Non-parametric calibration of jump–diffusion option pricing models

Rama Cont

Centre de Mathématiques Appliquées, CNRS – Ecole Polytechnique F 91128 Palaiseau, France

Peter Tankov

Centre de Mathématiques Appliquées, CNRS – Ecole Polytechnique F 91128 Palaiseau, France

We present a non-parametric method for calibrating jump–diffusion and, more generally, exponential Lévy models to a finite set of observed option prices. We show that the usual formulations of the inverse problem via non-linear least squares are ill-posed and propose a regularization method based on relative entropy: we reformulate our calibration problem into a problem of finding a risk-neutral exponential Lévy model that reproduces the observed option prices and has the smallest possible relative entropy with respect to a chosen prior model. Our approach allows us to reconcile the idea of calibration by relative entropy minimization with the notion of risk-neutral valuation in a continuous-time model. We discuss the numerical implementation of our method using a gradient-based optimization algorithm and show by simulation tests on various examples that the entropy penalty resolves the numerical instability of the calibration problem. Finally, we apply our method to data sets of index options and discuss the empirical results obtained.

1 Introduction

The inability of diffusion models to explain certain empirical properties of asset returns and option prices has led to the development, in option pricing theory, of a variety of models based on Lévy processes (Andersen and Andreasen, 2000; Eberlein, 2001; Eberlein, Keller and Prause, 1998; Cont, Bouchaud and Potters, 1997; Kou, 2002; Madan, 2001; Madan, Carr and Chang, 1998; Merton, 1976; Schoutens, 2002). A widely studied class is that of exponential Lévy processes in which the price of the underlying asset is written as $S_t = \exp(rt + X_t)$, where r is the discount rate and X is a Lévy process defined by its characteristic triplet (σ, ν, γ). While the main concern in the literature has been to obtain efficient analytical and numerical procedures for computing prices of various options, a preliminary step in using the model is to obtain model parameters – here the characteristic triplet of the Lévy process – from market data by calibrating the

Preliminary versions of this work were presented at the Bernoulli Society International Statistical Symposium (Taipei 2002), Maphysto (Aarhus and Copenhagen), Bachelier seminar (Paris, 2002), the University of Freiburg, the University of Warwick and INRIA. We thank Marco Avellaneda, Frédéric Bonnans, Mark Broadie, Stéphane Crépey, Dilip Madan and Monique Pontier for helpful remarks.

model to market prices of (liquid) call options. This amounts to solving the following *inverse problem*:

CALIBRATION PROBLEM 1 *Given the market prices of call options $C_0^*(T_i, K_i)$, $i \in I$ at $t = 0$, construct a Lévy process $(X_t)_{t \geq 0}$ such that the discounted asset price $S_t e^{-rt} = \exp X_t$ is a martingale and the market call option prices $C_0^*(T_i, K_i)$ coincide with the prices of these options computed in the exponential Lévy model driven by X:*

$$\forall i \in I, \quad C_0^*(T_i, K_i) = e^{-rT_i} E[(S_{T_i} - K_i)^+ \,|\, S_0] = e^{-rT_i} E[(S_0 e^{rT_i + X_{T_i}} - K_i)^+] \quad (1)$$

The index set I in the most general formulation need not be finite: for example, if we know market option prices for a given maturity *and* all strikes (which is, of course, unrealistic), the index set I will be continuous.

Note that, in order to price exotic options, we need to construct a risk-neutral *process*, not only its conditional densities (also called the state–price densities) as in Ait-Sahalia and Lo (1998).

Problem 1 can be seen as a generalized moment problem for the Lévy process X, which is typically an ill-posed problem: there may be no solution at all or an infinite number of solutions, and in the case where we use an additional criterion to choose one solution from many the dependence on input prices may be discontinuous, which results in numerical instabilities in the calibration algorithm.

To circumvent these difficulties we propose a *regularization* method based on the minimization of Kullback–Leibler information, or relative entropy, with respect to a prior model. Our method is based on the idea that, contrary to the diffusion setting where different volatility structures lead to singular (non-equivalent) probabilities on the path space (and therefore infinite relative entropy), two Lévy processes with different Lévy measures can define equivalent probabilities. It turns out that the relative entropy of exponential Lévy models is a simple function of their Lévy measures that can be used as a regularization criterion for solving the inverse problem 1 in stable way. Our approach leads to a non-parametric method for calibrating exponential Lévy models to option prices, extending similar methods previously developed for diffusion models (Samperi, 2002). However, the use of jump processes enables us to formulate the problem in a way that makes sense in a continuous-time framework without giving rise to singularities as in the diffusion calibration problem.

The paper is structured as follows. Section 2 defines the model setup and recalls some useful properties of Lévy processes and relative entropy. Section 3 proposes a well-posed formulation of the calibration problem as that of finding an exponential Lévy model that reproduces observed option prices and has the smallest possible relative entropy with respect to a prior jump–diffusion model. Section 4 discusses the numerical implementation of the calibration method in the framework of jump–diffusions,[1] the main ingredient of which is an explicit representation for the gradient of the criterion being minimized (Section 4.4).

[1] In this paper we use the term "jump–diffusion" to denote a Lévy process with a finite activity

To assess the performance of our method we first perform numerical experiments on simulated data: calibration is performed on a set of option prices generated from a given exponential Lévy model. Results are presented in Section 5: our algorithm enables us to calibrate the option prices with high precision and the resulting Lévy measure has little sensitivity to the initialization of the minimization algorithm. The precision of recovery of the Lévy measure is especially good for medium- and large-sized jumps, but small jumps are hard to distinguish from a continuous-diffusion component.

Section 6 presents empirical results obtained by applying our calibration method to a data set of DAX index options. Our tests reveal a density of jumps with strong negative skewness. While a small value of the jump intensity appears to be sufficient to calibrate the observed implied volatility patterns, the shape of the density of jump sizes evolves across maturities, indicating the need for departure from time-homogeneity.

2 Model setup

2.1 Lévy processes: definitions

A Lévy process is a stochastic process $(X_t)_{t\geq 0}$ with stationary independent increments, continuous in probability, having sample paths that are right-continuous with left limits ("cadlag") and satisfying $X_0 = 0$. The characteristic function of X_t has the following form, called the Lévy–Khinchin representation:

$$E\left[\mathrm{e}^{izX_t}\right] = \Phi_t(z) = \exp\left[t\psi(z)\right]$$

$$\psi(z) = -\frac{z^2\sigma^2}{2} + i\gamma z + \int_{-\infty}^{\infty} \left(\mathrm{e}^{izx} - 1 - izx\,1_{|x|\leq 1}\right) \nu(\mathrm{d}x) \tag{2}$$

where $\sigma \geq 0$, $\gamma \in \mathbb{R}$ and ν is a positive measure on $\mathbb{R}$ verifying

$$\nu(\{0\}) = 0 \qquad \int_{-1}^{+1} x^2 \nu(\mathrm{d}x) < \infty \qquad \int_{|x|>1} \nu(\mathrm{d}x) < \infty \tag{3}$$

We will denote the set of such measures by $\mathcal{L}(\mathbb{R})$. If the measure $\nu(\mathrm{d}x)$ admits a density with respect to the Lebesgue measure, we will call it the Lévy density of X and denote it by $\nu(x)$.

In general, ν is not a probability measure: $\int \nu(\mathrm{d}x)$ need not even be finite. In the case where $\lambda = \int \nu(\mathrm{d}x) < +\infty$, the Lévy process is said to be of finite activity, and the measure ν can then be normalized to define a *probability measure*, μ, on

of jumps, that is, a linear combination of a Brownian motion and a compound Poisson jump process.

$\mathbb{R}$, which can be interpreted as the distribution of jump sizes:

$$\mu(\mathrm{d}x) = \frac{\nu(\mathrm{d}x)}{\lambda} \tag{4}$$

In this case X is called a compound Poisson process and λ, which is the average number of jumps per unit time, is called the intensity of jumps. For compound Poisson processes it is not necessary to truncate small jumps and the Lévy–Khinchin representation reduces to

$$E\left[\mathrm{e}^{izX_t}\right] = \exp\left\{ t\left(-\frac{z^2\sigma^2}{2} + ibz + \int_{-\infty}^{\infty} (\mathrm{e}^{izx} - 1)\,\nu(\mathrm{d}x)\right)\right\} \tag{5}$$

where $b = \gamma - \int_{-1}^{1} x\nu(\mathrm{d}x)$. For further details of Lévy processes see Bertoin (1996), Jacod and Shiryaev (2003) and Sato (1999).

It is now time to say a few words about the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ on which the Lévy processes of interest are defined. Since the sample paths of $(X_t)_{t\in[0,T]}$ are cadlag, this process defines a probability measure of the space of cadlag functions on $[0, T]$. One can therefore choose Ω to be this space, $\mathcal{F}_t$ to be the history of the process between 0 and t completed by null sets and $\mathcal{F} = \mathcal{F}_T$.

2.2 Exponential Lévy models

Let $(S_t)_{t\in[0,T]}$ be the price of a financial asset modeled as a stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{Q})$. Under the hypothesis of absence of arbitrage there exists a measure equivalent to $\mathbb{Q}$ under which $(\mathrm{e}^{-rt}S_t)$ is a martingale. We will assume therefore without loss of generality that $\mathbb{Q}$ is already one such martingale measure.

We term "exponential Lévy model" a model where the dynamics of S_t under $\mathbb{Q}$ is represented as the exponential of a Lévy process:

$$S_t = \mathrm{e}^{rt+X_t} \tag{6}$$

Here X_t is a Lévy process with characteristic triplet (σ, ν, γ) and the interest rate, r, is included for ease of notation. Since the discounted price process $\mathrm{e}^{-rt}S_t = \mathrm{e}^{X_t}$ is a martingale, this gives a constraint on the triplet (σ, ν, γ):

$$\psi(-i) = 0 \Leftrightarrow \gamma = \gamma(\sigma, \nu) = -\frac{\sigma^2}{2} - \int \left(\mathrm{e}^y - 1 - y1_{|y|\le 1}\right)\nu(\mathrm{d}y) \tag{7}$$

We will assume that this relation holds in the sequel.

Different exponential Lévy models proposed in the financial modeling literature simply correspond to different parameterizations of the Lévy measure:

- Compound Poisson models: ν is a finite measure.
  - Merton model (Merton, 1976) – Gaussian jumps:

$$\nu = \frac{\lambda}{\delta\sqrt{2\pi}}\, e^{-\frac{(x-m)^2}{2\delta^2}}$$

  - Superposition of Poisson processes: $\nu = \sum_{k=1}^{n} \lambda_k \delta_{y_k}$, where δ_x is a measure that affects unit mass to point x.
  - Double exponential model (Kou, 2002): $\nu(x) = p\alpha_1 e^{-\alpha_1 x} 1_{x>0} + (1-p) \times \alpha_2 e^{\alpha_2 x} 1_{x<0}$.

- Infinite activity models.
  - Variance gamma (Madan, Carr and Chang, 1998): $\nu(x) = A|x|^{-1} \times \exp(-\eta_{\pm}|x|)$.
  - Tempered stable[2] processes (Koponen, 1995; Cont, Bouchaud and Potters, 1997): $\nu(x) = A_{\pm}|x|^{-(1+\alpha)} \exp(-\eta_{\pm}|x|)$.
  - Normal inverse gaussian process (Barndorff-Nielsen, 1998).
  - Hyperbolic and generalized hyperbolic processes (Eberlein, 2001; Eberlein, Keller and Prause, 1998).
  - Meixner process (Schoutens, 2002): $\nu(x) = Ae^{-ax}/x \sinh(x)$.

Detailed descriptions of many of the above models can be found in Cont and Tankov (2004a). The price of an option is computed as a discounted conditional expectation of its terminal payoff under the risk-neutral probability, $\mathbb{Q}$. By the stationarity and independence of increments of X_t, the value of a call option can be expressed as

$$\begin{aligned} C(t,S;T=t+\tau,K) &= e^{-r\tau} E[(S_T - K)^+ \,|\, S_t = S] \\ &= e^{-r\tau} E[(Se^{r\tau + X_\tau} - K)^+] = Ke^{-r\tau} E(e^{x+X_\tau} - 1)^+ \end{aligned} \tag{8}$$

where x is the *log forward moneyness*:

$$x = \ln(S/K) + r\tau \tag{9}$$

The rescaled option price, $u(\tau, x) = e^{r\tau} C(t,S;T=t+\tau,K)/K$, takes an especially simple form:

$$u(\tau, x) = E[(e^{x+X_\tau} - 1)^+] \tag{10}$$

This means that the entire pattern of call option prices for all dates, all values of the underlying, and all strikes and all maturities, which is *a priori* a four-

[2] Also called "truncated Lévy flights" in the physics literature (Koponen, 1995; Cont, Bouchaud and Potters, 1997) or CGMY processes in the finance literature.

parameter object, depends on only two parameters (log forward moneyness and time-to-maturity) in an exponential Lévy model.

2.3 Equivalence of measures for Lévy processes

One of the interesting properties of exponential Lévy models is that the class of martingale measures equivalent to a given exponential Lévy process is quite large. This remains true even if one restricts attention to the martingale measures under which the price process remains of exponential Lévy type. The following result gives a description of the set of Lévy processes equivalent to a given one. Similar results in a more general setting may be found in Jacod and Shiryaev (2003).

PROPOSITION 1 (Sato (1999), Theorems 33.1 and 33.2) *Let* $(X_t, \mathbb{P})$ *and* $(X_t, \mathbb{P}')$ *be two Lévy processes on* $(\Omega, \mathcal{F})$ *with characteristic triplets* (σ, ν, γ) *and* (σ', ν', γ'). *Then* $\mathbb{P}|_{\mathcal{F}_t}$ and $\mathbb{P}'|_{\mathcal{F}_t}$ *are mutually absolutely continuous for all t if and only if the following conditions are satisfied:*

- ❑ $\sigma = \sigma'$
- ❑ *The Lévy measures are mutually absolutely continuous with*

$$\int_{-\infty}^{\infty} \left(e^{\phi(x)/2} - 1\right)^2 \nu(dx) < \infty \tag{11}$$

where $\phi(x)$ *is the logarithm of the Radon–Nikodym density of* ν' *with respect to* ν: $e^{\phi(x)} = d\nu'/d\nu$.

- ❑ *If* $\sigma = 0$, *then in addition* γ' *must satisfy*

$$\gamma' - \gamma = \int_{-1}^{1} x(\nu' - \nu)(dx) \tag{12}$$

The Radon–Nikodym derivative is given by

$$\frac{d\mathbb{P}'|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = e^{U_t} \tag{13}$$

where $(U_t)_{t\geq 0}$ *is a Lévy process with characteristic triplet* $(\sigma_U, \nu_U, \gamma_U)$ *given by*

$$\sigma_U = \sigma\eta \tag{14}$$

$$\nu_U = \nu\phi^{-1} \tag{15}$$

$$\gamma_U = -\frac{1}{2}\sigma^2\eta^2 - \int_{-\infty}^{\infty} \left(e^y - 1 - y1_{|y|\leq 1}\right)(\nu\phi^{-1})(dy) \tag{16}$$

and η *is chosen so that*

$$\gamma' - \gamma - \int_{-1}^{1} x(\nu' - \nu)(\mathrm{d}x) = \sigma^2 \eta$$

Moreover, U_t *satisfies* $E^P[\mathrm{e}^{U_t}] = 1$ *for all* t.

