An overview of a nonparametric estimation method for Lévy processes

Enrique Figueroa-López
Department of Mathematics
Purdue University
150 N. University Street,
West Lafayette, IN 47906
jfiguero@math.purdue.edu

Christian Houdré*
Laboratoire d' Analyse et de
Mathématiques Appliquées
CNRS UMR 8050
Université Paris XII
94010 Créteil Cedex, France
and
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160, USA
houdre@math.gatech.edu

1st October 2005

*Research supported in part by NSF and NSA grants.
Key words and phrases: Lévy processes, Penalized projection estimator, Model selection, Oracle inequalities, Minimax risk on Besov spaces, Adaptive estimation
AMS Subject Classification (2000): 62G05, 62P05, 60G51, 60E07

Abstract

A nonparametric method for the estimation of the Lévy density of a process X is developed. Estimators that can be written in terms of the "jumps" of X are introduced, and so are discrete-data based approximations. A model selection approach made up of two steps is investigated. The first step consists in the selection of a good estimator from a linear model of proposed Lévy densities, while the second is a data-driven selection of a linear model among a given collection of linear models. By providing lower bounds for the minimax risk of estimation over Besov Lévy densities, our estimators are shown to achieve the "best" rate of convergence. A simulation study for the case of histogram estimators applied to the variance Gamma process, a classical model in mathematical finance, is presented.

1 Introduction

We present an overview of some new estimation methods for Lévy processes introduced in [13]. The main goal here is to give the key points behind the estimation paradigm and to summarize the main results. Proofs and further simulation experiments can be found in the paper just mentioned.

We propose efficient model-free statistical procedures that generate functional estimates for the Lévy density. Being nonparametric, we relax the dependency on the model and expect that data itself validates the best model. Three typed of theoretical results serve as foundations for our methodology: i) the characterization of the jumps associated with a Lévy process as a spatial Poisson process, ii) some recent methods for the nonparametric estimation of spatial Poisson processes introduced in [18], and iii) the short-term statistical properties of Lévy processes to approximate jump-dependent quantities. Our main motivation came from mathematical finance where Lévy-based models have received great attention due to their ability to capture stylized empirical features of financial data (see e.g. [9], [11], and [2]). It is relevant to point out that our procedures are suitable for high-frequency data, which is widely available nowadays. Furthermore, it is precisely for such data that the standard maximum-likelihood based statistical methods are not viable, the traditional *geometric Brownian motion* model is inaccurate, and general exponential Lévy models may be more relevant.

Let us describe the outline of the paper. We introduce the nonparametric estimators in Sections 2 and 3. These estimators can be written in terms of integrals of deterministic functions with respect to the random measure associated with the jumps of X. The proposed method follows the reasoning of the works on minimum contrast estimation on sieves and model selection developed in the context of density estimation and nonlinear regression in [7] (see also [5] and [8]), and recently extended to the estimation of intensity functions for Poisson processes in [18]. Concretely, the procedure addresses two problems: 1) the selection of a good estimator, called the *projection estimator*, from an approximating linear model $\mathcal{S}$ for the Lévy density, and 2) the selection of a linear model among a given collection of linear models using a penalization technique that lead to *penalized projection estimators* (p.p.e.). In Sections 4 and 5, we examine the fact that the Poisson jump measure cannot be retrieved from discrete observation, and suggests a natural approximation procedure for Poisson integrals based on equally spaced observations of the process. Section 6 illustrates how our methods works in a classical model used in mathematical finance: the *Variance Gamma model* of [10]. The Lévy processes are simulated using time series representations and "discrete skeletons", whereas the considered estimators are mainly regular histograms. In the last part we present some good theoretical qualities our methods enjoy. Concretely, a bound for the *risk* of the p.p.e. is found in Section 7. As a consequence, *Oracle inequalities*, that ensure to approximately attain the best expected error (using projection estimators) up to a constant, are obtained. We also assess the rate of convergence of the p.p.e. on *regular splines*, when the Lévy density belongs to some *Besov spaces*. By analyzing the *minimax risk* of estimation on these Besov spaces, it is actually proved in Section 9 that the p.p.e. attains the best possible rate in the minimax sense, when the estimation is based on jumps bounded away from the origin. We finish with some concluding remarks in Section 10.

2 A model-free estimation method

Consider a real Lévy process $X = \{X(t)\}_{t\geq 0}$ with Lévy density $p : \mathbb{R}\backslash\{0\} \to \mathbb{R}_+$. Let us denote

$$\mathcal{J}(B) \equiv \# \left\{ t > 0 : (t, X(t) - X(t^-)) \in B \right\}, \tag{2.1}$$

the random measure associated with the jumps of X, which is a Poisson measure on $\mathcal{B}([0,\infty)\times\mathbb{R}_0)$ with mean measure

$$\mu(B) = \iint_B p(x)\,dt\,dx. \tag{2.2}$$

We consider the problem of estimating the Lévy density p on a Borel set $D \in \mathcal{B}(\mathbb{R}_0)$ using a *projection estimation approach.* According to this paradigm, p is estimated by estimating its best approximating function in a finite-dimensional linear space $\mathcal{S}$. The linear space $\mathcal{S}$ is taken so that it has good approximation properties in general classes of functions. Typical choices are piecewise polynomials or wavelets. We take the following standing assumption.

Assumption 1 *The Lévy measure $\nu(dx) \equiv p(x)dx$ is absolutely continuous with respect to a known measure η on $\mathcal{B}(D)$ so that the Radon-Nikodym derivative*

$$\frac{d\nu}{d\eta}(x) = s(x), \quad x \in D, \tag{2.3}$$

is positive, bounded, and satisfies

$$\int_D s^2(x)\eta(dx) < \infty. \tag{2.4}$$

In that case, we call s the Lévy density on D of the process with respect to the reference measure η.

Remark 2.1 *Under the previous assumption, the measure $\mathcal{J}$ of (2.1), when restricted to $\mathcal{B}([0,\infty)\times D)$, is a Poisson process with mean measure*

$$\mu(B) = \iint_B s(x)\,dt\,\eta(dx), \quad B \in \mathcal{B}([0,\infty)\times D). \tag{2.5}$$

Our goal will be to estimate the Lévy density s, which itself could in turn be used to retrieve p on D via (2.3). To illustrate this strategy consider a continuous Lévy density p such that

$$p(x) = O\left(x^{-1}\right), \quad \text{as } x \to 0.$$

This type of densities satisfy the assumption above with respect to the measure $\eta(dx) = x^{-2}dx$ on domains of the form $D = \{x : 0 < |x| < b\}$. Clearly, each estimator $\hat{s}$ for s will induce the natural estimator $x^{-2}\hat{s}(x)$ for p.

Let us describe the main ingredients of our approach. Let $\mathcal{S}$ be a finite dimensional subspace of $L^2 \equiv L^2((D, \eta))$ equipped with the standard norm

$$\|f\|^2 \equiv \int_D f^2(x)\, \eta(dx).$$

The *projection estimator* of s on S is defined by

$$\hat{s}(x) \equiv \sum_{i=1}^{d} \hat{\beta}_i \varphi_i(x), \tag{2.6}$$

where $\{\varphi_1, \ldots, \varphi_d\}$ is an arbitrary orthonormal basis of $\mathcal{S}$ and

$$\hat{\beta}_i \equiv \frac{1}{T} \iint\limits_{[0,T]\times D} \varphi_i(x) \mathcal{J}(dt, dx). \tag{2.7}$$

This is a natural unbiased estimator for the *orthogonal projection* $s^\perp$ of s on the subspace $\mathcal{S}$, namely

$$s^\perp \equiv \sum_{i=1}^{d} \left(\int_D \varphi_i(y) s(y) \eta(dy) \right) \varphi_i(x). \tag{2.8}$$

The following proposition provides the first- and second-order properties $\hat{s}$

Proposition 2.2 *Under the Assumption 1, $\hat{s}$ is an unbiased estimator for $s^\perp$. Its integrated mean-square error $\chi^2 \equiv \|\hat{s} - s^\perp\|$ is*

$$\mathbb{E}\left[\chi^2\right] = \frac{1}{T} \sum_{i=1}^{d} \int_D \varphi_i^2(x) s(x)\, \eta(dx), \tag{2.9}$$

while its risk *admits the decomposition*

$$\mathbb{E}\left[\|s - \hat{s}\|^2\right] = \|s - s^\perp\|^2 + \mathbb{E}\left[\chi^2\right]. \tag{2.10}$$

The first term in (2.10), called the *bias term*, accounts for the distance of the unknown function s to the model $\mathcal{S}$, while the second term, called the *variance*, measures the error of estimation within within the linear model $\mathcal{S}$.

