ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 76, No. 4 (July, 2008), 727–761

FISHER'S INFORMATION FOR DISCRETELY SAMPLED LÉVY PROCESSES

YACINE AÏT-SAHALIA
Bendheim Center for Finance, Princeton University, Princeton, NJ 08540-5296, U.S.A. and NBER

JEAN JACOD
Institut de Mathématiques de Jussieu, 75013 Paris, France and Université P. et M. Curie, Paris-6, France

Econometrica, Vol. 76, No. 4 (July, 2008), 727–761

FISHER'S INFORMATION FOR DISCRETELY SAMPLED LÉVY PROCESSES

BY YACINE AÏT-SAHALIA AND JEAN JACOD[1]

This paper studies the asymptotic behavior of Fisher's information for a Lévy process discretely sampled at an increasing frequency. As a result, we derive the optimal rates of convergence of efficient estimators of the different parameters of the process and show that the rates are often nonstandard and differ across parameters. We also show that it is possible to distinguish the continuous part of the process from its jumps part, and even different types of jumps from one another.

KEYWORDS: Lévy process, jumps, rate of convergence, optimal estimation.

1. INTRODUCTION

CONTINUOUS TIME MODELS in finance often take the form of a semimartingale, because this is essentially the class of processes that precludes arbitrage. Among semimartingales, models allowing for jumps are becoming increasingly common. There is a relative consensus that, empirically, jumps are prevalent in financial data, especially if one incorporates not just large and infrequent Poisson-type jumps, but also infinite activity jumps that can generate smaller, more frequent, jumps. This evidence comes from direct observation of the time series of various asset prices, studying the distribution of primary asset and that of derivative returns, especially as high frequency data becomes more commonly available for a growing range of assets.

On the theory side, the presence of jumps changes many important properties of financial models. This is the case for option and derivative pricing, where most types of jumps will render markets incomplete, so the standard arbitrage-based delta hedging construction inherited from Brownian-driven models fails, and pricing must often resort to utility-based arguments or perhaps bounds. The implications of jumps are also salient for risk management, where jumps have the potential to alter significantly the tails of the distribution of asset returns. In the context of optimal portfolio allocation, jumps can generate the types of correlation patterns across markets, or contagion, that is often documented in times of financial crises. Even when jumps are not easily observable or even present in a given time series, the mere possibility that they could occur will translate into amended risk premia and asset prices: this is one aspect of the peso problem.

These theoretical differences between models driven by Brownian motions and those driven by processes that can jump are fairly well understood. How-

[1]Financial support from the NSF under Grants SBR-0350772 and DMS-0532370 is gratefully acknowledged. We are also grateful for comments from the editor, three anonymous referees, George Tauchen, and seminar and conference participants at Princeton University, The University of Chicago, MIT, Harvard, and the 2006 Econometric Society Winter Meetings.

ever, comparatively little is known about the estimation of models driven by jumps, even in the special case where the jump process has independent and identically distributed increments, that is, is a Lévy process (see, e.g., Woerner (2004) and the references cited therein), and among different inference procedures for parametrized Lévy processes, even less is known about their relative efficiency, especially for high frequency data.

In this paper, we take a few steps toward a better understanding of the econometric situation, in a parametric model that is far from being fully realistic, but has the advantage of being tractable and yet is rich enough to deliver some surprising results. In particular, we will see that jumps give rise to a nonstandard econometric situation where the rates of convergence of estimators of the various parameters of the model can depart from the usual $\sqrt{n}$ in a number of unexpected ways: another comparable, although unrelated, situation happens as one either stays on or crosses the unit root barrier in classical time series analysis.

We study a class of Lévy processes for a log-asset price X, indexed on three parameters. We split X into the sum of two independent Lévy processes as

$$X_t = \sigma W_t + \theta Y_t. \tag{1.1}$$

Here, we have $\sigma > 0$ and $\theta \in \mathbb{R}$, and W is a standard symmetric stable process with index $\beta \in (0, 2]$. We are often interested in the classical situation where $\beta = 2$ and so W is a Wiener process, hence continuous; σ and θ are scale parameters. In the case where W is a Wiener process, σ is the (Brownian or continuous) volatility of the process. As for Y, it is another Lévy process, viewed as a perturbation of W and "dominated" by W in a sense we will make precise below. In some applications, Y may represent frictions that are due to the mechanics of the trading process or, in the case of compound Poisson jumps, it may represent the infrequent arrival of relevant information related to the asset. In the latter case, W is then the driving process for the ordinary fluctuations of the asset value. Y is independent of W, and its law is either known or a nuisance parameter.

When $\beta = 2$, W has, as already mentioned, continuous paths, and we will then be working with a model that has both continuous and discontinuous components. But when $\beta < 2$, W is itself a jump process, and we will be able to study the effect on each other's parameters of including two separate jump components in the model, with different relative levels of jump activity. Our results cover both cases and we will describe the relevant differences.

While there is a large literature devoted to stable processes in finance (starting with the work of Mandelbrot in the 1960s and see, for example, Rachev and Mittnik (2000) for new developments), a model based on stable processes is arguably quite special. Rather than an attempt at a fully realistic model intended to be fitted to the data as is, we have chosen to capture only some of the salient features of financial data, such as the presence of jumps, in a reasonably tractable framework. Inevitably, we leave out some other important features such as stochastic volatility[2] or market microstructure noise, but in exchange we are able to derive explicitly some new and surprising econometric results in a fully parametric framework.

The log-asset price X is observed at n discrete instants separated by Δ_n units of time. In financial applications, we are often interested in the high frequency limit where $\Delta_n \to 0$. The length $T_n = n\Delta_n$ of the overall observation period may be a constant T or may go to infinity. With Y viewed as a perturbation of W, we studied in an earlier paper (Aït-Sahalia and Jacod (2007)) the estimation of the single parameter σ, showing in particular that it can be estimated with the same degree of accuracy as when the process Y is absent, at least asymptotically. We proposed estimators for σ designed to achieve the efficient rate despite the presence of Y.[3]

We now study the more general problem where the parameter vector of interest is either (σ, θ) or the full (σ, β, θ), viewing the law of Y nonparametrically. Our objective is to determine the optimal rate at which these parameters can be estimated. So we are led to consider the behavior of Fisher's information. Through the Cramer–Rao lower bound, this yields the rate at which an optimal sequence of estimators should converge and the lowest possible asymptotic variance.

Unlike the situation where Δ_n is fixed or the estimation of σ we previously studied, we are now in a nonstandard situation where the rates of convergence of optimal estimators of (β, θ) can depart from the usual $\sqrt{n}$ in a number of different and often unexpected ways. Even for such a simple class of models as (1.1), there is a wide range of different asymptotic behaviors for Fisher's information if we wish to estimate the three parameters σ, θ, and β or even σ and θ only when β is known, for example, when $\beta = 2$. We will show that different rates of convergence are achieved for the different parameters *and* for different types of Lévy processes.

The optimal rate for σ is $\sqrt{n}$ as seen in Aït-Sahalia and Jacod (2007). Now, for β, the optimal rate will be the faster $\sqrt{n}\log(1/\Delta_n)$, and is unaffected by the presence of Y (as was the case for σ). To estimate θ, the optimal rate is not $\sqrt{n}$ and, unlike the one for β, depends heavily on the specific nature of the Y process. Furthermore, unlike what happens for σ or β, which are immune to the presence of Y, the presence of the other process (W in this case) strongly affects the rate for θ. We study different examples of that situation; one in particular consists of the case where Y is also a stable process with index $\alpha < \beta$; another is the jump-diffusion situation where Y is a compound Poisson process.

[2]There has been much activity devoted to estimating the integrated volatility of the process in that context, with or without jumps: see, for example, Andersen, Bollerslev, and Diebold (2003), Barndorff-Nielsen and Shephard (2004), and Mykland and Zhang (2006).

[3]Another string of the recent literature focuses on disentangling the volatility from the jumps: see, for example, Aït-Sahalia (2004) and Mancini (2001). In our case, this is related to the estimation of σ when $\beta = 2$.

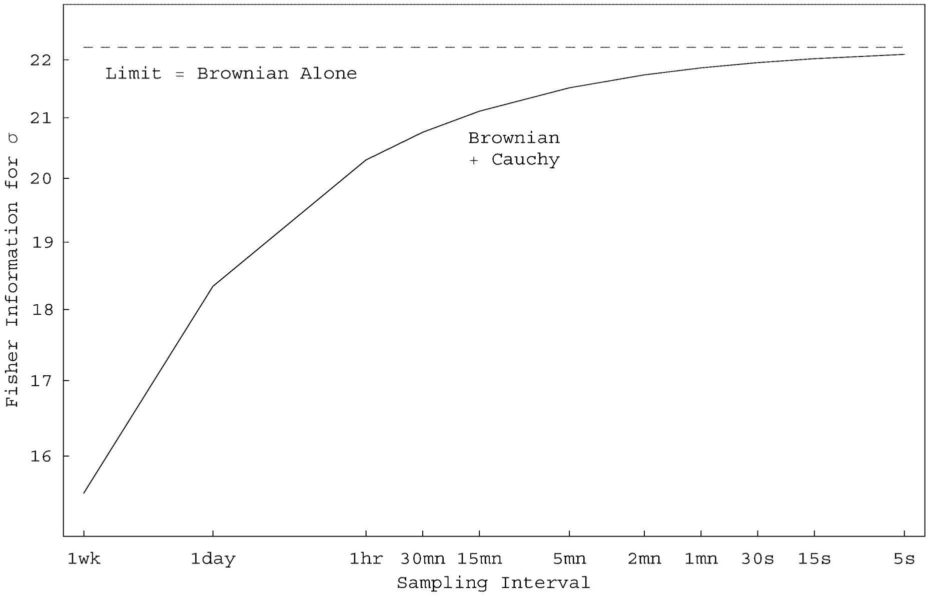

FIGURE 1.—Convergence as $\Delta \to 0$ of Fisher's information for the parameter σ from the model $X = \sigma W + \theta Y$, where W is Brownian motion ($\beta = 2$) and Y is a Cauchy process ($\alpha = 1$), to Fisher's information from the model $X = \sigma W$, which is $2/\sigma^2$ as we will show in Theorem 2. Fisher's information for the model $X = \sigma W + \theta Y$ (solid line) is computed numerically using the exact expressions for the densities of the two processes and numerical integration of their convolution. The figure is in log–log scale. The parameter values are $\sigma = 30\%$/year and $\theta = 0.3$.

To illustrate some of these effects numerically, consider the example where W is a Brownian motion ($\beta = 2$) and Y is a Cauchy process ($\alpha = 1$). We plot in Figures 1 and 2 the convergence, as the sampling interval decreases, of the (numerically computed) Fisher's information for σ and θ, respectively, to their limits derived in Theorems 2 and 7, respectively. While, as illustrated in Figure 1, we can estimate σ as well as if Y were absent when $\Delta_n \to 0$, the symmetric statement is not true for θ. The limiting behavior in Figure 2 is quite distinct from a convergence to Fisher's information for θ in the model without W, that is, $X = \theta Y$. The latter is simply $1/(2\theta^2)$. In fact, Fisher's information about θ in the presence of W tends to zero when $\Delta_n \to 0$. Consequently, any optimal estimator of θ will converge at a rate slower than $\sqrt{n}$.

The paper is organized as follows. In Section 2, we set up the model. In Section 3, we derive the asymptotic behavior of Fisher's information about the parameters (σ, β, θ). Then we show in Section 4 that the optimal rate for θ varies substantially according to the structure of the process Y, and we illustrate the versatility of the situation through several examples that display the variety of convergence rates that arise. Section 5 concludes. Proofs are in the Appendix.

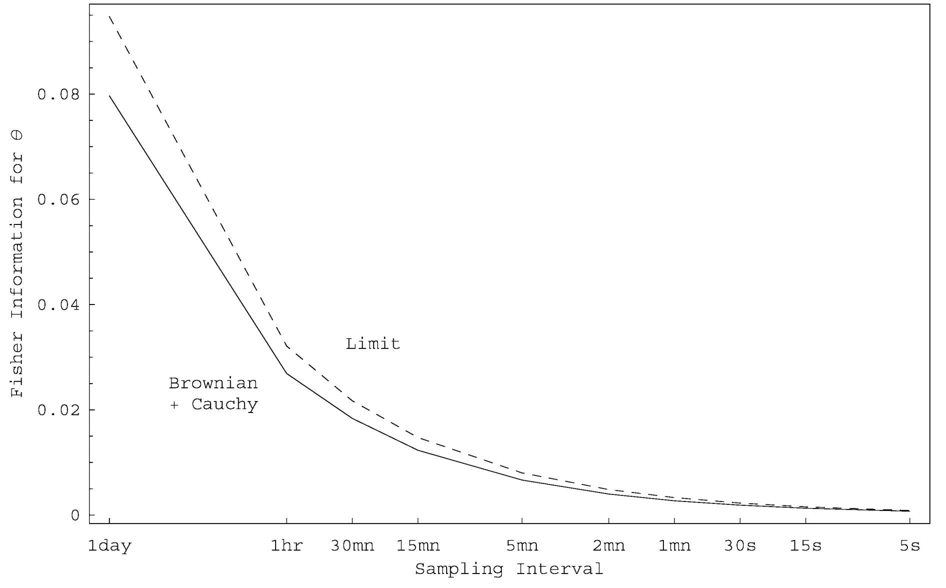

FIGURE 2.—Convergence as $\Delta \to 0$ of Fisher's information for the parameter θ from the model $X = \sigma W + \theta Y$, where W is a Brownian motion ($\beta = 2$) and Y is a Cauchy process ($\alpha = 1$), to the theoretical limiting curve $\Delta \mapsto (\pi\theta\sigma)^{-1}\Delta^{1/2}[\log(1/\Delta)]^{-1/2}$ derived in Theorem 7. Fisher's information for the model $X = \sigma W + \theta Y$ (solid line) is computed numerically using the exact expressions for the densities of the two processes, and numerical integration of their convolution. The x-axis is in log scale. The parameter values are identical to those of Figure 1.

2. THE MODEL

Because the components of X in (1.1) are all Lévy processes, they have an explicit characteristic function. Consider first the component W. The characteristic function of W_t is

$$\mathbb{E}(e^{iuW_t}) = e^{-t|u|^\beta/2}. \tag{2.1}$$

The factor 2 above is perhaps unusual for stable processes when $\beta < 2$, but we put it here to ensure continuity between the stable and the Gaussian cases.