Compound Poisson case A compound Poisson process is a pure jump Lévy process which has almost surely a finite number of jumps in every interval. This means that if two Lévy processes satisfy the conditions of mutual absolute continuity listed in Proposition 1 and one of them is of compound Poisson type, the other will also be of compound Poisson type since these processes must have the same almost sure behavior of sample functions. If the jump parts of both Lévy processes are of compound Poisson type, the conditions of the proposition are somewhat simplified:

COROLLARY 1 *Suppose that the jump part of* X_t *is of compound Poisson type. Then* $\mathbb{P}|_{\mathcal{F}_t}$ *and* $\mathbb{P}'|_{\mathcal{F}_t}$ *are mutually absolutely continuous for all t if and only if the following conditions are satisfied:*

- $\sigma = \sigma'$;
- *the jump part of* X'_t *is of compound Poisson type and the two jump size distributions are mutually absolutely continuous;*
- *if* $\sigma = 0$, *then we must in addition have* $b' = b$.

The Radon–Nikodym derivative is given by

$$\frac{\mathrm{d}\mathbb{P}'|_{\mathcal{F}_t}}{\mathrm{d}\mathbb{P}|_{\mathcal{F}_t}} = \mathrm{e}^{U_t} \tag{17}$$

where U_t *is a Lévy process with jump part of compound Poisson type. Its characteristic triplet is given by* (14)–(16).

PROOF First of all, the condition (11) is fulfilled automatically as

$$\int_{-\infty}^{\infty} \left(\mathrm{e}^{\phi(x)/2} - 1\right)^2 \nu(\mathrm{d}x) \leq 2 \int_{-\infty}^{\infty} \left(\nu(\mathrm{d}x) + \nu'(\mathrm{d}x)\right) < \infty$$

As can be seen from the form of its characteristic triplet (14)–(16), the Radon–Nikodym derivative process U_t also has a jump part of compound Poisson type because

$$\int_{-1}^{1} \nu_U(\mathrm{d}x) = \int_{-1}^{1} [\nu\phi^{-1}](\mathrm{d}x) = \int_{-1 \leq \phi(y) \leq 1} \nu(\mathrm{d}y) < \infty$$

□

2.4 Relative entropy for Lévy processes

The notion of *relative entropy* or *Kullback–Leibler distance* is often used as a measure of the closeness of two equivalent probability measures. In this section we recall its definition and properties and compute the relative entropy of the measures generated by two risk-neutral exponential Lévy models.

Let $\mathbb{P}$ and $\mathbb{Q}$ be two equivalent probability measures on $(\Omega, \mathcal{F})$. The relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined as

$$\mathcal{E}(\mathbb{Q}\,|\,\mathbb{P}) = E^{\mathbb{Q}}\left[\ln\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = E^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\ln\frac{d\mathbb{Q}}{d\mathbb{P}}\right]$$

If we introduce the strictly convex function $f(x) = x\ln x$, we can write the relative entropy as

$$\mathcal{E}(\mathbb{Q}\,|\,\mathbb{P}) = E^{\mathbb{P}}\left[f\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

It is readily observed that the relative entropy is a convex function of $\mathbb{Q}$. Jensen's inequality shows that it is always non-negative: $\mathcal{E}(\mathbb{Q}\,|\,\mathbb{P}) \geq 0$, with $\mathcal{E}(\mathbb{Q}\,|\,\mathbb{P}) = 0$ *only if* $d\mathbb{Q}/d\mathbb{P} = 1$ almost surely. The following result shows that, if $\mathbb{P}$ and $\mathbb{Q}$ correspond to exponential Lévy models, the relative entropy can be expressed in terms of the corresponding Lévy measures.

PROPOSITION 2 *Let $\mathbb{P}$ and $\mathbb{Q}$ be equivalent measures on $(\Omega, \mathcal{F})$ generated by exponential Lévy models with Lévy triplets $(\sigma, \nu^P, \gamma^P)$ and $(\sigma, \nu^Q, \gamma^Q)$ and suppose that $\sigma > 0$. The relative entropy $\mathcal{E}(\mathbb{Q}\,|\,\mathbb{P})$ is then given by*

$$\mathcal{E}(\mathbb{Q}|\mathbb{P}) = \frac{T}{2\sigma^2}\left\{\gamma^Q - \gamma^P - \int_{-1}^{1} x\left(\nu^Q - \nu^P\right)(dx)\right\}^2 + T\int_{-\infty}^{\infty}\left(\frac{d\nu^Q}{d\nu^P}\ln\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx) \tag{18}$$

If $\mathbb{P}$ and $\mathbb{Q}$ correspond to risk-neutral exponential Lévy models, ie, verify the condition (7), the relative entropy reduces to:

$$\mathcal{E}(\mathbb{Q}|\mathbb{P}) = \frac{T}{2\sigma^2}\left\{\int_{-\infty}^{\infty}\left(e^x - 1\right)\left(\nu^Q - \nu^P\right)(dx)\right\}^2 + T\int_{-\infty}^{\infty}\left(\frac{d\nu^Q}{d\nu^P}\ln\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx) \tag{19}$$

PROOF Consider two exponential Lévy models defined by (6). From the bijectivity of the exponential it is clear that the filtrations generated by X_t and S_t coincide. We can therefore equivalently compute the relative entropy of the log-price processes (which are Lévy processes). To compute the relative entropy of two Lévy processes we will use expression (13) for the Radon–Nikodym derivative:

$$\mathcal{E} = \int \ln\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\frac{d\mathbb{Q}}{d\mathbb{P}}\,d\mathbb{P} = E^P\left[U_T\,e^{U_T}\right] \tag{20}$$

where (U_t) is a Lévy process with characteristic triplet given by formulae (14)–(16). Let $\Phi_t(z)$ denote its characteristic function and $\psi(z)$ its characteristic exponent, that is,

$$\Phi_t(z) = E^P[e^{izU_t}] = e^{t\psi(z)}$$

Then we can write

$$E^P\left[U_T e^{U_T}\right] = -i\frac{d}{dz}\Phi_T(-i) = -iTe^{T\psi(-i)}\,\psi'(-i)$$

$$= -iT\psi'(-i)\,E^P\left[e^{U_T}\right] = -iT\psi'(-i)$$

From the Lévy–Khinchin formula we know that

$$\psi'(z) = -\sigma_U^2 z + i\gamma_U + \int_{-\infty}^{\infty}\left(ixe^{izx} - ix1_{|x|\le 1}\right)\nu_U(dx)$$

We can now compute the relative entropy as follows:

$$\mathcal{E} = \sigma_U^2 T + \gamma_U T + T\int_{-\infty}^{\infty}\left(xe^x - x1_{|x|\le 1}\right)\nu_U(dx)$$

$$= \frac{\sigma^2 T}{2}\eta^2 + T\int\left(ye^y - e^y + 1\right)\left(\nu^P\phi^{-1}\right)(dy)$$

$$= \frac{\sigma^2 T}{2}\eta^2 + T\int\left(\frac{d\nu^Q}{d\nu^P}\ln\frac{d\nu^Q}{d\nu^P} + 1 - \frac{d\nu^Q}{d\nu^P}\right)\nu^P(dx)$$

where η is chosen such that

$$\gamma^Q - \gamma^P - \int_{-1}^{1} x\left(\nu^Q - \nu^P\right)(dx) = \sigma^2\eta$$

Since we have assumed $\sigma > 0$, we can write

$$\frac{1}{2}\sigma^2\eta^2 = \frac{1}{2\sigma^2}\left\{\gamma^Q - \gamma^P - \int_{-1}^{1} x\left(\nu^Q - \nu^P\right)(\mathrm{d}x)\right\}^2$$

which leads to (18). If $\mathbb{P}$ and $\mathbb{Q}$ are martingale measures, we can express the drift, γ, using σ and ν:

$$\frac{\sigma^2}{2}\eta^2 = \frac{1}{2\sigma^2}\left\{\int_{-\infty}^{\infty}(\mathrm{e}^x - 1)(\nu^Q - \nu^P)(\mathrm{d}x)\right\}^2$$

Substituting the above in (18) yields (19). □

Observe that, due to time-homogeneity of the processes, the relative entropy (18) or (19) is a linear function of T: the relative entropy per unit time is finite and constant. The first term in the relative entropy (18) of the two Lévy processes penalizes the difference of drifts and the second one penalizes the difference of Lévy measures.

In the risk-neutral case the relative entropy depends only on the two Lévy measures ν^P and ν^Q. For a given reference measure ν^P, expression (19) viewed as a function of ν^Q defines a positive (possibly infinite) function on the set of Lévy measures $\mathcal{L}(\mathbb{R})$:

$$H : \mathcal{L}(\mathbb{R}) \to [0, \infty]$$

$$\nu^Q \to H(\nu^Q) = \mathcal{E}(\mathbb{Q}(\nu^Q, \sigma)), \mathbb{P}(\nu^P, \sigma)) \qquad (21)$$

We shall call H the relative entropy function. Its expression is given by (19). It is a positive convex function of ν^Q, equal to zero only when $\nu^Q \equiv \nu^P$.

Compound Poisson case When the jump parts of both Lévy processes are of compound Poisson type with jump intensities λ^Q and λ^P and jump size distributions μ^Q and μ^P, the relative entropy takes the following form in the risk-neutral case:

$$\frac{\mathcal{E}}{T} = \frac{\lambda^Q}{\lambda^P}\ln\frac{\lambda^Q}{\lambda^P} + \lambda^P - \lambda^Q + \frac{\lambda^Q}{\lambda^P}\int_{-\infty}^{\infty}\ln\left(\frac{\mu^Q(x)}{\mu^P(x)}\right)\mu^Q(x)\,\mathrm{d}x$$

$$+ \frac{1}{2\sigma^2}\left\{\int_{-\infty}^{\infty}\mathrm{d}x(\mathrm{e}^x - 1)\left(\lambda^P\mu^P(x) - \lambda^Q\mu^Q(x)\right)\right\}^2 \qquad (22)$$

Example 1

In many parametric models the relative entropy can be explicitly computed. As an example, let us consider two risk-neutral Merton models with the same volatility σ and with Lévy measures

$$\nu_P(x) = \frac{\lambda_P}{\delta_P\sqrt{2\pi}}\, e^{-\frac{(x-m_P)^2}{2\delta_P^2}} \quad \text{and} \quad \nu_Q(x) = \frac{\lambda_Q}{\delta_Q\sqrt{2\pi}}\, e^{-\frac{(x-m_Q)^2}{2\delta_Q^2}}$$

The relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ can be easily computed using formula (19):

$$T^{-1}\mathcal{E}(\mathbb{Q}|\mathbb{P}) = \lambda_Q \ln\frac{\lambda_Q}{\lambda_P} + \lambda_P - \lambda_Q + \lambda_Q\left\{\ln\frac{\delta_P}{\delta_Q} - \frac{1}{2} + \frac{(m_Q - m_P)^2 + \delta_Q^2}{2\delta_P^2}\right\}$$
$$+ \frac{1}{2\sigma^2}\left\{\lambda_Q\left(e^{m_Q+\delta_Q^2/2} - 1\right) - \lambda_P\left(e^{m_P+\delta_P^2/2} - 1\right)\right\} \tag{23}$$

Note that this expression is not a convex function of λ_Q, δ_Q and m_Q because the Lévy measure in the Merton model depends on the parameters in a non-linear way. Nevertheless, expression (23) inherits some nice properties of function (19): it is always finite and non-negative and is only equal to zero when the parameters of the two models coincide.

Example 2

In the previous example the probabilities $\mathbb{P}$ and $\mathbb{Q}$ were equivalent for all values of parameters and the relative entropy $\mathcal{E}(\mathbb{Q}\,|\,\mathbb{P})$ was always finite. However, the equivalence of measures is not a sufficient condition for the relative entropy to be finite. Let ν_Q be a Lévy measure with exponential tail decay (as, for example, in Kou's (2002) double exponential model) and let

$$\nu_P = \exp(-e^{x^2})\nu_Q$$

Then ν_P is also a Lévy measure. Its behavior in the neighborhood of zero is similar to that of ν_Q but its tails decay much faster. It can be easily seen that the relative entropy of a process with Lévy measure ν_Q with respect to a process with the same volatility and Lévy measure ν_P is always infinite. This means that unlike the equivalence of processes, which is not affected by the tails of Lévy measures, the *finiteness of relative entropy imposes some constraints on the tail behavior of Lévy measures*. We will observe this effect again in Section 6 in a non-parametric setting.

3 The calibration problem for exponential Lévy models

The calibration problem consists of identifying the Lévy measure ν and the volatility σ from a set of observations of call option prices. If we knew call option prices for one maturity and all strikes, we could deduce the volatility and the Lévy measure in the following way:

- ❑ Compute the risk-neutral distribution of log price from option prices using the Breeden–Litzenberger formula

$$q_T(k) = e^{-k}\{C''(k) - C'(k)\} \tag{24}$$

  where $k = \ln K$ is the log strike.
- ❑ Compute the characteristic function (2) of the stock price by taking the Fourier transform of q_T.
- ❑ Deduce σ and the Lévy measure from the characteristic function Φ_T. First, the volatility of the Gaussian component σ can be found as follows (see Sato (1999), p. 40):

$$\sigma^2 = \lim_{u\to\infty} -\frac{2\ln\Phi_T(u)}{Tu^2} \tag{25}$$

  Now, denoting $\psi(u) \equiv \ln\Phi_T(u)/T + \sigma^2u^2/2$, we can prove (see Sato (1999), equation 8.10) that

$$\int_{-1}^{1}\big(\psi(u) - \psi(u+z)\big)\mathrm{d}z = 2\int_{-\infty}^{\infty} e^{iux}\left(1 - \frac{\sin x}{x}\right)\nu(\mathrm{d}x) \tag{26}$$

  Therefore, the left-hand side of (26) is the Fourier transform of the positive finite measure $2(1 - \sin x/x)\,\nu(\mathrm{d}x)$. This means that this measure and, consequently, the Lévy measure ν can be uniquely determined from ψ by Fourier inversion.

Thus, if we knew with absolute precision a set of call option prices for all strikes and a *single* maturity, we could deduce all parameters of our model and hence compute option prices for other maturities. In this case, option price data for any other maturity can only contradict the information we already have but cannot give us any further information. However the procedure described above, which is similar to the Dupire (1994) formula in the case of diffusion models, is not applicable in practice for at least three different reasons.

First, call prices are only available for a finite number of strikes. This number may be quite small (between 10 and 40 in the empirical examples given below). Therefore the derivatives and limits in the formulae (24)–(26) are actually extrapolations and interpolations of the data and our inverse problem is largely underdetermined.

Second, if several maturities are present in the options data, the problem (1) with equality constraints will typically have no solution due to the model specification error: owing to the homogeneous nature of their increments, Lévy processes often fail to reproduce the term structure of implied volatilities (see the discussion of time-inhomogeneity in Section 6).

The third difficulty is due to the presence of observational errors (or simply bid–ask spreads) in the market data. Taking derivatives of observations as in equation (24) can amplify these errors, rendering unstable the result of the computation. Due to all these reasons, one can hope at best for a solution that approximately verifies the constraints, and it is necessary to reformulate problem (1) as an *approximation* problem.