3 Model selection via penalized projection estimator

A crucial issue in the above approach is the selection of the approximating linear model $\mathcal{S}$. In principle, "nice" Lévy density s can be approximated closely by general linear models such as splines or wavelet. However, a more robust model $\mathcal{S}'$ containing $\mathcal{S}$ will result in a better approximation of s, but in a larger variance. This rise the natural problem of selecting one model, out of a collection of linear models $\{\mathcal{S}_m, m \in \mathcal{M}\}$, that approximately realizes the best trade-off between the risk of estimation within the model and the distance of the unknown Lévy density to the approximating model.

Let $\hat{s}_m$ and $s_m^{\perp}$ be respectively the projection estimator and the orthogonal projection of s on $\mathcal{S}_m$. The following equation, readily derived from (2.10), gives insight on a sensible solution to the model selection problem:

$$\mathbb{E}\left[\|s - \hat{s}_m\|^2\right] = \|s\|^2 + \mathbb{E}\left[-\|\hat{s}_m\|^2 + \text{pen}(m)\right], \tag{3.1}$$

where $\text{pen}(m)$ is defined in terms of an orthonormal basis $\{\varphi_{1,m}, \ldots, \varphi_{d_m,m}\}$ for $\mathcal{S}_m$ by the equation:

$$\text{pen}(m) = \frac{2}{T^2} \iint\limits_{[0,T]\times D} \left(\sum_{i=1}^{d_m} \varphi_{i,m}^2(x) \right) \mathcal{J}(dt, dx). \tag{3.2}$$

Equation (3.1) shows that the risk of $\hat{s}_m$ moves "parallel" to the expectation of the *observable statistics* $-\|\hat{s}_m\|^2 + \text{pen}(m)$. This fact justifies to choose the model that minimizes such statistics. We will see later that other choices for $\text{pen}(\cdot)$ give also good results. Therefore, given a penalization function $\text{pen} : \mathcal{M} \to [0, \infty)$, we consider estimators of the form

$$\tilde{s} \equiv \hat{s}_{\hat{m}}, \tag{3.3}$$

where $\hat{s}_m$ is the projection estimator on $\mathcal{S}_m$ and

$$\hat{m} \equiv \text{argmin}_{m \in \mathcal{M}} \left\{ -\|\hat{s}_m\|^2 + \text{pen}(m) \right\}.$$

An estimator $\tilde{s}$ as in (3.3) is called a **penalized projection estimator** (p.p.e.) on the collection of linear models $\{\mathcal{S}_m, m \in \mathcal{M}\}$.

4 Calibration based on discrete time data

One drawback to the method outlined in Section 2 is that in general we do not observe the jumps of a Lévy process $X = \{X(t)\}_{t\geq 0}$. In practice, we can aspire to sample the process $X(t)$ at discrete times, but we are neither able to measure the size of the jumps $\Delta X(t) \equiv X(t) - X(t^-)$ nor the times of jumps $\{t : \Delta X(t) > 0\}$. In general, Poisson integrals of the type

$$I\,(f) \equiv \iint\limits_{[0,T]\times\mathbb{R}_0} f(x)\mathcal{J}(dt, dx) = \sum_{t\leq T} f(\Delta X(t)), \tag{4.1}$$

are not accessible. In this section, we discuss a very simple approximation of the integral (4.1) based on equally spaced observations of the process in the time interval $[0, T]$. Indeed, the following statistic is the most natural approximation to (4.1):

$$I_n\,(f) \equiv \sum_{k=1}^{n} f\,(\Delta_k X)\,, \tag{4.2}$$

where $\Delta_k X$ is the k^{th} increment of the process with time span $h_n := T/n$; that is,

$$\Delta_k X \equiv X\,(kh_n) - X\,((k-1)h_n)\,, \quad k = 1, \ldots, n.$$

How good is the approximation and in what sense? Under some conditions on f, we can readily prove the weak convergence of (4.2) to (4.1) using properties of the transition distributions of X in small time (see [6], Corollary 8.9 of [22], and [21]). Let us summarize some of these results.

Theorem 4.1 *Let $X = \{X(t)\}_{t\geq 0}$ be a Lévy process with Lévy measure ν. Then, the following statement holds true.*

(1) *For each $a > 0$,*

$$\lim_{t\to 0} \frac{1}{t}\mathbb{P}\,(X(t) > a) = \nu([a, \infty)), \tag{4.3}$$

$$\lim_{t\to 0} \frac{1}{t}\mathbb{P}\,(X(t) \leq -a) = \nu((-\infty, -a]). \tag{4.4}$$

(2) *For any continuous bounded function h vanishing on a neighborhood of the origin,*

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}\left[h\left(X(t)\right)\right]=\int_{\mathbb{R}_0} h(x)\nu(dx). \tag{4.5}$$

(3) *If h is continuous and bounded and if $\lim_{|x|\to 0} h(x)|x|^{-2}=0$, then*

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}\left[h\left(X(t)\right)\right]=\int_{\mathbb{R}_0} h(x)\nu(dx).$$

Moreover, if $\int_{\mathbb{R}_0}(|x|\wedge 1)\nu(dx)<\infty$, it suffices to have $h(x)(|x|\wedge 1)^{-1}$ is continuous and bounded.

Convergence results like (4.5) are useful to establish the convergence in distribution of $\mathrm{I}_n(f)$ since

$$\mathbb{E}\left[e^{\mathrm{i}uI_n(f)}\right]=\left(\mathbb{E}\left[e^{\mathrm{i}uf\left(X\left(\frac{T}{n}\right)\right)}\right]\right)^n=\left(1+\frac{a_n}{n}\right)^n,$$

where $a_n=n\mathbb{E}\left[h\left(X\left(\frac{T}{n}\right)\right)\right]$ with $h(x)=e^{\mathrm{i}uf(x)}-1$. So, if f is such that

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}\left[e^{\mathrm{i}uf(X(t))}-1\right]=\int_{\mathbb{R}_0}\left(e^{\mathrm{i}uf(x)}-1\right)\nu(dx), \tag{4.6}$$

then a_n converges to $a\equiv T\int_{\mathbb{R}_0} h(x)\nu(dx)$, and thus

$$\lim_{n\to\infty}\left(1+\frac{a_n}{n}\right)^n=\lim_{n\to\infty} e^{n\log\left(1+\frac{a_n}{n}\right)}=e^a.$$

We thus have the following result.

Proposition 4.2 *Let $X=\{X(t)\}_{t\geq 0}$ be a Lévy process with Lévy measure ν. Then,*

$$\lim_{n\to\infty}\mathbb{E}\left[e^{\mathrm{i}uI_n(f)}\right]=\exp\left\{T\int_{\mathbb{R}_0}\left(e^{\mathrm{i}uf(x)}-1\right)\nu(dx)\right\},$$

if f satisfies either one of the following:

(1) *$f(x)=\mathbf{1}_{(a,b]}(x)h(x)$ for an interval $[a,b]\subset\mathbb{R}_0$ and a continuous function h;*

(2) $f(x)$ is continuous on $\mathbb{R}_0$ and $\lim_{|x|\to 0} f(x)|x|^{-2} = 0$.

In particular, $I_n(f)$ converges in distribution to $I(f)$ under any one of the previous two conditions.

Remark 4.3 *Clearly, if f and f^2 satisfy (4.5), then the mean and variance of $I_n(f)$ obey the asymptotics:*

$$\lim_{n\to\infty} \mathbb{E}\left[I_n(f)\right] = T \int_{\mathbb{R}_0} f(x)\nu(dx);$$

$$\lim_{n\to\infty} \mathrm{Var}\left[I_n(f)\right] = T \int_{\mathbb{R}_0} f^2(x)\nu(dx).$$

Remark 4.4 *Very recently, [16] proposed a procedure to disentangle the jumps from the diffusion part in the case of jump-diffusion models driven by finite-jump activity Lévy processes. It is proved there that for certain functions $r : \mathbb{R}_+ \to \mathbb{R}_+$, there exists $N(\omega)$ such that for $n \geq N(\omega)$, a jump occurs in the interval $((k-1)h_n, kh_n]$ if and only if $(\Delta_k X)^2 > r(h_n)$. Here, $h_n = T/n$ and $\Delta_k X$ is the k^{th} increment of the process. These results suggest to use statistics of the form*

$$\sum_{k=1}^{n} f\left(\Delta_k X\right) \mathbf{1}\left[\left(\Delta_k X\right)^2 > r(h_n)\right]$$

instead of (4.2) to approximate the integral (4.1).