The essential difficulty in this class of problems is the fact that the density of W_Δ is not known in closed form except in special cases ($\beta = 1$ or 2).[4] Without the density of Y_Δ in closed form, it is of course hopeless to obtain the density of X_Δ, the convolution of the densities of σW_Δ and θY_Δ, in closed form. But when Δ goes to 0, the density of X_Δ and its derivatives degenerate in a way

[4]For fixed Δ, there exist representations in terms of special functions (see Zolotarev (1995) and Hoffmann-Jørgensen (1993)), whereas numerical approximations based on approximations of the densities may be found in DuMouchel (1971, 1973b, 1975) or Nolan (1997, 2001), or also Brockwell and Brown (1980).

which is controlled by the characteristics of the Lévy process, thus allowing us to explicitly describe, in *closed form*, the limiting behavior of Fisher's information.

To understand this behavior, we will need to be able to bound various functionals of the density of X_Δ. As is well known, when $\beta < 2$, we have $\mathbb{E}(|W_t|^\rho) < \infty$ if and only if $0 < \rho < \beta$. The density h_β of W_1 is C^∞ (by repeated integration of the characteristic function) for all $\beta \leq 2$, even, and its nth derivative $h_\beta^{(n)}$ behaves, as $|w| \to \infty$, as

$$\text{(2.2)} \qquad \left|h_\beta^{(n)}(w)\right| \sim \begin{cases} \dfrac{c_\beta}{|w|^{n+1+\beta}} \displaystyle\prod_{i=0}^{n-1}(\beta+i+1), & \text{if } \beta < 2, \\ |w|^n e^{-w^2/2}/\sqrt{2\pi}, & \text{if } \beta = 2 \end{cases}$$

(in the case $n = 0$, an empty product is defined to be 1), where c_β is the constant given by

$$\text{(2.3)} \qquad c_\beta = \begin{cases} \dfrac{\beta(1-\beta)}{4\Gamma(2-\beta)\cos(\beta\pi/2)}, & \text{if } \beta \neq 1, \beta < 2, \\ \dfrac{1}{2\pi}, & \text{if } \beta = 1. \end{cases}$$

This follows from the series expansion of the density due to Bergstrøm (1952), the duality property of the stable densities of order β and $1/\beta$ (see, e.g., Chapter 2 in Zolotarev (1986)), with an adjustment factor in c_β to reflect our definition of the characteristic function in (2.1). In addition, the cumulative distribution function of W_1 for $\beta < 2$ behaves according to $\int_w^\infty h_\beta(x)\,dx \sim c_\beta/(\beta w^\beta)$ when $w \to \infty$ (and symmetrically as $w \to -\infty$).

The second component of the model is Y. Its law as a process is entirely specified by the law G_Δ of the variable Y_Δ for any given $\Delta > 0$. We write $G = G_1$, and we recall that the characteristic function of G_Δ is given by the Lévy–Khintchine formula

$$\text{(2.4)} \qquad \mathbb{E}(e^{ivY_\Delta}) = \exp\Delta\left(ivb - \frac{cv^2}{2} + \int F(dx)\left(e^{ivx} - 1 - ivx1_{\{|x|\leq 1\}}\right)\right),$$

where (b, c, F) is the "characteristic triple" of G (or, of Y): $b \in \mathbb{R}$ is the drift of Y, and $c \geq 0$ is the local variance of the continuous part of Y, and F is the Lévy jump measure of Y, which satisfies $\int(1 \wedge x^2)F(dx) < \infty$ (see, e.g., Chapter II.2 in Jacod and Shiryaev (2003)).

We need to put some additional structure on the model. We define the "domination" of Y by W by the property that G belongs to one of the classes defined below, for some $\alpha \leq \beta$. Let first Φ be the class of all nonnegative continuous functions on $[0,1]$ with $\phi(0)=0$. If $\phi \in \Phi$, we set

$$\mathcal{G}(\phi,\alpha) = \text{the set of all infinitely divisible distributions with } c=0 \tag{2.5}$$

and

$$\forall x \in (0,1], \quad \begin{cases} x^\alpha F([-x,x]^c) \le \phi(x), & \text{if } \alpha < 2, \\ x^2 F([-x,x]^c) \le \phi(x) \quad \text{and} & \\ \displaystyle\int_{\{|y|\le x\}} |y|^2 F(dy) \le \phi(x), & \text{if } \alpha = 2, \end{cases} \tag{2.6}$$

$$\mathcal{G}_\alpha = \bigcup_{\phi\in\Phi} \mathcal{G}(\phi,\alpha).$$

We then have

$$\begin{cases} \alpha \in (0,2] \quad \Rightarrow \\ \quad \mathcal{G}_\alpha = \Big\{G \text{ is infinitely divisible}, c=0, \lim_{x\downarrow 0} x^\alpha F([-x,x]^c) = 0\Big\}, \\ \alpha = 2 \quad \Rightarrow \quad \mathcal{G}_2 = \{G \text{ is infinitely divisible}, c=0\}. \end{cases} \tag{2.7}$$

For example, if Y is a compound Poisson process or a gamma process, then G is in $\bigcap_{\in(0,2]} \mathcal{G}_\alpha$. If Y is an α-stable process, then G is in $\bigcap_{\alpha'>\alpha} \mathcal{G}_{\alpha'}$, but not in $\mathcal{G}_\alpha$. Recall also that the Blumenthal–Getoor index $\gamma(G)$ of Y (or G or F) is the infimum of all γ such that $\int_{\{|x|\le 1\}} |x|^\gamma F(dx) < \infty$, equivalently, $\sum_{s\le t} |Y_s - Y_{s-}|^\gamma$ is almost surely finite for all t. If $\mathcal{H}_\alpha$ denotes the set of all G such that $c=0$ and $\gamma(G) \le \alpha$, one may show that $\mathcal{G}_\alpha \subset \mathcal{H}_\alpha \subset \bigcap_{\alpha'>\alpha} \mathcal{G}_{\alpha'}$. Hence saying that $G \in \mathcal{G}_\alpha$ is a little bit more restrictive than saying that $G \in \mathcal{H}_\alpha$.

Our purpose is to understand the properties of the best possible estimators of the parameters of the model. The variables under consideration in the model (1.1) have densities which depend smoothly on the parameters, so in light of the Cramer–Rao lower bound, Fisher's information is an appropriate tool for studying the optimality of estimators. X is observed at n times $\Delta_n, 2\Delta_n, \dots, n\Delta_n$. Recalling that $X_0 = 0$, this amounts to observing the n increments $X_{i\Delta_n} - X_{(i-1)\Delta_n}$. So when $\Delta_n = \Delta$ is fixed, we observe n independent and identically distributed variables distributed as X_Δ. If, further, these variables have a density depending smoothly on a parameter η, a very weak assumption, we are on known grounds: the Fisher's information at stage n has the form $I_{n,\Delta}(\eta) = nI_\Delta(\eta)$, where $I_\Delta(\eta) > 0$ is the Fisher's information (an invertible matrix if η is multidimensional) of the model based on the observation of the single variable $X_\Delta - X_0$; we have the locally asymptotically normal property; the asymptotically efficient estimators $\widehat{\eta}_n$ are those for which $\sqrt{n}(\widehat{\eta}_n - \eta)$ converges in law to the normal distribution $\mathcal{N}(0, I_\Delta(\eta)^{-1})$, and the maximum likelihood estimator (MLE) does the job (see, e.g., DuMouchel (1973a)).

For high frequency data, though, things become more complicated since the time lag Δ_n becomes small. The corresponding asymptotics require $\Delta_n \to 0$ as the number n of observations goes to infinity. Fisher's information at stage n still has the form $I_{n,\Delta_n}(\eta) = nI_{\Delta_n}(\eta)$, but the asymptotic behavior of the information $I_{\Delta_n}(\eta)$, not to speak about the behavior of the MLE for example, is now far from obvious: the essential difficulty comes from the fact that the law of X_{Δ_n} becomes degenerate as $n \to \infty$.

In the model (1.1), the law of the observed process X depends on the three parameters (σ, β, θ) to be estimated, plus on the law of Y which is summarized by G. The law of the variable X_Δ has a density which depends smoothly on σ and θ, so that the 2×2 Fisher's information matrix (relative to σ and θ) of our experiment exists; it also depends smoothly on β when $\beta < 2$, so in this case the 3×3 Fisher's information matrix exists. In all cases we denote the information by $I_{n,\Delta_n}(\sigma, \beta, \theta, G)$, and it has the form $I_{n,\Delta_n}(\sigma, \beta, \theta, G) = nI_{\Delta_n}(\sigma, \beta, \theta, G)$, where $I_\Delta(\sigma, \beta, \theta, G)$ is the Fisher's information matrix associated with the observation of a single variable X_Δ. We denote the elements of the matrix $I_\Delta(\sigma, \beta, \theta, G)$ as $I_\Delta^{\sigma\sigma}(\sigma, \beta, \theta, G)$, $I_\Delta^{\sigma\beta}(\sigma, \beta, \theta, G)$, etcetera. G may appear as a nuisance parameter in the model, in which case we may wish to have estimates for the Fisher's information that are *uniform* in G, at least on some reasonable class of G's. Let us also mention that in many cases the parameter β is indeed known: this is particularly true when W is a Brownian motion, $\beta = 2$.

3. THE GENERAL CASE

We are first interested in determining the optimal rates (and constants) at which one can estimate (σ, β), while leaving the distribution $G \in \mathcal{G}_\beta$ as unspecified as possible. As we will see, the optimal rate for the estimation of θ is heavily dependent on the precise nature of G.

3.1. *Inference on* (σ, β)

It turns out that the limiting behavior of Fisher's information for (σ, β) is given by the corresponding limits in the situation where $Y = 0$, that is, we directly observe $X = \sigma W$. In our general framework, this corresponds to setting $G = \delta_0$, a Dirac mass at 0, and we set the (now unidentified) parameter θ to 0, or for that matter to any arbitrary value.

We start by studying whether the limiting behavior of $I_\Delta^{\sigma\sigma}(\sigma, \beta, \theta, G)$ when $\beta = 2$ and of the (σ, β) block of the matrix $I_\Delta(\sigma, \beta, \theta, G)$ when $\beta < 2$ is affected by the presence of Y. We have the intuitively obvious majoration of Fisher's information in the presence of Y by the one for which $Y = 0$. Note that in this result no assumption whatsoever is made on Y (except of course that it is independent of W):

THEOREM 1: *For any* $\Delta > 0$, *we have*

$$I_\Delta^{\sigma\sigma}(\sigma, 2, \theta, G) \le I_\Delta^{\sigma\sigma}(\sigma, 2, 0, \delta_0) \tag{3.1}$$

and, when $\beta < 2$, *the difference*

$$\begin{pmatrix} I_\Delta^{\sigma\sigma}(\sigma,\beta,0,\delta_0) & I_\Delta^{\sigma\beta}(\sigma,\beta,0,\delta_0) \\ I_\Delta^{\sigma\beta}(\sigma,\beta,0,\delta_0) & I_\Delta^{\beta\beta}(\sigma,\beta,0,\delta_0) \end{pmatrix} - \begin{pmatrix} I_\Delta^{\sigma\sigma}(\sigma,\beta,\theta,G) & I_\Delta^{\sigma\beta}(\sigma,\beta,\theta,G) \\ I_\Delta^{\sigma\beta}(\sigma,\beta,\theta,G) & I_\Delta^{\beta\beta}(\sigma,\beta,\theta,G) \end{pmatrix}$$

is a positive semidefinite matrix; *in particular we have*

$$I_\Delta^{\beta\beta}(\sigma,\beta,\theta,G) \le I_\Delta^{\beta\beta}(\sigma,\beta,0,\delta_0). \tag{3.2}$$

Next, how does the limit as $\Delta \to 0$ of $I_\Delta(\sigma,\beta,\theta,G)$ compare to that of $I_\Delta(\sigma,\beta,0,\delta_0)$? To answer this question, we need some further notation. For the first two derivatives of the density h_β of W_1, we write h'_β and h''_β, and we also introduce the functions

$$\breve{h}_\beta(w) = h_\beta(w) + w h'_\beta(w), \qquad \widetilde{h}_\beta(w) = \frac{\breve{h}_\beta(w)^2}{h_\beta(w)}. \tag{3.3}$$

Then $\tilde{h}_\beta$ is positive, even, continuous, and $\tilde{h}_\beta(w) = O(1/|w|^{1+\beta})$ as $|w| \to \infty$; hence $\tilde{h}_\beta$ is Lebesgue-integrable.

Let

$$\mathcal{I}(\beta) = \int \widetilde{h}_\beta(w)\, dw. \tag{3.4}$$

This is a positive number, which takes the value $\mathcal{I}(\beta) = 2$ when $\beta = 2$.

When $Y = 0$, we have $p_\Delta(x|\sigma,\beta,0,\delta_0) = (1/\sigma\Delta^{1/\beta}) h_\beta(x/\sigma\Delta^{1/\beta})$ and, therefore,

$$I_\Delta^{\sigma\sigma}(\sigma,\beta,0,\delta_0) = \frac{1}{\sigma^2}\mathcal{I}(\beta), \tag{3.5}$$

which does not depend on Δ. $\mathcal{I}(\beta)$ is simply the Fisher's information at point $\sigma = 1$ for the statistical model in which we observe σW_1.

