ECOLE POLYTECHNIQUE

CENTRE DE MATHÉMATIQUES APPLIQUÉES
UMR CNRS 7641

91128 PALAISEAU CEDEX (FRANCE). Tél: 01 69 33 41 50. Fax: 01 69 33 30 11

http://www.cmap.polytechnique.fr/

Recovering exponential Lvy models from option prices: regularization of an ill-posed inverse problem

Rama CONT & Peter TANKOV

R.I. N^0 549 *Aout 2004/ August 2004*

RETRIEVING EXPONENTIAL LÉVY MODELS FROM OPTION PRICES: REGULARIZATION OF AN ILL-POSED INVERSE PROBLEM *

RAMA CONT† AND PETER TANKOV‡

Abstract. We propose a stable nonparametric method for constructing an option pricing model of exponential Lévy type, consistent with a given data set of option prices. After demonstrating the ill-posedness of the usual and least squares version of this inverse problem, we suggest to regularize the calibration problem by reformulating it as the problem of finding an exponential Lévy model that minimizes the sum of the pricing error and the relative entropy with respect to a prior exponential Lévy model. We prove the existence of solutions for the regularized problem and show that it yields solutions which are continuous with respect to the data, stable with respect to the choice of prior and converge to the minimum-entropy least square solution of the initial problem.

Key words. inverse problem, entropy, Lévy process, model calibration, option pricing, regularization.

AMS subject classifications. 49N45 60G51 60J75 91B70

1. Introduction. The specification of an arbitrage-free option pricing model on a time horizon T_∞ involves the choice of a *risk-neutral* measure [25]: a probability measure Q on the set Ω of possible trajectories $(S_t)_{t\in[0,T_\infty]}$ of the underlying asset such that the discounted asset price $e^{-rt}S_t$ is a martingale (where r is the discount rate). Such a probability measure Q then specifies a pricing rule which attributes to an option with terminal payoff H_T at T the value $C(H_T) = e^{-rT}E^Q[H_T]$. For example, the value under the pricing rule Q of a call option with strike K and maturity T is given by $e^{-rT}E^Q[(S_T-K)^+]$. Given that data sets of option prices have become increasingly available, a common approach for selecting the pricing model Q is to choose, given option prices $(C(H^i))_{i\in I}$ with maturities T_i payoffs H^i, a risk-neutral measure Q compatible with the observed market prices, i.e. such that $C(H^i) = e^{-rT_i}E^Q[H^i]$. This *inverse problem* of determining a pricing model Q verifying these constraints is known as the "model calibration" problem. The number of observed options can be large ($\simeq 100-200$ for index options) and the Black-Scholes model has to be replaced with models with richer structure such as nonlinear diffusion models [18] or models with jumps [13]. The inverse problem is ill-posed in these settings [14, 34] and various methods have been proposed for solving it in a stable manner, mostly in the framework of diffusion models [1, 4, 5, 6, 8, 15, 18, 26, 33, 34].

We study in this paper the calibration problem for the class of option pricing models with jumps –exponential Lévy models– where the risk-neutral dynamics of the logarithm of the stock price is given by a Lévy process. The problem is then to choose the Lévy process –described by its Lévy measure– in a way compatible with a set of observed option prices. Option prices being evaluated as expectations, this inverse problem can also be interpreted as a (generalized) moment problem for a Lévy process: given a finite number of option prices, it is typically an ill-posed problem. The relation between the option prices and the Lévy measure being nonlinear, we face a nonlinear, infinite dimensional inverse problem. After demonstrating the ill-posedness of the usual and least squares version of this inverse problem, we show that

*The second authors work was supported by an AMX fellowship from Ecole Polytechnique.
†CMAP- Ecole Polytechnique, F 91128 Palaiseau, France. Email: Rama.Cont@polytechnique.org
‡CMAP-Ecole Polytechnique, F 91128 Palaiseau, France. Peter.Tankov@polytechnique.org

it can be regularized by using as penalization term the relative entropy of with respect to a prior exponential Lévy model. We show that our approach yields solutions which are continuous with respect to the data, stable with respect to the choice of prior and converge to the minimum-entropy least square solution of the initial problem.

1.1. Relation to previous literature. Several authors [22, 19, 17, 30] have investigated the minimal entropy martingale measure (MEMM) –the pricing measure Q that minimizes the relative entropy with respect to a reference probability P– as a method for pricing in incomplete markets. The specific case of exponential Lévy models has been treated in [30] (see also [10]). However, option prices computed using the MEMM are in general not consistent with the market-quoted prices of traded European options and can lead to arbitrage opportunities with respect to market-traded options.

The notion of minimal entropy measure consistent with observed market prices was introduced in a static (one-period) framework by Avellaneda [4, 3]: given prices of call options $\{C_M(T_i, K_i)\}_{i\in I}$ and a prior model P, it is obtained by minimizing relative entropy over all probability measures $Q \sim P$ such that

$$C_M(T_i, K_i) = E^Q[e^{-rT_i}(S_{T_i} - K_i)^+] \quad \text{for } i \in I \tag{1.1}$$

This approach is based on relative entropy minimization under constraints [16] and yields a computable result. This approach was extended to the case of stochastic processes by the Weighted Monte Carlo method of Avellaneda et al [5], but the martingale property is lost since it would correspond to an infinite number of constraints [31]. As a result, derivative prices computed with the weighted Monte Carlo algorithm may contain arbitrage opportunities, especially when applied to forward start contracts.

Goll and Rüschendorf [24] consider the notion of consistent (or calibrated) minimal entropy martingale measure (CMEMM), defined as the solution of

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

where the minimum is taken over all martingale measures $Q \sim P$ verifying (1.1). While this notion seems to conciliate the advantages of the MEMM and Avellaneda's entropy minimization under constraints, no algorithm is proposed in [24] to compute the CMEMM. In fact, the notion of CMEMM does not preserve in general the structure of the prior –e.g. the Markov property–and it may be difficult to represent.[1] We also note that relative entropy is not a convenient notion when dealing with one dimensional diffusion models since as soon as the model has a diffusion coefficient different from the prior their measures become singuar and the relative entropy is infinite. Other methods have been used to solve the calibration problem for the class of diffusion models: a stochastic control method under constraints [6, 34], Tikhonov regularization [15, 26, 29], stochastic particle methods [8].

In this paper we show that the shortcomings of the above approaches can be overcome by enlarging the class of models to include processes with jumps and using relative entropy as a regularization criterion rather than a selection criterion. On one hand, introducing jumps in the prior model allows to obtain a large class of equivalent martingale measures which also have finite relative entropy with respect to the prior,

[1] In particular, if X is a Lévy process under the prior P, it will in general no longer be a Lévy process under a consistent minimal entropy martingale measure.

avoiding the singularity which arises in diffusion models. On the other hand, by restricting the class of pricing models to exponential Lévy models (see section 2), we are able to go beyong existence and uniqueness solutions and obtain a computable alternative to the CMEMM. Also, unlike the Weighted Monte Carlo approach, our approach yields as solution a continuous-time price process whose discounted value is a martingale. Finally, the use of regularization yields a stable solution to the inverse problem for which a computational approach is possible [14].