5 Estimation Method

Let us summarize the previous sections and outline the proposed algorithm of estimation:

Model and data: The Lévy process $\{X(t)\}_{t\in[0,T]}$ satisfies the assumption 1, and is sampled at equally spaced times $t_k^n = k\frac{T}{n}$, $k = 1, \ldots, n$, during the time period $[0, T]$.

Statistician's parameters: The procedure is fed with a Borel *window of estimation* $D \subset \mathbb{R}_0$, a collection $\{\mathcal{S}_m\}_{m \in \mathcal{M}}$ of finite dimensional *linear models* of $L^2\left((D, \eta)\right)$, and a *level of penalization* $c > 1$.

Estimators: Inside the linear model $\mathcal{S}_m$, the estimator of s is the *approximated projection estimator*:

$$\hat{s}_m^n(x) \equiv \sum_{i=1}^{d_m} \hat{\beta}_{i,m}^n \varphi_{i,m}(x), \tag{5.1}$$

where $\{\varphi_{1,m}, \ldots, \varphi_{d_m,m}\}$ is an orthonormal basis for $\mathcal{S}_m$, and

$$\hat{\beta}_{i,m}^n \equiv \frac{1}{T} \sum_{k=1}^{n} \varphi_{i,m}\left(X\left(t_k^n\right) - X\left(t_{k-1}^n\right)\right). \tag{5.2}$$

Across the collection of linear models $\{\mathcal{S}_m : m \in \mathcal{M}\}$, the estimator $\hat{s}_m^n$ which minimizes

$$-\|\hat{s}_m^n\|^2 + c\,\mathrm{pen}^n(m)$$

is selected, where

$$\mathrm{pen}^n(m) = \frac{1}{T^2} \sum_{k=1}^{n} \left(\sum_{i=1}^{d_m} \varphi_{i,m}^2 \left(X\left(t_k^n\right) - X\left(t_{k-1}^n\right)\right) \right).$$

The following theorem shows that the above procedure, based on discrete samples of the process, achieves the same risk as if the estimators were based on time-continuous observation of the process. Let $\mathcal{R}(X)$ be the linear space of measurable functions h such that (4.5) is satisfied. For instance, $\mathcal{R}(X)$ contains the functions f satisfying conditions (1) or (2) in Proposition 4.2.

Proposition 5.1 *Let $s_m^\perp$ be the orthogonal projection of s on $\mathcal{S}_m$. If $\varphi_{i,m}$ and $\varphi_{i,m}^2$ belong to $\mathcal{R}(X)$ for every $m \in \mathcal{M}$ and $i = 1, \ldots, d_m$, then the approximated projection estimator $\hat{s}_m^n$ of s on $\mathcal{S}_m$ (based on n equally spaced observations) satisfies:*

$$\lim_{n \to \infty} \mathbb{E}\left[\|\hat{s}_m^n - s_m^\perp\|^2\right] = \mathbb{E}\left[\|\hat{s}_m - s_m^\perp\|^2\right]. \tag{5.3}$$

Moreover,

$$\lim_{n \to \infty} \mathbb{E}\left[\|\hat{s}_m^n - s\|^2\right] = \mathbb{E}\left[\|\hat{s}_m - s\|^2\right].$$

6 Numerical tests of projection estimators

In this section, we try to assess the performance of some penalized projection estimators based on simulated Lévy processes. Piecewise constant functions are considered, and for their intrinsic relevance in mathematical finance, two classes of Lévy processes are studied: Gamma and variance Gamma processes. A least-squares method is also applied to generate parametric Lévy densities that fit the nonparametric outputs. Simulations are based on the series representation for Lévy processes introduced by [19].

6.1 Estimation of the Gamma Lévy process

The approximating linear models considered here are the span of the indicator functions $\chi_{[x_0,x_1)},\ldots,\chi_{[x_{m-1},x_m)}$, where $x_0 < \cdots < x_n$ is a regular partition of an interval $D \equiv [a,b]$, with $0 < a$ or $b < 0$. Figure 1 shows the Gamma Lévy density $\frac{\alpha}{x}e^{x/\beta}$ with $\alpha = 1$ and $\beta = 1$, and the penalized projection estimator. A fit of the model $\frac{\alpha}{x}e^{x/\beta}$ to the non-parametric output using the least-square method yields the estimators $\hat{\alpha} = 0.932$ and $\hat{\beta} = 1.055$. The maximum likelihood estimators based on the increments of the sample path of time length 1 are 1.015 for α and 0.949 for β.

Another approach is to consider a different reference measure than the Lebesgue measure. In the notation of Assumption 1, we can estimate the Lévy density $s(x) = \alpha x e^{-x/\beta}$ with respect to the measure $\eta(dx) = \frac{1}{x^2}dx$. Once an estimator $\hat{s}$ for s has been obtained, $\hat{p}(x) = \hat{s}(x)/x^2$ can work as an estimator for the Lévy density $p(x) = \alpha e^{-x/\beta}/x$. Let us consider linear models of the form

$$\mathcal{S}_{\mathcal{C}} = \left\{ f(x) \equiv c_1 x \chi_{[x_0,x_1)}(x) + \sum_{i=2}^{m} c_i \, \chi_{[x_i,x_{i+1})}(x) : c_1,\ldots,c_m \in \mathbb{R} \right\},$$

where $x_0 < x_1 < \cdots < x_m = b$ is a regular partition of an interval $D = [0,b]$. We apply the above method to the simulated Lévy process used in Figure 1. Figure 2 shows the estimator $\hat{p}_2(x) = \hat{s}(x)/x^2$ and the actual Lévy density $p(x) = \exp(-x)/x$ for $x \in [0.02, 1]$ (we used regular partitions on $[0,1]$). From Figure 1, the improvement is notorious, and moreover, we accomplish a good estimation around the origin of $\hat{p}_2(x) = 0.9/x$, for $x \in [0, 0.2)$.

Remark 6.1 *The procedure just described is appropriate to estimate the density function $s(x) = \frac{\alpha}{x}\exp(-\frac{x}{\beta})$ around the origin as far as*

$$\hat{\alpha} \equiv \frac{1}{Tx_1}\sum_{t\leq T}\Delta X(t)\mathrm{I}\left[\Delta X(t) < x_1\right],$$

is a good estimator of α. It is not hard to check that the bias of $\hat{\alpha}$ tend to zero as $x_1 \searrow 0$. However, the variance of $\hat{\alpha}$ converges to $\frac{\alpha}{2T}$, suggesting that the method works better when T is "large" and α is "small".

6.2 Performance of projection estimation based on finitely many observation

In this part, we study the performance of the (approximate) projection estimators introduced in Section 4, and formally given in Section 5.

Table 1 compares the (approximate) penalized projection estimators with least-square errors (PPE-LSE) to the maximum likelihood estimators (MLE) for the Gamma Lévy process with $\alpha = \beta = 1$ using different time spans Δt. We also consider two types of simulations: jump-based and increment-based. The method based on jumps uses series representation with $n = 36500$ jumps occurring during the time period $[0, 365]$ (notice that if we think of 365 as days, the number of jumps corresponds to a rate of about 1 jump every 5 minute). The increment-based method is a *discrete skeleton* with mesh of 0.001. Notice that maximum likelihood estimation does not do well for small time spans when the approximate sample path is based on jumps. On the other hand, penalized projection estimation does not provide good results for long time spans when the approximate sample path is based on increments. The sampling distributions of the MLE for α and β are shown in Figures 5 and 6 in the case of $\Delta t = 0.1$. On the other hand, the sampling distributions of the estimates for α and β obtained from fitting the PPE are given in Figures 7 and 8. Even though, the MLE is superior, the estimates based PPE have good performance considering that they are model-free.

6.3 Estimation of variance Gamma processes.

Variance Gamma processes were proposed in [10] as substitutes to the Brownian Motion in the Black-Scholes model. This a pure-jump Lévy process with

	Jump-based Simulation				Increment-based Simulation			
Δt	PPE-LSE		MLE		PPE-LSE		MLE	
1	1.01	1.46	.997	.995	.73	1.78	1.09	.99
0.5	1.03	1.09	.972	.978	.9	1.49	1.01	1.06
0.1	.944	.995	1.179	.837	.923	1.03	.989	1.09
0.01	.969	.924	6.92	.5	.955	1.019	.9974	1.083

Table 1: Estimation of a Lévy Gamma process with $\alpha = \beta = 1$. Two types of simulation are considered: series-representation based and increments-based. The estimations are based on equally spaced sampling observation at the time span Δt. Results for the approximate penalized projection estimators with least-squares errors, and for the maximum likelihood estimators are given.