3.1 Non-linear least squares

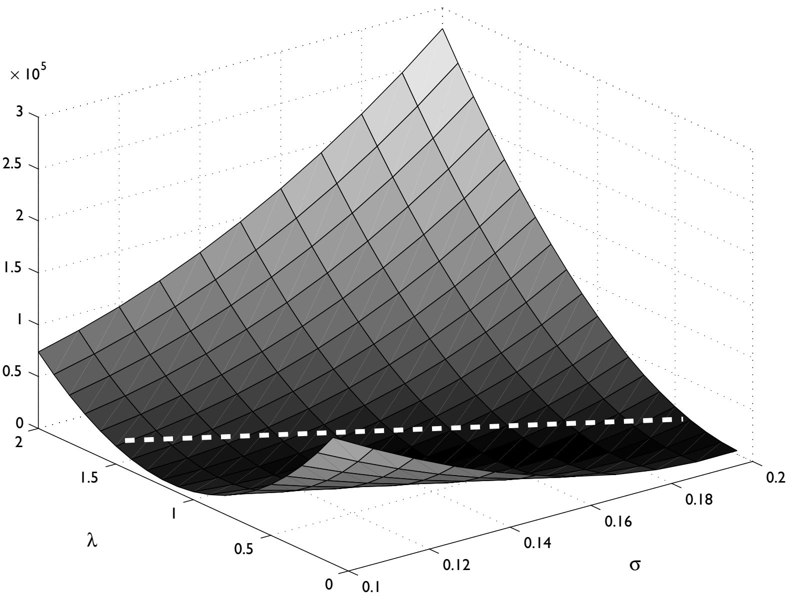
In order to obtain a practical solution to the calibration problem, many authors have resorted to minimizing the in-sample quadratic pricing error (see, for example, Andersen and Andreasen (2000) and Bates (1996a)):

$$(\sigma,\nu) = \arg\inf_{\sigma,\nu} \sum_{i=1}^{N} \omega_i \left| C^{\sigma,\nu}(t=0, S_0, T_i, K_i) - C_0^*(T_i, K_i) \right|^2 \tag{27}$$

where $C_0^*(K_i, T_i)$ is the market price of a call option observed at $t = 0$ and $C^{\sigma,\nu}(t = 0, S_0, T_i, K_i)$ is the price of this option computed in an exponential Lévy model with volatility σ and Lévy measure ν. The optimization problem (27) is usually solved numerically by a gradient-based method (Andersen and Andreasen, 2000; Bates, 1996a). While, contrarily to (1), one can always find some solution, the minimization function is non-convex, so a gradient descent may not succeed in locating the global minimum. Owing to the non-convex nature of the minimization function (27), two problems may arise, both of which reduce the quality of calibration algorithm.

The first issue is an identification problem: 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 will have flat regions in which the error has a low sensitivity to variations in model parameters. 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 in fact not true, as illustrated by the following empirical example. Figure 1 represents the magnitude of the quadratic pricing error for the Merton (1976) model 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 reducing 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 1). A gradient descent method

FIGURE 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 parameters being fixed. The dashed white line shows the "valley" along which the error function changes very little.

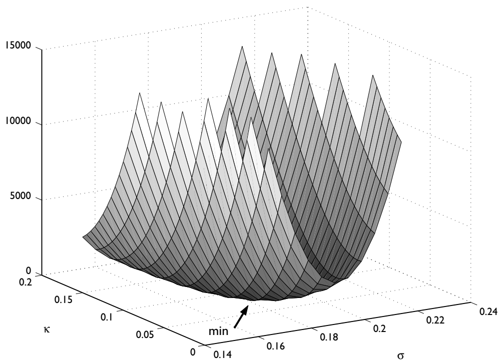


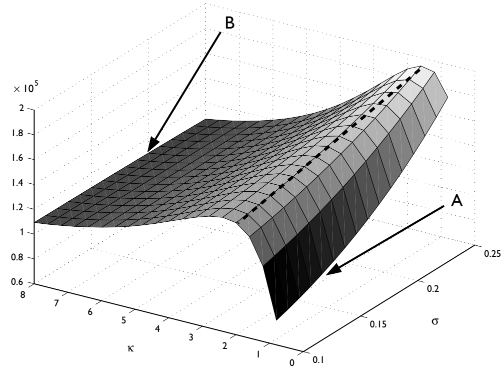
will typically succeed in locating the valley but will stop at a more or less random point in it. At first glance this does not seem to be too much of a problem: since the algorithm finds the valley's bottom, the best calibration quality will be achieved anyway. However, after a small change in option 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 even if the market option prices undergo only small changes. So, the issue is not so much the *precision* of the calibration but the *stability* of the parameters obtained.

The second problem is even more serious: since the the calibration function (27) is non-convex, it may have several local minima, and the gradient descent algorithm may stop in one of these local minima, leading to a much worse calibration quality than that of the true solution. Figure 2 illustrates this effect in the framework of the variance gamma model (Madan, Carr and Chang, 1998). In this model the Lévy process (X_t) is a pure jump one and its characteristic exponent is given by

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

FIGURE 2 Sum of squared differences between market prices (DAX options maturing in 10 weeks) and model prices in the variance gamma model as a function of σ and κ, the third parameter being fixed.

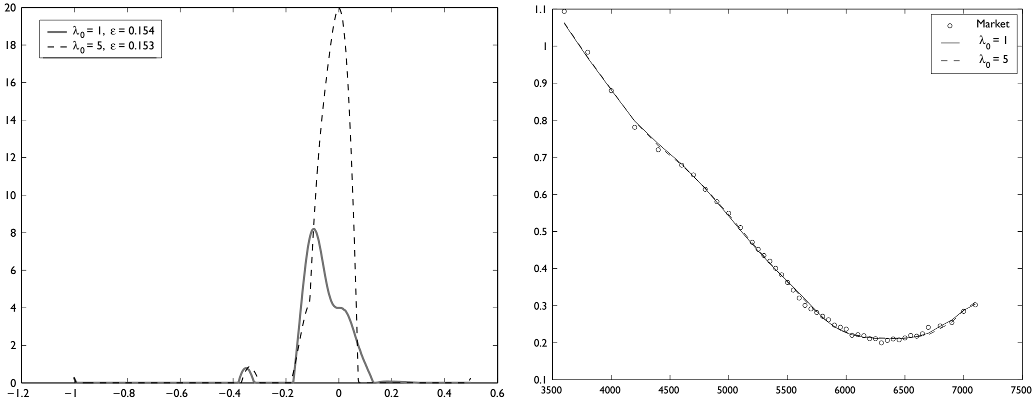




The left graph in Figure 2 shows the behavior of the objective function (27) in a small region around the global minimum. Since in this model there are only three parameters, the identification problem is not present, and a nice convex profile can be observed. However, when we look at the objective function on a larger scale (κ varies between 1 and 8), the convexity disappears completely and we observe a ridge (highlighted with a dashed black line) that separates two regions: if the minimization is initiated in region (A), the algorithm will eventually locate the minimum, but if we start in region (B), the gradient descent method will lead us away from the global minimum and the required calibration quality will never be achieved.

As a result the calibrated Lévy measure is very sensitive not only to the input prices but also to the numerical starting point in the minimization algorithm. Figure 3 shows an example of this instability 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. 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, in the second case (dashed line) a Merton model with intensity $\lambda = 5$. As can be seen in Figure 3 (left graph), the results of the minimization are totally different! However, although the calibrated measures are different, the prices to which they correspond are almost the same (see Figure 3, right graph), and the final values of the objective function for the two curves differ very little. This observation suggests that in the non-parametric setting we are more likely to find a flat "valley" in the error landscape rather than several distinct locally convex regions. The role of regularization is therefore to ensure continuous dependence of the calibrated measure on the data by helping to distinguish between measures

FIGURE 3 *Left*: Lévy measures calibrated to DAX option prices, maturity three months, using non-linear least-squares method. The starting jump measure for both graphs is Gaussian; the jump intensity λ_0 is initialized to one for the solid curve and to five for the dashed one. ε denotes the value of the objective function when the gradient descent algorithm stops. *Right*: implied volatility smiles corresponding to these two measures.



that give the same calibration quality. This comparison between parametric and non-parametric settings shows that the number of parameters is much less important from a numerical point of view than the *convexity* of the objective function to be minimized.

3.2 Regularization

The above remarks show that reformulating the calibration problem into a non-linear least-squares problem does not resolve the uniqueness and stability issues: the inverse problem remains ill-posed. To obtain a unique solution in a stable manner we introduce a regularization method (Engl, Hanke and Neubauer, 1996). One way to enforce uniqueness and stability of the solution is to add to the least-squares criterion (27) a convex *penalization* term:

$$(\sigma^*, \nu^*) = \arg\inf_{\sigma,\nu} \sum_{i=1}^{N} \omega_i \left| C^{\sigma,\nu}(t=0, S_0, T_i, K_i) - C_0^*(T_i, K_i) \right|^2 + \alpha F(\sigma,\nu) \tag{28}$$

where the term F, which is a measure of closeness of the model $\mathbb{Q}$ to a prior model $\mathbb{Q}_0$, is chosen such that the problem (28) becomes well-posed. Problem (28) can be understood as that of finding an exponential Lévy model satisfying the conditions (1), which is closest in some sense – defined by F – to a prior exponential Lévy model.

3.2.1 Choice of the regularization function

Several choices are possible for the penalization term. From the point of view of the uniqueness and stability of the solution, the criterion used should be convex with respect to the parameters (here, the Lévy measure). It is this convexity which was lacking in the non-linear least-squares criterion (27).

Two widely used regularization techniques for ill-posed inverse problems involve penalization by a quadratic function (also called Tikhonov regularization) (Engl, Hanke and Neubauer, 1996) and penalization by relative entropy with respect to a prior (see Engl, Hanke and Neubauer (1996), Sections 5.3 and 10.6). Tikhonov regularization uses the squared norm of the distance to a prior parameter as the regularization criterion: it is well suited when the parameters of interest form a Hilbert space with a natural notion of distance. Since Lévy measures do not form a Hilbert space and only positive measures are of interest, from the numerical point of view it would be desirable to find a regularization function that somehow incorporates this positivity requirement. The relative entropy (see Section 2.4) or Kullback–Leibler distance, $\mathcal{E}(\mathbb{Q}\,|\,\mathbb{Q}_0)$, of the the pricing measure $\mathbb{Q}$ with respect to some prior model $\mathbb{Q}_0$ has just this property: from the definition

$$\mathcal{E}(\mathbb{Q}\,|\,\mathbb{Q}_0) = E^{\mathbb{Q}}\left[\ln\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{Q}_0}\right]$$

it is clear that the relative entropy is only defined if $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{Q}_0$. Hence, if the prior Lévy measure is positive, the calibrated Lévy measure must also remain positive. From the numerical point of view this means that if, during the calibration process, the calibrated Lévy measure approaches zero, its gradient becomes arbitrarily large and is directed away from zero. So, one does not need to impose the positivity constraint explicitly.

Another advantage of the relative entropy, this time from a theoretical viewpoint, is that it is easily computable both in terms of probability measures on paths and in terms of the volatility, σ, and calibrated Lévy measure, ν: $\mathcal{E}(\mathbb{Q}\,|\,\mathbb{Q}_0) = F(\sigma, \nu) = H(\nu)$, where H, given by (21), is a convex function of the Lévy measure ν, with a unique minimum at $\nu = \nu_0$. On one hand, this enables one to use the probabilistic and information-theoretic interpretation of relative entropy: as explained in Section 2.4, this function plays the role of a pseudo-distance of the (risk-neutral) measure from the prior, and minimizing it corresponds to adding the least possible amount of information to the prior in order to correctly reproduce observed option prices (Engl, Hanke and Neubauer (1996), Section 5.3). On the other hand, the explicit expression of relative entropy in terms of the Lévy measure allows the construction of an efficient numerical method for finding the minimal entropy Lévy process. This process can be seen as a computable approximation to the minimal entropy martingale measure, which is a well-studied object in the financial literature (see Section 3.3).

3.2.2 Choice of the prior measure

The role of the prior probability measure, with respect to which relative entropy will be calculated, can hardly be overestimated. As shown by Figure 3, the option prices simply do not contain enough information to reproduce the Lévy measure in a stable way; new information must be added and this is done by introducing the prior. This improves the stability of the calibration but also introduces a bias of the calibrated measure towards the prior model. The calibration procedure can therefore be seen as a method to correct our initial knowledge of the Lévy measure (reflected by the prior) so that available option prices are reproduced correctly. The choice of the prior is thus a very important step in the algorithm.

One possible choice of prior is the exponential Lévy model estimated from historical data for the underlying. This choice ensures that the calibrated measure is absolutely continuous with respect to the historical measure, which is required by the absence of arbitrage in the market. In this case, even though the prior model is not risk-neutral, the calibrated model will be risk-neutral because of the martingale condition imposed on the calibrated Lévy process. When the historical data are not available, or considered unreliable, one can simply take a parametric model with "reasonable" parameter values, reflecting our views of the market. This parametric model will then be corrected by the calibration algorithm to incorporate the market prices of traded options. This choice of the prior measure will be discussed in more detail in Section 4.2. The third way to choose the prior is to take the calibrated measure of the day before, thus ensuring maximum stability of calibrated measures over time.

The calibration problem now takes the following form.

CALIBRATION PROBLEM 2 *Given a prior exponential Lévy model* $\mathbb{Q}_0$ *with characteristics* (σ_0, ν_0), *find a Lévy measure n which minimizes*

$$\mathcal{J}(\nu) = \alpha H(\nu) + \sum_{i=1}^{N} \omega_i \left(C_0^{\nu}(S_0, T_i, K_i) - C_0^{*}(T_i, K_i) \right)^2 \tag{29}$$

where $H(\nu)$ *is the relative entropy of the risk-neutral measure with respect to the prior, whose expression is given by (21).*

Here the weights ω_i are positive and sum to one; they reflect the pricing error tolerance for the option i. The choice of weights is addressed in more detail in Section 4.1.

The function (29) consists of two parts: the relative entropy function, which is convex in its argument ν, and the quadratic pricing error, which measures the precision of calibration. The coefficient α, called the "regularization parameter", defines the relative importance of the two terms: it characterizes the trade-off between prior knowledge of the Lévy measure and the information contained in option prices.

3.2.3 Discretized calibration problem

To implement the algorithm numerically without imposing some *a priori* parametric form on the Lévy measure, we discretize the Lévy measure on a grid. This is done by first localizing the Lévy measure on some bounded interval $[-M, M]$ and then choosing a partition, $\pi = (-M = x_1 < \ldots < x_N = M)$, of this interval. Define now L_π as the set of Lévy measures with support in π:

$$L_\pi = \left\{ \sum_{x \in \pi} a(x)\delta_x, \quad a \in (R^+)^\pi \right\} \tag{30}$$

where δ_x is a measure that affects unit mass to point x. Taking π to be a *finite* set of points, we implicitly assume that the Lévy measure is finite; that is, from now on we are working in the jump–diffusion framework (recall that in our terminology a jump–diffusion is a Lévy process with finite jump activity). Using the representation (30) means that we fix in advance the possible jump sizes $\{x_i\}$ of the Lévy processes and calibrate their intensities. In other words, our non-parametric Lévy process is a superposition of a (large number of) Poisson processes with different intensities. The discretized calibration problem now becomes

$$\inf_{\nu \in L_\pi} \mathcal{J}(\nu) \tag{31}$$

In the following we study this discretized problem, show that it is well-posed and develop a robust numerical method for solving it. The properties of the continuum version (29) and the convergence of the solutions of the discretized problem (31) are discussed in the companion paper (Cont and Tankov, 2004b). The following proposition shows that the use of entropy penalization makes our (discretized) problem well posed and hence numerically feasible.