Unlike linear inverse problems for which general results on regularization methods and their convergence properties are available [20], nonlinear inverse problems have been explored less systematically. Our study is an example of rigorous analysis of entropy-based regularization for a nonlinear, infinite-dimensional inverse problem. Previous results on regularization using entropy have been obtained in finite dimensional setting [21] by mapping the problem to a Tikhonov regularization problem. Using probabilistic methods, we are able to use a direct approach and extend these result to the infinite dimensional setting considered here.

1.2. Outline. The paper is structured as follows. Section 2 recalls basic facts about Lévy processes and exponential Lévy models. In Section 3 we formulate the calibration problem as that of finding a martingale measure Q, consistent with market-quoted prices of traded options, under which the logarithm of the stock price process remains a Lévy process. We show that both this problem and its least squares version are ill-posed: a solution may not exists and when it exists, may not be stable with respect to perturbations in the data. Section 4 discusses relative entropy in the case of Lévy processes, its use as a criterion for selecting solutions and introduces the notion of minimum-entropy least squares solution. Although this notion of solution may still lack uniqueness and stability, we show in Section 5 that it can be approximated in a stable manner using the method of regularization. We formulate the regularized version of the calibration problem and show that it always admits a solution which is stable with respect to market data. Moreover, we formulate conditions under which the solutions of the regularized problem converge to the minimum-entropy least squares solution.

In Section 6 we show that the solutions of the regularized calibration problem are stable with respect to small perturbations of the prior measure. This also implies that the solutions of the regularized calibration problem with any prior measure can be approximated (in the weak sense) by the solutions of regularized problems with discretized priors, which has implications for the discretization and the numerical solution of the regularized calibration problem: these issues are further discussed in the companion paper [14].

2. Definitions and notations. Consider a time horizon $T_\infty < \infty$ and denote by Ω the space of $\mathbb{R}^d$-valued cadlag functions on $[0, T_\infty]$, equipped with the Skorokhod topology [27]. 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 $s \in [0, T_\infty]$ and 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 $s \in [0, t]$. Weak convergence of measures will be denoted by $\Rightarrow$.

Lévy processes. A Lévy process $\{X_t\}_{t \geq 0}$ on $(\Omega, \mathcal{F}, P)$ is a cadlag stochastic process with stationary independent increments, satisfying $X_0 = 0$. The characteristic function of X_t has the following form, called the Lévy-Khinchin representation [35]:

$$E[e^{izX_t}] = e^{t\psi(z)}, \quad \text{with}$$
$$\psi(z) = -\frac{1}{2}Az^2 + i\gamma z + \int_{-\infty}^{\infty}(e^{izx} - 1 - izh(x))\nu(dx) \tag{2.1}$$

where $A \geq 0$ is the unit variance of the Brownian motion part of the Lévy process, $\gamma \in \mathbb{R}$, ν is a positive measure on $\mathbb{R}$ verifying $\nu(\{0\}) = 0$ and

$$\int_{-\infty}^{\infty}(x^2 \wedge 1)\nu(dx) < \infty,$$

and h is the truncation function: any bounded measurable function $\mathbb{R} \to \mathbb{R}$ such that $h(x) \equiv x$ on a neighborhood of zero. The most common choice of truncation function is $h(x) = x1_{|x|\leq 1}$ but sometimes in this paper we will need h to be continuous. The triplet (A, ν, γ) is called the characteristic triplet of X with respect to the truncation function h. Sometimes in financial literature, $\sigma := \sqrt{A}$ is called the volatility of X.

Model setup. In this paper we treat exponential Lévy models, where the stock price process S_t is modelled as the exponential of a Lévy process:

$$S_t = S_0 e^{rt+X_t} \tag{2.2}$$

where r is the interest rate. Under a risk-neutral probability Q, e^{X_t} must be a martingale. It follows from (2.1) that this is the case if and only if

$$\frac{A}{2} + \gamma + \int_{-\infty}^{\infty}(e^x - 1 - h(x))\nu(dx) = 0. \tag{2.3}$$

Under Q call option prices can be evaluated as discounted expectations of terminal payoffs:

$$C^Q(T,K) = e^{-rT}E^Q[(S_T - K)^+] = e^{-rT}E^Q[(S_0e^{rT+X_T} - K)^+]. \tag{2.4}$$

If X is a Lévy process under P then, unless X is almost surely increasing or almost surely decreasing under P, the exponential Lévy model corresponding to P is arbitrage-free: there exists a risk-neutral probability Q equivalent to P [28, 11].

Notation. In the sequel $\mathcal{P}(\Omega)$ denotes the set of probability measures (stochastic processes) on $(\Omega, \mathcal{F})$, $\mathcal{L}$ denotes the set of all probability measures $P \in \mathcal{P}(\Omega)$ under which the coordinate process X is a Lévy process and $\mathcal{M}$ stands for the set of all probability measures $P \in \mathcal{P}(\Omega)$, under which $\exp(X_t)$ is a martingale. $\mathcal{L}_{NA}$ is the set of all probability measures $P \in \mathcal{L}$ corresponding to Lévy processes describing markets with no arbitrage opportunity, that is, to Lévy processes that are not almost surely increasing nor almost surely decreasing. Furthermore for a constant $B > 0$ we define

$$\mathcal{L}_B^+ = \{P \in \mathcal{L},\ \forall t \in [0, T_\infty],\ P[\Delta X_t \leq B\ ] = 1\ \},$$

the set of Lévy processes with jumps bounded from above by B.

The following lemma shows the usefulness of the above definitions.

LEMMA 2.1. *The set $\mathcal{M} \cap \mathcal{L}_B^+$ is weakly closed for every $B > 0$.*

Proof. Let $\{Q_n\}_{n=1}^\infty \subset \mathcal{M} \cap \mathcal{L}_B^+$ with characteristic triplets (A_n, ν_n, γ_n) with respect to a continuous truncation function h and let Q be a Lévy process with characteristic triplet (A, ν, γ) 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 Corollary VII.3.6 in [27], $\int_{-\infty}^{\infty} f(x)\nu(dx) = \lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)\nu_n(dx) = 0$, which implies that the jumps of Q are bounded by B.