Lévy density

$$s(x) = \begin{cases} \frac{\alpha}{|x|} \exp\left(-\frac{|x|}{\beta_-}\right) & \text{if } x < 0, \\ \frac{\alpha}{x} \exp\left(-\frac{x}{\beta_+}\right) & \text{if } x > 0, \end{cases}$$

where $\alpha > 0$, $\beta_- \geq 0$, and $\beta_+ \geq 0$ (of course, $\beta_-^2 + \beta_+^2 > 0$).

From an algorithmic point of view, the estimation for the variance Gamma model using penalized projection is not different from the estimation for the Gamma process. We can simply estimate both tails of the variance Gamma process separately. However, from the point of view of maximum likelihood estimation (MLE), the problem is numerically challenging. Even though the marginal density functions have closed form expressions (see [10]), there are well-documented issues with MLE (see for instance [17]). The likelihood function is highly flat for a wide range of parameters and good starting values as well as convergence are critical. Also, the separation of parameters and the identification of the variance Gamma process from other classes of the generalized hyperbolic Lévy processes is difficult. In fact, difference between subclasses in terms of likelihood is small. It is important to mention that these issues worsen when dealing with "high-frequency" data.

Let us consider a numerical example motivated by the empirical findings of [10] based on daily returns on the S&P stock index from January 1992 to September 1994 (see their Table I). Using maximum likelihood methods, the annualized estimates of the parameters for the variance Gamma model were reported to be $\hat{\theta} = -0.00056256$, $\hat{\sigma}^2 = 0.01373584$, and $\hat{\nu} = 0.002$. After simulating the variance Gamma with these parameters and applying our methods, Figures 3 and 4 show respectively the left- and right- tails of the Lévy density and their penalized projection estimators as well as their corresponding best-fit variance Gamma Lévy densities using a least-squares method, and their marginal probability density functions (pdf) scaled by $1/\Delta t$ (the reciprocal of the time span between observations). The estimation was based on 5000 simulated increments with Δt equal to one-eight of a day. The figures seem quite comforting. To get a better picture, Figures 9 and 10 show the sampling distributions of the estimates for α_- and β_+ obtained from applying the least-square method to the penalized projection estimators. The histograms are based on 1000 samples of size 5000 with $\Delta t = 1/8$ of a day. This experiment shows a clear, though not very serious, underestimation of the parameter α, and a clear overestimation of the β parameters. A simple method of moments (based on the first four moments) yields better results (see Figures 11 and 12). Nonparametric methods are not free-lunches and usually the gain in robustness is paid by a loss in precision.

7 Risk bound and oracle inequalities

The model selection approach of Section 3 allows to choose an estimator out of the projection estimators $\{\hat{s}_m : m \in \mathcal{M}\}$ corresponding to a collection of approximating linear models $\{\mathcal{S}_m, m \in \mathcal{M}\}$. Of course, the method will not necessarily select the best model, namely

$$\bar{m} \equiv \text{argmin}_{m\in\mathcal{M}}\, \mathbb{E}\left[\|s - \hat{s}_m\|^2\right], \tag{7.1}$$

but how bad would be the risk of $\tilde{s}$ compared to the best risk that can be achieved by the class of projection estimators? In this section we show that certain penalized projection estimators $\tilde{s}$ satisfy the inequality

$$\mathbb{E}\left[\|s - \tilde{s}\|^2\right] \leq C \inf_{m\in\mathcal{M}} \mathbb{E}\left[\|s - \hat{s}_m\|^2\right] + \frac{C'}{T},$$

for some constants C, C' (remember that the time period of observations is $[0, T]$). Since the approximating model $\mathcal{S}_{\bar{m}}$ with minimal risk (measure by the risk of its projection estimator) is called the *oracle model*, inequalities of the above type are called *oracle inequalities*. The main tool in obtaining Oracle inequalities is an upper bound for the risk of the penalized projection estimator $\tilde{s}$ of (3.3).

Let us introduce some notation. Below, d_m denotes the dimension of the linear model $\mathcal{S}_m$, and $\{\varphi_{1,m}, \ldots, \varphi_{d_m,m}\}$ is an arbitrary orthonormal basis of $\mathcal{S}_m$. Define

$$D_m = \sup\left\{\|f\|_\infty^2 : f \in S_m, \|f\|^2 = 1\right\}, \tag{7.2}$$

which can be proved to be equal to $\|\sum_{i=1}^{d_m} \varphi_{i,m}^2\|_\infty$.

We take the following regularity condition, introduced in [18], that essentially controls the complexity of the linear models.

Assumption 2 *There exist constants $\Gamma > 0$ and $R \geq 0$ such that for every positive integer n,*

$$\#\{m \in \mathcal{M} : d_m = n\} \leq \Gamma n^R.$$

We now present the main result of this section.

Theorem 7.1 *Let $\{\mathcal{S}_m, m \in \mathcal{M}\}$ be a family of finite dimensional linear subspaces of $L^2((D,\eta))$ satisfying Assumption 2. Define $\mathcal{M}_T \equiv \{m \in \mathcal{M} : D_m \leq T\}$. If $\hat{s}_m$ and $s_m^\perp$ are respectively the projection estimator and the orthogonal projection of the Lévy density s on $\mathcal{S}_m$ then, the penalized projection estimator $\tilde{s}_T$ on $\{\mathcal{S}_m\}_{m\in\mathcal{M}_T}$ defined by (3.3) is such that*

$$\mathbb{E}\left[\|s - \tilde{s}_T\|^2\right] \leq C \inf_{m\in\mathcal{M}_T} \left\{\|s - s_m^\perp\|^2 + \mathbb{E}\left[\mathrm{pen}(m)\right]\right\} + \frac{C'}{T}, \tag{7.3}$$

whenever $\mathrm{pen} : \mathcal{M} \to [0,\infty)$ *takes one of the following forms for some fixed (but arbitrary) constants $c > 1$, $c' > 0$, and $c'' > 0$:*

(a) $\mathrm{pen}(m) \geq c\frac{D_m \mathcal{N}}{T^2} + c'\frac{d_m}{T}$, *where $\mathcal{N} \equiv \mathcal{J}([0,T]\times D)$ is the number of jumps prior to T with sizes in D and where it is assumed that $\rho \equiv \int_D s(x)\eta(dx) < \infty$;*

(b) $\mathrm{pen}(m) \geq c\frac{\hat{V}_m}{T}$, *where $\hat{V}_m$ is defined by*

$$\hat{V}_m \equiv \frac{1}{T} \iint_{[0,T]\times D} \left(\sum_{i=1}^{d_m} \varphi_{i,m}^2(x)\right) \mathcal{J}(dt, dx), \tag{7.4}$$

and where it is assumed $\beta \equiv \inf_{m\in\mathcal{M}} \frac{\mathbb{E}[\hat{V}_m]}{D_m} > 0$ and $\phi \equiv \inf_{m\in\mathcal{M}} \frac{D_m}{d_m} > 0$;

(c) $\text{pen}(m) \geq c\frac{\hat{V}_m}{T} + c'\frac{D_m}{T} + c''\frac{d_m}{T}$.

Moreover, the constant C depends only on c, c' and c'', while C' varies with c, c', c'', Γ, R, $\|s\|$, $\|s\|_\infty$, ρ, β, and ϕ.

Remark 7.2 *It can be shown that if $c \geq 2$, then for arbitrary $\varepsilon > 0$, there is a constant $C'(\varepsilon)$ (increasing) so that*

$$\mathbb{E}\|s - \tilde{s}\|^2 \leq (1+\varepsilon) \inf_{m \in \mathcal{M}} \left\{ \|s - s_m^{\perp}\|^2 + \mathbb{E}\left[\text{pen}(m)\right] \right\} + \frac{C'(\varepsilon)}{T}. \tag{7.5}$$

As a first application of the previous risk bound, the following oracle inequalities can be immediately derived:

Corollary 7.3 *In the setting of Theorem 7.1, if the penalty function is of the form* $\text{pen}(m) \equiv c\frac{\hat{V}_m}{T}$*, for every $m \in \mathcal{M}_T$, $\beta > 0$, and $\phi > 0$, then*

$$\mathbb{E}\left[\|s - \tilde{s}_T\|^2\right] \leq \tilde{C} \inf_{m \in \mathcal{M}_T} \left\{ \mathbb{E}\left[\|s - \hat{s}_m\|^2\right] \right\} + \frac{\tilde{C}'}{T}, \tag{7.6}$$

for a constant $\widetilde{C}$ depending only on c, and a constant $\widetilde{C}'$ depending on c, Γ, R, $\|s\|$, $\|s\|_\infty$, β, and ϕ.