PROPOSITION 3 (WELL-POSEDNESS OF THE REGULARIZED PROBLEM AFTER DISCRETIZATION) (*i*) *For any partition,* π, *of* $[-M, M]$, *the discretized calibration problem (29)–(31) admits a solution:*

$$\exists \nu^\pi \in L_\pi, \qquad J(\nu^\pi) = \min_{\nu \in L_\pi} \mathcal{J}(\nu) \tag{32}$$

If in addition the volatility coefficient s is non-zero, then for α *large enough the solution is unique.*

(ii) Every solution, ν^π, *of the regularized problem depends continuously on the vector of input prices* $(C^*(T_i, K_i),\ i = 1 \ldots n)$ *and for a suitable choice of* α *converges to a minimum-entropy least-squares solution when the error level on input prices tends to zero.*

PROOF See Appendices B and *C*. □

If α is large enough, the convexity properties of the entropy function stabilize the solution of problem (29). When $\alpha \to 0$, we recover the non-linear least-squares criterion (27). Therefore the correct choice of α is important: it cannot be fixed in advance but its "optimal" value depends on the data at hand and the level of error δ (see Section 4.3).

3.3 Relation to previous literature

3.3.1 Relation to minimal entropy martingale measures

The concept of relative entropy has been used in several contexts as a criterion for selecting pricing measures (Avellaneda, 1998; El Karoui and Rouge, 2000; Föllmer and Schied, 2002; Goll and Rüschendorf, 2001; Kallsen, 2001; Fritelli, 2000; Miyahara and Fujiwara, 2003). We briefly recall them here in relation to the present work.

In the absence of calibration constraints, the problem studied above reduces to that of identifying the equivalent martingale measure with minimal relative entropy with respect to a prior model. This problem has been widely studied and it is known that this unique pricing measure, called the minimal entropy martingale measure (MEMM) is related to the problem of portfolio optimization by maximization of exponential utility (El Karoui and Rouge, 2000; Föllmer and Schied, 2002; Fritelli, 2000; Miyahara and Fujiwara, 2003): if $u_\beta(X)$ is the utility indifference price for a random terminal payoff X for an investor with utility function $U(x) = \exp(-\beta x)$, then $E^{\mathbb{Q}}[e^{-rT}X]$ corresponds to the limit of $u_\beta(X)$ as $\beta \to 0$. Although we only consider here the class of measures corresponding to Lévy processes, if the prior measure is a Lévy process, then the MEMM is known to define again a Lévy process (Miyahara and Fujiwara, 2003). However, the notion of MEMM does not take into account the information obtained from observed option prices.

To take into account the prices of derivative products traded in the market, Kallsen (2001) introduced the notion of a consistent pricing measure, that is, a measure that correctly reproduces the market-quoted prices for a given number of derivative products. He studied the relation of the minimal entropy-consistent martingale measure (the martingale measure that minimizes the relative entropy distance to a given prior and respects a given number of market prices) to exponential hedging. He found that this MECMM defines the "least favorable consistent market completion" in the sense that it minimizes the exponential utility of the optimal trading strategy over all consistent martingale measures (see also El Karoui and Rouge, 2000). It satisfies

$$Q = \arg\min_{Q} \max_{X} \left\{ E^P\left(u\left(e + X - E^Q(X)\right)\right)\right\}$$

where the min is taken over all consistent equivalent martingale measures, the max is taken over all $\mathcal{F}_T$-measurable random variables, P is the prior/historical measure and e is the initial capital.

The minimal entropy measure studied in this article is not equivalent to the MECMM studied by Kallsen because we impose an additional restriction that the calibrated measure should stay in the class of measures corresponding to Lévy processes. It can be shown that the two measures only coincide in the case where there are no calibration constraints. However, where calibration constraints are present our measure can be seen as an approximation of the MECMM that stays in the class of Lévy processes. The usefulness of this approximation is clear: whereas the MECMM is an abstract notion for which one can at most assert existence and uniqueness, the one studied here is actually computable (see below) and can easily be used directly for pricing purposes. Therefore our framework can be regarded as a *computable* approximation of Kallsen's (2001) minimal entropy-constrained martingale measure.

3.3.2 Relation to the weighted Monte Carlo calibration method

Avellaneda (1998), Avellaneda *et al.* (2001) and Samperi (2002) and collaborators have proposed a non-parametric method based on relative entropy minimization for calibrating a pricing measure. In Avellaneda (1998) the calibration problem is formulated as one of finding a pricing measure which minimizes relative entropy with respect to a prior given calibration constraints:

Calibration problem 3

$$\mathbb{Q} = \arg\min_{Q \sim \mathbb{Q}_0} \mathcal{E}(Q, \mathbb{Q}_0) \quad \text{under} \quad E^Q\big(S(T_i) - K_i\big)^+ = C_t^*(T_i, K_i), \quad i = 1 \ldots n \tag{33}$$

where minimization is performed over all (not necessarily "risk-neutral") probability measures $\{Q\}$ equivalent to $\mathbb{Q}_0$. Problem (33) is still ill-posed because the equality constraints may be impossible to verify simultaneously due to model misspecification: a solution may not exist. However, it is not necessary to consider equality constraints like those in (33) since the market option prices are not exact but always quoted as bid–ask intervals. In a subsequent work, Avellaneda *et al.* (2001) consider a regularized version of problem (33) with quadratic penalization of constraints:

$$\mathbb{Q} = \arg\min_{Q \sim Q_0} \mathcal{E}(Q, \mathbb{Q}_0) + \sum_{i=1}^{n} \left| C^*(T_i, K_i) - E^Q\big(S(T_i) - K_i\big)^+ \right|^2 \tag{34}$$

In both cases the state space is discretized and the problem solved by a dual method: the result is a calibrated (but not necessarily "risk-neutral") probability distribution on a finite set of paths generated from the prior Q_0.

Although our formulation (29) looks quite similar to (34), there are several important differences.

First, note that the numerical solution of our problem (29) is done through discretization of the *parameter* space, not the state space Ω: the solution of (29) corresponds to a well-defined continuous-time process. By contrast, in Avellaneda *et al.* (2001) the discretization is applied to the state space: Ω is replaced by a finite set of sample paths generated by Monte Carlo simulation. Therefore the weighted Monte Carlo algorithm produces a measure $\mathbb{Q}_N$ on a finite set of paths Ω_N but which cannot be used to reconstruct a continuous-time process. The limit $N \to \infty$ is very subtle and not easy to describe.

Second, while the minimization in (34) is performed over all probability measures equivalent to the prior (the optimization variables are the probabilities themselves), in our case the minimization is performed over equivalent measures corresponding to jump–diffusion (exponential Lévy) models parameterized by their Lévy measure, ν. Although restricting the class of models, this approach has an advantage: it guarantees that we remain in the class of risk-neutral models, which is not the case in Avellaneda *et al.* (2001).

Third, while in Avellaneda *et al.* (2001) the optimization variable is the (discretized) probability measure Q itself, in our case the optimization variable is the Lévy measure ν. As a consequence, whereas the weighted Monte Carlo method yields a set of weights on trajectories, which is then used to price other options by Monte Carlo, our method yields a local description of the process (ie, its infinitesimal generator) through knowledge of ν. In particular, to price options one can use either Monte Carlo methods or solve the associated partial integro-differential equation, which may be preferable for American or barrier options.

Finally, even when Monte Carlo methods are used to price other options once the model is calibrated, it should be noted that in the weighted Monte Carlo method pricing is done using the original sample paths simulated from the prior model. Our approach has the advantage that we do not depend on the original set of paths to perform the Monte Carlo. Indeed, the posterior (calibrated) measure may be quite different from the prior, rendering many of the initial paths useless for computing expectations under the calibrated measure. Knowing the Lévy measure ν allows us to generate new paths under $\mathbb{Q}$.

4 Numerical implementation

As explained in Section 3, we tackle the ill-posedness of the initial calibration problem by transforming it into an optimization problem (29). We now describe a numerical algorithm for solving the discretized version (31) of this optimization problem. As mentioned above, for the numerical implementation we make the additional hypothesis that both the prior and the calibrated Lévy process have finite jump activity – that is, our numerical method allows us to calibrate a jump–diffusion model to market options data. This restriction is not as important as it may seem because, as we will see later, a jump–diffusion model allows to calibrate option prices with high precision even if they were generated by an infinite activity model. There are four main steps in the numerical solution:

- ❑ choice of the weights assigned to each option in the objective function;
- ❑ choice of the prior measure $\mathbb{Q}_0$ from the data;

❑ choice of the regularization parameter α;
❑ solution of the optimization problem for given α and $\mathbb{Q}_0$;

We shall describe each of these steps in detail below. This sequence of steps can be repeated a few times to minimize the influence of the choice of the prior.

4.1 The choice of weights in the minimization function

The relative weights, ω_i, of option prices in the minimization function should reflect our confidence in individual data points, which is determined by the liquidity of a given option. This can be assessed from the bid–ask spreads, but the bid and ask prices are not always available from option price databases. On the other hand, it is known that, at least for those options which are not too far from the money, the bid–ask spread is of the order of tens of basis points (<1%). This means that, to have errors proportional to the bid–ask spreads, one must minimize the differences of implied volatilities and not those of the option prices. However, this method involves many computational difficulties (numerical inversion of the Black–Scholes formula at each minimization step). A feasible solution to this problem is to minimize the squared differences of option prices weighted by the Black–Scholes "vegas" evaluated at the implied volatilities of the market option prices:

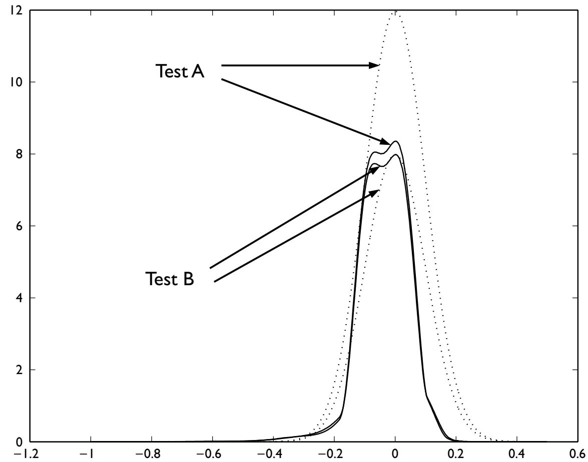
$$\sum_{i=1}^{N}\left(I\left(C^{\nu}(T_i,K_i)\right)-I_i\right)^2 \approx \sum_{i=1}^{N}\frac{\partial I}{\partial C}(I_i)\,|\,C^{\nu}(T_i,K_i)-C_i^*|^2$$
$$= \sum_{i=1}^{N}\frac{\left(C^{\nu}(T_i,K_i)-C_i\right)^2}{\text{Vega}^2(I_i)} \tag{35}$$

Here $I(\cdot)$ denotes the Black–Scholes implied volatility as a function of the option price and I_i denotes the market-implied volatilities.

4.2 Determination of the prior

As mentioned in Section 3.2, the prior reflects the user's view of the model. It is one of the most important ingredients of the method and the only one that cannot be determined completely automatically. The user should therefore specify a Lévy measure ν_0 and a diffusion coefficient σ_0. For example, these could be the result of the statistical estimation of a jump–diffusion model for the time series of asset returns. Alternatively, the prior can simply correspond to a model that seems "reasonable" to the user. Typically, however, the user may not have such detailed views and it is important to have a procedure to generate a reference measure $\mathbb{Q}_0$ automatically from options data. To do this we use an auxiliary jump–diffusion model (eg, Merton model) described by the volatility parameter σ_0 and a few other variables (denoted by θ) that parameterize the Lévy measure: $\nu_0 = \nu_0(\theta)$.

FIGURE 4 Sensitivity of implied Lévy densities to perturbations of prior model parameters. Solid curves represent calibrated Lévy densities and dotted curves depict the priors.



This model is then calibrated to data using the standard least-squares procedure (29):

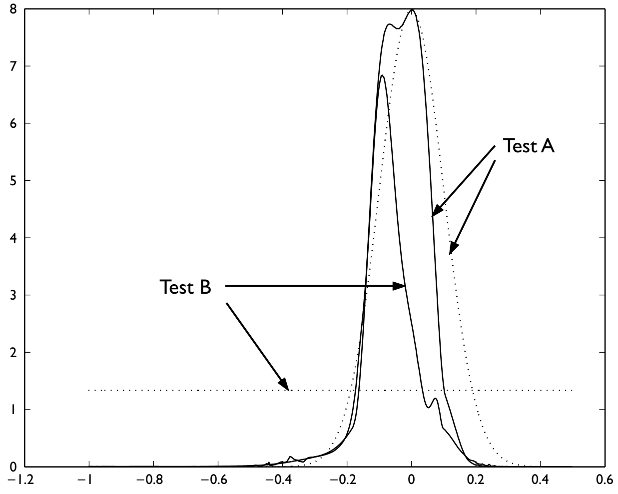
$$(\sigma_0, \nu_0) = \arg\inf_{\sigma,\theta} \epsilon(\sigma, \nu(\theta))$$

$$\epsilon(\sigma, \nu(\theta)) = \sum_{i=1}^{N} \omega_i \left| C^{\sigma,\nu(\theta)}(t_0 = 0, S_0, T_i, K_i) - C_0^*(T_i, K_i) \right|^2 \tag{36}$$

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 (eg, averages over a long period) and fix them once and for all. Since the objective function (36) is not convex, a simple gradient procedure may not give the global minimum. However, the solution (σ_0, ν_0) will be corrected at later stages and should only be viewed as a way of regularizing the optimization problem (29), so the minimization procedure at this stage need not be very precise.

To assess the influence of the prior on the results of calibration we carried out two series of numerical tests. In the first series the Lévy measure was calibrated twice to the same set of option prices using prior models that were different but closely similar. Namely, in test A we used a Merton model with diffusion volatility $\sigma = 0.2$, zero mean jump size, a jump standard deviation of 0.1 and a jump intensity $\lambda = 3$, whereas in test B all the parameters except the jump intensity had the same values and the intensity was equal to 2. The results of the tests are shown in Figure 4. The solid curves correspond to calibrated measures and the dotted ones depict the prior measures. Notice that there is very little difference

FIGURE 5 Sensitivity of implied Lévy measures to qualitative change of the prior model. Solid curves represent calibrated measures and dotted curves depict the priors.



between the calibrated measures, which means that the result of calibration is robust to minor variations of the parameters of the prior measure as long as its qualitative shape remains the same.

In the second series of tests we again calibrated the Lévy measure twice to the same set of option prices, this time taking two radically different priors. In test A we used a Merton model with diffusion volatility $\sigma = 0.2$, zero mean jump size, a jump standard deviation of 0.1 and jump 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 5) are still similar but exhibit many more differences than in the first series of tests. Not only do they differ in the middle, but the behavior of tails of the calibrated Lévy measure with uniform prior is also more erratic than when the Merton model was used as the prior.

Comparison of Figures 4 and 5 shows that the exact values of the parameters of the prior model are not very important but that it is crucial to find the right shape of the prior.

4.3 Determination of the regularization parameter

As remarked above, the regularization parameter α determines the trade-off between the accuracy of calibration and the numerical stability of the results with respect to the input option prices. It is therefore plausible that the right value of α should depend on the data at hand and should not be determined *a priori*.

One way to achieve this trade-off is by using the Morozov discrepancy principle (Morozov, 1966). First, we need to estimate the "intrinsic error", ϵ_0, present in the data, that is, the lower bound on the possible or desirable quadratic calibration error. Here we distinguish two cases, depending on the data that are available as input:

- If bid prices and ask prices are available for each calibration constraint, the *a priori* error level can be computed as:

$$\epsilon_0^2 = \sum_{i=1}^{N} \omega_i \left| C_i^{\text{bid}} - C_i^{\text{ask}} \right|^2 \tag{37}$$

- If confidence intervals/bid–ask intervals are not available, the *a priori* error level ϵ_0 must be estimated from the data themselves. In this case a possible solution is to substitute the "market error" with the "model error". First, we minimize the quadratic pricing error (27). The value of the calibration function at the minimum $\epsilon_0 \equiv \epsilon_{\alpha=0}$ can be interpreted as a measure of "model error": if $\epsilon_0 = 0$, it means that perfect calibration is achieved by the model; but, typically, due to the specification error or errors in the data, $\epsilon_0 > 0$. It can be seen as the "distance" of market prices from model prices, ie, it gives an *a priori* level of quadratic pricing error that one cannot really hope to improve on while keeping to the same class of models. Note that here we only need to find the minimum value of (27) and not to locate its minimum, so a rough estimate is sufficient and the presence of "flat" directions is not a problem.

