Thèse présentée pour obtenir le titre de

DOCTEUR DE L'ECOLE POLYTECHNIQUE

specialité: Mathématiques Appliquées

par

Peter TANKOV

Processus de Lévy en Finance: Problèmes Inverses et Modélisation de Dépendance

Lévy Processes in Finance: Inverse Problems and Dependence Modelling

Thèse soutenue le 21 septembre 2004 devant le jury composé de

Ole E. Barndorff-Nielsen	Rapporteur
Rama Cont	Directeur de thèse
Nicole El Karoui	
Monique Jeanblanc	Rapporteur
Bernard Lapeyre	
Agnès Sulem	

Résumé

Cette thèse traite de la modélisation de prix boursiers par les exponentielles de processus de Lévy. La première partie développe une méthode non-paramétrique stable de calibration de modèles exponentielle-Lévy, c'est-à-dire de reconstruction de ces modèles à partir des prix d'options cotées sur un marché financier. J'étudie les propriétés de convergence et de stabilité de cette méthode de calibration, décris sa réalisation numérique et donne des exemples de son utilisation. L'approche adoptée ici consiste à reformuler le problème de calibration comme celui de trouver un modèle exponentielle-Lévy risque-neutre qui reproduit les prix d'options cotées avec la plus grande précision possible et qui a l'entropie relative minimale par rapport à un processus "a priori" donné. Ce problème est alors résolu en utilisant la méthode de *régularisation*, provenant de la théorie de problèmes inverses mal posés. L'application de ma méthode de calibration aux données empiriques de prix d'options sur indice permet d'étudier certaines propriétés des mesures de Lévy implicites qui correspondent aux prix de marché

La deuxième partie est consacrée au développement d'une méthode permettant de caractériser les structures de dépendance entre les composantes d'un processus de Lévy multidimensionnel et de construire des modèles exponentielle-Lévy multidimensionnels. Cet objectif est atteint grâce à l'introduction de la notion de *copule de Lévy*, qui peut être considérée comme l'analogue pour les processus de Lévy de la notion de copule, utilisée en statistique pour modéliser la dépendance entre les variables aléatoires réelles. Les exemples de familles paramétriques de copules de Lévy sont donnés et une méthode de simulation de processus de Lévy multidimensionnels, dont la structure de dépendance est décrite par une copule de Lévy, est proposée.

Mots clefs: processus de Lévy, produits dérivés, calibration, problèmes inverses, régularisation, problèmes mal posés, entropie relative, copules, dépendance.

Abstract

This thesis deals with the modelling of stock prices by the exponentials of Lévy processes. In the first part we develop a non-parametric method allowing to calibrate exponential Lévy models, that is, to reconstruct such models from the prices of market-quoted options. We study the stability and convergence properties of this calibration method, describe its numerical implementation and give examples of its use. Our approach is first to reformulate the calibration problem as that of finding a risk-neutral exponential Lévy model that reproduces the option prices with the best possible precision and has the smallest relative entropy with respect to a given prior process, and then to solve this problem via the *regularization* methodology, used in the theory of ill-posed inverse problems. Applying this calibration method to the empirical data sets of index options allows us to study some properties of Lévy measures, implied by market prices.

The second part of this thesis proposes a method allowing to characterize the dependence structures among the components of a multidimensional Lévy process and to construct multidimensional exponential Lévy models. This is done by introducing the notion of *Lévy copula*, which can be seen as an analog for Lévy processes of the notion of copula, used in statistics to model dependence between real-valued random variables. We give examples of parametric families of Lévy copulas and develop a method for simulating multidimensional Lévy processes with dependence given by a Lévy copula.

Key words: Lévy processes, option pricing, calibration, inverse problems, regularization, ill-posed problems, relative entropy, copulas, dependence

Remerciements

Je voudrais, en premier lieu, exprimer toute ma gratitude à Rama Cont, mon directeur de thèse, non seulement pour sa disponibilité et son soutien au cours de ces trois années, mais aussi pour avoir fait preuve, à plusieurs reprises, de sa confiance en moi, et pour m'avoir aidé, en m'envoyant à des nombreuses conférences et séminaires, à faire connaître mon travail de recherche.

Je n'aurais jamais pensé à travailler dans le domaine de la finance mathématique si je n'avais pas suivi d'excellents cours de Nicole El Karoui, d'abord à l'Ecole Polytechnique et ensuite au DEA de Probabilités et Applications de l'Université Paris VI. Je la remercie pour cela mais aussi pour les nombreuses interactions qu'on a eu au cours de ma thèse, notamment à l'occasion du groupe de travail du CMAP.

Jan Kallsen m'a expliqué comment généraliser la notion de copule de Lévy, que j'ai introduite pour les processus de Lévy aux sauts positifs, au cas des processus de Lévy généraux. Je le remercie également pour avoir bien voulu écrire un article joint sur ce sujet et pour des nombreuses discussions qu'on a eu pendant ce travail.

Je remercie Ole Barndorff-Nielsen, Thomas Mikosch et Elisa Nicolato pour avoir fait de ma visite à MaPhySto en août-septembre 2002 un moment inoubliable et pour les discussions qu'on a eu pendant et après mon séjour au Danemark, qui ont été toutes très utiles.

Je voudrais également exprimer ma reconnaissance à Douglas N. Arnold and Scot Adams qui m'ont accueilli pendant ma visite à l'Institut de Mathématiques et Applications de l'Université de Minnesota, un séjour auquel cette thèse doit beaucoup.

Mes remerciements vont aussi à Philippe Durand et Emmanuel Bacry, qui ont été responsables de deux stages que j'ai effectués avant de commencer cette thèse, l'un au Crédit Lyonnais et l'autre au CMAP, et qui ont beaucoup fait pour ma formation scientifique.

Je suis reconnaissant à Christoph Schwab et Ana-Maria Matache d'avoir partagé avec moi leurs idées sur les méthodes numériques pour les processus de Lévy, au cours de leurs visites à Paris et de ma visite à Zurich.

Je remercie enfin tous les participants du groupe de travail du CMAP de toutes nos interactions scientifiques, mais aussi d'avoir contribué à une ambiance très agréable et chaleureuse dans notre équipe de modélisation financière.

Contents

Remarks on notation

Basic notation For a function F, $\operatorname{Dom} F$ and $\operatorname{Ran} F$ denote, respectively, the domain and the range of F.

δ_x is a measure given by $\delta_x(A) = 1_{x\in A}$ for every Borel set A and $\lambda|_A$ is the restriction of the Lebesgue measure onto a Borel set A.

$\mathbb{R}^d_+$ denotes $[0,\infty)^d \setminus \{0\}$ and $\bar{\mathbb{R}}$ denotes the extended real line: $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$

For two vectors $a, b \in \bar{\mathbb{R}}^d$, inequalities like $a \leq b$ should be interpreted as $a_i \leq b_i$ for $i = 1, \ldots, d$. For any such vectors, $[a, b)$ denotes a right-open left-closed interval of $\bar{\mathbb{R}}^d$: $[a, b) := [a_1, b_1) \times \cdots \times [a_d, b_d)$. Other types of intervals: $(a, b]$, $[a, b]$ and (a, b) are defined similarly. When we want to consider an interval as a set of its vertices, without specifying whether the boundary is included, we will use the notation $|a, b|$. Finally, $\langle a, b\rangle$ denotes the scalar product of a and b.

Measures and processes In this thesis, we fix a time horizon $T_\infty < \infty$ and Ω denotes the space of $\mathbb{R}^d$-valued càdlàg functions on $[0, T_\infty]$, equipped with the Skorokhod topology (see [54]). Unless otherwise mentioned, X is the coordinate process: for every $\omega \in \Omega$, $X_t(\omega) := \omega(t)$. $\mathcal{F}$ is the smallest σ-field, for which the mappings $\omega \in \Omega \mapsto \omega(s)$ are measurable for all $s \in [0, T_\infty]$. For any $t \in [0, T_\infty]$, $\mathcal{F}_t$ is the smallest σ-field, for which the mappings $\omega \in \Omega \mapsto \omega(s)$ are measurable for all $s \in [0, t]$.

For a semimartingale X, $[X]$ denotes is quadratic variation and $[X]^c$ is the continuous part of $[X]$.

$\mathcal{P}(\Omega)$ denotes the set of all probability measures (stochastic processes) on $(\Omega, \mathcal{F})$. For two probability measures $P, Q \in \mathcal{P}(\Omega)$, $Q \ll P$ means that Q is absolutely continuous with respect to P, that is, for every $B \in \mathcal{F}$, $P(B) = 0$ implies $Q(B) = 0$.

For $\{P_n\}_{n\geq 1}, P \in \mathcal{P}(\Omega)$ the notation $P_n \Rightarrow P$ means that the sequence $\{P_n\}_{n\geq 1}$ converges weakly to P, that is, for every continuous bounded function $f : \Omega \to \mathbb{R}$, $P_n(f) \to P(f)$.

The set $\mathcal{L}$ consists of all probability measures $P \in \mathcal{P}(\Omega)$, under which (X, P) is a Lévy process.

For a constant $B > 0$, $\mathcal{L}_B$ and $\mathcal{L}_B^+$ denote the sets of all probabilities $P \in \mathcal{L}$ such that (X, P) satisfies, $P[|\Delta X_t| \leq B \ \forall t : 0 \leq t \leq T_\infty] = 1$ for $\mathcal{L}_B$ (jumps are bounded by B in absolute value) and $P[\Delta X_t \leq B \ \forall t : 0 \leq t \leq T_\infty] = 1$ for $\mathcal{L}_B^+$ (jumps are bounded by B from above).

$\mathcal{L}_{NA}$ is the set of all probability measures $P \in \mathcal{L}$ corresponding to Lévy processes describing markets with no arbitrage opportunity, as defined in Proposition 1.8.

Finally, $\mathcal{M}$ stands for the set of all probability measures $P \in \mathcal{P}(\Omega)$, under which e^{X_t} is a martingale.

Introduction et principaux résultats

Un processus de Lévy est un processus stochastique aux accroissements indépendants et stationnaires: si $\{X_t\}_{t\geq 0}$ est un processus de Lévy, $X_t - X_s$ avec $t > s$ est indépendant de l'histoire du processus avant le temps s et sa loi ne dépend pas de t ou s séparément mais seulement de $t - s$. Cette propriété des accroissements évoque une analogie avec des fonctions linéaires: on peut dire que les processus de Lévy sont, dans un certain sens, des "processus linéaires" ou additifs.

Malgré cette simplicité apparente, les processus de Lévy ont des nombreuses propriétés intéressantes et constituent un domaine d'étude en plein développement: plusieurs ouvrages ont été publiés récemment [17, 87] et une série de conférences internationales dédiées aux processus de Lévy et applications a rencontré un grand succès [6].

Sur le plan de modélisation financière, les processus de Lévy fournissent une classe de modèles avec sauts qui est à la fois suffisamment riche pour bien décrire les données empiriques et assez simple pour faire beaucoup de calculs analytiquement. L'intérêt de tels modèles en finance est principalement dû aux trois facteurs suivants.

Premièrement, dans un modèle aux trajectoires continues comme un modèle de diffusion, le processus de prix se comporte localement comme un mouvement brownien et la probabilité que le prix bouge beaucoup pendant un temps court est très petite si la valeur de volatilité de volatilité n'est pas déraisonnablement grande. Cela implique que dans de tels modèles les prix des options "hors de la monnaie" doivent être beaucoup plus petits que ce que l'on observe en réalité. Par contre, si le processus de prix peut sauter, même pour une maturité courte on ne peut pas négliger la probabilité d'un mouvement inattendu du prix qui déplacerait l'option dans la monnaie.

Deuxièmement, du point de vue de la couverture, les modèles aux trajectoires continues

correspondent en général aux marchés complets ou aux marchés qui peuvent être complétés en ajoutant un ou deux actifs supplémentaires comme dans les modèles à volatilité stochastique. Ceci implique que le flux terminal de n'importe quel actif contingent est parfaitement repliquable, et les options cotées deviennent donc redondantes. Le "mystère" est facilement expliquée par la présence de sauts dans les prix: en présence de sauts la couverture parfaite est impossible et les options permettent aux participants du marché de couvrir les risques qui ne sont pas repliquables par une stratégie de trading ne faisant intervenir que le sous-jacent.

Le troisième argument en faveur de l'utilisation de modèles discontinus, qui est peut-être le plus fort, est la présence-même de sauts dans les prix. Figure 1 de page 28 trace l'évolution du taux de change DM/USD pendant une période de deux semaines en 1992. Sur ce graphique il y a au moins trois endroits ou le taux a bougé de plus de 100 points de base en moins de 5 minutes. Il est claire que des variations de prix comme celles-là ne peuvent pas être expliquées par un modèle aux trajectoires continues, mais elles doivent être prises en compte pour une gestion de risques fiable.

Quand cet étude a été commencé, plusieurs thèses de doctorat avaient déjà été soutenues [78, 79, 81, 86, 96] et quelques centaines d'articles avaient été publiés dans le domaine de modélisation financière avec des processus de Lévy. Cependant, deux questions majeures, qui apparaissent dans le titre de cette thèse étaient restées ouvertes.

Premièrement, alors que la piste de recherche, privilégiée dans la litterature, était le développement de méthodes efficaces de valorisation de produits dérivés dans les modèles exponentielle-Lévy, une étape essentielle d'utilisation d'un tel modèle est de trouver les valuers des paramètres, où, plus généralement, le triplet caractéristique du processus de Lévy sous-jacent, qui soit compatible avec les prix des options cotées sur le marché financier. Ce problème, connu sous le nom du problème de calibration, est un problème inverse à celui de valorisation des options européennes dans un modèle exponentielle-Lévy et il est nettement plus difficile à résoudre que ce dernier. Le problème de calibration a été traité par plusieurs auteurs (cf. par exemple [4, 35, 53, 85]) dans le cadre d'un modèle de diffusion markovienne, où le paramètre inconnu est la fonction de volatilité locale $\sigma(S_t, t)$. Cependant, dans le contexte de processus avec sauts, bien que plusieurs articles proposent des modèles paramétriques à base de processus de Lévy [5, 22, 36, 62, 67, 71], il n'existait pas, avant cet étude, de méthode permettant d'en choisir un parmi cette multitude de modèles ni d'algorithme stable de calibration de paramètres

de tels modèles à partir de prix d'options européennes. Dans la première partie de la thèse je développe donc une méthode non-paramétrique robuste de calibration de modèles exponentielle-Lévy, étudie ses propriétés de convergence et de stabilité, décris sa réalisation numérique et donne des exemples de son utilisation. Mon approche consiste à reformuler le problème de calibration comme celui de trouver un processus de Lévy risque neutre qui reproduit les prix d'options cotées avec la plus grande précision possible et qui a l'entropie relative minimale par rapport à un processus a priori donné. Ce problème est alors résolu en utilisant la méthode de *régularisation*, provenant de la théorie de problèmes inverses mal posés [40].

Le deuxième problème, traité dans cette thèse est celui de la modélisation multidimensionnelle avec des processus de Lévy. Alors que la plupart d'applications financières nécessitent un modèle multidimensionnel permettant de prendre en compte la dépendance entre les actifs, la quasi-totalité de modèles paramétriques disponibles dans la littérature ne s'appliquent qu'au cas d'un seul actif. Les deux méthodes permettant de construire des modèles multidimensionnels à base de processus de Lévy que l'on trouve dans la littérature concernent soit le cas d'un mouvement brownien multivarié changé de temps par un processus de Lévy croissant unidimensionnel [79] soit le cas où toutes les composantes ont une intensité finie de sauts [65]. Dans les deux cas, la gamme de structures de dépendance que l'on peut obtenir est très limitée. La deuxième partie de cette thèse est alors consacrée au développement d'une méthodologie générale permettant de caractériser les structures de dépendance entre les composantes d'un processus de Lévy multidimensionnel et de construire des modèles multidimensionnels à base de processus de Lévy. Cet objectif est atteint grâce à l'introduction de la notion de *copule de Lévy*, qui peut être considérée comme l'analogue pour les processus de Lévy de la notion de copule, utilisée en statistique pour modéliser la dépendance entre les variables aléatoires réelles [56, 76].

Principaux résultats

Le premier chapitre de la thèse commence par une brève présentation des processus de Lévy, complétée d'un recueil de leurs principales propriétés, utilisées dans la suite de la thèse, et des modèles exp-Lévy, c'est-à-dire, les modèles où le prix d'une action est décrit par l'exponentielle d'un processus de Lévy. Dans la section suivante je passe en revue les différents modèles exp-

Lévy paramétriques, disponibles dans la littérature financière. La dernière section de ce chapitre décrit une méthode, due à Carr et Madan [23], permettant de valoriser les options européennes dans les modèles exp-Lévy à l'aide de la transformée de Fourier. La méthode est fondée sur l'observation suivante: si on soustrait du prix de l'option call sa valeur intrinsèque:

$$z_T(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+] - (1 - e^{k-rT})^+,$$

alors la quantité qui reste est, sous certaines conditions, intégrable et on peut évaluer sa transformée de Fourier:

$$\zeta_T(v) := \int_{-\infty}^{+\infty} e^{ivk} z_T(k) dk = e^{ivrT} \frac{\Phi_T(v-i) - 1}{iv(1+iv)},$$

où Φ_T est la fonction caractéristique de X_T. Les prix d'options peuvent donc être évalués en calculant la transformée de Fourier inverse de ζ_T.

Cette section est la seule du premier chapitre à contenir des résultats originaux. Premièrement, j'ai démontré qu'on peut diminuer considérablement l'erreur de troncature dans le calcul de la transformée de Fourier inverse en remplaçant la valeur intrinsèque de l'option par son prix dans le modèle de Black et Scholes: si on dénote

$$\tilde{z}_T(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+] - C_{BS}^{\Sigma}(k),$$

où $C_{BS}^{\Sigma}(k)$ est le prix Black et Scholes d'une option call avec volatilité Σ et log-strike k, et $\Phi_T^{\Sigma}(v) = \exp(-\frac{\Sigma^2 T}{2}(v^2 + iv))$, alors la transformée de Fourier de $\tilde{z}_T(k)$ est donnée par

$$\tilde{\zeta}_T(v) = e^{ivrT} \frac{\Phi_T(v-i) - \Phi_T^{\Sigma}(v-i)}{iv(1+iv)}.$$

Pour presque tous les modèles paramétriques, discutés dans la littérature cette quantité converge vers zéro plus vite que toute puissance de $|v|$ lorsque $|v| \to \infty$ (notons que $\zeta_T(v)$ converge seulement comme $|v|^{-2}$).

Le deuxième apport de cette section est le développement d'une méthode de contrôle d'erreur pour l'algorithme de Carr et Madan. Supposons que la transformée de Fourier inverse de $\tilde{\zeta}_T$ est approchée comme suit:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \tilde{\zeta}_T(v) dv = \frac{L}{2\pi(N-1)} \sum_{m=0}^{N-1} w_m \tilde{\zeta}_T(v_m) e^{-ikv_m} + \varepsilon_T + \varepsilon_D, \tag{1}$$

où ε_T est l'erreur de troncature, ε_D est l'erreur de discrétisation, $v_m = -L/2 + m\Delta$, $\Delta = L/(N-1)$ est le pas de discrétisation et w_m sont les poids qui correspondent à la méthode d'intégration choisie. Alors, notant le triplet caractéristique de X par (A, ν, γ), l'erreur de troncature peut être évaluée comme suit:

1. Si on suppose $A > 0$ alors l'erreur de troncature satisfait:

$$|\varepsilon_T| \leq \frac{8}{\pi T \Sigma^2 L^3} e^{-\frac{TL^2\Sigma^2}{8}} + \frac{8}{\pi T A L^3} e^{-\frac{TL^2 A}{8}}$$

2. Si on suppose que ν a une densité de la forme $\nu(x) = \frac{e^{-x}}{|x|} f(x)$, où f est une fonction croissante sur $(-\infty, 0)$ et décroissante sur $(0, \infty)$ alors l'erreur de troncature satisfait:

$$|\varepsilon_T| \leq \frac{8}{\pi T \Sigma^2 L^3} e^{-\frac{TL^2\Sigma^2}{8}} + \frac{|\Phi_T(L/2 - i)| + |\Phi_T(-L/2 - i)|}{\pi L}.$$

Si l'intégrale dans l'équation (1) est approchée en utilisant la méthode de trapèzes, l'erreur de discrétisation ε_D satisfait

$$|\varepsilon_D| \leq \frac{\Delta^2}{6\pi} \sum_{l=0}^{2} \frac{C_{3-l} + C^{\Sigma}_{3-l}}{(3-l)!} \times \left\{ \left(\Delta + \frac{\pi}{2}\right) e^{|k - rT|} + \log\left(\frac{L}{2} + \sqrt{\frac{L^2}{4} + 1}\right) \frac{|k - rT|^l}{l!} \right\},$$

où

$$C_k = \begin{cases} |\Phi_T^{(k)}(-i)|, & k \text{ pair} \\ |\Phi_T^{(k+1)}(-i)|^{\frac{k}{k+1}}, & k \text{ impair} \end{cases}$$

et C_k^{Σ} sont calculés en utilisant la même formule à partir de la fonction caractéristique Φ_T^{Σ}.

Calibration de modèles exponentielle-Lévy

Le deuxième chapitre est consacré au traitement théorique du problème de calibration. Je commence par discuter la calibration au sens de moindres carrés linéaires, qui est la méthode couramment utilisée par les praticiens et chercheurs. Dans cette approche, pour trouver une solution, on minimise la somme pondérée des carrés des écarts entre les prix de marché et les prix du modèle:

Calibration au sens de moindres carrés non-linéaires. Etant donnés les prix C_M des options cotées sur le marché financier, trouver un modèle exponentielle-Lévy risque neutre Q^*, tel que

$$\|C_M - C^{Q^*}\|_w^2 = \inf_Q \|C_M - C^Q\|_w^2,$$

où le inf est calculé sur tous les modèles exponentielle-Lévy risque neutres, C^Q dénote les prix calculés dans le modèle Q et

$$\|C_M - C^Q\|_w^2 := \sum_{i=1}^N w_i (C_M(T_i, K_i) - C^Q(T_i, K_i))^2.$$

L'ensemble des solutions du problème de calibration au sens de moindres carrés sera noté $\mathcal{Q}^{LS}$. Je commence par démontrer que dans le contexte de calibration non-paramétrique de modèles exp-Lévy cette méthode ne permet pas toujours de trouver une solution et est, dans certaines situations, instable par rapport aux perturbations des données d'entrée C_M. Ensuite je donne des conditions suffisantes (assez restrictives), sous lesquelles le problème de calibration admet une solution continue par rapport aux données d'entrée (Théorème 2.1 et Proposition 2.5). Comme ces conditions sont rarement vérifiées en pratique, il est nécessaire de trouver une méthode plus fiable de résolution du problème de calibration. Le premier pas dans cette direction est de le reformuler comme le problème de trouver un processus de Lévy risque neutre qui reproduit les prix d'options cotées avec la plus grande précision possible et qui a l'entropie relative minimale par rapport à un processus a priori à donné parmi toutes les solutions du problème de calibration au sens de moindres carrés:

Calibration au sens de moindres carrés avec minimisation d'entropie Etant donnés les prix C_M des options cotées sur le marché financier, et un processus de Lévy a priori P, trouver un modèle exponentielle-Lévy risque neutre $Q^* \in \mathcal{Q}^{LS}$, tel que

$$I(Q^*|P) = \inf_{Q \in \mathcal{Q}^{LS}} I(Q|P).$$

Les solutions du problème ci-dessus seront appelées solutions au sens de moindres carrés

d'entropie minimale. Le lien de cette approche avec d'autres méthodes de calibration et de valorisation d'options à l'aide de l'entropie relative, disponibles dans la littérature, est discuté dans section 2.3.

Cette formulation résout le problème de non-identifiabilité, mais souffre toujours de manque de stabilité. Cependant, elle peut être transformée en un problème de calibration stable et admettant toujours une solution, en utilisant la méthode de régularisation, provenant de la théorie de problèmes inverses mal-posés. Pour régulariser le problème de calibration au sens de moindres carrés, je propose de minimiser la somme de l'erreur de valorisation et d'un terme de pénalisation donné par l'entropie relative du modèle par rapport à P:

Problème de calibration régularisé Etant donnés les prix C_M des options cotées sur le marché et un processus de Lévy a priori P, trouver un modèle exponentielle-Lévy risque neutre Q^*, tel que

$$J_\alpha(Q^*) = \inf_Q J_\alpha(Q),$$

où le inf est calculé sur tous les modèles exponentielle-Lévy risque neutres,

$$J_\alpha(Q) = \|C_M - C^Q\|_w^2 + \alpha I(Q|P)$$

et α est le paramètre de régularisation qui détermine l'intensité de pénalisation.

Dans le cadre de processus de Lévy, l'entropie relative peut être calculée explicitement (Théorème 2.9):

$$I(Q|P) = I(Q|_{\mathcal{F}_{T_\infty}}|P|_{\mathcal{F}_{T_\infty}}) = \frac{T_\infty}{2A}\left\{\gamma^Q - \gamma^P - \int_{-1}^{1} x(\nu^Q - \nu^P)(dx)\right\}^2 1_{A\neq 0} + T_\infty \int_{-\infty}^{\infty} \left(\frac{d\nu^Q}{d\nu^P}\log\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx).$$

Cette formule permet d'étudier les propriétés de solutions du problème de calibration régularisé et démontrer les résultats suivants, sous condition que les sauts du processus a priori P sont bornés supérieurement par une constante B et que P correspond à un modèle exponentielle-Lévy sans opportunité d'arbitrage

- Le problème de calibration régularisé admet au moins une solution.

- Si P satisfait une condition de régularité supplémentaire (2.28), alors la solution Q est une mesure équivalente à P (en général, Q est une mesure martingale, absolument continue par rapport à P).

- Les solutions du problème de calibration régularisé sont continues par rapport aux données de marché: si $\{C_M^n\}_{n\geq 1}$ est une suite de données telle que $\|C_M^n - C_M\|_w \to 0$, et pour chaque n, Q_n est la solution du problème de calibration régularisé avec donnée C_M^n, loi a priori P et paramètre de régularisation α, alors $\{Q_n\}_{n\geq 1}$ a une sous-suite faiblement convergente, et la limite de chaque sous-suite faiblement convergente de $\{Q_n\}_{n\geq 1}$ est la solution du problème de calibration régularisé avec donnée C_M, loi a priori P et paramètre de régularisation α.

- Lorsque le niveau d'erreur dans les données de marché tend vers 0, si le paramètre de régularisation α est choisi de façon appropriée, alors les solutions du problème régularisé convergent vers les solutions au sens de moindres carrés d'entropie minimale (lorsqu'il existe plusieurs solutions au sens de moindres carrés d'entropie minimale, la convergence est entendue dans le même sens que ci-dessus).

Le troisième chapitre traite la résolution numérique du problème de calibration régularisé, construit au chapitre précédent. Tout d'abord, pour se ramener dans un espace fini-dimensionnel, la mesure de Lévy du processus a priori P est discrétisée sur une grille:

$$\nu^P = \sum_{k=0}^{M-1} p_k \delta_{\{x_k\}}(dx), \tag{2}$$

ce qui implique que la mesure de Lévy de la solution Q a la même forme:

$$\nu^Q = \sum_{k=0}^{M-1} q_k \delta_{\{x_k\}}(dx).$$

Ensuite, je démontre que la solution du problème de calibration avec une loi a priori quelconque peut être approchée avec une précision arbitraire par une suite de solutions avec lois a priori de la forme (2) (théorème 3.2 et lemme 3.1). Section 3.2 discute le choix du processus a priori et estime son influence sur la solution, en effectuant des tests numériques. La conclusion de cette section est que la solution est peu sensible aux petites variations du processus a priori, mais qu'il est important de spécifier sa forme qualitative correctement.

Section 3.3 présente un algorithme de choix du paramètre de régularisation α "a posteriori" à partir des données. Ma méthode, qui est une adaptation du principe de discrépance, développé par Morozov [74] pour la régularisation de Tikhonov dans les espaces de Banach, consiste à choisir deux constantes c_1 et c_2 proches de 1, telles que $1 < c_1 < c_2$ et trouver la valeur de paramètre α pour laquelle $c_1\delta^2 \leq \varepsilon_\delta(\alpha) \leq c_2\delta^2$, où δ est le niveau de bruit dans les données et la discrépance $\varepsilon_\delta(\alpha)$ est définie par

$$\varepsilon_\delta(\alpha) := \|C^{Q_\alpha^\delta} - C_M^\delta\|^2,$$

avec Q_α^δ la solution du problème régularisé avec paramètre de régularisation α et niveau de bruit δ.

Section 3.5 est consacrée à l'algorithme numérique permettant de résoudre le problème de calibration discrétisé, une fois que le processus a priori et le paramètre de régularisation ont été choisis. Le problème est résolu en minimisant la fonctionnelle de calibration approchée $\hat{J}_\alpha(Q)$ par la méthode BFGS, faisant intervenir le gradient de la fonctionnelle à minimiser. La fonctionnelle approchée est donnée par

$$\hat{J}_\alpha(Q) = \sum_{i=1}^{N} w_i(\hat{C}^Q(T_i,K_i) - C_M(T_i,K_i))^2 + \frac{\alpha}{2A}\left(\frac{A}{2} + b^P + \sum_{j=0}^{M-1}(e^{x_j}-1)q_j\right)^2 + \alpha\sum_{j=0}^{M-1}\left(q_j\log(q_j/p_j) + 1 - q_j\right),$$

où $\hat{C}^Q(T_i,K_i)$ est l'approximation du prix d'une option call avec maturité T_i et prix d'exercice K_i, calculée dans le modèle Q en utilisant la méthode de transformée de Fourier. Le gradient de la fonctionnelle de calibration est évalué analytiquement en utilisant l'équation (3.32).

Dans section 3.6 l'algorithme est appliqué d'abord aux données simulées (à partir d'un modèle exp-Lévy avec une mesure de Lévy connue) et ensuite aux données réelles. Les tests sur des données artificielles mettent en évidence la stabilité de l'algorithme et sa capacité de reconstruire la vraie mesure de Lévy partout sauf dans un petit voisinage de 0 (parce que les sauts de taille 0 n'ont pas d'influence sur les prix d'options). Les tests sur les données réelles permettent de tirer un nombre de conclusions importantes.

- Premièrement, un modèle exp-Lévy permet de calibrer des prix d'options d'une seule maturité avec une grande précision. Cette conclusion contredit les résultats de Medvedev

et Scaillet [70] qui ont observé que “les sauts dans les prix ne permettent pas d'expliquer la pente de volatilité implicite à la monnaie”. Toutefois, il est important de remarquer que l'étude de Medvedev et Scaillet porte sur les options sur l'indice S&P 500 de courte maturité, tandis que les tests de cette thèse ont été effectués en utilisant les options sur l'indice DAX.

- De plus, la qualité de calibration est déjà excellente en utilisant seulement des modèles à intensité finie de sauts, ce qui remet en question la nécessité, du point de vue de modélisation du smile de volatilité, des modèles plus compliqués à intensité infinie.

- La troisième conclusion est que même dans le cadre non-paramétrique, il est impossible, en utilisant un modèle exp-Lévy, de calibrer des prix d'options de plusieurs maturités avec une précision satisfaisante. Les tests ont donc confirmé l'observation déjà faite par certains auteurs [13, 68] que le cadre de modèles exp-Lévy est trop rigide pour pouvoir décrire correctement la structure par terme de volatilités implicites.

Modélisation multidimensionnelle avec les processus de Lévy

Le quatrième chapitre commence par une revue de deux méthodes, disponibles dans la littérature, permettant de construire des modèles multidimensionnels à base de processus de Lévy. Dans la première approche il s'agit de changer l'échelle de temps d'un mouvement brownien multivarié par un processus de Lévy croissant, le même pour toutes les composantes [79], et la deuxième méthode, qui ne concerne que le cas de processus de Lévy à intensité finie de sauts, consiste à modéliser directement la dépendance entre les tailles des sauts dans les différentes composantes, en utilisant les copules. Malheureusement, dans ces méthodes la gamme de types de dépendance possibles est très limitée et des contraintes sont imposées sur le choix de modèles paramétriques pour les composantes (on ne peut pas coupler des composantes quelconques). La discussion des défauts de ces méthodes permet d'établir une liste de conditions qu'une bonne méthode de modélisation multidimensionnelle doit satisfaire:

- On doit pouvoir choisir librement un processus de Lévy unidimensionnel pour chacune de composantes.

- La gamme de structures de dépendance possibles doit inclure l'indépendance et la dépendance complète avec une transition continue entre ces deux cas extrêmes.
- La méthode doit permettre de construire des modèles de dépendance paramétriques.

Pour atteindre ces conditions, je propose de séparer la structure de dépendance d'un processus de Lévy des lois marginales de ces composantes. Dans le cadre plus restreint de variables aléatoires à valeurs dans $\mathbb{R}^d$, cette séparation est réalisée par la notion de copule. Je développe une notion analogue, applicable au cas des processus de Lévy.

Section 4.4 introduit les copules de Lévy dans le cas plus simple de processus de Lévy n'admettant que des sauts positifs dans chacune de composantes. Dans ce cas la copule de Lévy est définie comme une fonction $F : [0,\infty]^d \to [0,\infty]$ avec les propriétés suivantes:

1. $F(u_1,\dots,u_d) = 0$ dès que $u_i = 0$ pour au moins un $i \in \{1,\dots,d\}$,
2. F est une fonction croissante en d dimensions, c.-à-d., pour tout $a, b \in \operatorname{Dom} F$,
$$\sum (-1)^{N(c)} F(c) \geq 0,$$
où la somme est calculé sur tous les sommets c de $[a,b]$ et $N(c) := \#\{k : c_k = a_k\}$.
3. F a les fonctions marginales uniformes: $F(u_1,\dots,u_d)|_{u_i=\infty, i\neq k} = u_k$ pour tout $k \in \{1,\dots,d\}$, $u_k \in [0,\infty]$.

En représentant une mesure de Lévy ν sur $\mathbb{R}^d_+$ par son intégrale de queue:

$$U(x_1,\dots,x_d) = \nu([x_1,\infty) \times \cdots \times [x_d,\infty)),$$

je démontre que

- Pour toute mesure de Lévy ν sur $\mathbb{R}^d_+$ ayant l'intégrale de queue U et les mesures de Lévy marginales $\nu_1,\dots,\nu_d$, il existe une copule de Lévy F sur $[0,\infty]^d$, telle que
$$U(x_1,\dots,x_d) = F(U_1(x_1),\dots,U_d(x_d)), \quad (x_1,\dots,x_d) \in [0,\infty)^d, \tag{3}$$
où $U_1,\dots,U_d$ sont des intégrales de queue de $\nu_1,\dots,\nu_d$.
- Si F est une copule de Lévy sur $[0,\infty]^d$ et $\nu_1,\dots,\nu_d$ sont des mesures de Lévy sur $(0,\infty)$ avec les intégrales de queue $U_1,\dots,U_d$, alors équation (3) définit une intégrale de queue d'une mesure de Lévy sur $\mathbb{R}^d_+$ ayant les mesures de Lévy marginales $\nu_1,\dots,\nu_d$.

Ce résultat montre que pour chaque processus de Lévy il existe une copule de Lévy qui décrit sa structure de dépendance, et que pour chaque copule de Lévy et chaque combinaison de lois unidimensionnelles il existe un processus de Lévy avec des lois de composantes données et dont la dépendance est décrite par cette copule de Lévy.

Section 4.5 étend la notion de copule de Lévy et les résultats associés aux processus de Lévy généraux. Les résultats de cette section ont été obtenus dans un travail joint avec Jan Kallsen [59]. En particulier, la structure de dépendance d'un processus de Lévy général est caractérisé par une copule de Lévy sur $(-\infty,\infty]^d$, c.-à-d., une fonction $F : (-\infty,\infty]^d \to (-\infty,\infty]$ avec les propriétés suivantes:

1. $F(u_1,\dots,u_d) < \infty$ si $(u_1,\dots,u_d) \neq (\infty,\dots,\infty)$,
2. $F(u_1,\dots,u_d) = 0$ si $u_i = 0$ pour au moins un $i \in \{1,\dots,d\}$,
3. F est une fonction croissante en d dimensions,
4. F a les fonctions marginales uniformes: pour tout $k \in \{1,\dots,d\}$ et pour tout $x_k \in (-\infty,\infty]$,

$$\lim_{c\to\infty} \sum_{(x_j)_{j\neq k}\in\{-c,\infty\}^{d-1}} F(x_1,\dots,x_d) \prod_{j\neq k} \operatorname{sgn} x_j = x_k.$$

Dans la dernière section de ce chapitre je calcule les copules de Lévy qui correspondent aux types de dépendance particuliers. Les composantes d'un processus de Lévy multidimensionnel sans partie martingale continue sont indépendantes si et seulement s'il a une copule de Lévy suivante:

$$F_\perp(x_1,\dots,x_d) := \sum_{i=1}^{d} x_i \prod_{j\neq i} 1_{\{\infty\}}(x_j).$$

Les sauts d'un processus de Lévy d-dimensionnel sont dits complètement dépendants s'il existe un sous-ensemble strictement ordonné S de $K := \{x \in \mathbb{R}^d : \operatorname{sgn} x_1 = \cdots = \operatorname{sgn} x_d\}$, tel que $\Delta X_t := X_t - X_{t-} \in S$, $t \geq 0$. La copule de Lévy de la dépendance complète est

$$F_\parallel(x_1,\dots,x_d) := \min(|x_1|,\dots,|x_d|) 1_K(x_1,\dots,x_d) \prod_{i=1}^{d} \operatorname{sgn} x_i.$$

Le dernier chapitre de la thèse est dédié aux applications de copules de Lévy en finance. Dans les applications on n'a souvent pas assez d'information sur la dépendance pour employer

des méthodes non-paramétriques et je commence donc par donner des outils permettant de construire des familles paramétriques de copules de Lévy. Un exemple de famille de copules de Lévy en deux dimensions, dépendant de deux paramètres est donné par

$$F(u,v) = (|u|^{-\theta} + |v|^{-\theta})^{-1/\theta}(\eta 1_{uv\geq 0} - (1-\eta)1_{uv<0}). \tag{4}$$

Pour faire des calculs en utilisant la méthode de Monte Carlo il est crucial de pouvoir simuler les processus de Lévy avec la dépendance donnée par une copule de Lévy. Je présente une méthode de simulation de tels processus en utilisant une représentation en série qui peut être considérée comme une extension de résultats de Rosinski [84]. Chaque copule de Lévy F sur $(-\infty,\infty]^d$ définit une mesure positive μ sur $\mathbb{R}^d$, qui peut être décomposée comme

$$\mu(dx_1 \dots dx_d) = \lambda(dx_1) \otimes K(x_1, dx_2 \cdots dx_d),$$

où $K(x_1, dx_2 \cdots dx_d)$ est, pour presque tout x, une mesure de probabilité sur $\mathbb{R}^{d-1}$. Un processus de Lévy d-dimensionnel avec mesure de Lévy ν, telle que $\int(|x| \wedge 1)\nu(dx) < \infty$ et intégrales de queue marginales $U_i, i = 1, \dots, d$ peut alors être simulé comme suit.

- Soit $\{\Gamma_i^1\}_{i\geq 1}$ une suite de variables aléatoires telle que $\sum_{i=1}^{\infty} \delta_{\Gamma_i^1}$ est une mesure aléatoire de Poisson sur $\mathbb{R}$ ayant la mesure de Lebesgue comme mesure d'intensité.
- Pour chaque i, soit $(\Gamma_i^2, \dots, \Gamma_i^d)$ un vecteur aléatoire, qui est, conditionnellement à Γ_i^1, indépendant de Γ_j^k avec $i \neq j$ et tout k et distribué selon la loi $K(\Gamma_i^1, *)$.
- Soit $\{V_i\}_{i\geq 1}$ une suite indépendante de variables aléatoires indépendantes, distribuées uniformément sur $[0,1]$.

Alors

$$\{Z_t\}_{0\leq t\leq 1} \quad \text{où} \quad Z_t^k = \sum_{i=1}^{\infty} U_i^{(-1)}(\Gamma_i^k) 1_{[0,t]}(V_i), \quad k = 1, \dots, d,$$

est un processus de Lévy sur l'intervalle de temps $[0,1]$ avec fonction caractéristique

$$e^{i\langle u, Z_t\rangle} = \exp\left(t \int_{\mathbb{R}^d} (e^{i\langle u,z\rangle} - 1)\nu(dz)\right).$$

Enfin je construis un modèle exp-Lévy bidimensionnel paramétrique avec la structure de dépendance donnée par la copule de Lévy ((4)) et les composantes unidimensionnelles qui sont des processus variance gamma, et je montre comment on peut valoriser des options sur panier

dans un tel modèle en utilisant la méthode de Monte Carlo. Pour estimer l'importance de la dépendance dans les queues de distribution pour la valorisation des options, je prends deux jeux de paramètres η et θ, qui correspondent tous les deux à la corrélation de rendements de 50% mais ont les structures de dépendance différentes. Les prix des options sur panier pour les deux jeux de paramètres diffèrent de 10% à la monnaie, ce qui montre que la dépendance dans les queues de distribution est importante à prendre en compte pour la valorisation d'options sur panier, et que les copules de Lévy sont un outil adapté pour décrire cette dépendance.

Introduction

Lévy processes are defined as stochastic processes with stationary and independent increments: if $\{X_t\}_{t\geq 0}$ is a Lévy process, then $X_t - X_s$ with $t > s$ is independent of the history of the process up to time s, and its law only depends on $t - s$ but not on t or s separately. This property of increments suggests an analogy with linear functions: one can say that Lévy processes are, in some sense, "linear processes".

Despite this apparent simplicity, Lévy processes have many interesting properties and constitute an exciting field of study: this is shown by the recent publication of several monographs (see for example [17, 87]) and by the success of the series of international conferences on Lévy processes and applications (see [6]).

From the point of view of financial modelling, Lévy processes provide a class of models with jumps that is both sufficiently rich to reproduce empirical data and simple enough to do many computations analytically. The interest of jump models in finance is mainly due to three reasons.

First, in a model with continuous paths like a diffusion model, the price process behaves locally like a Brownian motion and the probability that the stock moves by a large amount over a short period of time is very small, unless one fixes an unrealistically high value of the volatility of volatility. Therefore, in such models the prices of short term out of the money options should be much lower than what one observes in real markets. On the other hand, if stock prices are allowed to jump, even when the time to maturity is very short, there is a non-negligible probability that after a sudden change in the stock price the option will move in the money.

Second, from the hedging point of view, continuous models of stock price behavior generally lead to a complete market or to a market that can be made complete by adding one or two

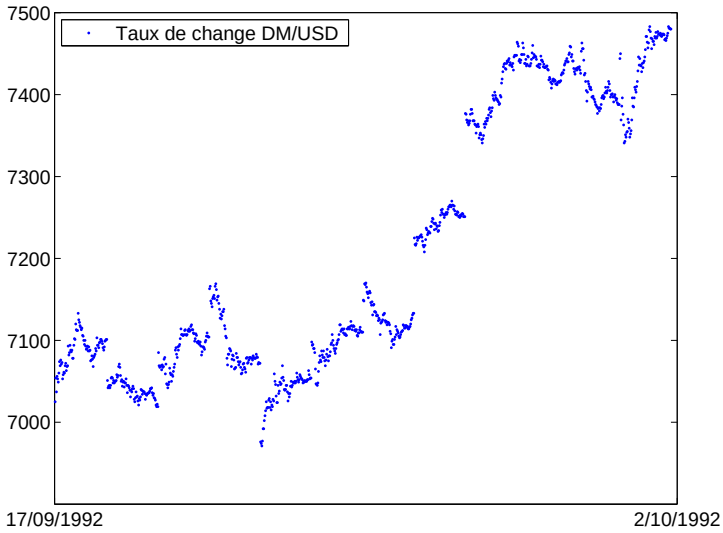

Figure 1: Jumps in the trajectory of DM/USD exchange rate, sampled at 5-minute intervals.

additional instruments, like in stochastic volatility models. Since in such a market every terminal payoff can be exactly replicated, the very existence of traded options becomes a puzzle. The mystery is easily solved by allowing for discontinuities: in real markets, due to the presence of jumps in the prices, perfect hedging is impossible and options enable the market participants to hedge risks that cannot be hedged by using the underlying only.

The third and the strongest argument for using discontinuous models is simply the presence of jumps in observed prices. Figure 1 depicts the evolution of the DM/USD exchange rate over a two-week period in 1992, and one can see at least three points where the rate moved by over 100 bp within a 5-minute period. Price moves like these ones clearly cannot be accounted for in the framework of a diffusion model with continuous paths, but they must be dealt with if the market risk is to be measured and managed correctly.

Although when this thesis was started, the field of financial modelling with Lévy processes was already a well-developed one, with several Ph. D. theses [78, 79, 81, 86, 96] and a few hundred research papers (see references in [27]) already written on the subject, two major issues, that appear in the title of the present study, remained unresolved.

First, while the main concern in the literature has been to find efficient analytical and numerical procedures for computing option prices in exponential Lévy models, an essential step in using such models is to obtain the parameters — here the characteristic triplet of the

underlying Lévy process — consistent with the market-quoted prices of traded options. This problem — called model calibration — is an inverse problem to that of pricing a European option in an exponential Lévy model and is much more difficult to solve than the latter. The calibration problem was addressed by several authors (see for example [4, 35, 53, 85]) in the framework of a markovian diffusion model, the unknown parameter being the local volatility function $\sigma(S_t, t)$. However, in the setting of processes with jumps, although many papers proposed parametric Lévy-based models [5, 22, 36, 62, 67, 71], prior to this study there existed no systematic approach to model selection, and no stable calibration method, that could be implemented in practice was available. In the first part of the thesis we develop a non-parametric method allowing to calibrate exponential-Lévy models, study its stability and convergence properties, describe its numerical implementation and give examples of its use. Our approach is first to reformulate the calibration problem as that of finding a risk-neutral exponential Lévy model that reproduces the observed option prices with the best possible precision and has the smallest relative entropy with respect to a given prior, and then to solve this problem via the *regularization* approach, used in the theory of ill-posed inverse problems [40].

The second issue, addressed in this thesis, is that of multidimensional modelling with Lévy process. Although most financial applications (basket option pricing, risk management, portfolio optimization etc.) require a multidimensional model taking into account the dependence between various assets, the vast majority of parametric models, available in the literature cover only the case of a single asset. The only attempts to construct multidimensional Lévy-based models that were made prior to this study concern either the case of a multivariate Brownian motion, time-changed with a one-dimensional increasing Lévy process [79] or the case where all components have finite jump intensity [65]. In both cases the scope of dependence structures that one can obtain is quite limited. In the second part of this thesis we therefore propose a general methodology to characterize dependence among the components of multidimensional Lévy processes and to construct multidimensional exp-Lévy models. This is done by introducing the notion of *Lévy copula*, which can be seen as an analog for Lévy processes of the notion of copula, used in statistics to model dependence between real random variables [56, 76].

This thesis is structured as follows. The first chapter contains a brief review of the properties of Lévy processes and exponential Lévy models, that are used in the sequel. The last section describes a method, due to Carr and Madan [23], for pricing European options in exponential

Lévy models by Fourier transform. This section is the only one of the first chapter to contain new results: we propose original improvements to this method and obtain estimates for the truncation and discretization errors, that are not given in the original reference [23].

The second chapter is dedicated to the theoretical treatment of the calibration problem. We start by discussing the least squares calibration method, commonly used by academics and practitioners. We show that in the context of non-parametric calibration of exponential Lévy models, least squares calibration does not always allow to find a solution and even if it does, the solution is typically very sensitive to small perturbations of input data. To solve the calibration problem in a stable manner, we first reformulate it as the problem of finding an exponential Lévy model that has the smallest relative entropy with respect to the prior among all solutions of the least squares calibration problem. This problem, called minimal entropy least squares calibration, is still ill-posed, so we regularize it using the technique of *minimal entropy regularization*, from the theory of ill-posed inverse problems.

The third chapter discusses the numerical implementation of our calibration algorithm. To solve the calibration problem numerically, it is expressed in terms of the characteristic triplets of the prior and the solution and the Lévy measure of the prior is discretized on a uniform grid so that the calibration problem becomes finite-dimensional.

In the fourth chapter we first review the two available methods to model dependence between components of Lévy processes. Understanding the drawbacks of these methods allows to formulate the desirable properties of a multidimensional modelling approach. These properties lead us to defining the notion of *Lévy copula*, first in the case of Lévy processes with only positive jumps in every component and then in the general case. After proving a representation theorem, which shows that Lévy copulas completely characterize dependence structures of Lévy processes, we compute Lévy copulas that correspond to various special types of dependence.

The fifth and the last chapter provides the tools necessary to apply Lévy copulas to finance. We first give methods to construct parametric families of Lévy copulas and then develop an efficient algorithm to simulate multidimensional Lévy processes with dependence structures given by Lévy copulas. The last section of this chapter contains an example of a multidimensional exponential Lévy model for option pricing, constructed using Lévy copulas.

Part I

Calibration of one-dimensional exp-Lévy models

Chapter 1

Lévy processes and exp-Lévy models

This introductory chapter serves essentially two purposes. First, in Section 1.1, we give an overview of the probabilistic properties of Lévy processes that will be used in the sequel. Second, we define exponential Lévy models (Section 1.2), review various parametrizations of the Lévy measure, proposed by other authors (Section 1.3) and discuss a method for option pricing in these models, based on Fourier transform (Section 1.4). This option pricing method is later used for the numerical solution of the calibration problem. While the results of the first three sections can be found in the literature, Section 1.4 contains new material: we improve the method due to Carr and Madan [23] in a number of ways and provide estimates of the truncation and discretization error, not found in the original paper.

1.1 Lévy processes

Proofs of the the results of this section can be found, unless otherwise mentioned in [87]. For additional details on Lévy processes the reader may consult [17] or [54]. The latter book treats a far more general class of semimartingales but properties of Lévy processes are often discussed as examples or corollaries of the general results.

Definition 1.1 (Lévy process). A stochastic process $\{X_t\}_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ such that $X_0 = 0$ is called a *Lévy process* if it possesses the following properties:

1. Independent increments: for every increasing sequence of times $t_0 \dots t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent.

2. Stationary increments: for every $h > 0$, the law of $X_{t+h} - X_t$ does not depend on t.

The law of a Lévy process is completely identified by its characteristic triplet (A, ν, γ), where A is a symmetric nonnegative-definite $d \times d$ matrix, $\gamma \in \mathbb{R}^d$ and ν is a Lévy measure, that is, a positive measure on $\mathbb{R}^d \setminus \{0\}$, satisfying

$$\int_{\mathbb{R}^d \setminus \{0\}} (|x|^2 \wedge 1)\nu(dx) < \infty.$$

In particular, the characteristic function of X_t can be computed from this triplet as follows.

Theorem 1.1 (Lévy-Khintchine representation). *Let $\{X_t\}_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A, ν, γ). Then*

$$\begin{aligned} &E[e^{i\langle z, X_t\rangle}] = e^{t\psi(z)}, \ z \in \mathbb{R}^d \\ &\text{with} \quad \psi(z) = -\frac{1}{2}\langle z, Az\rangle + i\langle \gamma, z\rangle + \int_{\mathbb{R}^d} (e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle 1_{|x|\leq 1})\nu(dx). \end{aligned} \tag{1.1}$$

Remark 1.1. Let $h : \mathbb{R}^d \to \mathbb{R}^d$ be a measurable function, such that for every z, $e^{i\langle z,x\rangle} - 1 - i\langle z, h(x)\rangle$ is integrable with respect to ν. Then the second equation in the Lévy-Khintchine formula (1.1) may be rewritten as

$$\begin{aligned} \psi(z) &= -\frac{1}{2}\langle z, Az\rangle + i\langle \gamma_h, z\rangle + \int_{\mathbb{R}^d} (e^{i\langle z,x\rangle} - 1 - i\langle z, h(x)\rangle)\nu(dx), \\ \text{where} \quad \gamma_h &= \gamma + \int_{-\infty}^{\infty} (h(x) - x1_{|x|\leq 1})\nu(dx). \end{aligned}$$

The triplet (A, ν, γ_h) is called the characteristic triplet of $\{X_t\}_{t\geq 0}$ with respect to the truncation function h. For instance, if $\int_{\mathbb{R}^d}(|x| \wedge 1)\nu(dx) < \infty$, one may take $h \equiv 0$ and the Lévy-Khintchine representation becomes

$$\psi(z) = -\frac{1}{2}\langle z, Az\rangle + i\langle \gamma_0, z\rangle + \int_{\mathbb{R}^d} (e^{i\langle z,x\rangle} - 1)\nu(dx).$$

The vector γ_0 is in this case called *drift* of the process X.

The tail behavior and moments of the distribution of a Lévy process at a given time are determined by the Lévy measure, as shown by the following proposition (see Proposition 2.5 and Theorems 25.3 and 25.17 in [87]).

Proposition 1.2 (Moments and cumulants of a Lévy process).

1. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A,ν,γ) and let $n \geq 1$. $E[|X_t|^n] < \infty$ for some $t > 0$ or equivalently for every t if and only if $\int_{|x|\geq 1} |x|^n \nu(dx) < \infty$. In this case $\Phi_t(z)$, the characteristic function of X_t, is of class C^n and the first n moments of X_t can be computed by differentiation:*

$$E[X_t^k] = \frac{1}{i^k}\frac{\partial^k}{\partial z^k}\Phi_t(z)|_{z=0}, \quad k = 1,\ldots,n.$$

The cumulants of X_t, defined by

$$c_k(X_t) := \frac{1}{i^k}\frac{\partial^k}{\partial z^k}\log \Phi_t(z)|_{z=0},$$

have a particularly simple structure:

$$\begin{aligned} c_1(X_t) &\equiv E[X_t] = t(\gamma + \int_{|x|\geq 1} x\nu(dx)),\\ c_2(X_t) &\equiv \operatorname{Var} X_t = t(A + \int_{-\infty}^{\infty} x^2\nu(dx)),\\ c_k(X_t) &= t\int_{-\infty}^{\infty} x^k\nu(dx) \quad \text{for } 3 \leq k \leq n. \end{aligned}$$

2. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A,ν,γ) and let $u \in \mathbb{R}$. $E[e^{uX_t}] < \infty$ for some t or, equivalently, for all $t > 0$ if and only if $\int_{|x|\geq 1} e^{ux}\nu(dx) < \infty$. In this case*

$$E[e^{uX_t}] = e^{t\psi(-iu)}.$$

where ψ is the characteristic exponent of the Lévy process defined by (1.1).

Corollary 1.1. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A,ν,γ).*

1. *$\{X_t\}_{t\geq 0}$ is a martingale if and only if $\int_{|x|\geq 1} |x|\nu(dx) < \infty$ and*

$$\gamma + \int_{|x|\geq 1} x\nu(dx) = 0.$$

2. *$\{\exp(X_t)\}_{t\geq 0}$ is a martingale if and only if $\int_{|x|\geq 1} e^x\nu(dx) < \infty$ and*

$$\frac{A}{2} + \gamma + \int_{-\infty}^{\infty} (e^x - 1 - x1_{|x|\leq 1})\nu(dx) = 0. \tag{1.2}$$

1.1.1 Stochastic exponential of Lévy processes

The stochastic (Doléans-Dade) exponential of a semimartingale $\{X_t\}_{t\geq 0}$ (see [54, Theorem I.4.61]) is defined as the unique càdlàg process $\{Z_t\}_{t\geq 0}$ such that

$$dZ_t = Z_{t-}dX_t, \qquad Z_0 = 1. \tag{1.3}$$

Z is given by:

$$Z_t = e^{X_t - X_0 - \frac{1}{2}[X]^c_t} \prod_{0\leq s\leq t} (1+\Delta X_s)e^{-\Delta X_s}. \tag{1.4}$$

The stochastic exponential of X is denoted by $\mathcal{E}(X)$. The following result, due to Goll and Kallsen [47] clarifies the relation between the stochastic and the ordinary exponential when X is a Lévy process.

Proposition 1.3 (Relation between ordinary and stochastic exponentials).

1. *Let $\{X_t\}_{t\geq 0}$ be a real-valued Lévy process with characteristic triplet (A,ν,γ) and $Z=\mathcal{E}(X)$ its stochastic exponential. If $Z>0$ a.s. then there exists another Lévy process $\{L_t\}_{t\geq 0}$ such that for all t, $Z_t = e^{L_t}$, where*

$$L_t = \log Z_t = X_t - \frac{At}{2} + \sum_{0\leq s\leq t} \left\{\log(1+\Delta X_s) - \Delta X_s\right\}. \tag{1.5}$$

Its characteristic triplet (A_L,ν_L,γ_L) is given by:

$$\begin{aligned} &A_L = A, \\ &\nu_L(B) = \nu(\{x : \log(1+x)\in B\}) = \int_{-\infty}^{\infty} 1_B(\log(1+x))\nu(dx), \quad B\in\mathcal{B}(\mathbb{R}\setminus\{0\}), \\ &\gamma_L = \gamma - \frac{A}{2} + \int_{-\infty}^{\infty} \nu(dx)\left\{\log(1+x)1_{[-1,1]}(\log(1+x)) - x1_{[-1,1]}(x)\right\}. \end{aligned} \tag{1.6}$$

2. *Let $\{L_t\}_{t\geq 0}$ be a real-valued Lévy process with characteristic triplet (A_L,ν_L,γ_L) and $S_t = \exp L_t$ its exponential. Then there exists a Lévy process $\{X_t\}_{t\geq 0}$ such that S_t is the stochastic exponential of X: $S=\mathcal{E}(X)$ where*

$$X_t = L_t + \frac{At}{2} + \sum_{0\leq s\leq t} \left\{e^{\Delta L_s} - 1 - \Delta L_s\right\}. \tag{1.7}$$

The characteristic triplet (A, ν, γ) of X is given by:

$$\begin{aligned} A &= A_L, \\ \nu(B) &= \nu_L(\{x : e^x - 1 \in B\}) = \int_{-\infty}^{\infty} 1_B(e^x - 1)\nu_L(dx), \quad B \in \mathcal{B}(\mathbb{R} \setminus \{0\}), \\ \gamma &= \gamma_L + \frac{A_L}{2} + \int_{-\infty}^{\infty} \nu_L(dx) \left\{(e^x - 1)1_{[-1,1]}(e^x - 1) - x1_{[-1,1]}(x)\right\}. \end{aligned} \tag{1.8}$$

Equation (1.3) shows that the stochastic exponential of a martingale is necessarily a local martingale. For Lévy processes, however, a stronger result is true, namely, the stochastic exponential of a martingale Lévy process is a (true) martingale. This fact seems to belong to the folklore of the Lévy process theory but we have been unable to find its proof in the literature.