8 Rate of convergence to smooth Lévy densities

We use the risk bound in the previous section to study the "long run" ($T \to \infty$) rate of convergence of penalized projection estimators on regular piecewise polynomials, when the Lévy density is "smooth". More precisely, restricted to the window of estimation $D \equiv [a, b] \subset \mathbb{R}_0$, the Lévy density s is assumed to belong to the *Besov space* or *Lipschitz* space $\mathcal{B}^{\alpha}_{\infty}(\mathbb{L}^p([a,b]))$ for some $p \in [2, \infty]$ and $\alpha > 0$ (see e.g. [12] for background on these spaces). An important reason for the choice of this class of functions is the availability of estimates for the error of approximation by *splines*[1], trigonometric polynomials, and wavelet expansions (see e.g. [12] and [5]). In particular, if $\mathcal{S}^k_m$

[1]Piecewise polynomial functions f such that on each compact interval, f is made up of only finitely many polynomial pieces.

denotes the space of piecewise polynomials of degree bounded by k, based on the regular partition of $[a, b]$ with m pieces ($m \geq 1$) and $s \in \mathcal{B}^{\alpha}_{\infty}\left(\mathbb{L}^p([a, b])\right)$ with $k > \alpha - 1$, then there exists a constant $C(s)$ such that

$$\mathrm{d}_p\left(s, \mathcal{S}^k_m\right) \leq C(s) m^{-\alpha}, \tag{8.1}$$

where d_p is the distance induced by the $\mathbb{L}^p$-norm on $([a, b], dx)$ (see [12]). The following give the rate of convergence of p.p.e. on regular splines.

Corollary 8.1 *Following the notation of Theorem 7.1, let* $\tilde{s}_T$ *be the penalized projection estimator on* $\{\mathcal{S}^k_m\}_{m \in \mathcal{M}_T}$ *with penalization*

$$\mathrm{pen}(m) \equiv c\frac{\hat{V}_m}{T} + c'\frac{D_m}{T} + c''\frac{d_m}{T},$$

for some fixed $c > 1$ *and* $c', c'' > 0$*. Then, if the restriction of the Lévy density* s *to* $[a, b]$ *belongs to* $\mathcal{B}^{\alpha}_{\infty}\left(\mathbb{L}^p([a, b])\right)$ *with* $2 \leq p \leq \infty$ *and* $0 < \alpha < k + 1$*, then*

$$\limsup_{T \to \infty} T^{2\alpha/(2\alpha+1)} \mathbb{E}\left[\|s - \tilde{s}_T\|^2\right] < \infty.$$

Moreover, for any $R > 0$ *and* $L > 0$*,*

$$\limsup_{T \to \infty} T^{2\alpha/(2\alpha+1)} \sup_{s \in \Theta(R,L)} \mathbb{E}\left[\|s - \tilde{s}_T\|^2\right] < \infty, \tag{8.2}$$

where $\Theta(R, L)$ *consists of all Lévy densities* s *such that* $\|s\|_{\mathbb{L}^{\infty}([a,b])} < R$*, and the restriction of* s *to* $[a, b]$ *is a member of* $\mathcal{B}^{\alpha}_{\infty}\left(\mathbb{L}^p([a, b])\right)$ *with seminorm* $|s|_{\mathcal{B}^{\alpha}_{\infty}(\mathbb{L}^p)} < L$*.*

The previous result implies that the p.p.e. on regular splines has a rate of convergence of order $T^{-2\alpha/(2\alpha+1)}$ for the class of Besov Lévy densities $\Theta(R, L)$. We will see in the next section that the rate cannot be improved (see Corollary 9.3 and Remark 9.4).

9 On the minimax risk of estimation for smooth Lévy densities

This section presents some results on the *minimax risk* of estimation for the family of smooth Lévy densities introduce in the previous section. In very general terms, the *minimax risk* on a given family Θ of “parameters” has the following general form:

$$\inf_{\hat{s}} \sup_{s\in\Theta} \mathbb{E}_s \left[d\left(s, \hat{s}\right)\right],$$

where the infimum is taken over all the estimators $\hat{s}$ and $d(s, \hat{s})$ is a function that measures how distant s and $\hat{s}$ are from each other. Comparison to the minimax risks is one of the most solicited measures of performance in statistical estimation. Minimax type results have been obtained in very general contexts (see e.g. [14] and [5] for the case of density estimation based on i.i.d. random variables, and [15] and [18] for the case of intensity estimation based on finite Poisson point processes).

Since the jumps of a Lévy process can be associated with a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}\backslash\{0\}$, many results and techniques for the statistical inference of Poisson processes can be translated into the context of Lévy processes. Following this approach, we adapt below a result of [15] (Theorem 6.5) on the asymptotic minimax risk for the estimation of “smooth” intensity functions of a Poisson point processes on $[0, 1]$, based on n independent copies. The idea of the proof is due to Ibragimov and Has'minskii and is based on the statistical tools for distributions satisfying the *Local Asymptotic Normality* (LAN) property (see Chapters II and Section IV.5 of [14]). Some generalizations and consequences are also presented.

Let us introduce a *loss function* $\ell : \mathbb{R} \to \mathbb{R}$ with the following properties:

(i) $\ell(\cdot)$ is nonnegative, $\ell(0) = 0$ but not identically 0, and continuous at 0;

(ii) it is symmetric: $\ell(u) = \ell(-u)$ for all u;

(iii) for any $c > 0$, $\{u : \ell(u) < c\}$ is a convex set;

(iv) $\ell(u) \exp\{\varepsilon|u|^2\} \to 0$ as $|u| \to \infty$, for any $\varepsilon > 0$.

The Lévy densities here considered satisfy a Hölder condition of order β on a given window of estimation. Concretely, fix an interval $[a, b] \subset \mathbb{R}\backslash\{0\}$, and let $k \in \{0, 1, \dots\}$ and $\beta \in (0, 1]$. Define the family $\Theta_{k+\beta}\left(L; [a, b]\right)$ of functions $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ such that f is k times differentiable on $[a, b]$ and

$$|f^{(k)}(x_1) - f^{(k)}(x_2)| \leq L|x_1 - x_2|^{\beta}, \;\; \forall \;\; x_1, x_2 \in [a, b]. \tag{9.1}$$

Below, $\mathcal{L}$ stands for the class of all Lévy densities; that is, all functions $s : \mathbb{R}_0 \to \mathbb{R}_+$ such that

$$\int_{\mathbb{R}_0} \left(x^2 \wedge 1\right) s(x) dx < \infty.$$

The following result gives the long-run behavior of the minimax risk on Θ at a fixed point.

Theorem 9.1 *If x_0 is an interior point of the interval $[a, b] \subset \mathbb{R} \backslash \{0\}$, then*

$$\liminf_{T \to \infty} \left\{ \inf_{\hat{s}_T} \sup_{s \in \Theta} \mathbb{E}_p \left[\ell \left(T^{\alpha/(2\alpha+1)} \left(\hat{s}_{T}(x_0) - s(x_0) \right) \right) \right] \right\} > 0, \tag{9.2}$$

where $\alpha := k + \beta$, $\Theta := \mathcal{L} \cap \Theta_\alpha (L; [a, b])$, and the infimum is over all the estimators $\hat{s}_T$ based on those jumps of the Lévy process $\{X(t)\}_{0 \le t \le T}$ with sizes falling in $[a, b]$.

The previous result can be strengthen to be uniform in $x_0 \in (a, b)$.

Corollary 9.2 *With the notation and hypothesis of Theorem 9.1,*

$$\liminf_{T \to \infty} \left\{ \inf_{\hat{s}_T} \inf_{x \in (a,b)} \sup_{s \in \Theta} \mathbb{E}_s \left[\ell \left(T^{\alpha/(2\alpha+1)} \left(\hat{s}_{T}(x) - s(x) \right) \right) \right] \right\} > 0. \tag{9.3}$$

As a consequence of the previous result, we obtain the long-run behavior of the minimax risk of estimators, under the $\mathbb{L}^2$-norm.

Corollary 9.3 *Let $[a, b]$ be a closed interval of $\mathbb{R} \backslash \{0\}$, then*

$$\liminf_{T \to \infty} T^{2\alpha/(2\alpha+1)} \left\{ \inf_{\hat{s}_T} \sup_{s \in \Theta} \mathbb{E}_s \left[\int_a^b \left(\hat{s}_{T}(x) - s(x) \right)^2 dx \right] \right\} > 0, \tag{9.4}$$

where $\alpha := k + \beta$, $\Theta := \mathcal{L} \cap \Theta_\alpha (L; [a, b])$, and the infimum is over all estimators $\hat{s}_T$ based on the jumps of the Lévy process $\{X(t)\}_{0 \le t \le T}$ with sizes falling on $[a, b]$.