The limiting behavior of the (σ,β) block of Fisher's information is given by the following theorem:

THEOREM 2: (a) *If* $G \in \mathcal{G}_\beta$, *we have*, *as* $\Delta \to 0$,

$$I_\Delta^{\sigma\sigma}(\sigma,\beta,\theta,G) \to \frac{1}{\sigma^2}\mathcal{I}(\beta) \tag{3.6}$$

and also, *when* $\beta < 2$,

$$\frac{I_\Delta^{\beta\beta}(\sigma,\beta,\theta,G)}{\log(1/\Delta)^2} \to \frac{1}{\beta^4}\mathcal{I}(\beta), \qquad \frac{I_\Delta^{\sigma\beta}(\sigma,\beta,\theta,G)}{\log(1/\Delta)} \to \frac{1}{\sigma\beta^2}\mathcal{I}(\beta). \tag{3.7}$$

(b) *For any* $\phi \in \Phi$ *and* $\alpha \in (0, \beta]$, *and* $K > 0$, *we have as* $\Delta \to 0$,

$$\sup_{G \in \mathcal{G}(\phi,\alpha), |\theta| \le K} \left| I_\Delta^{\sigma\sigma}(\sigma, \beta, \theta, G) - \frac{\mathcal{I}(\beta)}{\sigma^2} \right| \to 0, \tag{3.8}$$

$$\beta < 2 \quad \Rightarrow \quad \begin{cases} \displaystyle\sup_{G \in \mathcal{G}(\phi,\alpha), |\theta| \le K} \left| \frac{I_\Delta^{\beta\beta}(\sigma, \beta, \theta, G)}{(\log(1/\Delta))^2} - \frac{\mathcal{I}(\beta)}{\beta^4} \right| \to 0, \\ \displaystyle\sup_{G \in \mathcal{G}(\phi,\alpha), |\theta| \le K} \left| \frac{I_\Delta^{\sigma\beta}(\sigma, \beta, \theta, G)}{\log(1/\Delta)} - \frac{\mathcal{I}(\beta)}{\sigma\beta^2} \right| \to 0. \end{cases}$$

(c) *For each* n, *let* G^n *be the standard symmetric stable law of index* α_n, *with* α_n *a sequence strictly increasing to* β. *Then for any sequence* $\Delta_n \to 0$ *such that* $(\beta - \alpha_n)\log \Delta_n \to 0$ (*i.e.*, *the rate at which* $\Delta_n \to 0$ *is slow enough*), *the sequence of numbers* $I_{\Delta_n}^{\sigma\sigma}(\sigma, \beta, \theta, G^n)$ (*resp.* $I_{\Delta_n}^{\beta\beta}(\sigma, \beta, \theta, G^n)/(\log(1/\Delta_n))^2$ *when further* $\beta < 2$) *converges to a limit which is strictly less than* $\mathcal{I}(\beta)/\sigma^2$ (*resp.* $\mathcal{I}(\beta)/\beta^4$).

Since the result (a) is valid for all $G \in \mathcal{G}_\beta$, it is valid in particular when $G = \delta_0$, that is, when $Y = 0$. In other words, at their respective leading orders in Δ, the presence of Y has no impact on the information terms $I_\Delta^{\sigma\sigma}$, $I_\Delta^{\beta\beta}$, and $I_\Delta^{\sigma\beta}$, as soon as Y is dominated by W: so, in the limit where $\Delta_n \to 0$, the parameters σ and β can be estimated with the exact same degree of precision whether Y is present or not. Moreover, part (b) states the convergence of Fisher's information is *uniform* on the set $\mathcal{G}(\phi, \alpha)$ and $|\theta| \le K$ for all $\alpha \in [0, \beta]$; this settles the case where G and θ are considered as nuisance parameters when we estimate σ and β. But as α tends to β, the *convergence disappears*, as stated in part (c).

The part of the theorem related to the $\sigma\sigma$ term alone is discussed in Aït-Sahalia and Jacod (2007), where we use it to construct explicit $\sqrt{n}$-converging estimators of the parameter σ that are strongly efficient in the sense that not only do they converge at rate $\sqrt{n}$, but they also have the same asymptotic variance when Y is present as when it is not: $\sqrt{n}(\widehat{\sigma}_n - \sigma)$ converges in law to $\mathcal{N}(0, \sigma^2/\mathcal{I}(\beta))$ as soon as $\Delta_n \to 0$. This is the case even in the semiparametric case where nothing is known about Y, other than the fact that Y is sufficiently away from W, meaning that $G \in \mathcal{G}_\alpha$ for α small enough: $\alpha < \beta$ is not enough and the precise condition on α depends on the behavior of the pair (n, Δ_n); see Theorem 5 of that paper.

When it comes to estimating β, with σ known, things are different. When $\Delta_n \to 0$ and the true value is $\beta < 2$, we can hope for estimators converging to β at the *faster* rate $\sqrt{n}\log(1/\Delta_n)$, in view of the limit for $I_\Delta^{\beta\beta}$ given in (3.7). In fact, for a single observation, say X_Δ, there exists of course no consistent estimator for σ, but there is one for β when $\Delta \to 0$: namely $\widehat{\beta}_\Delta = -\log(1/\Delta)/\log|X_\Delta|$, which converges in probability to β as Δ, and the rate of convergence is $\log(1/\Delta)$; all these follow from the scaling property of W.

Estimators for β could be constructed using the jumps of the process of a size greater than some threshold (see Höpfner and Jacod (1994)). Estimating β was studied by DuMouchel (1973a), who computed numerically the term $I_\Delta^{\beta\beta}(\sigma, \beta, 0, \delta_0)$, including also an asymmetry parameter. Note that if one suspects that $\beta = 2$, then one should perform a test, as advised by DuMouchel (1973a), and in that case the behavior of the Fisher's information about β does not provide much insight.

The introduction of β as a parameter to be estimated makes the joint estimation of the pair (σ, β) a singular problem. The asymptotic (normalized) 2×2 Fisher's information matrix given by (3.6) and (3.7) is not invertible. In particular, there is no guarantee that the joint MLE for the pair (σ, β), for example, behaves well. However, the estimation of either parameter knowing the other remains a regular problem, albeit with a rate faster than $\sqrt{n}$ for β.

3.2. *Inference on θ*

For the entries of Fisher's information matrix involving the parameter θ, things are more complicated. First, observe that $I_\Delta^{\theta\theta}(0, \beta, \theta, G)$ (that is, the Fisher's information for the model $X = \theta Y$) does not necessarily exist, but of course if it does we have an inequality similar to (3.1) for all σ:

$$I_\Delta^{\theta\theta}(\sigma, \beta, \theta, G) \leq I_\Delta^{\theta\theta}(0, \beta, \theta, G). \tag{3.9}$$

Contrary to (3.1), however, this is a very rough estimate, which does not take into account the properties of W. The (θ, θ) Fisher's information is usually much smaller than what the right side above suggests, and we give below a more accurate estimate when Y has second moments, but without the domination assumption that $G \in \mathcal{G}_\beta$.

For this, we see another information appear, namely the Fisher's information associated with the estimation of the real number a for the model where one observes the single variable $W_1 + a$. This Fisher's information is

$$\mathcal{J}(\beta) = \int \frac{h'_\beta(w)^2}{h_\beta(w)}\, dw. \tag{3.10}$$

Observe in particular that $\mathcal{J}(\beta) = 1$ when $\beta = 2$.

THEOREM 3: *If Y_1 has finite variance v and mean m, we have*

$$I_\Delta^{\theta\theta}(\sigma, \beta, \theta, G) \leq \frac{\mathcal{J}(\beta)}{\sigma^2}\big(m^2 \Delta^{2-2/\beta} + v\Delta^{1-2/\beta}\big). \tag{3.11}$$

This estimate holds for all $\Delta > 0$. The asymptotic variant, which says that

$$\limsup_{\Delta \to 0} \Delta^{2/\beta - 1} I_\Delta^{\theta\theta}(\sigma, \beta, \theta, G) \leq \frac{v\mathcal{J}(\beta)}{\sigma^2}, \tag{3.12}$$

is sharp in some cases and not in others, as we will see in the examples below. These examples will also illustrate how the "translation" Fisher's information $\mathcal{J}(\beta)$ comes into the picture here.

Theorem 3 does not solve the problem in full generality: for a generic Y, we only obtain a bound. For some specific Y's, we will be able to give a full result in Section 4. It is indeed a feature of this problem that the estimation of θ depends heavily on the specific nature of the Y process, including the fact that the optimal rate at which θ can be estimated depends on the process Y. Hence it is not surprising that any general result regarding θ, such as Theorem 3, can at best state a bound: there is no single common rate that applies to all Y's.

4. SPECIFIC EXAMPLES OF LÉVY PROCESSES

The calculations of the previous section involving the parameter θ can be made fully explicit if we specify the distribution of the process Y. We can then illustrate the surprisingly wide range of situations that can occur to the rates of convergence. For simplicity, let us consider β as known in these examples and focus on the joint estimation of (σ, θ).

4.1. *Stable Process Plus Drift*

Here we assume that $Y_t = t$, so $G = \delta_1$ and θ is a drift parameter:

THEOREM 4: *The* 2×2 *Fisher's information matrix for estimating* (σ, θ) *is*

$$\begin{aligned}(4.1)\qquad &\begin{pmatrix} I_\Delta^{\sigma\sigma}(\sigma, \beta, \theta, \delta_1) & I_\Delta^{\sigma\theta}(\sigma, \beta, \theta, \delta_1) \\ I_\Delta^{\sigma\theta}(\sigma, \beta, \theta, \delta_1) & I_\Delta^{\theta\theta}(\sigma, \beta, \theta, \delta_1) \end{pmatrix} \\ &\quad = \frac{1}{\sigma^2}\begin{pmatrix} \mathcal{I}(\beta) & 0 \\ 0 & \Delta^{2-2/\beta}\mathcal{J}(\beta) \end{pmatrix}.\end{aligned}$$

This has several interesting consequences (we denote by $T_n = n\Delta_n$ the length of the *observation window*):

1. If θ is known, (4.1) suggests the existence of estimators $\widehat{\sigma}_n$ satisfying $\sqrt{n}(\widehat{\sigma}_n - \sigma) \xrightarrow{d} \mathcal{N}(0, \sigma^2/\mathcal{I}(\beta))$. This is the case and, in fact, since θ is known, observing $X_{i\Delta_n}$ and observing $W_{i\Delta_n}$ are equivalent.

2. If σ is known, one may hope for estimators $\widehat{\theta}_n$ satisfying $\sqrt{n}\Delta_n^{1-1/\beta}(\widehat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2/\mathcal{J}(\beta))$. If $\beta = 2$, the rate is thus $\sqrt{T_n}$: this is in accordance with the classical fact that for a diffusion the rate for estimating the drift coefficient is the square root of the total observation window, that is, $\sqrt{T_n}$ here. Moreover, in this case, the variable X_{T_n}/T_n is $\mathcal{N}(\theta, \sigma^2/T_n)$, so $\widehat{\theta}_n = X_{T_n}/T_n$ is an asymptotically efficient estimator for θ (recall that $\mathcal{J}(\beta) = 1$ when $\beta = 2$). When $\beta < 2$, we have $1 - 1/\beta < 1/2$, so the rate is bigger than $\sqrt{T_n}$ and it increases when β decreases; when $\beta < 1$, this rate is even bigger than $\sqrt{n}$.

3. Observe that here Y_1 has mean $m = 1$ and variance $v = 0$, so the bound (3.11) in Theorem 3 is indeed an equality. The fact that the translation Fisher's information $\mathcal{J}(\beta)$ appears here is transparent.

4. If both σ and θ are unknown, one may hope for estimators $\widehat{\sigma}_n$ and $\widehat{\theta}_n$ such that the pairs $(\sqrt{n}(\widehat{\sigma}_n - \sigma), \sqrt{n}\Delta_n^{1-1/\beta}(\widehat{\theta}_n - \theta))$ converge in law to the product $\mathcal{N}(0, \sigma^2/\mathcal{I}(\beta)) \otimes \mathcal{N}(0, \sigma^2/\mathcal{J}(\beta))$.

4.2. *Stable Process Plus Poisson Process*

Here we assume that Y is a standard Poisson process (for simplicity jumps of size 1, intensity 1) whose law we write as $G = P$. We can describe the limiting behavior of the (σ, θ) block of the matrix $I_\Delta(\sigma, \beta, \theta, P)$ as $\Delta \to 0$.

THEOREM 5: *If Y is a standard Poisson process, we have, as $\Delta \to 0$,*

$$I_\Delta^{\sigma\sigma}(\sigma, \beta, \theta, P) \to \frac{1}{\sigma^2}\mathcal{I}(\beta), \tag{4.2}$$

$$\Delta^{1/\beta-1/2} I_\Delta^{\sigma\theta}(\sigma, \beta, \theta, P) \to 0, \tag{4.3}$$

$$\Delta^{2/\beta-1} I_\Delta^{\theta\theta}(\sigma, \beta, \theta, P) \to \frac{1}{\sigma^2}\mathcal{J}(\beta). \tag{4.4}$$

Since $P \in \mathcal{G}_\beta$, (4.2) is nothing else than the first part of (3.6). One could prove more than (4.3), namely that $\sup_\Delta \Delta^{1/\beta-1} |I_\Delta^{\sigma\theta}(\sigma, \beta, \theta, P)| \le \infty$. Since Y_1 has mean $m = 1$ and variance $v = 1$, the asymptotic estimate (3.12) in Theorem 3 is sharp in view of (4.4). Here again, we deduce some interesting consequences for the estimation:

1. If σ is known, one may hope for estimators $\widehat{\theta}_n$ satisfying $\sqrt{n\Delta_n^{1-2/\beta}}(\widehat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2/\mathcal{J}(\beta))$. So the rate is *faster* than $\sqrt{n}$, except when $\beta = 2$. More generally, if both σ and θ are unknown, one may hope for estimators $\widehat{\sigma}_n$ and $\widehat{\theta}_n$ such that the pairs $(\sqrt{n}(\widehat{\sigma}_n - \sigma), \sqrt{n\Delta_n^{1-2/\beta}}(\widehat{\theta}_n - \theta))$ converge in law to the product $\mathcal{N}(0, \sigma^2/\mathcal{I}(\beta)) \otimes \mathcal{N}(0, \sigma^2/\mathcal{J}(\beta))$.

2. However, the above-described behavior of any estimator $\widehat{\theta}_n$ cannot be true when $T_n = n\Delta_n$ does not go to infinity, because in this case there is a positive probability that Y has no jump on the biggest observed interval, and so no information about θ can be drawn from the observations in that case. It is true, though, when $T_n \to \infty$, because Y will eventually have infinitely many jumps on the observed intervals. There is a large literature on this subject, starting with the base case where the Poisson process is directly observed, for which the same problem occurs.

4.3. *Stable Process Plus Compound Poisson Process*

Here we assume that Y is a compound Poisson process with arrival rate λ and law of jumps μ: that is, the characteristics of G are $b = \lambda \int_{\{|x|\le 1\}} x\mu(dx)$

and $c=0$ and $F=\lambda\mu$. We then write $G=P_{\lambda,\mu}$, which belongs to $\mathcal{G}_\beta$. If $\beta=2$, we are then in the classical context of jump diffusions in finance.