Proposition 1.4 (Martingale preserving property). *If $\{X_t\}_{t\geq 0}$ is a Lévy process and a martingale, then its stochastic exponential $Z = \mathcal{E}(X)$ is also a martingale.*

Proof. Let $\{X_t\}_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with triplet (A, ν, γ) such that $\gamma + \int_{|x|\geq 1} x\nu(dx) = 0$ (this implies that X is a martingale). First, suppose that $|\Delta X_s| \leq \varepsilon < 1$ a.s. Then

$$\log(1 + \Delta X_s) - \Delta X_s \geq -\frac{\Delta X_s^2}{2(1-\varepsilon)^2}$$

and Equation (1.4) implies that

$$Z_t \geq \exp\left(X_t - X_0 - \frac{1}{2}[X]_t^c - \frac{1}{2(1-\varepsilon)^2}\sum_{0\leq s\leq t} \Delta X_s^2\right) > 0 \quad \text{a.s.}$$

Therefore by Proposition 1.3 there exists a Lévy process L such that $e^{L_t} = Z_t$ for all t. Moreover, this process has bounded jumps and therefore admits all exponential moments. Again, by Proposition 1.3, we can write:

$$\begin{aligned} \gamma_L + \frac{A_L}{2} + \int_{-\infty}^{\infty} (e^z - 1 - z1_{|z|\leq 1})\nu_L(dz) &= \gamma + \int_{-1}^{1} \{z\nu_L(dz) - z\nu(dz)\} \\ &+ \int_{-\infty}^{\infty} (e^z - 1 - z1_{|z|\leq 1})\nu_L(dz) = \int_{-\infty}^{\infty} \{(e^z - 1)\nu_L(dz) - z\nu(dz)\} = 0 \end{aligned}$$

because $\Delta X_s = e^{\Delta L_s} - 1$ for all s. Therefore, by Corollary 1.1, $Z_t = e^{L_t}$ is a martingale.

The second step is to prove the proposition when X is a compensated compound Poisson process. In this case, the stochastic exponential has a very simple form:

$$Z_t = e^{bt} \prod_{0\leq s\leq t} (1 + \Delta X_s),$$

where $b = -\int_{-\infty}^{\infty} x\nu(dx)$. Denoting the jump intensity of X by λ, and the generic jump of X by ΔX, we obtain, conditioning on the number of jumps of X in the interval $[0,t]$:

$$E[Z_t] = e^{-\lambda t + bt} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} (1 + E[\Delta X])^n = 1.$$

Because the increments of X are independent and stationary, this proves that Z is a martingale.

Now let X be an arbitrary martingale Lévy process. It can be decomposed into a sum of a compensated compound Poisson process X' and an independent martingale Lévy process with jumps smaller than ε, denoted by X''. Since these two processes never jump together, $\mathcal{E}(X' + X'') = \mathcal{E}(X')\mathcal{E}(X'')$ (cf. Equation II.8.19 in [54]). Moreover, each of the factors is a martingale and they are independent, hence $\mathcal{E}(X' + X'')$ is a martingale. □

1.1.2 Change of measure and absolute continuity of Lévy processes

The following proposition (see [54], Theorem IV.4.39) provides a criterion of absolute continuity of one Lévy process with respect to another in terms of their characteristic triplets.

Proposition 1.5 (Absolute continuity of Lévy processes). *Let $\{X_t\}_{t\geq 0}$ be a real-valued Lévy process on $(\Omega, \mathcal{F}, Q)$ and on $(\Omega, \mathcal{F}, P)$ with respective characteristic triplets (A_Q, ν_Q, γ_Q) and (A_P, ν_P, γ_P). Then $Q|_{\mathcal{F}_t} \ll P|_{\mathcal{F}_t}$ for all $t \in [0, T_\infty]$ or, equivalently, $Q|_{\mathcal{F}_t} \ll P|_{\mathcal{F}_t}$ for some $t \in (0, T_\infty]$ if and only if the following conditions are satisfied:*

1. $A_Q = A_P := A$.
2. $\nu_Q \ll \nu_P$.
3. $\int_{-\infty}^{\infty} (\sqrt{\phi(x)} - 1)^2 \nu_P(dx) < \infty$, *where* $\phi := \frac{d\nu_Q}{d\nu_P}$.
4. $\int_{|x|\leq 1} |x(1 - \phi(x))| \nu_P(dx) < \infty$.
5. *If* $A = 0$ *then* $\gamma_Q - \gamma_P = \int_{|x|\leq 1} x(\phi(x) - 1)\nu_P(dx)$.

Remark 1.2. In fact, condition 4 above is a consequence of condition 3. Indeed, on the set $\{x : \phi(x) > 4\}$, we have $(\sqrt{\phi(x)} - 1)^2 > 1$ and $\frac{(\sqrt{\phi(x)}-1)^2}{\phi(x)} > \frac{1}{4}$. Therefore, condition 3 implies

$$\int_{\{\phi(x)>4\}} \nu_P(dx) \leq \int_{\{\phi(x)>4\}} (\sqrt{\phi(x)} - 1)^2 \nu_P(dx) < \infty$$
$$\text{and} \quad \int_{\{\phi(x)>4\}} \phi(x)\nu_P(dx) < 4 \int_{\{\phi(x)>4\}} (\sqrt{\phi(x)} - 1)^2 \nu_P(dx) < \infty.$$

On the other hand,

$$\int_{|x|\leq 1} |x(1-\phi(x))|\nu_P(dx) = \int_{|x|\leq 1;\phi(x)>4} |x(1-\phi(x))|\nu_P(dx) + \int_{|x|\leq 1;\phi(x)\leq 4} |x(1-\phi(x))|\nu_P(dx).$$

The first term in the right-hand side satisfies:

$$\int_{|x|\leq 1;\phi(x)>4} |x(1-\phi(x))|\nu_P(dx) \leq \int_{|x|\leq 1;\phi(x)>4} \nu_P(dx) + \int_{|x|\leq 1;\phi(x)>4} \phi(x)\nu_P(dx) < \infty,$$

and for the second term one can write:

$$\begin{aligned}\int_{|x|\leq 1;\phi(x)\leq 4} &|x(1-\phi(x))|\nu_P(dx) \\ &\leq \int_{|x|\leq 1;\phi(x)\leq 4} |x|^2\nu_P(dx) + \int_{|x|\leq 1;\phi(x)\leq 4} (\sqrt{\phi(x)}-1)^2(\sqrt{\phi(x)}+1)^2\nu_P(dx) < \infty.\end{aligned}$$

A real-valued Lévy process with characteristic triplet (A,ν,γ) is said to be of jump-diffusion type if $A>0$ and $\nu(\mathbb{R}\setminus\{0\}) < \infty$.

Corollary 1.2. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process of jump-diffusion type under P. Then $Q|_{\mathcal{F}_t} \ll P|_{\mathcal{F}_t}$ for all $t\in[0,T_\infty]$ or, equivalently, $Q|_{\mathcal{F}_t} \ll P|_{\mathcal{F}_t}$ for some $t\in(0,T_\infty]$ if and only if X is of jump-diffusion type under Q and conditions 1 and 2 of Proposition 1.5 are satisfied.*

Proof. Observe that for all u, v and $r>0$,

$$(1-r)u^2 + (1-1/r)v^2 \leq (u+v)^2 \leq (1+r)u^2 + (1+1/r)v^2.$$

Taking $u=\sqrt{\phi(x)}$, $v=-1$ and $r=1/2$, we obtain:

$$\frac{1}{2}\phi(x) - 1 \leq (\sqrt{\phi(x)}-1)^2 \leq \frac{3}{2}\phi(x) + 3. \tag{1.9}$$

Suppose that $Q|_{\mathcal{F}_t} \ll P|_{\mathcal{F}_t}$ for all t. Then (1.9) implies that

$$\nu_Q(\mathbb{R}\setminus\{0\}) \leq 2\nu_P(\mathbb{R}\setminus\{0\}) + 2\int_{-\infty}^{\infty} (\sqrt{\phi(x)}-1)^2\nu_P(dx) < \infty$$

and therefore $\{X_t\}_{t\geq 0}$ is of jump-diffusion type under Q.

Suppose that $\{X_t\}_{t\geq 0}$ is of jump-diffusion type under Q and conditions 1 and 2 are satisfied. On account of Remark 1.2, we only need to verify condition 3 of the proposition. Equation (1.9) implies that

$$\int_{-\infty}^{\infty} (\sqrt{\phi(x)}-1)^2\nu_P(dx) \leq 3\nu_P(\mathbb{R}\setminus\{0\}) + \frac{3}{2}\nu_Q(\mathbb{R}\setminus\{0\}) < \infty,$$

which completes the proof. □

1.1.3 Tightness and convergence of sequences of Lévy processes

Let E be a Polish space with Borel σ-field $\mathcal{F}$. The space of all probability measures on $(E, \mathcal{F})$ is denoted by $\mathcal{P}(E)$. The *weak topology* on $\mathcal{P}(E)$ is the coarsest topology for which the mappings $\mu \in \mathcal{P}(E) \mapsto \mu(f)$ are continuous for all bounded continuous functions f on E. We recall that a subset $B \subset \mathcal{P}(E)$ is called tight if for every $\varepsilon > 0$ there exists a compact subset $K \subset E$ such that $\mu(E \setminus K) \leq \varepsilon$ for all $\mu \in B$. Prohorov's theorem states that a subset $B \subset \mathcal{P}(E)$ is relatively compact for the weak topology if and only if it is tight.

The space of all càdlàg functions on $[0, T_\infty]$, equipped with Skorokhod topology (see Chapter VI in [54]) is a Polish space and all of the above remains valid. The following result is a corollary (an adaptation for the case Lévy processes) of Theorem VI.4.18 in [54].

Proposition 1.6. *Let $\{X^n\}_{n\geq 1}$ be a sequence of real-valued Lévy processes with characteristic triplets (A_n, ν_n, γ_n). For $\{X^n\}$ to be tight it suffices that*

1. *The sequence of Lévy measures $\{\nu_n\}$ satisfies*

$$\lim_{a\to\infty} \sup_n \nu_n(\{x : |x| > a\}) = 0.$$

2. *The sequence $\{\gamma_n\}$ is bounded.*

3. *The sequence $\{C_n\}$ is bounded, where $C_n = A_n + \int_{-\infty}^{\infty}(x^2 \wedge 1)\nu_n(dx)$.*

The next result characterizes the weak convergence of sequences of Lévy processes (cf. Corollary VII.3.6 in [54]).

Proposition 1.7. *Let $\{X^n\}_{n\geq 1}$ and X be real-valued Lévy processes with characteristic triplets $(A_n, \nu_n, \gamma_n^h)_{n\geq 1}$ and (A, ν, γ^h) with respect to a truncation function h (cf. Remark 1.1), that is continuous, bounded and satisfies $h(x) \equiv x$ on a neighborhood of 0. There is equivalence between*

1. *$X^n \xrightarrow{d} X$.*

2. *$X_1^n \xrightarrow{d} X_1$.*

3. *The following conditions are satisfied:*

   (a) $\gamma_n^h \to \gamma^h$.

(b) $C_n \to C$, *where* $C_* = A_* + \int_{-\infty}^{\infty} h^2(x)\nu_*(dx)$.

(c) $\int_{-\infty}^{\infty} f(x)\nu_n(dx) \to \int_{-\infty}^{\infty} f(x)\nu(dx)$ *for every continuous bounded function* f *such that* $f(x) \equiv 0$ *on a neighborhood of zero, or, equivalently, for every continuous bounded function* f *satisfying* $f(x) = o(|x|^2)$ *when* $x \to 0$.

1.2 Exponential Lévy models: definition and main properties

Exponential Lévy models are obtained by replacing the Brownian motion with drift in the classical Black-Scholes-Samuelson model of asset price, by a Lévy process:

$$S_t = S_0 e^{rt+X_t}, \tag{1.10}$$

where X is a Lévy process on $(\Omega, \mathcal{F}, P)$, and the interest rate term rt is introduced to simplify the notation below. When P is the probability that describes the evolution of stock prices in the real world, also called the historical probability, the model (1.10) is called a *historical exponential Lévy model*.

By the first fundamental theorem of asset pricing (see [34]), a financial market does not allow for arbitrage opportunity (more precisely, satisfies the No Free Lunch with Vanishing Risk condition) if there exists a probability measure Q, equivalent to P, such that the discounted prices $e^{-rt}V_t$ of all assets are Q-local martingales. Q is called a *risk-neutral* probability. The absence of arbitrage in the model (1.10) is therefore equivalent to the existence of a probability $Q \sim P$, such that e^X is a Q-local martingale. The following result shows that if X is a Lévy process under P, one can almost always find a probability $Q \sim P$, under which X is still a Lévy process and e^X is a martingale.

Proposition 1.8 (Absence of arbitrage in exp-Lévy models). *Let* $\{X_t\}_{t\geq 0}$ *be a Lévy process on* $(\Omega, \mathcal{F}, P)$ *with characteristic triplet* (A, ν, γ). *If the trajectories of* X *are neither almost surely increasing nor almost surely decreasing, then there exists a probability measure* Q *equivalent to* P *such that under* Q, $\{X_t\}_{t\geq 0}$ *is a Lévy process and* $\{e^{X_t}\}_{t\geq 0}$ *is a martingale.*

Q *may be chosen in such way that* (X, Q) *will have the characteristic triplet* (A^Q, ν^Q, γ^Q),

where

$$A^Q = A, \qquad \nu^Q = e^{-x^2+\theta x}\nu, \tag{1.11}$$

$$\gamma^Q = \gamma + A\theta + \int_{|x|\leq 1} x(e^{-x^2+\theta x} - 1)\nu(dx) \tag{1.12}$$

for some $\theta \in \mathbb{R}$.

In other words, this proposition shows that exponential Lévy models are arbitrage-free in the following (not mutually exclusive) cases:

- X has infinite variation $\int_{-1}^{1} |x|\nu(dx) = \infty$ and/or $A > 0$.
- X has both positive and negative jumps.
- X has positive jumps and negative drift or negative jumps and positive drift.

In the sequel, we will denote the set of Lévy processes satisfying at least one of these conditions by $\mathcal{L}_{NA}$ and the corresponding exponential Lévy models will be called *arbitrage-free* exponential Lévy models.

The first part of this proposition was proved in [55], see also [25, Theorem 4.6]. Since the second part will also be used below, we give a proof here.

Proof of Proposition 1.8. For any $\theta \in \mathbb{R}$ we denote the probability measure under which X has the characteristic triplet (1.11)–(1.12) by Q^θ. From Proposition 1.5 it follows easily that for every $\theta \in \mathbb{R}$, both $Q^\theta \ll P$ and $P \ll Q^\theta$. Therefore, we only need to prove that for some θ, the triplet $(A, \nu^{Q^\theta}, \gamma^{Q^\theta})$ satisfies the martingale condition (1.2), or equivalently, that the equation

$$f(\theta) + \gamma = 0 \tag{1.13}$$

has a solution, where

$$f(\theta) := \frac{A}{2} + A\theta + \int_{|x|\leq 1} x(e^{-x^2+\theta x} - 1)\nu(dx) + \int_{-\infty}^{\infty} (e^x - 1 - x1_{|x|\leq 1})e^{-x^2+\theta x}\nu(dx)$$

The dominated convergence theorem implies that f is continuous and differentiable and that $f'(\theta) = A + \int_{-\infty}^{\infty} x(e^x - 1)e^{-x^2+\theta x}\nu(dx) \geq 0$, which means that $f(\theta)$ is an increasing function. Moreover, if $A > 0$ or if $\nu((0,\infty)) > 0$ and $\nu((-\infty, 0)) > 0$ then f' is everywhere bounded from below by a positive number. Therefore in these cases $\lim_{\theta\to+\infty} f(\theta) = +\infty$, $\lim_{\theta\to-\infty} f(\theta) = -\infty$ and Equation (1.13) admits a solution.

It remains to treat the case when $A = 0$ and ν is supported by one of the half-axes. Suppose that $\nu((-\infty, 0)) = 0$ and $\nu((0, \infty)) > 0$. In this case $\lim_{\theta\to+\infty} f(\theta) = +\infty$ and

$$\lim_{\theta\to-\infty} f(\theta) = -\int_{0<x\leq 1} x\nu(dx).$$

Therefore, Equation (1.13) admits a solution unless $\gamma - \int_{0<x\leq 1} x\nu(dx) \geq 0$, which is the case when X has almost surely increasing trajectories (cf. Proposition 3.10 in [27]). By symmetry we can treat the case of decreasing trajectories and complete the proof. □

Risk-neutral exp-Lévy models An exponential Lévy model of type (1.10), where X is a Lévy process with characteristic triplet (A, ν, γ), satisfying condition (1.2) (that is, $\{e^{X_t}\}_{t\geq 0}$ is a martingale) is called a *risk-neutral exponential Lévy model*. Since γ is uniquely determined from A and ν by the martingale condition, a risk-neutral exponential Lévy model can be parametrized by A and ν only: $Q = Q(A, \nu)$.

Under a risk-neutral probability $Q(A, \nu)$, call option prices can be evaluated as discounted expectations of terminal payoffs:

$$C^{Q(A,\nu)}(T, K) = e^{-rT} E^{Q(A,\nu)}[(S_T - K)^+] = e^{-rT} E^{Q(A,\nu)}[(S_0 e^{rT+X_T} - K)^+]. \tag{1.14}$$

1.3 Exponential Lévy models in finance

This section reviews different exponential Lévy models found in the financial literature. These models can be used to describe stock price evolution under both historical and risk-neutral probability, but under the risk-neutral probability the drift parameter (γ) is fixed by the martingale condition.

Financial models with jumps fall into two categories. In the first category, called *jump-diffusion* models, the "normal" evolution of prices is given by a diffusion process, punctuated by jumps at random intervals. Here the jumps represent rare events — crashes and large drawdowns. Such an evolution can be represented by modelling the log-price as a Lévy process with a nonzero Gaussian component and a jump part, which is a compound Poisson process with finitely many jumps in every time interval:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \tag{1.15}$$

where $\{N_t\}_{t\geq 0}$ is the Poisson process counting the jumps of X and Y_i are jump sizes (i.i.d. variables).

In the *Merton model* [71], which is the first model of this type, suggested in the literature, jumps in the log-price X are assumed to have a Gaussian distribution: $Y_i \sim N(\mu, \delta^2)$. In the risk-neutral version the characteristic exponent of the log stock price takes the following form:

$$\psi(u) = -\frac{Au^2}{2} + \lambda\{e^{-\delta^2 u^2/2 + i\mu u} - 1\} - iu\left(\frac{A}{2} + \lambda(e^{\delta^2/2+\mu} - 1)\right). \tag{1.16}$$

In the *Kou model* [62], jump sizes are distributed according to an asymmetric Laplace law with a density of the form

$$\nu_0(dx) = [p\lambda_+ e^{-\lambda_+ x} 1_{x>0} + (1-p)\lambda_- e^{-\lambda_- |x|} 1_{x<0}]dx \tag{1.17}$$

with $\lambda_+ > 0$, $\lambda_- > 0$ governing the decay of the tails for the distribution of positive and negative jump sizes and $p \in [0, 1]$ representing the probability of an upward jump. The probability distribution of returns in this model has semi-heavy (exponential) tails.

In the above two models, the dynamical structure of the process is easy to understand and describe, since the distribution of jump sizes is known. They are easy to simulate and efficient Monte Carlo methods for pricing path-dependent options have been developed. Models of this type also perform quite well for the purposes of implied volatility smile interpolation. However, they do not lead to closed-form densities: statistical estimation and computation of moments or quantiles may be quite difficult.

The second category consists of models with infinite number of jumps in every interval, which we will call *infinite activity* or *infinite intensity* models. In these models, one does not need to introduce a Brownian component since the dynamics of jumps is already rich enough to generate nontrivial small time behavior [21] and it has been argued [21, 43, 66] that such models give a more realistic description of the price process at various time scales. In addition, many models from this class can be constructed via Brownian subordination, which gives them additional tractability compared to jump-diffusion models.

Table 1.1 compares the advantages and drawbacks of these two categories. It should be kept in mind that since the price process is observed on a discrete grid, it is difficult if not impossible to see empirically to which category the price process belongs. The choice is more a question of modelling convenience than an empirical one.

Table 1.1: Compound Poisson or infinite intensity: a comparison of two modelling approaches

Jump-diffusion models	**Infinite intensity models**
Must contain a Brownian component.	Brownian component is not needed.
Jumps are rare events.	The process moves only by jumps and deterministic linear drift.
Distribution of jump sizes is known.	"Distribution of jump sizes" does not exist: jumps arrive infinitely often.
Perform well for implied volatility smile interpolation.	Give a realistic description of the historical price process.
Densities not known in closed form.	Closed form densities available in some cases.
Easy to simulate.	In some cases can be represented via Brownian subordination, which gives additional tractability.

There are three standard ways to define a parametric Lévy process with infinite jump intensity, summarized in Table 1.2.

The first approach is to obtain a Lévy process by subordinating a Brownian motion with an independent increasing Lévy process (such a process is called a subordinator). Here the characteristic function of the resulting process can be obtained immediately, but an explicit formula for the Lévy measure is not always available. Due to the conditionally Gaussian structure of the process, simulation and some computations can be considerably simplified (for instance, call option price can be expressed as an integral involving Black-Scholes prices). The interpretation of the subordinator as "business time" [44] makes models of this type easier to understand and interpret. Multidimensional extensions are also possible: one can take a multivariate Brownian motion and change the time scale of all components with the same subordinator. Two examples of models from this class are the variance gamma process and the normal inverse Gaussian process. The variance gamma process [23, 67] is obtained by time-changing a Brownian motion with a gamma subordinator and has the characteristic exponent of the form:

$$\psi(u) = -\frac{1}{\kappa}\log(1 + \frac{u^2\sigma^2\kappa}{2} - i\theta\kappa u). \tag{1.18}$$

The density of the Lévy measure of the variance gamma process is given by

$$\nu(x) = \frac{c}{|x|}e^{-\lambda_-|x|}1_{x<0} + \frac{c}{x}e^{-\lambda_+ x}1_{x>0}, \tag{1.19}$$

where $c = 1/\kappa$, $\lambda_+ = \frac{\sqrt{\theta^2+2\sigma^2/\kappa}}{\sigma^2} - \frac{\theta}{\sigma^2}$ and $\lambda_- = \frac{\sqrt{\theta^2+2\sigma^2/\kappa}}{\sigma^2} + \frac{\theta}{\sigma^2}$.

The normal inverse Gaussian process [5, 7, 86] is the result of time-changing a Brownian motion with the inverse Gaussian subordinator and has the characteristic exponent:

$$\psi(u) = \frac{1}{\kappa} - \frac{1}{\kappa}\sqrt{1 + u^2\sigma^2\kappa - 2i\theta u\kappa}.$$

The second approach is to specify the Lévy measure directly. The main example of this category is given by the tempered stable process, introduced by Koponen [61] and used for financial modelling in [18, 26, 27], as well as in [22] (under the name of CGMY model) and in [19] (under the name of KoBoL process). The tempered stable process has a Lévy measure with density of the form:

$$\nu(x) = \frac{c_-}{|x|^{1+\alpha_-}}e^{-\lambda_-|x|}1_{x<0} + \frac{c_+}{x^{1+\alpha_+}}e^{-\lambda_+ x}1_{x>0} \tag{1.20}$$

with $\alpha_+ < 2$ and $\alpha_- < 2$. This rich parametric form of the Lévy measure is probably sufficient for most applications.

The third approach is to specify the density of increments of the process at a given time scale, say Δ, by taking an arbitrary infinitely divisible distribution. Generalized hyperbolic processes (see [36–38]) can be constructed in this way. In this approach it is easy to simulate the increments of the process at the same time scale and to estimate parameters of the distribution if data are sampled with the same period Δ, but, unless this distribution belongs to some parametric class closed under convolution, we do not know the law of the increments at other time scales. Also, given an infinitely divisible distribution, one may not know its Lévy-Khintchine representation, so it may not be easy to see whether the corresponding Lévy process has a Gaussian component, finite or infinite jump intensity, etc.

1.4 Pricing European options in exp-Lévy models via Fourier transform

In this section we present a method, adapted from [23], for pricing European call options in exp-Lévy models using Fourier transform and, in particular, the Fast Fourier transform algorithm [29]. We suggest several improvements to the original procedure and give a rigorous analysis of truncation and discretization errors. These results are of independent interest and will also be used for the numerical solution of the calibration problem in Chapter 3. Similar analysis has recently appeared in [63].

Let $\{X_t\}_{t\geq 0}$ be a Lévy process satisfying the martingale condition (1.2). To compute the price of a call option

$$C(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+], \tag{1.21}$$

we would like to express its Fourier transform in log strike in terms of the characteristic function $\Phi_T(v)$ of X_T and then find the prices for a range of strikes by Fourier inversion. However we cannot do this directly because $C(k)$ is not integrable (it tends to 1 as k goes to $-\infty$). The key idea is to instead compute the Fourier transform of the (modified) time value of the option, that is, the function

$$z_T(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+] - (1 - e^{k-rT})^+. \tag{1.22}$$

Table 1.2: Three approaches to building parametric exp-Lévy models

Brownian subordination	**Specifying the Lévy measure**	**Specifying probability density for $t = \Delta$**
Interpretation as "Brownian motion in business time".	Clear vision of the pathwise properties.	Structure of jumps is not known.
Simulation is easy if we know how to simulate the subordinator.	Simulation is quite involved.	Simulation is easy on a grid of size Δ.
Estimation via maximum likelihood may be difficult.	Estimation can be done by approximating the transition density.	Estimation is easy for data with sampling interval Δ.
Multivariate generalizations possible using multidimensional Brownian motion.	Rich variety of models.	The infinite divisibility of a given model may be difficult to prove.

Proposition 1.9 (Carr and Madan [23]). *Let $\{X_t\}_{t\geq 0}$ be a real-valued Lévy process satisfying the martingale condition* (1.2)*, such that*

$$E[e^{(1+\alpha)X_t}] < \infty \tag{1.23}$$

for some $\alpha > 0$. Then the Fourier transform in log-strike k of the time value of a call option is given by:

$$\zeta_T(v) := \int_{-\infty}^{+\infty} e^{ivk} z_T(k)dk = e^{ivrT}\frac{\Phi_T(v-i)-1}{iv(1+iv)} \tag{1.24}$$

Remark 1.3. Since typically $\Phi_T(z) \to 0$ as $\Re z \to \infty$, $\zeta_T(v)$ will behave like $|v|^{-2}$ at infinity which means that the truncation error in the numerical evaluation of the inverse Fourier transform will be large. The reason of such a slow convergence is that the time value (1.22) is not smooth; therefore its Fourier transform does not decay sufficiently fast at infinity. For most models the convergence can be improved by replacing the time value with a smooth function of strike. Instead of subtracting the (non-differentiable) intrinsic value of the option from its price, we suggest to subtract the Black-Scholes call price with a non-zero volatility (which is a smooth function). The resulting function will be both integrable and smooth. Suppose that hypothesis (1.23) is satisfied and denote

$$\tilde{z}_T(k) = e^{-rT}E[(e^{rT+X_T} - e^k)^+] - C_{BS}^{\Sigma}(k),$$

where $C_{BS}^{\Sigma}(k)$ is the Black-Scholes price of a call option with volatility Σ and log-strike k for the same underlying value and the same interest rate. Proposition 1.9 then implies that the Fourier transform of $\tilde{z}_T(k)$, denoted by $\tilde{\zeta}_T(v)$, satisfies

$$\tilde{\zeta}_T(v) = e^{ivrT}\frac{\Phi_T(v-i) - \Phi_T^{\Sigma}(v-i)}{iv(1+iv)}, \tag{1.25}$$

where $\Phi_T^{\Sigma}(v) = \exp(-\frac{\Sigma^2 T}{2}(v^2+iv))$. Since for most exp-Lévy models found in the literature (except variance gamma) the characteristic function decays faster than every power of its argument at infinity, this means that the expression (1.25) will also decay faster than every power of v as $\Re v \to \infty$, and the truncation error in the numerical evaluation of the inverse Fourier transform will be very small for *every* $\Sigma > 0$.

Figure 1.1 shows the behavior of $|\tilde{\zeta}_T|$ for different values of Σ compared to the behavior of $|\zeta_T|$ in Merton's jump-diffusion model with volatility 0.2, jump intensity equal to 5 and jump

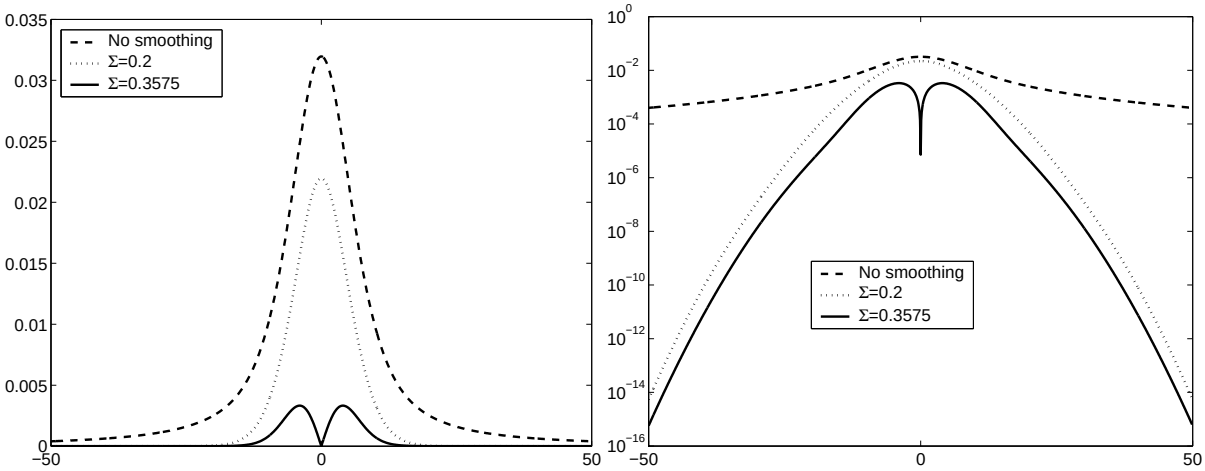

Figure 1.1: Convergence of Fourier transform of option's time value to zero in Merton model. Left graph: linear scale; right graph: logarithmic scale.

parameters $\mu = -0.1$ and $\delta = 0.1$ for the time horizon $T = 0.5$. The convergence of $\tilde{\zeta}_T$ to zero is faster than exponential for all values of Σ and it is particularly good for $\Sigma = 0.3575$, the value of Σ for which $\tilde{\zeta}(0) = 0$.

Proof of Proposition 1.9. Since the discounted price process is a martingale,

$$z_T(k) = e^{-rT} \int_{-\infty}^{\infty} \mu_T(dx)(e^{rT+x} - e^k)(1_{k\leq x+rT} - 1_{k\leq rT}),$$

where μ_T is the probability distribution of X_T. Condition (1.23) enables us to compute $\zeta_T(v)$ by interchanging integrals:

$$\begin{aligned}
\zeta_T(v) &= e^{-rT} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} \mu_T(dx) e^{ivk}(e^{rT+x} - e^k)(1_{k\leq x+rT} - 1_{k\leq rT}) \\
&= e^{-rT} \int_{-\infty}^{\infty} \mu_T(dx) \int_{x+rT}^{rT} e^{ivk}(e^k - e^{rT+x}) dk \\
&= \int_{-\infty}^{\infty} \mu_T(dx) \left\{ \frac{e^{ivrT}(1-e^x)}{iv+1} - \frac{e^{x+ivrT}}{iv(iv+1)} + \frac{e^{(iv+1)x+ivrT}}{iv(iv+1)} \right\}
\end{aligned}$$

The first term in braces disappears due to the martingale condition and the other two, after computing the integrals, yield (1.24). □

Numerical Fourier inversion. Option prices can be computed by evaluating numerically the inverse Fourier transform of $\tilde{\zeta}_T$:

$$\tilde{z}_T(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ivk}\tilde{\zeta}_T(v)dv \tag{1.26}$$

This integral can be efficiently computed for a range of strikes using the Fast Fourier transform, an algorithm due to Cooley and Tukey [29] which allows to compute the discrete Fourier transform $\mathrm{DFT}[f]_{n=0}^{N-1}$, defined by,

$$\mathrm{DFT}[f]_n := \sum_{k=0}^{N-1} f_k e^{-2\pi ink/N}, \quad n = 0 \dots N-1, \tag{1.27}$$

using only $O(N \log N)$ operations.

To approximate option prices, we truncate and discretize the integral (1.26) as follows:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ivk}\tilde{\zeta}_T(v)dv &= \frac{1}{2\pi}\int_{-L/2}^{L/2} e^{-ivk}\tilde{\zeta}_T(v)dv + \varepsilon_T \\ &= \frac{L}{2\pi(N-1)}\sum_{m=0}^{N-1} w_m\tilde{\zeta}_T(v_m)e^{-ikv_m} + \varepsilon_T + \varepsilon_D, \end{aligned}\tag{1.28}$$

where ε_T is the truncation error, ε_D is the discretization error, $v_m = -L/2 + m\Delta$, $\Delta = L/(N-1)$ is the discretization step and w_m are weights, corresponding to the chosen integration rule (for instance, for the trapezoidal rule $w_0 = w_{N-1} = 1/2$ and all other weights are equal to 1). Now, choosing $k_n = k_0 + \frac{2\pi n}{N\Delta}$ we see that the sum in the last term becomes a discrete Fourier transform:

$$\begin{aligned}\frac{L}{2\pi(N-1)}e^{ik_nL/2}\sum_{m=0}^{N-1} w_m\tilde{\zeta}_T(k_m)e^{-ik_0m\Delta}e^{-2\pi inm/N} \\ = \frac{L}{2\pi(N-1)}e^{ik_nL/2}\mathrm{DFT}_n[w_m\tilde{\zeta}_T(k_m)e^{-ik_0m\Delta}]\end{aligned}$$

Therefore, the FFT algorithm allows to compute $\tilde{z}_T$ and option prices for the log strikes $k_n = k_0 + \frac{2\pi n}{N\Delta}$. The log strikes are thus equidistant with the step d satisfying

$$d\Delta = \frac{2\pi}{N}.$$

Error control. We start with the truncation error.

Lemma 1.10. *Let $\{X_t\}_{t\geq 0}$ be a Lévy process with characteristic triplet (A, ν, γ) and characteristic function Φ_T. Then*

1. $\forall\, v \in \mathbb{R},\ |\Phi_T(v-i)| \le \exp(-\frac{TAv^2}{2})$.

2. *Suppose that* ν *has a density of the form* $\nu(x) = \frac{e^{-x}}{|x|} f(x)$, *where* f *is increasing on* $(-\infty, 0)$ *and decreasing on* $(0, \infty)$. *Then* $|\Phi_T(v-i)|$ *is increasing on* $v \in (-\infty, 0)$ *and decreasing on* $v \in (0, \infty)$.

Proof. The martingale condition implies that $\forall\, v \in \mathbb{R}$,

$$|\Phi_T(v-i)| = \exp\left\{-\frac{TAv^2}{2} - \int_{-\infty}^{\infty} e^x(1-\cos(vx))\nu(dx)\right\}. \tag{1.29}$$

This immediately entails the first part of the lemma. For the second part, let $0 > v_1 > v_2$ (the case $v_1 < v_2 < 0$ can be shown in the same way). Then, for all $t \in \mathbb{R}$,

$$\frac{e^{t/v_1}\nu(t/v_1)}{v_1} \le \frac{e^{t/v_2}\nu(t/v_2)}{v_2}.$$

Therefore,

$$\begin{aligned}\int_{-\infty}^{\infty} e^x(1-\cos v_1 x)\nu(dx) &= \int_{-\infty}^{\infty}(1-\cos t)\frac{e^{t/v_1}\nu(t/v_1)}{v_1}dt \\ &\le \int_{-\infty}^{\infty}(1-\cos t)\frac{e^{t/v_2}\nu(t/v_2)}{v_2}dt = \int_{-\infty}^{\infty} e^x(1-\cos v_2 x)\nu(dx),\end{aligned}$$

and it follows from Equation (1.29) that $|\Phi_T(v_1-i)| \ge |\Phi_T(v_2-i)|$. □

Proposition 1.11. *Let* $\{X_t\}_{t\ge 0}$ *be a real-valued Lévy process with triplet* (A, ν, γ) *and characteristic function* Φ_T *and let* $\Sigma > 0$.

1. *Suppose* $A > 0$. *Then the truncation error in Equation* (1.28) *satisfies:*

$$|\varepsilon_T| \le \frac{8}{\pi T\Sigma^2 L^3} e^{-\frac{TL^2\Sigma^2}{8}} + \frac{8}{\pi TAL^3} e^{-\frac{TL^2 A}{8}}. \tag{1.30}$$

2. *Suppose that* ν *has a density of the form* $\nu(x) = \frac{e^{-x}}{|x|} f(x)$, *where* f *is increasing on* $(-\infty, 0)$ *and decreasing on* $(0, \infty)$. *Then the truncation error in Equation* (1.28) *satisfies:*

$$|\varepsilon_T| \le \frac{8}{\pi T\Sigma^2 L^3} e^{-\frac{TL^2\Sigma^2}{8}} + \frac{|\Phi_T(L/2-i)| + |\Phi_T(-L/2-i)|}{\pi L}.$$

Proof. From Equation, (1.28),

$$\begin{aligned}|\varepsilon_T| &\le \frac{1}{2\pi}\int_{L/2}^{\infty} |\tilde{\zeta}_T(v)|dv + \frac{1}{2\pi}\int_{-\infty}^{-L/2} |\tilde{\zeta}_T(v)|dv \\ &\le \frac{1}{2\pi}\int_{(-\infty,-L/2)\cup(L/2,\infty)} \frac{|\Phi_T(v-i)|dv}{v^2} + \frac{1}{2\pi}\int_{(-\infty,-L/2)\cup(L/2,\infty)} \frac{|\Phi_T^{\Sigma}(v-i)|dv}{v^2}.\end{aligned}$$

By the first part of Lemma 1.10,

$$\frac{1}{2\pi}\int_{L/2}^{\infty}\frac{|\Phi_T(v-i)|}{v^2}dv \le \frac{1}{2\pi}\int_{L/2}^{\infty}\frac{dv}{v^2}e^{-TAv^2/2} = \frac{1}{4\pi}\int_{L^2/4}^{\infty}\frac{dz}{z^{3/2}}e^{-TAz/2}$$
$$\le \frac{1}{4\pi}\frac{8}{L^3}\int_{L^2/4}^{\infty}dze^{-TAz/2} = \frac{4}{\pi TAL^3}e^{-\frac{L^2AT}{8}},$$

and since a similar bound can be obtained for Φ_T^{Σ}, this proves the first part of the proposition. To prove the second statement, observe that from the second part of Lemma 1.10,

$$\frac{1}{2\pi}\int_{L/2}^{\infty}\frac{|\Phi_T(v-i)|}{v^2}dv \le \frac{|\Phi_T(L/2-i)|}{2\pi}\int_{L/2}^{\infty}\frac{dv}{v^2} = \frac{|\Phi_T(L/2-i)|}{\pi L}.$$

□

To compute a bound for the discretization (sampling) error ε_D in Equation (1.28), we define constants $C_k, k \ge 1$ by

$$C_k = \begin{cases} |\Phi_T^{(k)}(-i)|, & k \text{ is even}, \\ |\Phi_T^{(k+1)}(-i)|^{\frac{k}{k+1}}, & k \text{ is odd}, \end{cases} \tag{1.31}$$

and start by proving a technical lemma.

Lemma 1.12. *Let $\{X_t\}_{t\ge 0}$ be a real-valued Lévy process with triplet (A,ν,γ) and characteristic function Φ_T, such that the condition (1.23) is satisfied. Let $\Gamma_T(v) := \frac{\Phi_T(v-i)-1}{v}$. Then for all $n \ge 0$,*

$$|\Gamma_T^{(n)}(v)| \le \frac{C_{n+1}}{n+1} \quad \forall\, v \in \mathbb{R}$$

Proof. Under the condition (1.23), for all $n \ge 1$, $E[|X_T|^n e^{X_T}] < \infty$. Therefore, by Lebesgue's dominated convergence theorem, for all $n \ge 1$,

$$\Phi_T^{(n)}(v-i) = \int_{-\infty}^{\infty}(ix)^n e^{i(v-i)x}\mu_T(dx),$$

where μ_T is the probability distribution of X_T. For even $n > 1$ this implies:

$$|\Phi_T^{(n)}(v-i)| \le \int_{-\infty}^{\infty}x^n e^x \mu_T(dx) = C_n,$$

and for odd $n \ge 1$, using Jensen's inequality for the probability measure $\tilde{\mu}_T = e^x\mu_T$ yields:

$$|\Phi_T^{(n)}(v-i)| \le \int_{-\infty}^{\infty}|x|^n e^x \mu_T(dx) \le \left(\int_{-\infty}^{\infty}|x|^{n+1}e^x\mu_T(dx)\right)^{\frac{n}{n+1}} = C_n,$$

From the martingale condition,

$$\Gamma_T(v) = \frac{\Phi_T(v-i) - \Phi_T(-i)}{v} = \int_0^1 \Phi_T'(vt-i)dt.$$

Applying the dominated convergence theorem once again, we conclude that

$$\Gamma_T^{(n)}(v) = \int_0^1 t^n \Phi_T^{(n+1)}(vt-i)dt,$$

and therefore $|\Gamma_T^{(n)}(v)| \leq \frac{C_{n+1}}{n+1}$ for all $n \geq 0$. □

Proposition 1.13. *Suppose that the integral in* (1.28) *is approximated using the trapezoidal rule. Then the discretization error* ε_D *satisfies*

$$|\varepsilon_D| \leq \frac{\Delta^2}{6\pi} \sum_{l=0}^{2} \frac{C_{3-l} + C_{3-l}^{\Sigma}}{(3-l)!} \left\{ \left(\Delta + \frac{\pi}{2}\right) e^{|k-rT|} + \log\left(\frac{L}{2} + \sqrt{\frac{L^2}{4}+1}\right) \frac{|k-rT|^l}{l!} \right\}, \quad (1.32)$$

where C_k^{Σ} *is computed as in* (1.31) *for the characteristic function* Φ_T^{Σ}.

If this integral is approximated using the Simpson rule, ε_D *satisfies*

$$|\varepsilon_D| \leq \frac{\Delta^4}{5\pi} \sum_{l=0}^{4} \frac{C_{5-l} + C_{5-l}^{\Sigma}}{(5-l)!} \left\{ \left(2\Delta + \frac{\pi}{2}\right) e^{|k-rT|} + \log\left(\frac{L}{2} + \sqrt{\frac{L^2}{4}+1}\right) \frac{|k-rT|^l}{l!} \right\}, \quad (1.33)$$

Proof. The trapezoidal rule (cf. [32]) is defined by

$$\int_x^{x+h} f(\xi)d\xi = \frac{1}{2}h(f(x) + f(x+h)) + R \quad \text{with} \quad R = -\frac{1}{12}h^3 f''(x^*)$$

for some $x^* \in [x, x+h]$. The sampling error in (1.28) therefore satisfies:

$$\begin{aligned} |\varepsilon_D| &\leq \frac{\Delta^3}{24\pi} \sum_{m=0}^{N-2} \sup_{v \in [-L/2+m\Delta, -L/2+(m+1)\Delta]} \left| \frac{\partial^2}{\partial v^2} e^{-ivk} \tilde{\zeta}_T(v) \right| \\ &= \frac{\Delta^3}{24\pi} \sum_{m=0}^{N-2} \left\{ \sup \left| \frac{\partial^2}{\partial v^2} e^{-ivk} \zeta_T(v) \right| + \sup \left| \frac{\partial^2}{\partial v^2} e^{-ivk} \zeta_T^{\Sigma}(v) \right| \right\}, \quad (1.34) \\ &\quad \text{where} \quad \zeta_T^{\Sigma}(v) := e^{ivrT} \frac{\Phi_T^{\Sigma}(v-i) - 1}{iv(1+iv)}. \end{aligned}$$

Observe that $e^{-ivk}\zeta_T(v) = \frac{e^{-iv(k-rT)}}{i(1+iv)}\Gamma_T(v)$, and therefore

$$\begin{aligned} \left| \frac{\partial^n}{\partial v^n}(e^{ivk}\zeta_T(v)) \right| &\leq \sum_{l=0}^{n} |\Gamma_T^{(n-l)}(v)| \frac{n!}{(n-l)!} \sum_{j=0}^{l} \frac{|k-rT|^j}{j!} \left(\frac{1}{\sqrt{1+v^2}}\right)^{l-j+1} \\ &\leq n! \sum_{l=0}^{n} \frac{C_{n-l+1}}{(n-l+1)!} \sum_{j=0}^{l} \frac{|k-rT|^j}{j!} \left(\frac{1}{\sqrt{1+v^2}}\right)^{l-j+1} := \tilde{g}_n(v). \end{aligned}$$

Since the bound $\tilde{g}_n(v)$ is increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$,

$$\frac{\Delta^3}{24\pi}\sum_{m=0}^{N-2}\sup\left|\frac{\partial^2}{\partial v^2}e^{-ivk}\zeta_T(v)\right| \le \frac{\Delta^3}{12\pi}\tilde{g}_2(0) + \frac{\Delta^2}{24\pi}\int_{-L/2}^{L/2}\tilde{g}_2(v)dv$$

$$\le \frac{\Delta^2}{12\pi}\sum_{l=0}^{2}\frac{C_{3-l}}{(3-l)!}\left\{2\Delta\sum_{j=0}^{l}\frac{|k-rT|^j}{j!} + \sum_{j=0}^{l-1}\frac{|k-rT|^j}{j!}\int_{-L/2}^{L/2}\frac{dv}{1+v^2} + \frac{|k-rT|^l}{l!}\int_{-L/2}^{L/2}\frac{dv}{\sqrt{1+v^2}}\right\}$$

$$\le \frac{\Delta^2}{6\pi}\sum_{l=0}^{2}\frac{C_{3-l}}{(3-l)!}\left\{\left(\Delta+\frac{\pi}{2}\right)e^{|k-rT|} + \log\left(\frac{L}{2}+\sqrt{\frac{L^2}{4}+1}\right)\frac{|k-rT|^l}{l!}\right\}.$$

To complete the proof of (1.32) it remains to substitute this bound and a similar bound for $\frac{\Delta^3}{24\pi}\sum_{m=0}^{N-2}\sup\left|\frac{\partial^2}{\partial v^2}e^{-ivk}\zeta_T^{\Sigma}(v)\right|$ into (1.34).

The Simpson rule (cf. [32]) is defined by:

$$\int_x^{x+2h} f(\xi)d\xi = \frac{1}{3}h(f(x)+4f(x+h)+f(x+2h)) + R,$$
$$\text{where}\quad R = \frac{1}{90}h^4 f^{(4)}(x^*)$$

for some $x^* \in [x, x+2h]$. The proof of (1.33) can be carried out in the same way as for the trapezoidal rule. □

Example 1.1. Let us compute the truncation and discretization errors in the Merton model (1.16) with parameters $\sigma = 0.1$, $\lambda = 2$, $\delta = 0.1$, $\mu = 0$ for options with maturity $T = 0.25$. Taking $\Sigma = 0.1$, we obtain the following values of the coefficients C_k and C_k^{Σ}:

k	C_k	C_k^{Σ}
1	0.0871	0.05
2	0.0076	0.0025
3	0.0024	$2.85 \cdot 10^{-4}$
4	$3.29 \cdot 10^{-4}$	$1.88 \cdot 10^{-5}$
5	$1.82 \cdot 10^{-4}$	$2.99 \cdot 10^{-6}$

With 4048 points and the log strike step d equal to 0.01, the truncation error is extremely small: $\varepsilon_T = 2 \cdot 10^{-59}$, and the discretization error for at the money options is given by $\varepsilon_D = 0.0013$ for the trapezoidal rule and by $\varepsilon_D = 3.8 \cdot 10^{-5}$ for the Simpson rule. Since the call option price in this setting ($S = K = 1$, $r = q = 0$) is $C = 0.0313$, we conclude that the Simpson rule gives an acceptable pricing error for this choice of parameters.

Chapter 2

The calibration problem and its regularization

This chapter lays the theoretical foundations of our calibration method. Discretization of the regularized calibration problem and numerical implementation of the calibration algorithm are addressed in the next chapter.

The calibration problem consists, roughly speaking, of finding a risk-neutral exponential Lévy model consistent with market prices of traded options $\{C_M(T_i, K_i)\}_{i\in I}$ for some index set I. In other words, the solution of the calibration problem is a probability Q on the path space $(\Omega, \mathcal{F})$, such that (X, Q) is a Lévy process, satisfying the martingale condition (1.2) and such that the option prices, computed using (1.14) are in some sense close to market prices C_M.

Suppose first that the market data C_M are consistent with the class of exponential Lévy models. This is for example the case when the true model underlying market data is an exponential Lévy model, but this is not the only situation where the above is true: many models may give the same prices for a given set of European options. For instance, it is easy to construct, using Dupire's formula, a local volatility model that gives the same prices, for a set of European options, as a given exp-Lévy model. Denoting by $\mathcal{L}$ the set of all probabilities Q such that (X, Q) is a Lévy process and by $\mathcal{M}$ the set of probabilities Q such that $\{e^{X_t}\}_{t\geq 0}$ is a Q-martingale, the calibration problem assumes the following form:

Calibration problem for data, consistent with exp-Lévy model. Given market prices of call options $\{C_M(T_i, K_i)\}_{i\in I}$, find $Q^* \in \mathcal{M} \cap \mathcal{L}$, such that

$$\forall i \in I, C^{Q^*}(T_i, K_i) = C_M(T_i, K_i). \tag{2.1}$$

In most cases, however, Equations (2.1) cannot be solved exactly, either because the observed option prices contain errors or because the market in question cannot be described by an exponential Lévy model. In this case a common practice is to replace the exact constraints (2.1) by a nonlinear least squares calibration problem. Section 2.1 takes a critical look at this approach and establishes some limits of its applicability to non-parametric calibration of exponential Lévy models.

An important difficulty of least squares calibration is its lack of identification, meaning that a finite number of observed option prices does not allow to reconstruct the law of a Lévy process in a unique fashion. To address this issue, we suggest, in Section 2.2 to reformulate the calibration problem as that of finding the risk-neutral exponential Lévy model that has the smallest relative entropy with respect to a given prior probability measure among all solutions of the least squares calibration problem. Section 2.3 reviews the literature on the use of relative entropy for pricing and calibration and places the present study into the context of previous work on this subject.

The use of relative entropy for selection of solutions removes to some extent the identification problem but the resulting calibration problem is still ill-posed: small errors in market data may lead to large changes of its solution. The last section of this chapter uses the method of *regularization* to approximate the solutions of this problem in a stable manner in presence of data errors.

2.1 Least squares calibration

When the market data is not consistent with the class of exponential Lévy models, the exact calibration problem may not have a solution. In this case one may consider an approximate solution: instead of reproducing the market option prices exactly, one may look for a Lévy triplet which reproduces them in the best possible way in the least squares sense. Let w be a probability

measure on $[0, T_\infty] \times [0, \infty)$ (the weighting measure, determining the relative importance of different options). An option data set is defined as a mapping $C : [0, T_\infty] \times [0, \infty) \to [0, \infty)$ and the data sets that coincide w-almost everywhere are considered identical. One can introduce a norm on option data sets via

$$\|C\|_w^2 := \int_{[0,T_\infty]\times[0,\infty)} C(T,K)^2 w(dT \times dK). \tag{2.2}$$

The quadratic pricing error in model Q is then given by $\|C_M - C^Q\|_w^2$. If the number of constraints is finite then $w = \sum_{i=1}^N w_i \delta_{(T_i,K_i)}(dT \times dK)$ (we suppose that there are N constraints), where $\{w_i\}_{1\leq i\leq N}$ are positive weights that sum up to one. Therefore, in this case

$$\|C_M - C^Q\|_w^2 = \sum_{i=1}^N w_i (C_M(T_i,K_i) - C^Q(T_i,K_i))^2. \tag{2.3}$$

The calibration problem now takes the following form:

Least squares calibration problem. Given prices C_M of call options, find $Q^* \in \mathcal{M} \cap \mathcal{L}$, such that

$$\|C_M - C^{Q^*}\|_w^2 = \inf_{Q\in\mathcal{M}\cap\mathcal{L}} \|C_M - C^Q\|_w^2. \tag{2.4}$$

In the sequel, any such Q^* will be called a least squares solution and the set of all least squares solutions will be denoted by $\mathcal{Q}^{LS}$.

Several authors (see for example [2, 10]) have used the form (2.4) without taking into account that the least squares calibration problem is ill-posed in several ways. The principal difficulties of theoretical nature are the lack of identification (knowing the prices of a finite number of options is not sufficient to reconstruct the Lévy process), absence of solution (in some cases even the least squares problem may not admit a solution) and absence of continuity of solution with respect to market data. On the other hand, even if a solution exists, it is very difficult to find numerically, because the functional $\|C_M - C^Q\|^2$ is typically non-convex and has many local minima, preventing a gradient-based minimization algorithm from finding the true solution. In the rest of this section we discuss these difficulties in detail.

2.1.1 Lack of identification

If the data are consistent with an exponential Lévy model and call option prices are known for one maturity and all strikes, the characteristic triplet of the underlying Lévy process could be deduced in the following way:

- Compute the characteristic function Φ_T of log stock price by Fourier transform as in Equation (1.24).
- Deduce the unit variance of the Gaussian component A and the Lévy measure ν from the characteristic function Φ_T. First, A can be found as follows (see [87, page 40]):

$$A = \lim_{u\to\infty} -\frac{2\log\Phi_T(u)}{Tu^2} \tag{2.5}$$

  Now, denoting $\psi(u) \equiv \frac{\log\Phi_T(u)}{T} + \frac{Au^2}{2}$, it can be shown (see Equation (8.10) in [87]) that

$$\int_{-1}^{1}(\psi(u)-\psi(u+z))dz = 2\int_{-\infty}^{\infty} e^{iux}(1-\frac{\sin x}{x})\nu(dx) \tag{2.6}$$

  Therefore, the left-hand side of (2.6) is the Fourier transform of the positive finite measure $2(1-\frac{\sin x}{x})\nu(dx)$. This means that this measure, and, consequently, the Lévy measure ν is uniquely determined by ψ.

However, call prices are only available for a finite number of strikes. This number may be quite small (between 10 and 40 in real examples). Therefore, the above procedure cannot be applied and ν and A must be computed by minimizing the pricing error $\|C_M - C^Q\|_w^2$. Given that the number of calibration constraints (option prices) is finite and not very large, there may be many Lévy triplets which reproduce call prices with equal precision. This means that the error landscape may have flat regions in which the error has a low sensitivity to variations in A and ν.

One may think that in a parametric model with few parameters one will not encounter this problem since there are (many) more options than parameters. This is not true, as illustrated by the following empirical example. Figure 2.1 represents the magnitude of the quadratic pricing error for the Merton model (1.16) on a data set of DAX index options, as a function of the diffusion coefficient σ and the jump intensity λ, other parameters remaining fixed. It can be observed that if one increases the diffusion volatility while simultaneously decreasing the jump

intensity in a suitable manner, the calibration error changes very little: there is a long "valley" in the error landscape (highlighted by the dashed white line in Figure 2.1). A gradient descent method will typically succeed in locating the valley but will stop at a more or less random point in it. At first glance this does not seem to be too much of a problem: since the algorithm finds the valley's bottom, the best calibration quality will be achieved anyway. However, after a small change in option prices, the outcome of this calibration algorithm may shift a long way along the valley. This means that if the calibration is performed every day, one may come up with wildly oscillating parameters of the Lévy process, leading to important changes in the quantities computed from these parameters, like prices of exotic options, even if the market option prices change very little.

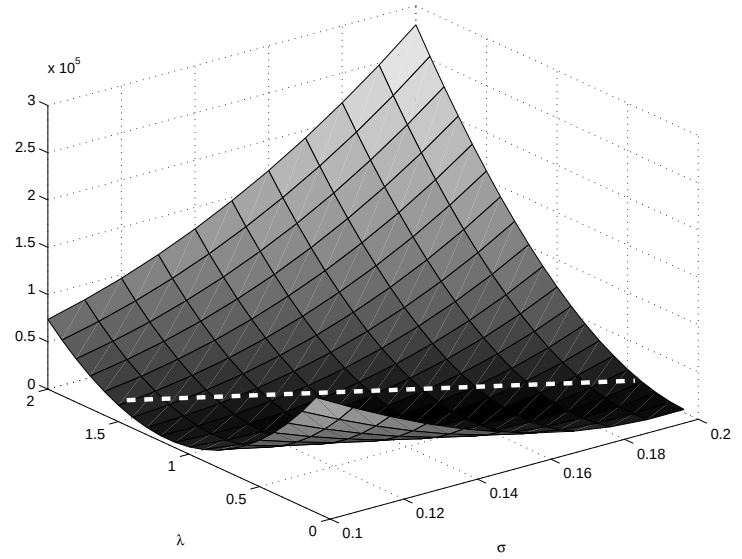

Figure 2.1: Sum of squared differences between market prices (DAX options maturing in 10 weeks) and model prices in Merton model, as a function of parameters σ and λ the other ones being fixed. Dashed white line shows the "valley" along which the error function changes very little.