Remark 9.4 *The proofs of the previous results can be readily modified to cover smaller classes of Lévy densities* Θ *such as*

$$\Theta = \mathcal{L} \cap \Theta_\alpha(L; [a, b]) \cap \{s : \|s\|_{\mathbb{L}^\infty([a,b])} < R\}.$$

The above class enjoys a very close relationship with the family of Besov densities $\Theta(R, L)$ *introduced in (8.2). Indeed, the class* $\Theta_\alpha(L; [a, b])$ *is contained in* $\mathcal{B}^\alpha_\infty(\mathbb{L}^\infty([a, b]))$ *(see Section 2.9 of [12]). Since* $\mathcal{B}^\alpha_\infty(\mathbb{L}^\infty) \subset \mathcal{B}^\alpha_\infty(\mathbb{L}^p)$*, (9.4) holds true on* $\Theta = \Theta(R, L)$*. Therefore, the p.p.e.'s on regular splines considered in the previous section, achieve the minimax rate of convergence on* $\Theta(R, L)$*. This type of property is called* adaptivity *in that, without knowing the smoothness of* s *(controlled by* α*), the p.p.e.'s reach asymptotically the minimax risk up to a constant. See for instance Section 4 of [5] for a discussion on adaptivity.*

10 Concluding Remarks

- In the present paper we have reviewed a new methodology for the estimation of the Lévy density of a Lévy process. Our methods are quite flexible in the sense that different type of estimating functions can be used; for instance, histograms, splines, trigonometric polynomials, and wavelets. The estimation is model free, easily implementable, and suitable for "high-frequency" data. Based on continuous-time data, the procedures enjoy good asymptotic properties. *Oracle inequalities* imply that, up to a constant, the procedure will achieve the best possible risk among the *projection estimators*. Moreover, it is proved that penalized projection estimators on splines achieve the optimal rate of convergence, from the minimax point of view, on some classes of smooth Lévy densities. Simulations show good results in Lévy models with infinite jump activity such as the variance Gamma model.

- Generalization to some multivariate Lévy models can be readily obtained, since the results behind the construction have multivariate versions. Indeed, the Lévy-Itô decomposition of the sample paths, the concentration inequalities for compensated Poisson integrals, the inference theory for Locally Asymptotically Normal distributions, and the short-term properties of the marginal distributions are valid in the multivariate setting. More precisely, consider a Lévy process $\mathbf{X} = \{\mathbf{X}(t)\}_{t\geq 0}$ on $\mathbb{R}^d$ with Lévy measure ν. Assume that, in a window of estimation $D \in \mathcal{B}(\mathbb{R}^d\backslash\{0\})$, ν is absolutely continuous with respect to a reference measure η and that $s \equiv d\nu/d\eta$ is bounded with also $\int_D s^2(\mathbf{x})\eta(d\mathbf{x}) < \infty$. Then, given a finite-dimensional subspace $\mathcal{S}$ of $L^2\left((D,\eta)\right)$, the projection estimator of s on $\mathcal{S}$ is defined as in Section 2 with $\mathcal{J}$ being the Poisson measure on $\mathbb{R}_+ \times \mathbb{R}^d$ associated with the jumps of $\mathbf{X}$. Similarly, penalized projection estimators can be constructed, and the risk bound of Theorem 7.1, along with the Oracle inequality (7.6), are satisfied. The results of Sections 4 and 5 are valid as well. However, let us point out that further specifications of our methods for some semiparametric models are desirable. Important examples of these models include multivariate stable processes, and the tempered stable Lévy processes, recently introduced in [20].

- The methods here focus on the estimation of the jump part of the Lévy process. It is natural to address the problem of estimating the continuous part too. In the one-dimensional case, this part is of the form $bt + \sigma W(t)$, where $\{W(t)\}_{t\geq 0}$ is a standard Brownian motion. In the multivariate case, it is characterized by a vector $\mathbf{b}$ and a symmetric nonnegative-definite matrix Σ. Several approaches deal with the estimation of Σ, from moment based methods to methods based on high-frequency data. A simple one is to use the following functional limit theorem:

$$\left\{\frac{1}{\sqrt{h}}\mathbf{X}(ht)\right\}_{t\geq 0} \stackrel{\mathfrak{D}}{\rightarrow} \{\mathbf{Y}(t)\}_{t\geq 0}, \quad h \to 0,$$

  where $\{\mathbf{Y}(t)\}_{t\geq 0}$ is a centered Gaussian Lévy process with variance-covariance matrix Σ. This result can be deduced from the proof of the uniqueness of the Lévy-Khintchine representation as in pp. 40 of [22]. Another simple method will be to consider empirical versions of the moments:

$$\mathbb{E}\left[(X_j(t) - X_j(t))(X_k(t) - X_k(t))\right] = t\left(\Sigma_{j,k} + \int_{\mathbb{R}_0^d} x_j x_k \nu(d\mathbf{x})\right),$$

  provided that $\int_{\|\mathbf{x}\|>1} \|\mathbf{x}\|^2 \nu(d\mathbf{x}) < \infty$ (see Section 25 [22]). Here, $X_j(t)$ and x_j refer to the j^{th} component of the vectors $\mathbf{X}(t)$ and $\mathbf{x}$, respectively, while $\Sigma_{j,k}$ is the (j,k) entry of Σ. The second term on the left hand side of the above expression can be estimated using our estimators for ν. In the one-dimensional case, another approach is to use "threshold estimators" of the form:

$$\sum_{k=1}^{n} (\Delta_k X)^2 \mathbf{1}\left((\Delta_k X)^2 \leq r(h)\right),$$

where $\Delta_k X \equiv X(t_k^n) - X(t_{k-1}^n)$ is the k^{th} increment of the process and $r(h)$ is an appropriate cutoff function (see [16] for details). For a class of semimartingales with finite jump activity, [3] provides another methodology based on the *bipower variation* (see also [4]). In the case of Lévy processes with finite jump activity, [1] disentangles the diffusion from the jumps using maximum likelihood and the *Generalized Method of Moments*. On the other hand, the estimation of the parameter $\mathbf{b}$ can be done by yet different methods. For instance, using the empirical version for

$$\mathbb{E}\left[\mathbf{X}(t)\right] = t\left(\mathbf{b} + \int_{\|\mathbf{x}\|>1} \mathbf{x}\nu(d\mathbf{x})\right),$$

valid if $\int_{\|\mathbf{x}\|>1} \|\mathbf{x}\|\nu(d\mathbf{x}) < \infty$. Another approach will be to estimate the "drift" $\mathbf{b}_0 \equiv \mathbf{b} - \int_{\|\mathbf{x}\|\leq 1} \mathbf{x}\nu(d\mathbf{x})$ (where the integration is componentwise) using the fact that

$$\mathbb{P}\left[\lim_{h\to 0} \frac{1}{h}\mathbf{X}(h) = \mathbf{b}_0\right] = 1.$$

The above result holds true if $\int_{\|\mathbf{x}\|\leq 1} \|\mathbf{x}\|\nu(d\mathbf{x}) < \infty$ and $\Sigma = 0$ (see [22]). Even though our methods are valid for non necessarily pure-jump Lévy processes, it is expected that the presence of a diffusion component will reduce the efficiency (in terms of speed of convergence and accuracy) of our estimators. It would be interesting to study in greater detail this phenomenon.

11 Figures

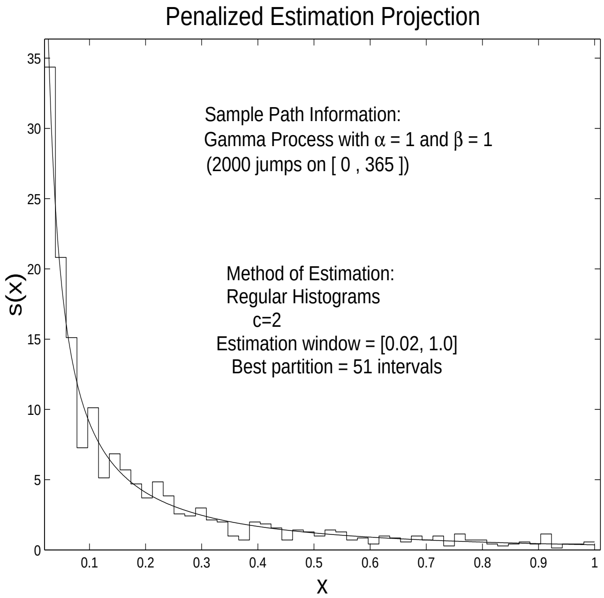

Figure 1: Penalized projection estimation of $\frac{e^{-x}}{x}$.