Define a function g by

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

Then, by Corollary VII.3.6 in [27] and because Q_n satisfies the martingale condition (2.3) for every n,

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

which shows that Q also satisfies the condition (2.3). □

3. The calibration problem and its least squares formulation. 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 market pricing rule is an exponential Lévy model but can hold more generally since many models may give the same prices for a given set of European options. For instance one can construct, using Dupire's formula [18], a diffusion model that gives the same prices, for a set of European options, as a given exp-Lévy model [12]. Using the notation, defined in the preceding section, the calibration problem assumes the following form:

PROBLEM 1 (Calibration problem with equality constraints). *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*

$$\forall i \in I, C^{Q^*}(T_i, K_i) = C_M(T_i, K_i). \tag{3.1}$$

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 data points). 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

$$\|C\|_w^2 := \int_{[0,T_\infty]\times[0,\infty)} C(T, K)^2 w(dT \times dK). \tag{3.2}$$

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 $w = \sum_{i=1}^N w_i \delta_{(T_i,K_i)}(dT \times dK)$ (with e.g. N constraints), where $\{w_i\}_{1\le i\le N}$ are positive weights that sum up to one. Therefore, in this case

$$\|C_M - C^Q\|_w^2 = \sum_{i=1}^N w_i (C_M(T_i,K_i) - C^Q(T_i,K_i))^2. \tag{3.3}$$

The following lemma establishes some useful properties of the pricing error functional.

Lemma 3.1. *The pricing error functional $Q \mapsto \|C_M - C^Q\|_w^2$ is uniformly bounded and weakly continuous on $\mathcal{M} \cap \mathcal{L}$.*

Proof. From Equation (2.4), $C^Q(T,K) \le S_0$. Absence of arbitrage in the market implies that the market option prices satisfy the same condition. Therefore, $(C_M(T,K) - C^Q(T,K))^2 \le S_0^2$ and since w is a probability measure, $\|C_M - C^Q\|_w^2 \le S_0^2$.

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

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

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

The calibration problem now takes the following form:

Problem 2 (Least squares calibration problem). *Given prices C_M of call options, find $Q^* \in \mathcal{M} \cap \mathcal{L}$, such that*

$$\|C_M - C^{Q^*}\|_w^2 = \inf_{Q \in \mathcal{M}\cap\mathcal{L}} \|C_M - C^Q\|_w^2. \tag{3.4}$$

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 $\mathcal{Q}^{LS}(C_M)$.

Several authors [2, 7] have used least squares formulations similar to (3.4) for calibrating parametric models without taking into account that the least squares calibration problem is ill-posed in several ways. The principal difficulties of theoretical nature are the following:

Lack of identification. Although knowing option prices for one maturity and all strikes allows to determine the characteristic triplet of the underlying Lévy process completely, in real data sets, prices are only available for a finite number of strikes (typically between 10 and 100) and knowing the prices of a finite number of options is not sufficient to reconstruct the Lévy process. This problem is discussed in detail in [14, 36].

Absence of solution. In some cases even the least squares problem may not admit a solution, as shown by the following (artificial) example.

Example 3.1. Suppose that $S_0 = 1$, there are no interest rates or dividends and the (equally weighted) market data consist of the following two observations:

$$C_M(T=1, K=1) = 1 - e^{-\lambda} \quad \text{and} \quad C_M(T=1, K=e^{\lambda}) = 0, \tag{3.5}$$

with some $\lambda > 0$. It is easy to see that these prices are, for example, compatible with the (martingale) asset price process $S_t = e^{\lambda t} 1_{t \le \tau_1}$, 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 (3.5), the calibration problem (3.4) 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 (3.5) 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 (3.5), 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, $\inf_{Q\in\mathcal{M}\cap\mathcal{L}}\|C_M-C^Q\|^2_w=0$. Indeed, let $\{N_t\}_{t\geq 0}$ be a Poisson process with intensity λ. Then for every n, the process

$$X_t^n := -nN_t+\lambda t(1-e^{-n}) \tag{3.6}$$

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 (3.4) does not admit a solution.

Lack of continuity of solutions with respect to market data. Market option prices are typically defined up to a bid-ask spread and the prices used for calibration may therefore be subject to perturbations of this order. 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. Even if we ignore errors in market data, 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, such as 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 [20] that applies to all calibration problems discussed in this paper.

Definition 3.2 (Continuity with respect to data). *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.*

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 $Q_n\Rightarrow Q$.

Remark 3.1. Note that the above definition can accommodate the presence of random errors ("noise") in the data. In this case the observational error can be described by a separate probability space $(\Omega_0,\mathcal{E},p_0)$. The continuity property must then be interpreted as almost-sure continuity with respect to the law p_0 of the observational errors: for every (random) sequence $\{C_M^n\}_{n\geq 0}$ such that $\|C_M^n -$

$C_M\|_w \xrightarrow[n\to\infty]{} 0$ almost surely, then any sequence of solution with data $\{C_M^n\}_{n\geq 0}$ must verify the properties of Definition 3.2 p_0-almost surely.

It is easy to construct an example of market data leading to a least squares calibration problem (3.4) that does not satisfy the above definition.

EXAMPLE 3.2. Assume $S_0 = 1$, 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 (3.6) 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. Corollary VII.3.6 in [27]).

In addition to these theoretical obstacles, even if a solution exists, it may be difficult to compute numerically since, as shown in [14, 36], the pricing error $\|C_M - C^Q\|^2$ is typically non-convex and can have many local minima, preventing a gradient-based minimization algorithm from finding the solution.

4. Relative entropy as a selection criterion. When constraints given by option prices 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 start from a reference Lévy process P and look for the solution of the problem (3.4) 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

$$I(Q|P) = \begin{cases} E^P\left[\frac{dQ}{dP}\log\frac{dQ}{dP}\right] & \text{if } Q \ll P \text{ and } \quad E^P[|\frac{dQ}{dP}\log\frac{dQ}{dP}|] < \infty \\ \infty & \text{otherwise,} \end{cases} \tag{4.1}$$

where by convention $x\log x = 0$ when $x = 0$.

PROBLEM 3 (Minimum entropy least squares calibration problem). *Given prices C_M of call options and a prior Lévy process P, find a least squares solution $Q^* \in \mathcal{Q}^{LS}(C_M)$, such that*

$$I(Q^*|P) = \inf_{Q\in\mathcal{Q}^{LS}(C_M)} I(Q|P). \tag{4.2}$$

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 MELSS(C_M).

The prior probability P reflects our a priori knowledge about the nature of possible trajectories of the underlying asset and their probabilities of occurrence. 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. Whether this choice is adopted or not does not affect our discussion below. Other possible ways to choose the prior model in practice are discussed in [14], which also gives an empirical analysis of the effect of the choice of prior on the solution of the calibration problem.

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

- Relative entropy can be interpreted as a (pseudo-)distance to the prior P: 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 $z = 1$.

- Relative entropy for Lévy processes is easily expressed in terms of their characteristic triplets (see Theorem A.1).
- Relative entropy has an information-theoretic interpretation and has been repeatedly used for model selection in finance (see Section 1).

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 [12]. However, the calibration problem (4.2) remains ill-posed: since the minimization of entropy is done over the results of least squares calibration, problem (4.2) may only admit a solution if problem (3.4) does. Also, $\mathcal{Q}^{LS}(C_M)$ is not necessarily a compact set, so even if it is nonempty, (4.2) may not have a solution. Other undesirable properties such as absence of continuity and numerical instability are also inherited from the least squares approach. In Section 5 we will propose a regularized version of problem (4.2) that does not suffer from these difficulties.