Figure 2.2 illustrates the same problem in the non-parametric setting. The two graphs represent the result of a non-linear least squares minimization where the variable is the vector of discretized values of ν on a grid (see Section 3.1). In both cases the same option prices are used, the only difference being the starting points of the optimization routines. In the first case (solid line) a Merton model with intensity $\lambda = 1$ is used and in the second case (dashed line)

we used a Merton model with intensity $\lambda = 5$. As can be seen in Figure 2.2 (left graph), the results of the minimization are totally different! However, although the calibrated measures are different, the prices that they correspond to are almost the same (see Figure 2.2, right graph), and the final values of the calibration functional for the two curves differ very little.

The identification problem will be addressed in the next section by adding information in the form of a prior Lévy process to the calibration problem.

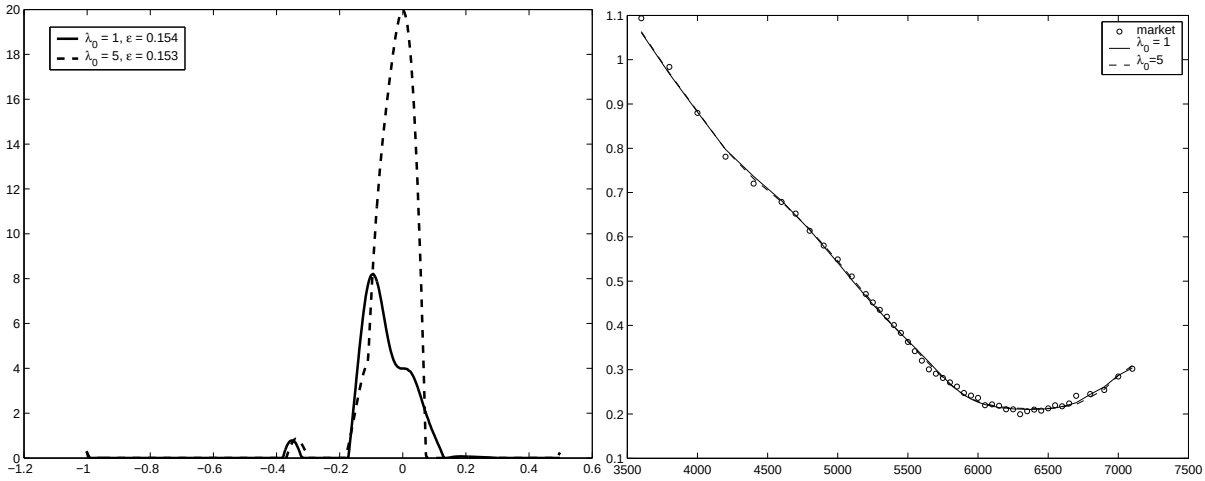

Figure 2.2: Left: Lévy measures calibrated to DAX option prices, maturity 3 months via non-linear least squares method. The starting jump measure for both graphs is Gaussian; the jump intensity λ_0 is initialized to 1 for the solid curve and to 5 for the dashed one. ε denotes the value of the calibration functional when the gradient descent algorithm stops. Right: implied volatility smiles corresponding to these two measures.

2.1.2 Existence of solutions

In this section, we will first give an (artificial) example which shows that the least squares formulation (2.4) does not in general guarantee the existence of solution. However, under additional conditions we are able to prove the existence of a solution of the least squares calibration problem.

Example 2.1. Suppose that $S_0 = 1$, there are no interest rates or dividends and the (equally weighted) market data consist of the following two observations:

$$C_M(T = 1, K = 1) = 1 - e^{-\lambda} \quad \text{and} \quad C_M(T = 1, K = e^{\lambda}) = 0, \tag{2.7}$$

with some $\lambda > 0$. It is easy to see that these prices are, for example, compatible with the (martingale) asset price process $S_t = e^{\lambda t} 1_{t \leq \tau_1}$, where τ_1 is the time of the first jump of a Poisson process with intensity λ. We will show that if the market data are given by (2.7), the calibration problem (2.4) does not admit a solution.

Absence of arbitrage implies that in every risk-neutral model Q, for fixed T, $C^Q(T,K)$ is a convex function of K and that $C^Q(T, K=0) = 1$. The only convex function which satisfies this equality and passes through the market data points (2.7) is given by $C(T=1,K) = (1-Ke^{-\lambda})^+$. Therefore, in every arbitrage-free model that is an exact solution of the calibration problem with market data (2.7), for every $K \geq 0$, $P[S_1 \leq K] = e^{-\lambda} 1_{K \leq e^\lambda}$. Since in an exponential Lévy model $P[S_1 > 0] = 1$, there is no risk-neutral exponential Lévy model for which $\|C_M - C^Q\|_w = 0$.

On the other hand, $\inf_{Q \in \mathcal{M} \cap \mathcal{L}} \|C_M - C^Q\|_w^2 = 0$. Indeed, let $\{N_t\}_{t \geq 0}$ be a Poisson process with intensity λ. Then for every n, the process

$$X_t^n := -nN_t + \lambda t(1 - e^{-n}) \tag{2.8}$$

belongs to $\mathcal{M} \cap \mathcal{L}$ and

$$\lim_{n \to \infty} E[(e^{X_t^n} - K)^+] = \lim_{n \to \infty} \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \left(e^{-nk + \lambda t(1-e^{-n})} - K \right)^+ = (1 - Ke^{-\lambda t})^+.$$

We have shown that $\inf_{Q \in \mathcal{M} \cap \mathcal{L}} \|C_M - C^Q\|^2 = 0$ and that for no Lévy process $Q \in \mathcal{M} \cap \mathcal{L}$, $\|C_M - C^Q\|^2 = 0$. Together this entails that the calibration problem (2.4) does not admit a solution.

This example makes clear that to solve the calibration problem (2.4), we must at least impose a bound on the jumps of the solution. The following theorem provides an existence result under this and another important condition (informally speaking, to control the variance of the solution we need to find a Lévy process for which the pricing error is already sufficiently low). In the theorem and below $\mathcal{L}_B$ denotes the sets of all probabilities $P \in \mathcal{L}$ such that $P[|\Delta X_t| \leq B \ \forall t : 0 \leq t \leq T_\infty] = 1$.

Theorem 2.1. *Suppose that for some $Q_0 \in \mathcal{M} \cap \mathcal{L}_B$ and some couple (T_0, K_0) with nonzero weight $w_0 := w(\{T_0, K_0\}) > 0$,*

$$\|C^{Q_0} - C_M\|_w < \sqrt{w_0}(S_0 - C_M(T_0, K_0)). \tag{2.9}$$

Then the problem (2.4) *has a solution in* $\mathcal{M} \cap \mathcal{L}_B$*: there exists* $Q^* \in \mathcal{M} \cap \mathcal{L}_B$*, such that*

$$\|C_M - C^{Q^*}\|_w^2 = \inf_{Q \in \mathcal{M} \cap \mathcal{L}_B} \|C_M - C^Q\|_w^2.$$

The proof is based on the following lemmas.

Lemma 2.2. *The pricing error functional* $Q \mapsto \|C_M - C^Q\|_w^2$*, defined by* (2.2)*, is uniformly bounded and weakly continuous on* $\mathcal{M} \cap \mathcal{L}$*.*

Proof. From Equation (1.14), $C^Q(T,K) \leq S_0$. Absence of arbitrage in the market implies that the market option prices satisfy the same condition. Therefore, $(C_M(T,K) - C^Q(T,K))^2 \leq S_0^2$ and since w is a probability measure, $\|C_M - C^Q\|_w^2 \leq S_0^2$.

Let $\{Q_n\}_{n\geq 1} \subset \mathcal{M} \cap \mathcal{L}$ and $Q \in \mathcal{M} \cap \mathcal{L}$ be such that $Q_n \Rightarrow Q$. For all T, K,

$$\begin{aligned}
\lim_n C^{Q_n}(T,K) &= e^{-rT} \lim_n E^{Q_n}[(S_0 e^{rT+X_T} - K)^+] \\
&= e^{-rT} \lim_n E^{Q_n}[S_0 e^{rT+X_T} - K] + e^{-rT} \lim_n E^{Q_n}[(K - S_0 e^{rT+X_T})^+] \\
&= S_0 - K e^{-rT} + e^{-rT} E^Q[(K - S_0 e^{rT+X_T})^+] = C^Q(T,K).
\end{aligned}$$

Therefore, by the dominated convergence theorem, $\|C_M - C^{Q_n}\|_w^2 \to \|C_M - C^Q\|_w^2$. □

Lemma 2.3. *For all* $B > 0$*,* $T > 0$*,* $K > 0$ *and all* C *with* $0 < C < S_0$*, the set of Lévy processes* $Q \in \mathcal{M} \cap \mathcal{L}_B$ *satisfying* $C^Q(T,K) \leq C$ *is relatively weakly compact.*

Proof. By Prohorov's theorem, weak relative compactness is implied by the tightness of this family of probability measures. Since the jumps are bounded, to prove the tightness, by Proposition 1.6 it is enough to show that there exist constants C_1 and C_2 such that for every Lévy process $Q \in \mathcal{M} \cap \mathcal{L}_B$ such that $C^Q(T,K) \leq C$, its characteristic triplet (A, ν, γ) satisfies $|\gamma| \leq C_1$ and $A + \int_{-1}^1 x^2 \nu(dx) \leq C_2$.

By the risk-neutrality condition (1.2),

$$\gamma = -\frac{A}{2} - \int_{-B}^{B} (e^x - 1 - x 1_{|x|\leq 1}) \nu(dx)$$

It is easy to see that $|e^x - 1 - x1_{|x|\leq 1}| \leq e^{B \wedge 1} x^2$ for $x \in (-\infty, B]$ and therefore $|\gamma| \leq e^{B\wedge 1}(A + \int_{-\infty}^{\infty} x^2 \nu(dx))$. This means that to prove the lemma, it is sufficient to show that

$A + \int_{-\infty}^{\infty} x^2 \nu(dx) < C_3$ for some constant C_3, independent on the choice of the Lévy process Q. Since for all $\alpha \in [0,1]$, $e^x \wedge 1 \leq e^{\alpha x}$, we have:

$$S_0 - C^Q(T,K) = Ke^{-rT} E[e^{X_T - m} \wedge 1] \leq Ke^{-rT} E[e^{\alpha(X_T - m)}] = S_0^\alpha (Ke^{-rT})^{1-\alpha} E[e^{\alpha X_T}],$$

where $m = \log(Ke^{-rT}/S_0)$. The right-hand side can be computed using the Lévy-Khintchine formula:

$$E[e^{\alpha X_T}] = \exp T \left\{ -\frac{A}{2}(\alpha - \alpha^2) - \int_{-\infty}^{\infty} (\alpha e^{\alpha x} - e^{\alpha x} - \alpha + 1)\nu(dx) \right\}.$$

Choosing $\alpha = 1/2$, the above reduces to

$$E[e^{X_T/2}] = \exp T\{ -\frac{A}{8} - \frac{1}{2} \int_{-\infty}^{\infty} (e^{x/2} - 1)^2 \nu(dx) \}.$$

On the other hand, it is easy to check that

$$\frac{A}{8} + \frac{1}{2} \int_{-\infty}^{\infty} (e^{x/2} - 1)^2 \nu(dx) \geq \frac{e^{-B}}{8} (A + \int_{-\infty}^{\infty} x^2 \nu(dx))$$

and therefore

$$A + \int_{-\infty}^{\infty} x^2 \nu(dx) \leq \frac{8e^B}{T} \log \frac{\sqrt{S_0 K e^{-rT}}}{S_0 - C^Q(T,K)} \leq \frac{8e^B}{T} \log \frac{\sqrt{S_0 K e^{-rT}}}{S_0 - C},$$

which finishes the proof of the lemma. □

In the following lemma, $\mathcal{L}_B^+$ denotes the sets of all probabilities $P \in \mathcal{L}$ such that $P[\Delta X_t \leq B \ \forall t : 0 \leq t \leq T_\infty] = 1$.

Lemma 2.4. *Both $\mathcal{M} \cap \mathcal{L}_B$ and $\mathcal{M} \cap \mathcal{L}_B^+$ are weakly closed for every $B > 0$.*

Proof. We give the proof for $\mathcal{M} \cap \mathcal{L}_B$; the proof for $\mathcal{M} \cap \mathcal{L}_B^+$ can be done in a similar fashion. Let $\{Q_n\}_{n=1}^\infty \subset \mathcal{M} \cap \mathcal{L}_B$ with characteristic triplets (A_n, ν_n, γ_n^h) with respect to a continuous bounded truncation function h, satisfying $h(x) = x$ in a neighborhood of 0, and let Q be a Lévy process with characteristic triplet (A, ν, γ^h) with respect to h, such that $Q_n \Rightarrow Q$. Note that a sequence of Lévy processes cannot converge to anything other than a Lévy process because due to convergence of characteristic functions, the limiting process must have stationary and independent increments. Define a function f by

$$f(x) := \begin{cases} 0, & |x| \leq B, \\ 1, & |x| \geq 2B, \\ \frac{|x| - B}{B} & B < |x| < 2B. \end{cases}$$

By Proposition 1.7, $\int_{-\infty}^{\infty} f(x)\nu(dx) = \lim_{n\to\infty}\int_{-\infty}^{\infty} f(x)\nu_n(dx) = 0$, which implies that the jumps of Q are bounded by B.

Define a function g by

$$g(x) := \begin{cases} e^x - 1 - h(x) - \frac{1}{2}h^2(x), & x \le B, \\ e^B - 1 - h(B) - \frac{1}{2}h^2(B), & x > B. \end{cases}$$

Then, by Proposition 1.7 and because Q_n satisfies the martingale condition (1.2) for every n,

$$\begin{aligned}\gamma^h + \frac{A}{2} + \int_{-\infty}^{\infty}(e^x - 1 - h(x))\nu(dx) &= \gamma^h + \frac{A + \int_{-\infty}^{\infty} h^2(x)\nu(dx)}{2} + \int_{-\infty}^{\infty} g(x)\nu(dx) \\ &= \lim_{n\to\infty}\left\{\gamma_n^h + \frac{A_n + \int_{-\infty}^{\infty} h^2(x)\nu_n(dx)}{2} + \int_{-\infty}^{\infty} g(x)\nu_n(dx)\right\} = 0,\end{aligned}$$

which shows that Q also satisfies the condition (1.2). □

Proof of Theorem 2.1. Let $\{Q_n\}_{n\ge1} \subset \mathcal{M}\cap\mathcal{L}_B$ be such that

$$\begin{aligned}&\lim_{n\to\infty}\|C_M - C^{Q_n}\|_w^2 = \inf_{Q\in\mathcal{M}\cap\mathcal{L}_B}\|C_M - C^Q\|_w^2 \\ &\text{and}\quad \|C_M - C^{Q_n}\|_w^2 \le \|C_M - C^{Q_0}\|_w^2 \quad \text{for all } n.\end{aligned}$$

Condition (2.9) implies that for every n,

$$\begin{aligned}|S_0 - C^{Q_n}(T_0,K_0)| &\ge |S_0 - C_M(T_0,K_0)| - |C_M(T_0,K_0) - C^{Q_n}(T_0,K_0)| \\ &\ge |S_0 - C_M(T_0,K_0)| - \frac{\|C_M - C^{Q_n}\|_w}{\sqrt{w_0}} \\ &\ge |S_0 - C_M(T_0,K_0)| - \frac{\|C_M - C^{Q_0}\|_w}{\sqrt{w_0}} > 0.\end{aligned}$$

Therefore, by Lemmas 2.3 and 2.4, there exists a subsequence $\{Q_{n_m}\}_{m\ge1}$ of $\{Q_n\}_{n\ge1}$ and $Q^* \in \mathcal{M}\cap\mathcal{L}_B$ such that $Q_{n_m} \Rightarrow Q$. By Lemma 2.2,

$$\|C_M - C^{Q^*}\|_w^2 = \lim_{m\to\infty}\|C_M - C^{Q_{n_m}}\|_w^2 = \inf_{Q\in\mathcal{M}\cap\mathcal{L}_B}\|C_M - C^Q\|_w^2,$$

which shows that Q^* is a solution of the least squares calibration problem (2.4). □

2.1.3 Continuity

Market option prices are typically defined up to a bid-ask spread and the close prices used for calibration may therefore contain numerical errors. If the solution of the calibration problem

is not continuous with respect to market data, these small errors may dramatically alter the result of calibration, rendering it completely useless. On the other hand, even if market data did not contain errors, in absence of continuity, small daily changes in prices could lead to large variations of calibrated parameters and of other quantities computed using these parameters (like prices of exotic options).

When the calibration problem has more than one solution, care should be taken in defining what is meant by continuity. In the sequel, we will use the following definition (see, e.g. [40]), that applies to all calibration problems, discussed in this chapter.

Definition 2.1. The solutions of a calibration problem are said to depend continuously on input data at the point C_M if for every sequence of data sets $\{C_M^n\}_{n\geq 0}$ such that $\|C_M^n - C_M\|_w \xrightarrow[n\to\infty]{} 0$, if Q_n is a solution of the calibration problem with data C_M^n then

1. $\{Q_n\}_{n\geq 1}$ has a weakly convergent subsequence $\{Q_{n_m}\}_{m\geq 1}$.

2. The limit Q of every weakly convergent subsequence of $\{Q_n\}_{n\geq 1}$ is a solution of the calibration problem with data C_M.

Note that if the solution of the calibration problem with the limiting data C_M is unique, this definition reduces to the standard definition of continuity, because in this case every subsequence of $\{Q_n\}$ has a further subsequence converging towards Q, which implies that $Q_n \Rightarrow Q$.

It is easy to construct an example of market data leading to a least squares calibration problem (2.4) that does not satisfy the above definition.

Example 2.2. Suppose that $S_0 = 1$, there are no interest rates or dividends and the market data for each n are given by a single observation:

$$C_M^n(T=1, K=1) = E[(e^{X_1^n} - 1)^+] \quad \text{for } n \geq 1 \quad \text{and} \quad C_M(T=1, K=1) = 1 - e^{-\lambda},$$

where X_t^n is defined by Equation (2.8) and $\lambda > 0$. Then $\|C_M^n - C_M\|_w \xrightarrow[n\to\infty]{} 0$ and X_t^n is clearly a solution for data C_M^n, but the sequence $\{X_t^n\}$ has no convergent subsequence (cf. Proposition 1.7).

Under conditions, similar to those of Theorem 2.1, Lemmas 2.2–2.4 allow to prove a continuity result for the least squares calibration problem (2.4).

Proposition 2.5. *Suppose that there exists a couple (T_0, K_0) with $w_0 := w(\{T_0, K_0\}) > 0$ and let C_M be such that the condition (2.9) is satisfied for some $Q_0 \in \mathcal{M} \cap \mathcal{L}_B$. Then the solutions of the least squares calibration problem (2.4) on $\mathcal{M} \cap \mathcal{L}_B$ depend continuously on the market data at the point C_M.*

Proof. Let $\{C_M^n\}_{n\geq 0}$ be a sequence of data such that $\|C_M^n - C_M\|_w \xrightarrow[n\to\infty]{} 0$, and for every n let Q_n be a solution of the calibration problem (2.4) on $\mathcal{M} \cap \mathcal{L}_B$ with data C_M^n (we can suppose without loss of generality that a solution exists for all n because it exists starting with a sufficiently large n by Theorem 2.1). Then, using the triangle inequality several times, we obtain:

$$\begin{aligned}
&|S_0 - C^{Q_n}(T_0, K_0)| \\
&\quad \geq |S_0 - C_M(T_0, K_0)| - |C_M^n(T_0, K_0) - C^{Q_n}(T_0, K_0)| - |C_M^n(T_0, K_0) - C_M(T_0, K_0)| \\
&\quad \geq |S_0 - C_M(T_0, K_0)| - \frac{\|C_M^n - C^{Q_0}\|_w}{\sqrt{w_0}} - |C_M^n(T_0, K_0) - C_M(T_0, K_0)| \\
&\quad \geq |S_0 - C_M(T_0, K_0)| - \frac{\|C_M - C^{Q_0}\|_w}{\sqrt{w_0}} - 2\frac{\|C_M^n - C_M\|_w}{\sqrt{w_0}} > C' > 0
\end{aligned}$$

for some C', starting from a sufficiently large n. Therefore, by Lemmas 2.3 and 2.4, $\{Q_n\}$ has a subsequence that converges weakly towards some $Q^* \in \mathcal{M} \cap \mathcal{L}_B$.

Let $\{Q_{n_m}\} \subseteq \{Q_n\}$ with $Q_{n_m} \Rightarrow Q^* \in \mathcal{M} \cap \mathcal{L}_B$ and let $Q \in \mathcal{M} \cap \mathcal{L}_B$. Using Lemma 2.2 and the triangle inequality, we obtain:

$$\begin{aligned}
\|C^{Q^*} - C_M\|_w = \lim_m \|C^{Q_{n_m}} - C_M\|_w &\leq \liminf_m \{\|C^{Q_{n_m}} - C_M^{n_m}\|_w + \|C_M^{n_m} - C_M\|_w\} \\
&\leq \liminf_m \|C^{Q_{n_m}} - C_M^{n_m}\|_w \leq \liminf_m \|C^Q - C_M^{n_m}\|_w \leq \|C^Q - C_M\|_w,
\end{aligned}$$

which shows that Q^* is indeed a solution of the calibration problem (2.4) with data C_M. □

2.1.4 Numerical difficulties of least squares calibration

A major obstacle for the numerical implementation of the least squares calibration is the non-convexity of the optimization problem (2.4), which is due to the non-convexity of the domain $(\mathcal{M} \cap \mathcal{L})$, where the pricing error functional $Q \mapsto \|C^Q - C_M\|^2$ is to be optimized. Due to this difficulty, the pricing error functional may have several local minima, and the gradient descent

algorithm used for numerical optimization may stop in one of these local minima, leading to a much worse calibration quality than that of the true solution.

Figure 2.3 illustrates this effect in the (parametric) framework of the variance gamma model (1.18). The left graph shows the behavior of the calibration functional in a small region around the global minimum. Since in this model there are only three parameters, the identification problem is not present, and a nice profile appearing to be convex can be observed. However, when we look at the calibration functional on a larger scale (κ changes between 1 and 8), the convexity disappears and we observe a ridge (highlighted with a dashed black line), which separates two regions: if the minimization is initiated in the region (A), the algorithm will eventually locate the minimum, but if we start in the region (B), the gradient descent method will lead us away from the global minimum and the required calibration quality will never be achieved.

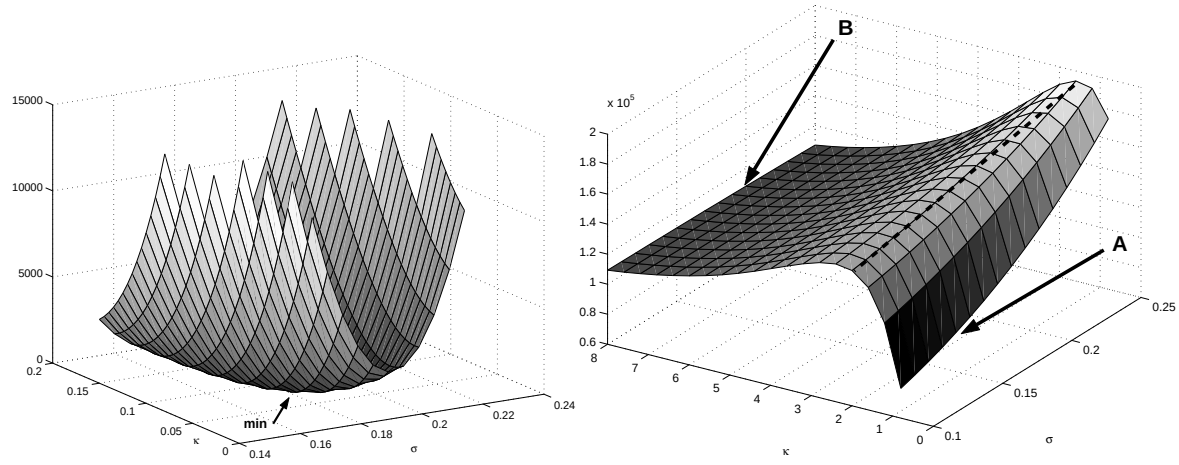

Figure 2.3: Sum of squared differences between market prices (DAX options maturing in 10 weeks) and model prices in the variance gamma model (1.18) as a function of σ and κ, the third parameter being fixed. Left: small region around the global minimum. Right: error surface on a larger scale.

2.2 Selection of solutions using relative entropy

When the option pricing constraints do not determine the exponential Lévy model completely (this is for example the case if the number of constraints is finite), additional information may

be introduced into the problem by specifying a *prior model*: we suppose given a Lévy process P and look for the solution of the problem (2.4) that has the smallest relative entropy with respect to P. For two probabilities P and Q on the same measurable space $(\Omega, \mathcal{F})$, the relative entropy of Q with respect to P is defined by

$$I(Q|P) = \begin{cases} E^P\left[\frac{dQ}{dP}\log\frac{dQ}{dP}\right], & \text{if } Q \ll P, \\ \infty, & \text{otherwise,} \end{cases} \tag{2.10}$$

where by convention $x \log x = 0$ when $x = 0$. For a time horizon $T \leq T_\infty$ we define $I_T(Q|P) := I(Q|_{\mathcal{F}_T}|P|_{\mathcal{F}_T})$.

Minimum entropy least squares calibration problem Given prices C_M of call options and a prior Lévy process P, find a least squares solution $Q^* \in \mathcal{Q}^{LS}$, such that

$$I(Q^*|P) = \inf_{Q \in \mathcal{Q}^{LS}} I(Q|P). \tag{2.11}$$

In the sequel, any such Q^* will be called a minimum entropy least squares solution (MELSS) and the set of all such solutions will be denoted by $\mathcal{Q}^{MELS}$.

The prior probability P must reflect our a priori knowledge about the risk-neutral distribution of the underlying. A natural choice of prior, ensuring absence of arbitrage in the calibrated model, is an exponential Lévy model, estimated from the time series of returns. The effect of the choice of prior on the solution of the calibration problem and the possible ways to choose it in practice are discussed in Section 3.2.

Using relative entropy for selection of solutions removes, to some extent, the identification problem of least-squares calibration. Whereas in the least squares case, this was an important nuisance, now, if two measures reproduce market option prices with the same precision and have the same entropic distance to the prior, this means that both measures are compatible with all the available information. Knowledge of many such probability measures instead of one may be seen as an advantage, because it allows to estimate model risk and provide confidence intervals for the prices of exotic options. However, the calibration problem (2.11) remains ill-posed: since the minimization of entropy is done over the results of least squares calibration, problem (2.11) may only admit a solution if problem (2.4) does. Also, $\mathcal{Q}^{LS}$ is not necessarily a compact set,

so even if it is nonempty, the least squares solution minimizing the relative entropy may not exist. Other undesirable properties like absence of continuity and numerical instability are also inherited from the least squares approach. In Section 2.5 we will propose a regularized version of problem (2.11) that does not suffer from these difficulties.

The choice of relative entropy as a method for selection of solutions of the calibration problem is driven by the following considerations:

- The relative entropy is a convenient notion of distance for probability measures. Indeed, it is convex, nonnegative functional of Q for fixed P, equal to zero if and only if $\frac{dQ}{dP} = 1$ P-a.s. To see this, observe that

$$E^P\left[\frac{dQ}{dP}\log\frac{dQ}{dP}\right] = E^P\left[\frac{dQ}{dP}\log\frac{dQ}{dP} - \frac{dQ}{dP} + 1\right],$$

  and that $z\log z - z + 1$ is a convex nonnegative function of z, equal to zero if and only if $z = 1$.

- The relative entropy of two Lévy processes is easily expressed in terms of their characteristic triplets (see Theorem 2.9).

- Relative entropy, also called Kullback-Leibler distance, is a well-studied object. It appears in many domains including the theory of large deviations and the information theory. It has also already been used in finance for pricing and calibration (see Section 2.3). We can therefore use the known properties of this functional (see e.g. [93]) as a starting point of our study.

The minimum entropy least squares solution need not always exist, but if the prior is chosen correctly, that is, if there exists at least one solution of problem (2.4) with finite relative entropy with respect to the prior, then the MELSS will also exist, as shown by the following lemma. We recall that $\mathcal{L}_{NA}$ stands for the set of Lévy processes, satisfying the no arbitrage conditions of Proposition 1.8.

Lemma 2.6. *Let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ for some $B > 0$ and suppose that the problem* (2.4) *admits a solution Q^+ with $I(Q^+|P) = C < \infty$. Then the problem* (2.11) *admits a solution.*

Proof. The proof of this lemma uses the properties of the relative entropy functional (2.10) that will be proven in Section 2.4 below. Under the condition of the lemma, it is clear that the

solution Q^* of problem (2.11), if it exists, satisfies $I(Q^*|P) \leq C$. This entails that $Q^* \ll P$, which means by Proposition 1.5 that $Q^* \in \mathcal{L}_B^+$. Therefore, Q^* belongs to the set

$$\mathcal{L}_B^+ \cap \{Q \in \mathcal{M} \cap \mathcal{L} : \|C^Q - C_M\| = \|C^{Q^+} - C_M\|\} \cap \{Q \in \mathcal{L} : I(Q|P) \leq C\}. \tag{2.12}$$

Lemma 2.10 and the Prohorov's theorem entail that the level set $\{Q \in \mathcal{L} : I(Q|P) \leq C\}$ is relatively weakly compact. On the other hand, by Corollary 2.1, $I(Q|P)$ is weakly lower semicontinuous with respect to Q for fixed P. Therefore, the set $\{Q \in \mathcal{P}(\Omega) : I(Q|P) \leq C\}$ is weakly closed and since by Lemma 2.4, $\mathcal{M} \cap \mathcal{L}_B^+$ is also weakly closed, the set $\mathcal{M} \cap \mathcal{L}_B^+ \cap \{Q \in \mathcal{L} : I(Q|P) \leq C\}$ is weakly compact. Lemma 2.2 then implies that the set (2.12) is also weakly compact. Since $I(Q|P)$ is weakly lower semicontinuous, it reaches its minimum on this set. □

2.3 The use of relative entropy for pricing and calibration: review of literature

2.3.1 Pricing with minimal entropy martingale measure

It is known (see [25, Theorem 4.7] or [27, Remark 10.3]) that all exponential Lévy models with the exception of the Black-Scholes model and the model driven by the compensated Poisson process correspond to incomplete markets meaning that there exists no *unique* equivalent martingale measure.

In the incomplete market setting many authors including [24, 41, 42, 49, 72, 73, 89] have investigated the minimal entropy martingale measure, that is, the pricing measure that minimizes the relative entropy with respect to the historical probability P. A probability measure $Q^* \in \mathcal{M}$ is called *minimal entropy martingale measure* if

$$I(Q^*|P) = \min_{Q \in \mathcal{M}} I(Q|P). \tag{2.13}$$

Frittelli [42] proves the existence of a minimal entropy martingale measure provided that the stock price process is bounded and that there exists a martingale measure with finite relative entropy with respect to P. He further shows that if there exists an *equivalent* martingale measure with finite relative entropy with respect to P, the MEMM is also equivalent to P.

Several authors including [24, 49, 73] discuss minimal entropy martingale measures for exponential Lévy models. The following result is proven in Miyahara and Fujiwara [73].

Theorem 2.7. *Let P be a Lévy process with characteristic triplet (A, ν, γ). If there exists a constant $\beta \in \mathbb{R}$ such that*

$$\int_{\{x>1\}} e^x e^{\beta(e^x-1)}\nu(dx) < \infty, \tag{2.14}$$
$$\gamma + \left(\frac{1}{2}+\beta\right)A + \int_{|x|\leq 1}\left\{(e^x-1)e^{\beta(e^x-1)} - x\right\}\nu(dx) + \int_{|x|>1}(e^x-1)e^{\beta(e^x-1)}\nu(dx) = 0,$$

then there exists a minimal entropy martingale measure Q^* with the following properties:

1. *The measure Q^* corresponds to a Lévy process: $Q^* \in \mathcal{L}$ with characteristic triplet*

$$A^* = A,$$
$$\nu^*(dx) = e^{\beta(e^x-1)}\nu(dx),$$
$$\gamma^* = \gamma + \beta A + \int_{|x|\leq 1} x(e^{\beta(e^x-1)} - 1)\nu(dx).$$

2. *The measure Q^* is an* equivalent *martingale measure: $Q^* \sim P$.*

3. *The minimal relative entropy is given by*

$$I(Q^*|P) = -T\left\{\frac{\beta}{2}(1+\beta)A + \beta\gamma + \int_{-\infty}^{\infty}\{e^{\beta(e^x-1)} - 1 - \beta x 1_{|x|\leq 1}\}\nu(dx)\right\}. \tag{2.15}$$

Remark 2.1. It is easy to show, along the lines of the proof of Proposition 1.8, that condition (2.14) is satisfied, in particular, if $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ for some $B > 0$, which corresponds to a stock price process with jumps bounded from above in a market without arbitrage opportunity.

Since large positive jumps do not happen very often in real markets, (2.14) turns out to be much less restrictive (and easier to check) than the general hypotheses in [42]. This shows that the notion of MEMM is especially useful and convenient in the context of exponential Lévy models.

In addition to its computational tractability, the interest of the minimal entropy martingale measure is due to its economic interpretation as the pricing measure that corresponds to the limit of utility indifference price for the exponential utility function when the risk aversion coefficient tends to zero. Consider an investor with initial endowment c, whose utility function is given by

$$U_\alpha(x) := 1 - e^{-\alpha x}, \tag{2.16}$$

where α is a risk-aversion coefficient. Given some set of admissible trading strategies Θ, the utility indifference price $p_\alpha(c, H)$ of a claim H for this investor is defined as the solution of the following equation:

$$\sup_{\theta\in\Theta} E[U_\alpha(c + p_\alpha(c, H) + \int_{(0,T]} \theta_u dS_u - H)] = \sup_{\theta\in\Theta} E[U_\alpha(c + \int_{(0,T]} \theta_u dS_u)].$$

Due to the special form (2.16) of the utility function, the initial endowment c cancels out of the above equation and we see that

$$p_\alpha(c, H) = p_\alpha(0, H) := p_\alpha(H).$$

Using the results in [33], Miyahara and Fujiwara [73] established the following properties of utility indifference price in exponential Lévy models. Similar results have been obtained by El Karoui and Rouge [39] in the setting of continuous processes.

Proposition 2.8. *Let (X, P) be a Lévy process such that $P \in \mathcal{L}_B^+ \cap \mathcal{L}_{NA}$, and let Q^* be the MEMM defined by* (2.13)*. Let $S_t := e^{X_t}$ and let Θ include all predictable S-integrable processes θ such that $\int_{(0,t]} \theta_u dS_u$ is a martingale for each local martingale measure Q, with $I(Q|P) < \infty$. Then the corresponding utility indifference price $p_\alpha(H)$ of a bounded claim H has the following properties:*

1. *$p_\alpha(H) \geq E^{Q^*}[H]$ for any $\alpha > 0$.*
2. *If $0 < \alpha < \beta$ then $p_\alpha(H) \leq p_\beta(H)$.*
3. *$\lim_{\alpha\downarrow 0} p_\alpha(H) = E^{Q^*}[H]$.*

The price of a claim H computed under the MEMM thus turns out to be the highest price at which all investors with exponential utility function will be willing to buy this claim.

Kallsen [58] defines *neutral* derivative prices for which the optimal trading strategy consists in having no contingent claim in one's portfolio. This approach to valuation in incomplete markets corresponds to the notion of *fair price* introduced by Davis [31]. Kallsen further shows that such prices are unique and correspond to a linear arbitrage-free pricing system, defined by an equivalent martingale measure (neutral pricing measure). If the utility function of an investor has the form (2.16), the neutral pricing measure coincides with the minimal entropy

martingale measure. The neutral pricing measure Q^* also corresponds to the *least favorable market completion* from the point of view of the investor in the following sense. Let

$$V(c,Q) := \sup\{E[U(c+X-E^Q[X])] : X \text{ is } \mathcal{F}_T\text{-measurable}\}$$

be the maximum possible expected utility that an investor with initial endowment c and utility function U can get by trading in a market where the price of every contingent claim X is equal to $E^Q[X]$. Then

$$V(c,Q^*) = \inf_{Q\in EMM(P)} V(c,Q),$$

that is, the neutral pricing measure minimizes the maximum possible expected utility that an investor can get in a completed market, over all possible arbitrage-free completions. The notion of neutral pricing measure thus coincides with the *minimax* measures studied in [16], [49] and other papers.

2.3.2 Calibration via relative entropy minimization

Despite its analytic tractability and interesting economic interpretation, the MEMM has an important drawback which makes it impossible to use for pricing in real markets: it does not take into account the information obtained from prices of traded options. To tackle this problem, Goll and Rüschendorf [48] introduced the notion of minimal distance martingale measure consistent with observed market prices. In particular, given prices of call options $\{C_M(T_i,K_i)\}_{i=1}^N$, a probability measure $Q^* \in \mathcal{M}$ is called consistent minimal entropy martingale measure (CMEMM) if

$$I(Q^*|P) = \min_Q I(Q|P),$$

where the minimum is taken over all martingale measures Q such that $C_M(T_i,K_i) = C^Q(T_i,K_i)$ for $i = 1,\dots,N$.

Kallsen [58] shows that consistent minimal distance martingale measures correspond to constant demand derivative pricing (when the optimal portfolio must contain not zero but a fixed amount of each traded derivative), thus providing an economic rationale for distance minimization under constraints, and, in particular, for entropy minimization.

A drawback of CMEMM is that it may be very difficult to compute numerically. In particular, if X is a Lévy process under the historical measure P, it will in general no longer be

a Lévy process under a consistent minimal entropy martingale measure and no analytic results similar to Theorem 2.7 are available for the constrained case. For example, let $\{X_t\}_{t\geq 0}$ be a real-valued Lévy process on (A, ν, γ) such that for every t,

$$X_t = N'_t - N''_t, \tag{2.17}$$

where N' and N'' are independent Poisson processes with intensity 1 under P. It follows from Proposition 1.5 that the set of Lévy processes equivalent to P and satisfying the martingale condition (1.2) contains all processes under which X has the form (2.17) where N' and N'' are still independent and have respective intensities $\lambda' = \lambda$ and $\lambda'' = e\lambda$ for some $\lambda > 0$. The set of all equivalent martingale measures under which X remains a Lévy process is thus parametrized by one parameter λ. If one allows X to be an additive process (process with independent increments, continuous in probability) under the new measure, the class of equivalent martingale measures is much larger: it follows from Theorem IV.4.32 in [54] that N' and N'' can now have variable (but deterministic) intensities $\lambda'(t) = \lambda(t)$ and $\lambda''(t) = e\lambda(t)$ for some function $\lambda : [0, T_\infty] \to (0, \infty)$. It is clear that many market data sets (involving several maturities) can be reproduced by an equivalent martingale measure, under which X is an additive process but not by a martingale measure under which X is a Lévy process, which implies that under the consistent minimal entropy martingale measure X will not be a Lévy process.

Stutzer [91] suggests a three-step algorithm for numerical evaluation of derivative prices under the CMEMM in a model with a single time horizon T. This method allows to compute prices of European options with maturity T, consistent with prices of market-quoted European options with the same maturity. First, possible asset price values at time T, $\{P_h\}_{h=1}^N$ and the corresponding probabilities $\hat{\pi}(h)$ are estimated nonparametrically using a histogram estimator from the historical return values. Second, one needs to find the probabilities $\pi^*(h)$, that satisfy the martingale constraint and the pricing constraints and have the smallest relative entropy with respect to $\hat{\pi}$. In this simplified one period model the martingale condition reduces to a single constraint that the discounted expectation of stock price under π^* be equal to its present value. The European options expiring at T can then be priced using these martingale probabilities π^*.

In this paper, Stutzer suggests an information theoretic rationale for using the relative entropy (also called Kullback-Leibler information criterion) for pricing and calibration. From a Bayesian point of view, the historical prices constitute a prior information about the future

risk-neutral distribution of asset prices. This prior must be updated to take into account the martingale constraint and the observed option prices, and it is natural to demand that the updated distribution incorporate no additional information other than the martingale constraint and the pricing constraints. One must therefore minimize some quantitative measure of relative information of the two distributions, and it has been shown (see [52]) that every measure of information satisfying a set of natural axioms must be proportional to Kullback-Leibler relative entropy.

Because it only allows to reconstruct the asset price distribution for one specified time horizon, and only takes into account the prices of options that expire at this horizon, Stutzer's method does not provide any information about the risk-neutral *process* and thus cannot be used to price any derivative that depends on the stock price at times other than T. Avellaneda et al. [3] pursue the same logic further and propose a method allowing to construct a discrete approximation of the law of the process underlying the observed option prices. They consider $N < \infty$ fixed trajectories $\{X^1, \ldots, X^N\}$ that the price process can take, simulated beforehand from a prior model. The new state space Ω' thus contains a finite number of elements: $\Omega' = \{X^1, \ldots, X^N\}$. The new prior P' is the uniform law on Ω': $P'(X^i) = \frac{1}{N}$ and the paper suggests to calibrate the weights (probabilities) of these trajectories $q_i := Q'(X^i)$ to reproduce market-quoted option prices correctly. Minimizing relative entropy $I(Q'|P')$ is then equivalent to maximizing the entropy of Q' and the calibration problem becomes:

$$\text{maximize } -\sum_{i=1}^{N} q_i \log q_i \text{ under constraints } E^{Q'}[H_j] = C_j,\ j = 1, \ldots, M,$$

where H_j are terminal payoffs and C_j the observed market prices of M traded options. Denoting by g_{ij} the payoff of j-th option on the i-th trajectory, and introducing Lagrange multipliers $\lambda_1, \ldots, \lambda_M$, Avellaneda et al. reformulate the calibration problem as a minimax problem:

$$\min_{\lambda} \max_{q} \left\{ -\sum_{i=1}^{N} q_i \log q_i + \sum_{j=1}^{M} \lambda_j \left(\sum_{i=1}^{N} q_i g_{ij} - C_j \right) \right\}.$$

The inner maximum can be computed analytically and, since it is taken over linear functions of $\lambda_1, \ldots, \lambda_M$, yields a convex function of Lagrange multipliers. The outer minimum can then be evaluated numerically using a gradient descent method.

This technique (weighted Monte Carlo) is attractive for numerical computations but has a number of drawbacks from the theoretical viewpoint. First, the result of calibration is a

probability measure on a finite set of paths chosen beforehand; it does not allow to reconstruct a process on the original space Ω. Second, the martingale condition is not imposed in this approach (because this would correspond to an infinite number of constraints). As a result, derivative prices computed with the weighted Monte Carlo algorithm may contain arbitrage opportunities, especially when applied to forward start contracts.

Nguyen [77] studies the convergence of the above method by Avellaneda et al. when the number of paths tends to infinity as well as the stability of the calibration procedure, minimization criteria other than relative entropy, possible ways to impose the (approximate) martingale condition and many other issues related to this calibration methodology. Statistical properties of weighted Monte Carlo estimators are also studied in [46].

Chapter 8 of [77] introduces an interesting theoretical approach to minimal entropy calibration, that is related to the present work and is worth being discussed in more detail. Starting with a very general prior jump-diffusion model P for the stock price S_t of the form

$$dS_t = S_{t-}[b(t, S_{t-})dt + \sigma(t, S_{t-})dW_t + \int_{z\in\mathbb{R}} \Phi(t, S_{t-}, z)(\Pi(dt, dz) - \pi(dt, dz))], \qquad (2.18)$$

where Π is a homogeneous Poisson random measure with intensity measure $\pi(dt, dz) = dt \times \rho(dz)$, and the coefficients satisfy some regularity hypotheses not listed here, Nguyen suggests to find a martingale measure Q^* which reproduces the observed option prices correctly and has the smallest relative entropy with respect to P in a subclass $\mathcal{M}'$ of all martingale measures on $(\Omega, \mathcal{F})$. This subclass $\mathcal{M}'$ contains all martingale measures Q^K under which S_t has the same volatility σ and the compensator of Π is given by $K(t, X_{t-}, z)\pi(dt, dz)$ with

$$K \in \mathcal{K}^H := \{K(t, x, z) : [0, T] \times \mathbb{R} \times \mathbb{R} \to (0, \infty) \text{ Borel with } |\log K| \leq \log(H)\phi_0\}$$

for some $\phi_0 \in L^2(\rho)$ positive with $\|\phi_0\|_\infty \leq 1$. In other words, the object of calibration here is the intensity of jumps of a Markovian jump diffusion.

The fact that the set

$$\mathcal{M}' := \left\{Q^K : K \in \mathcal{K}^H\right\}$$

is convex enables Nguyen to use classical convex analysis methods to study the calibration problem. In particular, he proves the existence and uniqueness of the solution of the (penalized) calibration problem under the condition that the coefficients appearing in (2.18) are sufficiently regular and the drift b and option price constraint C belong to some neighborhood of (r, C_r),

where r is the interest rate and C_r denotes the option prices computed for S_t given by (2.18) with $b = r$. Unfortunately, no insight is given as to the size of this neighborhood.

An important drawback of the calibration methodology advocated by Nguyen [77] is the difficulty of numerical implementation. The solution is obtained by minimizing a certain function of Lagrange multipliers (the value function), which is itself a solution of nonlinear partial integro-differential equation of HJB type, and the author gives no indication about the numerical methods that can be used. In addition, since a finite set of option prices does not typically contain enough information to identify a three-dimensional intensity function $K(t, x, z)$, the choice of the prior process, not discussed in [77], will play a crucial role.

2.3.3 The present study in the context of previous work on this subject

In this thesis, we suppose that the prior probability P is a Lévy process and restrict the class of martingale measures Q that we consider as possible solutions of the calibration problem to those under which (X, Q) remains a Lévy process. Since this class is rather narrow, it may not contain a measure that reproduces the observed option prices exactly. Therefore, in problem (2.11) we only require that these prices be reproduced in the least squares sense.

Restricting the calibration to the class of Lévy processes enables us to compute the solution of the calibration problem numerically using a relatively simple algorithm (see Chapter 3). Our method can thus be seen as a computable approximation of the consistent minimal entropy martingale measure, discussed by Kallsen [58] and Goll and Rüschendorf [48]. However, contrary to the approach by Avellaneda et al. [3], we do not discretize the space of sample paths: our method yields a Lévy process on the initial state space Ω and not on the finite space of trajectories simulated beforehand. The solution of calibration problem (2.11) and its regularized version (2.27) can therefore be used for all types of computations with arbitrary precision. Another advantage of restricting the class of martingale measures is that using a finite number of option prices it is easier to calibrate a one-dimensional object (the Lévy measure) than a three-dimensional object (the intensity function as in [77, Chapter 8]). Therefore, in our approach the influence of the prior on the solution is less important (see Chapter 3).

Unlike the class of all Markov processes and unlike the class of probabilities considered in [77, Chapter 8], the class of all probabilities Q, under which (X, Q) remains a Lévy process is

not convex. We are therefore unable to apply the methods of convex analysis to our formulation of the calibration problem and are limited to standard functional analysis tools. In particular, in our setting we cannot hope to obtain a uniqueness result for the solution of the calibration problem, but, as mentioned in Section 2.2, uniqueness is not always a desirable property from the financial viewpoint.

2.4 Relative entropy of Lévy processes

In this subsection we explicitly compute the relative entropy of two Lévy processes in terms of their characteristic triplets. Under additional assumptions this result was shown in [24] (where it is supposed that Q is equivalent to P and the Lévy process has finite exponential moments under P) and in [77] (where $\log \frac{d\nu^Q}{d\nu^P}$ is supposed bounded from above and below). We consider it important to give an elementary proof valid for all Lévy processes.

Theorem 2.9 (Relative entropy of Lévy processes). *Let $\{X_t\}_{t\geq 0}$ be a real-valued Lévy process on $(\Omega, \mathcal{F}, Q)$ and on $(\Omega, \mathcal{F}, P)$ with respective characteristic triplets (A_Q, ν_Q, γ_Q) and (A_P, ν_P, γ_P). Suppose that the conditions 1–5 of Proposition 1.5 are satisfied and denote $A := A_Q = A_P$. Then for every time horizon $T \leq T_\infty$ the relative entropy of $Q|_{\mathcal{F}_T}$ with respect to $P|_{\mathcal{F}_T}$ can be computed as follows:*

$$I_T(Q|P) = I(Q|_{\mathcal{F}_T}|P|_{\mathcal{F}_T}) = \frac{T}{2A}\left\{\gamma^Q - \gamma^P - \int_{-1}^{1} x(\nu^Q - \nu^P)(dx)\right\}^2 1_{A\neq 0} + T\int_{-\infty}^{\infty}\left(\frac{d\nu^Q}{d\nu^P}\log\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx). \quad (2.19)$$

Proof. Let $\{X_t^c\}_{t\geq 0}$ be the continuous martingale part of X under P (a Brownian motion), μ be the jump measure of X and $\phi := \frac{d\nu^Q}{d\nu^P}$. From [54, Theorem III.5.19], the density process $Z_t := \frac{dQ|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}}$ is the Doléans-Dade exponential of the Lévy process $\{N_t\}_{t\geq 0}$ defined by

$$N_t := \beta X_t^c + \int_{[0,t]\times\mathbb{R}} (\phi(x) - 1)\{\mu(ds \times dx) - ds\, \nu^P(dx)\},$$

where β is given by

$$\beta = \begin{cases} \frac{1}{A}\{\gamma^Q - \gamma^P - \int_{|x|\leq 1} x(\phi(x) - 1)\nu^P(dx)\} & \text{if } A > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Choose $0 < \varepsilon < 1$ and let $I := \{x : \varepsilon \leq \phi(x) \leq \varepsilon^{-1}\}$. We split N_t into two independent martingales:

$$N'_t := \beta X^c_t + \int_{[0,t]\times I} (\phi(x) - 1)\{\mu(ds \times dx) - ds\ \nu^P(dx)\} \quad \text{and}$$
$$N''_t := \int_{[0,t]\times(\mathbb{R}\setminus I)} (\phi(x) - 1)\{\mu(ds \times dx) - ds\ \nu^P(dx)\}.$$

Since N' and N'' never jump together, $[N', N'']_t = 0$ and $\mathcal{E}(N' + N'')_t = \mathcal{E}(N^1)_t\mathcal{E}(N^2)_t$ (cf. Equation II.8.19 in [54]). Moreover, since N' and N'' are Lévy processes and martingales, their stochastic exponentials are also martingales (Proposition 1.4). Therefore,

$$\begin{aligned} I_T(Q|P) &= E^P[Z_T \log Z_T] \\ &= E^P[\mathcal{E}(N')_T\mathcal{E}(N'')_T \log \mathcal{E}(N')_T] + E^P[\mathcal{E}(N')_T\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T] \\ &= E^P[\mathcal{E}(N')_T \log \mathcal{E}(N')_T] + E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T] \end{aligned} \tag{2.20}$$

if these expectations exist.

Since $\Delta N'_t > -1$ a.s., $\mathcal{E}(N')_t$ is almost surely positive. Therefore, from Proposition 1.3, $U_t := \log \mathcal{E}(N')_t$ is a Lévy process with characteristic triplet:

$$\begin{aligned} &A^U = \beta^2 A, \\ &\nu^U(B) = \nu^P(I \cap \{x : \log\phi(x) \in B\})\ \forall B \in \mathcal{B}(\mathbb{R}), \\ &\gamma^U = -\frac{\beta^2 A}{2} - \int_{-\infty}^{\infty} (e^x - 1 - x1_{|x|\leq 1})\nu^U(dx). \end{aligned}$$

This implies that e^{U_t} is a martingale and that U_t has bounded jumps and all exponential moments. Therefore, $E[U_T e^{U_T}] < \infty$ and can be computed as follows:

$$\begin{aligned} E^P[U_T e^{U_T}] &= -i\frac{d}{dz}E^P[e^{izU_T}]|_{z=-i} = -iT\psi'(-i)E^P[e^{U_T}] = -iT\psi'(-i) \\ &= T(A^U + \gamma^U + \int_{-\infty}^{\infty} (xe^x - x1_{|x|\leq 1})\nu^U(dx)) \\ &= \frac{\beta^2 AT}{2} + T\int_I (\phi(x)\log\phi(x) + 1 - \phi(x))\nu^P(dx) \end{aligned} \tag{2.21}$$

It remains to compute $E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T]$. Since N'' is a compound Poisson process, $\mathcal{E}(N'')_t = e^{bt}\prod_{s\leq t}(1 + \Delta N''_s)$, where $b = \int_{\mathbb{R}\setminus I}(1 - \phi(x))\nu^P(dx)$. Let ν'' be the Lévy measure of N'' and λ its jump intensity. Then

$$\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T = bT\mathcal{E}(N'')_T + e^{bT}\prod_{s\leq T}(1 + \Delta N''_s)\sum_{s\leq T}\log(1 + \Delta N''_s)$$

and

$$E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T] = bT + e^{bT} \sum_{k=0}^{\infty} e^{-\lambda T} \frac{(\lambda T)^k}{k!} E[\prod_{s \le T} (1 + \Delta N''_s) \sum_{s \le T} \log(1 + \Delta N''_s) | k \text{ jumps}]$$

Since, under the condition that N'' jumps exactly k times in the interval $[0, T]$, the jump sizes are independent and identically distributed, we find, denoting the generic jump size by $\Delta N''$:

$$\begin{aligned}
&E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T] \\
&\quad = bT + e^{bT} \sum_{k=0}^{\infty} e^{-\lambda T} \frac{(\lambda T)^k}{k!} k E[1 + \Delta N'']^{k-1} E[(1 + \Delta N'') \log(1 + \Delta N'')] \\
&\quad = bT + \lambda T E[(1 + \Delta N'') \log(1 + \Delta N'')] \\
&\quad = bT + T \int_{-\infty}^{\infty} (1 + x) \log(1 + x) \nu''(dx) \\
&\quad = T \int_{\mathbb{R} \setminus I} (\phi(x) \log \phi(x) + 1 - \phi(x)) \nu^P(dx).
\end{aligned}$$

In particular, $E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T]$ is finite if and only if the integral in the last line is finite. Combining the above expression with (2.21) and (2.20) finishes the proof. □

2.4.1 Properties of the relative entropy functional

Lemma 2.10. *Let $P, \{P_n\}_{n \ge 1} \subset \mathcal{L}_B^+$ for some $B > 0$, such that $P_n \Rightarrow P$. Then for every $r > 0$, the level set $L_r := \{Q \in \mathcal{L} : I(Q|P_n) \le r \text{ for some } n\}$ is tight.*

Proof. For any $Q \in L_r$, P_Q denotes any element of $\{P_n\}_{n \ge 1}$, for which $I(Q|P_Q) \le r$. The characteristic triplet of Q is denoted by (A^Q, ν^Q, γ^Q) and that of P_Q by $(A^{P_Q}, \nu^{P_Q}, \gamma^{P_Q})$. In addition, we define $\phi^Q := \frac{d\nu^Q}{d\nu^{P_Q}}$. From Theorem 2.9,

$$\int_{-\infty}^{\infty} (\phi^Q(x) \log \phi^Q(x) + 1 - \phi^Q(x)) \nu^{P_Q}(dx) \le r / T_\infty.$$

Therefore, for u sufficiently large,

$$\int_{\{\phi^Q > u\}} \phi^Q \nu^{P_Q}(dx) \le \int_{\{\phi^Q > u\}} \frac{2\phi^Q [\phi^Q \log \phi^Q + 1 - \phi^Q] \nu^{P_Q}(dx)}{\phi^Q \log \phi^Q} \le \frac{2r}{T_\infty \log u},$$

which entails that for u sufficiently large,

$$\int_{\{\phi^Q > u\}} \nu^Q(dx) \le \frac{2r}{T_\infty \log u}$$

uniformly with respect to $Q \in L_r$. Let $\varepsilon > 0$ and choose u such that $\int_{\{\phi^Q > u\}} \nu^Q(dx) \le \varepsilon/2$ for every $Q \in L_r$. By Proposition 1.7,

$$\int_{-\infty}^{\infty} f(x)\nu^{P_n}(dx) \to \int_{-\infty}^{\infty} f(x)\nu^P(dx)$$

for every continuous bounded function f that is identically zero on a neighborhood of zero. Since the measures ν^P and ν^{P_n} for all $n \ge 1$ are finite outside a neighborhood of zero, we can choose a compact K such that $\nu^{P_n}(\mathbb{R} \setminus K) \le \varepsilon/2u$ for every n. Then

$$\nu^Q(\mathbb{R} \setminus K) = \int_{(\mathbb{R}\setminus K)\cap\{\phi^Q \le u\}} \phi^Q \nu^{P_Q}(dx) + \int_{(\mathbb{R}\setminus K)\cap\{\phi^Q > u\}} \nu^Q(dx) \le \varepsilon,$$

which proves property 1 of Proposition 1.6.

It is easy to check by computing derivatives that for every $u > 0$, on the set $\{x : \phi^Q(x) \le u\}$,

$$(\phi^Q - 1)^2 \le 2u(\phi^Q \log \phi^Q + 1 - \phi^Q).$$

Therefore, for u sufficiently large and for all $Q \in L_r$,

$$\begin{aligned}
&\left|\int_{|x|\le 1} x(\phi^Q - 1)\nu^{P_Q}(dx)\right| \\
&\quad\le \left|\int_{|x|\le 1,\ \phi^Q \le u} x(\phi^Q - 1)\nu^{P_Q}(dx)\right| + \left|\int_{|x|\le 1,\ \phi^Q > u} x(\phi^Q - 1)\nu^{P_Q}(dx)\right| \\
&\quad\le \int_{|x|\le 1} x^2\nu^{P_Q}(dx) + \int_{|x|\le 1,\ \phi^Q \le u} (\phi^Q - 1)^2\nu^{P_Q}(dx) + 2\int_{\phi^Q > u} \phi^Q\nu^{P_Q}(dx) \\
&\quad\le \int_{|x|\le 1} x^2\nu^{P_Q}(dx) + 2u\int_{-\infty}^{\infty} (\phi^Q \log \phi^Q + 1 - \phi^Q)\nu^{P_Q}(dx) + \frac{4r}{T_\infty \log u} \\
&\quad\le \int_{|x|\le 1} x^2\nu^{P_Q}(dx) + \frac{3ru}{T_\infty}. \qquad (2.22)
\end{aligned}$$

By Proposition 1.6, applied to the sequence $\{P_n\}_{n\ge 1}$,

$$A^{P_n} + \int_{|x|\le 1} x^2\nu^{P_n}(dx) \qquad (2.23)$$

is bounded uniformly on n, which implies that the right hand side of (2.22) is bounded uniformly with respect to $Q \in L_r$. From Proposition 1.5, $A^Q = A^{P_Q}$ for all $Q \in L_r$ because for the relative entropy to be finite, necessarily $Q \ll P_Q$. From Theorem 2.9 and Proposition 1.5 it follows that

$$\left\{\gamma^Q - \gamma^P - \int_{-1}^{1} x(\nu^Q - \nu^P)(dx)\right\}^2 \le \frac{2A^{P_Q} r}{T_\infty}.$$

From (2.23), A^{P_n} is bounded uniformly on n. Therefore, inequality (2.22) shows that $|\gamma^Q|$ is bounded uniformly with respect to Q, which proves property 2 of Proposition 1.6.