We will further assume that μ has a density f satisfying

$$\lim_{|u|\to\infty} uf(u)=0,\qquad \sup_u\bigl(|f'(u)|(1+|u|)\bigr)<\infty. \tag{4.5}$$

We also suppose that the "multiplicative" Fisher's information associated with μ (that is, the Fisher's information for estimating θ in the model when one observes a single variable θU with U distributed according to μ) exists. It then has the form

$$\mathcal{L}=\int \frac{(xf'(x)+f(x))^2}{f(x)}\,dx. \tag{4.6}$$

We can describe the limiting behavior of the (σ,θ) block of the matrix $I_\Delta(\sigma,\beta,\theta,P_{\lambda,\mu})$ as $\Delta\to 0$.

THEOREM 6: *If Y is a compound Poisson process satisfying* (4.5) *and such that $\mathcal{L}$ in* (4.6) *is finite*, *we have*, *as* $\Delta\to 0$,

$$I_\Delta^{\sigma\sigma}(\sigma,\beta,\theta,P_{\lambda,\mu})\to\frac{1}{\sigma^2}\mathcal{I}(\beta), \tag{4.7}$$

$$\frac{1}{\sqrt{\Delta}}I_\Delta^{\sigma\theta}(\sigma,\beta,\theta,P_{\lambda,\mu})\to 0, \tag{4.8}$$

$$\frac{\lambda^2}{\theta^2}\int\frac{(xf'(x)+f(x))^2}{\lambda f(x)+\frac{c_\beta\sigma^\beta}{\theta^\beta|x|^{1+\beta}}}\,dx\le\liminf\frac{1}{\Delta}I_\Delta^{\theta\theta}(\sigma,\beta,\theta,P_{\lambda,\mu})\le\frac{1}{\theta^2}\mathcal{L}, \tag{4.9}$$

and also, *when* $\beta=2$,

$$\frac{1}{\Delta}I_\Delta^{\theta\theta}(\sigma,2,\theta,P_{\lambda,\mu})\to\frac{1}{\theta^2}\mathcal{L}. \tag{4.10}$$

As for the previous theorem, (4.7) is nothing else than the first part of (3.6). We could prove more than (4.8), namely that $\sup_\Delta\frac{1}{\Delta}|I_\Delta^{\sigma\theta}(\sigma,\beta,\theta,P_{\lambda,\mu})|<\infty$. Here again, we draw some interesting consequences:

1. Equations (4.9) and (4.10) suggest that there are estimators $\widehat{\theta}_n$ such that $\sqrt{T_n}(\widehat{\theta}_n-\theta)$ is tight (the rate is the same as for the case $Y_t=t$) and is even asymptotically normal when $\beta=2$. But this is of course wrong when T_n does not go to infinity, for the same reason as in the previous theorem.

2. When the measure μ has a second order moment, the right side of (3.11) is larger than the result of the previous theorem, so the estimate in Theorem 3 is not sharp.

The rates for estimating θ in the two previous theorems, and the limiting Fisher's information as well, can be explained as follows (supposing that σ is known and that we have n observations and that $T_n \to \infty$):

1. For Theorem 5: θ comes into the picture whenever the Poisson process has a jump. On the interval $[0, T_n]$ we have an average of T_n jumps, most of them being isolated in an interval $(i\Delta_n, (i+1)\Delta_n]$. So it essentially amounts to observing T_n (or rather the integer part $[T_n]$) independent variables, all distributed as $\sigma\Delta_n^{1/\beta}W_1 + \theta$. The Fisher's information for each of those (for estimating θ) is $\mathcal{J}(\beta)/\sigma^2\Delta^{2/\beta}$, and the "global" Fisher's information, namely $nI_{\Delta_n}^{\theta\theta}$, is approximately $T_n\mathcal{J}(\beta)/\sigma^2\Delta_n^{2/\beta} \sim \mathcal{J}(\beta)/\sigma^2\Delta_n^{2/\beta-1}$.

2. For Theorem 6: Again θ comes into the picture whenever the compound Poisson process has a jump. We have an average of λT_n jumps, so it essentially amounts to observing λT_n independent variables, all distributed as $\sigma\Delta_n^{1/\beta}W_1 + \theta V$, where V has the distribution μ. The Fisher's information for each of those (for estimating θ) is approximately $\mathcal{L}/\theta^2$ (because the variable $\sigma\Delta_n^{1/\beta}W_1$ is negligible), and the "global" Fisher's information $nI_{\Delta_n}^{\theta\theta}$ is approximately $\lambda T_n\mathcal{L}/\theta^2 \sim n\Delta_n\mathcal{L}/\theta^2$. This explains the rate in (4.9), and is an indication that (4.10) may be true even when $\beta < 2$, although we have been unable to prove it thus far.

4.4. *Two Stable Processes*

Our last example is about the case where Y is also a symmetric stable process with index α, $\alpha < \beta$. We write $G = S_\alpha$. Surprisingly, the results are quite involved, in the sense that for estimating θ we have different situations according to the relative values of α and β. We obviously still have (4.7), so we concentrate on the term $I_\Delta^{\theta\theta}$ and ignore the cross term in the statement of the following theorem. The constant c_β below is the one occurring in (2.2).

THEOREM 7: *If Y is a standard symmetric stable process with index $\alpha < \beta$, we have the following situations as $\Delta \to 0$:*

(a) *If $\beta = 2$,*

$$\frac{(\log(1/\Delta))^{\alpha/2}}{\Delta^{(\beta-\alpha)/\beta}} I_\Delta^{\theta\theta}(\sigma, \beta, \theta, S_\alpha) \to \frac{2\alpha c_\alpha \beta^{\alpha/2}}{\theta^{2-\alpha}\sigma^\alpha(2(\beta-\alpha))^{\alpha/2}}. \tag{4.11}$$

(b) *If $\beta < 2$ and $\alpha > \frac{\beta}{2}$,*

$$\begin{aligned} &\Delta^{(2(\beta-\alpha))/\beta} I_\Delta^{\theta\theta}(\sigma, \beta, \theta, S_\alpha) \\ &\quad \to \frac{\alpha^2 c_\alpha^2 \theta^{2\alpha-2}}{\sigma^{2\alpha}} \int \frac{(\int_{\mathbf{R}} |y|^{1-\alpha}\, dy \int_0^1 (1-v) h_\beta''(x-yv)\, dv)^2}{h_\beta(y)}\, dx. \end{aligned} \tag{4.12}$$

(c) *If* $\beta < 2$ *and* $\alpha = \frac{\beta}{2}$,

$$\Delta^{(2(\beta-\alpha))/\beta} \log\left(\frac{1}{\Delta}\right) I_\Delta^{\theta\theta}(\sigma, \beta, \theta, S_\alpha) \to \frac{2\alpha(\beta-\alpha)c_\alpha^2 \theta^{2\alpha-2}}{\beta c_\beta \sigma^{2\alpha}}. \tag{4.13}$$

(d) *If* $\beta < 2$ *and* $\alpha < \frac{\beta}{2}$,

$$\frac{1}{\Delta} I_\Delta^{\theta\theta}(\sigma, \beta, \theta, S_\alpha) \to \frac{\alpha^2 c_\alpha^2 \theta^{2\alpha-2}}{\sigma^{2\alpha}} \int \frac{dz}{c_\beta |z|^{1+2\alpha-\beta} + c_\alpha \theta^\alpha |z|^{1+\alpha}/\sigma^\alpha}. \tag{4.14}$$

Then if σ is known, one may hope to find estimators $\widehat{\theta}_n$ with $u_n(\widehat{\theta}_n - \theta) \xrightarrow{d} N(0, V)$, with

$$u_n = \begin{cases} \sqrt{n}\Delta_n^{(2-\alpha)/4}/(\log(1/\Delta_n))^{\alpha/4}, & \text{if } \beta = 2, \\ \sqrt{n}\Delta_n^{(\beta-\alpha)/\beta}, & \text{if } \beta < 2, \alpha > \beta/2, \\ \sqrt{n}\Delta_n^{1/2}\sqrt{\log(1/\Delta_n)}, & \text{if } \beta < 2, \alpha = \beta/2, \\ \sqrt{n\Delta_n}, & \text{if } \beta < 2, \alpha < \beta/2, \end{cases}$$

and of course the asymptotic variance V should be the inverse of the right-hand sides in (4.11)–(4.14).

5. CONCLUSIONS

We studied in this paper a tractable parametric model with jumps, also including a continuous component, and examined the impact of the presence of the different components of the model (continuous part vs. jump part and/or jumps with different relative levels of activity) on the various parameters of the model. We determined how optimal estimators should behave where the parameter vector of interest is (σ, β, θ), viewing the law of Y nonparametrically. While optimal estimators for σ should converge at rate $\sqrt{n}$ and have the same asymptotic variance with or without Y, the optimal rate for β is faster and given by $\sqrt{n}\log(1/\Delta_n)$, and is also unaffected by the presence of Y. But the rate of convergence of optimal estimators of θ is very dependent on the law of Y and is affected by the presence of W.

Given that we now know what can be achieved by efficient estimators, the natural next step is to exhibit actual, and tractable, estimators with those optimal properties. We did not attempt here to construct estimators of these parameters, and leave that question to future work. As a partial step in that direction, we studied in Aït-Sahalia and Jacod (2007) estimation procedures, for the parameter σ only, that satisfy the property that those estimators are asymptotically as efficient as when the process Y is absent.

Dept. of Economics, Bendheim Center for Finance, Princeton University, 26 Prospect Avenue, Princeton, NJ 08540-5296, U.S.A. and NBER; yacine@princeton.edu

and

Institut de Mathématiques de Jussieu, CNRS–UMR 7586, 175 Rue due Chevaleret, 75013 Paris, France and Université P. et M. Curie, Paris-6, France; jj@ccr.jussieu.fr.

Manuscript received May, 2006; final revision received October, 2007.

APPENDIX: PROOFS

A. *Fisher's Information When $X = \sigma W + \theta Y$*

By independence of W and Y, the density of X_Δ in (1.1) is the convolution (recall that G_Δ is the law of Y_Δ)

$$p_\Delta(x|\sigma,\beta,\theta,G) = \frac{1}{\sigma\Delta^{1/\beta}} \int G_\Delta(dy) h_\beta\left(\frac{x-\theta y}{\sigma\Delta^{1/\beta}}\right). \tag{A.1}$$

We now seek to characterize the entries of the full Fisher's information matrix. Since h'_β, $\breve{h}_\beta$, and $\dot{h}_\beta$ are continuous and bounded, we can differentiate under the integral in (A.1) to get

$$\partial_\sigma p_\Delta(x|\sigma,\beta,\theta,G) = -\frac{1}{\sigma^2\Delta^{1/\beta}} \int G_\Delta(dy) \breve{h}_\beta\left(\frac{x-\theta y}{\sigma\Delta^{1/\beta}}\right), \tag{A.2}$$

$$\partial_\beta p_\Delta(x|\sigma,\beta,\theta,G) = v_\Delta(x|\sigma,\beta,\theta,G) - \frac{\sigma\log\Delta}{\beta^2}\partial_\sigma p_\Delta(x|\sigma,\beta,\theta,G), \tag{A.3}$$

$$\partial_\theta p_\Delta(x|\sigma,\beta,\theta,G) = -\frac{1}{\sigma^2\Delta^{2/\beta}} \int G_\Delta(dy) y h'_\beta\left(\frac{x-\theta y}{\sigma\Delta^{1/\beta}}\right), \tag{A.4}$$

where

$$v_\Delta(x|\sigma,\beta,\theta,G) = \frac{1}{\sigma\Delta^{1/\beta}} \int G_\Delta(dy) \dot{h}_\beta\left(\frac{x-\theta y}{\sigma\Delta^{1/\beta}}\right). \tag{A.5}$$

The entries of the (σ,θ) block of the Fisher's information matrix are (leaving implicit the dependence on (σ,β,θ,G))

$$I_\Delta^{\sigma\sigma} = \int \frac{\partial_\sigma p_\Delta(x)^2}{p_\Delta(x)}\,dx, \quad I_\Delta^{\sigma\theta} = \int \frac{\partial_\sigma p_\Delta(x)\partial_\theta p_\Delta(x)}{p_\Delta(x)}\,dx, \tag{A.6}$$
$$I_\Delta^{\theta\theta} = \int \frac{\partial_\theta p_\Delta(x)^2}{p_\Delta(x)}\,dx.$$

When $\beta < 2$, the other entries are

$$I_\Delta^{\sigma\beta} = J_\Delta^{\sigma\beta} - \frac{\sigma \log \Delta}{\beta^2} I_\Delta^{\sigma\sigma}, \quad I_\Delta^{\beta\theta} = J_\Delta^{\beta\theta} - \frac{\sigma \log \Delta}{\beta^2} I_\Delta^{\sigma\theta}, \tag{A.7}$$

$$I_\Delta^{\beta\beta} = J_\Delta^{\beta\beta} - \frac{2\sigma \log \Delta}{\beta^2} J_\Delta^{\sigma\beta} + \frac{\sigma^2 (\log \Delta)^2}{\beta^4} I_\Delta^{\sigma\sigma}, \tag{A.8}$$

where

$$\begin{aligned} J_\Delta^{\sigma\beta} &= \int \frac{\partial_\sigma p_\Delta(x) v_\Delta(x)}{p_\Delta(x)}\, dx, \quad J_\Delta^{\beta\beta} = \int \frac{v_\Delta(x)^2}{p_\Delta(x)}\, dx, \\ J_\Delta^{\beta\theta} &= \int \frac{v_\Delta(x) \partial_\theta p_\Delta(x)}{p_\Delta(x)}\, dx. \end{aligned} \tag{A.9}$$

B. *Proof of Theorem 1*

The proof is standard and given for completeness, and given only in the case where $\beta < 2$ (when $\beta = 2$ take $v = 0$ below). What we need to prove is that, for any $u, v \in R$, we have

$$\begin{aligned} &\int \frac{(u\partial_\sigma p_\Delta(x|\sigma, \beta, \theta, G) + v\partial_\beta p_\Delta(x|\sigma, \beta, \theta, G))^2}{p_\Delta(x|\sigma, \beta, \theta, G)}\, dx \\ &\quad \le \int \frac{(u\partial_\sigma p_\Delta(x|\sigma, \beta, 0, \delta_0) + v\partial_\beta p_\Delta(x|\sigma, \beta, 0, \delta_0))^2}{p_\Delta(x|\sigma, \beta, 0, \delta_0)}\, dx. \end{aligned} \tag{B.1}$$

We set

$$\begin{aligned} &q(x) = p_\Delta(x|\sigma, \beta, \theta, G), \quad q_0(x) = p_\Delta(x|\sigma, \beta, 0, \delta_0), \\ &r(x) = u\partial_\sigma p_\Delta(x|\sigma, \beta, \theta, G) + v\partial_\beta p_\Delta(x|\sigma, \beta, \theta, G), \\ &r_0(x) = u\partial_\sigma p_\Delta(x|\sigma, \beta, 0, \delta_0) + v\partial_\beta p_\Delta(x|\sigma, \beta, 0, \delta_0). \end{aligned}$$

Observe that by (A.1), $q(x) = \int G_\Delta(dy) q_0(x - \theta y)$, hence $r(x) = \int G_\Delta(dy) \times r_0(x - \theta y)$ as well. Apply the Cauchy–Schwarz inequality to G_Δ with $r_0 = \sqrt{q_0}(r_0/\sqrt{q_0})$ to get

$$r(x)^2 \le q(x) \int G_\Delta(dy) \frac{r_0(x - \theta y)^2}{q_0(x - \theta y)}.$$

Then

$$\int \frac{r(x)^2}{q(x)}\, dx \le \int dx \int G_\Delta(dy) \frac{r_0(x - \theta y)^2}{q_0(x - \theta y)} = \int \frac{r_0(z)^2}{q_0(z)}\, dz$$

by Fubini's theorem and a change of variable: this is exactly (B.1).