The minimum entropy least squares solution does not always exist, but if the prior is chosen correctly, that is, if there exists at least one solution of problem (3.4) with finite relative entropy with respect to the prior, then minimum entropy least-squares solutions will also exist, as shown by the following lemma.

Lemma 4.1. *Let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ for some $B > 0$ and assume the problem* (3.4) *admits a solution Q^+ with $I(Q^+|P) = C < \infty$. Then the problem* (4.2) *admits a solution.*

Proof. Under the condition of the lemma, it is clear that the solution Q^* of problem (4.2), if it exists, satisfies $I(Q^*|P) \leq C$. This entails that $Q^* \ll P$, which means by Theorem IV.4.39 in [27] that $Q^* \in \mathcal{L}_B^+$. Therefore, Q^* belongs to the set

$$\mathcal{L}_B^+ \cap \{Q \in \mathcal{M} \cap \mathcal{L} : \|C^Q - C_M\| = \|C^{Q^+} - C_M\|\} \cap \{Q \in \mathcal{L} : I(Q|P) \leq C\}. \quad (4.3)$$

Lemma A.2 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 A.4, $I(Q|P)$ 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 2.1, $\mathcal{M} \cap \mathcal{L}_B^+$ 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 3.1 then implies that the set (4.3) is also weakly compact. Since $I(Q|P)$ is weakly lower semicontinuous, it reaches its minimum on this set. ☐

Remark 4.1. Notice that it is essential for our analysis that the model has discontinuous trajectories, i.e. the prior P corresponds to a Lévy process with jumps, not a diffusion process (which is, in this case, the Black Scholes model). More generally if P corresponds to the law of a diffusion model then the set of processes which have both the martingale property and finite entropy with respect to P is reduced to a single element and the solution to 4.2 is trivial.

5. Regularization using relative entropy. As observed in [14] and in Section 4, problem (4.2) 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 dispose of a noisy verion C_M^δ with $||C_M^\delta - C_M||_w \leq \delta$ and want to construct an approximation to MELSS(C_M), it is not a good idea to solve problem (4.2) with the noisy data C_M^δ because MELSS(C_M^δ) may be very different from MELSS(C_M). We therefore need to regularize the problem (4.2), 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 a minimum entropy least-squares solution as the noise level δ tends to zero if an appropriate parameter choice rule $\delta \mapsto \alpha(\delta)$ is used. 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 a classical approach to regularization of ill-posed problems [20], we suggest to construct a regularized version of (4.2) by using the relative entropy for penalization rather than for selection:

$$J_\alpha(Q) = \|C_M^\delta - C^Q\|_w^2 + \alpha I(Q|P), \tag{5.1}$$

where α is the regularization parameter and solve the following optimization problem:

Problem 4 (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*

$$J_\alpha(Q^*) = \inf_{Q \in \mathcal{M} \cap \mathcal{L}} J_\alpha(Q). \tag{5.2}$$

Problem (5.2) can be thought of in two ways:

- If the minimum entropy least squares solution with the true data C_M exists, (5.2) allows to construct a stable approximation of this solution using the noisy data.
- If the MELSS(C_M) $= \emptyset$, either because the set of least squares solutions is empty or because the least squares solutions are incompatible with the prior, the regularized problem (5.2) allows to achieve, in a stable manner, a trade-off between matching the 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 a sufficient condition on the prior P 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 (4.2).

5.1. Existence of solutions. The following result shows that, unlike the initial or the least squares formulation of the inverse problem, the regularized version always admits a solution:

Theorem 5.1. *Let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ for some $B > 0$. Then the calibration problem (5.2) has a solution $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$.*

Proof. By Lemma A.5, 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 A.2, $L_{J_\alpha(Q^0)}$ is tight and, by Prohorov's theorem, weakly relatively compact. Corollary A.4 entails that $I(Q|P)$ is weakly lower semicontinuous with respect to Q. Therefore $\{Q \in \mathcal{P}(\Omega) :$

$I(Q|P) \le J_\alpha(Q^0)\}$ is weakly closed and since by Lemma 2.1, $\mathcal{M} \cap \mathcal{L}_B^+$ is weakly closed, $\mathcal{M} \cap \mathcal{L}_B^+ \cap L_{J_\alpha(Q^0)}$ is weakly compact. Moreover, by Lemma 3.1, 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 (5.2) 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 "objective" probability measure, absence of arbitrage is guaranteed if options are priced using an *equivalent* martingale measure [25]. The following theorem gives a sufficient condition for this equivalence.

THEOREM 5.2. *Let $P \in \mathcal{L}_{NA} \cap \mathcal{L}_B^+$ and assume the characteristic function Φ_T^P of P satisfies*

$$\int_{-\infty}^{\infty} |\Phi_T^P(u)| du < \infty \tag{5.3}$$

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 (5.2) *satisfies $Q^* \sim P$.*

REMARK 5.1. Condition (5.3) 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 (5.3) for all T are

- Processes with non-zero Gaussian component ($A > 0$).
- Processes with stable-like behavior of small jumps whose Lévy measure satisfies

$$\exists \beta \in (0,2), \quad \liminf_{\varepsilon \downarrow 0} \varepsilon^{-\beta} \int_{-\varepsilon}^{\varepsilon} |x|^2 \nu(dx) > 0. \tag{5.4}$$

For a proof, see [35, Proposition 28.3]. This class includes tempered stable processes [13] with $\alpha_+ > 0$ and/or $\alpha_- > 0$.

To prove Theorem 5.2 we will use the following lemma:

LEMMA 5.3. *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*

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

Proof. From the Lévy-Khinchin formula and (2.3),

$$\psi(v-i) = -\frac{1}{2} A v(v-i) + \int_{-\infty}^{\infty} (e^{i(v-i)x} + iv - e^x - ive^x)\nu(dx). \tag{5.5}$$

Observe first that

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

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

$$\left| \frac{e^{i(v-i)x} + iv - e^x - ive^x}{(v-i)v} \right| \le xe^x + 1 - e^x + x^2 e^x. \tag{5.6}$$

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 $|v|\leq 2$,

$$\left|\frac{e^{i(v-i)x}+iv-e^x-ive^x}{(v-i)v}\right|\leq x(1-e^x)+\frac{x^2}{2}(v+\sqrt{1+v^2}e^x)\leq x(1-e^x)+x^2(1+2e^x). \tag{5.7}$$

Since the support of ν is bounded from above, the right-hand sides of (5.6) and (5.7) are ν-integrable and the proof of the lemma is completed. □

Proof. [Proof of Theorem 5.2] Let Q^* be a solution of (5.2) with prior P. Since $P\in\mathcal{L}_{NA}$, there exists $Q^0\in\mathcal{M}\cap\mathcal{L}$ such that $Q^0\sim P$ [11]. 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 (2.3), 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 $J_\alpha(Q_x)-J_\alpha(Q^*)\geq 0$ 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 (5.3) 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 IV.4.39 in [27], $\int_{-\infty}^{\infty}(\sqrt{\phi^*(x)}-1)^2\nu^P(dx)\leq K<\infty$ 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 $r>0$,

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

This implies that

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

for all $r>0$. Therefore, if the characteristic function of P satisfies the condition (5.3) 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 (11.15) in [13] is satisfied and option prices can be computed using Equation (11.20) of the above reference [2]:

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

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 $\Re\psi^0(v-i) \leq 0$ and $\Re\psi^*(v-i) \leq 0$ for all $v \in \mathbb{R}$ together with Lemma 5.3 implies that

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

for some constant C'. From the dominated convergence theorem and since Q^* satisfies (5.3), $\frac{\partial C^{Q_x}(T,K)}{\partial x}|_{x=0}$ 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

[2] This method for computing option prices by Fourier transform is originally due to Carr and Madan [9].