Once again, for u sufficiently large,

$$A^Q + \int_{-\infty}^{\infty} (x^2 \wedge 1)\phi^Q \nu^{P_Q}(dx) \leq A^Q + u \int_{\phi^Q \leq u} (x^2 \wedge 1)\nu^{P_Q}(dx) + \int_{\phi^Q > u} \phi^Q \nu^{P_Q}(dx) \leq A^{P_Q} + u \int_{-\infty}^{\infty} (x^2 \wedge 1)\nu^{P_Q}(dx) + \frac{2r}{T_\infty \log u}$$

and (2.23) implies that the right hand side is bounded uniformly with respect to $Q \in L_r$. Therefore, property 3 of Proposition 1.6 also holds and the proof is completed. □

Lemma 2.11. *Let Q and P be two probability measures on $(\Omega, \mathcal{F})$. Then*

$$I(Q|P) = \sup_{f \in C_b(\Omega)} \left\{ \int_\Omega f dQ - \int_\Omega (e^f - 1) dP \right\}, \tag{2.24}$$

where $C_b(\Omega)$ is space of bounded continuous functions on Ω.

Proof. Observe that

$$\phi(x) = \begin{cases} x \log x + 1 - x, & x > 0, \\ \infty, & x \leq 0 \end{cases}$$

and $\phi^*(y) = e^y - 1$ are proper convex functions on $\mathbb{R}$, conjugate to each other and apply Corollary 2 to Theorem 4 in [83]. □

Corollary 2.1. *The relative entropy functional $I(Q|P)$ is weakly lower semicontinuous with respect to Q for fixed P.*

Lemma 2.12. *Let $P, \{P_n\}_{n\geq 1} \subset \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ for some $B > 0$ such that $P_n \Rightarrow P$. There exists a sequence $\{Q_n\}_{n\geq 1} \subset \mathcal{M} \cap \mathcal{L}_B^+$ and a constant $C < \infty$ such that $I(Q_n|P_n) \leq C$ for every n.*

Proof. For every $n \geq 1$, by Remark 2.1, Theorem 2.7 can be applied to P_n. Let Q_n be the minimal entropy martingale measure and β_n be the corresponding constant, defined in (2.14). We must show that the minimal relative entropy, given by Equation (2.15), is bounded uniformly on n.

First, let us show that the sequence $\{\beta_n\}_{n\geq 1}$ is bounded. Let h be a continuous bounded truncation function, satisfying $h(x) = x$ in a neighborhood of zero and for any Lévy process Q

with characteristic triplet (A, ν, γ_h) with respect to the truncation function h, define

$$\begin{aligned} f(\beta, Q) &:= \gamma_h + \left(\frac{1}{2} + \beta\right) A + \int_{-\infty}^{\infty} \left\{(e^x - 1)e^{\beta(e^x-1)} - h(x)\right\} \nu(dx) \\ &= \gamma_h + \left(\frac{1}{2} + \beta\right) \left(A + \int_{-\infty}^{\infty} h^2(x)\nu(dx)\right) \\ &\quad + \int_{-\infty}^{\infty} \left\{(e^x - 1)e^{\beta(e^x-1)} - h(x) - \left(\frac{1}{2} + \beta\right) h^2(x)\right\} \nu(dx). \end{aligned} \tag{2.25}$$

Since $(e^x - 1)e^{\beta(e^x-1)} - x - \left(\frac{1}{2} + \beta\right) x^2 = o(|x|^2)$ and the integrand in the last term of (2.25) is bounded on $(-\infty, B]$, by Proposition 1.7, for every β, $\lim_n f(\beta, P_n) = f(\beta, P)$.

The support of ν^{P_n} is bounded from above by B, and the dominated convergence theorem allows to compute the derivative of $f(\beta, P_n)$ by interchanging the derivative and the integral:

$$f'_\beta(\beta, P_n) = A^{P_n} + \int_{-\infty}^{\infty} (e^x - 1)^2 e^{\beta(e^x-1)} \nu^{P_n}(dx) > 0.$$

Therefore, β_n is the *unique* solution of $f(\beta, P_n) = 0$. Let β^* be the solution of $f(\beta, P) = 0$. The support of ν^P is also bounded from above by B and $f'_\beta(\beta^*, P) > 0$. This means that there exist $\varepsilon > 0$ and finite constants $\beta_- < \beta^*$ and $\beta^+ > \beta^*$ such that $f(\beta_-, P) < -\varepsilon$ and $f(\beta_+, P) > \varepsilon$. One can then find N such that for all $n \geq N$, $f(\beta_-, P_n) < -\varepsilon/2$ and $f(\beta_+, P_n) > \varepsilon/2$, which means that $\beta_n \in [\beta_-, \beta_+]$ and the sequence $\{\beta_n\}$ is bounded.

To show that the sequence of relative entropies is bounded, observe that for $|x| \leq 1$,

$$\left|e^{\beta(e^x-1)} - 1 - \beta x\right| \leq \beta e^{\beta(e-1)+1}(1 + \beta e)|x|^2$$

and that for $x \leq B$,

$$\left|e^{\beta(e^x-1)} - 1 - \beta x 1_{|x|\leq 1}\right| \leq \beta e^{\beta(e^B+1)} + 1 + \beta B.$$

The uniform boundedness of the sequence of relative entropies $I(Q_n|P_n)$ now follows from Proposition 1.6 and Equation (2.15). □

2.5 Regularizing the calibration problem

As observed in Section 2.2, problem (2.11) is ill-posed and hard to solve numerically. In particular, its solutions, when they exist, may not be stable with respect to perturbations of market data. If we do not know the prices C_M exactly but only know the perturbed prices C_M^δ that

are within an error δ of C_M, and want to construct an approximation to MELSS(C_M), the solution of problem (2.11) with the true data, it is not a good idea to solve problem (2.11) with the noisy data C_M^δ because MELSS(C_M^δ) may be very far from MELSS(C_M). We therefore need to regularize the problem (2.11), that is, construct a family of continuous "regularization operators" $\{R_\alpha\}_{\alpha>0}$, where α is the parameter which determines the intensity of regularization, such that $R_\alpha(C_M^\delta)$ converges to MELSS of the calibration problem as the noise level δ tends to zero if, for each δ, the regularization parameter α is chosen appropriately. The approximation to MELSS(C_M) using the noisy data C_M^δ is then given by $R_\alpha(C_M^\delta)$ with an appropriate choice of α.

Following classical results on regularization of ill-posed problems (see [40]), we suggest to construct a regularized version of (2.11) by using the relative entropy for penalization rather than for selection, that is, to define

$$J_\alpha(Q) = \|C_M^\delta - C^Q\|_w^2 + \alpha I(Q|P), \tag{2.26}$$

where α is the regularization parameter, and solve the following regularized calibration problem:

Regularized calibration problem Given prices C_M of call options, a prior Lévy process P and a regularization parameter $\alpha > 0$, find $Q^* \in \mathcal{M} \cap \mathcal{L}$, such that

$$J_\alpha(Q^*) = \inf_{Q \in \mathcal{M} \cap \mathcal{L}} J_\alpha(Q). \tag{2.27}$$

Problem (2.27) can be thought of in two ways:

- If the minimum entropy least squares solution with the true data C_M exists, (2.27) allows to construct a stable approximation of this solution using the noisy data.
- If the MELSS with the true data does not exist, either because the set of least squares solutions is empty or because the least squares solutions are incompatible with the prior, the regularized problem (2.27) allows to find a "compromise solution", achieving a trade-off between the pricing constraints and the prior information.

In the rest of this section we study the regularized calibration problem. Under our standing hypothesis that the prior Lévy process has jumps bounded from above and corresponds to

an arbitrage free market ($P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$), we show that the regularized calibration problem always admits a solution that depends continuously on the market data. In addition, we give the condition that the prior P must satisfy in order for the solution to be an *equivalent* martingale measure, and show how the regularization parameter α must be chosen depending on the noise level δ if the regularized solutions are to converge to the solutions of the minimum entropy least squares calibration problem (2.11).

2.5.1 Properties of regularized solutions

We start with an existence theorem showing that if the prior Lévy process has jumps bounded from above and corresponds to an arbitrage-free market, the regularized calibration problem admits a solution.

Theorem 2.13. *Let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ for some $B > 0$. Then the calibration problem* (2.27) *has a solution $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$.*

Proof. By Lemma 2.12, there exists $Q^0 \in \mathcal{M} \cap \mathcal{L}$ with $I(Q^0|P) < \infty$. The solution, if it exists, must belong to the level set $L_{J_\alpha(Q^0)} := \{Q \in \mathcal{L} : I(Q|P) \leq J_\alpha(Q^0)\}$. Since $J_\alpha(Q^0) = \|C_M - C^{Q^0}\|_w^2 + I(Q^0|P) < \infty$, by Lemma 2.10, $L_{J_\alpha(Q^0)}$ is tight and, by Prohorov's theorem, weakly relatively compact. Corollary 2.1 entails that $I(Q|P)$ is weakly lower semicontinuous with respect to Q. Therefore $\{Q \in \mathcal{P}(\Omega) : I(Q|P) \leq J_\alpha(Q^0)\}$ is weakly closed and since by Lemma 2.4, $\mathcal{M} \cap \mathcal{L}_B^+$ is weakly closed, $\mathcal{M} \cap \mathcal{L}_B^+ \cap L_{J_\alpha(Q^0)}$ is weakly compact. Moreover, by Lemma 2.2, the squared pricing error is weakly continuous, which entails that $J_\alpha(Q)$ is weakly lower semicontinuous. Therefore, $J_\alpha(Q)$ achieves its minimum value on $\mathcal{M} \cap \mathcal{L}_B^+ \cap L_{J_\alpha(Q^0)}$, which proves the theorem. □

Since every solution Q^* of the regularized calibration problem (2.27) has finite relative entropy with respect to the prior P, necessarily $Q^* \ll P$. However, Q^* need not in general be equivalent to the prior. When the prior corresponds to the real world (historical) probability, absence of arbitrage is guaranteed if options are priced using an *equivalent* martingale measure. The next theorem gives a sufficient condition for this equivalence.

Theorem 2.14. *Let $P \in \mathcal{L}_{NA} \cap L_B^+$ and suppose that the characteristic function Φ_T^P of P*

satisfies

$$\int_{-\infty}^{\infty} |\Phi_T^P(u)| du < \infty \tag{2.28}$$

for some $T < T_0$, where T_0 is the shortest maturity, present in the market data. Then every solution Q^* of the calibration problem (2.27) *satisfies $Q^* \sim P$.*

Remark 2.2. Condition (2.28) implies that the prior Lévy process has a continuous density at time T and all subsequent times. Two important examples of processes satisfying the condition (2.28) for all T are

- Processes with non-trivial Gaussian component ($A > 0$). This follows directly from the Lévy-Khintchine formula (1.1).
- Processes with stable-like behavior of small jumps, that is, processes whose Lévy measure satisfies

$$\exists \beta \in (0,2), \quad \liminf_{\varepsilon \downarrow 0} \varepsilon^{-\beta} \int_{-\varepsilon}^{\varepsilon} |x|^2 \nu(dx) > 0. \tag{2.29}$$

  For proof, see Proposition 28.3 in [87]. This class includes tempered stable processes (1.20) with $\alpha_+ > 0$ and/or $\alpha_- > 0$.

Theorem 2.14 will be proven after the following technical lemma.

Lemma 2.15. *Let $P \in \mathcal{M} \cap \mathcal{L}_B^+$ with characteristic triplet (A, ν, γ) and characteristic exponent ψ. There exists $C < \infty$ such that*

$$\left| \frac{\psi(v-i)}{(v-i)v} \right| \leq C \quad \forall v \in \mathbb{R}.$$

Proof. From (1.1) and (1.2),

$$\psi(v-i) = -\frac{1}{2} A v (v-i) + \int_{-\infty}^{\infty} (e^{i(v-i)x} + iv - e^x - ive^x) \nu(dx). \tag{2.30}$$

Observe first that

$$e^{i(v-i)x} + iv - e^x - ive^x = iv(xe^x + 1 - e^x) + \frac{\theta v^2 x^2 e^x}{2} \quad \text{for some } \theta \text{ with } |\theta| \leq 1.$$

Therefore, for all v with $|v| \geq 2$,

$$\left| \frac{e^{i(v-i)x} + iv - e^x - ive^x}{(v-i)v} \right| \leq xe^x + 1 - e^x + x^2 e^x. \tag{2.31}$$

On the other hand,

$$\frac{e^{i(v-i)x}+iv-e^x-ive^x}{(v-i)v} = \frac{ie^x(e^{ivx}-1)}{v} - \frac{i(e^{i(v-i)x}-1)}{v-i}$$
$$= -xe^x - \frac{ivx^2}{2}e^{\theta_1 ivx} + x + \frac{i(v-i)x^2}{2}e^{\theta_2 i(v-i)x}$$

with some $\theta_1, \theta_2 \in [0,1]$. Therefore, for all v with $|v| \le 2$,

$$\left|\frac{e^{i(v-i)x}+iv-e^x-ive^x}{(v-i)v}\right| \le x(1-e^x) + \frac{x^2}{2}(v+\sqrt{1+v^2}e^x) \le x(1-e^x) + x^2(1+2e^x). \quad (2.32)$$

Since the support of ν is bounded from above, the right-hand sides of (2.31) and (2.32) are ν-integrable, and the proof of the lemma is completed. □

Proof of Theorem 2.14. Let Q^* be a solution of (2.27) with prior P. By Theorem 2.7, there exists $Q^0 \in \mathcal{M} \cap \mathcal{L}$ such that $Q^0 \sim P$. Denote the characteristic triplet of Q^* by (A, ν^*, γ^*) and that of Q^0 by (A, ν^0, γ^0).

Let Q_x be a Lévy process with characteristic triplet

$$(A, x\nu^0 + (1-x)\nu^*, x\gamma^0 + (1-x)\gamma^*).$$

From the linearity of the martingale condition (1.2), it follows that for all $x \in [0,1]$, $Q_x \in \mathcal{M} \cap \mathcal{L}$. Since Q^* realizes the minimum of $J_\alpha(Q)$, necessarily $J_\alpha(Q_x) - J_\alpha(Q^*) \ge 0$ for all $x \in [0,1]$. Our strategy for proving the theorem is first to show that $\frac{\|C_M - C^{Q_x}\|^2 - \|C_M - C^{Q^*}\|^2}{x}$ is bounded as $x \to 0$ and then to show that if $\frac{I(Q_x|P) - I(Q^*|P)}{x}$ is bounded from below as $x \to 0$, necessarily $Q^* \sim P$.

The first step is to prove that the characteristic function Φ^* of Q^* satisfies the condition (2.28) for some $T < T_0$. If $A > 0$, this is trivial; suppose therefore that $A = 0$. In this case, $|\Phi^*_T(u)| = \exp(T\int_{-\infty}^{\infty}(\cos(ux)-1)\nu^*(dx))$. Denote $\frac{d\nu^*}{d\nu^P} := \phi^*$. Since $Q^* \ll P$, by Theorem 1.5, $\int_{-\infty}^{\infty}(\sqrt{\phi^*(x)}-1)^2\nu^P(dx) \le K < \infty$ for some constant K. Therefore, there exists another constant $C > 0$ such that

$$\int_{\{\phi^*(x)>C\}} (1-\cos(ux))|\phi^* - 1|\nu^P(dx) < C$$

uniformly on u. For all $r > 0$,

$$\begin{aligned}\int_{-\infty}^{\infty}(1-\cos(ux))|\phi^*-1|\nu^P(dx) &\le C + \int_{\{\phi^*(x)\le C\}}(1-\cos(ux))|\phi^*-1|\nu^P(dx)\\ &\le C + \frac{r}{2}\int_{\{\phi^*(x)\le C\}}(1-\cos(ux))^2\nu^P(dx) + \frac{1}{2r}\int_{\{\phi^*(x)\le C\}}(\phi^*-1)^2\nu^P(dx)\\ &\le C + r\int_{-\infty}^{\infty}(1-\cos(ux))\nu^P(dx) + \frac{K(\sqrt{C}+1)^2}{2r}.\end{aligned}$$

This implies that

$$\int_{-\infty}^{\infty}(\cos(ux)-1)\nu^*(dx) \le (1+r)\int_{-\infty}^{\infty}(\cos(ux)-1)\nu^P(dx) + C + \frac{K(\sqrt{C}+1)^2}{2r}$$

for all $r > 0$. Therefore, if the characteristic function of P satisfies the condition (2.28) for some T, the characteristic function of Q^* will satisfy it for every $T' > T$.

Since $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$, $Q_x \in \mathcal{M} \cap \mathcal{L}_B^+$ for all $x \in [0,1]$. Therefore, condition (1.23) is satisfied and option prices can be computed by inverting the Fourier transform (1.24):

$$C^{Q_x}(T,K) = (1-Ke^{-rT})^+ + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iv\log K + ivrT}\frac{\exp(T(1-x)\psi^*(v-i) + Tx\psi^0(v-i)) - 1}{iv(1+iv)}dv,$$

where ψ^0 and ψ^* denote the characteristic exponents of Q_0 and Q^*. It follows that

$$\frac{C^{Q_x}(T,K) - C^{Q^*}(T,K)}{x} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iv\log K + ivrT}\frac{e^{T(1-x)\psi^*(v-i)+Tx\psi^0(v-i)} - e^{T\psi^*(v-i)}}{iv(1+iv)x}dv$$

The fact that $\Re\psi^0(v-i) \le 0$ and $\Re\psi^*(v-i) \le 0$ for all $v \in \mathbb{R}$ together with Lemma 2.15 implies that

$$\begin{aligned}&\left|e^{-iv\log K + ivrT}\frac{e^{T(1-x)\psi^*(v-i)+Tx\psi^0(v-i)} - e^{T\psi^*(v-i)}}{iv(1+iv)x}\right|\\ &\qquad\le T\frac{|e^{T(1-x)\psi^*(v-i)}||\psi^0(v-i)-\psi^*(v-i)|}{|v(1+iv)|} \le T|e^{T(1-x)\psi^*(v-i)}|C'\end{aligned}$$

for some constant C'. From the dominated convergence theorem and since Q^* satisfies (2.28), $\frac{\partial C^{Q_x}(T,K)}{\partial x}|_{x=0}$ exists and is bounded uniformly on T and K in the market data set. This in turn means that $\frac{\|C_M - C^{Q_x}\|^2 - \|C_M - C^{Q^*}\|^2}{x}$ is bounded as $x \to 0$.

To complete the proof, it remains to show that if $\frac{I(Q_x|P)-I(Q^*|P)}{x}$ is bounded from below as $x \to 0$, necessarily $Q^* \sim P$. Using the convexity (with respect to ν^Q and γ^Q) of the two terms

in the expression (2.19) for relative entropy, we have:

$$
\begin{aligned}
&\frac{I(Q_x|P) - I(Q^*|P)}{x} \\
&= \frac{T_\infty}{2Ax}\left\{x\gamma^0 + (1-x)\gamma^* - \gamma^P - \int_{|z|\le 1} z(x\nu^0 + (1-x)\nu^* - \nu^P)\nu^P(dz)\right\}^2 1_{A\neq 0} \\
&\qquad - \frac{T_\infty}{2Ax}\left\{\gamma^* - \gamma^P - \int_{|z|\le 1}(\nu^* - \nu^P)\nu^P(dz)\right\}^2 1_{A\neq 0} \\
&\qquad + \frac{T_\infty}{x}\int_{-\infty}^{\infty}\{(x\phi^0 + (1-x)\phi^*)\log(x\phi^0 + (1-x)\phi^*) - x\phi^0 - (1-x)\phi^* + 1\}\nu^P(dz) \\
&\qquad - \frac{T_\infty}{x}\int_{-\infty}^{\infty}\{\phi^*\log(\phi^*) - \phi^* + 1\}\nu^P(dz) \\
&\le \frac{T_\infty}{2A}\left\{\gamma^0 - \gamma^P - \int_{|z|\le 1}(\nu^0 - \nu^P)\nu^P(dz)\right\}^2 1_{A\neq 0} \\
&\qquad - \frac{T_\infty}{2A}\left\{\gamma^* - \gamma^P - \int_{|z|\le 1}(\nu^* - \nu^P)\nu^P(dz)\right\}^2 1_{A\neq 0} \\
&\qquad + T_\infty\int_{\{\phi^*>0\}}\{\phi^0\log(\phi^0) - \phi^0 + 1\}\nu^P(dz) - T_\infty\int_{\{\phi^*>0\}}\{\phi^*\log(\phi^*) - \phi^* + 1\}\nu^P(dz) \\
&\qquad + T_\infty\int_{\{\phi^*=0\}}\{\phi^0\log(x\phi^0) - \phi^0\}\nu^P(dz) \le I(Q_0|P) + T_\infty\int_{\{\phi^*=0\}}(\phi_0\log x - 1)\nu^P(dx)
\end{aligned}
$$

Since $J_\alpha(Q_x) - J_\alpha(Q^*) \ge 0$, this expression must be bounded from below. Therefore, $\nu^P(\{\phi^* = 0\}) = 0$, and Proposition 1.5 entails that $P \ll Q^*$. □

Continuity of solutions with respect to data

Theorem 2.16. *Let $\{C_M^n\}_{n\ge 1}$ and C_M be data sets of option prices such that*

$$\lim_n \|C_M^n - C_M\|_w \to 0.$$

Let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$, $\alpha > 0$, and for each n, let Q_n be a solution of the calibration problem (2.27) *with data C_M^n, prior Lévy process P and regularization parameter α. Then $\{Q_n\}_{n\ge 1}$ has a subsequence, converging weakly to a process $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$, and the limit of every converging subsequence of $\{Q_n\}_{n\ge 1}$ is a solution of calibration problem* (2.27) *with data C_M, prior P and regularization parameter α.*

Proof. By Lemma 2.12, there exists $Q^0 \in \mathcal{M} \cap \mathcal{L}$ with $I(Q^0|P) < \infty$. Since, by Lemma 2.2, $\|C^{Q^0} - C_M^n\|^2 \le S_0^2$ for all n, $\alpha I(Q_n|P) \le S_0^2 + \alpha I(Q^0|P)$ for all n. Therefore, by Lemmas 2.4

and 2.10 and Prohorov's theorem, $\{Q_n\}_{n\geq 1}$ is weakly relatively compact, which proves the first part of the theorem.

Choose any subsequence of $\{Q_n\}_{n\geq 1}$, converging weakly to a process $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$. , To simplify notation, this subsequence is denoted again by $\{Q_n\}_{n\geq 1}$. The triangle inequality and Lemma 2.2 imply that

$$\|C^{Q_n} - C_M^n\|^2 \xrightarrow[n\to\infty]{} \|C^{Q^*} - C_M\|^2 \tag{2.33}$$

Since, by Lemma 2.11, the relative entropy functional is weakly lower semicontinuous in Q, for every $Q \in \mathcal{M} \cap \mathcal{L}_B^+$,

$$\begin{aligned}
\|C^{Q^*} - C_M\| + \alpha I(Q|P) &\leq \liminf_n \{\|C^{Q_n} - C_M^n\|^2 + \alpha I(Q_n|P)\} \\
&\leq \liminf_n \{\|C^{Q} - C_M^n\|^2 + \alpha I(Q|P)\} \\
&= \lim_n \|C^{Q} - C_M^n\|^2 + \alpha I(Q|P) \\
&= \|C^{Q} - C_M\|^2 + \alpha I(Q|P),
\end{aligned}$$

where the second inequality follows from the fact that Q_m is the solution of the calibration problem with data C_M^m and the last line follows from the triangle inequality. □

2.5.2 Convergence of regularized solutions

In this section we study the convergence of solutions of the regularized calibration problem (2.27) to the solutions of the minimum entropy least squares calibration problem (2.11) when the noise level in the data tends to zero.

Theorem 2.17. *Let $\{C_M^\delta\}$ be a family of data sets of option prices such that $\|C_M - C_M^\delta\| \leq \delta$, let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ and suppose that there exist a solution Q of problem (2.4) with data C_M (a least squares solution) such that $I(Q|P) < \infty$.*

If $\|C^Q - C_M\| = 0$ (the constraints are reproduced exactly), let $\alpha(\delta)$ be such that $\alpha(\delta) \to 0$ and $\frac{\delta^2}{\alpha(\delta)} \to 0$ as $\delta \to 0$. Otherwise, let $\alpha(\delta)$ be such that $\alpha(\delta) \to 0$ and $\frac{\delta}{\alpha(\delta)} \to 0$ as $\delta \to 0$.

Then every sequence $\{Q^{\delta_k}\}$, where $\delta_k \to 0$ and Q^{δ_k} is a solution of problem (2.27) with data $C_M^{\delta_k}$, prior P and regularization parameter $\alpha(\delta_k)$, has a weakly convergent subsequence. The limit of every convergent subsequence is a solution of problem (2.11) (MELSS) with data C_M and prior P. If such a MELSS Q^+ is unique then $\lim_{\delta\to 0} Q^\delta = Q^+$.

Proof. By Lemma 2.6, there exists at least one MELSS with data C_M and prior P, that has finite relative entropy with respect to the prior. Let Q^+ be any such MELSS. Since Q^{δ_k} is the solution of the regularized problem, for every k,

$$\|C^{Q^{\delta_k}} - C_M^{\delta_k}\|^2 + \alpha(\delta_k) I(Q^{\delta_k}|P) \leq \|C^{Q^+} - C_M^{\delta_k}\|^2 + \alpha(\delta_k) I(Q^+|P).$$

Using the fact that for every $r > 0$ and for every $x, y \in \mathbb{R}$,

$$(1-r)x^2 + (1-1/r)y^2 \leq (x+y)^2 \leq (1+r)x^2 + (1+1/r)y^2,$$

we obtain that

$$(1-r)\|C^{Q^{\delta_k}} - C_M\|^2 + \alpha(\delta_k) I(Q^{\delta_k}|P) \\ \leq (1+r)\|C^{Q^+} - C_M\|^2 + \frac{2\delta_k^2}{r} + \alpha(\delta_k) I(Q^+|P), \quad (2.34)$$

and since Q^+ is a least squares solution with data C_M, this implies for all $r \in (0,1)$ that

$$\alpha(\delta_k) I(Q^{\delta_k}|P) \leq 2r\|C^{Q^+} - C_M\|^2 + \frac{2\delta_k^2}{r} + \alpha(\delta_k) I(Q^+|P). \quad (2.35)$$

If the constraints are reproduced exactly, then $\|C^{Q^+} - C_M\| = 0$ and with the choice $r = 1/2$, the above expression yields:

$$I(Q^{\delta_k}|P) \leq \frac{4\delta_k^2}{\alpha(\delta_k)} + I(Q^+|P).$$

Since, by the theorem's statement, in the case of exact constraints $\frac{\delta_k^2}{\alpha(\delta_k)} \to 0$, this implies that

$$\limsup_k \{I(Q^{\delta_k}|P)\} \leq I(Q^+|P). \quad (2.36)$$

If $\|C^{Q^+} - C_M\| > 0$ (misspecified model) then the right-hand side of (2.35) achieves its maximum when $r = \delta_k \|C^{Q^+} - C_M\|^{-1}$, in which case we obtain

$$I(Q^{\delta_k}|P) \leq \frac{4\delta_k}{\alpha(\delta_k)}\|C^{Q^+} - C_M\| + I(Q^+|P).$$

Since in the case of approximate constraints, $\frac{\delta_k}{\alpha(\delta_k)} \to 0$, we obtain (2.36) once again.

Inequality (2.36) implies in particular that $I(Q^{\delta_k}|P)$ is uniformly bounded, which proves, by Lemmas 2.10 and 2.4, that $\{Q^{\delta_k}\}$ is relatively weakly compact in $\mathcal{M} \cap \mathcal{L}_B^+$.

Choose a subsequence of $\{Q^{\delta_k}\}$, converging weakly to $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$. To simplify notation, this subsequence is denoted again by $\{Q^{\delta_k}\}_{k\geq 1}$. Substituting $r = \delta$ into Equation (2.34) and making k tend to infinity shows that

$$\limsup_k \|C^{Q^{\delta_k}} - C_M\|^2 \leq \|C^{Q^+} - C_M\|^2.$$

Together with Lemma 2.2 this implies that

$$\|C^{Q^*} - C_M\|^2 \leq \|C^{Q^+} - C_M\|^2,$$

hence Q^* is a least squares solution. By weak lower semicontinuity of I (cf. Lemma 2.11) and using (2.36),

$$I(Q^*|P) \leq \liminf_k I(Q^{\delta_k}|P) \leq \limsup_k I(Q^{\delta_k}|P) \leq I(Q^+|P),$$

which means that Q^* is a MELSS. The last assertion of the theorem follows from the fact that in this case every subsequence of $\{Q^{\delta_k}\}$ has a further subsequence converging toward Q^+. □

Chapter 3

Numerical implementation of the calibration algorithm

Before solving the calibration problem (2.27) numerically, we reformulate it as follows:

- The calibration problem (2.27) is expressed in terms of the characteristic triplets (A, ν^Q, γ^Q) and (A, ν^P, γ^P) of the prior Lévy process and the solution. This can be done using Equations (1.1), (1.25) and (1.26) (option pricing by Fourier transform) and Equation (2.19) (expression of the relative entropy in terms of characteristic triplets).

- The Lévy measure ν^Q is discretized (approximated by a finite-dimensional object). This discussed in detail in Section 3.1.

The prior Lévy process P is a crucial ingredient that must be specified by the user. Section 3.2 suggests different ways to do this and studies the effect of a misspecification of the prior on the solutions of the calibration problem. Section 3.3 discusses the methods to choose the regularization parameter α based on the data and Section 3.4 treats the choice of weights w_i of different options. Section 3.5 details the numerical algorithm that we use to solve the calibration problem once the prior, the regularization parameter and the weights have been fixed.

The calibration algorithm, including the automatic choice of the regularization parameter, has been implemented in a computer program `levycalibration` and various tests have been carried out, both on simulated option prices (computed in a known exponential Lévy model) and using real market data. Section 3.6 discusses the results of these tests.

3.1 Discretizing the calibration problem

A convenient way to discretize the calibration problem is to take a prior Lévy process P with Lévy measure supported by a finite number of points:

$$\nu^P = \sum_{k=0}^{M-1} p_k \delta_{\{x_k\}}(dx). \tag{3.1}$$

In this case, by Proposition 1.5, the Lévy measure of the solution necessarily satisfies $\nu^Q \ll \nu^P$, therefore

$$\nu^Q = \sum_{k=0}^{M-1} q_k \delta_{\{x_k\}}(dx), \tag{3.2}$$

that is, the solution belongs to a finite-dimensional space and can be computed using a numerical optimization algorithm. The advantage of this discretization approach is that we are solving the same problem (2.27), only with a different prior measure, so all results of Section 2.5 (existence of solution, continuity etc.) hold in the finite-dimensional case.

Taking Lévy measures of the form (3.1) we implicitly restrict the class of possible solutions to Lévy processes with bounded jumps and finite jump intensity. However, in this section we will see that this restriction is not as important as it seems: the solution of a calibration problem with any prior can be approximated (in the weak sense) by a sequence of solutions of calibration problems with priors having Lévy measures of the form (3.1). Moreover, in Section 3.6.1 we will observe empirically that smiles produced by infinite intensity models can be calibrated with arbitrary precision by such jump-diffusion models.

We start with a lemma showing that every Lévy process can be approximated by Lévy processes with atomic Lévy measures.

Lemma 3.1. *Let P be a Lévy process with characteristic triplet (A, ν, γ) with respect to a continuous bounded truncation function h, satisfying $h(x) = x$ in a neighborhood of 0, and for every n, let P_n be a Lévy process with characteristic triplet (A, ν_n, γ) (with respect to the same truncation function) where*

$$\nu_n := \sum_{k=1}^{2n} \delta_{\{x_k\}}(dx) \frac{\mu([x_k - 1/\sqrt{n}, x_k + 1/\sqrt{n}))}{1 \wedge x_k^2},$$

$x_k := (2(k-n)-1)/\sqrt{n}$ and μ is a finite measure on $\mathbb{R}$, defined by $\mu(B) := \int_B (1 \wedge x^2)\nu(dx)$ for all $B \in \mathcal{B}(\mathbb{R})$. Then $P_n \Rightarrow P$.

Proof. For a function $f \in C_b(\mathbb{R})$, define

$$f_n(x) := \begin{cases} 0, & x \geq 2\sqrt{n}, \\ 0, & x < -2\sqrt{n}, \\ f(x_i), & x \in [x_i - 1/\sqrt{n}, x_i + 1/\sqrt{n}) \quad \text{with } 1 \leq i \leq 2n, \end{cases}$$

Then clearly

$$\int (1 \wedge x^2) f(x) \nu_n(dx) = \int f_n(x) \mu(dx).$$

Since $f(x)$ is continuous, $f_n(x) \to f(x)$ for all x and since f is bounded, the dominated convergence theorem implies that

$$\lim_n \int (1 \wedge x^2) f(x) \nu_n(dx) = \lim_n \int f_n(x) \mu(dx) = \int f(x) \mu(dx) = \int (1 \wedge x^2) f(x) \nu(dx). \quad (3.3)$$

With $f(x) \equiv \frac{h^2(x)}{1 \wedge x^2}$ the above yields:

$$\lim_n \int h^2(x) \nu_n(dx) = \int h^2(x) \nu(dx).$$

On the other hand, for every $g \in C_b(\mathbb{R})$ such that $g(x) \equiv 0$ on a neighborhood of 0, $f(x) := \frac{g(x)}{1 \wedge x^2}$ belongs to $C_b(\mathbb{R})$. Therefore, from Equation (3.3), $\lim_n \int g(x) \nu_n(dx) = \int g(x) \nu(dx)$, and by Proposition (1.7), $P_n \Rightarrow P$. □

Theorem 3.2. *Let $P, \{P_n\}_{n \geq 1} \subset \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ such that $P_n \Rightarrow P$, let $\alpha > 0$, let C_M be a data set of option prices and for each n let Q_n be a solution of the calibration problem* (2.27) *with prior P_n, regularization parameter α and data C_M. Then the sequence $\{Q_n\}_{n \geq 1}$ has a weakly convergent subsequence and the limit of every weakly convergent subsequence of $\{Q_n\}_{n \geq 1}$ is a solution of the calibration problem* (2.27) *with prior P.*

Proof. By Lemma 2.12, there exists $C < \infty$ such that for every n, one can find $\tilde{Q}_n \in \mathcal{M} \cap \mathcal{L}$ with $I(\tilde{Q}_n | P_n) \leq C$. Since, by Lemma 2.2, $\|C^{\tilde{Q}_n} - C_M\|_w^2 \leq S_0^2$ for every n and Q_n is the solution of the calibration problem, $I(Q_n | P_n) \leq S_0^2/\alpha + C < \infty$ for every n. Therefore, by Lemma 2.10, $\{Q_n\}$ is tight and, by Prohorov's theorem and Lemma 2.4, weakly relatively compact in $\mathcal{M} \cap \mathcal{L}_B^+$. Choose a subsequence of $\{Q_n\}$, converging weakly to $Q \in \mathcal{M} \cap \mathcal{L}_B^+$. To simplify notation, this subsequence is also denoted by $\{Q_n\}_{n \geq 1}$. It remains to show that Q is indeed a solution of the calibration problem with prior P.

Lemma 2.11 entails that

$$I(Q,P) \leq \liminf_n I(Q_n, P_n), \tag{3.4}$$

and since, by Lemma 2.2, the pricing error is weakly continuous, we also have

$$\|C^Q - C_M\|_w^2 + \alpha I(Q,P) \leq \liminf_n \{\|C^{Q_n} - C_M\|_w^2 + \alpha I(Q_n, P_n)\}. \tag{3.5}$$

Let $\phi \in C_b(\Omega)$ with $\phi \geq 0$ and $E^P[\phi] = 1$. Without loss of generality we can suppose that for every n, $E^{P_n}[\phi] > 0$ and therefore Q'_n, defined by $Q'_n(B) := \frac{E^{P_n}[\phi 1_B]}{E^{P_n}[\phi]}$, is a probability measure on Ω. Clearly, Q'_n converges weakly to Q' defined by $Q'(B) := E^P[\phi 1_B]$. Therefore, by Lemma 2.2,

$$\lim_n \|C^{Q'_n} - C_M\|_w^2 = \|C^{Q'} - C_M\|_w^2. \tag{3.6}$$

Moreover,

$$\begin{aligned}\lim_n I(Q'_n|P_n) &= \lim_n \int_\Omega \frac{\phi}{E^{P_n}[\phi]} \log \frac{\phi}{E^{P_n}[\phi]} dP_n \\ &= \lim_n \frac{1}{E^{P_n}[\phi]} \int_\Omega \phi \log \phi dP_n - \lim_n \log \int_\Omega \phi dP_n = \int_\Omega \phi \log \phi dP. \end{aligned} \tag{3.7}$$

For the rest of this proof, for every $\phi \in L^1(P)$ with $\phi \geq 0$ and $E^P[\phi] = 1$ let Q_ϕ denote the probability measure on Ω, defined by $Q_\phi(B) := E^P[\phi 1_B]$ for every $B \in \mathcal{F}$. Using (3.5–3.7) and the optimality of Q_n, we obtain that for every $\phi \in C_b(\Omega)$ with $\phi \geq 0$ and $E^P[\phi] = 1$,

$$\|C^Q - C_M\|_w^2 + I(Q,P) \leq \|C^{Q_\phi} - C_M\|_w^2 + I(Q_\phi|P) \tag{3.8}$$

To complete the proof of the theorem, we must generalize this inequality to all $\phi \in L^1(P)$ with $\phi \geq 0$ and $E^P[\phi] = 1$.

First, let $\phi \in L^1(P) \cap L^\infty(P)$ with $\phi \geq 0$ and $E^P[\phi] = 1$. Then there exists a sequence $\{\phi_n\} \subset C_b(\Omega)$ such that $\phi_n \to \phi$ in $L^1(P)$, $\phi_n \geq 0$ for all n and ϕ_n are bounded in L^∞ norm uniformly on n. Moreover, $\phi'_n := \phi_n / E^P[\phi_n]$ also belongs to $L^1(P)$, is positive and $\phi'_n \xrightarrow{L^1(P)} \phi$ because by the triangle inequality,

$$\|\phi'_n - \phi\|_{L^1} \leq \frac{1}{E^P[\phi_n]} \left(\|\phi_n - \phi\|_{L^1} + \|\phi - \phi E^P[\phi_n]\|_{L^1}\right) \xrightarrow[n\to\infty]{} 0.$$

In addition, it is easy to see that $Q_{\phi'_n} \Rightarrow Q_\phi$. Therefore,

$$\lim_n \|C^{Q_{\phi'_n}} - C_M\|_w^2 = \|C^{Q_\phi} - C_M\|_w^2$$

Since ϕ'_n are bounded in L^∞ norm uniformly on n, $\phi'_n \log \phi'_n$ is also bounded and the dominated convergence theorem implies that $\lim_n I(Q_{\phi'_n}|P) = I(Q_\phi|P)$. Passing to the limit in (3.8), we obtain that this inequality holds for every $\phi \in L^1(P) \cap L^\infty(P)$ with $\phi \geq 0$ and $E^P[\phi] = 1$.

Let us now choose a nonnegative $\phi \in L^1(P)$ with $E^P[\phi] = 1$. If $I(Q_\phi|P) = \infty$ then surely (3.8) holds, therefore we can suppose $I(Q_\phi|P) < \infty$. Let $\phi_n = \phi \wedge n$. Then $\phi_n \to \phi$ in $L^1(P)$ because

$$\|\phi_n - \phi\|_{L^1} \leq \int_{\phi \geq n} \phi dP = \int_{\phi \geq n} \frac{\phi \log \phi}{\log \phi} dP \leq \frac{I(Q_\phi|P)}{\log n} \to 0.$$

Denoting $\phi'_n := \phi_n / E^P[\phi_n]$ as above, we obtain that

$$\lim_n \|C^{Q_{\phi'_n}} - C_M\|_w^2 = \|C^{Q_\phi} - C_M\|_w^2$$

Since, for a sufficiently large n, $|\phi_n(x) \log \phi_n(x)| \leq |\phi(x) \log \phi(x)|$, we can once again apply the dominated convergence theorem:

$$\lim_n \int \phi'_n \log \phi'_n dP = \frac{1}{\lim_n E^P[\phi_n]} \lim_n \int \phi_n \log \phi_n dP - \lim_n \log E^P[\phi_n] = \int \phi \log \phi dP$$

Therefore, by passage to the limit, (3.8) holds for all $\phi \in L^1(P)$ with $\phi \geq 0$ and $E^P[\phi] = 1$, which completes the proof of the theorem. □

To approximate numerically the solution of the calibration problem (2.27) with a given prior P, we need to construct, using Lemma 3.1, an approximating sequence $\{P_n\}$ of Lévy processes with atomic measures such that $P_n \Rightarrow P$. The sequence $\{Q_n\}$ of solutions corresponding to this sequence of priors will converge (in the sense of Theorem 3.2) to a solution of the calibration problem with prior P.

3.2 Choice of the prior Lévy process

The prior Lévy process must, generally speaking, reflect the user's view of the model. It is one of the most important ingredients of the calibration method and cannot be determined completely automatically because the choice of the prior has a strong influence on the outcome of the calibration. The user should therefore specify a characteristic triplet (A^P, ν^P, γ^P) of the prior P. A natural solution, justified by the economic considerations of Section 2.3 is to take the historical probability, resulting from statistical estimation of an exponential Lévy

model from the time series of asset returns. Since, under the conditions of Theorem 2.14, the calibration procedure yields a martingale probability equivalent to the prior, the choice of historical probability as prior ensures the absence of arbitrage opportunity involving stock and options in the calibrated model. Estimation of exponential Lévy models is not discussed here and interested reader is referred to [27, Chapter 7]. Specific exponential Lévy models are discussed in [57] (Merton model), [79] (generalized hyperbolic model) and [21] (CGMY or tempered stable model). [9] and [12] treat more general jump-diffusion type models.

To ensure the stability of calibrated Lévy measures over time, one can systematically choose the calibrated exponential Lévy model of the previous day as today's prior. This choice guarantees that the current calibrated Lévy measure will be altered in the least possible way, to accommodate the arrival of new option pricing information.

Alternatively, the prior can simply correspond to a model that seems "reasonable" to the user (e.g. an analyst). In this case our calibration method may be seen as a way to make the smallest possible change in this model to take into account the observed option prices.

When the user does not have such detailed views, it should be possible to generate a reference measure P automatically from options data. To do this we choose an auxiliary model $Q(\theta)$ (e.g. Merton model) described by a finite-dimensional parameter vector θ and calibrate it to data using the least squares procedure:

$$\theta^* = \arg\min_{\theta} \|C^{Q(\theta)} - C_M\|_w^2 \tag{3.9}$$

It is generally not a good idea to recalibrate this parametric model every day, because in this case the prior will completely lose its stabilizing role. On the contrary, one should try to find typical parameter values for a particular market (e.g. averages over a long period) and fix them once and for all. Since the objective function in (3.9) is usually not convex, a simple gradient descent procedure may not give the global minimum. However, the solution $Q(\theta^*)$ will be corrected at later stages and should only be viewed as a way to regularize the optimization problem (2.4) so the minimization procedure at this stage need not be precise.

Theorem 3.2 shows that the calibrated solutions are continuous with respect to the prior, that is, small changes in the prior process induce small changes in the solutions. To assess empirically the influence of finite changes in the prior on the result of calibration, we have carried out two series of numerical tests. In the first series of tests, Lévy measure was calibrated

twice to the same set of option prices using prior models that were different but close to each other (see Section 3.5.3 for the description of the calibration algorithm). Namely, in the test A we used Merton model with diffusion volatility $\sigma = 0.2$, zero mean jump size, jump standard deviation of 0.1 and intensity $\lambda = 3$, whereas in the test B all the parameters except intensity had the same values and the intensity was equal to 2. The result of the test is shown in Figure 3.1. The solid curves correspond to calibrated measures and the dotted ones depict the prior measures. Notice that there is very little difference between the calibrated measures, which means that, in harmony with Theorem 3.2, the result of calibration is robust to minor variations of the parameters of prior measure, as long as its qualitative shape remains the same.

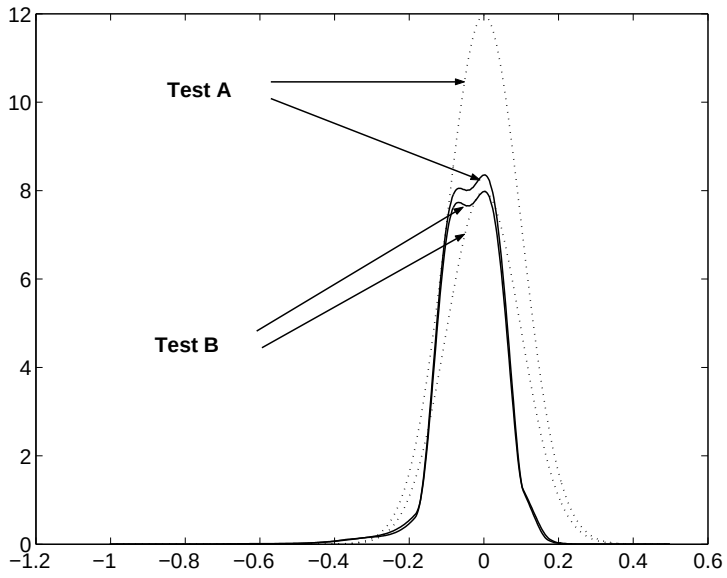

Figure 3.1: Lévy measures calibrated to the same data set using two prior measures similar to each other. Solid curves correspond to calibrated measures and dotted ones depict the priors.

In the second series of tests we have again calibrated the Lévy measure twice to the same set of option prices, this time taking two radically different priors. Namely, in test A we used Merton model with diffusion volatility $\sigma = 0.2$, zero mean jump size, jump standard deviation of 0.1 and intensity $\lambda = 2$, whereas in test B we took a uniform Lévy measure on the interval $[-1, 0.5]$ with intensity $\lambda = 2$. The calibrated measures (solid lines in Figure 3.2) are still similar but exhibit much more differences than in the first series of tests. Not only they are different

in the middle, but also the behavior of tails of the calibrated Lévy measure with uniform prior is more erratic than in the case where Merton model was used as prior (see Figure 3.2, right graph).

Comparison of Figures 3.1 and 3.2 shows that the exact values of parameters of the prior model are not very important, but it is crucial to find the right shape of the prior.

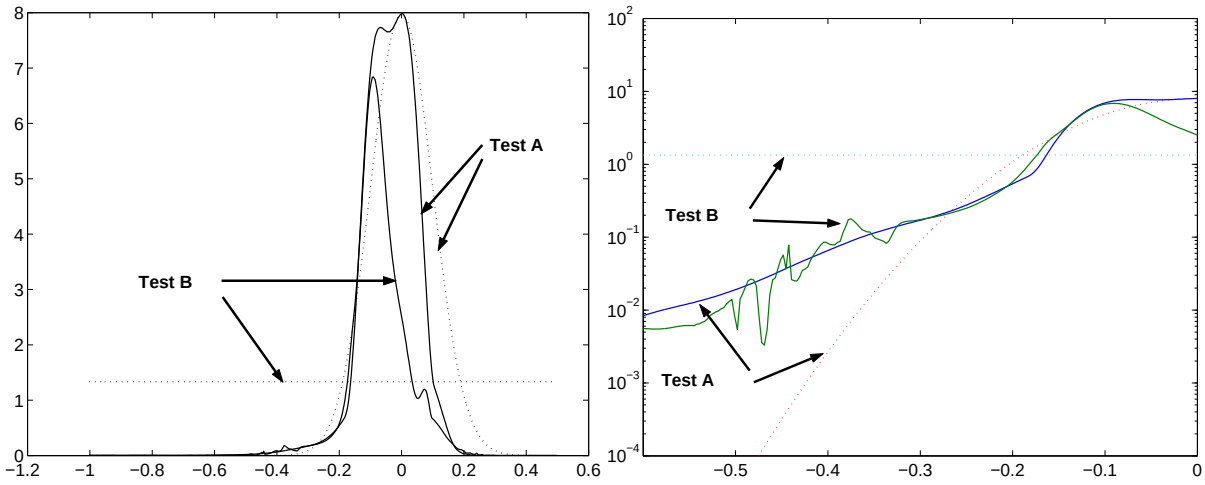

Figure 3.2: Lévy measures calibrated to the same data set using two qualitatively different priors. Solid curves correspond to calibrated measures and dotted ones depict the priors. Right graph: zoom of the left tail of Lévy densities on log scale.

3.3 Choice of the regularization parameter

Theorem 2.17 of the preceding chapter suggests that for the regularization method to converge, the parameter choice strategy $\alpha(\delta)$ must satisfy the following *limiting* relations, when the data error δ goes to zero:

- $\alpha(\delta) \to 0$,
- If the exact data C_M is attainable by the model, one must have $\frac{\delta^2}{\alpha(\delta)} \to 0$ and if not, $\frac{\delta}{\alpha(\delta)} \to 0$.

A typical example of a parameter choice rule that works in both attainable and unattainable cases would be

$$\alpha(\delta) = C\delta^{\mu} \tag{3.10}$$

with any $C > 0$ and any $\mu \in (0, 1)$.

However, in real calibration problems the error level δ is fixed and finite and rules of type (3.10) do not tell us how we should choose α in this case. Good performance and error control for finite values of δ is achieved by using the so called a posteriori parameter choice rules (α depends not only on δ but also on the data C_M^δ), the most widely used of which is the *discrepancy principle*, originally developed by Morozov for Tikhonov regularization in Banach spaces [74, 75], see also [40]. In the rest of this section we apply this principle and its variants that is better suited for nonlinear problems to our entropic regularization method.

3.3.1 A posteriori parameter choice rules for attainable calibration problems

In this subsection we make the following standing assumptions:

1. The prior Lévy process corresponds to an arbitrage-free model with jumps bounded from above by B: $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ (this implies the existence of a minimal entropy martingale measure Q^*).

2. There exists a solution Q^+ of problem (2.11) with data C_M (minimum entropy least squares solution) such that

$$I(Q^+|P) < \infty \quad \text{and}$$
$$\|C^{Q^+} - C_M\|_w = 0 \quad \text{(the data is attainable by an exp-Lévy model).}$$

3. There exists δ_0 such that

$$\varepsilon^{\max} := \inf_{\delta \le \delta_0} \|C^{Q^*} - C_M^\delta\|_w^2 > \delta_0^2. \tag{3.11}$$

Remark 3.1. In the condition (3.11), δ_0 can be seen as the highest possible noise level of all data sets that we consider. If, for some δ, $\|C^{Q^*} - C_M^\delta\|_w < \delta_0$ then either Q^* is already sufficiently close to the solution or the noise level is so high that all the useful information in the data is

completely blurred and the data does not allow to construct a better approximation of the true solution than Q^*.

Let Q_α^δ denote a solution of the regularized problem (2.27) with data C_M^δ and regularization parameter α. The function

$$\varepsilon_\delta(\alpha) := \|C^{Q_\alpha^\delta} - C_M^\delta\|_w^2$$

is called the discrepancy function of the calibration problem (2.27). Note that since this problem can have many solutions, $\varepsilon_\delta(\alpha)$ is a priori a multivalued function. Given two constants c_1 and c_2 satisfying

$$1 < c_1 \leq c_2 < \frac{\varepsilon^{\max}}{\delta_0^2}, \tag{3.12}$$

the discrepancy principle can be stated as follows:

Discrepancy principle For a given noise level δ, choose $\alpha > 0$ that satisfies

$$\delta^2 < c_1\delta^2 \leq \varepsilon_\delta(\alpha) \leq c_2\delta^2, \tag{3.13}$$

If, for a given α, the discrepancy function has several possible values, the above inequalities must be satisfied by each one of them.

The intuition behind this principle is as follows. We would like to find a solution Q of the equation $C^Q = C_M$. Since the error level in the data is of order δ, it is the best possible precision that we can ask for in this context, so it does not make sense to calibrate the noisy data C_M^δ with a precision higher than δ. Therefore, we try to solve $\|C^{Q_\alpha^\delta} - C_M^\delta\|_w^2 \leq \delta^2$. In order to gain stability we must sacrifice some precision compared to δ, therefore, we choose a constant c with $1 \lesssim c$, for example, $c = 1.1$ and look for Q_α^δ in the level set

$$\|C^{Q_\alpha^\delta} - C_M^\delta\|_w^2 \leq c\delta^2. \tag{3.14}$$

Since, on the other hand, by increasing precision, we decrease the stability, the highest stability is obtained when the inequality in (3.14) is replaced by equality and we obtain

$$\varepsilon_\delta(\alpha) \equiv \|C^{Q_\alpha^\delta} - C_M^\delta\|_w^2 = c\delta^2.$$

To make the numerical solution of the equation easier, we do not impose a strict value of the discrepancy function but allow it to lie between two bounds, obtaining (3.13).

Supposing that an α solving (3.13) exists, it is easy to prove the convergence of the regularized solutions to the minimal entropy least squares solution, when the regularization parameter is chosen using the discrepancy principle.

Proposition 3.3. *Suppose that the hypotheses 1–3 of page 103 are satisfied and let c_1 and c_2 be as in* (3.12). *Let $\{C_M^{\delta_k}\}_{k\geq 1}$ be a sequence of data sets such that $\|C_M - C_M^{\delta_k}\|_w < \delta_k$ and $\delta_k \to 0$.*

Then the sequence $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$ where $Q_{\alpha_k}^{\delta_k}$ is a solution of problem (2.27) *with data $C_M^{\delta_k}$, prior P and regularization parameter α_k chosen according to the discrepancy principle, has a weakly convergent subsequence. The limit of every such subsequence of $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$ is a MELSS with data C_M and prior P.*

Proof. Using the optimality of $Q_{\alpha_k}^{\delta^k}$, we can write:

$$\varepsilon_{\delta_k}(\alpha_k) + \alpha_k I(Q_{\alpha_k}^{\delta^k}|P) \leq \|C^{Q^+} - C_M^{\delta_k}\|_w^2 + \alpha_k I(Q^+|P) \leq \delta_k^2 + \alpha_k I(Q^+|P).$$

The discrepancy principle (3.13) then implies that

$$I(Q_{\alpha_k}^{\delta^k}|P) \leq I(Q^+|P), \tag{3.15}$$

By Lemma 2.10, the sequence $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$ is tight and therefore, by Prohorov's theorem and Lemma 2.4, relatively weakly compact in $\mathcal{M} \cap \mathcal{L}_B^+$.

Choose a subsequence of $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$, converging weakly to a limit Q and denoted, to simplify notation, again by $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$. Inequality (3.13) and the triangle inequality yield:

$$\|C^{Q_{\alpha_k}^{\delta_k}} - C_M\|_w \leq \|C^{Q_{\alpha_k}^{\delta_k}} - C_M^{\delta_k}\|_w + \delta_k \leq \delta_k(1 + \sqrt{c_2}) \xrightarrow[k\to\infty]{} 0.$$

By Lemma 2.2 this means that $\|C^Q - C_M\|_w = 0$ and therefore Q is a solution. By weak lower semicontinuity of I (cf. Corollary 2.1) and using (3.15),

$$I(Q|P) \leq \liminf_k I(Q_{\alpha_k}^{\delta_k}|P) \leq \limsup_k I(Q_{\alpha_k}^{\delta_k}|P) \leq I(Q^+|P),$$

which means that Q is a MELSS. □

The discrepancy principle performs well for regularizing linear operators but may fail in nonlinear problems like (2.27), because Equation (3.13) may have no solution due to discontinuity of the discrepancy function $\varepsilon_\delta(\alpha)$. Examples of nonlinear ill-posed problems to which the

discrepancy principle cannot be applied are given in [95] and [50]. However, in our numerical tests (see Section 3.6) we have always been able to find a solution to (3.13)

We will now give a simple sufficient condition, adapted from [94], under which (3.13) admits a solution.

Proposition 3.4. *Suppose that the hypotheses 1–3 of page 103 are satisfied and let c_1 and c_2 satisfy* (3.12)*. If $\varepsilon_\delta(\alpha)$ is a single-valued function then there exists an α satisfying* (3.13)*.*

This proposition is a direct consequence of the following lemma.

Lemma 3.5. *The function $\varepsilon_\delta(\alpha)$ is non-decreasing and satisfies the following limit relations:*

$$\lim_{\alpha\downarrow 0}\varepsilon_\delta(\alpha)\leq\delta^2,\qquad \lim_{\alpha\to\infty}\varepsilon_\delta(\alpha)=\|C^{Q^*}-C_M^\delta\|_w^2.$$

If, at some point $\alpha>0$, $\varepsilon_\delta(\alpha)$ is single-valued, then it is continuous at this point.