C. *Proof of Theorem 2*

It is easily seen from the characteristic function that $\beta \mapsto h_\beta(w)$ is differentiable on $(0, 2]$, and we denote by $\dot{h}_\beta(w)$ its derivative. However, instead of (2.2), one has

$$\text{(C.1)} \qquad |\dot{h}_\beta(w)| \sim \begin{cases} \dfrac{c_\beta \log|w|}{|w|^{1+\beta}}, & \text{if } \beta < 2, \\ \dfrac{1}{|w|^3}, & \text{if } \beta = 2, \end{cases}$$

as $|w| \to \infty$, by differentiation of the series expansion for the stable density. Therefore the quantity

$$\text{(C.2)} \qquad \mathcal{K}(\beta) = \int \frac{\dot{h}_\beta(w)^2}{h_\beta(w)}\, dw$$

is finite when $\beta < 2$ and infinite for $\beta = 2$. This is the Fisher's information for estimating β, upon observing the single variable W_1.

First, part (b) of the theorem implies (a). Second, the claim about the $\sigma\sigma$ term

$$\text{(C.3)} \qquad I_{\Delta_n}^{\sigma\sigma}(\sigma, \beta, \theta_n, G^n) \to \frac{\mathcal{I}(\beta)}{\sigma^2}$$

and its uniformity are proved in Aït-Sahalia and Jacod (2007), so we only need to prove the claims about $I^{\beta\beta}$ and $I^{\sigma\beta}$ in (b) and (c).

We assume $\beta < 2$. If (b) fails, one can find $\phi \in \Phi$, $\alpha \in (0, \beta]$, and $\varepsilon > 0$, and also a sequence $\Delta_n \to 0$ and a sequence G^n of measures in $G(\phi, \alpha)$ and a sequence of numbers θ_n converging to a limit θ, such that

$$\left| \frac{I_{\Delta_n}^{\beta\beta}(\sigma, \beta, \theta_n, G^n)}{(\log(1/\Delta_n))^2} - \frac{\mathcal{I}(\beta)}{\beta^4} \right| + \left| \frac{I_{\Delta_n}^{\sigma\sigma}(\sigma, \beta, \theta_n, G^n)}{\log(1/\Delta_n)} - \frac{\mathcal{I}(\beta)}{\sigma\beta^2} \right| \geq \varepsilon$$

for all n. In other words, we will get a contradiction and (b) will be proved as soon as we show that

$$\text{(C.4)} \qquad \frac{I_{\Delta_n}^{\beta\beta}(\sigma, \beta, \theta_n, G^n)}{(\log(1/\Delta))^2} \to \frac{\mathcal{I}(\beta)}{\beta^4}, \qquad \frac{I_{\Delta_n}^{\sigma\beta}(\sigma, \beta, \theta_n, G^n)}{\log(1/\Delta)} \to \frac{\mathcal{I}(\beta)}{\sigma\beta^2}.$$

Recall (A.7) and (A.8): with the notation $J_\Delta^{\sigma\beta}(\sigma, \beta, \theta, G), \ldots,$ of (A.9), we see that

$$\text{(C.5)} \qquad J_\Delta^{\sigma\beta}(\sigma, \beta, \theta, G) \leq \sqrt{I_\Delta^{\sigma\sigma}(\sigma, \beta, \theta, G) J_\Delta^{\beta\beta}(\sigma, \beta, \theta, G)}$$

by a first application of the Cauchy–Schwarz inequality. A second application of the same plus (A.1) and (A.5) yields

$$v_\Delta(x|\sigma,\beta,\theta,G)^2 \leq \frac{p_\Delta(x|\sigma,\beta,\theta,G)}{\sigma\Delta^{1/\beta}}\int G_\Delta(dy)\frac{\dot{h}_\beta^2}{h_\beta}\left(\frac{x-\theta y}{\sigma\Delta^{1/\beta}}\right).$$

Then Fubini's theorem and the change of variable $x \leftrightarrow (x-\theta y)/\sigma\Delta^{1/\beta}$ in (A.9) leads to

$$J_\Delta^{\beta\beta}(\sigma,\beta,\theta,G) \leq \mathcal{K}(\beta). \tag{C.6}$$

Next, (C.4) readily follows from (C.5), (C.6), and (C.3), and also (A.7) and (A.8).

Finally, it remains to prove the part of (c) concerning $I^{\beta\beta}$. We already know that the sequence $I_{\Delta_n}^{\sigma\sigma}(\sigma,\beta,\theta,G^n)$ converges to a limit which is strictly less than $\mathcal{I}(\beta)/\sigma^2$. Then by (A.8), (C.5), (C.6), and (C.3), it is clear that $I_{\Delta_n}^{\beta\beta}(\sigma,\beta,\theta,G^n)/(\log(1/\Delta_n))^2$ also converges to a limit which is strictly less than $\mathcal{I}(\beta)/\beta^4$.

D. *Proof of Theorem 3*

The Cauchy–Schwarz inequality gives us, by (A.1) and (A.4),

$$\begin{aligned}&|\partial_\theta p_\Delta(x|\sigma,\beta,\theta,G)|^2\\ &\quad\leq \frac{1}{\sigma^3\Delta^{3/\beta}}p_\Delta(x|\sigma,\beta,\theta,G)\int G_\Delta(dy)y^2\frac{[h'_\beta((x-\theta y)/\sigma\Delta^{1/\beta})]^2}{h_\beta((x-\theta y)/\sigma\Delta^{1/\beta})}.\end{aligned}$$

Plugging this into (A.6), applying Fubini's theorem, and doing the change of variable $x \leftrightarrow (x-\theta y)/\sigma\Delta^{1/\beta}$ leads to

$$I_\Delta^{\theta\theta} \leq \frac{1}{\sigma^2\Delta^{2/\beta}}\int G_\Delta(dy)y^2\int\frac{h'_\beta(x)^2}{h_\beta(x)}\,dx. \tag{D.1}$$

Since $E(Y_\Delta^2) = m\Delta^2 + v\Delta$, we readily deduce (3.11).

E. *Proof of Theorem 4*

When $Y_t = t$ we gave $G_\Delta = \delta_\Delta$. Then (4.1) follows directly from applying the formulae (A.2) and (A.4), and from the change of variable $x \leftrightarrow (x-\theta)/\sigma\Delta^{1/\beta}$ in (A.6), after observing that the function $\breve{h}_\beta h'_\beta/h_\beta$ is integrable and odd, hence has a vanishing Lebesgue integral.

F. *Proofs of Theorems 5 and 6*

F.1. *Preliminaries*

We suppose that Y is a compound Poisson process with arrival rate λ and law of jumps μ, and that μ_k is the kth fold convolution of μ. So we have

$$G_\Delta = e^{-\lambda\Delta} \sum_{k=0}^{\infty} \frac{(\lambda\Delta)^k}{k!} \mu_k.$$

Set

$$\gamma_\Delta^{(1)}(k,x) = \frac{1}{\sigma\Delta^{1/\beta}} \int \mu_k(du) h_\beta\left(\frac{x-\theta u}{\sigma\Delta^{1/\beta}}\right), \tag{F.1}$$

$$\gamma_\Delta^{(2)}(k,x) = \frac{1}{\sigma^2\Delta^{1/\beta}} \int \mu_k(du) \check{h}_\beta\left(\frac{x-\theta u}{\sigma\Delta^{1/\beta}}\right), \tag{F.2}$$

$$\gamma_\Delta^{(3)}(k,x) = \frac{1}{\sigma^2\Delta^{2/\beta}} \int \mu_k(du) u h'_\beta\left(\frac{x-\theta u}{\sigma\Delta^{1/\beta}}\right). \tag{F.3}$$

We have (recall that $\mu_0 = \delta_0$)

$$p_\Delta(x|\sigma,\beta,\theta,G) = e^{-\lambda\Delta} \sum_{k=0}^{\infty} \frac{(\lambda\Delta)^k}{k!} \gamma_\Delta^{(1)}(k,x), \tag{F.4}$$

$$\partial_\sigma p_\Delta(x|\sigma,\beta,\theta,G) = -e^{-\lambda\Delta} \sum_{k=0}^{\infty} \frac{(\lambda\Delta)^k}{k!} \gamma_\Delta^{(2)}(k,x), \tag{F.5}$$

$$\partial_\theta p_\Delta(x|\sigma,\beta,\theta,G) = -e^{-\lambda\Delta} \sum_{k=1}^{\infty} \frac{(\lambda\Delta)^k}{k!} \gamma_\Delta^{(3)}(k,x). \tag{F.6}$$

Omitting the mention of (σ,β,θ,G), we also set

$$i = 2,3, \quad \Gamma_\Delta^{(i)}(k,k') = \int \frac{\gamma_\Delta^{(i)}(k,x)\gamma_\Delta^{(i)}(k',x)}{p_\Delta(x)}\, dx, \tag{F.7}$$

$$\Gamma_\Delta^{(4)}(k,k') = \int \frac{\gamma_\Delta^{(2)}(k,x)\gamma_\Delta^{(3)}(k',x)}{p_\Delta(x)}\, dx. \tag{F.8}$$

By the Cauchy–Schwarz inequality, we have

$$i = 2,3, \quad \left|\Gamma_\Delta^{(i)}(k,k')\right| \le \sqrt{\Gamma_\Delta^{(i)}(k,k)\Gamma_\Delta^{(i)}(k',k')}, \tag{F.9}$$

$$\left|\Gamma_\Delta^{(4)}(k,k')\right| \le \sqrt{\Gamma_\Delta^{(2)}(k,k)\Gamma_\Delta^{(3)}(k',k')}. \tag{F.10}$$

For any $k \geq 0$ we have $p_\Delta(x) \geq e^{-\lambda\Delta}((\lambda\Delta)^k)/k!\gamma_\Delta^{(1)}(k,x)$. Therefore,

$$i = 2, 3, \quad \Gamma_\Delta^{(i)}(k,k) \leq \frac{e^{\lambda\Delta}k!}{(\lambda\Delta)^k} \int \frac{\gamma_\Delta^{(i)}(k,x)^2}{\gamma_\Delta^{(1)}(k,x)}\,dx. \tag{F.11}$$

Finally, if we plug (F.5) and (F.6) into (A.6), we get

$$I_\Delta^{\sigma\theta} = e^{-2\lambda\Delta} \sum_{k=0}^\infty \sum_{l=1}^\infty \frac{(\lambda\Delta)^{k+l}}{k!l!} \Gamma_\Delta^{(4)}(k,l), \tag{F.12}$$

$$I_\Delta^{\theta\theta} = e^{-2\lambda\Delta} \sum_{k=1}^\infty \sum_{l=1}^\infty \frac{(\lambda\Delta)^{k+l}}{k!l!} \Gamma_\Delta^{(3)}(k,l). \tag{F.13}$$

F.2. *Proof of Theorem 5*

When Y is a standard Poisson process, we have $\lambda = 1$ and $\mu_k = \varepsilon_k$. Therefore, we get

$$\gamma_\Delta^{(1)}(k,x) = \frac{1}{\sigma\Delta^{1/\beta}} h_\beta\left(\frac{x-\theta k}{\sigma\Delta^{1/\beta}}\right), \tag{F.14}$$

$$\gamma_\Delta^{(2)}(k,x) = \frac{1}{\sigma^2\Delta^{1/\beta}} \breve{h}_\beta\left(\frac{x-\theta k}{\sigma\Delta^{1/\beta}}\right), \tag{F.15}$$

$$\gamma_\Delta^{(3)}(k,x) = \frac{k}{\sigma^2\Delta^{2/\beta}} h'_\beta\left(\frac{x-\theta k}{\sigma\Delta^{1/\beta}}\right). \tag{F.16}$$

Plugging this into (F.11) yields

$$\Gamma_\Delta^{(2)}(k,k) \leq \frac{e^\Delta k!}{\sigma^2\Delta^k}\mathcal{I}(\beta), \quad \Gamma_\Delta^{(3)}(k,k) \leq \frac{e^\Delta k^2 k!}{\sigma^2\Delta^{k+2/\beta}}\mathcal{J}(\beta). \tag{F.17}$$

Recall that (4.2) follows from Theorem 2, so we need to prove (4.3) and (4.4). In view of (F.12) and (F.13), this amounts to proving the two properties

$$\sum_{k=0}^\infty \sum_{l=1}^\infty \frac{\Delta^{k+l+1/\beta-1/2}}{k!l!} \Gamma_\Delta^{(4)}(k,l) \to 0,$$

$$\sum_{k=1}^\infty \sum_{l=1}^\infty \frac{\Delta^{k+l+2/\beta-1}}{k!l!} \Gamma_\Delta^{(3)}(k,l) \to \frac{1}{\sigma^2}\mathcal{J}(\beta).$$