γ^Q) of the two terms in the expression (A.1) for relative entropy, we have:

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

Since $J_\alpha(Q_x) - J_\alpha(Q^*) \geq 0$, this expression must be bounded from below. Therefore, $\nu^P(\{\phi^* = 0\}) = 0$ and Theorem IV.4.39 in [27] entails that $P \ll Q^*$. □

5.2. Continuity of solutions with respect to data.

THEOREM 5.4 (Continuity of solutions with respect to data). *Let $\{C_M^n\}_{n\geq 1}$ and C_M be data sets of option prices such that*

$$\|C_M^n - C_M\|_w \underset{n\to\infty}{\to} 0.$$

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 (5.2) *with data C_M^n, prior Lévy process P and regularization parameter α. Then $\{Q_n\}_{n\geq 1}$ has a subsequence, converging weakly to $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$ and the limit of every converging subsequence of $\{Q_n\}_{n\geq 1}$ is a solution of calibration problem* (5.2) *with data C_M, prior P and regularization parameter α.*

Proof. By Lemma A.5, there exists $Q^0 \in \mathcal{M} \cap \mathcal{L}$ with $I(Q^0|P) < \infty$. Since, by Lemma 3.1, $\|C^{Q^0} - C_M^n\|^2 \leq S_0^2$ for all n, $\alpha I(Q_n|P) \leq S_0^2 + \alpha I(Q^0|P)$ for all n. Therefore, by Lemmas 2.1 and A.2 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 $Q^* \in \mathcal{M} \cap \mathcal{L}_B^+$. To simplify notation, this subsequence is denoted again by $\{Q_n\}_{n\geq 1}$. The triangle inequality and Lemma 3.1 imply that

$$\|C^{Q_n} - C_M^n\|^2 \xrightarrow[n\to\infty]{} \|C^{Q^*} - C_M\|^2 \tag{5.8}$$

Since, by Lemma A.3, the relative entropy functional is weakly lower semi-continuous 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. □

5.3. Convergence analysis. The convergence analysis of regularization methods for inverse problems usually involves the study of the solution of the regularized problem (5.2) as the noise level δ vanishes, the regularization parameter being chosen as a function $\alpha(\delta)$ of the noise level using some parameter choice rule. The following result gives conditions on the parameter choice rule $\delta \mapsto \alpha(\delta)$ under which the solutions of the regularized problem (5.2) converge to minimum entropy least squares solutions defined by (4.2):

THEOREM 5.5. *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* (3.4) *with data C_M (a least squares solution) such that $I(Q|P) < \infty$.*

In the case where the constraints are attainable i.e. $\|C^Q - C_M\| = 0$ 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 (5.2) *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* (4.2) *with data C_M and prior P. If the mimimum entropy least squares solution is unique MELSS(C_M) $= \{Q^+\}$ then*

$$Q^\delta \underset{\delta \to 0}{\Rightarrow} Q^+$$

Proof. By Lemma 4.1, there exists at least one MELSS with data C_M and prior P, with finite relative entropy with respect to the prior. Let $Q^+ \in$ MELSS(C_M). 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

$$\begin{aligned}&(1-r)\|C^{Q^{\delta_k}} - C_M\|^2 + \alpha(\delta_k) I(Q^{\delta_k}|P) \\ &\qquad \leq (1+r)\|C^{Q^+} - C_M\|^2 + \frac{2\delta_k^2}{r} + \alpha(\delta_k) I(Q^+|P), \quad (5.9)\end{aligned}$$

and since $Q^+ \in \mathcal{Q}^{LS}(C_M)$, this implies for all $r \in (0,1)$ that

$$\alpha(\delta_k)I(Q^{\delta_k}|P) \leq 2r\|C^{Q^+} - C_M\|^2 + \frac{2\delta_k^2}{r} + \alpha(\delta_k)I(Q^+|P). \tag{5.10}$$

If the constraints are met exactly $\|C^{Q^+} - C_M\| = 0$ 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

$$\limsup_k \{I(Q^{\delta_k}|P)\} \leq I(Q^+|P). \tag{5.11}$$

If $\|C^{Q^+} - C_M\| > 0$ (misspecified model) then the right-hand side of (5.10) 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 (5.11) once again.

Inequality (5.11) implies in particular that $I(Q^{\delta_k}|P)$ is uniformly bounded, which proves, by Lemmas A.2 and 2.1, 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 (5.9) 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 3.1 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 A.3) and using (5.11),

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

which means that $Q^* \in \mathrm{MELSS}(C_M)$. 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^+. □

Remark 5.2 (Random errors). In line with Remark 3.1, it is irrelevant whether the noise in the data is "deterministic" or "random", as long the error level δ is interpreted as a worst-case error level i.e. an almost sure bound on the error:

$$p_0(||C_M^\delta - C_M|| \leq \delta) = 1. \tag{5.12}$$

In this case, Theorem 5.5 holds for random errors, all convergences being interpreted as almost-sure convergence with respect to the law p_0 of the errors.

6. Stability with respect to the prior. A convenient way to discretize the calibration problem (5.2) is to take a prior Lévy process P with a finite number of jump sizes:

$$\nu^P = \sum_{k=0}^{M-1} p_k \delta_{\{x_k\}}(dx). \tag{6.1}$$

In this case, since the solution Q satisfies $Q \ll P$, by Theorem IV.4.39 in [27], the Lévy measure of the solution necessarily satisfies $\nu^Q \ll \nu^P$, therefore

$$\nu^Q = \sum_{k=0}^{M-1} q_k \delta_{\{x_k\}}(dx), \tag{6.2}$$

Thus, the calibration problem (5.2) becomes a finite-dimensional optimization problem and can be solved using a numerical optimization algorithm [14]. The advantage of this discretization approach is that we are solving the same problem (5.2), only with a different prior measure, so all results of Section 5 (existence of solution, continuity etc.) hold in the finite-dimensional case.