The function

$$J_\delta(\alpha):=\|C^{Q_\alpha^\delta}-C_M^\delta\|_w^2+\alpha I(Q_\alpha^\delta|P).$$

is non-decreasing, continuous, and satisfies the following limit relations:

$$\lim_{\alpha\downarrow 0}J_\delta(\alpha)\leq\delta^2,\qquad \lim_{\alpha\to\infty}J_\delta(\alpha)\geq\|C^{Q^*}-C_M^\delta\|_w^2.$$

Proof. Let $\gamma_\delta(\alpha):=I(Q_\alpha^\delta|P)$ and let $0<\alpha_1<\alpha_2$. By the optimality of $Q_{\alpha_1}^\delta$ and $Q_{\alpha_2}^\delta$ we have:

$$\varepsilon_\delta(\alpha_1)+\alpha_1\gamma_\delta(\alpha_1)\leq\varepsilon_\delta(\alpha_2)+\alpha_1\gamma_\delta(\alpha_2),$$
$$\varepsilon_\delta(\alpha_2)+\alpha_2\gamma_\delta(\alpha_2)\leq\varepsilon_\delta(\alpha_1)+\alpha_2\gamma_\delta(\alpha_1)$$

and therefore

$$\varepsilon_\delta(\alpha_2)-\varepsilon_\delta(\alpha_1)\geq\alpha_1(\gamma_\delta(\alpha_1)-\gamma_\delta(\alpha_2))$$
$$\varepsilon_\delta(\alpha_2)-\varepsilon_\delta(\alpha_1)\leq\alpha_2(\gamma_\delta(\alpha_1)-\gamma_\delta(\alpha_2)),$$

which implies that $\varepsilon_\delta(\alpha_2)\geq\varepsilon_\delta(\alpha_1)$ and $\gamma_\delta(\alpha_1)\geq\gamma_\delta(\alpha_2)$. To prove the first limit relation for $\varepsilon_\delta(\alpha)$, observe that for all $\alpha>0$,

$$\varepsilon_\delta(\alpha)\leq\|C^{Q^+}-C_M^\delta\|_w^2+\alpha I(Q^+|P)\leq\delta^2+\alpha I(Q^+|P)\xrightarrow[\alpha\to 0]{}\delta^2.$$

To prove the second limit relation for $\varepsilon_\delta(\alpha)$, one can write, using the optimality of Q_α^δ:

$$\varepsilon_\delta(\alpha)+\alpha\gamma_\delta(\alpha)\leq\|C^{Q^*}-C_M^\delta\|_w^2+\alpha I(Q^*|P).$$

From [30, Theorem 2.2],

$$\gamma_\delta(\alpha) \geq I(Q_\alpha^\delta|Q^*) + I(Q^*|P). \tag{3.16}$$

Therefore

$$I(Q_\alpha^\delta|Q^*) \leq \frac{\|C^{Q^*} - C_M^\delta\|_w^2}{\alpha} \xrightarrow[\alpha\to\infty]{} 0.$$

Using the inequality

$$|P - Q| \leq \sqrt{2I(P|Q)}, \tag{3.17}$$

where $|P - Q|$ denotes the total variation distance (see [30, Equation (2.3)]), this implies that Q_α^δ converges to Q^* in total variation distance (and therefore also weakly) as α goes to infinity. The limit relation now follows from Lemma 2.2.

To prove the continuity of $\varepsilon_\delta(\alpha)$, let $\{\alpha_n\}$ be a sequence of positive numbers, converging to $\alpha > 0$. By the optimality of $Q_{\alpha_n}^\delta$, $I(Q_{\alpha_n}^\delta|P)$ is bounded and one can choose a subsequence of $\{Q_{\alpha_n}^\delta\}$, converging weakly toward some measure Q', and denoted, to simplify notation, again by $\{Q_{\alpha_n}^\delta\}_{n\geq 1}$. We now need to prove that Q' is the solution of the calibration problem with regularization parameter α. By weak continuity of the pricing error (Lemma 2.2) and weak lower semicontinuity of the relative entropy (Lemma 2.11), we have for any other measure Q:

$$\begin{aligned}
\|C^{Q'} - C_M^\delta\|_w^2 + \alpha I(Q'|P) &\leq \liminf_n \{\|C^{Q_{\alpha_n}^\delta} - C_M^\delta\|_w^2 + \alpha I(Q_{\alpha_n}^\delta|P)\} \\
&= \liminf_n \{\|C^{Q_{\alpha_n}^\delta} - C_M^\delta\|_w^2 + \alpha_n I(Q_{\alpha_n}^\delta|P) + (\alpha - \alpha_n) I(Q_{\alpha_n}^\delta|P)\} \\
&\leq \liminf_n \{\|C^{Q} - C_M^\delta\|_w^2 + \alpha_n I(Q|P) + (\alpha - \alpha_n) I(Q_{\alpha_n}^\delta|P)\} \\
&= \|C^{Q} - C_M^\delta\|_w^2 + \alpha I(Q|P).
\end{aligned}$$

Therefore, the sequence $\{\varepsilon_\delta(\alpha_n)\}$ converges to one of the possible values of $\varepsilon_\delta(\alpha)$. If this function is single-valued in α, this means that every subsequence of the original sequence $\{\varepsilon_\delta(\alpha_n)\}$ has a further subsequence that converges toward $\varepsilon_\delta(\alpha)$, and therefore, the original sequence also converges toward $\varepsilon_\delta(\alpha)$.

The fact that $J_\delta(\alpha)$ is nondecreasing as a function of α is trivial. To show the continuity, observe that for $0 < \alpha_1 < \alpha_2$,

$$\begin{aligned}
J_\delta(\alpha_2) - J_\delta(\alpha_1) &\leq \varepsilon_\delta(\alpha_1) + \alpha_2\gamma_\delta(\alpha_1) - \varepsilon_\delta(\alpha_1) - \alpha_1\gamma_\delta(\alpha_1) \\
&= (\alpha_2 - \alpha_1)\gamma_\delta(\alpha_1) \leq (\alpha_2 - \alpha_1) I(Q^+|P).
\end{aligned}$$

The limit relations for $J_\delta(\alpha)$ follow from the relations for $\varepsilon_\delta(\alpha)$. □

We will now present an alternative a posteriori parameter choice rule, which reduces to the discrepancy principle when inequality (3.13) has a solution but also works when this is not the case. However, if the parameter is chosen according to the alternative rule, the sequence of regularized solutions does not necessarily converge to a minimum entropy solution as in Proposition 3.3 but to a solution with bounded entropy (see Proposition 3.7). Our treatment partly follows [50] where this parameter choice rule is applied to Tikhonov regularization.

Alternative principle For a given noise level δ, if there exists $\alpha > 0$ that satisfies

$$c_1\delta^2 \leq \varepsilon_\delta(\alpha) \leq c_2\delta^2, \tag{3.18}$$

choose one such α; otherwise, choose an $\alpha > 0$ that satisfies

$$\varepsilon_\delta(\alpha) \leq c_1\delta^2, \qquad J_\delta(\alpha) \geq c_2\delta^2. \tag{3.19}$$

Proposition 3.6. *Suppose that the hypotheses 1–3 of page 103 are satisfied and let c_1 and c_2 be as in* (3.12)*. Then there exists $\alpha > 0$ satisfying either* (3.18) *or* (3.19)*.*

Proof. Suppose that (3.18) does not admit a solution. We need to prove that there exists $\alpha > 0$ satisfying (3.19). Let

$$B := \{\alpha > 0 : \varepsilon_\delta(\alpha) \leq c_1\delta^2\} \qquad \text{and} \qquad U := \{\alpha > 0 : \varepsilon_\delta(\alpha) > c_2\delta^2\}.$$

The limit relations of Lemma 3.5 imply that both sets are nonempty. Moreover, since we have assumed that (3.18) does not admit a solution, necessarily $\sup B = \inf U$. Let $\alpha^* := \sup B \equiv \inf U$. Now we need to show that

$$J_\delta(\alpha^*) > c_2\delta^2. \tag{3.20}$$

By continuity of $J_\delta(\alpha)$,

$$J_\delta(\alpha^*) \geq c_2\delta^2 + \lim_{\alpha\downarrow\alpha^*} \gamma_\delta(\alpha).$$

If $\lim_{\alpha\downarrow\alpha^*} \gamma_\delta(\alpha) > 0$ then (3.20) holds. Otherwise from (3.16), P is the minimal entropy martingale measure and (3.17) implies that $Q_\alpha^\delta \Rightarrow P$ as $\alpha \downarrow \alpha^*$. Therefore, $J_\delta(\alpha^*) = \lim_{\alpha\downarrow\alpha^*} \varepsilon_\delta(\alpha) = \|C^{Q^*} - C_M^\delta\|_w^2 > c_2\delta^2$ and (3.20) is also satisfied.

By continuity of $J_\delta(\alpha)$, there exists $\Delta > 0$ such that $J_\delta(\alpha^* - \Delta) > c_2\delta^2$. However, since $\alpha^* = \sup B$ and $\varepsilon_\delta(\alpha)$ is nondecreasing, necessarily $\varepsilon_\delta(\alpha^* - \Delta) \leq c_1\delta^2$. Therefore, $\alpha - \Delta$ is a solution of (3.19). □

Remark 3.2. If $c_1 < c_2$, one can show, along the lines of the above proof, that there exists not a single α that satisfies either (3.18) or (3.19) but an interval of nonzero length (α_1, α_2), such that each point inside this interval satisfies one of the two conditions. From the computational viewpoint this means that a feasible α can be found by bisection with a finite number of iterations.

Proposition 3.7. *Suppose that the hypotheses 1–3 of page 103 are satisfied and that c_1 and c_2 are chosen according to* (3.12). *Let $\{C_M^{\delta_k}\}_{k\geq 1}$ be a sequence of data sets such that $\|C_M - C_M^{\delta_k}\|_w < \delta_k$ and $\delta_k \to 0$.*

Then the sequence $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$ where $Q_{\alpha_k}^{\delta_k}$ is a solution of problem (2.27) *with data $C_M^{\delta_k}$, prior P and regularization parameter α_k chosen according to the alternative principle, has a weakly convergent subsequence. The limit Q' of every such subsequence of $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$ satisfies*

$$\|C^{Q'} - C_M\|_w = 0$$
$$I(Q'|P) \leq \frac{c_2}{c_2 - 1} I(Q^+|P).$$

Proof. Using the optimality of $Q_{\alpha_k}^{\delta^k}$, we can write:

$$\varepsilon_{\delta_k}(\alpha_k) + \alpha_k I(Q_{\alpha_k}^{\delta^k}|P) \leq \|C^{Q^+} - C_M^{\delta_k}\|_w^2 + \alpha_k I(Q^+|P) \leq \delta^2 + \alpha_k I(Q^+|P). \tag{3.21}$$

If α_k satisfies (3.18), the above implies that

$$I(Q_{\alpha_k}^{\delta^k}|P) \leq I(Q^+|P),$$

otherwise (3.19) entails that

$$\varepsilon_{\delta_k}(\alpha_k) + \alpha_k I(Q_{\alpha_k}^{\delta^k}|P) \geq c_2\delta^2,$$

and together with (3.21) this gives

$$I(Q_{\alpha_k}^{\delta^k}|P) \leq \frac{c_2}{c_2 - 1} I(Q^+|P).$$

In both cases $I(Q_{\alpha_k}^{\delta^k}|P)$ is uniformly bounded, which means, by Lemma 2.10, that the sequence $\{Q_{\alpha_k}^{\delta_k}\}_{k\geq 1}$ is tight and therefore, by Prohorov's theorem, relatively weakly compact. The rest of the proof follows the lines of the proof of Proposition 3.3. □

3.3.2 A posteriori parameter choice rule for non-attainable calibration problems

In this subsection we suppose that hypothesis 1 of page 103 is satisfied and that hypotheses 2 and 3 are replaced with the following ones:

2a. There exists a solution Q^+ of problem (2.11) with data C_M such that

$$I(Q^+|P) < \infty \quad \text{and}$$
$$\|C^{Q^+} - C_M\|_w > 0 \quad \text{(the data is not attainable by an exp-Lévy model).}$$

3a. There exists δ_0 such that

$$\varepsilon^{\max} := \inf_{\delta \le \delta_0} \{\|C^{Q^*} - C_M^\delta\|_w^2 - \|C^{Q^+} - C_M^\delta\|_w^2\} > 0.$$

Given two constants c_1 and c_2 satisfying

$$0 < c_1 \le c_2 < \frac{\varepsilon^{\max}}{\delta_0^2}, \tag{3.22}$$

the discrepancy principle can be stated as follows:

Discrepancy principle for non-attainable data For a given noise level δ, choose $\alpha > 0$ that satisfies

$$c_1\delta^2 \le \varepsilon_\delta(\alpha) - \varepsilon_\delta(0) \le c_2\delta^2, \tag{3.23}$$

Proofs of the following two results are similar to the proofs of Propositions 3.4 and 3.3 of the preceding subsection and are therefore omitted.

Proposition 3.8. *Suppose that hypothesis 1 of page 103 and hypotheses 2a and 3a of page 110 are satisfied and let c_1 and c_2 satisfy* (3.22)*. If $\varepsilon_\delta(\alpha)$ is a single-valued function then there exists an α satisfying* (3.23)*.*

Proposition 3.9. *Suppose that hypothesis 1 of page 103 and hypotheses 2a and 3a of page 110 are satisfied and let c_1 and c_2 satisfy* (3.22) *for all δ. Let $\{C_M^{\delta_k}\}_{k\ge1}$ be a sequence of data sets such that $\|C_M - C_M^{\delta_k}\|_w \le \delta_k$ and $\delta_k \to 0$.*

Then the sequence $\{Q^{\delta_k}_{\alpha_k}\}_{k\geq 1}$ *where* $Q^{\delta_k}_{\alpha_k}$ *is a solution of problem* (2.27) *with data* $C^{\delta_k}_M$*, prior* P *and regularization parameter* α_k *chosen according to the discrepancy principle, has a weakly convergent subsequence. The limit of every such subsequence of* $\{Q^{\delta_k}_{\alpha_k}\}_{k\geq 1}$ *is a MELSS with data* C_M *and prior* P*.*

The alternative principle of the preceding section does not carry over to non-attainable problems as easily and is not discussed here.

3.3.3 Computing the noise level

If the bid and ask prices are known, the noise level δ and an estimate of option prices C^δ_M can be computed directly using

$$C^\delta_M(T,K) := \frac{C^{bid}_M(T,K) + C^{ask}(T,K)}{2}, \quad \forall\, T,K,$$
$$\delta := \frac{\|C^{bid}_M - C^{ask}_M\|_w}{2}.$$

Since for all i, the true option prices satisfy $C_M(T_i,K_i) \in (C^{bid}_M, C^{ask}_M)$, we clearly have $\|C^\delta_M - C_M\|_w \leq \delta$.

If the bid and ask prices are unavailable, one can assess the order of magnitude of δ directly from C^δ_M, using the following heuristic argument. Suppose that the exact data C_M is attainable and that *the errors in different option prices are independent*. Then it is reasonable to assume that the main part of error, present in C^δ_M will lead to violations of arbitrage constraints on option prices (e.g. convexity). Since the least squares solution Q^+ is an arbitrage-free model, these violations of no-arbitrage constraints will contribute to the discrepancy $\|C^{Q^+} - C^\delta_M\|^2_w \equiv \varepsilon_\delta(0)$, as shown in Figure 3.3. Therefore, $\varepsilon_\delta(0)$ will have the same order of magnitude as δ^2 and one can approximately take $\delta \sim \sqrt{\varepsilon_\delta(0)}$. Since the noise level δ is only used to choose the regularization parameter α, we do not need to know it with high precision and this rough estimate is sufficient.

3.4 Choice of the weights of option prices

The relative weights w_i of option prices in the pricing error term (2.3) should reflect our confidence in individual data points, which is determined by the liquidity of a given option. This

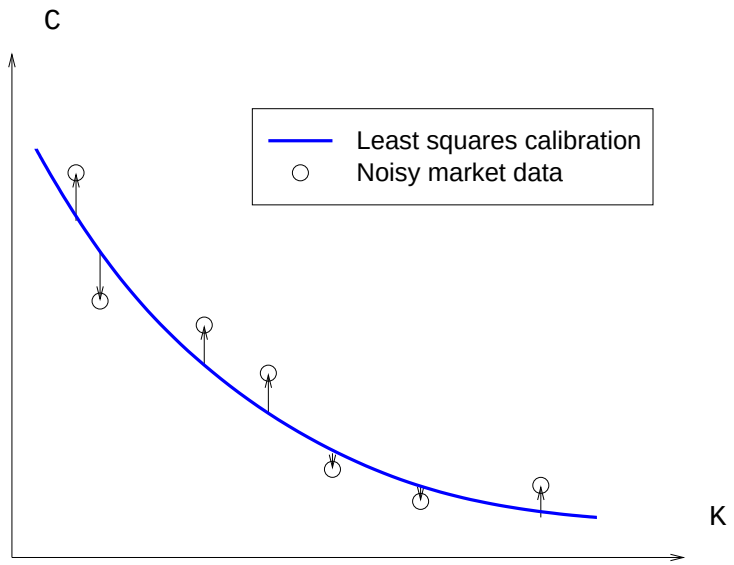

Figure 3.3: Estimating the noise level in the data.

can be assessed from the bid-ask spreads by choosing

$$w_i := \frac{w}{(C_M^{bid}(T_i, K_i) - C_M^{ask}(T_i, K_i))^2},$$

where w is the normalization constant, determined from $\sum_i w_i = 1$. However, the bid and ask quotes are not always available from option price data bases. On the other hand, it is known that at least for the options that are not too far from the money, the bid-ask spread in implied volatility units is of order of 1%. This means that to have errors proportional to bid-ask spreads, one must minimize the differences of implied volatilities and not those of the option prices. However, this would prohibitively increase the computational burden since one would have to invert numerically the Black-Scholes formula once for each data point at each minimization step. each A feasible solution is to minimize the squared differences of option prices weighted by the Black Scholes "vegas" evaluated at the points corresponding to market option prices. Denoting the implied volatility computed in the model Q for strike K and maturity T by $\Sigma^Q(T, K)$ and the corresponding market implied volatility by $\Sigma_M(T, K)$, we

have the following approximation:

$$\begin{aligned}
&\sum_{i=1}^{N}(\Sigma^Q(T_i,K_i)-\Sigma_M(T_i,K_i))^2 \\
&\quad\approx \sum_{i=1}^{N}\left(\frac{\partial\Sigma}{\partial C}(C_M(T_i,K_i))|C^Q(T_i,K_i)-C_M(T_i,K_i)|\right)^2 \\
&\quad= \sum_{i=1}^{N}\frac{(C^Q(T_i,K_i)-C_M(T_i,K_i))^2}{\mathrm{Vega}^2(\Sigma_M(T_i,K_i))},
\end{aligned} \tag{3.24}$$

where Vega denotes the derivative of the Black-Scholes option price with respect to volatility:

$$\mathrm{Vega}(\sigma)=S\sqrt{T}n\left(\frac{1}{\sigma\sqrt{T}}\log\left(\frac{S}{Ke^{-rT}}\right)+\frac{1}{2}\sigma\sqrt{T}\right),$$

with n denoting the CDF of the standard normal distribution. Therefore, by using $w_i=\frac{w}{\mathrm{Vega}_i^2}$ one can achieve the correct weighting of option prices without increasing the computational burden[1] (because w_i can be computed in advance.)

3.5 Numerical solution of calibration problem

Once the (discrete) prior $P=P(A,\nu^P,\gamma^P)$, the regularization parameter α and the weights w_i have been fixed, it remains to find the numerical solution of the calibration problem (2.27).

We construct an approximate solution of the calibration problem (2.27) by minimizing an approximation of the calibration functional $J_\alpha(Q)$, denoted by

$$\hat{J}_\alpha(Q):=\|C_M-\hat{C}^Q\|_w^2+\alpha I(Q|P), \tag{3.25}$$

where $\hat{C}^Q$ is the approximate option price defined by Equation (3.31) below. The minimization is done over all Lévy processes $Q\in\mathcal{M}\cap\mathcal{L}_B^+$ with Lévy measures of the form (3.2). The calibration functional therefore becomes a function of a finite number of arguments:

$$\hat{J}_\alpha(Q)=\hat{J}_\alpha(q_0,\ldots,q_{M-1}), \tag{3.26}$$

To minimize (3.26) we use a variant of the popular Broyden-Fletcher-Goldfarb-Shanno (BFGS) variable metric algorithm. For a description of the algorithm see [80]. This algorithm requires

[1] Since vegas can be very small for out of the money options, to avoid giving them too much weight one should impose a lower bound on the vegas used for weighting, i.e., use weights of the form $w_i=\frac{w}{(\mathrm{Vega}_i\wedge C)^2}$ with some constant $C\sim 10^{-2}\div 10^{-1}$.

the user to supply the function value and its gradient and tries, iteratively, to build up a good approximation of the inverse Hessian matrix of the functional being optimized. In our numerical examples we used the LBFGS implementation by Jorge Nocedal et al. [20]. The algorithm is typically initialized with the prior Lévy measure.

Since gradient-based methods only allow to find one local minimum of the objective function and the calibration functional (3.26) is not convex, there is no guarantee that the BFGS algorithm will converge to its true global minimum. However, starting the optimization procedure from different initializers, we have empirically observed (see Section 3.6.1) that in presence of even a small regularization, the minimizers do not depend on the starting point of the minimization algorithm and that using a gradient-based optimization procedure produces an acceptable calibration quality.

In the rest of this section we show how to compute numerically the functional (3.26) and its gradient. For simplicity, from now on we will suppose that the prior Lévy process has a non-zero diffusion part ($A > 0$).

3.5.1 Computing the calibration functional

Substituting the discrete expressions (3.1) and (3.2) into formula (3.25), we obtain:

$$\hat{J}_\alpha(Q) = \sum_{i=1}^{N} w_i (\hat{C}^Q(T_i, K_i) - C_M(T_i, K_i))^2 + \frac{\alpha}{2A}\left(\frac{A}{2} + b^P + \sum_{j=0}^{M-1}(e^{x_j} - 1)q_j\right)^2 + \alpha \sum_{j=0}^{M-1}\left(q_j \log(q_j/p_j) + 1 - q_j\right), \quad (3.27)$$

where $b^P = \gamma^P - \int_{|x|\leq 1} x\nu^P(dx)$ is the drift of the prior process. The last two terms of the above equation can be evaluated directly using a finite number of computer operations, we will therefore concentrate on the first term.

The approximate option prices $\hat{C}^Q(T_i, K_i)$ are computed using the Fourier transform algorithm of Section 1.4 as follows: for each maturity date, present in the data, prices are first computed on a fixed grid of strikes and then interpolated to obtain prices at strikes K_i corresponding to this maturity date. The number of Fourier transforms is thus determined by the number of maturity dates present in the data, which is typically smaller than 10.

To use the Fourier transform method, the first step is to compute the characteristic exponent

of X_t under Q at the points $u-i$ for real u:

$$\psi^Q(u-i) = -\frac{A}{2}u(u-i) - (1+iu)\sum_{j=0}^{M-1}(e^{x_j}-1)q_j - \sum_{j=0}^{M-1} q_j + \sum_{j=0}^{M-1} e^{iux_j}e^{x_j}q_j.$$

Introducing uniform grids

$$x_j = x_0 + jd \qquad \text{and} \qquad u_k = u_0 - k\Delta,$$

the last term becomes:

$$\sum_{j=0}^{M-1} e^{iu_k x_j}e^{x_j}q_j = e^{-ik\Delta x_0}\sum_{j=0}^{M-1} e^{-ikjd\Delta}e^{(iu_0+1)x_j}q_j$$

and comparing this to Equation (1.27), we see that if $d\Delta = \frac{2\pi}{M}$, the last term of $\{\psi^Q(u_k-i)\}_{k=0}^{M-1}$ becomes a discrete Fourier transform (cf. Equation (1.27)):

$$\psi_k := \psi^Q(u_k-i) = -\frac{A}{2}u_k(u_k-i) - (1+iu_k)\sum_{j=0}^{M-1}(e^{x_j}-1)q_j - \sum_{j=0}^{M-1} q_j + e^{-ik\Delta x_0}\mathrm{DFT}_k[e^{(iu_0+1)x_j}q_j]. \quad (3.28)$$

This expression can therefore be computed using the fast Fourier transform algorithm. Note that at this stage all computations are exact: there are no truncation or discretization errors.

For a given maturity date T we now need to compute the Fourier transform $\tilde{\zeta}_T$ of the modified time value of options (see Equation (1.25)) at the points $\{u_k\}_{k=0}^{M-1}$:

$$\tilde{\zeta}_T(u_k) = e^{iu_k rT}\frac{e^{T\psi_k} - e^{-\frac{AT}{2}(u_k(u_k-i))}}{iu_k(1+iu_k)}. \quad (3.29)$$

Option time values are then approximated using Equation (1.28) on the same grid of log-strikes $\{x_j\}_{j=0}^{M-1}$ that was used to discretize the Lévy measure.[2] In the following equation and below we denote approximated quantities by putting a hat over the corresponding variables.

$$\hat{\tilde{z}}_T(x_j) = \frac{\Delta}{2\pi}e^{-ix_j u_{M-1}}\sum_{k=0}^{M-1} w_k\tilde{\zeta}_T(u_{M-1-k})e^{-ix_0\Delta k}e^{-2\pi ijk/M} = \frac{\Delta}{2\pi}e^{-ix_j u_{M-1}}\mathrm{DFT}_j[w_k\tilde{\zeta}_T(u_{M-1-k})e^{-ix_0\Delta k}], \quad (3.30)$$

[2] Actually, the grid of log-strikes may be shifted but we do not do this here to simplify the formulas.

where w_k are the weights corresponding to the chosen integration rule. For a given strike K such that $x_0 \leq \log K \leq x_{M-1}$, the option price can now be computed by linear interpolation (other interpolation methods can also be used):

$$\hat{C}^Q(T,K) = C_{BS}^{\sqrt{A}}(T,K) + \frac{\log K - x_{n_K}}{x_{n_K+1} - x_{n_K}} \hat{\tilde{z}}_T(x_{n_K+1}) + \frac{x_{n_K+1} - \log K}{x_{n_K+1} - x_{n_K}} \hat{\tilde{z}}_T(x_{n_K}), \tag{3.31}$$

where $n_K := \sup\{n : x_n \leq \log K\}$ and $C_{BS}^{\sqrt{A}}(T,K)$ denotes the Black Scholes price of a call option with time to maturity T, strike K and volatility $\sqrt{A}$.

Error control The above method of evaluating option prices contains three types of numerical errors: truncation and discretization errors appear when the integral is replaced by a finite sum using Equation (1.28), and interpolation error appears in Equation (3.31). Supposing that the grid in Fourier space is centered, that is, $u_0 = L/2$ and $\Delta = L/(M-1)$ for some constant L, Section 1.4 allows to obtain the following bounds for the first two types of error.

Since we have supposed that $A > 0$, Equation (1.30) provides a uniform (with respect to Q) bound for the truncation error:

$$|\varepsilon_T| \leq \frac{16 e^{-TAL^2}}{\pi T A L^3}.$$

Proposition 1.13 does not give uniform bounds for the discretization (sampling) error, however, since the calibrated Lévy measures are usually similar to the Lévy measures of the prior process (see Section 3.6), the order of magnitude of the discretization error can be estimated by computing the bounds of Proposition 1.13 for the prior process. The exact form of the error depends on the integration rule used; for example for Simpson's rule one has

$$|\varepsilon_D| = O\left(\frac{L^4 \log L}{M^4}\right).$$

A uniform bound for the interpolation error depends on the interpolation method that one is using. For linear interpolation using Equation (3.31) one easily obtains

$$|\varepsilon_I| \leq \frac{d^2}{8} \max \tilde{z}_T'',$$

and the second derivative of the time value is uniformly bounded under the hypothesis $A > 0$

since

$$\left|\tilde{z}''_T(k)\right| = \frac{1}{2\pi}\left|\int_{-\infty}^{\infty} v^2 e^{-ivk}\tilde{\zeta}_T(v)dv\right| \\ \leq \frac{1}{2\pi}\int_{-\infty}^{\infty}|\Phi_T(v-i)|dv + \frac{1}{2\pi}\int_{-\infty}^{\infty}|\Phi_T^{\sqrt{A}}(v-i)|dv \leq \sqrt{\frac{2}{\pi AT}},$$

where we have used Equation (1.25) and Lemma 1.10. Combining the above two expressions, one obtains:

$$|\varepsilon_I| \leq \frac{1}{L^2}\sqrt{\frac{\pi^3}{2AT}}.$$

Taking L and M sufficiently large so that L/M becomes small, one can make the total error $\varepsilon = \varepsilon_T + \varepsilon_D + \varepsilon_I$ arbitrarily small. In practice, the parameters M and L of the discretization scheme should be chosen such that the total numerical error is of the same order as the noise level in the data. The continuity result (Theorem 2.16) then guarantees that the numerical approximation of the calibrated measure will be close to the true solution.

3.5.2 Computing the gradient of the calibration functional

We emphasize that for the optimization algorithm to work correctly, we must compute the *exact* gradient of the approximate functional (3.26), rather than an approximation of the gradient of the exact functional. The gradient of the approximate calibration functional is computed as follows:

$$\frac{\partial \hat{J}_\alpha(q_0,\ldots,q_{M-1})}{\partial q_k} = \sum_{i=1}^{N} w_i(\hat{C}^Q(T_i,K_i) - C_M(T_i,K_i))\frac{\partial \hat{C}^Q(T_i,K_i)}{\partial q_k} \\ + \frac{\alpha(e^{x_k}-1)}{2A}\left(\frac{A}{2} + b^P + \sum_{j=0}^{M-1}(e^{x_j}-1)q_j\right) + \alpha\log(q_k/p_k).$$

The nontrivial part is therefore to compute the gradient of the approximate option price $\hat{C}^Q(T_i,K_i)$ and for this, due to the linear structure of the interpolating formula (3.31) it suffices to know the gradient of $\hat{\tilde{z}}_T$ at the points $\{x_i\}_{i=0}^{M-1}$. From Equation (3.30),

$$\frac{\partial \hat{\tilde{z}}_T(x_j)}{\partial q_m} = \frac{\Delta}{2\pi}e^{-ix_j u_{M-1}}\mathrm{DFT}_j\left[w_k\frac{\partial\tilde{\zeta}_T(u_{M-1-k})}{\partial q_m}e^{-ix_0\Delta k}\right],$$

and formulas (3.28–3.29) yield:

$$\begin{aligned}\frac{\partial\tilde{\zeta}_T(u_k)}{\partial q_m} &= \frac{Te^{iu_krT+T\psi_k}}{iu_k(1+iu_k)}\left\{e^{x_m+iu_kx_m}-1-(1+iu_k)(e^{x_m}-1)\right\}\\ &= T(e^{x_m}-1)\frac{e^{iu_krT+T\psi_k}}{1+iu_k}+Te^{x_m}\frac{e^{iu_krT}(e^{T\psi_k}-e^{T\psi_k^A})}{iu_k(1+iu_k)}e^{iu_kx_m}\\ &- Te^{x_m}\frac{e^{iu_krT}(e^{T\psi_k}-e^{T\psi_k^A})}{iu_k(1+iu_k)}+Te^{x_m}\frac{e^{iu_krT+T\psi_k^A}(e^{iu_kx_m}-1)}{iu_k(1+iu_k)},\end{aligned}$$

where $\psi_k^A=-\frac{A}{2}u_k(u_k-i)$. Suppose that the grid in strike space is such that $x_0=m_0d$ for some $m_0\in\mathbb{Z}$. Then for every $j\in\{0,\dots,M-1\}$,

$$\begin{aligned}\mathrm{DFT}_j[f_ke^{iu_{M-1-k}x_m}] &= e^{iu_{M-1}x_m}\sum_{k=0}^{M-1}e^{-2\pi ik(j-m-m_0)/M}f_k\\ &= e^{iu_{M-1}x_m}\mathrm{DFT}_{(j-m_0-m)\bmod M}[f_k].\end{aligned}$$

Introducing additional notation:

$$H_k=\frac{e^{iu_krT+\psi_kT}}{1+iu_k}\qquad G_k=\frac{e^{iu_krT+\psi_k^AT}}{iu_k(1+iu_k)},$$

we obtain the final formula for computing the gradient:

$$\begin{aligned}\frac{\partial\hat{\tilde{z}}_T(x_j)}{\partial q_m} &= \frac{\Delta}{2\pi}e^{-ix_ju_{M-1}}T(e^{x_m}-1)\mathrm{DFT}_j\left[w_ke^{-ix_0\Delta k}H_{M-1-k}\right]\\ &+ Te^{x_m}(e^{iu_{M-1}x_m}\hat{\tilde{z}}_T(x_{(j-m_0-m)\bmod M})-\hat{\tilde{z}}_T(x_j))\\ &+ \frac{\Delta}{2\pi}e^{i(x_m-x_j)u_{M-1}}Te^{x_m}\mathrm{DFT}_{(j-m_0-m)\bmod M}\left[w_ke^{-ix_0\Delta k}G_{M-1-k}\right]\\ &- \frac{\Delta}{2\pi}e^{-ix_ju_{M-1}}Te^{x_m}\mathrm{DFT}_j\left[w_ke^{-ix_0\Delta k}G_{M-1-k}\right].\end{aligned}\tag{3.32}$$

The time values $\hat{\tilde{z}}_T$ have already been computed in all points of the grid when evaluating option prices, and the Fourier transforms of G_k do not need to be reevaluated at each step of the algorithm because G_k do not depend on q_i. Therefore, *compared to evaluating the functional alone, to compute the gradient of $\hat{J}_\alpha$ one only needs one additional fast Fourier transform per maturity date.* The complexity of evaluating the gradient using the above formula is thus only about 1.5 times higher than that of evaluating the functional itself (because evaluating option prices requires two Fourier transforms per maturity date). If the gradient was to be evaluated numerically, the complexity would typically be M times higher. The analytic formula (3.32) therefore allows to reduce the overall running time of the calibration algorithm from several hours to less than a minute on a standard PC.

3.5.3 Overview of the algorithm

Here is the final numerical algorithm, as implemented in the computer program `levycalibration`, used to run the tests of Section 3.6.

- Fix the prior using one of the methods described in Section 3.2. In the tests below, a user-specified prior was taken.
- Compute the weights of market option prices (Section 3.4) and estimate the noise level (Section 3.3.3).
- Use one of the a posteriori methods of Section 3.3.1 to compute an optimal regularization parameter α^* achieving trade-off between precision and stability. The optimal α^* is computed by bisection, minimizing $\tilde{J}_\alpha$ several times for different values of α with low precision. In the tests below we have always been able to choose α^* using the discrepancy principle (3.13) with $c_1 = 1.1$ and $c_2 = 1.3$, so there was no need to resort to the alternative scheme (3.19).
- Minimize $\tilde{J}_{\alpha^*}$ with high precision to find the regularized solution Q^*.

3.6 Numerical and empirical tests

Our numerical tests, performed using the `levycalibration` program, fall into two categories. First, to assess the accuracy and numerical stability of our method, we tested it on option prices produced by a known exponential-Lévy model (Section 3.6.1). We then applied our algorithm to real options data, using the prices of European options on different European stock market indices, provided by Thomson Financial® and studied the properties of Lévy measures, implied by market data.

3.6.1 Tests on simulated data

A compound Poisson example: Kou's model In the first series of tests, option prices were generated using Kou's jump diffusion model [62] with a diffusion part with volatility $\sigma_0 = 10\%$ and a Lévy density given by Equation (1.17).

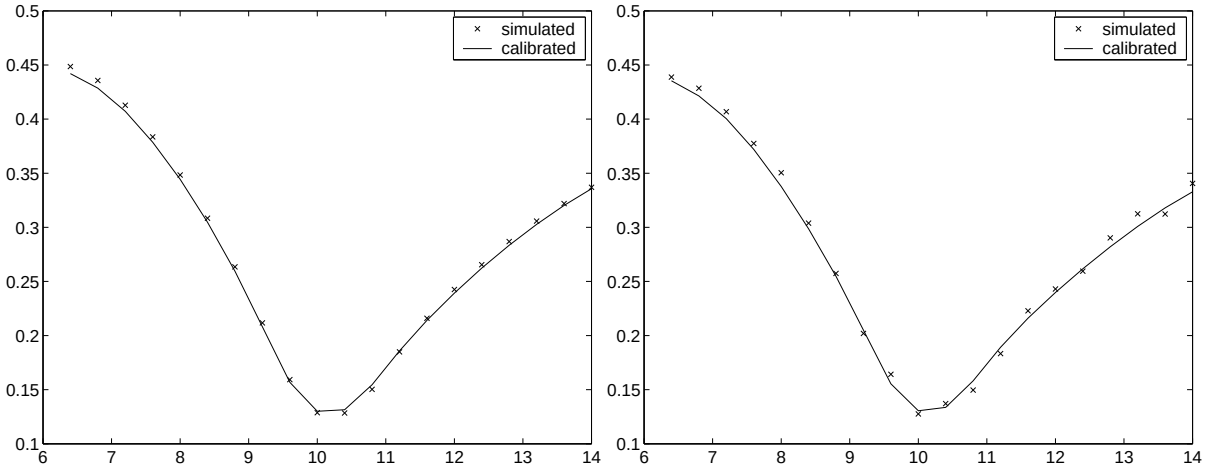

Figure 3.4: Calibration quality (implied volatilities as a function of strike prices) for Kou's jump diffusion model. Left: no perturbations. Right: 1% perturbations were added to the data

The Lévy density used in the tests was asymmetric with the left tail heavier than the right one ($\alpha_1 = 1/0.07$ and $\alpha_2 = 1/0.13$). The jump intensity λ was equal to 1 and the last parameter p was chosen such that the Lévy density is continuous at $x = 0$. The option prices were computed using the Fourier transform method of Section 1.4. The maturity of the options was 5 weeks and we used 21 equidistant strikes ranging from 6 to 14 (the spot being at 10). To capture tail behavior it is important to have strikes quite far in and out of the money. Merton's jump diffusion model [71] was used as prior.

After generating sets of call option prices from Kou's model, in the first test, the algorithm described in Section 3.5.3 was applied to these prices directly, and in the second test, to model the presence of data errors, random perturbations of order of 1% of implied volatility were added to the simulated prices. As can be observed from Figure 3.4, in both cases the accuracy of calibration at the level of implied volatilities is satisfying with only 21 options.

Figure 3.5 compares the non-parametric reconstruction of the Lévy density to the true Lévy density which, in this case, is known to be of the form (1.17). It can be observed that even in presence of data errors we retrieve successfully the main features of the true density with our non-parametric approach. The only region in which we observe a detectable error is near zero: very small jumps have a small impact on option prices. The gradient of option's time value given by Equation (3.32) vanishes at zero (i.e. for $m = m_0$) which means that the Lévy density

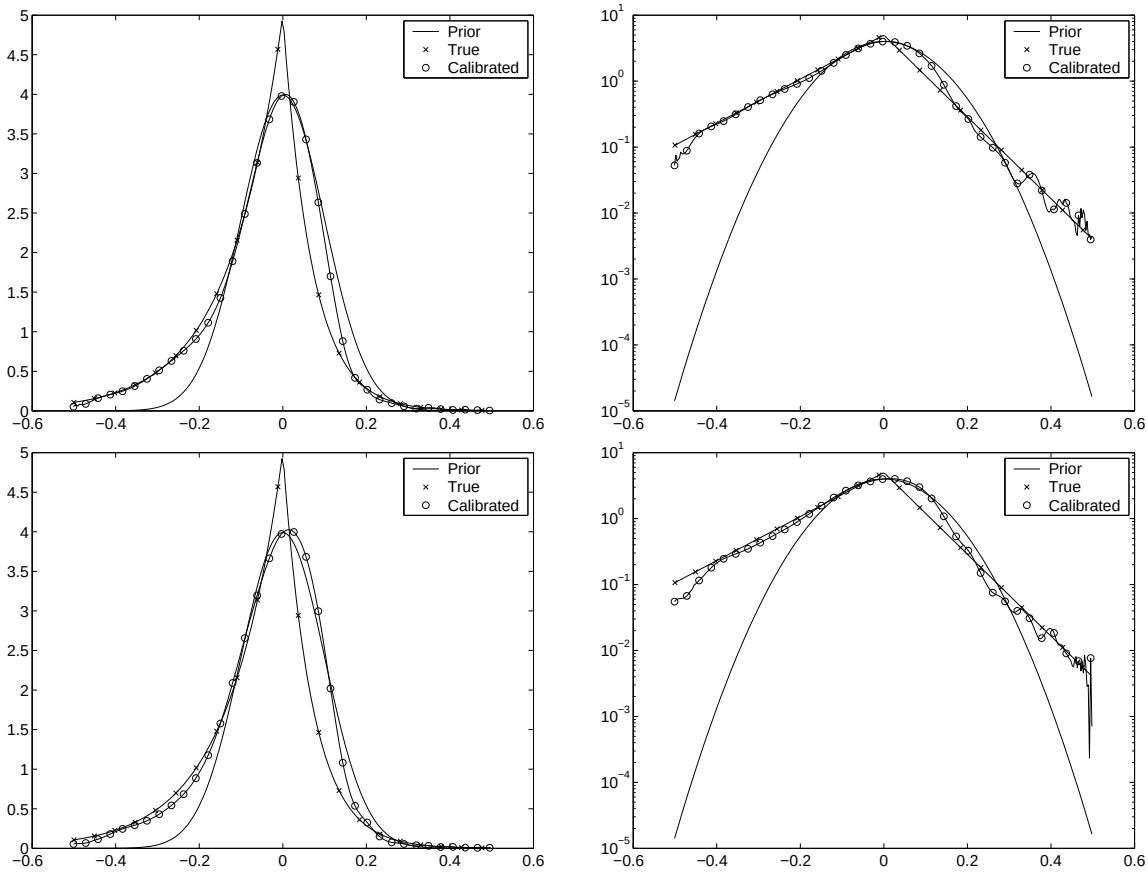

Figure 3.5: Lévy measure calibrated to option prices simulated from Kou's jump diffusion model with $\sigma_0 = 10\%$, on linear and logarithmic scale. Top: No perturbations were added. Bottom: 1% perturbations were added to the data.

at or near zero is determined exclusively by the prior: the intensity of small jumps cannot be retrieved accurately.

Figure 3.6 illustrates the redundancy of small jumps and diffusion: the two graphs were calibrated to the same prices and only differ in the diffusion coefficients. Comparing the two graphs shows that the algorithm compensates the error in the diffusion coefficient by adding jumps to the Lévy density so that, overall, the accuracy of calibration is maintained (the standard deviation in both graphs is 0.2%). The redundancy of small jumps and diffusion component has been pointed out by other authors in the context of statistical estimation on time series [15, 69]. Here we retrieve another version of this redundancy in a context of calibration to a cross sectional data set of options.

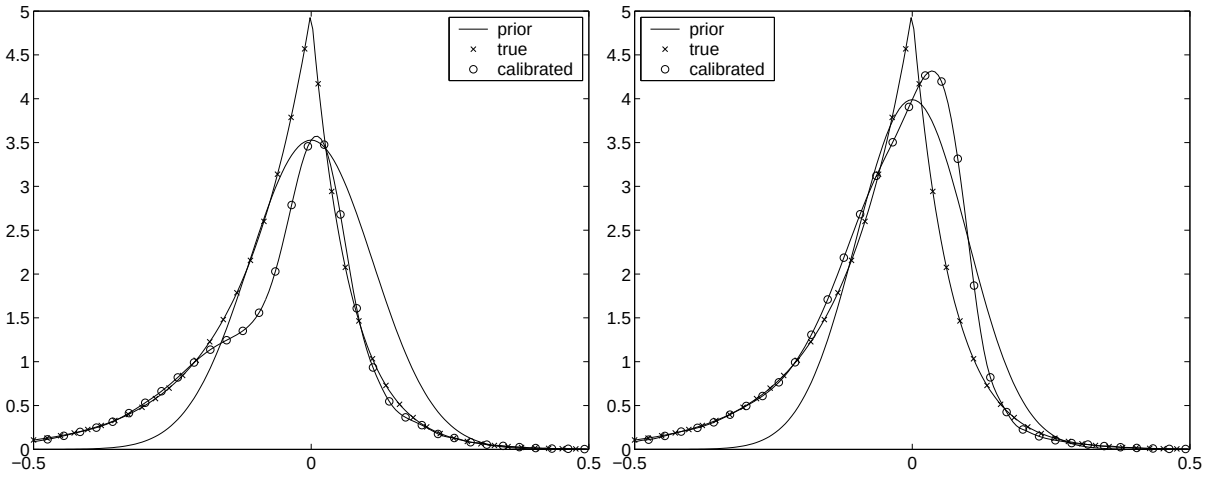

Figure 3.6: Lévy measure calibrated to option prices simulated from Kou's jump diffusion model with $\sigma_0 = 10\%$. Left: $\sigma = 10.5\% > \sigma_0$. Right: $\sigma = 9.5\% < \sigma_0$.

The stability of the algorithm with respect to initial conditions can be tested by altering the starting point of the optimization routine and examining the effect on the output. As illustrated in Figure 3.7, the entropy penalty removes the sensitivity observed in the non-linear least squares algorithm (see Figure 2.2 and Section 2.1). The only minor difference between the two calibrated measures is observed in the neighborhood of zero i.e. the region which, as remarked above, has little influence on option prices.

Variance gamma model In a second series of tests we examine how our method performs when applied to option prices generated by an infinite intensity process such as the variance gamma model. We assume that the user, ignoring that the data generating process has infinite jump intensity, chooses a (misspecified) prior which has a finite jump intensity (e.g. the Merton model).

We generated option prices for 21 equidistant strikes between 0.6 and 1.4 (the spot being at 1) using the variance gamma model [67] with no diffusion component and applied the calibration algorithm using a Merton jump diffusion model as prior to these prices. The parameters of the variance gamma process (cf Equation (1.18)) were $\sigma = 0.3$, $\theta = -0.2$ and $k = 0.04$. All the options were maturing in five weeks. The left graph in Figure 3.8 shows that even though the prior is misspecified, the calibrated model reproduced the simulated implied volatilities with

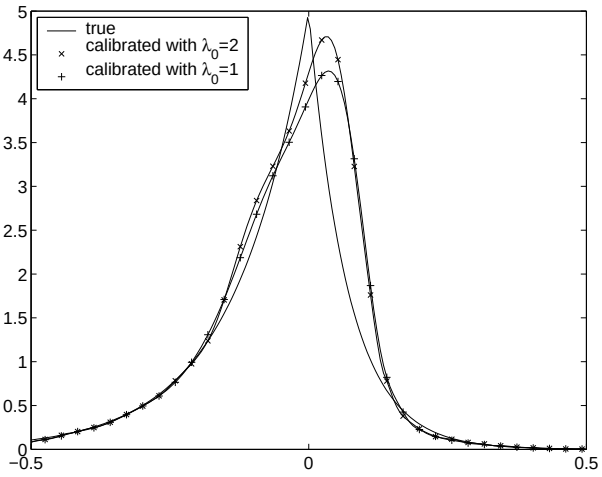

Figure 3.7: Levy densities calibrated to option prices generated from Kou model, using two different initial measures with intensities $\lambda = 1$ and $\lambda = 2$.

good precision (for both values of σ_0 that we took, the standard deviation is less than 0.5% in implied volatility units).

The calibrated Lévy densities for two different values of prior diffusion volatility σ_0 are shown in the right graph of Figure 3.8: a smaller value of the volatility parameter leads to a greater intensity of small jumps.

Here we observe once again the redundancy of volatility and small jumps, this time in an infinite intensity context. More precisely this example shows that call option prices generated from an infinite intensity exponential Lévy model can be reproduced with arbitrary precision using a jump-diffusion with finite jump intensity. This leads us to conclude that given a finite (but realistic) number of option prices, the shape of implied volatility skews/smiles does not allow to distinguish infinite activity models like variance gamma from jump diffusions.

3.6.2 Empirical properties of implied Lévy measures

To illustrate our calibration method we have applied it to a data set spanning the period from 1999 to 2001 and containing daily prices of options on the DAX (German stock market index) for a range of strikes and maturities.

Figure 3.10 shows the calibration quality for three different maturities. Note that here each maturity has been calibrated separately (three different exponential Lévy models). These

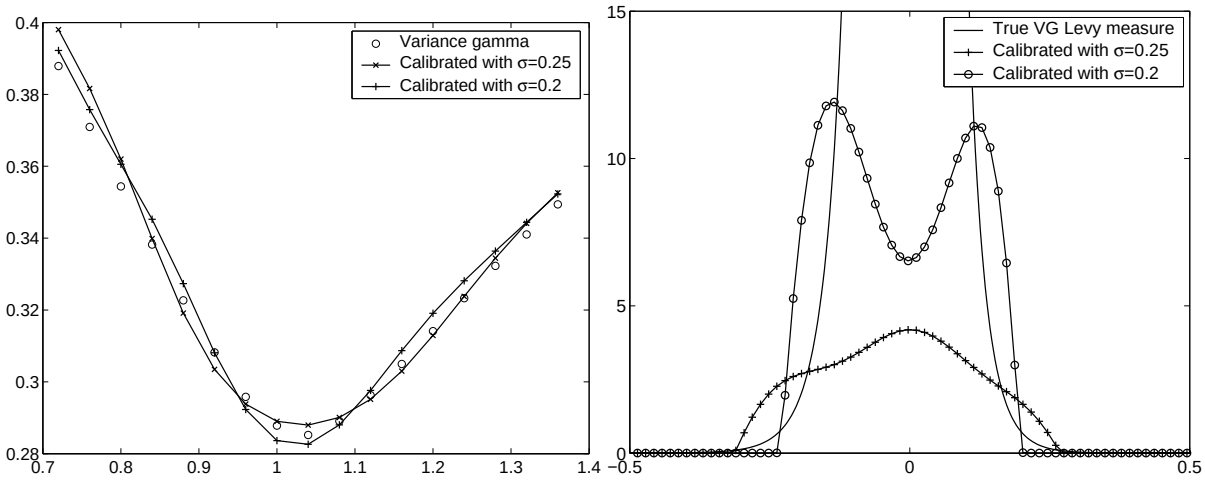

Figure 3.8: Calibration of options simulated from a variance gamma model using a finite intensity prior with different values of diffusion volatility σ_0. Left: calibration quality (implied volatilities as a function of strike prices). Right: Calibrated Lévy measures. Increasing the diffusion coefficient decreases the intensity of small jumps in the calibrated measure.

graphs show that using an exponential Lévy model one can calibrate with high precision the prices of a set of options with common maturity. This conclusion contradicts the findings of Medvedev and Scaillet [70] who observe that "jumps in returns do not help explaining the slope of implied volatilities at the money". It is important to note, however, that Medvedev and Scaillet used short term options on S&P 500 index whereas our analysis is based on DAX index options.

Figure 3.9, left graph, illustrates the typical shape of risk neutral Lévy densities obtained from our data set. It is readily seen that the Lévy measures obtained are far from being symmetric: the distribution of jump sizes is highly skewed toward negative values. Figure 3.11 shows the same result across calendar time, showing that this asymmetry persists across time. This effect also depends on the maturity of options in question: for longer maturities (see Figure 3.11, right graph) the support of the Lévy measure extends further to the left. Most of the densities we obtained are bimodal with one mode corresponding to small jumps that are hard to separate from a diffusion component and another clearly distinguishable mode corresponding to a large negative jump that can reflect the market participants' "fear of crash" [11].

The logarithmic scale in the right graph of Figure 3.9 allows the tails to be seen more clearly.

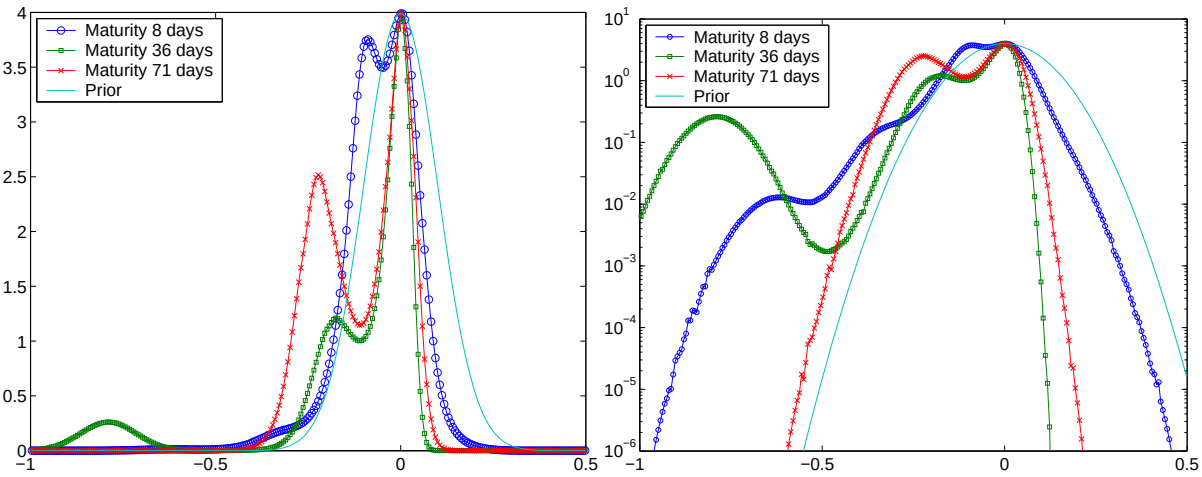

Figure 3.9: Lévy measures calibrated to DAX options on the same calendar date for three different maturities, linear and logarithmic scale.

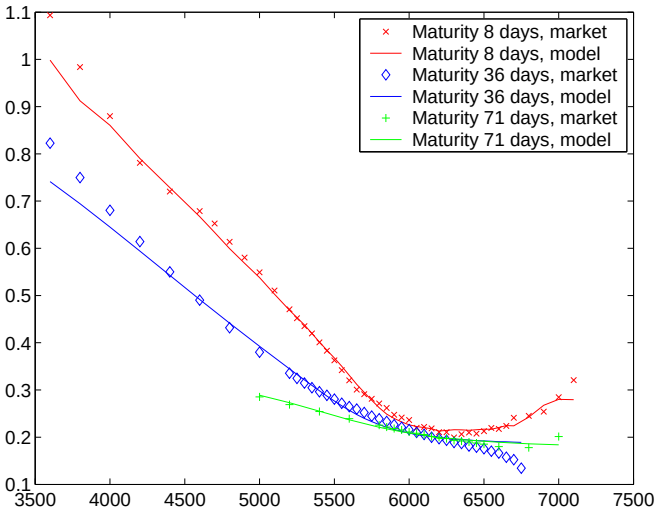

Figure 3.10: Calibration quality for different maturities: market implied volatilities for DAX options against model implied volatilities. Each maturity has been calibrated separately.

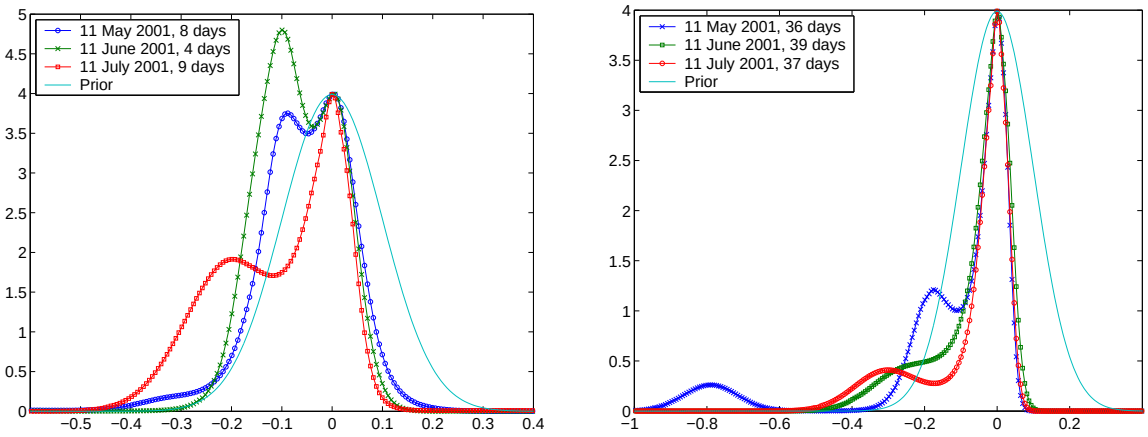

Figure 3.11: Results of calibration at different dates for shortest (left) and second shortest (right) maturity. DAX index options.

Recall that the prior density is Gaussian, which shows up as a symmetric parabola on log scales.

The *area* under the curves shown here is to be interpreted as the (risk neutral) jump intensity. While the shape of the curve does vary slightly across calendar time, the intensity stays surprisingly stable: its numerical value is empirically found to be $\lambda \simeq 1$, which means around one jump a year. Of course note that this is the *risk neutral* intensity: jump intensities are *not* invariant under equivalent change of measures. Moreover this illustrates that a small intensity of jumps λ can be sufficient for explaining the shape of the implied volatility skew for small maturities. Therefore the motivation of infinite activity processes from the point of view of option pricing does not seem clear to us. On the other hand, from the modelling perspective infinite intensity processes do present a particular interest since they do not have a diffusion component and therefore do not suffer from the problem of redundancy of small jumps and diffusion, described above, leading to more parsimonious and identifiable models.