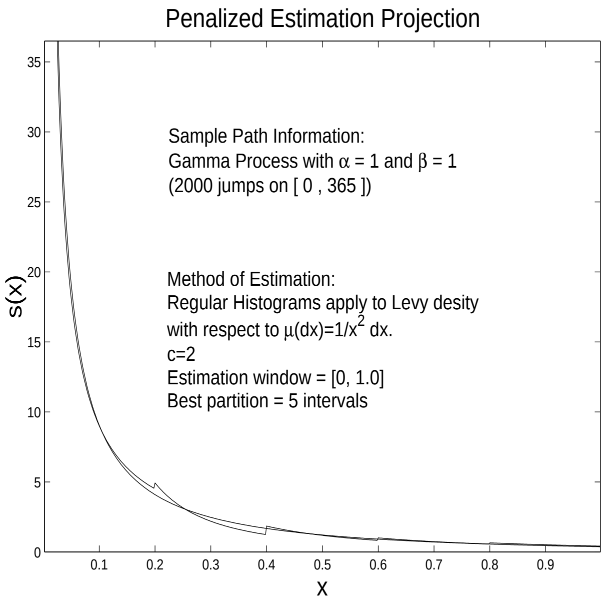

Figure 2: Regularized penalized projection estimation of $\frac{e^{-x}}{x}$.

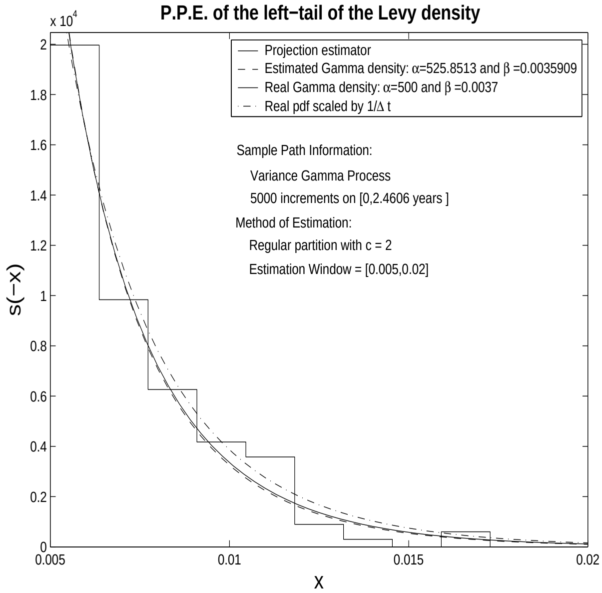

Figure 3: Penalized projection estimation of the left-tail of the variance Gamma Lévy density.

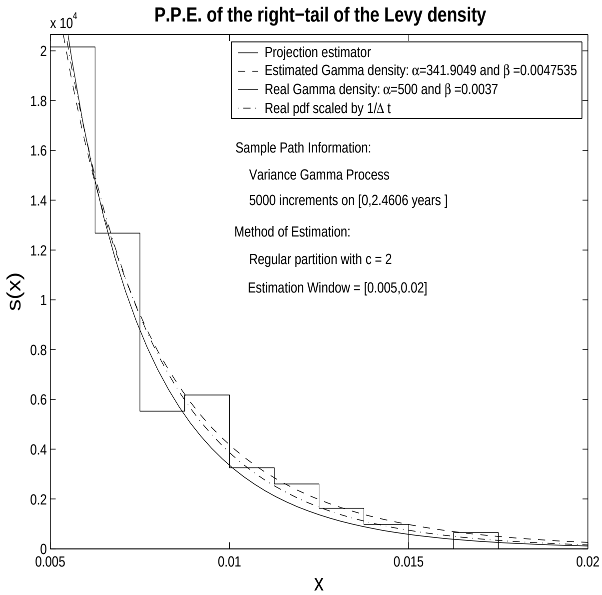

Figure 4: Penalized projection estimation of the right-tail of the variance Gamma L'evy density.

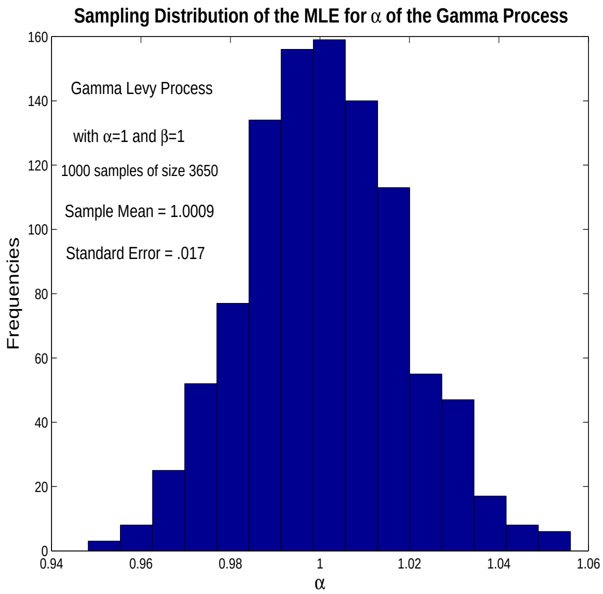

Figure 5: Sampling Distribution for the MLE of the α of the Gamma Lévy Process.

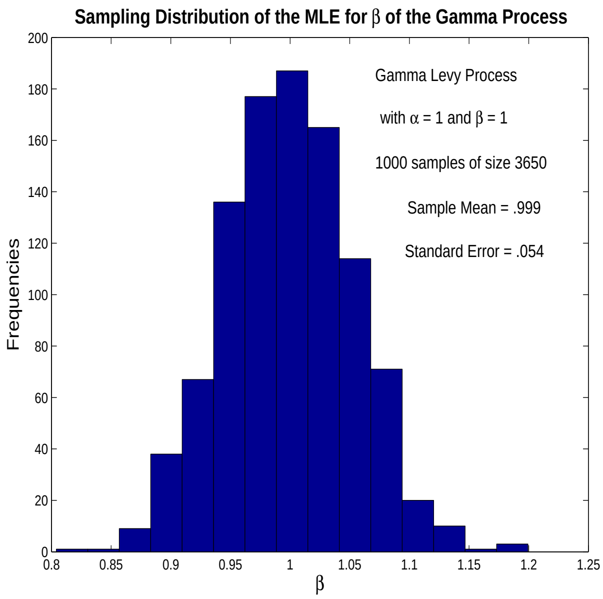

Figure 6: Sampling Distribution for the MLE of the β of the Gamma Lévy Process.

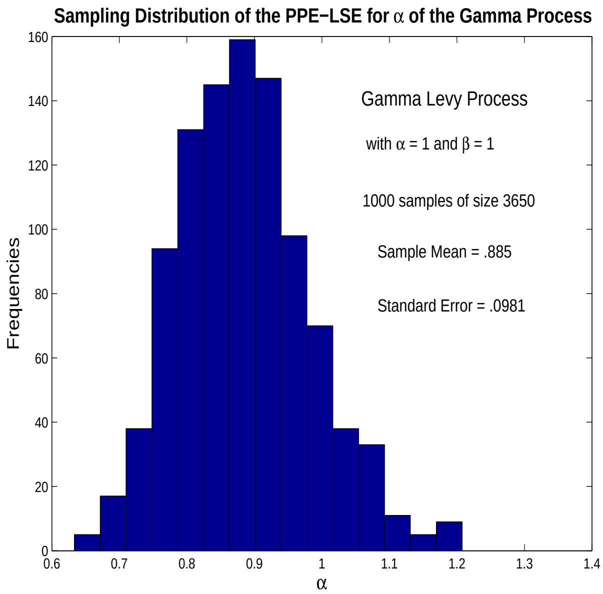

Figure 7: Sampling Distribution for the Estimates of the α of a Gamma Lévy process obtained from the PPE and the LSE method.