If we use (F.9) and (F.17), it is easily seen that the sum of all summands in the first (resp. second) left side above, except the one for $k = 0$ and $l = 1$ (resp. $k = l = 1$), goes to 0. So we are left to prove

$$\Delta^{1/2+1/\beta}\Gamma_\Delta^{(4)}(0,1) \to 0, \quad \Delta^{1+2/\beta}\Gamma_\Delta^{(3)}(1,1) \to \frac{1}{\sigma^2}\mathcal{J}(\beta). \tag{F.18}$$

Let $\omega = \frac{1}{2(1+\beta)}$, so $(1+\frac{1}{\beta})(1-\omega) = 1 + 1/2\beta$. Assume first that $\beta < 2$. Then if $|x| \le |\theta/\sigma\Delta^{1/\beta}|^{\omega}$, we have for some constant $C \in (0,\infty)$, possibly depending on θ, σ, and β, that changes from one occurrence to the other and provided $\Delta \le (2\theta/\sigma)^{\beta}$,

$$h_\beta(x) \ge C\Delta^{\omega(1+1/\beta)}, \quad h_\beta\bigl(x + r\theta/\sigma\Delta^{1/\beta}\bigr) \le C\Delta^{1+1/\beta}$$

when $r \in Z\backslash\{0\}$, and thus

$$|x| \le \bigl|\theta/\sigma\Delta^{1/\beta}\bigr|^{\omega}$$
$$\Rightarrow \quad h_\beta\bigl(x + r\theta/\sigma\Delta^{1/\beta}\bigr) \le Ch_\beta(x)\Delta^{(1+1/\beta)(1-\omega)} = Ch_\beta(x)\Delta^{1+1/2\beta}.$$

When $\beta = 2$, a simple computation on the normal density shows that the above property also holds. Therefore, in view of (F.4) and (F.14) we deduce that in all cases

$$\left|\frac{x-\theta}{\sigma\Delta^{1/\beta}}\right| \le \left(\frac{\theta}{\sigma\Delta^{1/\beta}}\right)^{\omega}$$

implies

$$\begin{aligned} p_\Delta(x) &\le \frac{e^{-\Delta}}{\sigma\Delta^{1/\beta}} h_\beta\left(\frac{x-\theta}{\sigma\Delta^{1/\beta}}\right)\left(\Delta + C\Delta^{1+\beta/2}\left(1 + \sum_{k=2}^{\infty}\frac{\Delta^k}{k!}\right)\right) \\ &\le \frac{e^{-\Delta}}{\sigma\Delta^{1/\beta-1}} h_\beta\left(\frac{x-\theta}{\sigma\Delta^{1/\beta}}\right)\bigl(1 + C\Delta^{\beta/2}\bigr). \end{aligned}$$

By (F.7) it follows that

$$\begin{aligned} \Gamma_\Delta^{(3)}(1,1) &\ge \frac{e^{\Delta}}{\sigma^3\Delta^{1+3/\beta}} \int_{\{|(x-\theta)/\sigma\Delta^{1/\beta}| \le (\theta/\sigma\Delta^{1/\beta})^{\omega}\}} \frac{h'_\beta((x-\theta)/\sigma\Delta^{1/\beta})^2}{h_\beta((x-\theta)/\sigma\Delta^{1/\beta})}\,dx \\ &\ge \frac{e^{\Delta}}{\sigma^2\Delta^{1+2/\beta}} \int_{\{|x| \le (\theta/\sigma\Delta^{1/\beta})^{\omega}\}} \frac{h'_\beta(x)^2}{h_\beta(x)}\,dx. \end{aligned}$$

We readily deduce that $\liminf_{\Delta\to 0} \Delta^{1+2/\beta}\Gamma_\Delta^{(3)}(1,1) \ge \mathcal{J}(\beta)/\sigma^2$. On the other hand, (F.17) yields $\limsup_{\Delta\to 0} \Delta^{1+2/\beta}\Gamma_\Delta^{(3)}(1,1) \le \mathcal{J}(\beta)/\sigma^2$, and thus the second part of (F.18) is proved.

Finally $\breve{h}_\beta/h_\beta$ is bounded, so (F.15), (F.16), and the fact that

$$p_\Delta(x) \ge e^{-\Delta} h_\beta\bigl(x/\sigma\Delta^{1/\beta}\bigr)/\sigma\Delta^{1/\beta}$$

yield

$$\bigl|\Gamma_\Delta^{(4)}(0,1)\bigr| \le \frac{e^{\Delta}}{\sigma^3\Delta^{2/\beta}} \int \left|h'_\beta\left(\frac{x-\theta}{\sigma\Delta^{1/\beta}}\right)\right| dx \le \frac{e^{\Delta}}{\sigma^2\Delta^{1/\beta}} \int |h'_\beta(x)|\,dx$$

and the first part of (F.18) readily follows.

F.3. *Proof of Theorem 6*

We use the same notation as in the previous proof, but here the measure μ_k has a density f_k for all $k \geq 1$, which further is differentiable and satisfies (4.5) uniformly in k, while we still have $\mu_0 = \delta_0$. Exactly as in (3.3), we set

$$\breve{f}_k(u) = uf_k'(u) + f_k(u).$$

Recall the Fisher's information $\mathcal{L}$ defined in (4.6), which corresponds to estimating θ in a model where one observes a variable θU, with U having the law μ. Now if we have n independent variables U_i with the same law μ, the Fisher's information associated with the observation of θU_i for $i = 1, \dots, n$ is of course $n\mathcal{L}$; if instead one observes only $\theta(U_1 + \cdots + U_n)$, one gets a smaller Fisher's information $\mathcal{L}_n \leq n\mathcal{L}$. In other words, we have

$$\text{(F.19)}\qquad \mathcal{L}_n := \int \frac{\breve{f}_n(u)^2}{f_n(u)}\,du \leq n\mathcal{L}.$$

Taking advantage of the fact that μ_k has a density, for all $k \geq 1$, we can rewrite $\gamma_\Delta^{(i)}(k,x)$ as (using further an integration by parts when $i = 3$ and the fact that each f_k satisfies (4.5))

$$\text{(F.20)}\qquad \gamma_\Delta^{(1)}(k,x) = \frac{1}{\theta}\int h_\beta(y) f_k\left(\frac{x - y\sigma\Delta^{1/\beta}}{\theta}\right)dy,$$

$$\text{(F.21)}\qquad \gamma_\Delta^{(2)}(k,x) = \frac{1}{\sigma\theta}\int \breve{h}_\beta(y) f_k\left(\frac{x - y\sigma\Delta^{1/\beta}}{\theta}\right)dy,$$

$$\text{(F.22)}\qquad \gamma_\Delta^{(3)}(k,x) = \frac{1}{\theta^2}\int h_\beta(y) \breve{f}_k\left(\frac{x - y\sigma\Delta^{1/\beta}}{\theta}\right)dy.$$

Since the f_k's satisfy (4.5) uniformly in k, we readily deduce that

$$\text{(F.23)}\qquad k \geq 1, i = 1,2,3 \quad\Rightarrow\quad \left|\gamma_\Delta^{(i)}(k,x)\right| \leq C,$$

$$\text{(F.24)}\qquad \gamma_\Delta^{(1)}(1,x) \to \frac{1}{\theta} f\left(\frac{x}{\theta}\right), \quad \gamma_\Delta^{(3)}(1,x) \to \frac{1}{\theta^2}\breve{f}\left(\frac{x}{\theta}\right).$$

Let us start with the lower bound. Since $\gamma_\Delta^{(1)}(0,x) = (1/\sigma\Delta^{1/\beta})h_\beta(x/\sigma\Delta^{1/\beta})$, we deduce from (2.2), (F.23), and (F.4) that

$$\text{(F.25)}\qquad \frac{1}{\Delta}p_\Delta(x) \to \frac{c_\beta\sigma^\beta}{|x|^{1+\beta}} + \frac{\lambda}{\theta}f\left(\frac{x}{\theta}\right)$$

as soon as $x \neq 0$, and with the convention $c_2 = 0$. In a similar way, we deduce from (F.23) and (F.6) that

$$\text{(F.26)} \qquad \frac{1}{\Delta}\partial_\theta p_\Delta(x) \to -\frac{\lambda}{\theta^2}\breve{f}\left(\frac{x}{\theta}\right).$$

Then plugging (F.25) and (F.26) into the last equation in (A.6), we conclude by Fatou's lemma and after a change of variable that

$$\liminf_{\Delta\to 0}\frac{I_\Delta^{\theta\theta}}{\Delta} \geq \frac{\lambda^2}{\theta^2}\int \frac{\breve{f}(x)^2}{\lambda f(x) + c_\beta\sigma^\beta/\theta^\beta|x|^{1+\beta}}\,dx.$$

It remains to prove (4.8) and the upper bound in (4.9) (including when $\beta = 2$). By the Cauchy–Schwarz inequality, we get (using successively the two equivalent versions for $\gamma_\Delta^{(i)}(k,x)$)

$$\gamma_\Delta^{(2)}(k,x)^2 \leq \frac{1}{\sigma^3}\gamma_\Delta^{(1)}(k,x)\int \mu_k(du)\widetilde{h}_\beta\big((x-\theta u)/\sigma\Delta^{1/\beta}\big),$$

$$\gamma_\Delta^{(3)}(k,x)^2 \leq \frac{1}{\theta^3}\gamma_\Delta^{(1)}(k,x)\int h_\beta(y)\frac{\breve{f}_k((x-y\sigma\Delta^{1/\beta})/\theta)^2}{f_k((x-y\sigma\Delta^{1/\beta})/\theta)}\,dy.$$

Then it follows from (F.11) and (F.19) that

$$\text{(F.27)} \qquad \Gamma_\Delta^{(2)}(k,k) \leq \frac{e^{\lambda\Delta}k!}{\sigma^2(\lambda\Delta)^k}\mathcal{I}(\beta), \qquad \Gamma_\Delta^{(3)}(k,k) \leq \frac{e^{\lambda\Delta}kk!}{\theta^2(\lambda\Delta)^k}\mathcal{L}.$$

We also need an estimate for $\Gamma_\Delta^{(4)}(0,1)$. By (F.1), (F.2), and (F.4) we obtain $p_\Delta(x) \geq e^{-\lambda\Delta}h_\beta(x/\sigma\Delta^{1/\beta})/\sigma\Delta^{1/\beta}$ and $\gamma_\Delta^{(2)}(0,x) = \breve{h}_\beta(x/\sigma\Delta^{1/\beta})/\sigma^2\Delta^{1/\beta}$. Then use (F.22) and the definition (F.7) to get

$$\text{(F.28)} \qquad \begin{aligned} \big|\Gamma_\Delta^{(4)}(0,1)\big| &\leq \frac{1}{\sigma\theta^2}\int\left|\frac{\breve{h}_\beta}{h_\beta}\left(\frac{x}{\sigma\Delta^{1/\beta}}\right)\right|dx\int h_\beta(y)\left|\breve{f}\left(\frac{x-y\sigma\Delta^{1/\beta}}{\theta}\right)\right|dy \\ &= \frac{1}{\sigma\theta^2}\int h_\beta(y)\,dy\int\left|\frac{\breve{h}_\beta}{h_\beta}\left(\frac{x}{\sigma\Delta^{1/\beta}}\right)\breve{f}\left(\frac{x-y\sigma\Delta^{1/\beta}}{\theta}\right)\right|dx \\ &\leq C, \end{aligned}$$

where the last inequality comes from the facts that $\breve{h}_\beta/h_\beta$ is bounded and that $\breve{f}$ is integrable (due to (4.5)).

At this stage we use (F.9) together with (F.12) and (F.13), and the fact that $2|xy| \leq ax^2 + y^2/a$ for all $a > 0$. Taking arbitrary constants $a_{kl} > 0$, we deduce from (F.27) that

$$I_\Delta^{\theta\theta} \leq \frac{\mathcal{L}}{\theta^2}e^{-\lambda\Delta}\left(\sum_{k=1}^\infty\frac{(\lambda\Delta)^k k}{k!} + 2\sum_{k=1}^\infty\sum_{l=k+1}^\infty\frac{(\lambda\Delta)^{k+l}}{k!l!}\left(\frac{kk!}{(\lambda\Delta)^k}\frac{ll!}{(\lambda\Delta)^l}\right)^{1/2}\right)$$

$$\leq \frac{\mathcal{L}}{\theta^2}\left(\lambda\Delta + \sum_{k=1}^{\infty}\sum_{l=k+1}^{\infty}\frac{(\lambda\Delta)^l k}{l!}a_{kl} + \sum_{k=1}^{\infty}\sum_{l=k+1}^{\infty}\frac{(\lambda\Delta)^k l}{k!}\frac{1}{a_{kl}}\right).$$

Then if we take $a_{kl} = (\lambda\Delta)^{(k-l)/2}$ for $l > k$, a simple computation shows that indeed

$$I_\Delta^{\theta\theta} \leq \frac{\mathcal{L}}{\theta^2}\big(\lambda\Delta + C\Delta^{3/2}\big)$$

for some constant C, and thus we get the upper bound in (4.9). In a similar way, replacing $\mathcal{L}/\theta^2$ above by the supremum between $\mathcal{L}/\theta^2$ and $\mathcal{I}(\beta)/\sigma^2$, we see that in (F.12) the sum of the absolute values of all summands except the one for $k=0$ and $l=1$ is smaller than a constant times $\Lambda\Delta$. Finally, the same holds for the term for $k=0$ and $l=1$, because of (F.28), and this proves (4.8).