3.6.3 Testing time homogeneity

While the literature on jump processes in finance has focused on time homogeneous (Lévy) models, practitioners have tended to use time dependent jump or volatility parameters. Despite the fact that, as our non-parametric analysis and several empirical studies by other authors [21, 67] have shown, Lévy processes reproduce the implied volatility smile for a single maturity

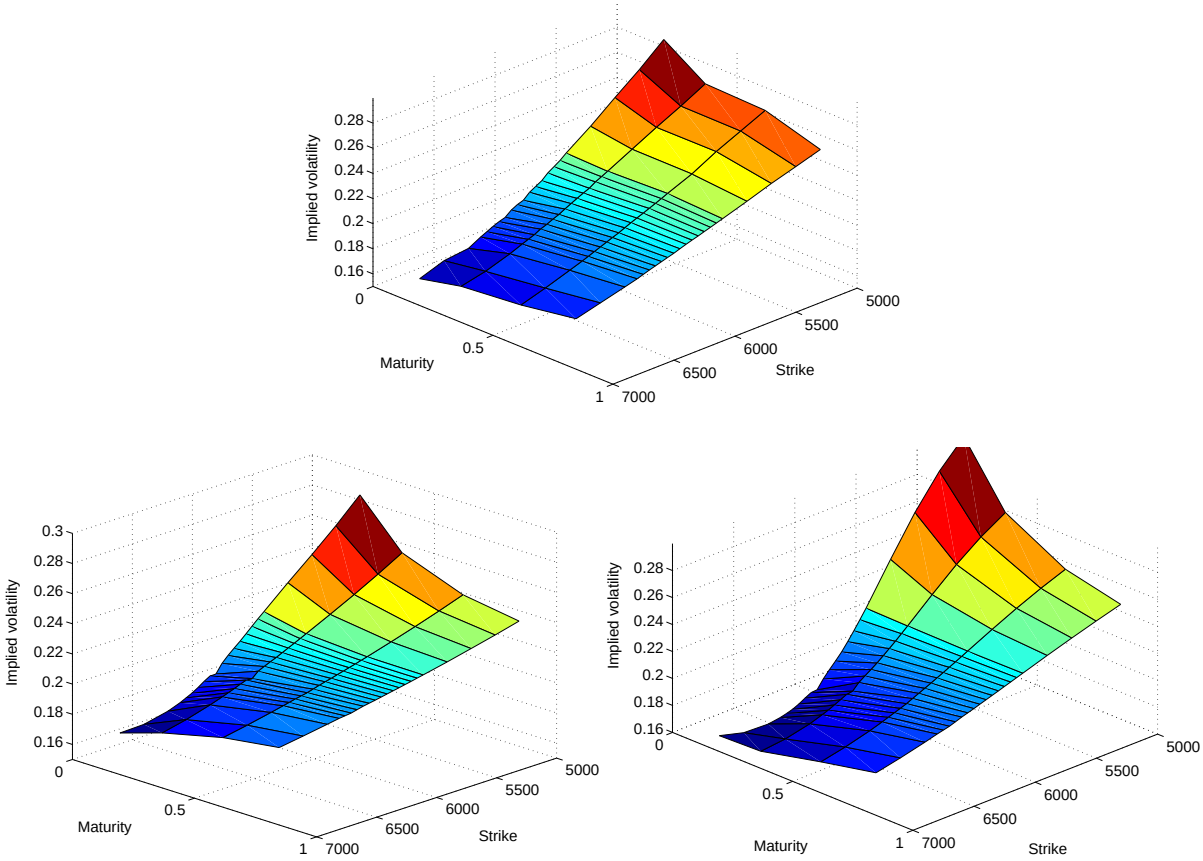

Figure 3.12: Top: Market implied volatility surface. Bottom left: implied volatility surface in an exponential Lévy model, calibrated to market prices of the first maturity. Bottom right: implied volatility surface in an exponential Lévy model, calibrated to market prices of the last maturity.

quite well, when it comes to calibrating several maturities at the same time, the calibration by Lévy processes becomes much less precise. This is clearly seen from the three graphs of Figure 3.12. The top graph shows the market implied volatilities for four maturities and different strikes. The bottom left graphs depicts implied volatilities, computed in an exponential Lévy model calibrated using our nonparametric algorithm to the first maturity present in the market data. One can see that while the calibration quality is acceptable for the first maturity, it quickly deteriorates as the time to maturity increases: the smile in an exponential Lévy model flattens too fast. The same effect can be observed in the bottom right graph: here, the model was calibrated to the last maturity, present in the data. As a result, the calibration quality is poor for the first maturity: the smile in an exponential Lévy model is more pronounced and its shape does not resemble that of the market.

It is difficult to calibrate an exponential Lévy model to options of several maturities because due to independence and stationarity of their increments, Lévy processes have a very rigid term structure of cumulants. In particular, the skewness of a Lévy process is proportional to the inverse square root of time and the excess kurtosis is inversely proportional to time [68]. A number of empirical studies have compared the term structure of skewness and kurtosis implied in market option prices to the skewness and kurtosis of Lévy processes. Bates [13], after an empirical study of implicit kurtosis in $\$/DM$ exchange rate options concludes that "while implicit excess kurtosis does tend to increase as option maturity shrinks, ..., the magnitude of maturity effects is not as large as predicted [by a Lévy model]". For stock index options, Madan and Konikov [68] report even more surprising results: both implied skewness and kurtosis actually decrease as the length of the holding period becomes smaller. It should be mentioned, however, that implied skewness/kurtosis cannot be computed from a finite number of option prices with high precision.

Our non-parametric approach allows to investigate time homogeneity by calibrating the Lévy measure separately to various option maturities. Figure 3.9 shows Lévy measures obtained by running the algorithm for options of different maturity. The hypothesis of time homogeneity would imply that all the curves are the same, which is apparently not the case here. However, computing the areas under the curves yields similar jump intensities across maturities: this result can be interpreted by saying that the risk neutral jump intensity is relatively stable through time while the shape of the (normalized) jump size density can actually change. The

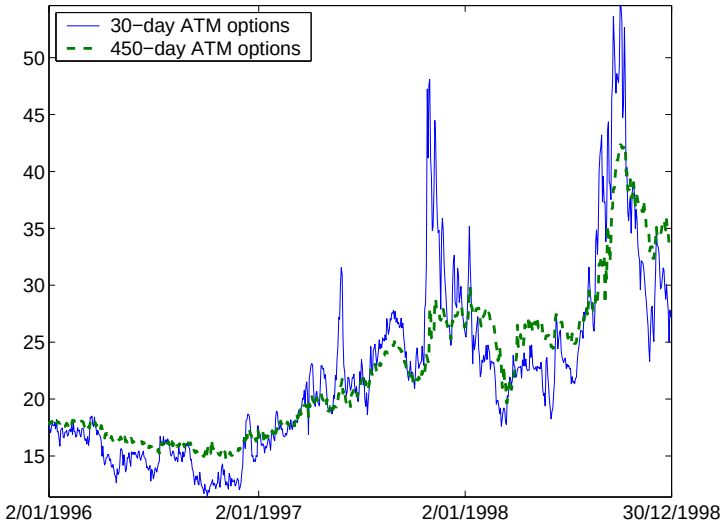

Figure 3.13: Implied volatility of at the money European options on CAC40 index.

type of time dependence is therefore more complicated than simply a time dependent intensity.

A second major difficulty arising while trying to calibrate an exponential Lévy model is the time evolution of the smile. Exponential Lévy models belong to the class of so called "sticky moneyness" models, meaning that in an exponential Lévy model, the implied volatility of an option with given moneyness (strike price to spot ratio) does not depend on time. This can be seen from the following simple argument. In an exponential Lévy model Q, the implied volatility σ of a call option with moneyness m, expiring in τ years, satisfies:

$$e^{-r\tau}E^Q[(S_t e^{r\tau+X_\tau} - mS_t)^+|\mathcal{F}_t] = e^{-r\tau}E[(S_t e^{r\tau+\sigma W_\tau - \frac{\sigma^2}{2}\tau} - mS_t)^+|\mathcal{F}_t]$$

Due to the independent increments property, S_t cancels out and we obtain an equation for the implied volatility σ which does not contain t or S_t. Therefore, in an exp-Lévy model this implied volatility does not depend on date t or stock price S_t. This means that once the smile has been calibrated for a given date t, its shape is fixed for all future dates. Whether or not this is true in real markets can be tested in a model-free way by looking at the implied volatility of at the money options with the same maturity for different dates. Figure 3.13 depicts the behavior of implied volatility of two at the money options on the CAC40 index, expiring in 30 and 450 days. Since the maturities of available options are different for different dates, to obtain the implied volatility of an option with fixed maturity T for each date, we have taken two maturities, present in the data, closest to T from above and below: $T_1 \leq T$ and $T_2 > T$.

The implied volatility $\Sigma(T)$ of the hypothetical option with maturity T was then interpolated using the following formula:

$$\Sigma^2(T) = \Sigma^2(T_1)\frac{T_2 - T}{T_1 - T} + \Sigma^2(T_2)\frac{T - T_1}{T_2 - T_1}.$$

As we have seen, in an exponential Lévy model the implied volatility of an option which is at the money and has fixed maturity must not depend on time or stock price. Figure 3.13 shows that in reality this is not so: both graphs are rapidly varying random functions.

This simple test shows that real markets do not have the "sticky moneyness" property: arrival of new information can alter the form of the smile. The exponential Lévy models are therefore "not random enough" to account for the time evolution of the smile. Moreover, models based on additive processes, that is, time-inhomogeneous processes with independent increments, although they perform well in calibrating the term structure of implied volatilities for a given date [27], are not likely to describe the time evolution of the smile correctly since in these models the future form of the smile is still a deterministic function of its present shape [27]. To describe the time evolution of the smile in a consistent manner, one may need to introduce additional stochastic factors (e.g. stochastic volatility) [9, 10, 22].

One of the important conclusions of this chapter is that at least in the domain of stock index options, to which we have applied our tests, exponential Lévy models, although they provide a considerable improvement compared to Black-Scholes model, do not allow to calibrate the term structure of implied volatilities with sufficient precision. Our calibration algorithm therefore should not be seen as a method to find an exponential Lévy model consistent with all available market data but rather as a way to construct a Lévy measure, implied by option prices of a particular maturity. It can therefore be extended to other models where jumps are represented, in a nonparametric way, by a Lévy measure. Its many possible applications include, among others,

- Interpolation of option prices for a single maturity;
- Calibration, using short-maturity data, of the jump part of hybrid models, including both jumps and stochastic volatility;
- Separate calibration of Lévy densities for different maturities in order to determine the correct pattern of time dependence of Lévy measure for additive processes.

Part II

Multidimensional modelling with Lévy processes

Chapter 4

Characterization of dependence of multidimensional Lévy processes

4.1 Introduction to dependence modelling

Many financial applications require a multidimensional model with jumps, taking into account the dependence between components. However, such models are more difficult to construct than one-dimensional ones and the applications continue to be dominated by (geometric) Brownian motion.

A simple method to introduce jumps into a multidimensional model is to take a multivariate Brownian motion and time change it with a one-dimensional increasing Lévy process. This approach, advocated in [36, 79], allows to construct multidimensional versions of many popular one-dimensional models, including variance gamma, normal inverse Gaussian and generalized hyperbolic process. The principal advantage of this method is its simplicity and analytic tractability; in particular, processes of this type are easy to simulate. However, the range of dependence patterns that one can obtain using this approach is quite limited (for instance, independence is not included), and all components must follow the same parametric model (e.g., either all of the components are variance gamma or all of the components are normal inverse Gaussian etc.)

To be more specific, suppose that two stock price processes $\{S_t^1\}_{t\geq 0}$ and $\{S_t^2\}_{t\geq 0}$ are mod-

elled as follows:

$$S_t^1 = \exp(X_t^1), \quad X_t^1 = B^1(Z_t) + \mu_1 Z_t,$$
$$S_t^2 = \exp(X_t^2), \quad X_t^2 = B^2(Z_t) + \mu_2 Z_t,$$

where B^1 and B^2 are two components of a planar Brownian motion, with variances σ_1^2 and σ_2^2 and correlation coefficient ρ, and $\{Z_t\}_{t\geq 0}$ is the stochastic time change (an increasing Lévy process). The correlation of returns, $\rho(X_t^1, X_t^2)$, can be computed by conditioning with respect to Z_t:

$$\rho(X_t^1, X_t^2) = \frac{\sigma_1\sigma_2\rho E[Z_t] + \mu_1\mu_2 \operatorname{Var} Z_t}{(\sigma_1^2 E[Z_t] + \mu_1^2 \operatorname{Var} Z_t)^{1/2}(\sigma_2^2 E[Z_t] + \mu_2^2 \operatorname{Var} Z_t)^{1/2}}.$$

In the completely symmetric case ($\mu_1 = \mu_2 = 0$) and in this case only $\rho(X_t^1, X_t^2) = \rho$: the correlation of returns equals the correlation of Brownian motions that are being subordinated. However, the distributions of real stocks are skewed and in the skewed case the correlation of returns will be different from the correlation of Brownian motions that we put into the model. Even if the Brownian motions are independent, the covariance of returns is equal to $\mu_1\mu_2 \operatorname{Var} Z_t$ and if the distributions of stocks are not symmetric, they are correlated.

In the symmetric case, if Brownian motions are independent, the two stocks are decorrelated but not independent. Since the components of the Brownian motion are time changed with the same subordinator, large jumps in the two stocks (that correspond to large jumps of the subordinator) will tend to arrive together, which means that absolute values of returns will be correlated. If $\mu_1 = \mu_2 = 0$ and $\rho = 0$ then the covariance of squares of returns is

$$\operatorname{Cov}((X_t^1)^2, (X_t^2)^2) = \sigma_1\sigma_2 \operatorname{Cov}((B^1(Z_t))^2, (B^2(Z_t))^2) = \sigma_1\sigma_2 \operatorname{Var} Z_t,$$

which means that squares of returns are correlated unless the subordinator Z_t is deterministic. This phenomenon can lead to mispricing and errors in evaluation of risk measures.

In finite activity models, a more accurate modelling of dependence may be achieved by specifying directly the dependence of individual jumps in one-dimensional compound Poisson processes (see [65]). This approach is useful in presence of few sources of jump risk (e.g., when all components jump at the same time) because in this case it allows to achieve a precise description of dependence within a simple model. Suppose that we want to improve a d-dimensional Black-Scholes model by allowing for "market crashes." The dates of market crashes can be modelled

as jump times of a standard Poisson process $\{N_t\}_{t\geq 0}$. This leads us to the following model for the log-price processes of d assets:

$$X_t^i = \mu^i t + B_t^i + \sum_{j=1}^{N_t} Y_j^i, \quad i = 1 \dots d,$$

where (B_t) is a d-dimensional Brownian motion with covariance matrix Σ, and $\{Y_j\}_{j=1}^{\infty}$ are i.i.d. d-dimensional random vectors which determine the sizes of jumps in individual assets during a market crash. This model contains only one driving Poisson shock because we only account for jump risk of one type (global market crash affecting all assets). To construct a parametric model, we need to specify the distribution of jumps in individual assets during a crash (distribution of Y_*^i for all i) and the dependence between jumps in assets. If we make a simplifying assumption that $\{Y_j^i\}_{i=1}^d$ are Gaussian random vectors, then we need to specify their covariance matrix Σ' and the mean vector m, thus obtaining a multivariate version of Merton's model [71].

If the jumps are not Gaussian, we must specify the distribution of jumps in each component and the copula[1] describing their dependence. The model is thus completely specified by a covariance matrix Σ, d jump size distributions, a d-dimensional copula C and a jump intensity parameter λ. However, sometimes it is necessary to have several independent shocks to account for events that affect individual companies or individual sectors rather than the entire market. In this case we need to introduce several driving Poisson processes into the model, which now takes the following form:

$$X_t^i = \mu^i t + B_t^i + \sum_{k=1}^{M} \sum_{j=1}^{N_t^k} Y_{j,k}^i, \quad i = 1 \dots d,$$

where $N_t^1, \dots, N_t^M$ are Poisson processes driving M independent shocks and $Y_{j,k}^i$ is the size of jump in i-th component after j-th shock of type k. The vectors $\{Y_{j,k}^i\}_{i=1}^d$ for different j and/or k are independent. To define a parametric model completely, one must specify a one-dimensional distribution for each component for each shock type — because different shocks influence the same stock in different ways — and one d-dimensional copula for each shock type. This adds up to $M \times d$ one-dimensional distributions and M one-dimensional copulas. How many different

[1] For an introduction to copulas see [76] — this monograph treats mostly the bivariate case — and [56] for the multivariate case.

shocks do we need to describe sufficiently rich dependence structures? The answer depends on the particular problem, but to describe *all* possible dependence structures, such that the d-dimensional process remains a Lévy process of compound Poisson type, one needs a total of $2^M - 1$ shocks (M shocks that affect only one stock, $\frac{M(M-1)}{2}$ shocks that affect two stocks etc., adding up to $2^M - 1$). It is clear that as the dimension of the problem grows, this kind of modelling quickly becomes infeasible. Not only the number of parameters grows exponentially, but also, when the number of shocks is greater than one, one cannot specify directly the laws of components because the laws of jumps must be given separately for each shock.

Comparison of advantages and drawbacks of these two methods leads to an understanding of the required properties of a multidimensional modelling approach for Lévy processes. In general, such an approach must satisfy the following conditions:

- One should be able to choose any one-dimensional Lévy process for each of the components. In particular, it should be possible to couple a compound Poisson process with a process which has infinite jump intensity. This is particularly important for financial applications because information about margins and about dependence may come from different sources. For instance, for pricing basket options the market practice is to estimate the correlations (dependence) from historical data while using implied volatilities, computed from the prices of traded options.

- The range of possible dependence structures should include complete dependence and complete independence with a "smooth" transition between these two extremes.

- It should be possible to model dependence in a parametric fashion (e.g., not through the entire multidimensional Lévy measure). A parsimonious description of dependence is especially important because one typically does not have enough information about the dependence structure to estimate many parameters or proceed with a nonparametric approach.

To implement this program, we suggest to model separately the dependence structure of a Lévy process and the behavior of its components (margins). It has long been known that the dependence structure of a random vector on $\mathbb{R}^d$ can be disentangled from its margins via the notion of copula. Since the law of a Lévy process is completely determined by its

distribution at time t for any fixed $t > 0$, the dependence structure of a multidimensional Lévy process $X \equiv \{X_t^i\}_{t\geq 0}^{i=1\ldots d}$ can be parametrized by the copula C_t of the random vector $\{X_t^i\}^{i=1\ldots d}$ for some $t > 0$. However, this approach has a number of drawbacks. First, for given infinitely divisible one-dimensional laws $\mu_t^1, \ldots, \mu_t^d$, it is unclear, which copulas C_t will yield a d-dimensional infinitely divisible law. Second, the copula C_t may depend on t and C_s for some $s \neq t$ cannot in general be computed from C_t alone; to compute it one also needs to know the marginal distributions at time t and at time s. We will now construct an explicit example of a Lévy process with nonconstant copula.

Example 4.1. Let $Z := \{Z_t\}_{t\geq 0}$ be a 2-dimensional Cauchy process, that is, a Lévy process with characteristic triplet $(0, \nu^Z, 0)$, where ν^Z has a density, also denoted by ν^Z, given by

$$\nu^Z(x, y) = \frac{1}{(x^2 + y^2)^{3/2}}.$$

The probability distribution of Z_t for every $t > 0$ has a density

$$p_t^Z(x, y) = \frac{1}{2\pi} \frac{t}{\{(x^2 + y^2)^2 + t^2\}^{3/2}}.$$

The copula of Z_t does not depend on time and can be computed explicitly:

$$C^Z(u, v) = -\frac{1}{4} + \frac{u}{2} + \frac{v}{2} + \frac{1}{2\pi} \arctan \left\{ \frac{\tan \pi(u - \frac{1}{2}) \tan \pi(v - \frac{1}{2})}{\sqrt{1 + \tan^2 \pi(u - \frac{1}{2}) + \tan^2 \pi(v - \frac{1}{2})}} \right\},$$

which is clearly different from the independence copula $C^\perp(u, v) = uv$.

Let $W := \{W_t\}_{t\geq 0}$ be a standard planar Brownian motion, independent from Z. Since the components of W are independent, the copula of W_t for each $t > 0$ is the independence copula $C^\perp(u, v) = uv$. For every t, let $X_t = Z_t + W_t$. $\{X_t\}_{t\geq 0}$ is a Lévy process with characteristic triplet $(Id_2, \nu^Z, 0)$. Since Z is a 1-stable process and W is $\frac{1}{2}$-stable, $\frac{X_t}{t} \stackrel{d}{=} Z_1 + W_{1/t}$ and $\frac{X_t}{\sqrt{t}} \stackrel{d}{=} Z_{\sqrt{t}} + W_1$. Therefore, the *random variable* $\frac{X_t}{t}$ is infinitely divisible with characteristic triplet $(\frac{1}{t} Id_2, 0, \nu^Z)$ and the random variable $\frac{X_t}{\sqrt{t}}$ is infinitely divisible with characteristic triplet $(Id_2, 0, \sqrt{t}\nu^Z)$. From Proposition 1.7 it follows that $\frac{X_t}{t} \xrightarrow[t\to\infty]{d} Z_1$ and $\frac{X_t}{\sqrt{t}} \xrightarrow[t\to 0]{d} W_1$. Since the copula is invariant with respect to transformations of margins by strictly increasing functions, X_t, $\frac{X_t}{\sqrt{t}}$ and $\frac{X_t}{t}$ have the same copula. Therefore, Theorem 2.1 in [64] implies that the copula C_t of X_t has the following properties

$$\forall u, v, \quad C_t(u, v) \to C^Z(u, v) \text{ as } t \to \infty,$$

$$\forall u, v, \quad C_t(u, v) \to C^\perp(u, v) \text{ as } t \to 0.$$

Since $C^Z(u,v)$ is different from $C^\perp(u,v)$, we conclude that C_t is not constant over time.

Every Lévy process $X := \{X_t\}_{t\geq 0}$ is described in a time-dependent fashion by its characteristic triplet (A,ν,γ). It seems therefore natural to describe the dependence between components of X also in terms of its characteristic triplet. Since the continuous martingale component of X is completely described by the covariance matrix A and is independent from the jump part, we will focus on the dependence of the jump part of X, that is, we only consider Lévy processes with $A = 0$. For such a process, separate modelling of margins and dependence will be achieved by introducing Lévy copulas, which play the same role for Lévy measures as copulas for probability measures. After showing how basic dependence patterns for Lévy processes are expressed in terms of their characteristic triplets in Section 4.2 and introducing the necessary notation and definitions in Section 4.3, first, in Section 4.4 we treat the conceptually simpler case of Lévy processes, admitting only positive jumps in every component or, equivalently, having Lévy measures supported by $[0,\infty)^d$. Lévy copulas for general Lévy processes, introduced in a joint work of the present author with Jan Kallsen [59], are discussed in Sections 4.5 and 4.6 of this chapter.

4.2 Dependence concepts for multidimensional Lévy processes

In this section, $X := \{X_t^i\}_{t\geq 0}^{i=1,\dots,d}$ denotes a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A,ν,γ). For $I \subset \{1,\dots,d\}$ we define $I^c := \{1,\dots,d\} \setminus I$ and $|I|$ denotes the number of elements in I. We start by defining the *margins* of a Lévy process.

Definition 4.1. Let $I \subset \{1,\dots,d\}$ nonempty. The I-margin of X is the Lévy process $X^I := \{X_t^i\}_{t\geq 0}^{i\in I}$.

The following lemma explains that the Lévy measure of X^I only depends on the Lévy measure of X and shows how it can be computed.

Lemma 4.1 (Marginal Lévy measures). *Let $I \subset \{1,\dots,d\}$ nonempty. Then the Lévy process X^I has Lévy measure ν^I given by*

$$\nu^I(B) = \nu(\{x \in \mathbb{R}^d : (x_i)_{i\in I} \in B\}), \quad \forall B \in \mathcal{B}(\mathbb{R}^{|I|} \setminus \{0\}). \tag{4.1}$$

Proof. This lemma is a direct consequence of Proposition 11.10 in [87]. □

In view of the above lemma, for a given Lévy measure ν we will refer to the Lévy measure ν^I defined by Equation (4.5) as the I-margin of ν. To simplify notation, when $I = \{k\}$ for some k, the I-margin of ν will be denoted by ν_k and called simply k-th margin of ν.

Next we would like to characterize the independence of Lévy processes in terms of their characteristic triplets.

Lemma 4.2. *The components $X^1, \ldots, X^d$ of an $\mathbb{R}^d$-valued Lévy process X are independent if and only if their continuous martingale parts are independent and the Lévy measure ν is supported by the coordinate axes. ν is then given by*

$$\nu(B) = \sum_{i=1}^{d} \nu_i(B_i) \quad \forall B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}), \tag{4.2}$$

where for every i, ν_i denotes the i-th margin of ν and

$$B_i = \{x \in \mathbb{R} : (\underbrace{0, \ldots, 0}_{i-1 \text{ times}}, x, 0, \ldots, 0) \in B\}.$$

Proof. Since the continuous martingale part and the jump part of X are independent, we can assume without loss of generality that X has no continuous martingale part, that is, its characteristic triplet is given by $(0, \nu, \gamma)$.

The "if" part. Suppose ν is supported by the coordinate axes. Then necessarily for every $B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, $\nu(B) = \sum_{i=1}^d \tilde{\nu}_i(B_i)$ with some measures $\tilde{\nu}_i$, and Lemma 4.1 show that these measures coincide with the margins of ν: $\tilde{\nu}_i = \nu_i \ \forall i$. Using the Lévy-Khintchine formula for the process X, we obtain:

$$\begin{aligned} E[e^{i\langle u, X_t\rangle}] &= \exp t\{i\langle \gamma, u\rangle + \int_{\mathbb{R}^d \setminus \{0\}} (e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle 1_{|x|\leq 1})\nu(dx)\} \\ &= \exp t \sum_{k=1}^{d} \{i\gamma_k u_k + \int_{\mathbb{R}\setminus\{0\}} (e^{iu_k x_k} - 1 - iu_k x_k 1_{|x_k|\leq 1})\nu_k(dx_k)\} = \prod_{k=1}^{d} E[e^{iu_k X_t^k}], \end{aligned}$$

which shows that the components of X are independent Lévy processes.

The "only if" part. Define a measure $\tilde{\nu}$ on $\mathbb{R}^d \setminus \{0\}$ by $\tilde{\nu}(B) = \sum_{i=1}^d \nu_i(B_i)$, where ν_i is the i-th marginal Lévy measure of X and B_i is as above. It is straightforward to check that $\tilde{\nu}$ is a Lévy measure. Since the components of X are independent, applying the Lévy-Khintchine formula to each component of X, we find:

$$E[e^{i\langle u, X_t\rangle}] = \exp t\{i\langle \gamma, u\rangle + \int_{\mathbb{R}^d \setminus \{0\}} (e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle 1_{|x|\leq 1})\tilde{\nu}(dx)\}.$$

Now from the uniqueness of Lévy-Khintchine representation we conclude that $\tilde{\nu}$ is the Lévy measure of X. □

The complete dependence of Lévy processes is a new notion that is worth being discussed in detail. First, the following definition is in order.

Definition 4.2. A subset S of $\mathbb{R}^d$ is called *ordered* if, for any two vectors $v, u \in S$, either $v_k \leq u_k$, $k = 1, \ldots, d$ or $v_k \geq u_k$, $k = 1, \ldots, d$. S is called *strictly ordered* if, for any two different vectors $v, u \in S$, either $v_k < u_k$, $k = 1, \ldots, d$ or $v_k > u_k$, $k = 1, \ldots, d$.

We recall that random variables $Y_1, \ldots, Y_d$ are said to be completely dependent or comonotonic if there exists a strictly ordered set $S \subset \mathbb{R}^d$ such that $(Y_1, \ldots, Y_d) \in S$ with probability 1. However, saying that the components of a Lévy process are completely dependent only if they are completely dependent for every fixed time is too restrictive; the components of a Lévy process can be completely dependent as processes without being completely dependent as random variables for every fixed time. The following example clarifies this point.

Example 4.2 *(Dynamic complete dependence for Lévy processes).* Let $\{X_t\}_{t\geq 0}$ be a Lévy process with characteristic triplet (A, ν, γ) such that $A = 0$ and $\gamma = 0$ and let $\{Y_t\}_{t\geq 0}$ be a Lévy process, constructed from the jumps of X: $Y_t = \sum_{s\leq t} \Delta X_s^3$. From the dynamic point of view X and Y are completely dependent, because the trajectory of any one of them can be reconstructed from the trajectory of the other. However, the copula of X_t and Y_t is not that of complete dependence because Y_t is not a deterministic function of X_t. Indeed, if X is a compound Poisson process having jumps of size 1 and 2 and $X_t = 3$ for some t, this may either mean that X has three jumps of size 1 in the interval $[0, t]$, and then $Y_t = 3$, or that X has one jump of size 1 and one jump of size 2, and then $Y_t = 9$.

This example motivates the following definition. In this definition and below,

$$K := \{x \in \mathbb{R}^d : \operatorname{sgn} x_1 = \cdots = \operatorname{sgn} x_d\}. \tag{4.3}$$

Definition 4.3. Let X be a $\mathbb{R}^d$-valued Lévy process. Its jumps are said to be *completely dependent* or *comonotonic* if there exists a strictly ordered subset $S \subset K$ such that $\Delta X_t := X_t - X_{t-} \in S$, $t \geq 0$ (except for a set of paths having zero probability).

Clearly, an element of a strictly ordered set is completely determined by one coordinate only. Therefore, if the jumps of a Lévy process are completely dependent, the jumps of all components can be determined from the jumps of any single component. If the Lévy process has no continuous martingale part, then the trajectories of all components can be determined from the trajectory of any one component, which indicates that Definition 4.3 is a reasonable dynamic notion of complete dependence for Lévy processes. The condition $\Delta X_t \in K$ means that if the components of a Lévy process are comonotonic, they always jump in the same direction.

For any $\mathbb{R}^d$-valued Lévy process X with Lévy measure ν and for any $B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ the number of jumps in the time interval $[0,t]$ with sizes in B is a Poisson random variable with parameter $t\nu(B)$. Therefore, Definition 4.3 can be equivalently restated in terms of the Lévy measure ν of X as follows:

Definition 4.4. Let X be a $\mathbb{R}^d$-valued Lévy process with Lévy measure ν. Its jumps are said to be *completely dependent* or *comonotonic* if there exists a strictly ordered subset S of K such that $\nu(\mathbb{R}^d \setminus S) = 0$.

4.3 Increasing functions

In this section we sum up some aspects of the theory of increasing functions of several variables. Most definitions are well known (see, e.g. [88]) but some are introduced either here for the first time or in [59].

Let $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ denote the extended real line. In the sequel, for $a, b \in \bar{\mathbb{R}}^d$ such that $a \leq b$ (inequalities like this one are to be interpreted componentwise), let $[a,b)$ denote a right-open left-closed interval of $\bar{\mathbb{R}}^d$ (d-box):

$$[a,b) := [a_1,b_1) \times \cdots \times [a_d,b_d),$$

In the same way we define other types of intervals: $(a,b]$, $[a,b]$ and (a,b). To simplify notation it is convenient to introduce a "general" type of interval: when we want to consider a d-dimensional interval as a set of its vertices, without specifying whether the boundary is included or not, we will write $|a,b|$.

Definition 4.5 (F-volume). Let $F : S \subset \bar{\mathbb{R}}^d \to \bar{\mathbb{R}}$. For $a, b \in S$, with $a \leq b$, the F-volume of $|a, b|$ is defined by

$$V_F(|a, b|) := \sum (-1)^{N(c)} F(c), \tag{4.4}$$

where the sum is taken over all vertices c of $|a, b|$, and $N(c) := \#\{k : c_k = a_k\}$.

The notion of F-volume should only be seen as a convenient notation for the sum in the right-hand side of (4.4). It does not in general correspond to any measure because the measure of $|a, b|$ will depend on whether the boundary is included or not.

In particular, in two dimensions ($d = 2$) the F-volume of a rectangle $B = |x_1, x_2| \times |y_1, y_2|$ satisfies

$$V_F(B) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1).$$

If $F(u) = \prod_{i=1}^d u_i$, the F-volume of any interval is equal to its Lebesgue measure.

Definition 4.6 (d-increasing function). A function $F : S \subset \bar{\mathbb{R}}^d \to \bar{\mathbb{R}}$ is called d-increasing if for all $a, b \in S$ with $a \leq b$, $V_F(|a, b|) \geq 0$.

Definition 4.7 (Grounded function). For each k, let $S_k \subset \bar{\mathbb{R}}$ be such that $\inf S_k \in S_k$. A function $F : S_1 \times \cdots \times S_d \to \bar{\mathbb{R}}$ is grounded if $F(x_1, \ldots, x_d) = 0$ whenever $x_k = \inf S_k$ for at least one k.

Definition 4.8 (margins of a d-increasing grounded function). For each k, let $S_k \subset \bar{\mathbb{R}}$ be such that $\inf S_k \in S_k$ and $\sup S_k \in S_k$ and let $F : S_1 \times \cdots \times S_d \to \bar{\mathbb{R}}$ be d-increasing and grounded. Then for $I \subset \{1, \ldots, d\}$ nonempty, the I-margin of F is a function $F^I : \prod_{i \in I} S_i \to \bar{\mathbb{R}}$ defined by

$$F^I((x_i)_{i \in I}) := F(x_1, \ldots, x_d)|_{x_i = \sup S_i,\, i \in I^c}. \tag{4.5}$$

When $I = \{k\}$, to simplify notation, the (one-dimensional) I-margin of F is denoted by F_k.

Example 4.3. The distribution function F of a random vector $X \in \mathbb{R}^d$ is usually defined by

$$F(x_1, \ldots, x_d) := P[X_1 \leq x_1, \ldots, X_d \leq x_d]$$

for $x_1, \dots, x_d \in \bar{\mathbb{R}}$. F is then clearly increasing because for every $a, b \in \bar{\mathbb{R}}^d$ with $a \leq b$,

$$V_F(|a, b|) = P[X \in (a, b]] \tag{4.6}$$

It is grounded because $F(x_1, \dots, x_d) = 0$ if $x_k = -\infty$ for some k, and the margins of F are the distribution functions of the margins of X: for example, $F_1(x) = F(x, \infty, \dots, \infty) = P[X_1 \leq x]$.

The following technical lemma will be useful in the sequel.

Lemma 4.3. *For each k, let $S_k \subset \bar{\mathbb{R}}$ be such that $\inf S_k \in S_k$ and let $F, H : S_1 \times \cdots \times S_d \to \bar{\mathbb{R}}$ be d-increasing and grounded. Then their product FH is d-increasing and grounded.*

Proof. We will prove this lemma by induction on d. The product of two increasing grounded functions on $\bar{\mathbb{R}}$ is clearly increasing and grounded. Suppose $d \geq 2$, and for each k let $a_k, b_k \in S_k$ with $a_k \leq b_k$. Consider the function

$$\begin{aligned}\tilde{F}(u_2, \dots, u_d) &:= F(b_1, u_2, \dots, u_d)H(b_1, u_2, \dots, u_d) - F(a_1, u_2, \dots, u_d)H(a_1, u_2, \dots, u_d) \\ &= H(a_1, u_2, \dots, u_d)[F(b_1, u_2, \dots, u_d) - F(a_1, u_2, \dots, u_d)] \\ &\qquad + F(b_1, u_2, \dots, u_d)[H(b_1, u_2, \dots, u_d) - H(a_1, u_2, \dots, u_d)]\end{aligned}$$

Since H is grounded,

$$\begin{aligned}V_{H(a_1,*)}(|a_2, b_2| \times \cdots \times |a_d, b_d|) &= V_H(|0, a_1| \times |a_2, b_2| \times \cdots \times |a_d, b_d|) \geq 0, \\ V_{H(b_1,*)-H(a_1,*)}(|a_2, b_2| \times \cdots \times |a_d, b_d|) &= V_H(|a_1, b_1| \times |a_2, b_2| \times \cdots \times |a_d, b_d|) \geq 0,\end{aligned}$$

and the same is true for F. Therefore $\tilde{F}$ is increasing by the induction hypothesis. Since

$$V_{FH}(|a_1, b_1| \times \cdots \times |a_d, b_d|) = V_{\tilde{F}}(|a_2, b_2| \times \cdots \times |a_d, b_d|),$$

this finishes the proof of the lemma. □

To develop the theory of Lévy copulas for general Lévy processes, we will need to define the notion of margins for a function that is not grounded. The following example gives an intuition of how this can be done.

Example 4.4. Consider the following "alternative" definition of a distribution function of a random vector X:

$$\tilde{F}(x_1, \dots, x_d) := P[X_1 \in (x_1 \wedge 0, x_1 \vee 0]; \dots; X_d \in (x_d \wedge 0, x_d \vee 0]] \prod_{i=1}^{d} \operatorname{sgn} x_i \tag{4.7}$$

for $x_1, \dots x_d \in \bar{\mathbb{R}}$. This function satisfies Equation (4.6) and can therefore play the role of a distribution function. However, the margins of $\tilde{F}$ (e.g. the distribution functions computed using Equation (4.7) for the components of X) are no longer given by (4.5). It can be shown that

$$\tilde{F}^I((x_i)_{i\in I}) := P[X_i \in (x_i \wedge 0, x_i \vee 0], i \in I] \prod_{i\in I} \operatorname{sgn} x_i$$
$$= \sum_{(x_i)_{i\in I^c} \in \{-\infty,\infty\}^{|I^c|}} \tilde{F}(x_1, \dots, x_d) \prod_{i\in I^c} \operatorname{sgn} x_i.$$

The above example motivates the following definitions. First, we need to generalize the notion of "increasing and grounded" function.

Definition 4.9 (Volume function). Let $S_k \subset \bar{\mathbb{R}}$ for $k = 1, \dots, d$. A function $F : S_1 \times \cdots \times S_d$ is called *volume function* if it is d-increasing and there exists $x^* \in S_1 \times \cdots \times S_d$ such that $F(x_1, \dots, x_d) = 0$ whenever $x_k = x_k^*$ for some k.

The term "volume function" is due to the fact that an increasing function is a volume function if and only if there exists $x^* \in S_1 \times \cdots \times S_d$ such that

$$F(x_1, \dots, x_d) = V_F(|x_1 \wedge x_1^*, x_1 \vee x_1^*| \times \cdots \times |x_d \wedge x_d^*, x_d \vee x_d^*|) \prod_{i=1}^{d} \operatorname{sgn}(x_i - x_i^*)$$

for all $x \in S_1 \times \cdots \times S_d$. Every increasing grounded function is a volume function. The function $\tilde{F}$, defined in Example 4.4 is a volume function but is not grounded.

Definition 4.10 (Margins of a volume function). Let $S_k \subset \bar{\mathbb{R}}$ for $k = 1, \dots, d$ and let $F : S_1 \times \cdots \times S_d \to \bar{\mathbb{R}}$ be a volume function. Then for $I \subset \{1, \dots, d\}$ nonempty, the I-margin of F is a function $F^I : \prod_{i\in I} S_i \to \bar{\mathbb{R}}$ defined by

$$F^I((x_i)_{i\in I}) := \left(\prod_{i\in I} \operatorname{sgn}(x_i - x_i^*) \right)$$
$$\times \sup_{a_i, b_i \in S_i : i \in I^c} \sum_{(x_i)_{i\in I^c} \in \prod_{j\in I^c} \{a_j, b_j\}} (-1)^{N((x_i)_{i\in I^c})} F(x_1, \dots, x_d), \quad (4.8)$$

where $N((x_i)_{i\in I^c}) = \#\{i \in I^c : x_i = a_i\}$.

In particular, for $d = 2$ we have $F_1(x) = \operatorname{sgn}(x - x^*) \sup_{y_1, y_2 \in S_2} \{F(x, y_2) - F(x, y_1)\}$. Equation (4.8) looks so complicated because it applies without modification to all cases that

are of interest to us: distribution functions, copulas, Lévy copulas with closed or open domains. In particular cases we will obtain simplifications that are easier to work with. Using the F-volume notation, Equation (4.8) can be rewritten as follows:

$$F^I((x_i)_{i\in I}) = \left(\prod_{i\in I}\operatorname{sgn}(x_i - x_i^*)\right) \sup_{a_i,b_i\in S_i: a_i\le b_i, i\in I^c} V_F\left(\prod_{i=1}^d \begin{cases} |x_i\wedge x_i^*, x_i\vee x_i^*|, & i\in I \\ |a_i,b_i|, & i\in I^c \end{cases}\right). \tag{4.9}$$

When F satisfies the conditions of Definition 4.8, since F is d-increasing, the sup is achieved when for every $i\in I^c$, $a_i = \inf S_i$ and $b_i = \sup S_i$. Therefore, in this case the above formula yields $F^I((x_i)_{i\in I}) = F(x_1,\dots,x_d)|_{x_i=\sup S_i,\, i\in I^c}$ and the two definitions coincide.

The following important property of increasing functions will be useful in the sequel.

Lemma 4.4. *Let $S_k\subset\bar{\mathbb{R}}$ for $k=1,\dots,d$ and let $F: S_1\times\cdots\times S_d\to\bar{\mathbb{R}}$ be a volume function. Let $(x_1,\dots,x_d)$ and $(y_1,\dots,y_d)$ be any points in* $\operatorname{Dom} F$*. Then*

$$|F(x_1,\dots,x_d) - F(y_1,\dots,y_d)| \le \sum_{k=1}^d |F_k(x_k) - F_k(y_k)|. \tag{4.10}$$

Proof. From the triangle inequality,

$$\begin{aligned}|F(x_1,\dots,x_d) - F(y_1,\dots,y_d)| \le |F(x_1,\dots,x_d) - F(y_1,x_2\dots,x_d)| + \dots \\ + |F(y_1,\dots,y_{d-1},x_d) - F(y_1,\dots,y_d)|,\end{aligned}$$

hence it suffices to prove that

$$|F(x_1,\dots,x_d) - F(y_1,x_2\dots,x_d)| \le |F_1(x_1) - F_1(y_1)|. \tag{4.11}$$

Without loss of generality suppose that $x_1\ge y_1\ge x_1^*$ and that $x_k\ge x_k^*$ for $k=2,\dots,d$ and let $B := |x_2^*,x_2|\times\cdots\times|x_d^*,x_d|$. Then

$$F(x_1,\dots,x_d) - F(y_1,x_2\dots,x_d) = V_F(|y_1,x_1|\times B). \tag{4.12}$$

Moreover, for all $x\ge x_1^*$,

$$F_1(x) = \sup_{a_i,b_i\in S_i: a_i\le b_i, i=2,\dots,d} V_F(|x_1^*,x|\times|a_2,b_2|\times\cdots\times|a_d,b_d|).$$

Fix $\varepsilon > 0$ and choose $a_i, b_i \in S_i : a_i \le b_i, i = 2, \dots, d$ such that

$$F_1(y_1) \le \varepsilon + V_F(|x_1^*, y_1| \times |a_2, b_2| \times \cdots \times |a_d, b_d|).$$

For $i = 2, \dots, d$ let $\tilde{a}_i := a_i \wedge x_i^*$ and $\tilde{b}_i := b_i \vee x_i$ and denote $\tilde{B} := |\tilde{a}_2, \tilde{b}_2| \times \cdots \times |\tilde{a}_d, \tilde{b}_d|$. Then, since F is d-increasing, $V_F(|y_1, x_1| \times B) \le V_F(|y_1, x_1| \times \tilde{B})$ and $V_F(|x_1^*, y_1| \times |a_2, b_2| \times \cdots \times |a_d, b_d|) \le V_F(|x_1^*, y_1| \times \tilde{B})$. Therefore

$$V_F(|y_1, x_1| \times B) \le V_F(|x_1^*, x_1| \times \tilde{B}) - F_1(y_1) + \varepsilon \le F_1(x_1) - F_1(y_1) + \varepsilon.$$

Since the above is true for all $\varepsilon > 0$, in view of (4.12), the proof of (4.11) is complete. □

We close this section with the definition of (ordinary) copula and the Sklar's theorem, which relates copulas to distribution functions. The proof of Sklar's theorem can be found in [90].

Definition 4.11 (Copula). A d-dimensional copula (a d-copula) is a function $C : [0,1]^d \to [0,1]$ such that

1. C is grounded and d-increasing.
2. C has margins $C_k, k = 1, 2, \dots, d$, which satisfy $C_k(u) = u$ for all u in $[0,1]$.

Theorem 4.5 (Sklar). *Let F be a d-dimensional distribution function with margins $F_1, \dots, F_d$. Then there exists a d-dimensional copula C such that for all $x \in \bar{\mathbb{R}}^d$,*

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)). \tag{4.13}$$

If $F_1, \dots, F_d$ are continuous then C is unique; otherwise, C is uniquely determined on $\operatorname{Ran} F_1 \times \cdots \times \operatorname{Ran} F_d$. *Conversely, if C is a d-copula and $F_1, \dots, F_d$ are distribution functions, then the function F defined by (4.13) is a d-dimensional distribution function with margins $F_1, \dots, F_d$.*

4.4 Lévy copulas for spectrally positive Lévy processes

This section discusses the notion of Lévy copula for Lévy processes with only positive jumps in each component. This notion was introduced by the present author in [92]. Examples of Lévy copulas for spectrally positive Lévy processes and methods to construct them will be given in Sections 4.6 and 5.1 together with the examples of general Lévy copulas. Further properties of

Lévy copulas in the spectrally positive case, including a characterization of the convergence of Lévy measures in terms of Lévy copulas can be found in a recent paper by Barndorff-Nielsen and Lindner [8].

As the laws of random variables are represented by their distribution functions, Lévy measures can be represented by their tail integrals.

Definition 4.12. Let ν be a Lévy measure on $\mathbb{R}^d_+ := [0,\infty)^d \setminus \{0\}$. The tail integral U of ν is a function $[0,\infty)^d \to [0,\infty]$ such that

1. $U(0,\dots,0) = \infty$.

2. For $(x_1,\dots,x_d) \in \mathbb{R}^d_+$,

$$U(x_1,\dots,x_d) = \nu([x_1,\infty) \times \cdots \times [x_d,\infty)).$$

The Lévy measure is uniquely determined by its tail integral, because the above definition implies that for every $x, y \in \mathbb{R}^d_+$ with $x \le y$,

$$V_U(|x,y|) = (-1)^d \nu([x_1,y_1) \times \cdots \times [x_d,y_d)). \tag{4.14}$$

Definition 4.13. Let X be a $\mathbb{R}^d$-valued Lévy process and let $I \subset \{1,\dots,d\}$ non-empty. The *I-marginal tail integral* U^I of X is the tail integral of the process $X^I := (X^i)_{i\in I}$. The one-dimensional margins are, as usual, denoted by $U_i := U^{\{i\}}$.

Lemma 4.1 entails that for $I \subset \{1,\dots,d\}$ nonempty, the I-marginal tail integral of a Lévy measure ν on $\mathbb{R}^d_+$ can be computed from the tail integral U of ν by substituting 0 instead of arguments with indices not in I:

$$U^I((x_i)_{i\in I}) = U(x_1,\dots,x_d)|_{(x_i)_{i\in I^c}=0}. \tag{4.15}$$

A Lévy copula is defined similarly to ordinary copula but on a different domain.

Definition 4.14. A function $F : [0,\infty]^d \to [0,\infty]$ is a *Lévy copula* if

1. $F(u_1,\dots,u_d) < \infty$ for $(u_1,\dots,u_d) \neq (\infty,\dots,\infty)$,

2. F is grounded: $F(u_1,\dots,u_d) = 0$ whenever $u_i = 0$ for at least one $i \in \{1,\dots,d\}$,

3. F is d-increasing,

4. $F_i(u) = u$ for any $i \in \{1, \dots, d\}$, $u \in [0, \infty]$.

The following theorem is an analog of the well-known Sklar's theorem for copulas. The proof is done using multilinear interpolation and is inspired by Sklar's proof of his theorem in [90].

Theorem 4.6. *Let ν be a Lévy measure on $\mathbb{R}^d_+$ with tail integral U and marginal Lévy measures $\nu_1, \dots, \nu_d$. There exists a Lévy copula F on $[0, \infty]^d$ such that*

$$U(x_1, \dots, x_d) = F(U_1(x_1), \dots, U_d(x_d)), \quad (x_1, \dots, x_d) \in [0, \infty)^d, \tag{4.16}$$

where $U_1, \dots, U_d$ are tail integrals of $\nu_1, \dots, \nu_d$. This Lévy copula is unique on $\prod_{i=1}^d \overline{\operatorname{Ran} U_i}$.

Conversely, if F is a Lévy copula on $[0, \infty]^d$ and $\nu_1, \dots, \nu_d$ are Lévy measures on $(0, \infty)$ with tail integrals $U_1, \dots, U_d$ then Equation (4.16) defines a tail integral of a Lévy measure on $\mathbb{R}^d_+$ with marginal Lévy measures $\nu_1, \dots, \nu_d$.

Remark 4.1. In particular, the Lévy copula F is unique if the marginal Lévy measures $\nu_1, \dots, \nu_d$ are infinite and have no atoms, because in this case $\overline{\operatorname{Ran} U_k} = [0, \infty]$ for every k.

The first part of this theorem states that all types of dependence of Lévy processes (with only positive jumps), including complete dependence and independence, can be represented with Lévy copulas. The second part shows that one can construct multivariate Lévy process models by specifying separately jump dependence structure and one-dimensional laws for the components. The laws of components can have very different structure, in particular, it is possible to construct examples of Lévy processes with some components being compound Poisson and others having an infinite jump intensity.

Proof of theorem 4.6. For the purposes of this proof, we introduce some auxiliary functions and measures. First, for every $k = 1, \dots, d$ and every $x \in [0, \infty]$ we define

$$\tilde{U}_k(x) := \begin{cases} U_k(x), & x \neq \infty \\ 0, & x = \infty. \end{cases}$$

$$\tilde{U}_k^{(-1)}(t) := \sup\{x \geq 0 : \tilde{U}_k(x) \geq t\}.$$

The following properties of $\tilde{U}_k^{(-1)}$ follow directly from its definition:

$$\tilde{U}_k^{(-1)}(t) \text{ is nonincreasing in } t,$$
$$\tilde{U}_k^{(-1)}(\infty) = 0, \tag{4.17}$$
$$\tilde{U}_k(\tilde{U}_k^{(-1)}(t)) = t \ \forall t \in \operatorname{Ran} \tilde{U}_k. \tag{4.18}$$

For every $(x_1, \dots, x_d) \in [0, \infty]^d$, we define

$$\tilde{U}(x_1, \dots, x_d) = \begin{cases} 0, & x_k = \infty \quad \text{for some } k \\ U(x_1, \dots, x_d), & \text{otherwise.} \end{cases}$$

Finally, introduce a measure $\tilde{\nu}$ on $[0, \infty]^d \setminus \{0\}$ by $\tilde{\nu}(B) := \nu(B \cap \mathbb{R}_+^d)$ for all $B \in \mathcal{B}([0, \infty]^d \setminus \{0\})$. Clearly, Equation (4.14) still holds for all $x, y \in [0, \infty]^d \setminus \{0\}$ with $x \le y$, if U and ν are replaced by $\tilde{U}$ and $\tilde{\nu}$.

First part. To prove the existence of a Lévy copula, we construct the required Lévy copula in two stages.

1. First consider the function $\tilde{F} : D := \operatorname{Ran} \tilde{U}_1 \times \cdots \times \operatorname{Ran} \tilde{U}_d \to [0, \infty]$ defined by

$$\tilde{F}(x_1, \dots, x_d) = \tilde{U}(\tilde{U}_1^{(-1)}(x_1), \dots, \tilde{U}_d^{(-1)}(x_d)).$$

Suppose that $x_k = 0$ for some k. Without loss of generality we can take $k = 1$. Then

$$\tilde{F}(0, x_2, \dots, x_d) = \tilde{U}(z, \tilde{U}_2^{(-1)}(x_2), \dots, \tilde{U}_d^{(-1)}(x_d)),$$

where z is such that $\tilde{U}_1(z) = 0$. Since $\tilde{U}$ is nonnegative and nonincreasing in each argument,

$$0 \le \tilde{U}(z, \tilde{U}_2^{(-1)}(x_2), \dots, \tilde{U}_d^{(-1)}(x_d)) \le \tilde{U}(z, 0, \dots, 0) = \tilde{U}_1(z) = 0.$$

Therefore, $\tilde{F}$ is grounded. Let $a, b \in D$ with $a_k \le b_k$, $k = 1, \dots, d$ and denote

$$B := (a_1, b_1] \times \cdots \times (a_d, b_d] \quad \text{and}$$
$$\tilde{B} := [\tilde{U}_1^{(-1)}(b_1), \tilde{U}_1^{(-1)}(a_1)) \times \cdots \times [\tilde{U}_d^{(-1)}(b_d), \tilde{U}_d^{(-1)}(a_d)).$$

Since $\tilde{U}_k^{(-1)}(b_k) \le \tilde{U}_k^{(-1)}(a_k)$ for every k, formula (4.14) entails that

$$V_F(B) = (-1)^d V_{\tilde{U}}(\tilde{B}) = \tilde{\nu}(\tilde{B}) \ge 0,$$

which means that $\tilde{F}$ is a d-increasing. Equations (4.17) and (4.18) show that for every k, $\tilde{F}(\infty,\dots,\infty,x_k,\infty,\dots,\infty) = x_k,\ \forall x_k \in \operatorname{Ran}\tilde{U}_k$. Therefore, $\tilde{F}$ has uniform margins.

2. The second stage is to extend $\tilde{F}$ to the function F defined on $[0,\infty]^d$. The extension can be carried out step by step, constructing a sequence of $d+1$ grounded d-increasing functions $F^0,\dots,F^d$ with uniform margins such that $F^0 \equiv \tilde{F}$, $F^d \equiv F$,

$$\operatorname{Dom} F^k = [0,\infty]^k \times \operatorname{Ran}\tilde{U}_{k+1} \times \cdots \times \operatorname{Ran}\tilde{U}_d,$$

and for each $k = 1,\dots,d-1$, $F^k \equiv F^{k+1}$ on $\operatorname{Dom} F^k$. In other words, at each step, we extend $\tilde{F}$ along one coordinate only. Since all d steps are performed in the same way, we need only to describe one of them. To simplify the notation and without loss of generality, we describe the first one, that is, we show how to extend $\tilde{F}$ to a function defined on $[0,\infty]\times\operatorname{Ran}\tilde{U}_2\times\cdots\times\operatorname{Ran}\tilde{U}_d$.

First, let us extend $\tilde{F}$ by continuity to a function defined on $\overline{\operatorname{Ran}\tilde{U}_1}\times\operatorname{Ran}\tilde{U}_2\times\cdots\times\operatorname{Ran}\tilde{U}_d$. Given a sequence of real numbers $\{\xi_i\}_{i=1}^{\infty}$ such that $\xi_i \in \operatorname{Ran}\tilde{U}_1$ for all i and $\lim_i \xi_i \notin \operatorname{Ran}\tilde{U}_1$, we define

$$F(\lim_i \xi_i, x_2,\dots,x_d) := \lim_i F(\xi_i, x_2,\dots,x_d).$$

Since $\infty \in \operatorname{Ran}\tilde{U}_1$, $\lim_i \xi_i$ is finite. Therefore, by Lemma 4.4, the limit in the right-hand side of the above expression exists and is uniform in other coordinates so F is well-defined. It is also clear that F is d-increasing, grounded and has uniform margins.

Now we can suppose that $\operatorname{Ran}\tilde{U}_1$ is closed and extend $\tilde{F}$ using linear interpolation to a function defined on $([0,\lambda_1]\cup\{\infty\})\times\operatorname{Ran}\tilde{U}_2\times\cdots\times\operatorname{Ran}\tilde{U}_d$ where $\lambda_1 = \lim_{x\downarrow 0}\tilde{U}_1(x)$. For any $x \le \lambda_1$ such that $x \notin \operatorname{Ran}\tilde{U}_1$, we introduce $\underline{x} := \sup\{\xi \in \operatorname{Ran}\tilde{U}_1, \xi \le x\}$, $\bar{x} := \inf\{\xi \in \operatorname{Ran}\tilde{U}_1, \xi \ge x\}$. Since $\operatorname{Ran}\tilde{U}_1$ is closed, it contains both $\underline{x}$ and $\bar{x}$. Define

$$F(x,x_2,\dots,x_d) := \tilde{F}(\underline{x},x_2,\dots,x_d)\frac{\bar{x}-x}{\bar{x}-\underline{x}} + \tilde{F}(\bar{x},x_2,\dots,x_d)\frac{x-\underline{x}}{\bar{x}-\underline{x}}. \tag{4.19}$$

The function F is clearly grounded and has uniform margins; we only need to prove that it is d-increasing. Fix a d-box $B = |x_1,y_1|\times\cdots\times|x_d,y_d|$ with vertices in the domain of F and denote $\tilde{B} = |x_2,y_2|\times\cdots\times|x_d,y_d|$.

If both x_1 and y_1 belong to $\operatorname{Ran}\tilde{U}_1$ then all vertices of B are in $\operatorname{Dom}\tilde{F}$ and $V_F(B) = V_{\tilde{F}}(B) \ge 0$.

If $x_1 \notin \operatorname{Ran}\tilde{U}_1$, $y_1 \notin \operatorname{Ran}\tilde{U}_1$ and between x_1 and y_1 there are no points that belong to

$\operatorname{Ran}\tilde{U}_1$, then a straightforward computation using Equation (4.19) shows that

$$V_F(B) = \frac{y_1 - x_1}{\bar{x}_1 - \underline{x}_1} V_{\tilde{F}}(|\underline{x}_1, \bar{x}_1| \times \tilde{B}) \geq 0.$$

Suppose that $x_1 \notin \operatorname{Ran}\tilde{U}_1$ but $y_1 \in \operatorname{Ran}\tilde{U}_1$. Let $z = \inf\{\zeta \geq x_1, \zeta \in \operatorname{Ran}\tilde{U}_1\}$. Because $\operatorname{Ran}\tilde{U}_1$ is closed, $z \in \operatorname{Ran}\tilde{U}_1$. The F-volume of B can be decomposed as follows:

$$V_F(B) = V_F(|x_1, z| \times \tilde{B}) + V_F(|z, y_1| \times \tilde{B}).$$

The second term is positive because all vertices of the corresponding d-box are in the domain of $\tilde{F}$. The first term can again be computed using formula (4.19):

$$V_F(|x_1, z| \times \tilde{B}) = \frac{z - x_1}{z - \underline{x}_1} V_{\tilde{F}}(|\underline{x}_1, z| \times \tilde{B}) \geq 0.$$

The case where $x_1 \in \operatorname{Ran}\tilde{U}_1$ and $y_1 \notin \operatorname{Ran}\tilde{U}_1$ can be treated in this same way and if $x_1 \notin \operatorname{Ran}\tilde{U}_1$, $y_1 \notin \operatorname{Ran}\tilde{U}_1$ and between x_1 and y_1 there are points that belong to $\operatorname{Ran}\tilde{U}_1$, the interval $|x_1, y_1|$ can be split onto two intervals of types that we have already discussed.