The discretized calibration problem and the numerical methods to solve it have been discussed in detail in our previous paper [14]. Here we will complement these results by a theorem showing that 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 discretized priors. We start by showing that every Lévy process can be approximated by Lévy processes with Lévy measures of the form (6.1):

LEMMA 6.1. *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

$$\lim_n \int (1 \wedge x^2) f(x)\nu_n(dx) = \lim_n \int f_n(x)\mu(dx) = \int f(x)\mu(dx) = \int (1 \wedge x^2) f(x)\nu(dx). \tag{6.3}$$

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 (6.3), $\lim_n \int g(x)\nu_n(dx) = \int g(x)\nu(dx)$ and by Corollary VII.3.6 in [27], $P_n \Rightarrow P$. □

Theorem 6.2. *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* (5.2) *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* (5.2) *with prior P.*

Remark 6.1. To approximate numerically the solution of the calibration problem (5.2) with a given prior P, we need to construct, using Lemma 6.1, 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 the above theorem) to a solution of the calibration problem with prior P.

The second implication of the above theorem is that small changes in the prior Lévy process lead to small changes in the solution: this means that the solution is not very sensitive to minor errors in the determination of the prior measure. This result confirms the empirical observations made in [14].

Proof. By Lemma A.5, 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 3.1, $\|C^{\tilde{Q}_n} - C_M\|_w^2 \leq S_0^2$ 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 A.2, $\{Q_n\}$ is tight and, by Prohorov's theorem and Lemma 2.1, 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 A.3 entails that

$$I(Q, P) \leq \liminf_n I(Q_n, P_n), \tag{6.4}$$

and since, by Lemma 3.1, the pricing error is weakly continuous, we also have

$$\|C^Q - C_M\|_w^2 + \alpha I(Q, P) \leq \liminf_n \{\|C^{Q_n} - C_M\|_w^2 + \alpha I(Q_n, P_n)\}. \tag{6.5}$$

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 3.1,

$$\lim_n \|C^{Q'_n} - C_M\|_w^2 = \|C^{Q'} - C_M\|_w^2. \tag{6.6}$$

Moreover,

$$\begin{aligned}\lim_n I(Q'_n|P_n) &= \lim_n \int_\Omega \frac{\phi}{E^{P_n}[\phi]} \log \frac{\phi}{E^{P_n}[\phi]} dP_n \\ &= \lim_n \frac{1}{E^{P_n}[\phi]} \int_\Omega \phi \log \phi dP_n - \lim_n \log \int_\Omega \phi dP_n = \int_\Omega \phi \log \phi dP. \end{aligned} \tag{6.7}$$

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 (6.5–6.7) and the optimality of Q_n, we obtain that for every $\phi \in C_b(\Omega)$ with $\phi \geq 0$ and $E^P[\phi] = 1$,

$$\|C^Q - C_M\|_w^2 + I(Q, P) \leq \|C^{Q_\phi} - C_M\|_w^2 + I(Q_\phi|P) \tag{6.8}$$

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 (6.8), 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 (6.8) 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, (6.8) holds for all $\phi \in L^1(P)$ with $\phi \geq 0$ and $E^P[\phi] = 1$, which completes the proof of the theorem. □

7. Conclusion. We have proposed here a stable method for constructing an option pricing model of exponential Lévy type, consistent with a given data set of option prices. Our approach is based on the regularization of the calibration problem using the relative entropy with respect to a prior exp-Lévy model as penalization term. The regularization restores existence and stability of solutions; the use of relative entropy links our approach to previous work using relative entropy as a criterion for selection of pricing rules. This technique is readily amenable to numerical implementation, as shown in [14], where empirical applications to financial data are also discussed.

The problem studied here is an example of regularization of a nonlinear, infinite-dimensional inverse problem with noisy data. above may also be useful for other nonlinear inverse problems where positivity constraints on the unknown parameter make regularization by relative entropy suitable.

Finally, although we have considered the setting of Lévy processes, this approach can also be adapted to other models with jumps –such as stochastic volatility models with jumps (see [13, Chapter 15] for a review)– where the jump structure is described by a Lévy measure, to be retrieved from observations.

Appendix A. Relative entropy for Lévy processes. In this appendix we explicitly compute the relative entropy of two Lévy processes in terms of their characteristic triplets and establish some properties of the relative entropy viewed as a functional on Lévy processes. Under additional assumptions the relative entropy of two Lévy processes was computed in [10] (where it is supposed that Q is equivalent to P and the Lévy process has finite exponential moments under P) and in [31] (where $\log \frac{d\nu^Q}{d\nu^P}$ is supposed bounded from above and below). We give here an elementary proof valid for all Lévy processes.

THEOREM A.1 (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 $Q \ll P$ (by Theorem IV.4.39 in [27], this implies that $A^Q = A^P$ and $\nu^Q \ll \nu^P$) 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:*

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

Proof. Let $\{X^c_t\}_{t\geq 0}$ 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 [27, 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

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

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 [27]). Moreover, since N' and N'' are Lévy processes and martingales, their stochastic exponentials are also martingales (Proposition 8.23 in [13]). Therefore,

$$\begin{aligned} I_T(Q|P) &= E^P[Z_T \log Z_T] \\ &= E^P[\mathcal{E}(N')_T\mathcal{E}(N'')_T \log \mathcal{E}(N')_T] + E^P[\mathcal{E}(N')_T\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T] \\ &= E^P[\mathcal{E}(N')_T \log \mathcal{E}(N')_T] + E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T] \end{aligned} \tag{A.2}$$

if these expectations exist.

Since $\Delta N'_t > -1$ a.s., $\mathcal{E}(N')_t$ is almost surely positive. Therefore, from Lemma 5.8 in [23], $U_t := \log \mathcal{E}(N')_t$ is a Lévy process with characteristic triplet:

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

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:

$$\begin{aligned} E^P[U_T e^{U_T}] &= -i\frac{d}{dz}E^P[e^{izU_T}]|_{z=-i} = -iT\psi'(-i)E^P[e^{U_T}] = -iT\psi'(-i) \\ &= T(A^U + \gamma^U + \int_{-\infty}^{\infty}(xe^x - x1_{|x|\le 1})\nu^U(dx)) \\ &= \frac{\beta^2 AT}{2} + T\int_I(\phi(x)\log\phi(x) + 1 - \phi(x))\nu^P(dx) \end{aligned} \tag{A.3}$$

It remains to compute $E^P[\mathcal{E}(N'')_T \log \mathcal{E}(N'')_T]$. Since N'' is a compound Poisson process, $\mathcal{E}(N'')_t = e^{bt}\prod_{s\le t}(1+\Delta N''_s)$, 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

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

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 $\Delta N''$:

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

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 (A.3) and (A.2) finishes the proof. $\square$

LEMMA A.2. *Let $P, \{P_n\}_{n\geq 1} \subset \mathcal{L}_B^+$ for some $B > 0$, such that $P_n \Rightarrow P$. Then for every $r > 0$, the level set $L_r := \{Q \in \mathcal{L} : I(Q|P_n) \leq r$ for some $n\}$ is tight.*

