochastic Differential Equations with reflection and Dynkin Games, *Annals of Probability* 24 (4), 2024-2056 (1996).

[49] Davis, M. and Lo, V.: Modelling default correlation in bond portfolios, in *Mastering Risk Volume 2: Applications*, ed. Carol Alexander, Financial Times Prentice Hall 2001, pp 141-151.

[50] Davis, M.H. and Varaiya, P.: The multiplicity of an increasing family of σ-fields, *Annals of Probab.*, 2(1974), 958–963.

[51] Delbaen, F. and Schachermayer, W.: The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen* 312 (1997), 215–250.

[52] Dellacherie, C. and Meyer, P.-A.: *Probabilities and Potential*, North-Holland, 1982.

[53] Del Moral, P.: *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Springer-Verlag, 2004.

[54] Del Moral, P. and Garnier, J.: Genealogical particle analysis of rare events. *Annals of Applied Probability,* 15(4):2496–2534, November 2005.

[55] Derman, E.: Regimes of Volatility. *Risk magazine*, April 1999.

[56] Derman, E. and Kani, I.: Riding on a smile, *Risk*, 7 (1994), pp. 32–39.

[57] Duffie, D. and Garleanu, N.: Risk and valuation of collateralized debt olbligations. *Financial Analysts Journal*, 57 (1), 41–59, 2001.

[58] Dupire, B.: Pricing with a smile, *Risk*, 7 (1994), pp. 18–20.

[59] Dynkin, E.B.: Game variant of a problem on optimal stopping. *Soviet Math. Dokl.* 10 (1969), 270–274.

[60] El Karoui, N., Hamadène, S. and Matoussi, A.: Backward Stochastic Differential Equations and Applications, in *Paris-Princeton Lecture Notes on Mathematical Finance*, Edts. R. Carmona, Springer-Verlag, Forthcoming.

[61] El Karoui, N., Kapoudjian, E., Pardoux, C., Peng, S., and Quenez, M.-C.: Reflected solutions of backward SDEs, and related obstacle problems for PDEs. *Annals of Probability* 25 (1997), 702–737.

[62] El Karoui, N., Peng, S., and Quenez, M.-C.: Backward stochastic differential equations in finance. *Mathematical Finance* 7 (1997), 1–71.

[63] El Karoui, N. and Quenez, M.-C.: Nonlinear pricing theory and backward stochastic differential equations. In: *Financial Mathematics, Bressanone, 1996,* W. Runggaldier, ed. Lecture Notes in Math. 1656, Springer, Berlin Heidelberg New York, 1997, pp. 191–246.

[64] Engl, H. W., Hanke, M. and Neubauer, A.: *Regularization of Inverse Problems.* Kluwer, Dordrecht, 1996.

[65] Engl, H. W., Kunisch, K. and Neubauer,A.: Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems. *Inverse Problems*, 5 (1989), 4, pp. 523–540.

[66] Ethier, H.J. and Kurtz, T.G.: *Markov Processes. Characterization and Convergence*. Wiley, 1986.

[67] Fabes, E.: Singular integrals and partial differential equations of parabolic type, *Studia Math.*, XXVIII (1966), pp. 6–131.

[68] W. Fleming and H. Soner. *Controlled Markov processes and viscosity solutions*, Second edition, Springer, 2006.

[69] Föllmer, H., and Sondermann, D. Hedging of non-redundant contingent claims, *Contributions to Mathematical Economics*, A. Mas-Colell and W. Hildenbrand ed., North-Holland, Amsterdam, 205–223, 1986.

[70] Frey, R., Backhaus, J.: Dynamic Hedging of Synthetic CDO Tranches with Spread Risk and Default Contagion. *Working Paper*, 2007.

[71] Frey, R., Backhaus, J.: Pricing and Hedging of Portfolio Credit Derivatives with Interacting Default Intensities. *Working Paper*, 2007.

[72] Frey, R., Backhaus, J.: Portfolio credit risk models with interacting default intensities: a Markovian approach. *Working Paper*, 2004.

[73] Gallo, A., Galiani, S., Jonsson, J. and Kakodkar, A.: Pricing and Hedging of Portfolio Credit Derivatives with Interacting Default Intensities. *Merrill Lynch Credit Derivatives Handbook*, Vol. 2, 2006.

[74] Giesecke, K., Goldberg, L.: A Top Down Approach to Multi-Name Credit. *Working Paper*, 2007.

[75] Glasserman, P.: *Monte Carlo Methods in Financial Engineering*, Springer, 2004.

[76] Gobet, E., Lemor, J.-P. and Warin, X.: *A regression-based Monte-Carlo method for backward stochastic differential equations*, *Ann. Appl. Probab.*, 15, 2172-2202, 2005.

[77] Gyöngy, I.: Mimicking the one dimensional distributions of processes having an Ito differential. *Probability theory and related fields*, Vol. 71, No. 4, pp. 501-516, 1986.

[78] Halperin, I. and Tomecek, P.: Climbing Down from the Top: Single Name Dynamics in Credit Top Down Models. *Working paper, Quantitative Research JP Morgan*, September 2008.

[79] Hamadène, S.: Mixed zero-sum differential game and American game options. *SIAM J. Control Optim.* 45 (2006), 496–518.

[80] Hamadène, S. and Hassani, M.: BSDEs with two reflecting barriers driven by a Brownian motion and an independent Poisson noise and related Dynkin game. *Electronic Journal of Probability*, vol. 11 (2006), paper no. 5, pp. 121-145.

[81] Hamadène, S. and Ouknine, Y.: Reflected backward stochastic differential equation with jumps and random obstacle. *Electronic Journal of Probability*, Vol. 8 (2003), p. 1–20.