Now let (σ, ν_α) be the solution of (31) for a given regularization parameter $\alpha > 0$. Then the *a posteriori* quadratic pricing error is given by $\epsilon(\sigma, \nu_\alpha)$, which one would expect to be a bit larger than ϵ_0 since, by adding the entropy term, we have sacrificed some precision for a gain in stability. The Morozov discrepancy principle consists in authorizing a loss of precision that is of the same order as the model error by choosing α such that

$$\epsilon_0 \simeq \epsilon(\sigma, \nu_\alpha) \tag{38}$$

In practice we fix some $\delta > 1$, $\delta \simeq 1$ (for example, $\delta = 1.1$) and numerically solve

$$\delta\epsilon_0 = \epsilon(\sigma, \nu_\alpha) \tag{39}$$

The left-hand side is a differentiable function of α, so the solution can be obtained with few iterations – for example, by Newton's (or a bisection) method.

4.4 Computation of the gradient

In order to minimize the function (31) using a BFGS gradient descent method,[3] the essential step is the computation of the gradient of the calibration function with respect to the discretized values of the Lévy measure. The discretization grid

[3] For a description of the algorithm see Press *et al.* (1992). In our numerical examples we used the LBFGS implementation by Byrd, Lu and Nocedal (1995).

for the Lévy measure ν is of the form (x_i, $i = 1 \ldots N$), where $x_i = x_0 + i\Delta x$. The grid must be uniform for the FFT algorithm to be used for option pricing.

If we were to compute the gradient numerically, the complexity would increase by a factor equal to the number of grid points. A crucial point of our method is that we are able to compute the gradient of the option prices with only a twofold increase of complexity compared to computing prices alone. Due to this optimization, the execution time of the program reduces on average from several hours to about a minute on a standard PC.

Below, to simplify the formulae, all computations are carried out in the continuous case (we compute the variational derivative). In the discretized case the idea is the same, but the Fréchet derivative is replaced by the usual gradient and the formulae become more cumbersome.

To emphasize the dependence of all quantities on the Lévy measure, it will appear explicitly as an argument in square brackets below. The main step is to compute the variational derivative of the option price, $DC_T(K)[\nu]$. Since the intrinsic value of the option does not depend on the Lévy measure, computing the derivative of the option price is equivalent to computing the derivative of the time value, $z_T(k)[\nu]$, defined by formula (A5) of Appendix A. The function which maps the Lévy measure ν into the time value $z_T(k)[\nu]$ is a superposition of the Lévy–Khinchin formula (2) and equation (A9) of Appendix A. Let us take an admissible test function h and compute the directional derivative of $z_T(k)[\nu]$ in the direction h. By definition

$$D_h z_T(k)[\nu] = \frac{\partial}{\partial \varepsilon}\left\{z_T(k)[\nu + \varepsilon h]\right\}\Big|_{\varepsilon=0}$$

Under sufficient integrability conditions on the stock price process we can now combine (2) and (A9) and find that

$$D_h z_T(k)[\nu] = \frac{1}{2\pi}\int_{-\infty}^{\infty} dv\, \mathrm{e}^{-ivk-rT}\,\frac{T\mathrm{e}^{T\psi(v-i)}}{iv(1+iv)}\int_{-\infty}^{\infty} dx\, h(x)\left\{\mathrm{e}^{ivx} - 1 - iv\mathrm{e}^{x} + iv\right\}$$

By interchanging the two integrals we can compute, again under sufficient integrability conditions, the Fréchet derivative, Dz_T, of the time value:

$$Dz_T(k)[\nu] = \frac{1}{2\pi}\int_{-\infty}^{\infty} dv\, \mathrm{e}^{-ivk-rT}\,\frac{T\mathrm{e}^{T\psi(v-i)}}{iv(1+iv)}\left\{\mathrm{e}^{ivx} - 1 - iv\mathrm{e}^{x} + iv\right\} \tag{40}$$

By rearranging terms and separating integrals we have

$$Dz_T(k)[\nu] = \frac{T}{2\pi}\int_{-\infty}^{\infty} dv\, \mathrm{e}^{-iv(k+x)}\,\frac{\mathrm{e}^{-rT}\mathrm{e}^{T\psi(v-i)} - \mathrm{e}^{ivrT}}{iv(1+iv)} -$$

$$
\begin{aligned}
&- \frac{T}{2\pi} \int_{-\infty}^{\infty} dv e^{-ivk} \frac{e^{-rT} \exp(T\psi(v-i)) - e^{ivrT}}{iv(1+iv)} \\
&+ \frac{T}{2\pi} \int_{-\infty}^{\infty} dv e^{-ivk} \frac{e^{-ivx+ivrT} - e^{ivrT}}{iv(1+iv)} \\
&- \frac{T(e^x-1)}{2\pi} \int_{-\infty}^{\infty} dv e^{-ivk-rT} \frac{e^{T\psi(v-i)}}{1+iv}
\end{aligned} \tag{41}
$$

Here the first two terms may be expressed in terms of the option price function, the third term does not depend on the Lévy measure and can be computed analytically, and the last is a product of a simple function of x and some auxiliary function that does not depend on x (and therefore has to be computed only once for each gradient evaluation). Finally, we obtain

$$
\begin{aligned}
Dz_T(k)[\nu] &= Tz_T(k+x) - Tz_T(k) + T(1-e^{k+x-rT})^+ - T(1-e^{k-rT})^+ \\
&\quad - \frac{T(e^x-1)}{2\pi} \int_{-\infty}^{\infty} dv e^{-ivk-rT} \frac{\exp(T\psi(v-i))}{1+iv} \\
&= T\big(C_T(k+x) - C_T(k)\big) - (e^x-1) \frac{T}{2\pi} \int_{-\infty}^{\infty} dv e^{-ivk-rT} \frac{\exp(T\psi(v-i))}{1+iv}
\end{aligned} \tag{42}
$$

Fortunately, this expression may be represented in terms of the option price and one auxiliary function. Since we are using a fast Fourier transform (FFT) to compute option prices for the whole price sheet, we already know these prices for the whole range of strikes. As the auxiliary function will also be computed using the FFT algorithm, the computational time will only increase by a factor of two.

4.5 The algorithm

Here is the final numerical algorithm as implemented in the examples below.

1. Calibrate an auxiliary jump–diffusion model (Merton model) to obtain an estimate of volatility, σ_0, and a candidate for the prior Lévy measure, ν_0.
2. Estimate the “noise level” ϵ_0 from option prices as explained in Section 4.3:

$$
\epsilon_0^2 = \inf_{\nu} \sum_{i=1}^{N} \omega_i \left| C_i^{\sigma_0,\nu} - C_i^* \right|^2 \tag{43}
$$

3. Use a *posteriori* method described in Section 4.3 to compute an optimal

regularization parameter, α^*, that achieves a trade-off between precision and stability:

$$\epsilon^2(\alpha^*) = \sum_{i=1}^{N} \omega_i |C_i^{\sigma,\nu} - C_i^*|^2 \simeq \delta\epsilon_0^2 \qquad (44)$$

with delta slightly greater than 1. The optimal α^* is found by running the gradient descent method (BFGS) several times with low precision.

4. Minimize $\mathcal{J}(\nu)$ with α^* by gradient-based method (BFGS) with high precision using either a user-specified prior or result of 1) as prior.

5 Numerical tests

To verify the accuracy and numerical stability of our algorithm, we first tested it on simulated data sets of option prices generated using a known jump–diffusion model. This allowed us to explore the magnitude of finite sample effects and to assess the importance of the two different stages of the calibration procedure described in Section 4.

5.1 A compound Poisson example: the Kou model

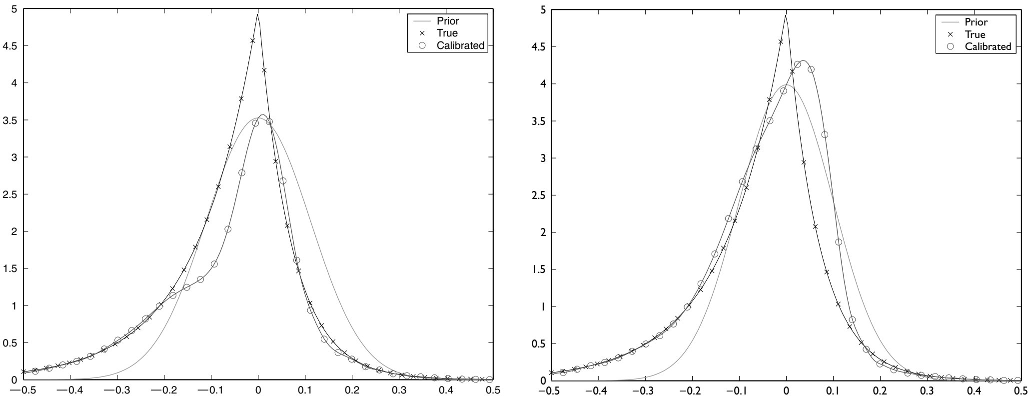
In the first series of tests, option prices were generated using Kou's (2002) jump–diffusion model with a diffusion element $\sigma_0 = 10\%$ and a Lévy density given by

$$\nu(x) = \lambda[1_{x>0}p\alpha_1 e^{-\alpha_1 x} + (1-p)\alpha_2 e^{-\alpha_2|x|}1_{x<0}] \qquad (45)$$

In the tests we took an asymmetric density, with the left tail heavier than the right ($\alpha_1 = 1/0.07$ and $\alpha_2 = 1/0.13$). The intensity was taken to be $\lambda = 1$, and the last constant, p, was chosen such that the density was continuous at $x = 0$. The option prices were computed using the Fourier transform method described in the appendix. The maturity of the options was five 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 was used as the prior.

After generating sets of call option prices from Kou's model, the algorithm described in Section 4 was applied to these prices. Figure 6 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 (45). As seen in Figure 7, the accuracy of calibration at the level of option prices and/or implied volatilities is satisfying with only 21 options. Comparing the jump size densities obtained with the true value, we observe that we have successfully retrieved 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. In fact, the gradient of our calibration criterion (computed in

FIGURE 6 Lévy measure calibrated to option prices simulated from Kou's (2002) jump–diffusion model with σ_0 = 10%. *Left*: σ calibrated in a separate step (σ = 10.5%). *Right*: σ fixed at 9.5% < σ_0.



Section 4.4) vanishes at zero, which means that the algorithm does not modify the Lévy density in this region: the intensity of small jumps cannot be retrieved accurately. The redundancy of small jumps and the diffusion component is well known in the context of statistical estimation on time series (Beckers, 1981;

FIGURE 7 Calibrated vs. simulated (true) implied volatilities corresponding to Figure 6 for the Kou (2002) model.

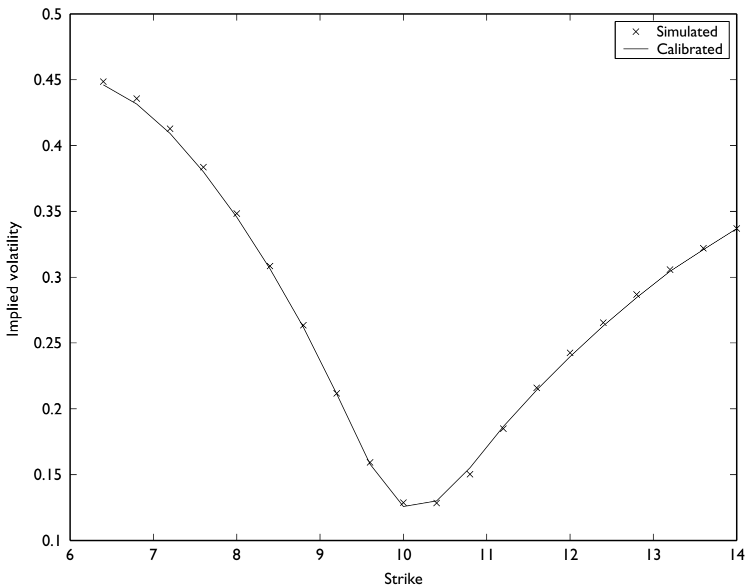
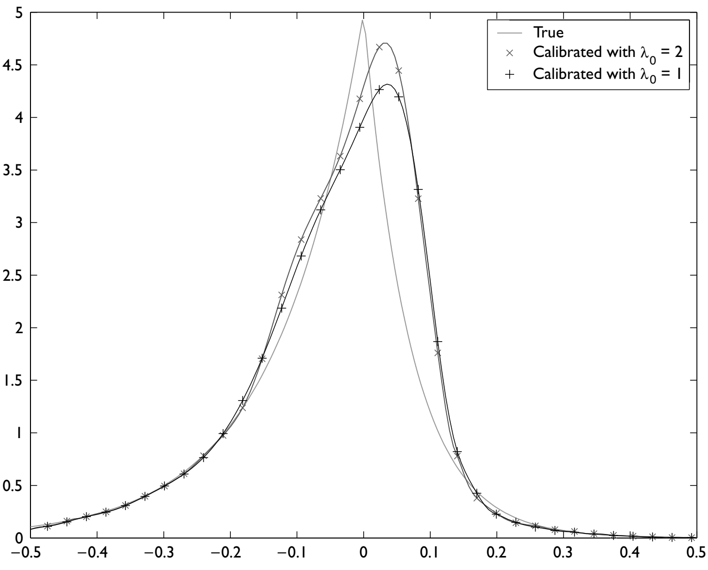


FIGURE 8 Lévy densities calibrated to option prices generated from the Kou model using two different initial measures with intensities λ = 1 and λ = 2.



Mancini, 2001). Here we retrieve another version of this redundancy in a context of calibration to a cross-sectional data set of options.

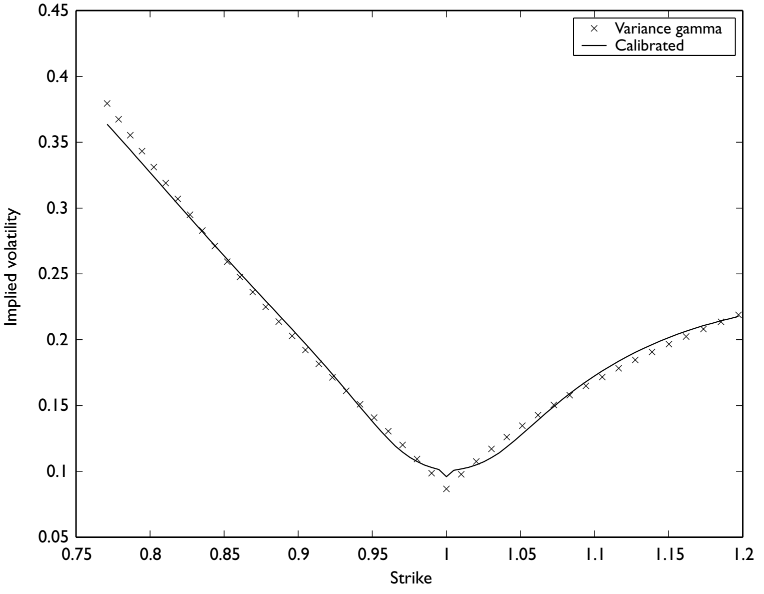
Comparison of the left and right graphs in Figure 6 further illustrates the redundancy of small jumps and diffusion: the two graphs were calibrated to the same prices and only differ in the diffusion coefficients. Comparing the two graphs shows that the algorithm compensates for the error in the diffusion coefficient by adding jumps to the Lévy density such that, overall, the accuracy of calibration is maintained (the standard deviation is 0.2%).