Proof. For any $Q \in L_r$, P_Q denotes any element of $\{P_n\}_{n\geq 1}$, for which $I(Q|P_Q) \leq r$. 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 A.1,

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

Therefore, for u sufficiently large,

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

which entails that for u sufficiently large,

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

uniformly with respect to $Q \in L_r$. Let $\varepsilon > 0$ and choose u such that $\int_{\{\phi^Q>u\}} \nu^Q(dx) \leq \varepsilon/2$ for every $Q \in L_r$. By Corollary VII.3.6 in [27],

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

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 $n \geq 1$ are finite outside a neighborhood of zero, we can choose a compact K such that $\nu^{P_n}(\mathbb{R} \setminus K) \leq \varepsilon/2u$ for every n. Then

$$\nu^Q(\mathbb{R} \setminus K) = \int_{(\mathbb{R}\setminus K)\cap\{\phi^Q\leq u\}} \phi^Q \nu^{P_Q}(dx) + \int_{(\mathbb{R}\setminus K)\cap\{\phi^Q>u\}} \nu^Q(dx) \leq \varepsilon \tag{A.4}$$

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

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

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

$$\begin{aligned}
&\left|\int_{|x|\leq 1} x(\phi^Q - 1)\nu^{P_Q}(dx)\right| \\
&\quad \leq \left|\int_{|x|\leq 1,\ \phi^Q\leq u} x(\phi^Q - 1)\nu^{P_Q}(dx)\right| + \left|\int_{|x|\leq 1,\ \phi^Q>u} x(\phi^Q - 1)\nu^{P_Q}(dx)\right| \\
&\quad \leq \int_{|x|\leq 1} x^2\nu^{P_Q}(dx) + \int_{|x|\leq 1,\ \phi^Q\leq u} (\phi^Q - 1)^2\nu^{P_Q}(dx) + 2\int_{\phi^Q>u} \phi^Q\nu^{P_Q}(dx) \\
&\quad \leq \int_{|x|\leq 1} x^2\nu^{P_Q}(dx) + 2u\int_{-\infty}^{\infty} (\phi^Q \log \phi^Q + 1 - \phi^Q)\nu^{P_Q}(dx) + \frac{4r}{T_\infty \log u} \\
&\quad \leq \int_{|x|\leq 1} x^2\nu^{P_Q}(dx) + \frac{3ru}{T_\infty}.
\end{aligned} \tag{A.5}$$

By Proposition VI.4.18 in [27], the tightness of $\{P_n\}_{n\geq 1}$ implies that

$$A^{P_n} + \int_{|x|\leq 1} x^2 \nu^{P_n}(dx) \tag{A.6}$$

is bounded uniformly on n, which means that the right hand side of (A.5) is bounded uniformly with respect to $Q \in L_r$. From Theorem IV.4.39 in [27], $A^Q = A^{P_Q}$ for all $Q \in L_r$ because for the relative entropy to be finite, necessarily $Q \ll P_Q$. Theorem A.1 then implies that

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

From (A.6), A^{P_n} is bounded uniformly on n. Therefore, inequality (A.5) shows that $|\gamma^Q|$ is bounded uniformly with respect to Q.

Once again, for u sufficiently large,

$$\begin{aligned} A^Q + \int_{-\infty}^{\infty} (x^2 \wedge 1)\phi^Q \nu^{P_Q}(dx) &\leq A^Q + u \int_{\phi^Q \leq u} (x^2 \wedge 1)\nu^{P_Q}(dx) \\ &+ \int_{\phi^Q > u} \phi^Q \nu^{P_Q}(dx) \leq A^{P_Q} + u \int_{-\infty}^{\infty} (x^2 \wedge 1)\nu^{P_Q}(dx) + \frac{2r}{T_\infty \log u} \end{aligned} \tag{A.7}$$

and (A.6) implies that the right hand side is bounded uniformly with respect to $Q \in L_r$. By Proposition VI.4.18 in [27], (A.4), (A.7) and the fact that $|\gamma^Q|$ is bounded uniformly with respect to Q entail that the set L_r is tight. □

Lemma A.3. *Let Q and P be two probability measures on $(\Omega, \mathcal{F})$. Then*

$$I(Q|P) = \sup_{f \in C_b(\Omega)} \left\{ \int_\Omega f dQ - \int_\Omega (e^f - 1) dP \right\}, \tag{A.8}$$

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 [32, Theorem 4]. □

Corollary A.4. *The relative entropy functional $I(Q|P)$ is weakly lower semi-continuous with respect to Q for fixed P.*

Lemma A.5. *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. Let $h : \mathbb{R} \to \mathbb{R}$ be a bounded continuous function such that $h(x) \equiv x$ on a neighborhood of 0. For every n, let (A_n, ν_n, γ_n) be the characteristic triplet of P_n with respect to truncation function h and let

$$f(\beta, P_n) := \gamma_n + \left(\frac{1}{2} + \beta\right) A_n + \int_{-\infty}^{\infty} \left\{(e^x - 1)e^{\beta(e^x - 1)} - h(x)\right\} \nu_n(dx).$$

The first step is to show that for every n, there exists a unique β_n such that $f(\beta_n, P_n) = 0$ and that the sequence $\{\beta_n\}_{n\geq 1}$ is bounded.

Since for every n, $P_n \in \mathcal{L}_B^+$, the dominated convergence theorem yields:

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

and since $P_n \in \mathcal{L}_{NA}$, the Lévy process (X, P_n) is not a.s. increasing nor a.s. decreasing, which means that at least one of the following conditions holds:

1. $A_n > 0$,
2. $\nu_n((-\infty, 0)) > 0$ and $\nu_n(0, \infty) > 0$,
3. $A_n = 0$, $\nu_n((-\infty, 0)) = 0$ and $\gamma_n - \int_{-\infty}^{\infty} h(x)\nu_n(dx) < 0$,
4. $A_n = 0$, $\nu_n((0, \infty)) = 0$ and $\gamma_n - \int_{-\infty}^{\infty} h(x)\nu_n(dx) > 0$.

Since clearly $f'_\beta(\beta, P_n) \geq A_n + \min\left(\int_{-\infty}^0 (e^x-1)^2\nu_n(dx), \int_0^\infty (e^x-1)^2\nu_n(dx)\right)$, if conditions 1 or 2 above hold, $f'_\beta(\beta, P_n)$ is bounded from below by a positive constant therefore

$$\exists! \beta_n : f(\beta_n, P_n) = 0. \tag{A.9}$$

If condition 3 above holds, $\lim_{\beta\to-\infty} f(\beta, P_n) = \gamma_n - \int_{-\infty}^{\infty} h(x)\nu_n(dx) < 0$ and $\lim_{\beta\to\infty} f(\beta, P_n) = \infty$, which means that (A.9) also holds. The case when condition 4 above is satisfied may be treated similarly.