[82] Herbertsson, A.: Pricing kth-to-default swaps under default contagion: the matrix-analytic approach. *The Journal of Computational Finance*, (1-30) Volume 12/Number 1, Fall 2008.

[83] Herbertsson, A.: Pricing portfolio credit derivatives. *PhD Thesis*, Göteborg University, 2007.

[84] Jacod, J. and Shiryaev, A.N.: *Limit Theorems for Stochastic Processes.* Springer, Berlin Heidelberg New York, 2003.

[85] Jakobsen, E. R., Karlsen, K. H. and La Chioma, C.: Error estimates for approximate solutions to Bellman equations associated with controlled jump-diffusions. *Numer. Math.*, Forthcoming.

[86] Jarrow, R.A. and Turnbull, S.: Pricing Derivatives on Financial Securities Subject to Credit Risk. *Journal of Finance* vol. 50, March, 1995.

[87] Jarrow, R.A. and Yu, F.: Counterparty risk and the pricing of defaultable securities. *Journal of Finance* 56, 1765-1799, 2001.

[88] Johansen, A. M., Del Moral, P. , and Doucet, A.: Sequential Monte Carlo Samplers for Rare Events. *Working Paper.*

[89] Kallsen, J. and Kühn, C.: Convertible bonds: financial derivatives of game type. In: *Exotic Option Pricing and Advanced Lévy Models*, Edited by Kyprianou, A., Schoutens, W. and Wilmott, P., Wiley, 2005, pp. 277–288.

[90] Karatzas, I.: On the pricing of American options. *Appl. Math. Optim.* 17 (1988), 37–60.

[91] Kifer, Y.: Game options. *Finance and Stochastics* 4 (2000), 443–463.

[92] Krylov, N.V.: *Controlled Diffusion Processes*, Springer Verlag, Berlin, 1980.

[93] Kusuoka, S.: A remark on default risk models. *Advances in Mathematical Economics* 1 (1999), 69–82.

[94] Lagnado, R. and Osher, S.: A Technique for Calibrating Derivative Security Pricing Models: Numerical Solution of an Inverse Problem, *Journal of Computational Finance*, 1 (1997), pp. 13–25.

[95] Lando, D.: Three Essays on Contingent Claims Pricing. *Ph.D. Thesis*, Cornell University, 1994.

[96] Last, G., Brandt, A.: *Marked Point Processes on the Real Line: The Dynamical Approach*, Springer, 1995.

[97] Laurent, J.-P., Cousin, A. and Fermanian, J-D.: Hedging default risks of CDOs in Markovian contagion models. *Working Paper*, 2007.

[98] Laurent, J.-P.: Basket default swaps, CDOs and factor copulas. *Journal of Risk*, 2003.

[99] Lepeltier, J.-P. and Maingueneau, M.: Le jeu de Dynkin en théorie générale sans l'hypothèse de Mokobodski. *Stochastics* 13 (1984), 25–44.

[100] Li, D.-X.: On default correlation: a copula approach. *Journal on Fixed Income* 9, March, 43–54, 2000.

[101] Longstaff, F.A. and Schwartz, E.S.: Valuing American options by simulations: a simple least-squares approach. *Review of Financial Studies* 14 (2001), 113–147.

[102] Ma, J. and Cvitanić, J.: Reflected Forward-Backward SDEs and obstacle problems with boundary conditions. *Journal of Applied Mathematics and Stochastic Analysis* 14(2) (2001), 113–138.

[103] Merton, R.C.: The theory of rational option pricing, *Bell J. Econ. Man. Sc.*, 4 (1973), pp. 141–183.

[104] Nikeghbali, A. and Yor, M.: A definition and some characteristic properties of pseudo-stopping times. *Annals of Probability* 33 (2005), 1804–1824.

[105] Nocedal, J. and Wright, S.J.: *Numerical Optimization, 2^{nd} edition*, Springer, 2006.

[106] O'Kane, D. and Livesey, M.: Base Correlation Explained, *Lehman Brothers Fixed Income Quantitative Credit Research*, Nov 2004.

[107] Pardoux, E., Pradeilles, F. and Rao, Z.: Probabilistic interpretation of systems of semilinear PDEs. *Annales de l'Institut Henri Poincaré, série Probabilités-Statistiques*, 33, 467–490, 1997.

[108] Pham, H.: Optimal stopping of controlled jump diffusion processes: A viscosity solution approach, *Journal of Mathematical Systems, Estimation and Control*, 8, 1–27, 1998.

[109] Phillips, D.: A technique for the numerical solution of certain integral equations of the first kind, *J Assoc Comput Mach*, 9, 84-97 (1962).

[110] Protter, P.E.: *Stochastic Integration and Differential Equations, Second Edition.* Springer, Berlin Heidelberg New York, 2004.

[111] Schönbucher, P.: Portfolio Losses And The Term Structure Of Loss Transition Rates: A New Methodology For The Pricing Of Portfolio Credit Derivatives. *Working Paper*, 2006.

[112] Schweizer, M.: Option hedging for semimartingales. *Stochastic Processes and their Applications* 37 (1991), 339–363.

[113] Sidenius, J., Piterbarg, V. and Andersen, L.: A New Framework For Dynamic Credit Portfolio Loss Modelling. *Working Paper*, 2006.

[114] Stroock, D.W. and Varadhan, S.R.S.: *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin, 1979.

[115] Tikhonov, A.: Solution of incorrectly formulated problems and the regularization method, *Soviet Math Dokl* 4, 1035-1038, English translation of *Dokl Akad Nauk SSSR* 151, 501-504, 1963.

[116] Yu, F.: Correlated defaults in Intensity-Based Models. *Mathematical Finance*, Vol. 17, No. 2 (April 2007), pp. 155-173.