The stability of the algorithm with respect to initial conditions can be examined by perturbing the starting point of the optimization routine and examining the effect on the output. As illustrated in Figure 8, the entropy penalty removes the sensitivity observed in the non-linear least-squares algorithm (see Figure 3 and Section 3.1). The only minor difference between the two calibrated measures is observed in the neighborhood of zero, ie, the region which, as remarked above, has little influence on option prices.

5.2 Variance gamma model

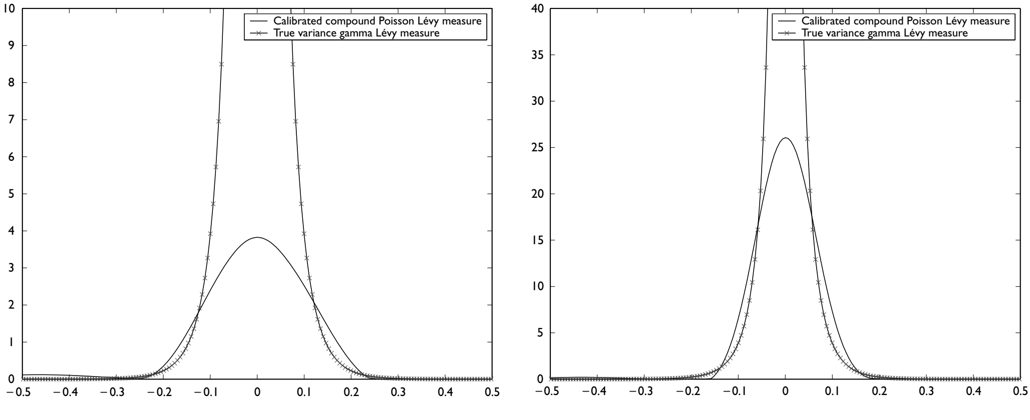
In a second series of tests we examined how our method performs when applied to option prices generated by an infinite activity process such as the variance gamma model. We assume that the user, ignoring the fact that the data-generating process has infinite activity, chooses a (misspecified) prior which has a finite jump intensity (eg, the Merton model).

FIGURE 9 Implied volatility smile for variance gamma model with $\sigma_0 = 0$ compared with smile generated from the calibrated Lévy measure. Calibration yields $\sigma = 7.5\%$.



Option prices for 30 strike values were generated using the variance gamma model (Madan, Carr and Chang, 1998) with no diffusion component ($\sigma_0 = 0$) and the calibration algorithm was applied using as prior a Merton jump–diffusion model. Figure 9 shows that even though the prior is misspecified, the result repro-

FIGURE 10 Lévy measure calibrated to variance gamma option prices with $\sigma = 0$ using a compound Poisson prior with $\sigma = 10\%$ (*left*) and $\sigma = 7.5\%$ (*right*). Increasing the diffusion coefficient reduces the intensity of small jumps in the calibrated measure.



duces the implied volatilities with good precision (the error is less than 0.5% in implied volatility units). The calibrated value of the diffusion coefficient of σ is 7.5%, while near zero the Lévy density has been truncated to a finite value (Figure 10, left): the algorithm has substituted a non-zero diffusion part for the small jumps which are the origin of infinite activity. Figure 10 further compares the Lévy measures obtained when σ is fixed to two different values: we observe that 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 activity context. More precisely, this example shows that call option prices generated from an infinite activity exponential Lévy model can be reproduced with arbitrary precision using a compound Poisson model with finite jump intensity. This leads us to conclude that, for a finite (but realistic) number of observations, infinite activity models such as variance gamma are hard to distinguish from finite activity compound Poisson models on the basis of option prices.

6 Empirical results

To illustrate our calibration method we have applied it to a data set consisting of a daily series of prices and implied volatilities for options on the DAX (German index) from 1999 to 2001. A detailed description of data formats and filtering procedures can be found in Cont and da Fonseca (2002). Some of the results obtained with this data set are described below.

6.1 Empirical properties of implied Lévy densities

Figure 11 illustrates the typical shape of a risk-neutral Lévy density obtained from our data set: it is peaked at zero and highly skewed towards negative values. The effect of including the entropy penalty can be assessed by comparing the results obtained when changing the initialization in the algorithm. Figure 12 compares the Lévy measures obtained with different initializers: in this case the jump intensity of the initial Lévy measure (a Merton model) was shifted from $\lambda = 1$ to $\lambda = 5$. Compared to the high sensitivity observed in the non-linear least-squares algorithm (Figure 3), we observe that adding the entropic penalty term has stabilized our algorithm.

The logarithmic scale in Figure 12 allows the tails to be seen more clearly. Recall that the prior density is gaussian, which shows up as a symmetric parabola on log scales. By contrast, it is readily observed that the Lévy measures obtained are far from symmetric: the distribution of jump sizes is highly skewed towards negative values. Figure 16 shows the same result across calendar time, showing that this asymmetry persists across time. This effect also depends on the maturity of the options in question: for longer maturities (see Figure 17) the support of the Lévy measure extends further to the left.

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

FIGURE 11 Lévy density implied from DAX option prices with maturity of three months.

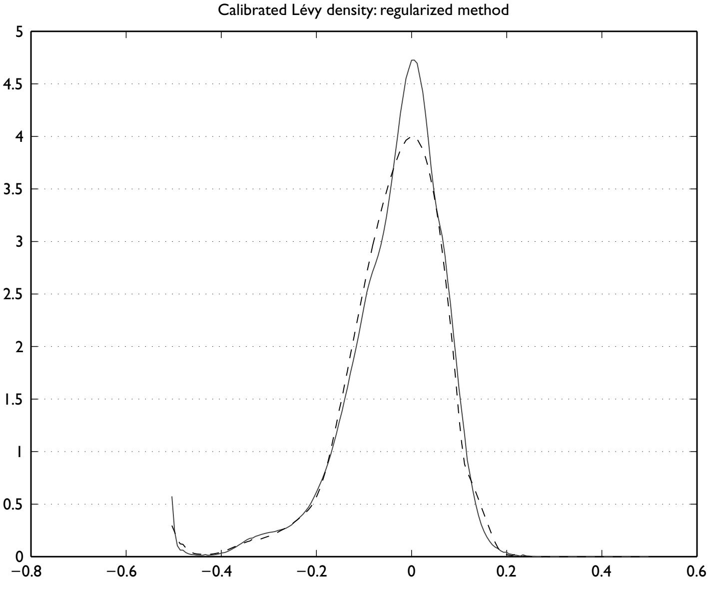


FIGURE 12 Lévy density implied by DAX option prices with maturity of three months.

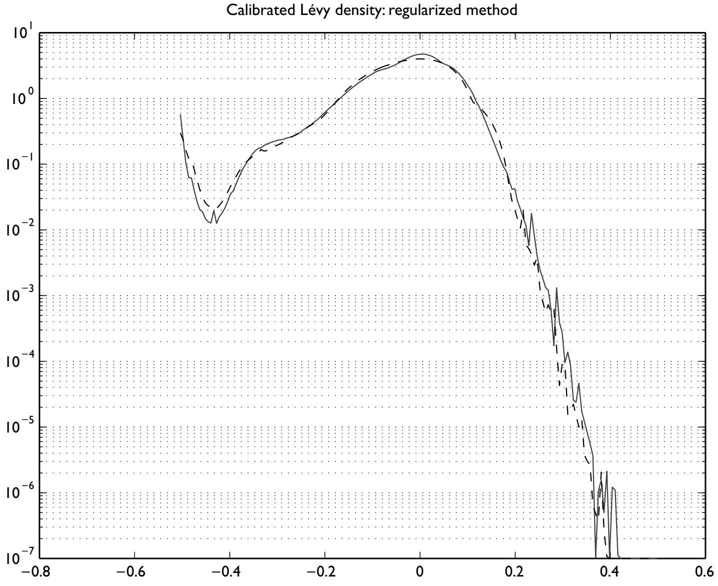
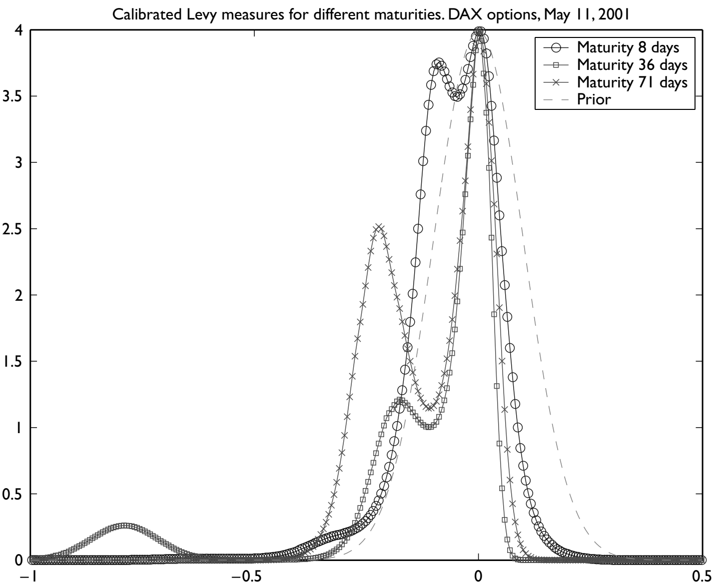


FIGURE 13 Lévy measures implied from DAX options on the same calendar date for three different maturities.



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 to explain the shape of the implied volatility skew for small maturities. Therefore, the motivation of infinite activity processes does not seem clear to us, at least from the viewpoint of option pricing.

6.2 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 several empirical studies have shown that Lévy processes reproduce the implied volatility smile for a single maturity quite well (Carr *et al.*, 2002; Madan, Carr and Chang, 1998), when it comes to calibrating several maturities at the same time the calibration by Lévy processes becomes much less precise. The reason is that, due to the stationary nature of their increments, Lévy processes have a 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. 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.

FIGURE 14 Calibration quality for different maturities: market-implied volatilities for DAX options against model implied volatilities. Each maturity has been calibrated separately.

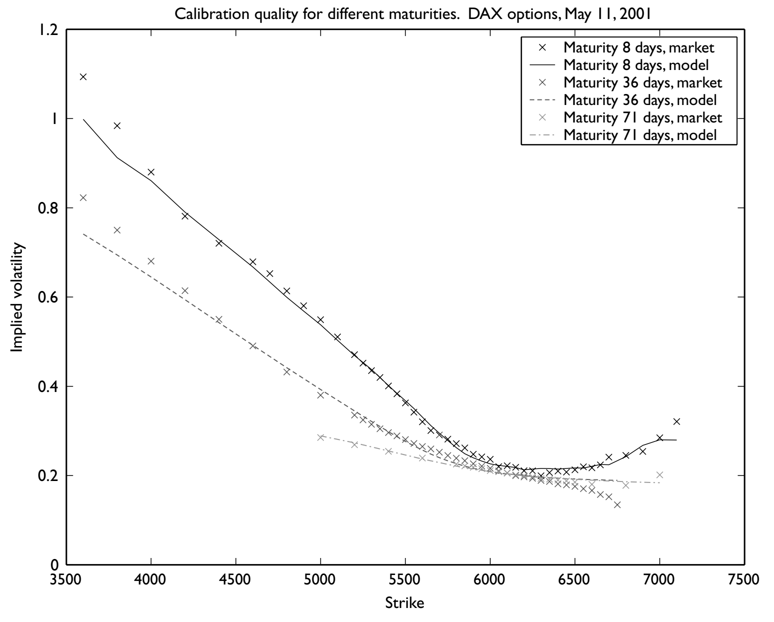


FIGURE 15 Lévy measures implied from DAX options (logarithmic scale).

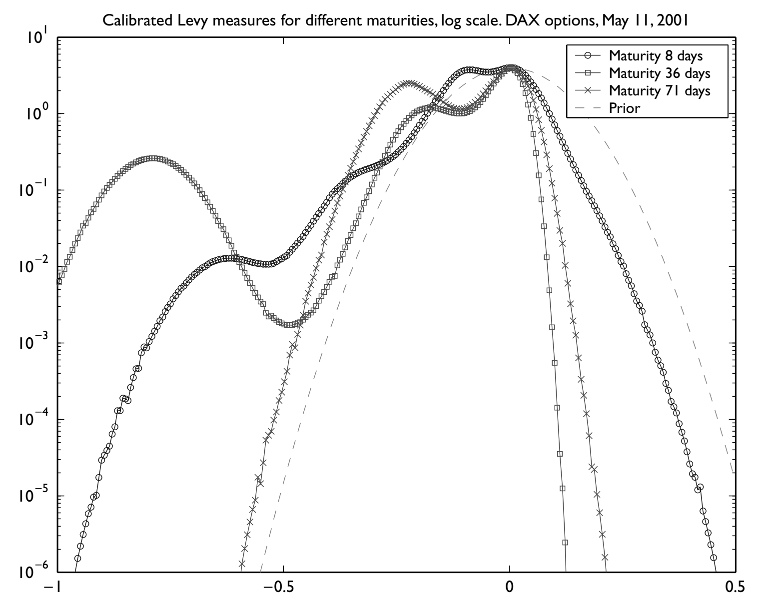


FIGURE 16 Results of calibration at different dates for shortest-maturity DAX index options.

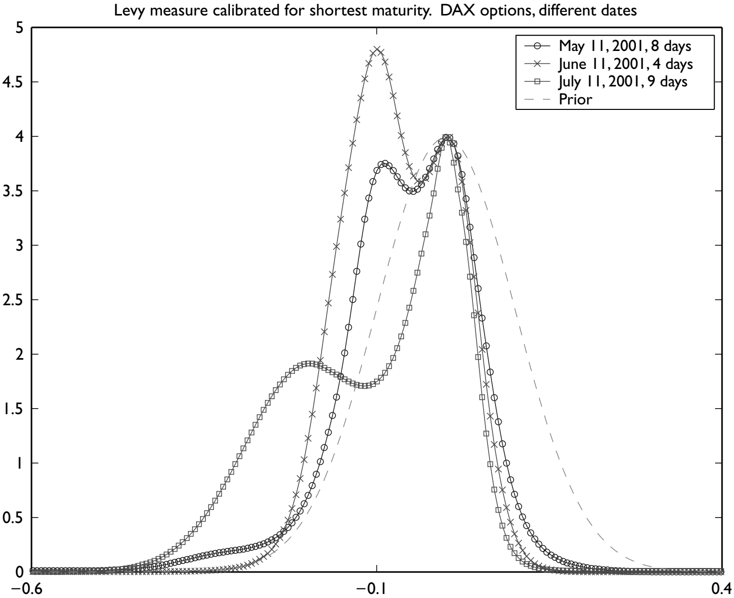
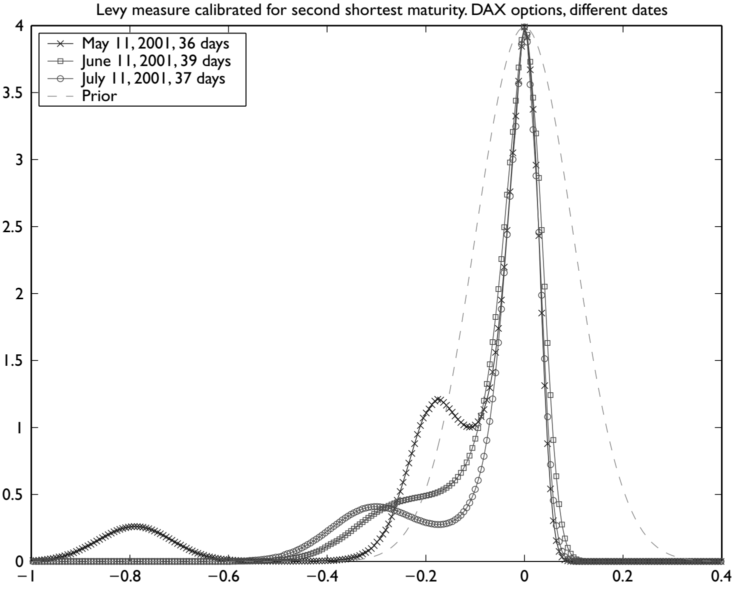


FIGURE 17 Results of calibration at different dates for second shortest maturity DAX index options.