We have thus extended $\tilde{F}$ to a function defined on $([0, \lambda_1] \cup \{\infty\}) \times \operatorname{Ran}\tilde{U}_2 \times \cdots \times \operatorname{Ran}\tilde{U}_d$. If $\lambda_1 = \infty$ then we are done; otherwise we extend $\tilde{F}$ to $[0, \infty] \times \operatorname{Ran}\tilde{U}_2 \times \cdots \times \operatorname{Ran}\tilde{U}_d$ by defining

$$F(x_1, \ldots, x_d) := \tilde{F}(g_{\lambda_1}(x_1), x_2, \ldots, x_d) + (x_1 - \lambda_1)^+ 1_{x_2=\infty} \ldots 1_{x_d=\infty}, \tag{4.20}$$

where

$$g_{\lambda_1}(x) := \begin{cases} x, & x \leq \lambda_1 \\ \lambda_1, & \lambda_1 < x < \infty \\ \infty, & x = \infty. \end{cases}$$

The function F, defined by (4.20) is an increasing function because it is a sum of two increasing functions. The groundedness and marginal properties can be verified by direct substitution. This completes the construction of the Lévy copula.

To prove the uniqueness, assume that there exist two functions with required properties, i.e., we have

$$F^1(U_1(x_1), \ldots, U_d(x_d)) = F^2(U_1(x_1), \ldots, U_d(x_d)), \forall x_1, \ldots, x_d.$$

For each vector $(t_1, \ldots, t_d) \in \operatorname{Ran} U_1 \times \cdots \times \operatorname{Ran} U_d$ there exists a vector $(x_1, \ldots, x_d) \in [0, \infty]^d$ such that $U_1(x_1) = t_1, \ldots, U_d(x_d) = t_d$. This means that for every $(t_1, \ldots, t_d) \in \operatorname{Ran} U_1 \times \cdots \times$

$\operatorname{Ran} U_d$ we have $F^1(t_1,\dots,t_d) = F^2(t_1,\dots,t_d)$. The uniqueness statement now follows from Lemma 4.4.

The converse statement. Since $U_1,\dots,U_d$ are left-continuous (as tail integrals of Lévy measures), Lemma 4.4 entails that U is left-continuous in each variable. Therefore, there exists a unique positive measure ν on $\mathcal{B}(\mathbb{R}^d_+)$ such that for every right-open left-closed interval I of $\mathbb{R}^d_+$, $\nu(I) = (-1)^d V_U(I)$ (see Section 4.5 in [60]).

It remains to prove that $\int_{|x|\le 1} |x|^2 \nu(dx) < \infty$. Let us fix $k \in \{1 \dots d\}$ and consider the measure ν'_k defined by

$$\nu'_k(A) := \int_{\{x\in\mathbb{R}^d_+ : x_k \in A\}} \nu(dx) \quad \text{for } A \in \mathcal{B}((0,\infty)).$$

By the uniqueness of the extension this measure coincides with ν_k because it is straightforward to check that the two measures coincide on right-open left-closed intervals of $(0,\infty)$. Therefore

$$\int_{[0,1]^d} |x|^2 \nu(dx) = \sum_{k=1}^d \int_{[0,1]^d} x_k^2 \nu(dx) \le \sum_{k=1}^d \int_{[0,1]} x_k^2 \nu'_k(dx_k) = \sum_{k=1}^d \int_{[0,1]} x_k^2 \nu_k(dx_k).$$

Since ν_k for every k is by the theorem's statement a Lévy measure, the right-hand side of the above expression is finite, which shows that ν is indeed a Lévy measure. □

4.5 Lévy copulas for general Lévy processes

This section discusses the notion of Lévy copula for general Lévy processes, introduced in a joint work of the present author with Jan Kallsen [59]. In the sequel, we will need a special interval associated with any $x \in \mathbb{R}$:

$$\mathcal{I}(x) := \begin{cases} [x,\infty), & x \ge 0, \\ (-\infty, x), & x < 0. \end{cases} \tag{4.21}$$

Similarly to Lévy processes with positive jumps, the Lévy measure of a general Lévy process will be represented by its tail integral.

Definition 4.15. Let X be a $\mathbb{R}^d$-valued Lévy process with Lévy measure ν. The *tail integral* of ν is the function $U : (\mathbb{R}\setminus\{0\})^d \to \mathbb{R}$ defined by

$$U(x_1,\dots,x_d) := \nu\left(\prod_{j=1}^d \mathcal{I}(x_j)\right) \prod_{i=1}^d \operatorname{sgn}(x_i) \tag{4.22}$$

Note that the tail integral is defined so that $(-1)^d U$ is d-increasing and left-continuous in each orthant as was the case for the tail integral of Lévy measure on $\mathbb{R}^d_+$.

Since the tail integral is only defined on $(\mathbb{R} \setminus \{0\})^d$, it does not determine the Lévy measure uniquely (unless we know that the latter does not charge the coordinate axes). However, we will see from the following lemma that the Lévy measure is completely determined by its tail integral and all its marginal tail integrals (cf. Definition 4.13).

Lemma 4.7. *Let X be a $\mathbb{R}^d$-valued Lévy process. Its marginal tail integrals $\{U^I : I \subset \{1, \dots, d\}$ non-empty$\}$ are uniquely determined by its Lévy measure ν. Conversely, its Lévy measure is uniquely determined by the set of its marginal tail integrals.*

An important difference between the spectrally positive case and the general case is that in the spectrally positive case the Lévy measure is determined by a single tail integral U, defined on $[0, \infty)^d$, whereas in the general case the Lévy measure is determined jointly by all marginal tail integrals, each one being defined on $(\mathbb{R} \setminus \{0\})^I$. In the spectrally positive case Equation 4.15 shows that specifying the values of U when some of the arguments equal 0 is equivalent to specifying the marginal tail integrals. In principle, we could do the same in the general case and define the tail integral on $\mathbb{R}^d$ by Equation (4.22). However, due to a more complicated structure of margins of a Lévy copula in the general case, with this definition of the tail integral the representation formula (4.16) does not hold on $\mathbb{R}^d$. Therefore, we keep the Definition 4.15 and describe the Lévy measure using its tail integral, defined on $(\mathbb{R} \setminus \{0\})^d$ and all its marginal tail integrals.

Proof of Lemma 4.7. The fact that marginal tail integrals are uniquely determined by the Lévy measure ν follows from Lemma 4.1.

The converse statement. It is sufficient to prove that $\nu([a, b))$ is completely determined by the tail intergals for any $a, b \in \mathbb{R}^d$ with $a \leq b$ and $0 \notin [a, b)$. We prove by induction on $k = 0, \dots, d$ that $\nu^I(\prod_{i \in I}[a_i, b_i))$ is determined by the tail integrals for any $a, b \in \mathbb{R}^d$ such that $a \leq b$ and $a_i b_i \leq 0$ for at most k indices and any non-empty $I \subset \{1, \dots, d\}$ with $0 \notin \prod_{i \in I}[a_i, b_i)$.

If $k = 0$, Definitions 4.15 and 4.13 entail that

$$\nu^I \left(\prod_{i \in I} [a_i, b_i) \right) = (-1)^{|I|} V_{U^I} \left(\prod_{i \in I} [a_i, b_i) \right).$$

Let $a, b \in \mathbb{R}^d$ such that $a_i b_i \leq 0$ for at most k indices. For ease of notation we suppose that $a_i b_i \leq 0$ for $i = 1, \ldots, k$. Let $I \subset \{1, \ldots, d\}$ non-empty with $0 \notin \prod_{i \in I}[a_i, b_i)$. By induction hypothesis, $\nu^I(\prod_{i \in I}[a_i, b_i))$ is uniquely determined if $k \notin I$. Suppose that $k \in I$. If $b_k = 0$, then

$$\nu^I\left(\prod_{i \in I}[a_i, b_i)\right) = \lim_{\alpha \uparrow 0} \nu^I\left(\prod_{i \in I, i<k}[a_i, b_i) \times [a_k, \alpha) \times \prod_{i \in I, i>k}[a_i, b_i)\right)$$

and the right-hand side is uniquely determined by the induction hypothesis. If $b_k \neq 0$, then

$$\begin{aligned}\nu^I\left(\prod_{i \in I}[a_i, b_i)\right) &= \nu^{I \setminus \{k\}}\left(\prod_{i \in I \setminus \{k\}}[a_i, b_i)\right) \\ &\quad - \lim_{c \uparrow \infty} \nu^I\left(\prod_{i \in I, i<k}[a_i, b_i) \times [b_k, c) \times \prod_{i \in I, i>k}[a_i, b_i)\right) \\ &\quad - \lim_{\alpha \uparrow a_k; c \downarrow -\infty} \nu^I\left(\prod_{i \in I, i<k}[a_i, b_i) \times [c, a_k) \times \prod_{i \in I, i>k}[a_i, b_i)\right),\end{aligned}$$

which is uniquely determined as well. □

Lévy copulas in the general case are defined similarly to the spectrally positive case.

Definition 4.16. A function $F : (-\infty, \infty]^d \to (-\infty, \infty]$ is a *Lévy copula* if

1. $F(u_1, \ldots, u_d) < \infty$ for $(u_1, \ldots, u_d) \neq (\infty, \ldots, \infty)$,
2. $F(u_1, \ldots, u_d) = 0$ if $u_i = 0$ for at least one $i \in \{1, \ldots, d\}$,
3. F is d-increasing,
4. $F_i(u) = u$ for any $i \in \{1, \ldots, d\}$, $u \in (-\infty, \infty]$.

Remark 4.2. Since F is d-increasing, the sup in Equation (4.8) for the margins of F may be computed by taking $b_i = \infty$ and $a_i \to -\infty$ for every $i \in I^c$. Therefore, Equation (4.8) reduces to

$$F^I((x_i)_{i \in I}) := \lim_{c \to \infty} \sum_{(x_j)_{j \in I^c} \in \{-c, \infty\}^{|I^c|}} F(x_1, \ldots, x_d) \prod_{j \in I^c} \operatorname{sgn} x_j. \tag{4.23}$$

When F is a Lévy copula on $[0, \infty]^d$ (Definition 4.14), it can be extended to a Lévy copula F_{ext} on $(-\infty, \infty]^d$ by taking

$$F_{ext}(x_1, \ldots, x_d) := \begin{cases} F(x_1, \ldots, x_d), & (x_1, \ldots, x_d) \in [0, \infty]^d \\ 0 & \text{otherwise.} \end{cases}$$

In the following, the term *Lévy copula* without specifying the domain refers to a Lévy copula on $(-\infty, \infty]^d$.

Example 4.5. It is easy to check that the function

$$F(u, v) := \frac{1}{2} \frac{uv}{1 + |u| + |v|}$$

defines a Lévy copula on $(-\infty, \infty]^2$.

In view of Lemma 4.7, the following theorem is analogous to Theorem 4.6 for Lévy measures on $\mathbb{R}^d_+$ and to Sklar's theorem for probability measures.

Theorem 4.8. *Let ν be a Lévy measure on $\mathbb{R}^d \setminus \{0\}$. Then there exists a Lévy copula F such that the tail integrals of ν satisfy:*

$$U^I((x_i)_{i \in I}) = F^I((U_i(x_i))_{i \in I}) \tag{4.24}$$

for any non-empty $I \subset \{1, \dots, d\}$ and any $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$. The Lévy copula F is unique on $\prod_{i=1}^d \overline{\operatorname{Ran} U_i}$.

Conversely, if F is a d-dimensional Lévy copula and $\nu_1, \dots, \nu_d$ are Lévy measures on $\mathbb{R} \setminus \{0\}$ with tail integrals $U_i, i = 1, \dots, d$ then there exists a unique Lévy measure on $\mathbb{R}^d \setminus \{0\}$ with one-dimensional marginal tail integrals $U_1, \dots, U_d$ and whose marginal tail integrals satisfy Equation (4.24) for any non-empty $I \subset \{1, \dots, d\}$ and any $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$.

Remark 4.3. In particular, the Lévy copula F is unique if for each k, the marginal Lévy measure ν_k has no atoms and both $\nu_k((-\infty, 0)) = \infty$ and $\nu_k((0, \infty)) = \infty$, because in this case $\overline{\operatorname{Ran} U_k} = [-\infty, \infty]$ for each k.

Definition 4.17. For a Lévy process X on $\mathbb{R}^d$ with characteristic triplet (A, ν, γ), any Lévy copula as in statement of Theorem 4.8 is called the *Lévy copula of X*.

Proof of Theorem 4.8. First part. Denote the marginal Lévy measures of ν by $\nu_1, \dots, \nu_d$. For the purposes of this proof we set for $x \in (-\infty, \infty]$, $i = 1, \dots, d$,

$$\tilde{U}_i(x) := \begin{cases} U_i(x) & \text{for } x \neq 0 \text{ and } x \neq \infty, \\ 0 & \text{for } x = \infty, \\ \infty & \text{for } x = 0 \end{cases}$$

and

$$\Delta U_i(x) := \begin{cases} \lim_{\xi \downarrow x} U_i(\xi) - U_i(x) & \text{for } x \neq 0 \text{ and } x \neq \infty, \\ 0 & \text{for } x = \infty \text{ or } x = 0. \end{cases}$$

Let m be the measure on $((-\infty,\infty]^d \setminus \{0\}) \times [0,1]^d \times \mathbb{R}$ defined by

$$m := \tilde{\nu} \otimes \lambda|_{(0,1)^d} \otimes \delta_0 + \sum_{i=1}^{d} \delta_{(\underbrace{0,\ldots,0}_{i-1},\infty,0,\ldots,0)} \otimes \delta_{(\underbrace{0,\ldots,0}_{d})} \otimes \lambda|_{(\nu_i((0,\infty)),\infty) \cup (-\infty,-\nu_i((-\infty,0)))},$$

where $\tilde{\nu}$ is the extension of ν to $(-\infty,\infty]^d \setminus \{0\}$, i.e. $\tilde{\nu}(B) := \nu(B \cap \mathbb{R}^d)$. Let

$$g_i : (-\infty,\infty] \times [0,1] \times \mathbb{R} \to (-\infty,\infty], \quad (x,y,z) \mapsto \tilde{U}_i(x) + y\Delta U_i(x) + z$$

and define a measure $\tilde{m}$ on $(-\infty,\infty]^d \setminus \{\infty,\ldots,\infty\}$ via

$$\tilde{m}(B) := m(\tilde{g}^{-1}(B))$$

with

$$\tilde{g}(x_1,\ldots,x_d,y_1,\ldots,y_d,z) := (g_1(x_1,y_1,z),\ldots,g_d(x_d,y_d,z)).$$

Finally, let F be given by

$$F(u_1,\ldots,u_d) := \begin{cases} \tilde{m}\left(\prod_{i=1}^{d}(u_i \wedge 0, u_i \vee 0]\right) \prod_{i=1}^{d} \operatorname{sgn} u_i, & (u_1,\ldots,u_d) \in (-\infty,\infty]^d \setminus (\infty,\ldots,\infty) \\ \infty, & (u_1,\ldots,u_d) = (\infty,\ldots,\infty). \end{cases}$$

Properties 1 and 2 in Definition 4.16 are obvious. From the fact that $\tilde{m}$ is a positive measure it follows immediately that F is d-increasing. Let $I \subset \{1,\ldots,d\}$ nonempty and $(u_i)_{i\in I} \in (-\infty,\infty]^I$. For ease of notation, we consider only the case of non-negative u_i. The general case follows analogously. By definition of F we have

$$\begin{aligned} F^I((u_i)_{i\in I}) &= \lim_{c\to\infty} \sum_{(u_j)_{j\in I^c} \in \{-c,\infty\}^{I^c}} F(u_1,\ldots,u_d) \prod_{j\in I^c} \operatorname{sgn} u_j \\ &= \tilde{m}\left(\prod_{i\in I}(0,u_i] \times (-\infty,\infty]^{I^c}\right) \\ &= m\Bigg(\Big\{(x_1,\ldots,x_d,y_1,\ldots,y_d,z) \in ((-\infty,\infty]^d \setminus \{0\}) \times [0,1]^d \times \mathbb{R} : \\ &\qquad \tilde{U}_i(x_i) + y_i\Delta U_i(x_i) + z \in (0,u_i] \text{ for } i \in I\Big\}\Bigg). \end{aligned}$$

If $I = \{i\}$, then the definition of m implies that this equals

$$\left(\nu_i \otimes \lambda|_{(0,1)}\right)\left(\{(x,y) \in \mathbb{R} \times [0,1] : \tilde{U}_i(x) + y\Delta U_i(x) \in (0,u_i]\}\right) + (u_i - \nu_i((0,\infty)))1_{\{u_i > \nu_i((0,\infty))\}}.$$

Introducing $x^* := \inf\{x \geq 0 : \tilde{U}_i(x) + \Delta U_i(x) \leq u_i\}$, this can be expressed as

$$\nu_i((x^*,\infty)) + (u_i - \tilde{U}_i(x^*) - \Delta U_i(x^*))1_{\{x^* \neq 0\}} + (u_i - \nu_i((0,\infty)))1_{\{x^*=0\}} = u_i,$$

i.e. property 4 in Definition 4.16 is met.

Now, let $(x_i)_{i\in I} \in (\mathbb{R} \setminus \{0\})^I$. Again, we consider only the case where all the x_i are nonnegative. Then

$$\begin{aligned} F^I((U_i(x_i))_{i\in I}) &= m\Bigg(\Big\{(\tilde{x}_1,\ldots,\tilde{x}_d,y_1,\ldots,y_d,z) \in (-\infty,\infty]^d \times [0,1]^d \times \mathbb{R} : \\ &\qquad \tilde{U}_i(\tilde{x}_i) + y_i\Delta U_i(\tilde{x}_i) + z \in (0,U_i(x_i)] \text{ for } i \in I\Big\}\Bigg) \\ &= \nu\left(\prod_{i\in I}[x_i,\infty) \times \mathbb{R}^{I^c}\right) \\ &= \nu^I\left(\prod_{i\in I}[x_i,\infty)\right) \\ &= U^I((x_i)_{i\in I}) \end{aligned}$$

as claimed. The uniqueness statement follows from (4.24) and Lemma 4.4.

The converse statement. Since F is d-increasing and continuous (by Lemma 4.4), there exists a unique measure μ on $(-\infty,\infty]^d \setminus \{\infty,\ldots,\infty\}$ such that $V_F(|a,b|) = \mu((a,b])$ for any $a,b \in (-\infty,\infty]^d \setminus \{\infty,\ldots,\infty\}$ with $a \leq b$. (see [60], Section 4.5). For a one-dimensional tail integral $U(x)$, we define

$$U^{(-1)}(u) := \begin{cases} \sup\{x > 0 : U(x) \geq u\} \vee 0, & u \geq 0 \\ \sup\{x < 0 : U(x) \geq u\}, & u < 0. \end{cases} \tag{4.25}$$

Let $\tilde{\nu} := f(\mu)$ be the image of μ under

$$f : (u_1,\ldots,u_d) \mapsto (U_1^{(-1)}(u_1),\ldots,U_d^{(-1)}(u_d))$$

and let ν be the restriction of $\tilde{\nu}$ to $\mathbb{R}^d \setminus \{0\}$. We need to prove that ν is a Lévy measure and that its marginal tail integrals U_ν^I satisfy

$$U_\nu^I((x_i)_{i\in I}) = F^I((U_i(x_i))_{i\in I})$$

for any non-empty $I \subset \{1, \dots, d\}$ and any $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$. Suppose for ease of notation that $x_i > 0$, $i \in I$. Then

$$\begin{aligned} U_\nu^I((x_i)_{i\in I}) &= \nu(\{\xi \in \mathbb{R}^d : \xi_i \in [x_i, \infty), i \in I\}) \\ &= \mu(\{u \in (-\infty, \infty]^d : U_i^{(-1)}(u_i) \in [x_i, \infty), i \in I\}) \\ &= \mu(\{u \in (-\infty, \infty]^d : 0 < u_i \leq U(x_i), i \in I\}). \\ &= F^I((U_i(x_i))_{i\in I}). \end{aligned}$$

This proves in particular that the one-dimensional marginal tail integrals of ν equal $U_1, \dots, U_d$.

Since the marginals ν_i of ν are Lévy measures on $\mathbb{R}$, we have $\int (x_i^2 \wedge 1)\nu_i(dx_i) < \infty$ for $i = 1, \dots, d$. This implies

$$\int (|x|^2 \wedge 1)\nu(dx) \leq \int \sum_{i=1}^d (x_i^2 \wedge 1)\nu(dx) \sum_{i=1}^d \int (x_i^2 \wedge 1)\nu_i(dx_i) < \infty$$

and hence ν is a Lévy measure on $\mathbb{R}^d$. The uniqueness of ν follows from the fact that it is uniquely determined by its marginal tail integrals (cf. Lemma 4.7). □

4.6 Examples of Lévy copulas

In this section we derive the form of Lévy copulas corresponding to special dependence structures of Lévy processes: independence, complete dependence and the dependence of stable processes. Examples of parametric families of Lévy copulas will be given in the next chapter. To characterize independence of components of a multidimensional Lévy process in terms of its Lévy copula, we need to restate Lemma 4.2 in terms of tail integrals.

Lemma 4.9. *The components $X^1, \dots, X^d$ of an $\mathbb{R}^d$-valued Lévy process X are independent if and only if their continuous martingale parts are independent and the tail integrals of the Lévy measure satisfy $U^I((x_i)_{i\in I}) = 0$ for all $I \subset \{1, \dots, d\}$ with $\operatorname{card} I \geq 2$ and all $(x_i)_{i\in I} \in (\mathbb{R}\setminus\{0\})^I$.*

Proof. The "only if" part. Let $I \subset \{1, \dots, d\}$ with $\operatorname{card} I \geq 2$ and $(x_i)_{i\in I} \in (\mathbb{R} \setminus \{0\})^I$. Then the components of the Lévy process $(X^i)_{i\in I}$ are independent as well. Applying Lemma 4.2 to this process, we conclude, using Equation (4.2), that $U^I((x_i)_{i\in I}) = 0$.

The "if" part. Let ν be defined by Equation (4.2), where ν_i is the Lévy measure of X^i for $i = 1, \dots, d$. Then all marginal tail integrals of ν coincide with those of the Lévy measure of

X. Therefore, ν is the Lévy measure of X (cf. Lemma 4.7), which entails by Lemma 4.2 that $X^1, \dots, X^d$ are independent. □

Theorem 4.10. *The components $X^1, \dots, X^d$ of an $\mathbb{R}^d$-valued Lévy process X are independent if and only if their Brownian motion parts are independent and X has a Lévy copula of the form*

$$F_\perp(x_1, \dots, x_d) := \sum_{i=1}^{d} x_i \prod_{j \neq i} 1_{\{\infty\}}(x_j) \tag{4.26}$$

Proof. It is straightforward to see that Equation (4.26) defines a Lévy copula.

The "if" part. Let $I \subset \{1, \dots, d\}$ with $\operatorname{card} I \geq 2$. Equation (4.23) entails that $F_\perp^I((u_i)_{i \in I}) = 0$ for all $(u_i)_{i \in I} \in \mathbb{R}^I$. Therefore, the tail integrals of X satisfy $U^I((x_i)_{i \in I}) = 0$ for all $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$ by (4.24). From Lemma 4.9 we conclude that $X^1, \dots, X^d$ are independent.

The "only if" part. By Lemma 4.9, the tail integrals of X satisfy $U^I((x_i)_{i \in I}) = 0$ for all $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^I$ and all $I \subset \{1, \dots, d\}$ with $\operatorname{card} I \geq 2$. Since also $F_\perp^I((u_i)_{i \in I}) = 0$, we conclude that $F_\perp$ is a Lévy copula for X. □

Observing that $F_\perp|_{[0,\infty]^d}$ is a Lévy copula in the sense of Definition 4.14, we obtain the following easy corollary.

Corollary 4.1. *The components $X^1, \dots, X^d$ of an $\mathbb{R}^d$-valued spectrally positive Lévy process X are independent if and only if their Brownian motion parts are independent and X has a Lévy copula on $[0, \infty]^d$ of the form $F_\perp|_{[0,\infty]^d}$, where $F_\perp$ is defined by* (4.26).

The following theorem describes complete jump dependence in terms of Lévy copulas.

Theorem 4.11. *Let X be a $\mathbb{R}^d$-valued Lévy process whose Lévy measure is supported by an ordered set $S \subset K$, where K is as in Equation* (4.3). *Then the* complete dependence Lévy copula *given by*

$$F_\parallel(x_1, \dots, x_d) := \min(|x_1|, \dots, |x_d|) 1_K(x_1, \dots, x_d) \prod_{i=1}^{d} \operatorname{sgn} x_i \tag{4.27}$$

is a Lévy copula of X.

Conversely, if $F_\parallel$ is a Lévy copula of X, then the Lévy measure of X is supported by an ordered subset of K. If, in addition, the tail integrals U_i of X^i are continuous and satisfy $\lim_{x \to 0} U_i(x) = \infty$, $i = 1, \dots, d$, then the jumps of X are completely dependent.

The proof is based on the following representation of an ordered set as a union of a strictly ordered set and countable many segments that are perpendicular to some coordinate axis.

Lemma 4.12. *Let $S \subset \mathbb{R}^d$ be an ordered set. It can be written as*

$$S = S^* \cup \bigcup_{n=1}^{\infty} S_n, \tag{4.28}$$

where $S^* \subset \mathbb{R}^d$ is strictly ordered and for every n, $S_n \subset \mathbb{R}^d$ and there exist $k(n)$ and $\xi(n)$ such that $x_{k(n)} = \xi(n)$ for all $x \in S_n$.

Proof. For the purposes of this proof we define the length of an ordered set S' by $|S'| := \sup_{a,b \in S'} \sum_{i=1}^{d} (a_i - b_i)$. Let

$$S(\xi, k) = \{x \in \mathbb{R}^d : x_k = \xi\} \cap S. \tag{4.29}$$

First, we want to prove that there is at most a countable number of such segments with non-zero length. Consider N different segments of this type $(S_i := S(\xi_i, k_i))_{i=1}^{N}$ with $k_i = k$ and $|S_i| \geq \varepsilon > 0$ for all i. Since S_i are different, ξ_i must all be different and we can suppose without loss of generality that $\xi_i < \xi_{i+1}$ for all i. Then $\overline{x}_i \leq \underline{x}_{i+1}$ for all i, where $\overline{x}_i$ and $\underline{x}_i$ are the upper and the lower bounds of S_i. Since all S_i are subsets of S, which is an ordered set, this implies that $|\bigcup_{i=1}^{N} S_i| \geq N\varepsilon$. Therefore, for all $A > 0$ and for all $\varepsilon > 0$, the set $[-A, A]^d$ contains a finite number of segments of type (4.29) with length greater or equal to ε. This means that there is at most a countable number of segments of non-zero length, and one can enumerate them in a sequence $\{S_n\}_{n=1}^{\infty}$ with $S_n := S(\xi(n), k(n))$.

Now let $S^* = S \setminus \bigcup_{n=1}^{\infty} S_n$. S^* is ordered because it is a subset of S. Let $x, y \in S^*$. If $x_k = y_k$ for some k then either x and y are the same or they are in some segment of type (4.29) hence not in S^*. Therefore, either $x_k < y_k$ for every k or $x_k > y_k$ for every k, which entails that S^* is strictly ordered and we have obtained the desired representation for S. □

Proof of Theorem 4.11. We start by proving that $F_{\|}$ is indeed a Lévy copula in the sense of Definition 4.16. Properties 1 and 2 are obvious. To show property 3, introduce a positive measure μ on $\bar{\mathbb{R}}^d$ by

$$\mu(B) = \lambda(\{x \in \mathbb{R} : (\underbrace{x, \ldots, x}_{d \text{ times}}) \in B\}), \quad B \in \mathcal{B}(\bar{\mathbb{R}}^d),$$

where λ denotes the Lebesgue measure on $\mathbb{R}$. Then $V_{F_\parallel}((a,b]) = \mu((a,b])$ for any $a \le b$, and therefore $F_\parallel$ is d-increasing. The margins of F have the same form as F, namely

$$F_\parallel^I((x_i)_{i\in I}) = \min_{i\in I} |x_i| 1_{\{(-1,\dots,-1),(1,\dots,1)\}}((\operatorname{sgn} x_i)_{i\in I}) \prod_{i\in I} \operatorname{sgn} x_i. \tag{4.30}$$

Therefore, the one-dimensional margins satisfy $F^{\{i\}}(u) = u$.

The first part. Let $x \in (0,\infty)^d$. Clearly, $U(x) \le U_k(x_k)$ for any k. On the other hand, since S is an ordered set, we have

$$\{y \in \mathbb{R}^d : x_k \le y_k\} \cap S = \{y \in \mathbb{R}^d : x \le y\} \cap S$$

for some k. Indeed, suppose that this is not so. Then there exist points $z^1, \dots, z^d \in S$ and indices $j_1, \dots, j_d$ such that $z^k_k \ge x_k$ and $z^k_{j_k} < x_{j_k}$ for $k = 1, \dots, d$. Choose the greatest element among $z^1, \dots, z^d$ (this is possible because they all belong to an ordered set) and call it z^k. Then $z^k_{j_k} < x_{j_k}$. However, by construction of $z^1, \dots, z^d$ we also have $z^{j_k}_{j_k} \ge x_{j_k}$, which is a contradiction to the fact that z^k is the greatest element. Therefore,

$$U(x) = \min(U_1(x_1), \dots, U_d(x_d)).$$

Similarly, it can be shown that for every $x \in (-\infty, 0)^d$,

$$U(x) = (-1)^d \min(|U_1(x_1)|, \dots, |U_d(x_d)|).$$

Since $U(x) = 0$ for any $x \notin K$, we have shown that

$$U(x) = F_\parallel(U_1(x_1), \dots, U_d(x_d))$$

for any $x \in (\mathbb{R} \setminus \{0\})^d$. Since the marginal Lévy measures of X are also supported by ordered sets and the margins of $F_\parallel$ have the same form as $F_\parallel$, we have

$$U^I((x_i)_{i\in I}) = F_\parallel^I((U_i(x_i))_{i\in I}) \tag{4.31}$$

for any $I \subset \{1, \dots, d\}$ and any $(x_i)_{i\in I} \in (\mathbb{R} \setminus \{0\})^I$.

The converse statement. Let $S := \operatorname{supp} \nu$. Let us first show that $S \subseteq K$. Suppose that this is not so. Then there exists $x \in S$ such that for some m and n, $x_m < 0$ and $x_n > 0$ and for every neighborhood N of x, $\nu(N) > 0$. This implies that $U^{\{m,n\}}(x_m/2, x_n/2) > 0$, which contradicts Equation (4.31).

Suppose now that S is not an ordered set. Then there exist two points $u, v \in S$ such that $u_m > v_m$ and $u_n < v_n$ for some m and n. Moreover, either $u_i \geq 0$ and $v_i \geq 0$ for all i or $u_i \leq 0$ and $v_i \leq 0$ for all i. Suppose that $u_i \geq 0$ and $v_i \geq 0$, the other case being analogous. Let $x = \frac{u+v}{2}$. Since $u, v \in S$, we have $\nu(\{z \in \mathbb{R}^d : z_m < x_m, z_n \geq x_n\}) > 0$ and $\nu(\{z \in \mathbb{R}^d : z_m \geq x_m, z_n < x_n\}) > 0$. However

$$\begin{aligned} \nu(\{z \in \mathbb{R}^d : z_m < x_m, z_n \geq x_n\}) &= U_n(x_n) - U^{\{m,n\}}(x_m, x_n) \\ &= U_n(x_n) - \min(U_m(x_m), U_n(x_n)) \end{aligned}$$

and

$$\nu(\{z \in \mathbb{R}^d : z_m \geq x_m, z_n < x_n\}) = U_m(x_m) - \min(U_m(x_m), U_n(x_n)),$$

which is a contradiction because these expressions cannot be simultaneously positive.

For the last assertion, we assume that the tail integrals U_i of X^i are continuous and satisfy $\lim_{x \to 0} U_i(x) = \infty$, $i = 1, \ldots, d$. It suffices to show that $\nu(S_n) = 0$ for any n in decomposition (4.28). If $\xi(n) \neq 0$, then

$$\nu(S_n) = \lim_{\varepsilon \downarrow 0} (U_{k(n)}(\xi(n) - \varepsilon) - U_{k(n)}(\xi(n))) = 0,$$

because $U_{k(n)}$ is continuous. Suppose now that $\xi(n) = 0$. Since S_n does not reduce to a single point, we must have either $x_m > 0$ or $x_m < 0$ for some $x \in S_n$ and some m. Suppose that $x_m > 0$, the other case being analogous. Since S is ordered, we have

$$\nu(\{x \in \mathbb{R}^d : x_{k(n)} \geq \varepsilon\} \cap S) \leq \nu(\{\xi \in \mathbb{R}^d : \xi_m \geq x_m\} \cap S) < \infty$$

uniformly in $\varepsilon > 0$. This implies $\lim_{x \downarrow 0} U_{k(n)}(x) < \infty$ in contradiction to $\lim_{x \to 0} U_{k(n)}(x) = \infty$. Hence, $\xi(n) > 0$ for any n. Therefore, $\nu(\mathbb{R}^d \setminus S^*) = 0$ and the proof is completed. □

A characterization of complete dependence of spectrally positive Lévy processes can be obtained as a corollary of Theorem 4.11.

Corollary 4.2. *Let X be a $\mathbb{R}^d$-valued spectrally positive Lévy process whose Lévy measure is supported by an ordered set $S \subset \mathbb{R}^d_+$. Then*

$$F_{\|}(x_1, \ldots, x_d)|_{[0,\infty]^d} \equiv \min(x_1, \ldots, x_d)$$

is a Lévy copula of X.

Conversely, if $F_{\|}|_{[0,\infty]^d}$ is a Lévy copula of a spectrally positive Lévy process X, then the Lévy measure of X is supported by an ordered subset of $\mathbb{R}^d_+$. If, in addition, the tail integrals U_i of X^i are continuous and satisfy $\lim_{x\downarrow 0} U_i(x) = \infty$, $i = 1, \dots, d$, then the jumps of X are completely dependent.

Lévy copulas provide a simple characterization of possible dependence patterns of multidimensional stable processes.

Theorem 4.13. *Let $X := (X^1, \dots, X^d)$ be a Lévy process on $\mathbb{R}^d$ and let $\alpha \in (0, 2)$. X is α-stable if and only if its components $X^1, \dots, X^d$ are α-stable and it has a Lévy copula F that is a homogeneous function of order 1:*

$$\forall r > 0,\ \forall u_1, \dots, u_d,\ F(ru_1, \dots, ru_d) = rF(u_1, \dots, u_d). \tag{4.32}$$

Proof. The "only if" part. Let X be α-stable. For each $i = 1, \dots, d$, three situations are possible: $\operatorname{Ran} U_i = (-\infty, 0]$ (only negative jumps), $\operatorname{Ran} U_i = [0, \infty)$ (only positive jumps) or $\operatorname{Ran} U_i = (-\infty, 0) \cup (0, \infty)$ (jumps of both signs). We exclude the trivial case of a component having no jumps at all. Let $I_1 = \{i : \operatorname{Ran} U_i = (-\infty, 0]\}$ and $I_2 = \{i : \operatorname{Ran} U_i = [0, \infty)\}$ and for each i, let $\bar{X}^i$ be a copy of X^i, independent from X and from $\bar{X}^k$ for $k \neq i$. Define a Lévy process $\tilde{X}$ on $\mathbb{R}^d$ by

$$\tilde{X}^i = \begin{cases} X^i, & i \notin I_1 \cup I_2, \\ X^i - \bar{X}^i, & i \in I_1 \cup I_2. \end{cases}$$

Let $\tilde{\nu}$ be the Lévy measure of $\tilde{X}$, $\tilde{U}$ be its tail integral and $\tilde{F}$ be its Lévy copula (it exists by Theorem 4.8). The process $\tilde{X}$ is clearly α-stable and each component of this process has jumps of both signs ($\operatorname{Ran} \tilde{U}_i = \mathbb{R} \setminus \{0\}$). By Theorem 14.3 in [87], for every $B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ and for every $r > 0$,

$$\tilde{\nu}(B) = r^\alpha \tilde{\nu}(rB). \tag{4.33}$$

Therefore, for every $I \subset \{1, \dots, d\}$ nonempty and for every $(x_i)_{i\in I} \in (\mathbb{R} \setminus \{0\})^{|I|}$,

$$\tilde{U}^I((x_i)_{i\in I}) = r^\alpha \tilde{U}^I((rx_i)_{i\in I}). \tag{4.34}$$

By Theorem 4.8 this implies that for all $(u_1, \dots, u_d) \in (\mathbb{R} \setminus \{0\})^d$,

$$\tilde{F}^I((u_i)_{i\in I}) = r^{-1} \tilde{F}^I((ru_i)_{i\in I}),$$

and therefore (4.32) holds for $\tilde{F}$. However, $\tilde{F}$ is also the copula of X. Indeed, let $I \subset \{1, \dots, d\}$ nonempty and $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^{|I|}$. Two situations are possible:

- For every $i \in I$, $\tilde{U}_i(x_i) = U(x_i)$. Then it is easy to see that Equation (4.24) holds with F replaced by $\tilde{F}$.
- For some $k \in I$, $U_k(x_k) = 0$. Then $\tilde{F}^I((U_i(x_i))_{i \in I}) = 0$, but on the other hand $|U^I((x_i)_{i \in I})| \leq |U_k(x_k)| = 0$, and (4.24) also holds.

The "if" part. Let X have α-stable margins and a homogeneous Lévy copula. Then the marginal tail integrals of X satisfy (4.34), and since by Lemma 4.7, the Lévy measure of every set of the form $[a, b)$ can be expressed as a linear combination of tail integrals, the Lévy measure of X has the property (4.33) and we conclude by Theorem 14.3 in [87] that X is α-stable. □

Chapter 5

Applications of Lévy copulas

To apply Lévy copulas to multidimensional financial problems with dependence between assets, three kinds of tools are required:

1. Parametric families of Lévy copulas. Parsimonious models are needed because one typically does not have enough information about the dependence to estimate many parameters or proceed with a non-parametric approach.

2. Algorithms allowing to compute various quantities within a Lévy copula model (e.g., option prices, risk measures etc.)

3. Estimation methods for Lévy copula models.

In this chapter we give an answer to the first two questions. Section 5.1 describes several methods to construct parametric families of Lévy copulas and gives examples of such families. In Section 5.2 we show how Lévy copulas can be used to simulate multidimensional Lévy processes with dependence between components. This simulation algorithm enables us to compute the quantities of interest using the Monte Carlo method. Estimation methods for Lévy copulas are the topic of our current research and we do not discuss them here.

A fundamental advantage of the Lévy copula approach compared to ordinary copulas is the possibility to work with several time scales at the same time. When the time scale is fixed, the dependence structure of returns can be described with an ordinary copula. However, on a different time scale the copula will not be the same (cf. Example 4.1). On the other hand, a

Lévy copula of a Lévy process is not linked to any given time scale; it describes the dependence structure of the entire process.

One example of a financial modelling problem where Lévy copulas naturally appear is the pricing of multi-asset options on underlyings, described by exponential Lévy models. In this setting the parameters of the marginal Lévy processes can be calibrated from the prices of European options, quoted at the market, and the dependence structure will typically be estimated from the historical time series of returns. Since the sampling rate of returns is different from the maturity of traded options as well as from the maturity of the basket option that one wants to price, ordinary copulas are not well suited for this problem. Pricing of basket options in a Lévy copula model is discussed in more detail in Section 5.3.

Lévy copulas can also be useful outside the realm of financial modelling. Other contexts where modelling dependence in jumps is required are portfolios of insurance claims and models of operational risk.

Consider an insurance company with two subsidiaries, in France and in Germany. The aggregate loss process of the French subsidiary is modelled by the subordinator $\{X_t\}_{t\geq 0}$ and the loss process of the German one is $\{Y_t\}_{t\geq 0}$. The nature of processes X and Y may be different because the subsidiaries may not be working in the same sector and many risks that cause losses are local. However, common risks like floods and pan-European windstorms will lead to a certain degree of dependence between the claims. In this setting it is convenient to model the dependence between X and Y using a Lévy copula on $[0,\infty]^2$. In this modelling approach, the two-dimensional Lévy measure of (X,Y) is known and the overall loss distribution and ruin probability can be computed.

Another example where jump processes naturally appear is given by models of operational risk. The 2001 Basel agreement defines the operational risk as "the risk of direct and indirect loss resulting from inadequate or failed internal processes, people and systems or from external events" and allows banks to use internal loss data to compute regulatory capital requirements. Taking into account the dependence between different business lines in this computation, due to a diversification effect, may lead to substantial reduction of regulatory capital [14]. Aggregate loss processes from different business lines can be dynamically modelled by subordinators and the dependence between them can be accounted for using a Lévy copula on $[0,\infty]^2$.

The applications of Lévy copulas are not limited to constructing exponential Lévy models. Lévy copula models can be time changed to obtain multidimensional analogs of stochastic volatility models discussed in [22]. More generally, since a large class of Markov processes or even semimartingales behaves locally as a Lévy process in the sense that its dynamics can be described by a drift rate, a covariance matrix, and a Lévy measure, which may all change randomly through time (cf. e.g. [54], II.2.9, II.4.19), Lévy copulas could be used to describe dependence between processes of these types.

5.1 Parametric families of Lévy copulas

Our first result is a method to construct Lévy copulas on $[0,\infty]^d$ from ordinary copulas.

Theorem 5.1 (Construction of Lévy copulas from ordinary copulas). *Let C be a copula on $[0,1]^d$ and $\phi : [0,1] \to [0,\infty]$ be a strictly increasing continuous function with $\phi(1) = \infty$, $\phi(0) = 0$, having nonnegative derivatives of orders up to d on $(0,1)$. Then*

$$F(u_1,\dots,u_d) := \phi(C(\phi^{-1}(u_1),\dots,\phi^{-1}(u_d)))$$

is a Lévy copula on $[0,\infty]^d$.

Proof. First, note that ϕ^{-1} is well defined and satisfies $\phi^{-1}(0) = 0$ and $\phi^{-1}(\infty) = 1$. Therefore, properties 1 and 2 of Definition 4.14 are clear, and in view of Equation (4.5), property 4 also holds. It remains to show that F is a d-increasing function, and since ϕ^{-1} is strictly increasing, it suffices to prove that the function $\phi(C(u_1,\dots,u_d))$ is d-increasing on $[0,1]^d$.

To this end, let us show by induction on d that if $H : [0,1]^d \to [0,1]$ is d-increasing and grounded and $\psi : [0,1] \to [0,\infty]$ is a continuous increasing function with $\psi(0) = 0$ and positive derivatives of orders up to d on $(0,1)$ then $\psi(H)$ is also d-increasing and grounded. For $d = 1$, the result is clear. Suppose $d \geq 2$. For $k = 1,\dots,d$ let $a_k, b_k \in [0,1]$ with $a_k \leq b_k$. The function $\tilde{H}(u_2,\dots,u_d) := \psi(H(b_1,u_2,\dots,u_d)) - \psi(H(a_1,u_2,\dots,u_d))$ satisfies

$$V_{\psi(H)}(|a_1,b_1| \times \cdots \times |a_d,b_d|) = V_{\tilde{H}}(|a_2,b_2| \times \cdots \times |a_d,b_d|),$$

hence, it remains to prove that $\tilde{H}$ is $d-1$-increasing. It can be represented as follows ("$*$"

stands for $u_2, \dots, u_d$):

$$\begin{aligned}\tilde{H}(*) &= \int_{H(a_1,*)}^{H(b_1,*)} \psi'(t)dt \\ &= (H(b_1,*) - H(a_1,*)) \int_0^1 \psi'(H(a_1,*) + t(H(b_1,*) - H(a_1,*)))dt \\ &= (H(b_1,*) - H(a_1,*))\psi'_+(0) \\ &+ (H(b_1,*) - H(a_1,*)) \int_0^1 \{\psi'(H(a_1,*) + t(H(b_1,*) - H(a_1,*))) - \psi'_+(0)\}dt.\end{aligned}$$

Note that $\psi'_+(0)$, the right derivative of ψ at 0, is well defined because we have supposed that $d \geq 2$ and therefore ψ'' exists on $(0,1)$ and ψ' is increasing. In the right-hand side of the above equation:

- The integrand is $d-1$-increasing by the induction hypothesis, because $H(a_1,*)+t(H(b_1,*)-H(a_1,*))$ is $d-1$-increasing and grounded for every t and $\tilde{\psi}(t) := \psi'(t) - \psi'_+(0)$ is an increasing continuous function with $\tilde{\psi}(0) = 0$ and nonnegative derivatives of orders up to $d-1$ on $(0,1)$.
- The second term is $d-1$-increasing by Lemma 4.3,
- The first term is $d-1$-increasing because H is,

hence $\tilde{H}$ is $d-1$-increasing as sum of $d-1$-increasing functions and the proof is complete. □

One example of an absolutely monotonic (with positive derivatives of all orders) function that maps $[0,1]$ into $[0,\infty]$ is $\phi(x) = \frac{x}{1-x}$.

The following result allows to construct Lévy copulas on $(-\infty,\infty]^d$, analogous to the Archimedean copulas (cf. [76]). It can be used to build parametric families of Lévy copulas in arbitrary dimension, where the number of parameters does not depend on the dimension.

Theorem 5.2 (Archimedean Lévy copulas). *Let $\phi : [-1,1] \to [-\infty,\infty]$ be a strictly increasing continuous function with $\phi(1) = \infty$, $\phi(0) = 0$, and $\phi(-1) = -\infty$, having derivatives of orders up to d on $(-1,0)$ and $(0,1)$, and satisfying*

$$\frac{d^d\phi(e^x)}{dx^d} \geq 0, \quad \frac{d^d\phi(-e^x)}{dx^d} \leq 0, \quad x \in (-\infty,0). \tag{5.1}$$

Let

$$\tilde{\phi}(u) := 2^{d-2}\{\phi(u) - \phi(-u)\}$$

for $u \in [-1,1]$. Then

$$F(u_1,\dots,u_d) := \phi\left(\prod_{i=1}^{d} \tilde{\phi}^{-1}(u_i)\right), \quad (u_1,\dots,u_d) \in (-\infty,\infty]^d$$

defines a Lévy copula.

Proof. Firstly, note that $\tilde{\phi}$ is a strictly increasing continuous function from $[-1,1]$ to $[-\infty,\infty]$, satisfying $\tilde{\phi}(1) = \infty$ and $\tilde{\phi}(-1) = -\infty$, which means that $\tilde{\phi}^{-1}$ exists for all $u \in (-\infty,\infty]$ and F is well defined. Properties 1 and 2 of Definition 4.16 are clearly satisfied. For $k = 1,\dots,d$ and $u_k \in \mathbb{R}$ we have

$$\begin{aligned}
F^{\{k\}}(u_k) &= \lim_{c\to\infty} \sum_{(u_i)_{i\neq k}\in\{-c,\infty\}^{d-1}} F(u_1,\dots,u_d) \prod_{i\neq k} \operatorname{sgn} u_i \\
&= \sum_{(u_i)_{i\neq k}\in\{-\infty,\infty\}^{d-1}} \phi\left(\tilde{\phi}^{-1}(u_k)\prod_{i\neq k}\operatorname{sgn} u_i\right)\prod_{i\neq k}\operatorname{sgn} u_i \\
&= \sum_{i=0}^{d-1}\binom{d-1}{i}(-1)^i \phi\left(\tilde{\phi}^{-1}(u_k)(-1)^i\right) \\
&= 2^{d-2}\{\phi(\tilde{\phi}^{-1}(u_k)) - \phi(-\tilde{\phi}^{-1}(u_k))\} = u_k,
\end{aligned}$$

which proves property 4. It remains to show that F is d-increasing. Since $\tilde{\phi}^{-1}$ is increasing, we only need to show that $(u_1,\dots,u_d) \mapsto \phi(\prod_{i=1}^d u_i)$ is d-increasing on $(-1,1]^d$, and for this, because $\phi(\prod_{i=1}^d u_i) = \phi(\prod_{i=1}^d |u_i| \prod_{i=1}^d \operatorname{sgn} u_i)$, it suffices to prove that both $(u_1,\dots,u_d) \mapsto \phi(\prod_{i=1}^d u_i)$ and $(u_1,\dots,u_d) \mapsto -\phi(-\prod_{i=1}^d u_i)$ are d-increasing on $[0,1]^d$ or, equivalently, on $(0,1)^d$ (since ϕ is continuous). The first condition of (5.1) implies that

$$\frac{\partial^d \psi(z_1,\dots,z_d)}{\partial z_1 \dots \partial z_d} \geq 0$$

on $(-\infty,0)^d$ for $\psi(z_1,\dots,z_d) := \phi(e^{z_1+\dots+z_d})$. From Definition 4.5 it follows easily that

$$V_\psi(B) = \int_B \frac{\partial^d \psi(z_1,\dots,z_d)}{\partial z_1 \dots \partial z_d} dz_1 \dots dz_d.$$

Therefore, ψ is increasing on $(-\infty,0)^d$, which implies that $(u_1,\dots,u_d) \mapsto \phi(\prod_{i=1}^d u_i)$ is d-increasing on $(0,1)^d$. The second condition of (5.1) entails similarly that $(u_1,\dots,u_d) \mapsto -\phi(-\prod_{i=1}^d u_i)$ is d-increasing on $(0,1)^d$ as well. □

Remark 5.1. Condition (5.1) is satisfied in particular if for any $k = 1,\dots,d$,

$$\frac{d^k\phi(u)}{du^k} \geq 0, \quad u \in (0,1) \quad \text{and} \quad (-1)^k \frac{d^k\phi(u)}{du^k} \leq 0, \quad u \in (-1,0).$$

Corollary 5.1 (Archimedean Lévy copulas for spectrally positive processes). *Let $\psi : [0,\infty] \to [0,\infty]$ be a strictly decreasing continuous function with $\psi(0) = \infty$, $\psi(\infty) = 0$, having derivatives of orders up to d on $(0,\infty)$ and satisfying*

$$(-1)^d \frac{d^d \psi(x)}{dx^d} \geq 0$$

for all $x \in (0,\infty)$. Then

$$F(u_1,\dots,u_d) := \psi(\sum_{i=1}^{d} \psi^{-1}(u_i)) \tag{5.2}$$

is a Lévy copula on $[0,\infty]^d$.

Proof. Let

$$\phi(u) = \begin{cases} \psi(-\log u), & u \geq 0 \\ -\psi(-\log(-u)), & u < 0 \end{cases}, \qquad u \in [-1,1].$$

Then ϕ satisfies the conditions of Theorem 5.2 and therefore

$$\bar{F}(u_1,\dots,u_d) := \phi\left(\prod_{i=1}^{d} \phi^{-1}(u_i/2^{d-1})\right)$$

is a Lévy copula on $(-\infty,\infty]^d$. This implies that the function $\tilde{F} := \bar{F}|_{[0,\infty]^d}$ has properties 1, 2 and 3 of Definition 4.14. However, it is easy to check that $\tilde{F}(u_1,\dots,u_d)|_{u_i=\infty, i\neq k} = u_k/2^{d-1}$ and therefore

$$2^{d-1}\tilde{F}(u_1,\dots,u_d) = 2^{d-1}\psi(\sum_{i=1}^{d} \psi^{-1}(u_i/2^{d-1}))$$

is a Lévy copula on $[0,\infty]^d$, which means that (5.2) also defines a positive Lévy copula. □

Example 5.1. Let

$$\phi(x) := \eta(-\log|x|)^{-1/\theta} 1_{x\geq 0} - (1-\eta)(-\log|x|)^{-1/\theta} 1_{x<0}$$

with $\theta > 0$ and $\eta \in (0,1)$. Then

$$\tilde{\phi}(x) = 2^{d-2}(-\log|x|)^{-1/\theta}\operatorname{sgn} x, \quad \text{and}$$
$$\tilde{\phi}^{-1}(u) = e^{-|2^{2-d}u|^{-\theta}}\operatorname{sgn} u,$$

and therefore

$$F(u_1,\dots,u_d) = 2^{2-d}\left(\sum_{i=1}^{d}|u_i|^{-\theta}\right)^{-1/\theta}(\eta 1_{u_1\cdots u_d\geq 0} - (1-\eta)1_{u_1\cdots u_d<0}) \tag{5.3}$$

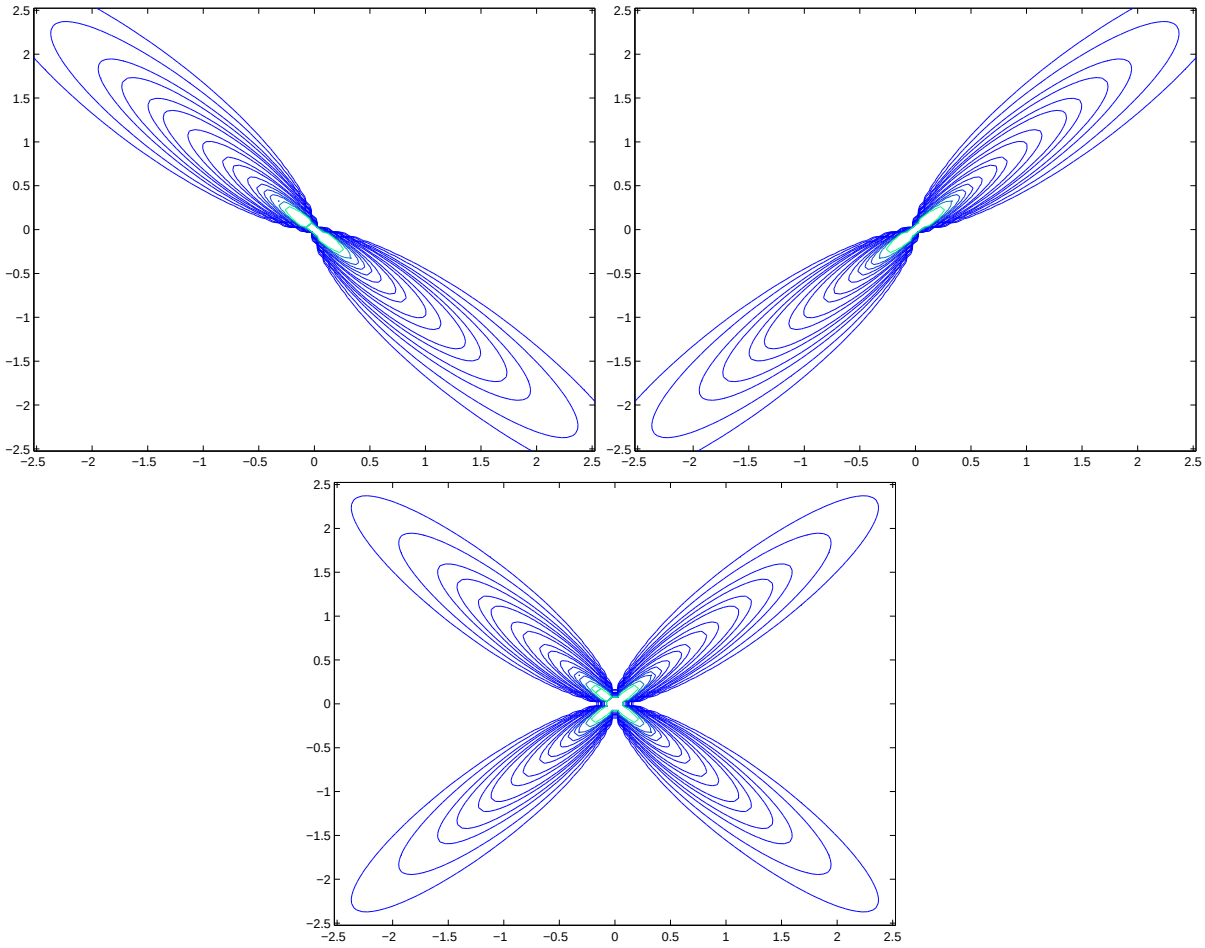

Figure 5.1: Contour plots of Lévy density with 1-stable margins and Lévy copula (5.4) with $\theta = 5$ and different values of η. Top left: $\eta = 0$. Top right: $\eta = 1$. Bottom: $\eta = 0.5$. On each curve the Lévy density $\nu(x, y)$ is constant and it increases from outer curves to inner curves.

defines a two-parameter family of Lévy copulas (F is in fact a Lévy copula for all $\theta > 0$ and $\eta \in [0, 1]$). The role of parameters is most easy to analyze in the case $d = 2$, when (5.3) becomes

$$F(u, v) = (|u|^{-\theta} + |v|^{-\theta})^{-1/\theta}(\eta 1_{uv \geq 0} - (1 - \eta)1_{uv<0}). \tag{5.4}$$

To explain the meaning of the parameters θ and η, we have drawn the contour plots of a Lévy density with 1-stable margins and the dependence structure given by F above. Figure 5.1 shows that the parameter η determines the dependence of the *sign* of jumps: when $\eta = 1$, the two components always jump in the same direction, and when $\eta = 0$, positive jumps in one component are accompanied by negative jumps in the other and vice versa. For intermediate

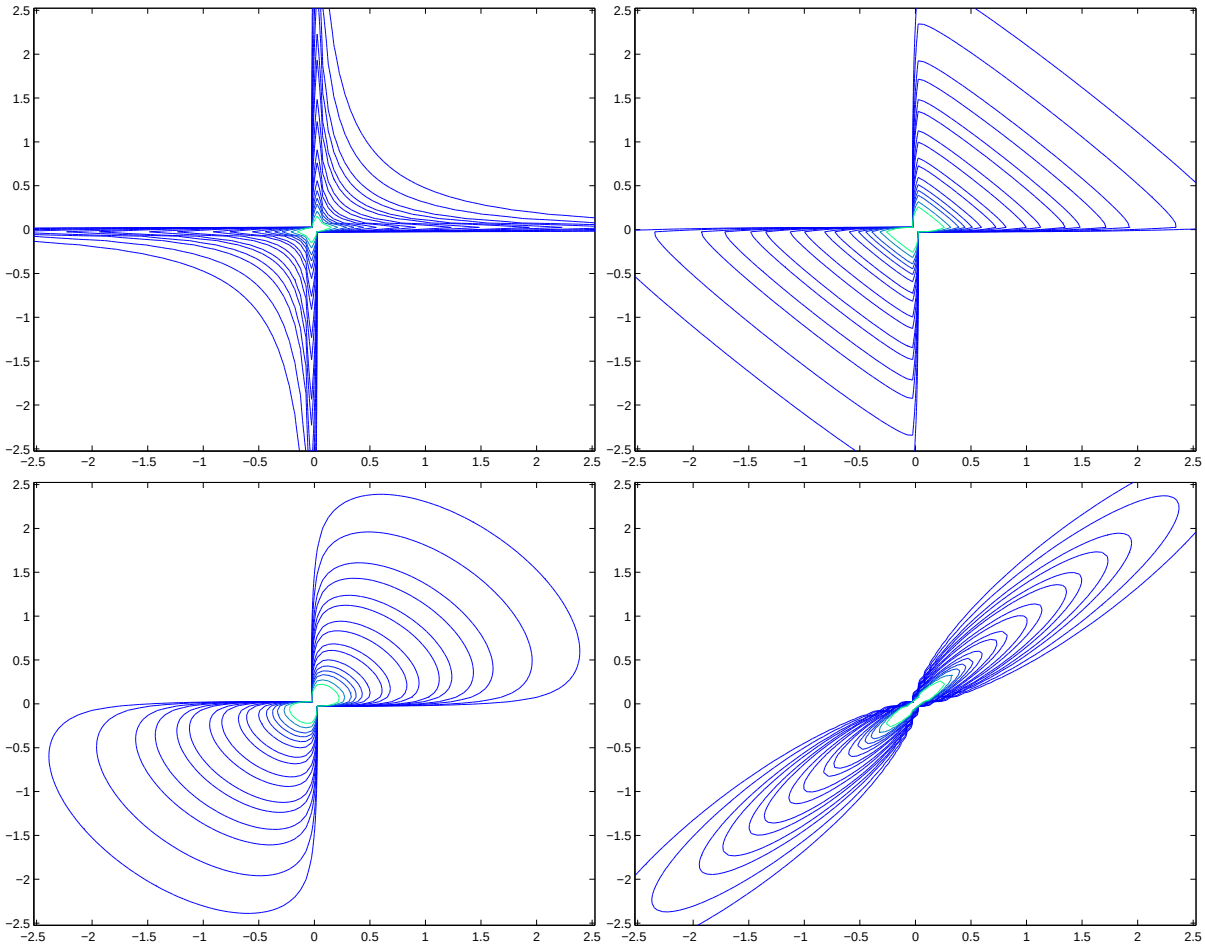

Figure 5.2: Contour plots of Lévy density with 1-stable margins and Lévy copula (5.4) with $\eta = 1$. Top left: $\theta = 0.2$. Top right: $\theta = 0.7$. Bottom left: $\theta = 1$. Bottom right: $\theta = 5$. On each curve the Lévy density $\nu(x, y)$ is constant and it increases from outer curves to inner curves.

values of η, positive jumps in one component can correspond to both positive and negative jumps in the other component. The parameter θ is responsible for the dependence of absolute values of jumps in different components: from Figure 5.2 it is seen that as θ grows from 0.2 to 5, the dependence structure changes from independence to almost complete dependence.