G. *Proof of Theorem 7*

G.1. *Preliminaries*

In the setting of Theorem 7, the measure G_Δ admits the density $y \mapsto h_\alpha(y/\Delta^{1/\alpha})/\Delta^{1/\alpha}$. For simplicity, we set

$$u = \Delta^{1/\alpha - 1/\beta} \tag{G.1}$$

(so $u \to 0$ as $\Delta \to 0$), and a change of variable allows to write (A.1) as

$$p_\Delta(x|\sigma,\beta,\theta,G) = \frac{1}{\theta\Delta^{1/\alpha}}\int h_\beta\left(\frac{x}{\sigma\Delta^{1/\beta}} - y\right)h_\alpha\left(\frac{\sigma y}{\theta u}\right)dy.$$

Therefore,

$$\partial_\theta p_\Delta(x|\sigma,\beta,\theta,G) = -\frac{1}{\theta^2\Delta^{1/\alpha}}\int h_\beta\left(\frac{x}{\sigma\Delta^{1/\beta}} - y\right)\breve{h}_\alpha\left(\frac{\sigma y}{\theta u}\right)dy$$

and another change of variable in (A.6) leads to

$$I_\Delta^{\theta\theta} = \frac{\theta^{2\alpha-2}u^{2\alpha}}{\sigma^{2\alpha}}J_u, \tag{G.2}$$

where

$$J_u = \int \frac{R_u(x)^2}{S_u(x)}\,dx, \tag{G.3}$$

$$R_u(x) = \left(\frac{\sigma}{\theta u}\right)^{1+\alpha}\int h_\beta(x-y)\breve{h}_\alpha\left(\frac{\sigma y}{\theta u}\right)dy$$

$$= \left(\frac{\sigma}{\theta u}\right)^{\alpha} \int h_\beta\left(x - \frac{y\theta u}{\sigma}\right) \breve{h}_\alpha(y)\, dy,$$

$$S_u(x) = \frac{\sigma}{\theta u} \int h_\beta(x-y) h_\alpha\left(\frac{\sigma y}{\theta u}\right) dy = \int h_\beta\left(x - \frac{y\theta u}{\sigma}\right) h_\alpha(y)\, dy.$$

Below, we denote by K a constant and by ϕ a continuous function on R_+ with $\phi(0) = 0$, both of them changing from line to line and possibly depending on the parameters α, β, σ, θ; if they depend on another parameter η, we write them as K_η or ϕ_η. Recalling (2.2), we have, as $|x| \to \infty$,

$$\text{(G.4)} \qquad h_\alpha(x) \sim \frac{c_\alpha}{|x|^{1+\alpha}}, \quad \int_{\{|y|>|x|\}} h_\alpha(y)\, dy \sim \frac{2c_\alpha}{\alpha |x|^\alpha},$$
$$\breve{h}_\alpha(x) \sim -\frac{\alpha c_\alpha}{|x|^{1+\alpha}}, \quad \int_{\{|y|>|x|\}} \breve{h}_\alpha(y)\, dy \sim -\frac{2c_\alpha}{|x|^\alpha}.$$

Another application of (2.2) when $\beta < 2$ and of the explicit form of h_2 gives

$$\text{(G.5)} \qquad |y| \le 1 \quad \Longrightarrow \quad |h''_\beta(x-y)| \le \overline{h}_\beta(x) := \begin{cases} K h_\beta(x), & \text{if } \beta < 2, \\ K(1+x^2)e^{-x^2/3}, & \text{if } \beta = 2. \end{cases}$$

We will now introduce an auxiliary function

$$\text{(G.6)} \qquad D_\eta(x) = \int_{\{|y|>\eta\}} h_\beta(x-y) \frac{1}{|y|^{1+\alpha}}\, dy,$$

which will help us control the behavior of R_u and S_u, as stated in the following lemma.

LEMMA 1: *We have*

$$\text{(G.7)} \qquad \left| S_u(x) - \left(h_\beta(x) + c_\alpha \frac{(\theta u)^\alpha}{\sigma^\alpha} D_1(x)\right)\right|$$
$$\le (h_\beta(x) + u^\alpha D_1(x))\phi(u) + K u^\alpha \overline{h}_\beta(x),$$
$$\left| R_u(x) - c_\alpha\left(\frac{2h_\beta(x)}{\eta^\alpha} - \alpha D_\eta(x)\right)\right|$$
$$\le \left(\frac{h_\beta(x)}{\eta^\alpha} + D_\eta(x)\right)\phi_\eta(u) + K\overline{h}_\beta(x)\eta^{2-\alpha}$$

with

$$\text{(G.8)} \qquad D_\eta(x) \sim \frac{1}{|x|^{1+\alpha}} \quad \textit{as} \quad |x| \to \infty.$$

PROOF: We split the first two integrals in R_u and S_u into sums of integrals on the two domains $\{|y| \le \eta\}$ and $\{|y| > \eta\}$ for some $\eta \in (0,1]$ to be chosen later. We have $|h_\beta(x-y) - h_\beta(x) + h'_\beta(x)y| \le \overline{h}_\beta(x)y^2$ as soon as $|y| \le 1$, so the fact that both $f = \breve{h}_\alpha$ and $f = h_\alpha$ are even functions gives

$$\left|\int_{\{|y|\le\eta\}} h_\beta(x-y) f\left(\frac{\sigma y}{\theta u}\right) dy - h_\beta(x) \int_{\{|y|\le\eta\}} f\left(\frac{\sigma y}{\theta u}\right) dy\right|$$
$$\le \overline{h}_\beta(x) \int_{\{|y|\le\eta\}} y^2 \left| f\left(\frac{\sigma y}{\theta u}\right)\right| dy.$$

On the one hand we have with $f = \breve{h}_\alpha$ or $f = h_\alpha$, and in view of (G.4),

$$\int_{\{|y|\le\eta\}} y^2 \left| f\left(\frac{\sigma y}{\theta u}\right)\right| dy = \frac{\theta^3 u^3}{\sigma^3} \int_{\{|z|\le\sigma\eta/\theta u\}} z^2 |f(z)|\, dz \le K u^{1+\alpha} \eta^{2-\alpha}.$$

On the other hand, the integrability of h_α and $\breve{h}_\alpha$, and the fact that $\int \breve{h}_\alpha(y)\, dy = 0$ yield

$$\int_{\{|y|\le 1\}} h_\alpha\left(\frac{\sigma y}{\theta u}\right) dy = \frac{\theta u}{\sigma} \int_{\{|z|\le\sigma/\theta u\}} h_\alpha(z)\, dz = \frac{\theta u}{\sigma}(1 + \phi(u)),$$
$$\int_{\{|y|\le\eta\}} \breve{h}_\alpha\left(\frac{\sigma y}{\theta u}\right) dy = \frac{\theta u}{\sigma} \int_{\{|z|\le\sigma\eta/\theta u\}} \breve{h}_\alpha(z)\, dz$$
$$= -\frac{\theta u}{\sigma} \int_{\{|z|>\sigma\eta/\theta u\}} \breve{h}_\alpha(z)\, dz$$
$$= 2c_\alpha \frac{(\theta u)^{1+\alpha}}{\sigma^{1+\alpha}\eta^\alpha}(1 + \phi_\eta(u)).$$

Putting all these facts together yields

$$\left|\int_{\{|y|\le 1\}} h_\beta(x-y) h_\alpha\left(\frac{\sigma y}{\theta u}\right) dy - \frac{\theta u}{\sigma} h_\beta(x)\right| \le u\phi(u) h_\beta(x) + K u^{1+\alpha} \overline{h}_\beta(x) \tag{G.9}$$

and

$$\left|\int_{\{|y|\le\eta\}} h_\beta(x-y) \breve{h}_\alpha\left(\frac{\sigma y}{\theta u}\right) dy - 2c_\alpha \frac{(\theta u)^{1+\alpha}}{\sigma^{1+\alpha}\eta^\alpha} h_\beta(x)\right| \le u^{1+\alpha}\left(h_\beta(x)\frac{\phi_\eta(u)}{\eta^\alpha} + K\overline{h}_\beta(x)\eta^{2-\alpha}\right). \tag{G.10}$$

For the integrals on $\{|y| > \eta\}$ we observe that by (G.4) we have

$$\left|h_\alpha(\sigma y/\theta u) - c_\alpha(\theta u/\sigma|y|)^{1+\alpha}\right| \le (\theta u/\sigma y)^{1+\alpha}\phi(u),$$

and the same for $\breve{h}_\alpha$ except that c_α is substituted with $-\alpha c_\alpha$. Given the definition of D_η, we readily get

$$\text{(G.11)}\quad \begin{aligned} &\left|\int_{\{|y|>1\}} h_\beta(x-y)h_\alpha\left(\frac{\sigma y}{\theta u}\right)dy - c_\alpha\frac{(\theta u)^{1+\alpha}}{\sigma^{1+\alpha}}D_1(x)\right| \\ &\quad \le D_1(x)u^{1+\alpha}\phi(u), \\ &\left|\int_{\{|y|>\eta\}} h_\beta(x-y)\breve{h}_\alpha\left(\frac{\sigma y}{\theta u}\right)dy + \alpha c_\alpha\frac{(\theta u)^{1+\alpha}}{\sigma^{1+\alpha}}D_\eta(x)\right| \\ &\quad \le D_\eta(x)u^{1+\alpha}\phi(u), \end{aligned}$$

and if we put together (G.9) and (G.11), we obtain (G.7).

Finally, we prove (G.8). We split the integral in (G.6) into the sum of the integrals, say $D_\eta^{(1)}$ and $D_\eta^{(2)}$, over the two domains $\{|y| > \eta, |y-x| \le |x|^\gamma\}$ and $\{|y| > \eta, |y-x| > |x|^\gamma\}$, where $\gamma = 4/5$ if $\beta = 2$ and $\gamma \in ((1+\alpha)/(1+\beta), 1)$ if $\beta < 2$. On the one hand, $D_\eta^{(2)}(x) \le K_\eta h_\beta(|x|^\gamma)$, so with our choice of γ we obviously have $|x|^{1+\alpha}D_\eta^{(2)}(x) \to 0$. On the other hand, $|x|^{1+\alpha}D_\eta^{(1)}(x)$ is clearly equivalent, as $|x| \to \infty$, to $\int_{\{|y-x|\le|x|^\gamma\}} h_\beta(x-y)\,dy$, which equals $\int_{\{|z|\le|x|^\gamma\}} h_\beta(z)\,dz$, which in turn goes to 1. Hence we get (G.8) for all $\eta > 0$. This concludes the proof of the lemma. *Q.E.D.*

Using the lemma, we can now obtain the behavior of R_u and S_u as $u \to 0$. First, an application of (G.4) to the last formula in (G.3) and Lebesgue's theorem readily give

$$\text{(G.12)}\quad \lim_{u\to 0} S_u(x) = h_\beta(x).$$

Also, by (G.4), (G.5), (G.7), and (G.8), we get

$$\text{(G.13)}\quad S_u(x) \ge \begin{cases} \dfrac{C}{1+|x|^{1+\beta}}, & \text{if } \beta < 2, \\ C\left(e^{-x^2/2} + \dfrac{u^\alpha}{1+|x|^{1+\alpha}}\right), & \text{if } \beta = 2, \end{cases}$$

for some $C > 0$ depending on the parameters and all u small enough. In the same way, we see that $u^\alpha R_u(x) \to 0$, but this not enough. However, if $r_\eta = 2h_\beta/\eta^\alpha - \alpha D_\eta$, we deduce from (G.7) that for all $\eta \in (0, 1]$,

$$\text{(G.14)}\quad \limsup_{u\to 0} |R_u(x) - c_\alpha r_\eta(x)| \le K\overline{h}_\beta(x)\eta^{2-\alpha}.$$

A simple computation and the second order Taylor expansion with integral remainder for h_β yield

$$\begin{aligned} r_\eta(x) &= \alpha \int_{\{|y|>\eta\}} \frac{h_\beta(x) - h_\beta(x-y)}{|y|^{1+\alpha}}\,dy \\ &= -\alpha \int_{\{|y|>\eta\}} |y|^{1-\alpha}\,dy \int_0^1 (1-v) h''_\beta(x-yv)\,dv. \end{aligned}$$

By Lebesgue's theorem, r_η converges pointwise as $\eta \to 0$ to the function r given by

$$r(x) = -\alpha \int_{\mathbb{R}} |y|^{1-\alpha}\,dy \int_0^1 (1-v) h''_\beta(x-yv)\,dv.$$

Then, taking into account (G.14) and using once more (G.7) together with (G.4), (G.5), and (G.8), and also $\alpha < \beta$, we get

$$\text{(G.15)} \qquad \lim_{u\to 0} R_u(x) = c_\alpha r(x), \qquad |R_u(x)| \le \frac{K}{1+|x|^{1+\alpha}}.$$

G.2. *The Case (b) in Theorem 7*

We are now in a position to prove (4.12), where $\beta < 2$ and $\alpha > \beta/2$ (and of course $\alpha < \beta$). Since $\alpha > \beta/2$, we see that (G.12), (G.13), and (G.15) allow us to apply Lebesgue's theorem in the definition (G.3) to get that $J_u \to \int dx\,(c_\alpha r(x))^2/h_\beta(x)$: this is (4.12) (obviously $|r(x)| \le K/(1+|x|^{1+\alpha})$, while $h_\beta(x) > C/(1+|x|^{1+\beta})$ for some $C > 0$; hence the integral in (4.12) converges).