Based on an empirical study of the kurtosis implied by US dollar/Deutschmark exchange rate options, Bates (1996b) concluded 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]". In the field of stock index options, Madan and Konikov (2002) have reported even more surprising results: both implied skewness and kurtosis actually decrease as the length of the holding period becomes smaller. Whereas these analyses rely on implied moments and cumulants, here we can investigate the evolution of the entire implied Lévy measure with maturity in a non-parametric way by separately calibrating the Lévy measure to various option maturities. Figure 13 shows Lévy measures obtained by running the algorithm separately for options of different maturity. The null 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. Of course, this is a more complicated form of time-dependence than simply having a time-dependent intensity.

These results can be further used to investigate what form of time-dependence is appropriate to introduce so as to capture the empirically observed term structure of implied volatilities. This line of thought naturally leads to the domain of additive processes, that is, processes with independent but not necessarily stationary increments (Cont and Tankov, 2004a).

7 Conclusion

We have proposed a non-parametric method for identifying, in a numerically stable fashion, a risk-neutral exponential Lévy model consistent with market prices of options and equivalent to a prior model. We have also presented a stable computational implementation and tested its performance on simulated and empirical data. Theoretically our method can be seen as a computable approximation to the notions of minimal-entropy martingale measures made consistent with observed market prices of options. Computationally, it is a stable alternative to current least-squares calibration methods for exponential Lévy models. The jump part is retrieved in a non-parametric fashion: we do not assume shape restrictions on the Lévy measure. Our approach allows us to reconcile the idea of calibration by relative entropy minimization (Avellaneda, 1998) with the notion of risk-neutral valuation, enabling us to use relative entropy as a model selection criterion without preliminary discretization of the state space.

Our method can complement the existing literature on parametric exponential Lévy models in option pricing in various ways. For many parametric models the relative entropy can be explicitly computed (see example 1 in Section 2.4). Since in parametric models the Lévy measure is typically a non-linear function of model parameters, the relative entropy will no longer be convex but may still be a well-behaved function with one global minimum, locally convex around this minimum. Therefore, relative entropy calibration of parametric models is still possible, but in this setting one should also consider other regularizing functions. On the other hand, using a non-parametric calibration is not necessarily incompatible with using a parametric model for pricing. Our method can be used as a specification test for choosing the correct parametric class of exponential Lévy models. Our computational approach for estimating risk-neutral jump processes from options data can potentially be applied to other models where jump processes have to be deduced from observation of contingent claims: credit risk models are typically such examples. Separate calibration of the jump density to various option maturities can be used to investigate time-inhomogeneity in a non-parametric way. Finally, our approach can be extended to other inverse problems in which an unknown jump process has to be identified, such as calibration problems for stochastic volatility models with jumps (Barndorff-Nielsen, Mikosch and Resnick, 2001; Bates, 1996a). We intend to pursue these issues in future research.

Appendix A Option pricing by Fourier transform

We recall here the expression, due to Carr and Madan (1998) of option prices in terms of the characteristic function of the Lévy process. Due to the special structure of the characteristic function in these models, it is convenient to express option prices in terms of the characteristic function. In particular, for a European call option with log strike k,

$$C_T(k) = \mathrm{e}^{-rT} E^Q[(\mathrm{e}^{s_T} - \mathrm{e}^k)^+] \tag{A1}$$

where s_T is the terminal log price with density $q_T(s)$. The characteristic function of this density is defined by

$$\phi_T(u) \equiv \int_{-\infty}^{\infty} \mathrm{e}^{ius} q_T(s)\,\mathrm{d}s \tag{A2}$$

On the other hand, as remarked above, the characteristic function of the log price is given by the Lévy–Khinchin formula (here we limit ourselves to the compound Poisson case):

$$\phi_T(u) = \exp\left\{T\left(-\frac{1}{2}\sigma^2 u^2 + i\gamma(\nu)u + \int_{-\infty}^{\infty}(\mathrm{e}^{iux} - 1)\nu(x)\,\mathrm{d}x\right)\right\} \tag{A3}$$

$$\gamma(\nu) = r - \frac{\sigma^2}{2} - \int_{-\infty}^{\infty}(\mathrm{e}^x - 1)\nu(x)\,\mathrm{d}x \tag{A4}$$

In some important cases this characteristic function is known analytically; otherwise one can discretize the Lévy measure and use (in the compound Poisson case) the fast Fourier transform to compute the characteristic function.

Following Carr and Madan (1998), we use Fourier transform methods to evaluate the expression (A1) for a given Lévy measure. To do so we observe that although the call price as a function of log strike is not square integrable, the time value of the option, that is, the function

$$z_T(k) = E[(\mathrm{e}^{s_T} - \mathrm{e}^k)^+] - (1 - \mathrm{e}^{k-rT})^+ \tag{A5}$$

equal to the price of the option (call or put), which is for given k out-of-the-money (forward), may be square integrable. Here we have assumed without loss of generality that $s_0 = 0$. Let $\zeta_T(v)$ denote the Fourier transform of the time value:

$$\zeta_T(v) = \int_{-\infty}^{\infty} \mathrm{e}^{ivk} z_T(k)\,\mathrm{d}k \tag{A6}$$

It can be expressed in terms of the characteristic function of the log-price in the following way. First, we note that since the discounted price process is a martingale, we can write

$$z_T(k) = \mathrm{e}^{-rT} \int_{-\infty}^{\infty} q_T(s)\,\mathrm{d}s(\mathrm{e}^s - \mathrm{e}^k)\big(1_{k\le s} - 1_{k\le rT}\big)$$

Next, we compute $\zeta(v)$ by interchanging integrals:

$$\zeta_T(v) = \mathrm{e}^{-rT} \int_{-\infty}^{\infty} \mathrm{d}k \int_{-\infty}^{\infty} \mathrm{d}s\,\mathrm{e}^{ivk} q_T(s)(\mathrm{e}^s - \mathrm{e}^k)\big(1_{k\le s} - 1_{k\le rT}\big)$$

$$= \mathrm{e}^{-rT} \int_{-\infty}^{\infty} q_t(s)\,\mathrm{d}s \int_{s}^{rT} \mathrm{e}^{ivk}(\mathrm{e}^k - \mathrm{e}^s)\,\mathrm{d}k$$

A sufficient condition allowing us to justify the interchange of integrals is that the stock prices have a moment of order $1 + \alpha$ for some positive alpha or

$$\exists \alpha > 0: \quad \int_{-\infty}^{\infty} q_T(s)\,\mathrm{e}^{(1+\alpha)s}\,\mathrm{d}s < \infty \tag{A7}$$

We can write for the inner integral

$$\int_{s}^{rT} |\mathrm{e}^k - \mathrm{e}^s|\,\mathrm{d}k \le \mathrm{e}^{rT} - \mathrm{e}^s, \qquad \text{if } rT \ge s \tag{A8}$$

and

$$\int_{rT}^{s} |e^k - e^s| \, dk \le e^s (s - rT) 1_{s > rT}, \qquad \text{if } rT < s$$

We see that under the condition (A7) both expressions, when multiplied by $q_T(s)$, are integrable with respect to s and we can apply Fubini's theorem to justify the interchange. The inner integral is computed in a straightforward fashion, and after computing the outer integral for some terms and re-expressing it in terms of the characteristic function of the log stock price, we obtain

$$\zeta_T(v) = \frac{e^{-rT} \phi_T(v - i) - e^{ivrT}}{iv(1 + iv)} \tag{A9}$$

The martingale condition guarantees that the numerator is equal to zero for $v = 0$. Under the condition (A7), we see that the numerator becomes an analytic function and the fraction has a finite limit for $v \to 0$. The option prices can now be found by inverting the Fourier transform:

$$z_T(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ivk} \zeta_T(v) \, dv \tag{A10}$$

Appendix B Existence of a solution for the discretized problem

In this section we present a proof of Proposition 3. First, we will establish the continuity of J on $\mathcal{L}_\pi$. Then we will establish a lower bound for $H(\nu)$ which will enable us to show that the minimum of $J(\nu)$ is reached on a bounded subset of $\mathcal{L}_\pi$. Finally, we will show that for α large enough this minimum is unique.

B1 Continuity of the relative entropy function in the discretized case

Consider the function (19) for a Lévy measure which belongs to the class (30). In this case, all integrals become finite sums and the relative entropy takes the following form:

$$H(\nu) = \frac{T}{2\sigma^2} \left\{ \sum_{i=1}^{N} (e^{x_i} - 1)\left(\nu_i - \nu^P(x_i)\right) \right\}^2 + T \sum_{i=1}^{N} \left\{ \nu_i \ln \frac{\nu_i}{\nu^P(x_i)} + \nu^P(x_i) - \nu_i \right\} \tag{B1}$$

where we denote $\nu_i \equiv \nu(x_i)$ The first term is continuous because it is a continuous function (square) of a finite sum of continuous (linear) functions of ν_i. To treat the second term, consider the one-dimensional function $f(x) = x \ln(x/x_0) + x_0 - x$. It is continuous for $x \ge 0$ if we take by definition $f(0) = x_0$ (in fact, it is even uniformly continuous on this set). Thus, the second term is continuous for $\nu_i \ge 0$, $\forall i$.

B2 Continuity of the pricing error

To establish the continuity of the pricing error function, it is sufficient to prove that one option price is a continuous function of the Lévy measure or, equivalently, to prove that the time value $z_T(k, \nu)$ defined by equation (A5) is continuous for all k. Let ν^1 and ν^2 be two Lévy measures in $\mathcal{L}_\pi$ and take $\varepsilon > 0$. We will prove that there exists $\delta > 0$ such that if $\max_i |\nu_i^1 - \nu_i^2| < \delta$, then $|z_T(k, \nu^1) - z_T(k, \nu^2)| < \varepsilon$. From equations (A10) and (A9),

$$|z_T(k,\nu^1) - z_T(k,\nu^2)| \le \int_{-\infty}^{\infty} |\zeta_T(u,\nu^1) - \zeta_T(u,\nu^2)|\, dv$$

$$\le e^{-rT} \int_{-\infty}^{\infty} \frac{|\phi^{\nu^1}(v-i) - \phi^{\nu^1}(v-i)|}{|v|\sqrt{1+v^2}}\, dv$$

for all k. Equations (A3) and (A4) entail that $|\phi^\nu(v-i)| \le e^{rT}$ for all ν and for all $v \in \mathbb{R}$. Hence, one can find A such that

$$e^{-rT} \int_{|v|\ge A} \frac{|\phi^{\nu^1}(v-i) - \phi^{\nu^1}(v-i)|}{|v|\sqrt{1+v^2}}\, dv \le \frac{\epsilon}{2}$$

Recalling that $\phi^\nu(u) = \exp T\psi^\nu(v)$ and using once again the upper bound for the modulus of ϕ along with the mean value theorem, we find

$$e^{-rT} \int_{|v|\le A} \frac{|\phi^{\nu^1}(v-i) - \phi^{\nu^1}(v-i)|}{|v|\sqrt{1+v^2}}\, dv \le \int_{|v|\le A} \frac{|\phi^{\nu^1}(v-i) - \phi^{\nu^1}(v-i)|}{|v|\sqrt{1+v^2}}\, dv$$

$$\le \sum_{j=1}^{N} |e^{x_j} - 1||\nu_j^1 - \nu_j^2| \int_{v\le A} \frac{dv}{\sqrt{1+v^2}} + \sum_{j=1}^{N} e^{x_j} |\nu_j^1 - \nu_j^2| \int_{|v|\le A} \frac{|e^{ivx_k} - 1|}{|v|\sqrt{1+v^2}}$$

Since all the integrals in the last term are integrals of bounded functions over bounded domains, one can choose $\delta > 0$ such that if $\max_i |\nu_i^1 - \nu_i^2| < \delta$, then the last term is smaller than $\varepsilon/2$ and hence $|z_T(k, \nu^1) - z_T(k, \nu^2)|$ is smaller than ε. This proves the continuity of the pricing function uniformly on k.

B3 Lower bound for regularized function

Using the following trivial inequality

$$x \ln \frac{x}{x_0} + x_0 - x \ge x - x_0(e-1)$$

we obtain the following bound for the function $J(\nu)$:

$$J(\nu) \geq \alpha H(\nu) \geq \alpha T \sum_{i=1}^{N} \nu_i - \alpha T(\mathrm{e}-1) \sum_{i=1}^{N} \nu_i^P \tag{B2}$$

The regularized function is thus bounded below by the L^1 norm of ν minus some constant.

B4 Existence of solution for the regularized problem

To prove that the regularized problem has a solution, consider a compact $C \subset \mathcal{L}_\pi$ defined by

$$C = \left\{ \nu : \alpha T \sum_{i=1}^{N} \nu_i \leq \alpha T(\mathrm{e}-1) \sum_{i=1}^{N} \nu_i^P + J(\nu^P) + 1 \right\} \tag{B3}$$

Since J is continuous,

$$\exists \nu^* \in L_\pi, J(\nu^*) = \inf_{\nu \in C} J(\nu)$$

However, using the bound (B2) for all $\nu \in \mathrm{L}_\pi \backslash C$, we find $J(\nu) \geq J(\nu^P) + 1 > J(\nu^*)$. Hence, ν^* is the solution of the regularized problem.

B5 Uniqueness of the solution for large α

Making the additional hypothesis that $\sigma > 0$, we will now show that for any compact K there exists an α_0 such that for any $\alpha \geq \alpha_0$, $J(\nu)$ is convex on K. Since the size of compact C in (B3) decreases when α grows, this entails uniqueness of the solution for $\alpha > \alpha_0$.