Theorem 5.2 can be used to construct parsimonious models of dependence; this is typically useful when one has little information about the dependence structure of the problem. If a more precise vision is necessary, a possible strategy is to model the dependence in each corner of the Lévy measure separately. A d-dimensional Lévy copula can be constructed from 2^d positive

Lévy copulas (one for each orthant) as follows:

Theorem 5.3. *For each* $\{\alpha_1, \ldots, \alpha_d\} \in \{-1, 1\}^d$ *let* $g^{(\alpha_1,\ldots,\alpha_d)}(u) : [0, \infty] \to [0, 1]$ *be a nonnegative, increasing function satisfying*

$$\sum_{\alpha\in\{-1,1\}^d \text{ with } \alpha_k=-1} g^{(\alpha_1,\ldots,\alpha_d)}(u) = 1 \quad \text{and} \quad \sum_{\alpha\in\{-1,1\}^d \text{ with } \alpha_k=1} g^{(\alpha_1,\ldots,\alpha_d)}(u) = 1$$

for all $u \in [0, \infty]$ *and all* $k \in \{1, \ldots, d\}$. *Moreover, let* $F^{(\alpha_1,\ldots,\alpha_d)}$ *be a positive Lévy copula that satisfies the following continuity property at infinity: for all* $I \subset \{k : \alpha_k = -1\}$, $(u_i)_{i\in I^c} \in [0, \infty]^{I^c}$ *we have*

$$\lim_{\{u_i\}_{i\in I}\to(\infty,\ldots,\infty)} F^{(\alpha_1,\ldots,\alpha_d)}(u_1, \ldots, u_d) = F^{(\alpha_1,\ldots,\alpha_d)}(v_1, \ldots, v_d),$$

where $v_i = u_i$ *for* $i \in I^c$ *and* $v_i = \infty$ *otherwise. Then*

$$F(u_1, \ldots, u_d) := F^{(\operatorname{sgn} u_1,\ldots,\operatorname{sgn} u_d)}\left(|u_1|g^{(\operatorname{sgn} u_1,\ldots,\operatorname{sgn} u_d)}(|u_1|), \ldots, |u_d|g^{(\operatorname{sgn} u_1,\ldots,\operatorname{sgn} u_d)}(|u_d|)\right)\prod_{i=1}^{d}\operatorname{sgn} u_i$$

defines a Lévy copula.

Proof. Properties 1 and 2 of Definition 4.16 are obvious. Property 3 follows after observing that $u \mapsto ug^{(\alpha_1,\ldots,\alpha_d)}(u)$ is increasing on $[0, \infty]$ for any $\{\alpha_1, \ldots, \alpha_d\} \in \{-1, 1\}^d$. To prove property 4, note that

$$F^{(\alpha_1,\ldots,\alpha_d)}\left(|u_1|g^{(\alpha_1,\ldots,\alpha_d)}(|u_1|), \ldots, |u_d|g^{(\alpha_1,\ldots,\alpha_d)}(|u_d|)\right) = |u_k|g^{(\alpha_1,\ldots,\alpha_d)}(|u_k|)$$

for any $\{\alpha_1, \ldots, \alpha_d\} \in \{-1, 1\}^d$ and any $\{u_1, \ldots, u_d\} \in \bar{\mathbb{R}}^d$ with $u_i = \infty$ for all $i \neq k$. Therefore,

$$\begin{aligned} F^{\{k\}}(u) &= \begin{cases} \displaystyle\sum_{\alpha\in\{-1,1\}^d \text{ with } \alpha_k=1} ug^{(\alpha_1,\ldots,\alpha_d)}(u)\prod_{i=1}^{d}\alpha_i\prod_{j\neq k}\alpha_j & \text{if } u \geq 0 \\ \displaystyle\sum_{\alpha\in\{-1,1\}^d \text{ with } \alpha_k=-1} |u|g^{(\alpha_1,\ldots,\alpha_d)}(|u|)\prod_{i=1}^{d}\alpha_i\prod_{j\neq k}\alpha_j & \text{if } u < 0 \end{cases} \\ &= u. \end{aligned}$$

□

Example 5.2. Let

$$g^{(\alpha_1,\ldots,\alpha_d)}(u) = \begin{cases} 1 & \text{for } \alpha_1 = \cdots = \alpha_d \\ 0 & \text{otherwise.} \end{cases}$$

Then the Lévy copula F in Theorem 5.3 satisfies $F(u_1,\ldots,u_d) = 0$ if $u_i u_j < 0$ for some i, j. This means that the Lévy measure is supported by the positive and the negative orthant: either all components of the process jump up or all components jump down.

5.2 Simulation of multidimensional dependent Lévy processes

Lévy copulas turn out to be a convenient tool for simulating multidimensional Lévy processes with specified dependence. In this section we first give the necessary definitions and two auxiliary lemmas and then prove two theorems which show how multidimensional Lévy processes with dependence structures given by Lévy copulas can be simulated in the finite variation case (Theorem 5.6) and in the infinite variation case (Theorem 5.7).

To simulate a Lévy process $\{X_t\}_{0\leq t\leq 1}$ on $\mathbb{R}^d$ with Lévy measure ν, we will first simulate a Poisson random measure on $[0,1]\times\mathbb{R}^d$ with intensity measure $\lambda_{[0,1]}\otimes\nu$, where λ denotes the Lebesgue measure. The Lévy process can then be constructed via the Lévy-Itô decomposition.

Let F be a Lévy copula on $(-\infty,\infty]^d$ such that for every $I\in\{1,\ldots,d\}$ nonempty,

$$\lim_{(x_i)_{i\in I}\to\infty} F(x_1,\ldots,x_d) = F(x_1,\ldots,x_d)|_{(x_i)_{i\in I}=\infty}. \tag{5.5}$$

This Lévy copula defines a positive measure μ on $\mathbb{R}^d$ with Lebesgue margins such that for each $a, b\in\mathbb{R}^d$ with $a\leq b$,

$$V_F(|a,b|) = \mu((a,b]). \tag{5.6}$$

For a one-dimensional tail integral U, the inverse tail integral $U^{(-1)}$ was defined in Equation (4.25). In the sequel we will need the following technical lemma.

Lemma 5.4. *Let ν be a Lévy measure on $\mathbb{R}^d$ with marginal tail integrals U_i, $i=1,\ldots,d$, and Lévy copula F on $(-\infty,\infty]^d$, satisfying* (5.5)*, let μ be defined by* (5.6) *and let*

$$f:(u_1,\ldots,u_d)\mapsto(U_1^{(-1)}(u_1),\ldots,U_d^{(-1)}(u_d)).$$

Then ν is the image measure of μ by f.

Proof. We must prove that for each $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\nu(A) = \mu(\{u \in \mathbb{R}^d : f(u) \in A\}),$$

but in view of Lemma 4.7, it is sufficient to show that for each $I \subset \{1, \dots, d\}$ nonempty and for all $(x_i)_{i \in I} \in (\mathbb{R} \setminus \{0\})^{|I|}$,

$$U^I((x_i)_{i \in I}) = \mu(\{u \in \mathbb{R}^d : U_i^{(-1)}(u_i) \in \mathcal{I}(x_i), i \in I\}),$$

where $\mathcal{I}(x)$ was defined in (4.21). However, since U_i is left-continuous, for every i, $U_i^{(-1)}(u) \in \mathcal{I}(x)$ if and only if $u \in (U_i(x) \wedge 0, U_i(x) \vee 0]$. Therefore,

$$\begin{aligned}\mu(\{u \in \mathbb{R}^d : U_i^{(-1)}(u_i) \in \mathcal{I}(x_i), i \in I\}) \\ = \mu(\{u \in \mathbb{R}^d : u_i \in (U_i(x_i) \wedge 0, U_i(x_i) \vee 0], i \in I\}) = F^I((U_i(x_i))_{i \in I}),\end{aligned}$$

and an application of Theorem 4.8 completes the proof. □

By Theorem 2.28 in [1], there exists a family, indexed by $\xi \in \mathbb{R}$, of positive Radon measures $K(\xi, dx_2 \cdots dx_d)$ on $\mathbb{R}^{d-1}$, such that

$$\xi \mapsto K(\xi, dx_2 \cdots dx_d)$$

is Borel measurable and

$$\mu(dx_1 \dots dx_d) = \lambda(dx_1) \otimes K(x_1, dx_2 \cdots dx_d). \tag{5.7}$$

In addition, $K(\xi, \mathbb{R}^{d-1}) = 1$ λ-almost everywhere, that is, $K(\xi, *)$ is, almost everywhere, a probability distribution. In the sequel we will call $\{K(\xi, *)\}_{\xi \in \mathbb{R}}$ the family of conditional probability distributions associated to the Lévy copula F.

Let F_ξ be the distribution function of the measure $K(\xi, *)$:

$$F_\xi(x_2, \dots, x_d) := K(\xi, (-\infty, x_2] \times \cdots \times (-\infty, x_d]), \quad (x_2, \dots, x_d) \in \mathbb{R}^{d-1}. \tag{5.8}$$

The following lemma shows that it can be computed in a simple manner from the Lévy copula F.

Lemma 5.5. *Let F be a Lévy copula on $(-\infty,\infty]^d$, satisfying (5.5), and F_ξ be the corresponding conditional distribution function, defined by (5.8). Then, there exists $N \subset \mathbb{R}$ with $\lambda(N) = 0$ such that for every fixed $\xi \in \mathbb{R} \setminus N$, $F_\xi(*)$ is a probability distribution function, satisfying*

$$F_\xi(x_2,\dots,x_d) = \operatorname{sgn}(\xi)\frac{\partial}{\partial \xi} V_F((\xi \wedge 0, \xi \vee 0] \times (-\infty, x_2] \times \cdots \times (-\infty, x_d]) \tag{5.9}$$

in every point $(x_2,\dots,x_d)$, where F_ξ is continuous.

Remark 5.2. Since the law of a random variable is completely determined by the values of its distribution function at the continuity points of the latter, being able to compute F_ξ at all points where it is continuous is sufficient for all practical purposes.

Proof. Since it has already been observed that $K(\xi, \mathbb{R}^{d-1}) = 1$ λ-almost everywhere, we only need to prove the second part of the lemma. Let

$$G(x_1,\dots,x_d) := \operatorname{sgn} x_1 V_F((x_1 \wedge 0, x_1 \vee 0] \times (-\infty, x_2] \times \cdots \times (-\infty, x_d])$$

By Theorem 2.28 in [1], for each $f \in L^1(\mathbb{R}^d, \mu)$,

$$\int_{\mathbb{R}^d} f(x_1,\dots,x_d)\mu(dx_1 \cdots dx_d) = \int_{-\infty}^{\infty} dx_1 \int_{\mathbb{R}^d} f(x_1,\dots,x_d)K(x_1, dx_2 \cdots dx_d), \tag{5.10}$$

which implies that

$$G(x_1,\dots,x_d) = \operatorname{sgn} x_1 \int_{(x_1\wedge 0, x_1 \vee 0]} d\xi F_\xi(x_2,\dots,x_d),$$

Therefore, for fixed $(x_2,\dots,x_d)$, (5.9) holds ξ-almost everywhere. Since a union of countably many sets of zero measure is again a set of zero measure, there exists a set $N \subset \mathbb{R}$ with $\lambda(N) = 0$ such that for every $\xi \in \mathbb{R} \setminus N$, (5.9) holds for all $(x_2,\dots,x_d) \in \mathbb{Q}^d$, where $\mathbb{Q}$ denotes the set of rational numbers.

Fix $\xi \in \mathbb{R} \setminus N$ and let $x \in \mathbb{R}^{d-1}$ and $\{x_n^+\}$ and $\{x_n^-\}$ be two sequences of $d-1$-dimensional vectors with coordinates in $\mathbb{Q}$, converging to x from above and from below (componentwise). Since F_ξ is increasing in each coordinate (as a probability distribution function), the limits $\lim_n F_\xi(x_n^+)$ and $\lim_n F_\xi(x_n^-)$ exist. Suppose that

$$\lim_n F_\xi(x_n^+) = \lim_n F_\xi(x_n^-) = F^* \tag{5.11}$$

and observe that for every $\delta \neq 0$,

$$\frac{G(\xi+\delta, x_n^-) - G(\xi, x_n^-)}{\delta} \leq \frac{G(\xi+\delta, x) - G(\xi, x)}{\delta} \leq \frac{G(\xi+\delta, x_n^+) - G(\xi, x_n^+)}{\delta}.$$

For every $\varepsilon > 0$, in view of (5.11), there exists N_0 such that for every $n \geq N_0$, $F_\xi(x_n^+) - F^* \leq \varepsilon/2$ and $F^* - F_\xi(x_n^-) \leq \varepsilon/2$. Since G is differentiable with respect to the first variable at points (ξ, x_n^+) and (ξ, x_n^-), we can choose δ small enough so that

$$\left| \frac{G(\xi + \delta, x_n^-) - G(\xi, x_n^-)}{\delta} - F_\xi(x_n^-) \right| \leq \varepsilon/2$$

and

$$\left| \frac{G(\xi + \delta, x_n^+) - G(\xi, x_n^+)}{\delta} - F_\xi(x_n^+) \right| \leq \varepsilon/2$$

This proves that

$$\lim_{\delta \to 0} \frac{G(\xi + \delta, x) - G(\xi, x)}{\delta} = F^*.$$

We have thus shown that F_ξ satisfies Equation (5.9) in all points where (5.11) holds, that is, where F_ξ is continuous. □

In the following two theorems we show how Lévy copulas may be used to simulate multidimensional Lévy processes with specified dependence. Our results can be seen as an extension to Lévy processes, represented by Lévy copulas, of the series representation results, developed by Rosinski and others (see [84] and references therein). The first result concerns the simpler case when the Lévy process has finite variation on compacts.

Theorem 5.6. (Simulation of multidimensional Lévy processes, finite variation case) *Let ν be a Lévy measure on $\mathbb{R}^d$, satisfying $\int (|x| \wedge 1)\nu(dx) < \infty$, with marginal tail integrals U_i, $i = 1, \ldots, d$ and Lévy copula $F(x_1, \ldots, x_d)$, such that the condition (5.5) is satisfied, and let $K(x_1, dx_2 \cdots dx_d)$ be the corresponding conditional probability distributions, defined by (5.8). Let $\{V_i\}$ be a sequence of independent random variables, uniformly distributed on $[0, 1]$. Introduce d random sequences $\{\Gamma_i^1\}, \ldots, \{\Gamma_i^d\}$, independent from $\{V_i\}$ such that*

- *$N = \sum_{i=1}^\infty \delta_{\{\Gamma_i^1\}}$ is a Poisson random measure on $\mathbb{R}$ with intensity measure λ.*
- *Conditionally on Γ_i^1, the random vector $(\Gamma_i^2, \ldots, \Gamma_i^d)$ is independent from Γ_j^k with $j \neq i$ and all k and is distributed on $\mathbb{R}^{d-1}$ with law $K(\Gamma_i^1, dx_2 \cdots dx_d)$.*

Then

$$\{Z_t\}_{0 \leq t \leq 1} \quad \text{where} \quad Z_t^k = \sum_{i=1}^\infty U_i^{(-1)}(\Gamma_i^k) 1_{[0,t]}(V_i), \quad k = 1, \ldots, d, \tag{5.12}$$

is a Lévy process on the time interval $[0,1]$ *with characteristic function*

$$e^{i\langle u,Z_t\rangle} = \exp\left(t\int_{\mathbb{R}^d}(e^{i\langle u,z\rangle}-1)\nu(dz)\right). \tag{5.13}$$

Remark 5.3. The probability distribution function of $(\Gamma_i^2,\dots,\Gamma_i^d)$ conditionally on Γ_i^1 is known from Lemma 5.9.

Remark 5.4. The sequence $\{\Gamma_i^1\}_{i\geq 1}$ can be constructed, for example, as follows. Let $\{X_i\}_{i\geq 1}$ be a sequence of jump times of a Poisson process with jump intensity equal to 2. Then it is easy to check that one can define Γ_i^1 by $\Gamma_i^1 = X_i(-1)^i$.

Proof. First note that $\{\Gamma_i^k\}$ are well defined since by Lemma 5.5, $K(x_1,*)$ is a probability distribution for almost all x_1. Let

$$Z_{\tau,t}^k = \sum_{-\tau\leq\Gamma_i^1\leq\tau} U_k^{(-1)}(\Gamma_i^k)1_{V_i\leq t}, \quad k=1,\dots,d.$$

By Proposition 3.8 in [82],

$$Z_{\tau,t}^k = \int_{[0,t]\times[-\tau,\tau]\times\mathbb{R}^{d-1}} U_k^{(-1)}(x_k)M(ds\times dx_1\cdots dx_d),$$

where M is a Poisson random measure on $[0,1]\times\mathbb{R}^d$ with intensity measure $\lambda_{[0,1]}(dt)\otimes\mu(dx_1\cdots dx_d)$, and the measure μ was defined in Equation (5.6).

By Lemma 5.4 and Proposition 3.7 in [82],

$$Z_{\tau,t}^k = \int_{[0,t]\times\mathbb{R}^d} x_k N_\tau(ds\times dx_1\cdots dx_d), \tag{5.14}$$

for some Poisson random measure N_τ on $[0,1]\times\mathbb{R}^d$ with intensity measure $\lambda_{[0,1]}(ds)\otimes\nu_\tau(dx_1\cdots dx_d)$, where

$$\nu_\tau := 1_{(-\infty,U_1^{(-1)}(-\tau)]\cup[U_1^{(-1)}(\tau),\infty)}(x_1)\nu(dx_1\cdots dx_d) \tag{5.15}$$

The Lévy-Itô decomposition [87, Theorem 19.2] implies that $Z_{\tau,t}$ is a Lévy process on the time interval $[0,1]$ with characteristic function

$$e^{i\langle u,Z_{\tau,t}\rangle} = \exp\left(t\int_{\mathbb{R}^d}(e^{i\langle u,z\rangle}-1)\nu_\tau(dz)\right).$$

Let h be a bounded continuous truncation function such that $h(x)\equiv x$ on a neighborhood of 0. Since $\lim_{\tau\to\infty}U_1^{(-1)}(\tau)=0$ and $\lim_{\tau\to\infty}U_1^{(-1)}(-\tau)=0$, by dominated convergence,

$$\int_{\mathbb{R}^d}h^2(x)\nu_\tau(dx)\xrightarrow[\tau\to\infty]{}\int_{\mathbb{R}^d}h^2(x)\nu(dx) \qquad\text{and}\qquad \int_{\mathbb{R}^d}h(x)\nu_\tau(dx)\xrightarrow[\tau\to\infty]{}\int_{\mathbb{R}^d}h(x)\nu(dx).$$

Moreover, for every $f \in C_b(\mathbb{R}^d)$ such that $f(x) \equiv 0$ on a neighborhood of 0,

$$\int_{\mathbb{R}^d} f(x)\nu_\tau(dx) = \int_{\mathbb{R}^d} f(x)\nu(dx)$$

starting from a sufficiently large τ. Therefore, Proposition 1.7 allows to conclude that $\{Z_{\tau,t}\}_{0\le t\le 1}$ converges in law to a Lévy process with characteristic function given by (5.13). □

If the Lévy process has paths of infinite variation on compacts, it can no longer be represented as the sum of its jumps and we have to introduce a centering term into the series (5.12).

Theorem 5.7. (Simulation of multidimensional Lévy processes, infinite variation case)

Let ν be a Lévy measure on $\mathbb{R}^d$ with marginal tail integrals U_i, $i = 1, \dots, d$ and Lévy copula $F(x_1, \dots, x_d)$, such that the condition (5.5) *is satisfied. Let $\{V_i\}$ and $\{\Gamma_i^1\}, \dots, \{\Gamma_i^d\}$ be as in Theorem 5.6. Let*

$$A_k(\tau) = \int_{|x|\le 1} x_k \nu_\tau(dx_1 \cdots dx_d), \quad k = 1 \dots d,$$

where ν_τ is given by (5.15)*. Then the process*

$$\{Z_{\tau,t}\}_{0\le t\le 1}, \quad \textit{where} \quad Z^k_{\tau,t} = \sum_{-\tau\le\Gamma_i^1\le\tau} U_k^{(-1)}(\Gamma_i^k) 1_{V_i\le t} - tA_k(\tau),$$

converges in law as $\tau \to \infty$ to a Lévy process $\{Z_t\}_{0\le t\le 1}$ on the time interval $[0,1]$ with characteristic function

$$e^{i\langle u, Z_t\rangle} = \exp\left(t \int_{\mathbb{R}^d} (e^{i\langle u,z\rangle} - 1 - i\langle u, z\rangle) 1_{|z|\le 1} \nu(dz)\right). \tag{5.16}$$

Proof. The proof is essentially the same as in Theorem 5.6. Similarly to Equation (5.14), $Z^k_{\tau,s}$ can now be represented as

$$Z^k_{\tau,s} = \int_{[0,s]\times\{x\in\mathbb{R}^d:|x|\le 1\}} x_k \left\{N_\tau(ds \times dx_1 \cdots dx_d) - ds\nu_\tau(dx_1 \cdots dx_d)\right\} + \int_{[0,s]\times\{x\in\mathbb{R}^d:|x|>1\}} x_k N_\tau(ds \times dx_1 \cdots dx_d),$$

where N_τ is a Poisson random measure on $[0,1] \times \mathbb{R}^d$ with intensity measure $\lambda_{[0,1]}(ds) \otimes \nu_\tau$, and ν_τ is defined by (5.15). This entails that $Z_{\tau,s}$ is a Lévy process (compound Poisson) with

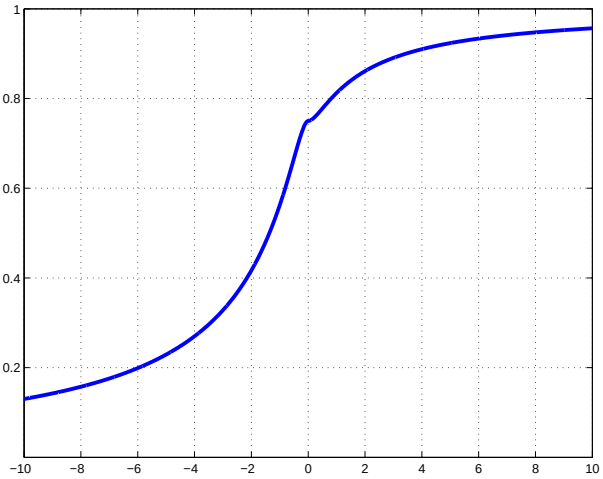

Figure 5.3: Conditional distribution function F_ξ, corresponding to the Lévy copula of Example 5.1 with $\xi = 1$, $\theta = 1$ and $\eta = 0.25$.

characteristic function

$$e^{i\langle u, Z_{\tau,t}\rangle} = \exp\left(t \int_{\mathbb{R}^d} (e^{i\langle u,z\rangle} - 1 - i\langle u, z\rangle 1_{|z|\leq 1})\nu_\tau(dz)\right).$$

Proposition 1.7 once again allows to conclude that $(Z_{\tau,s})_{0\leq s\leq 1}$ converges in law to a Lévy process with characteristic function (5.16). □

Example 5.3. Let $d = 2$ and F be the Lévy copula of Example 5.1. A straightforward computation yields:

$$F_\xi = \left\{(1-\eta) + \left(1 + \left|\frac{\xi}{x_2}\right|^\theta\right)^{-1-1/\theta} (\eta - 1_{x_2<0})\right\} 1_{\xi\geq 0} + \left\{\eta + \left(1 + \left|\frac{\xi}{x_2}\right|^\theta\right)^{-1-1/\theta} (1_{x_2\geq 0} - \eta)\right\} 1_{\xi<0}. \quad (5.17)$$

This distribution function is plotted in Figure 5.3. This conditional distribution function can be inverted analytically:

$$F_\xi^{-1}(u) = B(\xi, u)|\xi| \left\{C(\xi, u)^{-\frac{\theta}{\theta+1}} - 1\right\}^{-1/\theta}$$

with $B(\xi, u) = \operatorname{sgn}(u - 1 + \eta)1_{\xi\geq 0} + \operatorname{sgn}(u - \eta)1_{\xi<0}$

and $C(\xi, u) = \left\{\frac{u-1+\eta}{\eta} 1_{u\geq 1-\eta} + \frac{1-\eta-u}{1-\eta} 1_{u<1-\eta}\right\} 1_{\xi\geq 0} + \left\{\frac{u-\eta}{1-\eta} 1_{u\geq \eta} + \frac{\eta-u}{\eta} 1_{u<\eta}\right\} 1_{\xi<0}.$

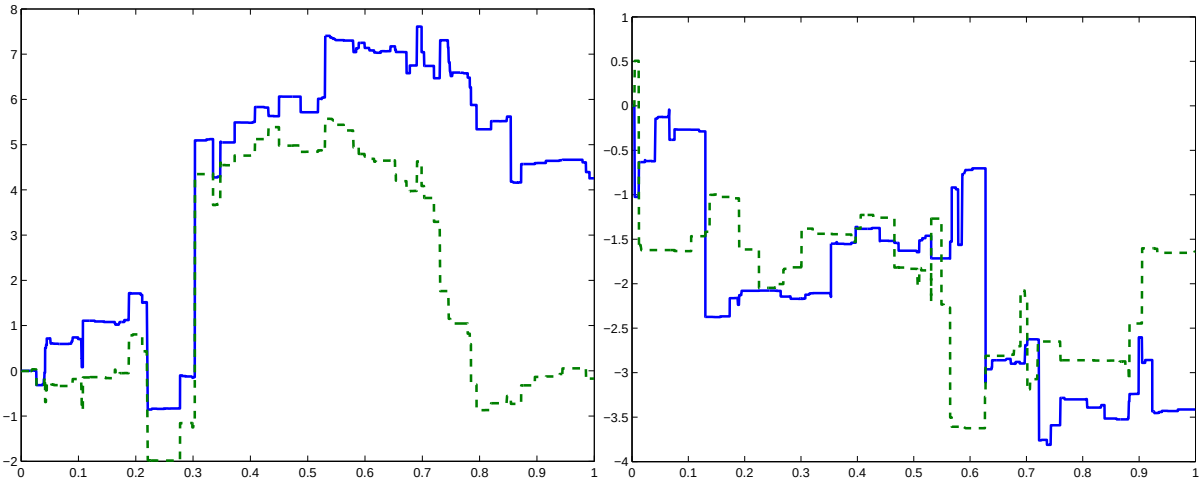

Figure 5.4: Trajectories of two variance gamma processes with dependence structure given by the Lévy copula of Example 5.1. In both graphs both variance gamma processes are driftless and have parameters $c = 10$, $\lambda_- = 1$ and $\lambda_+ = 1$ (cf. Equation (1.19)). In the left graph, the dependence between the two components is strong both in terms of sign and absolute value ($\eta = 0.9$ and $\theta = 3$): the processes jump mostly in the same direction and the sizes of jumps are similar. In the right graph the dependence of absolute values is weak ($\theta = 0.5$) and the dependence of jump sign is negative ($\eta = 0.25$).

If ν is a Lévy measure on $\mathbb{R}^2$, satisfying $\int (|x| \wedge 1)\nu(dx) < \infty$ with marginal tail integrals U_1, U_2 and Lévy copula F of Example 5.1, the Lévy process with characteristic function (5.13) can be simulated as follows. Let $\{V_i\}$ and $\{\Gamma_i^1\}$ be as in Theorem 5.6 and let $\{W_i\}$ be an independent sequence of independent random variables, uniformly distributed on $[0, 1]$. For each i, let $\Gamma_i^2 = F_{\Gamma_i^1}^{-1}(W_i)$. Then the Lévy process that we want to simulate is given by Equation (5.12).

Figure 5.4 shows the simulated trajectories of variance gamma processes with dependence structure given by the Lévy copula of Example 5.1 with different values of parameters. The number of jumps for each trajectory was limited to 2000 and the inverse tail integral of the variance gamma Lévy measure was computed by inverting numerically the exponential integral function (function `expint` available in MATLAB). Simulating two trajectories with 2000 jumps each takes about 1 second on a Pentium III computer running MATLAB, but this time could be reduced by several orders of magnitude if the inverse exponential integral function is tabulated and a lower-level programming language (e.g. C++) is used.

5.3 A two-dimensional variance gamma model for option pricing

In this section we present a case study showing how one particular model, constructed using Lévy copulas, can be used to price multi-asset options.

The model We suppose that under the risk-neutral probability, the prices $\{S_t^1\}_{t\geq 0}$ and $\{S_t^2\}_{t\geq 0}$ of two risky assets satisfy

$$S_t^1 = e^{rt+X_t^1}, \qquad S_t^2 = e^{rt+X_t^2}, \tag{5.18}$$

where (X^1, X^2) is a Lévy process on $\mathbb{R}^d$ with characteristic triplet $(0, \nu, b)$ with respect to zero truncation function. X^1 and X^2 are supposed to be variance gamma processes, that is, the margins ν_1 and ν_2 of ν are of the form (1.19) with parameters $c^1, \lambda_+^1, \lambda_-^1$ and $c^2, \lambda_+^2, \lambda_-^2$. The Lévy copula F of ν is supposed to be of the form (5.4) with parameters θ and η. The no-arbitrage condition imposes that for $i = 1, 2$, $\lambda_+^i > 1$ and the drift coefficients satisfy

$$b_i = c_i \log\left(1 - \frac{1}{\lambda_+^i} + \frac{1}{\lambda_-^i} - \frac{1}{\lambda_+^i \lambda_-^i}\right).$$

The problem In the rest of this section, model (5.18) will be used to price two different kinds of multi-asset options: the option on weighted average, whose payoff at expiration date T is given by

$$H_T = \left(\sum_{i=1}^{2} w_i S_T^i - K\right)^+ \quad \text{with} \quad w_{1,2} \geq 0 \quad \text{and} \quad w_1 + w_2 = 1,$$

and the best-of or alternative option with payoff structure

$$H_T = \left(N \max\left(\frac{S_T^1}{S_0^1}, \frac{S_T^2}{S_0^2}\right) - K\right)^+$$

Option pricing by Monte Carlo Basket options, described above can be priced by Monte Carlo method using European options on individual stocks as control variates. Denote the discounted payoffs of European options by

$$V_T^i = e^{-rT}(S_T^i - K)^+ \quad \text{for } i = 1, 2.$$

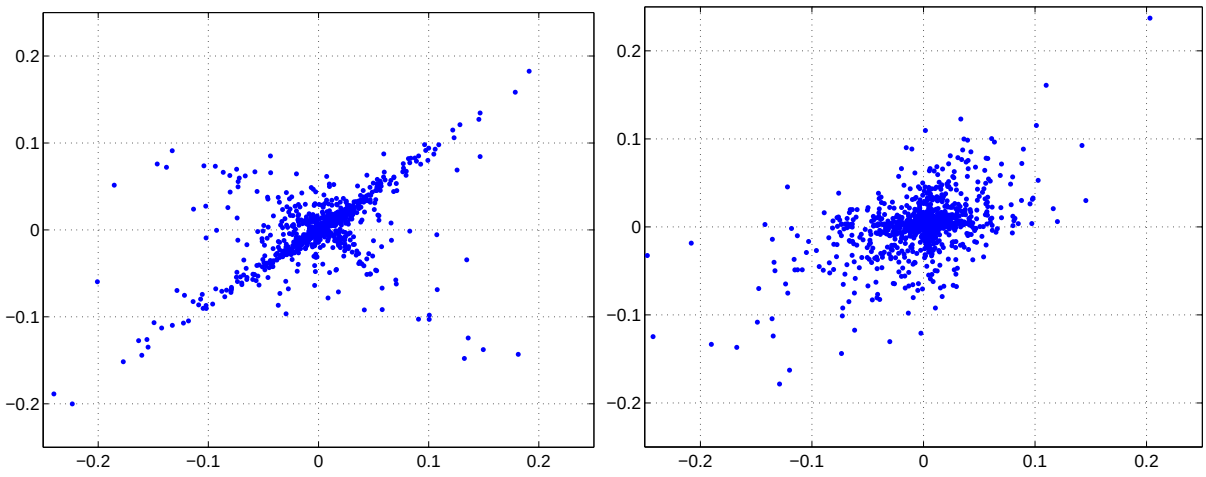

Figure 5.5: Scatter plots of returns in a 2-dimensional variance gamma model with correlation $\rho = 50\%$ and different tail dependence. Left: strong tail dependence ($\eta = 0.75$ and $\theta = 10$). Right: weak tail dependence ($\eta = 0.99$ and $\theta = 0.61$).

and the discounted payoff of the basket option by $V_T = e^{-rT} H_T$. Then the Monte Carlo estimate of basket option price is given by

$$\widehat{E[V_T]} = \bar{V}_T + a_1(E[V_T^1] - \bar{V}_T^1) + a_2(E[V_T^2] - \bar{V}_T^2),$$

where a bar over a random variable denotes the sample mean over N i.i.d. realizations of this variable, that is, $\bar{V}_T = \frac{1}{N}\sum_{i=1}^N V_T^{(i)}$, where $V_T^{(i)}$ are independent and have the same law as V_T. The coefficients a_1 and a_2 should be chosen in order to minimize the variance of $\widehat{E[V_T]}$. It is easy to see that this variance is minimal if $a = \Sigma a^0$, where $\Sigma_{ij} = \mathrm{Cov}(V_T^i, V_T^j)$ and $a_i^0 = \mathrm{Cov}(V_T, V_T^i)$. In practice these covariances are replaced by their in-sample estimates; this may introduce a bias into the estimator $\widehat{E[V_T]}$, but for sufficiently large samples this bias is small compared to the Monte Carlo error [45].

To illustrate the option pricing procedure, we fixed the following parameters of the marginal distributions of the two assets: $c^1 = c^2 = 25$, $\lambda_+^1 = 28.9$, $\lambda_-^1 = 21.45$, $\lambda_+^2 = 31.66$ and $\lambda_-^2 = 25.26$. In the parametrization (1.18) this corresponds to $\theta^1 = \theta^2 = -0.2$, $\kappa^1 = \kappa^2 = 0.04$, $\sigma^1 = 0.3$ and $\sigma^2 = 0.25$. To emphasize the importance of tail dependence for pricing multi-asset options, we used two sets of dependence parameters, which correspond both to a correlation of 50% (the correlation is computed numerically) but lead to returns with very different tail dependence structures:

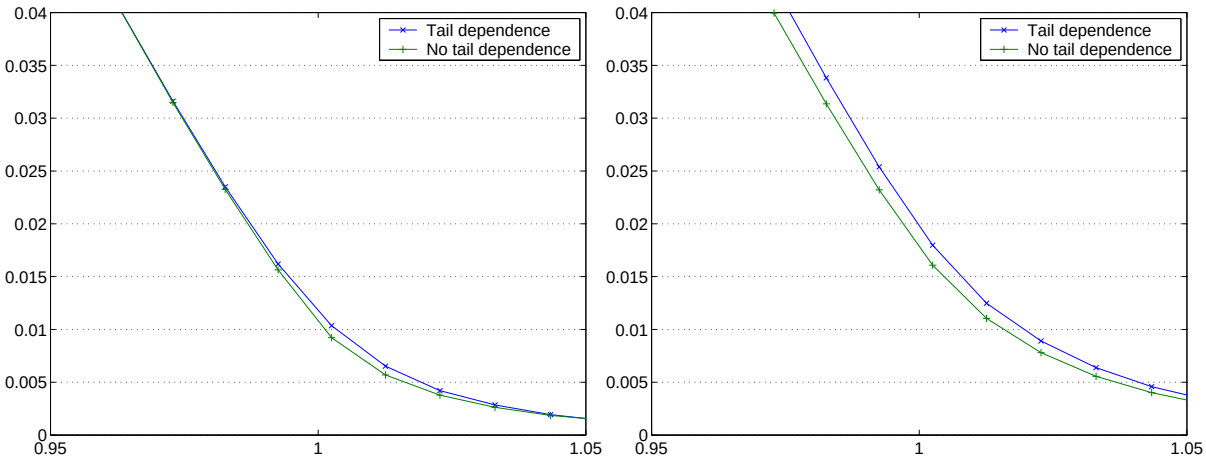

Figure 5.6: Prices of options on weighted average (left) and of best-of options (right) for two different dependence patterns.

Pattern 1 Strong tail dependence: $\theta = 10$ and $\eta = 0.75$. The scatter plot of returns is shown in Figure 5.5, left graph. Although the signs of returns may be different, the probability that the returns will be large in absolute value simultaneously in both components is very high.

Pattern 2 Weak tail dependence: $\theta = 0.61$ and $\eta = 0.99$. The scatter plot of returns in shown in Figure 5.5, right graph. With this dependence structure the returns typically have the same sign but their absolute values are not correlated.

In each of the two cases, a sample of 1000 realizations of the couple (X_T^1, X_T^2) with $T = 0.02$ (one-week options) was simulated using the procedure described in Example 5.3. The cutoff parameter τ (see Equation (5.14)) was taken equal to 1000, which lead to limiting the average number of jumps for each trajectory to about 40. For this value of τ, $U_i^{-1}(\tau)$ is of order of 10^{-19} for both assets. Since for the variance gamma model the convergence of U^{-1} to zero as $\tau \to \infty$ is exponential, the error resulting from the truncation of small jumps is of the same order, hence, negligible.

Figure 5.6 shows the prices of basket options, computed for different strikes with dependence patterns given above. The initial asset prices were $S_0^1 = S_0^2 = 1$, and the interest rate was taken to be $r = 0.03$. For the option on weighted average, the weights w_i were both equal to 0.5 and for the best-of option the coefficient was $N = 1$. The prices of European options, used for

variance reduction, were computed using the Fourier transform algorithm described in Chapter 1. The standard deviation of Monte Carlo estimates of option prices was below $2 \cdot 10^{-4}$ at the money in all cases.

The difference between option prices computed with and without tail dependece is clearly important for both types of options: as seen from Figure 5.6, neglecting tail dependence may easily lead to a 10% error on the option price at the money. On the other hand, this example shows that using Lévy copulas allows to take into account the tail dependence and discriminate between two situations that would be undistinguishable in a log-normal framework.

Conclusions and perspectives

In the first part of this thesis we have solved, using entropic regularization, the ill-posed problem of calibrating an exponential Lévy model to options data and proposed a stable numerical method for computing this solution. Applying our method to prices of index options allowed us to estimate the risk-neutral Lévy measures, implied by market prices. This object is the analog, for exponential Lévy models, of implied volatility, used in the Black-Scholes framework. Our empirical results allow to make a number of important conclusions. First, using an exponential Lévy model one can calibrate with high precision the prices of a set of options with common maturity. Moreover, high quality of calibration is achieved already by using finite-intensity Lévy processes. Therefore, from the point of view of option pricing the imperative for using more complex infinite-intensity models is not clear. The third conclusion is that even in the non-parametric setting it is impossible, using an exponential Lévy model, to calibrate accurately the prices of stock index options of several maturities at the same time: options of different maturities produce different implied Lévy measures. This confirms the observation already made by several authors [13, 68] that the framework of exponential Lévy models is not sufficiently flexible to reproduce the term structure of implied volatilities correctly.

In view of the above conclusions, we plan to continue the line of research initiated by this thesis, by extending its results to models of stock price behavior that do allow to describe the entire term structure of implied volatilities, e.g. models based on additive processes (processes with independent but not stationary increments) and hybrid models including both jumps and stochastic volatility. The second important direction of future research is to investigate the impact of our calibration methodology on the methods of hedging in presence of jumps in stock prices.

In the second part of this thesis we introduced the notion of Lévy copula, providing a

general framework in which the dependence structures of multidimensional Lévy processes can be described. Lévy copulas completely characterize the possible dependence patterns of Lévy processes in the sense that for every Lévy process, there exists a Lévy copula that describes its dependence structure and for every Lévy copula and every n one-dimensional Lévy processes, there exists an n-dimensional Lévy process with this Lévy copula and with margins given by these one-dimensional processes. Multidimensional Lévy process models for applications can thus be constructed by taking any n one-dimensional processes and a Lévy copula from a (possibly parametric) family. The simulation methods, developed in the last chapter of this thesis, allow to compute various quantities of interest in a Lévy copula model using the Monte Carlo method.

The scope of potential applications of Lévy copula models in finance and other domains is large. Financial applications include basket option pricing, and portfolio management. Lévy copula models are also useful in insurance and in risk management, to model the dependence between loss processes of different business lines, and more generally, in all multivariate problems where dependence between jumps needs to be taken into account.

From the point of view of applications, the next step is to develop the methods of estimating Lévy copula models from the data, using, for example, simulation-based techniques of statistical inference [51]. A more theoretical research direction that we currently pursue in collaboration with Jan Kallsen is to investigate the relation between the Lévy copula of a Lévy process and its probabilistic copula at a given time.

Bibliography

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, 2000.

[2] L. Andersen and J. Andreasen, *Jump-diffusion models: Volatility smile fitting and numerical methods for pricing*, Rev. Derivatives Research, 4 (2000), pp. 231–262.

[3] M. Avellaneda, R. Buff, C. Friedman, N. Grandchamp, L. Kruk, and J. Newman, *Weighted Monte Carlo: a new technique for calibrating asset-pricing models*, Int. J. Theor. Appl. Finance, 4 (2001), pp. 91–119.

[4] M. Avellaneda, C. Friedman, R. Holmes, and D. Samperi, *Calibrating volatility surfaces via relative entropy minimization*, Applied Mathematical Finance, 4 (1997), pp. 37–64.

[5] O. Barndorff-Nielsen, *Processes of normal inverse Gaussian type*, Finance Stoch., 2 (1998), pp. 41–68.

[6] O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds., *Lévy Processes — Theory and Applications*, Birkhäuser, Boston, 2001.

[7] O. E. Barndorff-Nielsen, *Normal inverse Gaussian distributions and stochastic volatility modelling*, Scand. J. Statist., 24 (1997), pp. 1–13.

[8] O. E. Barndorff-Nielsen and A. Lindner, *Some aspects of Lévy copulas*. Download from `www-m4.ma.tum.de/m4/pers/lindner`, 2004.

[9] O. E. Barndorff-Nielsen and N. Shephard, *Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial econometrics*, J. R. Statistic. Soc. B, 63 (2001), pp. 167–241.

[10] D. Bates, *Jumps and stochastic volatility: the exchange rate processes implicit in Deutschemark options*, Rev. Fin. Studies, 9 (1996), pp. 69–107.

[11] D. Bates, *Post-87 crash fears in the S&P 500 futures option market*, J. Econometrics, 94 (2000), pp. 181–238.

[12] D. Bates, *Maximum likelihood estimation of latent affine processes.* Working paper, available from `http://www.biz.uiowa.edu/faculty/dbates/`, 2004.

[13] D. S. Bates, *Testing option pricing models*, in Statistical Methods in Finance, vol. 14 of Handbook of Statistics, North-Holland, Amsterdam, 1996, pp. 567–611.

[14] N. Baud, A. Frachot, and T. Roncalli, *How to avoid over-estimating capital charge in operational risk.* Download from `gro.creditlyonnais.fr/content/fr/home_mc.htm`, 2002.

[15] S. Beckers, *A note on estimating parameters of a jump-diffusion process of stock returns*, J. Fin. and Quant. Anal., 16 (1981), pp. 127–140.

[16] F. Bellini and M. Frittelli, *On the existence of minimax martingale measures*, Math. Finance, 12 (2002), pp. 1–21.

[17] J. Bertoin, *Lévy Processes*, Cambridge University Press, Cambridge, 1996.

[18] J. Bouchaud and M. Potters, *Théorie des Risques Financiers*, Aléa, Saclay, 1997.

[19] S. Boyarchenko and S. Levendorskiĭ, *Non-Gaussian Merton-Black-Scholes Theory*, World Scientific, River Edge, NJ, 2002.

[20] R. Byrd, P. Lu, and J. Nocedal, *A limited memory algorithm for bound constrained optimization*, SIAM Journal on Scientific and Statistical Computing, 16 (1995), pp. 1190–1208.

[21] P. Carr, H. Geman, D. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Business, 75 (2002), pp. 305–332.

[22] P. Carr, H. Geman, D. Madan, and M. Yor, *Stochastic volatility for Lévy processes*, Math. Finance, 13 (2003), pp. 345–382.

[23] P. Carr and D. Madan, *Option valuation using the fast Fourier transform*, J. Comput. Finance, 2 (1998), pp. 61–73.

[24] T. Chan, *Pricing contingent claims on stocks driven by Lévy processes*, Ann. Appl. Probab., 9 (1999), pp. 504–528.

[25] A. S. Cherny and A. N. Shiryaev, *Change of time and measure for Lévy processes.* MaPhySto lecture notes series 2002-13, 2002.

[26] R. Cont, J.-P. Bouchaud, and M. Potters, *Scaling in financial data: Stable laws and beyond*, in Scale Invariance and Beyond, B. Dubrulle, F. Graner, and D. Sornette, eds., Springer, Berlin, 1997.

[27] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall / CRC Press, 2004.

[28] R. Cont and P. Tankov, *Non-parametric calibration of jump-diffusion option pricing models*, J. Comput. Finance, 7 (2004).

[29] J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp., 19 (1965), pp. 297–301.

[30] I. Csiszar, *I-divergence geometry of probability distributions and minimization problems*, Annals of Probability, 3 (1975), pp. 146–158.

[31] M. H. A. Davis, *Option pricing in incomplete markets*, in Mathematics of Derivative Securities, M. H. A. Dempster and S. R. Pliska, eds., 1997.

[32] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, 1975.

[33] F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker, *Exponential hedging and entropic penalties*, Math. Finance, 12 (2002), pp. 99–123.

[34] F. Delbaen and W. Schachermayer, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Math. Ann., 312 (1998), pp. 215–250.

[35] B. Dupire, *Pricing with a smile*, RISK, 7 (1994), pp. 18–20.

[36] E. Eberlein, *Applications of generalized hyperbolic Lévy motion to Finance*, in Lévy Processes — Theory and Applications, O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds., Birkhäuser, Boston, 2001, pp. 319–336.

[37] E. Eberlein, U. Keller, and K. Prause, *New insights into smile, mispricing and Value at Risk: The hyperbolic model*, Journal of Business, 71 (1998), pp. 371–405.

[38] E. Eberlein and S. Raible, *Some analytic facts on the generalized hyperbolic model*, in European Congress of Mathematics, Vol. II (Barcelona, 2000), vol. 202 of Progr. Math., Birkhäuser, Basel, 2001, pp. 367–378.

[39] N. El Karoui and R. Rouge, *Pricing via utility maximization and entropy*, Math. Finance, 10 (2000), pp. 259–276.

[40] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

[41] H. Föllmer and M. Schweizer, *Hedging of contingent claims under incomplete information*, in Applied Stochastic Analysis, M. Davis and R. Elliott, eds., vol. 5, Gordon and Breach, London, 1991, pp. 389–414.

[42] M. Frittelli, *The minimal entropy martingale measure and the valuation problem in incomplete markets*, Math. Finance, 10 (2000), pp. 39–52.

[43] H. Geman, *Pure jump Lévy processes for asset price modeling*, Journal of Banking and Finance, 26 (2002), pp. 1297–1316.

[44] H. Geman, D. Madan, and M. Yor, *Time changes for Lévy processes*, Math. Finance, 11 (2001), pp. 79–96.

[45] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer, New York, 2003.

[46] P. Glasserman and B. Yu, *Large sample properties of weighted Monte Carlo estimators.* Preprint, available from `www.gsb.columbia.edu/faculty/pglasserman/Other/`, 2002.

[47] T. Goll and J. Kallsen, *Optimal portfolios for logarithmic utility*, Stochastic Process. Appl., 89 (2000), pp. 31–48.

[48] T. Goll and L. Rüschendorf, *Minimal distance martingale measures and optimal portfolios consistent with observed market prices.* Preprint, available from `www.stochastik.uni-freiburg.de/~goll/start.html`, 2000.

[49] ——, *Minimax and minimal distance martingale measures and their relationship to portfolio optimization*, Finance Stoch., 5 (2001), pp. 557–581.

[50] A. V. Goncharsky, A. S. Leonov, and A. G. Yagola, *Applicability of the disparity principle in the case of non-linear incorrectly posed problems, and a new regularizing algorithm for solving them*, USSR Comp. Math. Math. Phys, 15 (1975), pp. 8–16.

[51] C. Gouriéroux and A. Monfort, *Simulation-Based Econometric Methods*, Oxford University Press, 1996.

[52] A. Hobson, *Concepts in Statistical Mechanics*, Gordon and Breach, Langhorne, Pennsylvania, 1971.

[53] N. Jackson, E. Süli, and S. Howison, *Computation of deterministic volatility surfaces*, Journal of Computational Finance, 2 (1999), pp. 5–32.

[54] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer, Berlin, 2nd ed., 2003.

[55] P. Jakubenas, *On option pricing in certain incomplete markets*, Proceedings of the Steklov Mathematical Institute, 237 (2002).

[56] H. Joe, *Multivariate Models and Dependence Concepts*, Chapman & Hall, London, 1997.

[57] P. Jorion, *On jump processes in the foreign exchange and stock markets*, Rev. Fin. Studies, 1 (1988), pp. 427–445.

[58] J. Kallsen, *Utility-based derivative pricing*, in Mathematical Finance — Bachelier Congress 2000, Springer, Berlin, 2001.

[59] J. Kallsen and P. Tankov, *Characterization of dependence of multidimensional Lévy processes using Lévy copulas.* Preprint, available from `www.cmap.polytechnique.fr/~tankov`, 2004.

[60] J. F. C. Kingman and S. J. Taylor, *Introduction to Measure and Probability*, Cambridge University Press, Cambridge, 1966.

[61] I. Koponen, *Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process.*, Physical Review E, 52 (1995), pp. 1197–1199.

[62] S. Kou, *A jump-diffusion model for option pricing*, Management Science, 48 (2002), pp. 1086–1101.

[63] R. W. Lee, *Option pricing by transform methods: extensions, unification and error control*, J. Comput. Finance, 7 (2004).

[64] A. Lindner and A. Szimayer, *A limit theorem for copulas.* Download from `www-m4.ma.tum.de/m4/pers/lindner`, 2003.

[65] F. Lindskog and A. McNeil, *Common Poisson shock models: Applications to insurance and credit risk modelling.* Available from `www.risklab.ch`, September 2001.

[66] D. Madan, *Financial modeling with discontinuous price processes*, in Lévy Processes — Theory and Applications, O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds., Birkhäuser, Boston, 2001.

[67] D. Madan, P. Carr, and E. Chang, *The variance gamma process and option pricing*, European Finance Review, 2 (1998), pp. 79–105.

[68] D. Madan and M. Konikov, *Option pricing using variance gamma Markov chains*, Rev. Derivatives Research, 5 (2002), pp. 81–115.

[69] C. Mancini, *Disentangling the jumps from the diffusion in a geometric jumping Brownian motion*, Giornale dell'Istituto Italiano degli Attuari, LXIV (2001), pp. 19–47.

[70] A. Medvedev and O. Scaillet, *A simple calibration procedure of stochastic volatility models with jumps by short term asymptotics.* Download from `www.hec.unige.ch/professeurs/SCAILLET_Olivier/pages_web/Home_Page.htm`, 2004.

[71] R. Merton, *Option pricing when underlying stock returns are discontinuous*, J. Financial Economics, 3 (1976), pp. 125–144.

[72] Y. Miyahara, *Canonical martingale measures in incomplete assets markets*, in Proceedings of the seventh Japan-Russian symposium (Probability theory and mathematical statistics, Tokyo 1995), World Scientific, 1996, pp. 343–352.

[73] Y. Miyahara and T. Fujiwara, *The minimal entropy martingale measures for geometric Lévy processes*, Finance Stoch., 7 (2003), pp. 509–531.

[74] V. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Doklady, 7 (1966), pp. 414–417.

[75] V. A. Morozov, *The error principle in the solution of operational equations by the regularization method*, USSR Comp. Math. Math. Phys., 8 (1968), pp. 63–87.

[76] R. Nelsen, *An Introduction to Copulas*, Springer, New York, 1999.

[77] L. Nguyen, *Calibration de Modèles Financiers par Minimisation d'Entropie Relative et Modèles aves Sauts*, PhD thesis, Ecole Nationale de Ponts et Chaussées, 2003.

[78] E. Nicolato, *Stochastic Volatility Models of Ornstein-Uhlenbeck Type*, PhD thesis, Aarhus University, Aarhus, 2000.

[79] K. Prause, *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, PhD thesis, Universität Freiburg i. Br., 1999.

[80] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: the Art of Scientific Computing*, Cambridge University Press, Cambridge, 1992.

[81] S. Raible, *Lévy Processes in Finance: Theory, Numerics and Empirical Facts*, PhD thesis, Freiburg University, 1998.

[82] S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, 1987.

[83] R. T. Rockafellar, *Integrals which are convex functionals*, Pacific journal of mathematics, 24 (1968), pp. 525–539.

[84] J. Rosiński, *Series representations of Lévy processes from the perspective of point processes*, in Lévy Processes — Theory and Applications, O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds., Birkhäuser, Boston, 2001.

[85] M. Rubinstein, *Implied binomial trees*, Journal of Finance, 49 (1994), pp. 771–819.

[86] T. H. Rydberg, *The normal inverse Gaussian Lévy process: simulation and approximation*, Comm. Statist. Stochastic Models, 13 (1997), pp. 887–910.

[87] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK, 1999.

[88] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, North Holland, 1983.

[89] M. Schweizer, *On the minimal martingale measure and the Föllmer-Schweizer decomposition*, Stochastic Anal. Appl., 13 (1995), pp. 573–599.

[90] A. Sklar, *Random variables, distribution functions, and copulas — a personal look backward and forward*, in Distributions with Fixed Marginals and Related Topics, L. Rüschendorf, B. Schweizer, and M. D. Taylor, eds., Institute of Mathematical Statistics, Hayward, CA, 1996.

[91] M. Stutzer, *A simple nonparametric approach to derivative security valuation*, Journal of Finance, 51 (1996), pp. 1633–1652.

[92] P. Tankov, *Dependence structure of spectrally positive multidimensional Lévy processes*. Download from `www.cmap.polytechnique.fr/~tankov`, 2003.

[93] M. Teboulle and I. Vajda, *Convergence of best ϕ-entropy estimates*, IEEE transactions on information theory, 39 (1993), pp. 297–301.

[94] A. N. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*, Wiley, New York, 1977.

[95] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer Academic Publishers, 1995.

[96] X. Zhang, *Analyse Numérique des Options Américaines dans un Modèle de Diffusion avec Sauts*, PhD thesis, Ecole Nationale des Ponts et Chaussées, 1994.

Index