Let us now show that the sequence $\{\beta_n\}_{n\geq 1}$ is bounded. For every n, $f(\beta, P_n)$ may be rewritten as follows:

$$\begin{aligned} f(\beta, P_n) := \gamma_n &+ \left(\frac{1}{2} + \beta\right)\left(A_n + \int_{-\infty}^{\infty} h^2(x)\nu_n(dx)\right) \\ &+ \int_{-\infty}^{\infty} \left\{(e^x-1)e^{\beta(e^x-1)} - h(x) - \left(\frac{1}{2}+\beta\right)h^2(x)\right\}\nu_n(dx). \end{aligned} \tag{A.10}$$

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 (A.10) is bounded on $(-\infty, B]$, by Corollary VII.3.6 in [27], for every β, $\lim_n f(\beta, P_n) = f(\beta, P)$.

Since P also belongs to $\mathcal{L}_B^+ \cap \mathcal{L}_{NB}$, by the same argument as above, there exists a unique β^* such that $f(\beta, P) = 0$ 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.

For every n, let (X, Q_n) be the Lévy process with characteristic triplet (with respect to h)

$$\begin{aligned} A_n^Q &= A_n, \\ \nu_n^Q &= e^{\beta_n(e^x-1)}\nu_n, \\ \gamma_n^Q &= \gamma_n + A_n\beta_n + \int_{-\infty}^{\infty} h(x)(e^{\beta(e^x-1)} - 1)\nu_n(dx). \end{aligned}$$

The measure Q_n is in fact the minimal entropy martingale measure for P_n, computed in [30]. From Theorem A.1,

$$I(Q_n|P_n) = -T\left\{\frac{\beta_n}{2}(1+\beta_n)A_n + \beta_n\gamma_n + \int_{-\infty}^{\infty} \{e^{\beta_n(e^x-1)} - 1 - \beta_n h(x)\}\nu_n(dx)\right\}.$$

To show that the sequence $\{I(Q_n|P_n)\}_{n\geq 1}$ 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 now follows from Theorem VI.4.18 in [27]. □

REFERENCES

[1] Y. Achdou, *An inverse problem for parabolic variational inequalities in the calibration of american options*, SIAM Journal of Control and Optimization, (to appear).

[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, *The minimum-entropy algorithm and related methods for calibrating asset-pricing models*, in Proceedings of the International Congress of Mathematicians, (Berlin, 1998), vol. III, Documenta Mathematica, 1998, pp. 545–563.

[4] ———, *Minimum entropy calibration of asset pricing models*, Int. J. Theor. Appl. Finance, 1 (1998), pp. 447–472.

[5] 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.

[6] M. Avellaneda, C. Friedman, R. Holmes, and D. Samperi, *Calibrating volatility surfaces via relative entropy minimization*, Applied Mathematical Finance, 4 (1997), pp. 37–64.

[7] D. Bates, *Jumps and stochastic volatility: the exchange rate processes implicit in Deutschemark options*, Rev. Fin. Studies, 9 (1996), pp. 69–107.

[8] S. Ben Hamida and R. Cont, *Recovering volatility from option prices by evolutionary optimization*, Rapport Interne 534, CMAP, Ecole Polytechnique, 2004.

[9] P. Carr and D. Madan, *Option valuation using the fast Fourier transform*, J. Comput. Finance, 2 (1998), pp. 61–73.

[10] T. Chan, *Pricing contingent claims on stocks driven by Lévy processes*, Ann. Appl. Probab., 9 (1999), pp. 504–528.

[11] A. S. Cherny and A. N. Shiryaev, *Change of time and measure for Lévy processes*. MaPhySto lecture notes series 2002-13, 2002.

[12] R. Cont, *Model uncertainty and its impact on the pricing of derivative instruments*, Rapport Interne 543, CMAP, Ecole Polytechnique, 2004. http://ssrn.com/abstract=562721.

[13] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall / CRC Press, 2004.

[14] R. Cont and P. Tankov, *Non-parametric calibration of jump-diffusion option pricing models*, J. Comput. Finance, 7 (2004).

[15] S. Crépey, *Calibration of the local volatility in a trinomial tree using tikhonov regularization*, Inverse Problems, 19 (2003), pp. 91–127.

[16] I. Csiszar, *I-divergence geometry of probability distributions and minimization problems*, Annals of Probability, 3 (1975), pp. 146–158.

[17] 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.

[18] B. Dupire, *Pricing with a smile*, RISK, 7 (1994), pp. 18–20.

[19] N. El Karoui and R. Rouge, *Pricing via utility maximization and entropy*, Math. Finance, 10 (2000), pp. 259–276.

[20] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

[21] H. W. Engl and G. Landl, *Convergence rates for maximum entropy regularization*, SIAM J. Numer. Anal., 30 (1993), pp. 1509–1536.

[22] M. Frittelli, *The minimal entropy martingale measure and the valuation problem in incomplete markets*, Math. Finance, 10 (2000), pp. 39–52.

[23] T. Goll and J. Kallsen, *Optimal portfolios for logarithmic utility*, Stochastic Process. Appl., 89 (2000), pp. 31–48.

[24] T. Goll and L. Rüschendorf, *Minimal distance martingale measures and optimal portfolios consistent with observed market prices*, in Stochastic processes and related topics (Siegmundsburg, 2000), vol. 12 of Stochastics Monogr., Taylor & Francis, London, 2002, pp. 141–154.

[25] J. M. Harrison and S. R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Process. Appl., 11 (1981), pp. 215–260.

[26] N. Jackson, E. Süli, and S. Howison, *Computation of deterministic volatility surfaces*, Journal of Computational Finance, 2 (1999), pp. 5–32.

[27] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer, Berlin, 2nd ed., 2003.

[28] P. Jakubenas, *On option pricing in certain incomplete markets*, Proceedings of the Steklov Mathematical Institute, 237 (2002).

[29] R. Lagnado and S. Osher, *A technique for calibrating derivative security pricing models: numerical solution of the inverse problem*, J. Comput. Finance, 1 (1997).

[30] Y. Miyahara and T. Fujiwara, *The minimal entropy martingale measures for geometric Lévy processes*, Finance Stoch., 7 (2003), pp. 509–531.

[31] L. Nguyen, *Calibration de Modèles Financiers par Minimisation d'Entropie Relative et Modèles aves Sauts*, PhD thesis, Ecole Nationale de Ponts et Chaussées, 2003.

[32] R. T. Rockafellar, *Integrals which are convex functionals*, Pacific journal of mathematics, 24 (1968), pp. 525–539.

[33] M. Rubinstein, *Implied binomial trees*, Journal of Finance, 49 (1994), pp. 771–819.

[34] D. Samperi, *Calibrating a diffusion pricing model with uncertain volatility: Regularization and stability*, Math. Finance, 12 (2002), pp. 71–87.

[35] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK, 1999.

[36] P. Tankov, *Lévy Processes in Finance: Inverse Problems and Dependence Modelling*, PhD thesis, Ecole Polytechnique, France, 2004.