Consider the regularized function $J(\nu) = \alpha H(\nu) + \epsilon(\nu)$, where $\epsilon(\nu)$ denotes the sum of squared pricing errors. From equation (B1),

$$H^{jk} \equiv \frac{\partial^2 H}{\partial \nu_j \, \partial \nu_k} = \frac{T}{a}(\mathrm{e}^{x_j} - 1)(\mathrm{e}^{x_k} - 1) + \frac{T}{\nu_j} \delta_{jk}$$

Suppose that we can prove that the second derivative of $\epsilon(\nu)$ can be represented as follows:

$$\epsilon^{jk} \equiv \frac{\partial^2 \epsilon}{\partial \nu_j \, \partial \nu_k} = P^{jk} + B^{jk} \tag{B4}$$

where P is a positive definite matrix and the elements of B are bounded by some constant c. Let $K = \{\nu : \nu_i \leq \nu_0\}$ for some $\nu_0 > 0$. Then for every $\alpha > Nc\nu_0/T$, J is convex on K. Indeed, for every vector $\Delta\nu$ and for every $\nu \in K$,

$$\sum_{jk} \frac{\partial^2 J(\nu)}{\partial\nu_j\,\partial\nu_k}\Delta\nu_j\,\Delta\nu_k = \frac{\alpha T}{a}\left(\sum_j (e^{x_j}-1)\Delta\nu_j\right)^2 + \alpha T\sum_j \frac{\Delta\nu_j^2}{\nu_j} + \sum_{jk} P^{jk}\Delta\nu_j\,\Delta\nu_k + \sum_{jk} B^{jk}\Delta\nu_j\,\Delta\nu_k$$

$$\geq \alpha T\sum_j \frac{\Delta\nu_j^2}{\nu_j} + \sum_{jk} B^{jk}\Delta\nu_j\,\Delta\nu_k \geq \alpha T\sum_j \frac{\Delta\nu_j^2}{\nu_0} + Nc\sum_j \Delta\nu_j^2 \geq 0$$

It remains therefore to prove the representation (B4). Without loss of generality we can consider that there is only one option; hence, $\epsilon(\nu) = (C(\nu) - C_0)^2$. Furthermore,

$$\frac{\partial^2 \epsilon(\nu)}{\partial\nu_j\,\partial\nu_k} = 2\frac{\partial C(\nu)}{\partial\nu_j}\frac{\partial C(\nu)}{\partial\nu_k} + 2\big(C(\nu) - C_0\big)\frac{\partial^2 C(\nu)}{\partial\nu_j\,\partial\nu_k}$$

The first term in this expression is a positive definite matrix and we must prove the boundedness of the right term. Since the price of an option is always bounded by the current stock price, it is sufficient to check the boundedness of $\partial^2 C(\nu)/\partial\nu_j\partial\nu_k$ or, equivalently, of $\partial^2 z(\nu)/\partial\nu_j\partial\nu_k$ (here we omit the variable k, which is now irrelevant). From formula (A10),

$$\frac{\partial^2 z(\nu)}{\partial\nu_j\,\partial\nu_k} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ivk}\,\frac{\partial^2 \zeta_v(\nu)}{\partial\nu_j\,\partial\nu_k}\,dv$$

Thus, for $z(\nu)$ to be bounded $\zeta_v(\nu)$ must be integrable as a function of v uniformly on ν. By formula (A9) this is equivalent to uniform integrability of

$$\left|\frac{1}{v(1+iv)}\frac{\partial^2 \phi(v-i)}{\partial\nu_j\,\partial\nu_k}\right| \tag{B5}$$

By the Lévy–Khinchin formula,

$$\phi_T(v-i) = e^{T\psi(v-i)} \tag{B6}$$

where

$$\psi(v-i) = -\frac{1}{2}\sigma^2 v^2 + \frac{1}{2}iv\sigma^2 + ivr + r - iv\sum_j (e^{x_j}-1)\nu_j + \sum_j e^{x_j}(e^{ivx_j}-1)\nu_j$$

Hence

$$\frac{\partial^2 \phi(v-i)}{\partial \nu_j \partial \nu_k} = T^2 \phi(v-i) \frac{\partial \psi(v-i)}{\partial \nu_j} \frac{\partial \psi(v-i)}{\partial \nu_k}$$

Since

$$\frac{\partial \psi(v-i)}{\partial \nu_j} = -iv(e^{x_j}-1) + e^{x_j}(e^{ivx_j}-1)$$

both $\partial\psi(v-i)/\partial\nu_j$ and $(1/v)\left(\partial\psi(v-i)/\partial\nu_j\right)$ are bounded above by polynomials in v, independent on ν. Formula (B6) entails that $|\phi_T(v-i)| \leq e^{r-1/2\sigma^2 v^2}$. Therefore, the integral of (B5) is bounded by a number which does not depend on ν and the proof is completed.

Appendix C Properties of regularized solutions

We present here some properties of the solutions of our regularized problem in the discretized case (ie, the Lévy measure is concentrated on a discrete grid). This case is of most interest from the point of view of numerical implementation. We shall denote by H the relative entropy function defined in (21): $H(\nu) = \mathcal{E}(\mathbb{Q}(\sigma|\nu), \mathbb{Q}(\sigma|\nu_0))$. Define $\delta > 0$ as the observational error on the data C^*: $\|C^* - C\| \leq \delta$, where C^* is the vector of observed option prices and C is a vector of arbitrage-free ("true") prices.

The solution of (29) is in general not unique due to non-convexity of the pricing function. It depends continuously on the data in the following sense.

PROPOSITION 4 *Let $\alpha > 0$ and let $\{C^k\}$ and $\{\nu^k\}$ be sequences where $C^k \to C^*$ and ν^k is the solution of problem* (29) *with C^* replaced by C^k. Then there exists a convergent subsequence of* $\{\nu^k\}$ *and the limit of every convergent subsequence is a solution of (29).*

REMARK If the solution of (29) is unique, this is just the definition of continuity.

PROOF To simplify the notation we write $F(\nu)$ for a set of model prices and $\|F(\nu) - C^*\|^2$ for the sum of squared differences of model prices corresponding to Lévy measure ν and market prices C^*. Let $\{C_k\}$ be a sequence of data sets converging to C^* and $\{\nu_k\}$ be the corresponding sequence of solutions:

$$\nu_k = \arg\inf\left\{\|F(\nu_k) - C_k\|^2 + \alpha H(\nu)\right\}$$

By construction we have

$$\|F(\nu_k) - C_k\|^2 + \alpha H(\nu) \leq \|F(\nu) - C_k\|^2 + \alpha H(\nu), \qquad \forall \nu \in \mathcal{L}(\mathbb{R}) \qquad \text{(C1)}$$

Hence the sequences $\|\nu_k\|$ and $\|F(\nu_k)\|$ are bounded. Since we work in a finite-dimensional space, we can find a convergent subsequence $\nu_m \to \nu^*$ of $\{\nu_k\}$. Now let $\{\nu_m\}$ be any convergent subsequence. Using the continuity of the pricing function, we have $F(\nu_m) \to F(\nu^*)$. This, together with (C1) and the continuity of the relative entropy function implies

$$\forall \nu \in \mathcal{L}(\mathbb{R}), \quad \|F(\nu^*) - C^*\|^2 + \alpha H(\nu) = \lim\left\{\|F(\nu_m) - C_m\|^2\right\} + \alpha H(\nu)$$

$$\leq \lim\left\{\|F(\nu) - C_m\|^2\right\} + \alpha H(\nu) = \|F(\nu^*) - C^*\|^2 + \alpha H(\nu)$$

Hence, we have proven that ν^* is a minimizer of $\|F(\nu) - C^*\|^2 + \alpha H(\nu)$. □

Let M be the set of discretized Lévy measures ν which solve the least-squares calibration problem (27). Assume that

$$\exists \nu \in M, \ \mathcal{E}(\mathbb{Q}(\sigma,\nu)\,|\,\mathbb{Q}_0) < \infty \tag{C2}$$

Then a minimum-entropy least-squares solution is defined as a solution of

$$\inf_{\nu \in M} \mathcal{E}\big(\mathbb{Q}(\sigma,\nu)\,|\,\mathbb{Q}_0\big) \tag{C3}$$

The next proposition describes how the solutions of (29) converge towards minimum-entropy least-squares solutions as the error level δ decreases.

Proposition 5 *Suppose that the calibration problem with data C^* admits a minimum-entropy least-squares solution C^*. Let*

$$\|C^* - C\| \leq \delta$$

and let $\alpha(\delta)$ be such that $\alpha(\delta) \to 0$ and $\delta/\alpha(\delta) \to 0$ as $\delta \to 0$. Then every sequence $\{\nu^{\delta_k}_{\alpha(\delta_k)}\}$ where $\delta_k \to 0$ and $\nu^{\delta_k}_{\alpha(\delta_k)}$ is a solution of problem (29) with data C^*_δ has a convergent subsequence. The limit of every convergent subsequence is a a minimum-entropy least-squares solution. If the minimum-entropy least-squares solution ν^* is unique, then

$$\lim_{\delta \to 0} \nu^{\delta}_{\alpha(\delta)} = \nu^*$$

where ν^* is the solution of (C3).

Proof Let the sequences $\{\nu^{\delta_k}_{\alpha(\delta_k)}\}$ and δ_k be as above and ν^* be a minimum entropy least-squares solution. Then by definition of $\nu^{\delta_k}_{\alpha(\delta_k)}$ we have

$$\|F(\nu^{\delta_k}_{\alpha(\delta_k)}) - C^*_{\delta_k}\|^2 + \alpha(\delta_k) H(\nu^{\delta_k}_{\alpha(\delta_k)}) \leq \|F(\nu^*) - C^*_{\delta_k}\|^2 + \alpha(\delta_k) H(\nu^*)$$

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

$$\begin{aligned} &\| F(\nu^{\delta_k}_{\alpha(\delta_k)}) - C^* \|^2 (1-r) + \alpha(\delta_k) H(\nu^{\delta_k}_{\alpha(\delta_k)}) \\ &\qquad \leq \| F(\nu^*) - C^* \|^2 (1+r) + 2\delta_k^2/r + \alpha(\delta_k) H(\nu^*) \end{aligned} \tag{C4}$$

which implies that for all $r \in (0,1)$,

$$\alpha(\delta_k) H(\nu^{\delta_k}_{\alpha(\delta_k)}) \leq 2r \| F(\nu^*) - C^* \|^2 + 2\delta_k^2/r + \alpha(\delta_k) H(\nu^*)$$

Taking $r = \delta_k$ yields

$$H(\nu^{\delta_k}_{\alpha(\delta_k)}) \leq 2\delta_k / \alpha(\delta_k)(1 + \| F(\nu^*) - C^* \|^2) + H(\nu^*)$$

and therefore

$$\limsup_{k \to \infty} H(\nu^{\delta_k}_{\alpha(\delta_k)}) \leq H(\nu^*) \tag{C5}$$

Hence, $H(\nu^{\delta_k}_{\alpha(\delta_k)})$ is bounded and therefore $\{\nu^{\delta_k}_{\alpha(\delta_k)}\}_{k \geq 1}$ has a convergent subsequence. Let $\{\nu^{\delta_m}_{\alpha(\delta_m)}\}_{m \geq 1}$ be any such subsequence, converging to a measure ν. Substituting $r = \delta_k$ in (C4) and making k tend to ∞ yields that

$$\limsup_{k \to \infty} \| F(\nu^{\delta_k}_{\alpha(\delta_k)}) - C^* \| = \| F(\nu^*) - C^* \|$$

This, together with (C5) and the continuity of the calibration function implies that ν is a minimum-entropy least-squares solution. The last assertion of the proposition follows from the fact that in this case every subsequence of $\nu^{\delta_k}_{\alpha(\delta_k)}$ has a further subsequence converging towards ν^*. □

REFERENCES

Ait-Sahalia, Y., and Lo, A. (1998). Nonparametric estimation of state–price densities implicit in financial asset prices. *Journal of Finance* **53**, 499–547.

Andersen, L., and Andreasen, J. (2000). Jump diffusion models: volatility smile fitting and numerical methods for pricing. *Review of Derivatives Research* **4**, 231–62.

Avellaneda, M. (1998). Minimum entropy calibration of asset pricing models. *International Journal of Theoretical and Applied Finance* **1**, 447–72.

Avellaneda, M., Buff, R., Friedman, C., Grandchamp, N., Kruk, L., and Newman, J. (2001). Weighted Monte Carlo: a new technique for calibrating asset-pricing models. *International Journal of Theoretical and Applied Finance* **4**, 91–119.

Barndorff-Nielsen. O. E. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics*, 41–68.

Barndorff-Nielsen, O. E., Mikosch, T., and Resnick, S. (eds) (2001). *Lévy processes – theory and applications.* Birkhauser, Boston.

Bates, D. S. (1996a). Jumps and stochastic volatility: the exchange rate processes implicit in Deutschmark options. *Review of Financial Studies* **9**, 69–107.

Bates, D. S. (1996b). Testing option pricing models. In *Statistical methods in finance*, vol. 14 of *Handbook of Statististics*, pp. 567–611. North-Holland, Amsterdam.

Beckers, S. (1981). A note on estimating parameters of a jump–diffusion process of stock returns. *Journal of Financial and Quantitative Analysis* **16**, 127–40.

Bertoin, J. (1996). *Lévy Processes.* Cambridge University Press.

Byrd, R., Lu, P., and Nocedal, J. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing* **16**, 1190–208.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business* **75**, 305–32.

Carr, P., and Madan, D. (1998). Option valuation using the fast Fourier transform. *Journal of Computational Finance* **2**, 61–73.

Cont, R., Bouchaud, J.-P., and Potters, M. (1997). Scaling in financial data: stable laws and beyond. In B. Dubrulle, F. Graner and D. Sornette (eds), *Scale invariance and beyond*. Springer, Berlin.

Cont, R., and da Fonseca, J. (2002). Dynamics of implied volatility surfaces. *Quantitative Finance* **2**, 45–60.

Cont, R., and Tankov, P. (2004a). *Financial modelling with jump processes.* Chapman & Hall/CRC Press.

Cont, R. and Tankov, P. (2004b). Identifying exponential Lévy models from option prices using relative entropy. Working paper, CMAP.

Dupire, B. (1994). Pricing with a smile. *Risk* **7**, 18–20.

Eberlein, E. (2001). *Applications of generalized hyperbolic Lévy motion to finance.* In O. E. Barndorff-Nielsen, T. Mikosch and S. Resnick (eds), *Lévy processes – theory and applications*, pp. 319–36. Boston: Birkhauser.

Eberlein, E., Keller, U., and Prause, K. (1998). New insights into smile, mispricing and value at risk: the hyperbolic model. *Journal of Business* **71**, 371–405.

El Karoui, N., and Rouge, R. (2000). Pricing via utility maximization and entropy. *Mathematical Finance* **10**, 259–76.

Engl, H. W., Hanke, M., and Neubauer, A. (1996). *Regularization of inverse problems.* Vol. 375 of *Mathematics and its Applications*, Kluwer Academic Publishers Group, Dordrecht.

Föllmer, H., and Schied, A. (2002). *Stochastic finance.* Berlin: De Gruyter.

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

Goll, T., and Rüschendorf, L. (2001). Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics* **5**, 557–81.

Jacod, J., and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes,* 2nd edition. Springer, Berlin.

Kallsen, J. (2001). Utility-based derivative pricing. In *Mathematical Finance – Bachelier Congress 2000*, Springer, Berlin.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated Lévy fiights towards the Gaussian stochastic process. *Physical Review E* **52**, 1197–99.

Kou, S. G. (2002). A jump diffusion model for option pricing. *Management Science* **48**, 1086–101.

Madan, D. (2001). Financial modeling with discontinuous price processes. In O. E. Barndorff-Nielsen, T. Mikosch and S. Resnick (eds), *Lévy processes – theory and applications*. Birkhauser, Boston.

Madan, D., and Konikov, M. (2002). Option pricing using variance gamma Markov chains. *Review of Derivatives Research* **5**, 81–115.

Madan, D. B., Carr, P., and Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review* **2**, 79–105.

Mancini, C. (2001). Disentangling the jumps from the diffusion in a geometric jumping Brownian motion. *Giornale dell 'Istituto Italiano degli Attuari* **LXIV**, 19–47.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* **3**, 125–44.

Miyahara, Y., and Fujiwara, T. (2003). The minimal entropy martingale measures for geometric Lévy processes. *Finance and Stochastics* **7**, 509–31.

Morozov, V. (1966). On the solution of functional equations by the method of regularization. *Soviet Mathematics Doklady* **7**, 414–17.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C: the art of scientific computing.* Cambridge University Press.

Samperi, D. (2002). Calibrating a diffusion pricing model with uncertain volatility: regularization and stability. *Mathematical Finance* **12**, 71–87.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press.

Schoutens, W. (2002). Meixner processes: theory and applications in finance. EURANDOM report 004.