G.3. *The Other Cases*

The other cases are a bit more involved, because Lebesgue's theorem does not apply and we will see that J_u goes to infinity. First, we introduce the functions

$$R'(x) = \frac{\alpha c_\alpha}{|x|^{1+\alpha}}, \qquad S'_u(x) = \begin{cases} \dfrac{c_\beta}{|x|^{1+\beta}} + \dfrac{c_\alpha \theta^\alpha u^\alpha}{\sigma^\alpha |x|^{1+\alpha}}, & \text{if } \beta < 2, \\[2ex] \dfrac{e^{-x^2/2}}{\sqrt{2\pi}} + \dfrac{c_\alpha \theta^\alpha u^\alpha}{\sigma^\alpha |x|^{1+\alpha}}, & \text{if } \beta = 2. \end{cases}$$

Below we denote by $\psi(u,\Gamma)$ for $u \in (0,1]$ and $\Gamma \ge 1$ the sum $\phi'(u) + \phi''(1/\Gamma)$ for any two functions ϕ' and ϕ'', changing from line to line, that are continuous on R_+ with $\phi'(0) = \phi''(0) = 0$. We deduce from (G.4), (G.5), (G.7) for $\eta = 1$, and (G.8) that

$$\text{(G.16)} \quad |x| > \Gamma \\ \Rightarrow \quad |R_u(x) + R'(x)| \le \psi(u, \Gamma)R'(x) + \begin{cases} \dfrac{K}{|x|^{1+\beta}}, & \text{if } \beta < 2, \\ Kx^2 e^{-x^2/3}, & \text{if } \beta = 2. \end{cases}$$

In a similar way, we get

$$|x| > \Gamma \\ \Rightarrow \quad |S_u(x) - S'_u(x)| \le \begin{cases} \psi(u, \Gamma)S'_u(x), & \text{if } \beta < 2, \\ \psi(u, \Gamma)S'_u(x) + K_0 u^\alpha x^2 e^{-x^2/3}, & \text{if } \beta = 2, \end{cases}$$

where K_0 is some constant. Then for any $\varphi > 0$, we denote by Γ_φ the smallest number bigger than 1, such that $K_0 x^2 e^{-x^2/3} \le \varphi \frac{c_\alpha \theta^\alpha}{\sigma^\alpha |x|^{1+\alpha}}$ for all $|x| > \Gamma_\varphi$. The last estimate above for $\beta = 2$ reads as $|S_u - S'_u| \le S'_u(\psi(u, \Gamma) + \varphi)$, so in all cases we have, for some fixed function ψ_0 as above,

$$\text{(G.17)} \quad S_u(x) = S'_u(x)(1 + \rho_u(x)),$$

where

$$|\rho_u(x)| \le \begin{cases} \psi_0(u, \Gamma), & \text{if } \beta < 2, |x| > \Gamma, \\ \psi_0(u, \Gamma) + \varphi, & \text{if } \beta = 2, |x| > \Gamma > \Gamma_0. \end{cases}$$

At this stage, we set

$$\text{(G.18)} \quad J_{u,\Gamma} = \int_{\{|x|>\Gamma\}} \frac{R'(x)^2}{S'_u(x)}\, dx.$$

Observe that $J_u = J_{u,\Gamma} + \sum_{i=1}^4 J_{u,\Gamma}^{(i)}$, where

$$J_{u,\Gamma}^{(1)} = \int_{\{|x|\le\Gamma\}} \frac{R_u(x)^2}{S_u(x)}\, dx,$$

$$J_{u,\Gamma}^{(2)} = \int_{\{|x|>\Gamma\}} \frac{(R_u(x) + R'(x))^2}{S_u(x)}\, dx,$$

$$J_{u,\Gamma}^{(3)} = -2\int_{\{|x|>\Gamma\}} \frac{R'(x)(R_u(x) + R'(x))}{S_u(x)}\, dx,$$

$$J_{u,\Gamma}^{(4)} = \int_{\{|x|>\Gamma\}} \frac{R'(x)^2}{S_u(x)}\, dx - \int_{\{|x|>\Gamma\}} \frac{R'(x)^2}{S'_u(x)}\, dx.$$

In light of (G.13), (G.15), and (G.16), we get, for some $u_0 > 0$,

$$\sup_{u\in(0,u_0]} J^{(1)}_{u,\Gamma} < \infty, \qquad u \in (0, u_0]$$
$$\Rightarrow \quad J^{(2)}_{u,\Gamma} \le K + \psi(u,\Gamma)\big(J^{(4)}_{u,\Gamma} + J_{u,\Gamma}\big).$$

The Cauchy–Schwarz inequality yields

$$\big|J^{(3)}_{u,\Gamma}\big| \le 2\big(J^{(2)}_{u,\Gamma}\big(J^{(4)}_{u,\Gamma} + J_{u,\Gamma}\big)\big)^{1/2}.$$

Finally, (G.13) and (G.17), and the definition of R' yield (with ψ_0 as in (G.17))

$$\big|J^{(4)}_{u,\Gamma}\big| \le \begin{cases} 2\psi_0(u,\Gamma)J_{u,\Gamma}, & \text{if } \beta < 2,\ \psi_0(u,\Gamma) < 1/2, \\ 2(\psi_0(u,\Gamma)+\varphi)J_{u,\Gamma}, & \text{if } \beta = 2,\ \Gamma > \Gamma_\varphi,\ \psi_0(u,\Gamma)+\varphi < 1/2. \end{cases}$$

Therefore, we get

$$\left|\frac{J_u}{J_{u,\Gamma}} - 1\right| \le \frac{K}{J_{u,\Gamma}} + \psi(u,\Gamma) + 2(\psi_0(u,\Gamma)+\varphi)$$

as soon as $\psi_0(u,\Gamma)+\varphi < 1/2$ and $\Gamma > \Gamma_\varphi$, and with the convention that $\varphi = 0$ and $\Gamma_0 = 1$ when $\beta < 2$. Then, remembering that $\lim_{u\to 0}\lim_{\Gamma\to\infty}\psi(u,\Gamma) = 0$, and the same for ψ_0, and that $\varphi = 0$ when $\beta < 2$ and φ is arbitrarily small when $\beta = 2$, we readily deduce the following fact: suppose that for some function $u \mapsto \gamma(u)$ going to $+\infty$ as $u \to 0$ and independent of Γ we have proved that

$$\text{(G.19)} \qquad J_{u,\Gamma} \sim \gamma(u) \quad \text{as} \quad u \to 0\ \forall \Gamma > 1.$$

Then $J_u \sim \gamma(u)$ and, therefore, by (G.2) we get

$$\text{(G.20)} \qquad I^{\theta\theta}_\Delta \sim \frac{\theta^{2\alpha-2}u^{2\alpha}\gamma(u)}{\sigma^{2\alpha}}.$$

G.3.1. *The Case (d) in Theorem 7:* Now let $\beta < 2$ and $\alpha < \beta/2$. The change of variables $z = xu^{\alpha/(\beta-\alpha)}$ in (G.18) yields

$$\text{(G.21)} \qquad J_{u,\Gamma} = \alpha^2 c_\alpha^2 \int_{\{|x|>\Gamma\}} \frac{1}{c_\beta|x|^{1+2\alpha-\beta} + c_\alpha\theta^\alpha u^\alpha|x|^{1+\alpha}/\sigma^\alpha}\,dx$$
$$= \frac{\alpha^2 c_\alpha^2}{u^{(\alpha(\beta-2\alpha))/(\beta-\alpha)}} \int_{\{|z|>\Gamma u^{\alpha/(\beta-\alpha)}\}} \frac{1}{c_\beta|z|^{1+2\alpha-\beta} + c_\alpha\theta^\alpha|z|^{1+\alpha}/\sigma^\alpha}\,dz$$

and we have (G.19) with

$$\text{(G.22)}\qquad \gamma(u)=\frac{\alpha^2c_\alpha^2}{u^{(\alpha(\beta-2\alpha))/(\beta-\alpha)}}\int\frac{1}{c_\beta|z|^{1+2\alpha-\beta}+c_\alpha\theta^\alpha|z|^{1+\alpha}/\sigma^\alpha}\,dz.$$

So if we combine this with (G.20), we get (4.14).

G.3.2. *The Case (c) in Theorem 7:* Suppose now that $2\alpha=\beta<2$. Then

$$\text{(G.23)}\qquad J_{u,\Gamma}=2\alpha^2c_\alpha^2\int_\Gamma^\infty\frac{1}{x(c_\beta+c_\alpha\theta^\alpha u^\alpha x^\alpha/\sigma^\alpha)}\,dx.$$

For $v>0$ we let H_v be the unique point $x>0$ such that $c_\alpha\theta^\alpha v^\alpha x^\alpha/\sigma^\alpha=c_\beta$, so in fact $H_v=\rho/v$ for some $\rho>0$. We have

$$J_{u,\Gamma}\le\frac{2\alpha^2c_\alpha^2}{c_\beta}\int_\Gamma^{H_u}\frac{1}{x}\,dx+\frac{K}{u^\alpha}\int_{H_u}^\infty\frac{1}{x^{1+\alpha}}\,dx\le\frac{2\alpha^2c_\alpha^2}{c_\beta}\log\left(\frac{1}{u}\right)+K.$$

On the other hand, if $\mu>0$ and $\Gamma<x<H_{u\mu}$, we have $c_\beta+c_\alpha\theta^\alpha u^\alpha x^\alpha/\sigma^\alpha<c_\beta(1+1/\mu^\alpha)$, hence

$$J_{u,\Gamma}>\frac{2\alpha^2c_\alpha^2}{c_\beta}\int_\Gamma^{H_{u\varphi}}\frac{1}{x(1+1/\mu^\alpha)}\,dx\ge\frac{2\alpha^2c_\alpha^2}{c_\beta}\frac{1}{1+1/\mu^\alpha}\log\left(\frac{1}{u}\right)-K_\varphi.$$

Putting together these two estimates and choosing μ big gives (G.19) with

$$\gamma(u)=\frac{2\alpha^2c_\alpha^2}{c_\beta}\log\left(\frac{1}{u}\right)$$

and we readily deduce (4.13).

G.3.3. *The Case (a) in Theorem 7:* Finally, consider the case $\beta=2$. We have

$$\text{(G.24)}\qquad J_{u,\Gamma}=2\alpha^2c_\alpha^2\int_\Gamma^\infty\frac{1}{x^{1+\alpha}(x^{1+\alpha}e^{-x^2/2}/\sqrt{2\pi}+c_\alpha\theta^\alpha u^\alpha/\sigma^\alpha)}\,dx.$$

Suppose that Γ is large enough for $x\mapsto x^{2+2\alpha}e^{-x^2/2}$ to be decreasing on $[\Gamma,\infty)$. For $v>0$ small enough, there is a unique number $H_v=x>\Gamma$ such that $c_\alpha\theta^\alpha v^\alpha/\sigma^\alpha=x^{1+\alpha}e^{-x^2/2}/\sqrt{2\pi}$, so in fact $H_v\sim\sqrt{2\alpha\log(1/v)}$ when $v\to0$. Then

$$\begin{aligned}\text{(G.25)}\qquad J_{u,\Gamma}&\le K\int_\Gamma^{H_u}\frac{e^{x^2/2}}{x^{2+2\alpha}}\,dx+\frac{2\alpha^2c_\alpha\sigma^\alpha}{\theta^\alpha u^\alpha}\int_{H_u}^\infty\frac{1}{x^{1+\alpha}}\,dx\\&\le\frac{K}{H_u^\alpha}+\frac{2\alpha c_\alpha\sigma^\alpha}{\theta^\alpha u^\alpha H_u^\alpha}\le\frac{2\alpha c_\alpha\sigma^\alpha}{\theta^\alpha u^\alpha(2\alpha)^{\alpha/2}(\log(1/u))^{\alpha/2}}(1+o(1)).\end{aligned}$$

On the other hand, if $\mu > 1$ and $x > H_{u\mu}$, we have $x^{1+\alpha}e^{-x^2/2}/\sqrt{2\pi} + c_\alpha\theta^\alpha u^\alpha/\sigma^\alpha < +c_\alpha\theta^\alpha u^\alpha(1+\mu^\alpha)/\sigma^\alpha$, hence,

$$J_{u,\Gamma} > \frac{2\alpha^2 c_\alpha \sigma^\alpha}{\theta^\alpha u^\alpha}\int_\Gamma^{H_{u\varphi}} \frac{1}{x^{1+\alpha}(1+\mu^\alpha)}\,dx \geq \frac{2\alpha c_\alpha \sigma^\alpha}{\theta^\alpha u^\alpha (2\alpha)^{\alpha/2}(\log(1/u))^{\alpha/2}(1+\mu^\alpha)}(1+o(1)). \tag{G.26}$$

So again we see, by choosing μ close to 1, that the desired result holds with

$$\gamma(u) = \frac{2\alpha c_\alpha \sigma^\alpha}{\theta^\alpha u^\alpha (2\alpha)^{\alpha/2}(\log(1/u))^{\alpha/2}} \tag{G.27}$$

and we deduce (4.11).

REFERENCES

AÏT-SAHALIA, Y. (2004): "Disentangling Diffusion From Jumps," *Journal of Financial Economics*, 74, 487–528.

AÏT-SAHALIA, Y., AND J. JACOD (2007): "Volatility Estimators for Discretely Sampled Lévy Processes," *The Annals of Statistics*, 35, 355–392.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2004): "Power and Bipower Variation With Stochastic Volatility and Jumps" (With Discussion), *Journal of Financial Econometrics*, 2, 1–48.

BERGSTRØM, H. (1952): "On Some Expansions of Stable Distribution Functions," *Arkiv für Mathematik*, 2, 375–378.

BROCKWELL, P., AND B. BROWN (1980): "High-Efficiency Estimation for the Positive Stable Laws," *Journal of the American Statistical Association*, 76, 626–631.

DUMOUCHEL, W. H. (1971): "Stable Distributions in Statistical Inference," Ph.D. Thesis, Department of Statistics, Yale University.

——— (1973a): "On the Asymptotic Normality of the Maximum-Likelihood Estimator When Sampling From a Stable Distribution," *The Annals of Statistics*, 1, 948–957.

——— (1973b): "Stable Distributions in Statistical Inference: 1. Symmetric Stable Distributions Compared to Other Symmetric Long-Tailed Distributions," *Journal of the American Statistical Association*, 68, 469–477.

——— (1975): "Stable Distributions in Statistical Inference: 2. Information From Stably Distributed Samples," *Journal of the American Statistical Association*, 70, 386–393.

HOFFMANN-JØRGENSEN, J. (1993): "Stable Densities," *Theory of Probability and Its Applications*, 38, 350–355.

HÖPFNER, R., AND J. JACOD (1994): "Some Remarks on the Joint Estimation of the Index and the Scale Parameters for Stable Processes," in *Proceedings of the Fifth Prague Symposium on Asymptotic Statistics*, ed. by P. Mandl and M. Huskova. Heidelberg: Physica-Verlag, 273–284.

JACOD, J., AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes* (Second Ed.). New York: Springer-Verlag.

MANCINI, C. (2001): "Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion," *Giornale dell'Istituto Italiano degli Attuari*, LXIV, 19–47.

MYKLAND, P. A., AND L. ZHANG (2006): "ANOVA for Diffusions and Itô, Processes," *The Annals of Statistics*, 34, 1931–1963.

NOLAN, J. P. (1997): "Numerical Computation of Stable Densities and Distribution Functions," *Communications in Statistics—Stochastic Models*, 13, 759–774.

——— (2001): "Maximum Likelihood Estimation and Diagnostics for Stable Distributions," in *Lévy Processes: Theory and Applications*, ed. by O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick. Boston: Birkhäuser, 379–400.

RACHEV, S., AND S. MITTNIK (2000): *Stable Paretian Models in Finance*. Chichester, U.K.: Wiley.

WOERNER, J. H. (2004): "Estimating the Skewness in Discretely Observed Lévy Processes," *Econometric Theory*, 20, 927–942.

ZOLOTAREV, V. M. (1986): *One-Dimensional Stable Distributions*, Vol. 65. Providence, RI: Translations of Mathematical Monographs, American Mathematical Society.

——— (1995): "On Representation of Densities of Stable Laws by Special Functions," *Theory of Probability and Its Applications*, 39, 354–362.