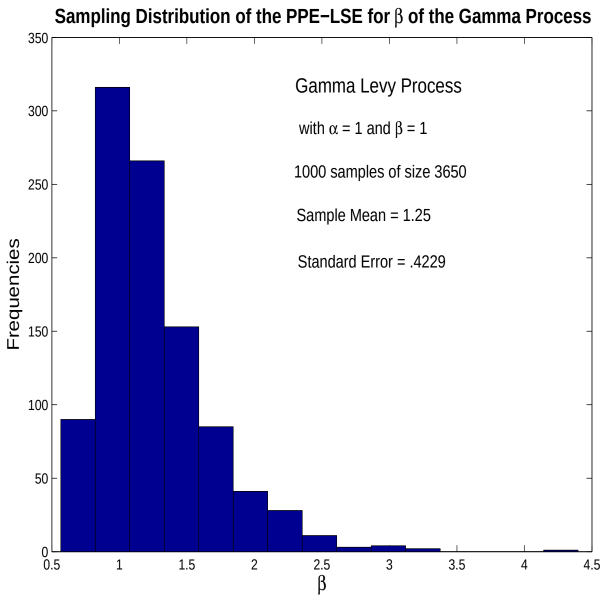

Figure 8: Sampling Distribution for the Estimates of the β of a Gamma Lévy process obtained from the PPE and the LSE method.

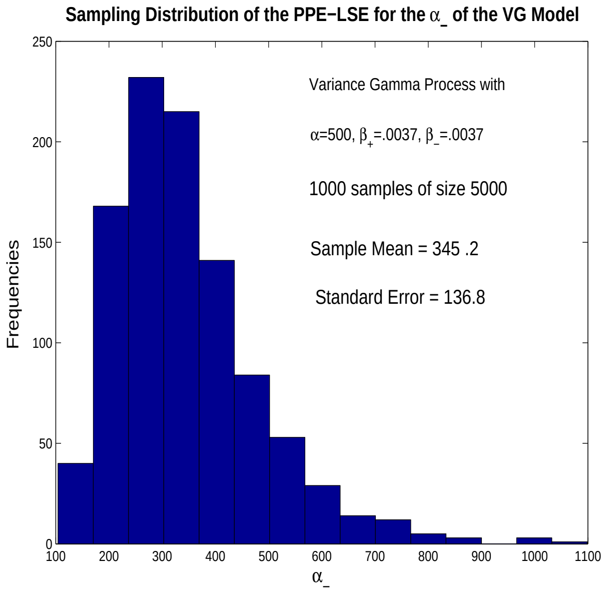

Figure 9: Sampling Distribution for the Estimates of α_- obtained from the PPE and the LSE method.

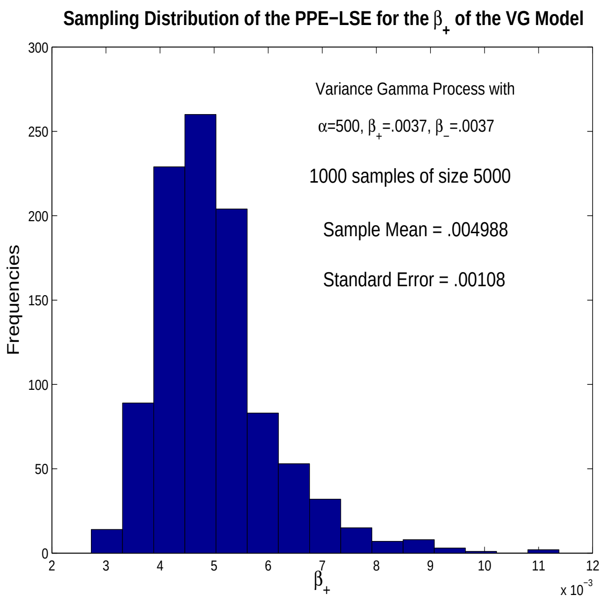

Figure 10: Sampling Distribution for the Estimates of β_+ obtained from the PPE and the LSE method.

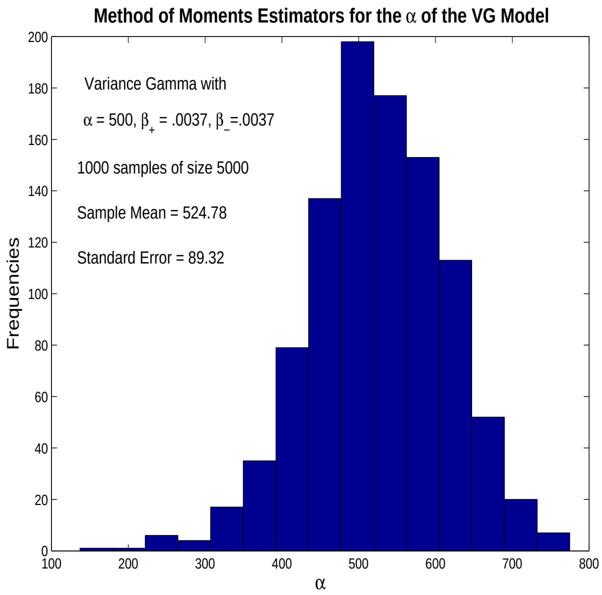

Figure 11: Sampling Distribution for the Estimator of α obtained by the Method of Moments.

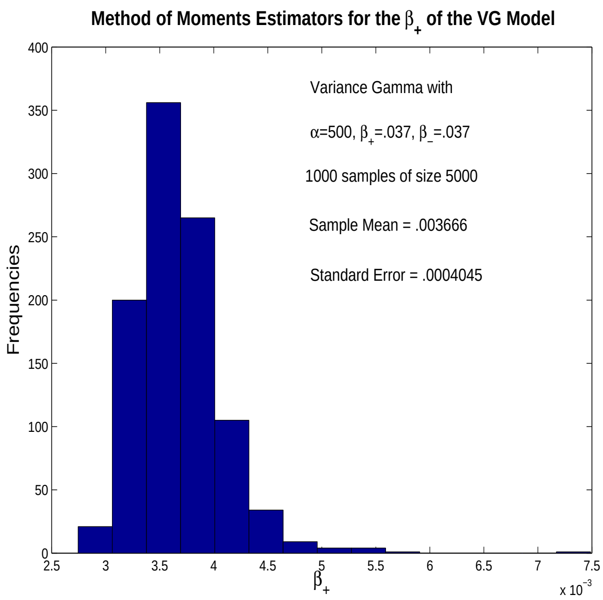

Figure 12: Sampling Distribution for the Estimator of β_+ obtained by the Method of Moments.

References

[1] Y. Aït-Sahalia. Disentangling difussion from jumps. *Journal of Financial Economics*, July 2004.

[2] O.E. Barndorff-Nielsen and N. Shephard. Modelling by Lévy processess for financial economics. *Lévy Processes. Theory and Applications, by O.E. Barndorff-Nielsen, T. Mikosch, and S.I. Resnick*, pages 283–318, 2001.

[3] O.E. Barndorff-Nielsen and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. Technical Report W21, Economics Group, Nuffield College, University of Oxford, 2003.

[4] O.E. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2:1–48, 2004.

[5] A.R. Barron, L. Birgé, and P. Massart. Model selection via penalization. Technical Report 54, Université Paris-Sud, 1995.

[6] J. Bertoin. *Lévy processes.* Cambridge University Press, 1996.

[7] L. Birgé and P. Massart. Minimum contrast estimation on Sieves. Technical Report 34, Université Paris-Sud, 1994.

[8] L. Birgé and P. Massart. From model selection to adaptive estimation. *In Festschrift for Lucien Le Cam*, pages 55–87, 1997.

[9] P. Carr, H. Geman, D. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, April 2002.

[10] P. Carr, D. Madan, and E. Chang. The variance Gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[11] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[12] R.A. DeVore and G.G. Lorentz. *Constructive Approximation.* Springer-Verlag, 1993.

[13] E. Figueroa-Lopez and C. Houdré. Nonparametric estimation of Lévy processes with a view towards mathematical finance. November 2004. Available at [ArXiv math.ST/0412351].

[14] I.A. Ibraginov and R.Z. Has'minskii. *Statistical estimation. Asymptotic Theory.* Springer-Verlag,Berlin, New York, Heidelberg, 1981.

[15] Y.A. Kutoyants. *Statistical Inference for Spatial Poisson Processes.* Springer, 1998.

[16] C. Mancini. Estimating the integrated volatility in stochastic volatility models with Lévy type jumps. Technical report, Dipartimento di Matematica per le Decisioni, Universita di Firenze, July 2004. Presented at the Bachelier Finance Society, 3rd World Congress.

[17] K. Prause. *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures.* PhD thesis, University of Freiburg, October 1999.

[18] P. Reynaud-Bouret. Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. *Probab. Theory Related Fields*, 126(1):103–153, 2003.

[19] J. Rosiński. Series representations of Lévy processes from the perspective of point processes. *In Lévy processes-Theory and Applications*, pages 401–415, 2001.

[20] J. Rosiński. Tempered stable processes. Technical report, University of Tennessee, 2002.

[21] L. Rüschendorf and J. Woerner. Expansion of transition distributions of Lévy processes in small time. *Bernoulli*, 8:81–96, 2002.

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