IDENTIFICATION OF LÉVY SYSTEMS IN FINANCIAL MATHEMATICS*

L. GERENCSÉR †, M. MÁNFAY ‡, AND B. GERENCSÉR §

Abstract. We propose a new class of processes for modeling financial data. The novelty of our approach is that the standard assumption of independent returns is relaxed, thus allowing to model the effect of feedback or friction in financial markets. The formal model is a linear stochastic system driven by a Lévy process. Widely used Lévy processes are the stable, VG or the CGMY process. We present a novel method for the identification of such systems, which is a combination of standard techniques in systems identification with a specific statistical technique widely used in the context in finance, called the empirical characteristic function (ECF) method.

Key words. linear stochastic systems, Lévy processes, system identification, empirical characteristic function

AMS subject classifications. 62P05, 93E12

1. Introduction. Lévy processes have become a widely used tool in modeling price processes of financial instruments, such as stock prices or indices [10]. A Lévy process (Z_t) is much like a Wiener process: a process with stationary an independent increments, but discontinuities or jumps are allowed. For an excellent introduction to Lévy processes see [1]. A key building block in the theory of Lévy processes is the compound Poisson process. A more general class of pure jump Lévy process is formally obtained via

$$Z_t = \int_0^t \int_{\mathbf{R}^1} x N(ds, dx)$$

where $N(dt, dx)$ is a time-homogeneous, space-time Poisson point process, counting the number of jumps of size x at time dt. The intensity of $N(dt, dx)$ is defined by

$$\mathbb{E}N(dt, dx) = dt \cdot \nu(dx),$$

where $\nu(dx)$ is the Lévy-measure. The above representation is mathematically rigorous if

$$\int_{\mathbf{R}^1} \min(|x|, 1)\nu(dx) < \infty.$$

Under this condition the sample paths of Z_t are of *finite variation*, a property supported by empirical evidence for most indices. The characteristic function of a Lévy process can be written in the form

$$\mathbb{E}e^{iuZ_t} = e^{t\psi(u)}, \tag{1.1}$$

where $\psi(u)$ is the characteristic exponent.

The standard model of a price process within this framework is then

$$S_t = S_0 \exp Z_t. \tag{1.2}$$

*This work has been partially supported by Morgan Stanley Analytics Ltd., Budapest, Hungary.
†MTA SZTAKI, Budapest, Hungary
‡MTA SZTAKI, Budapest, Hungary; Central European University, Budapest, Hungary
§Alfréd Rényi Institute of Mathematics, Budapest, Hungary

(S_t) is called a geometric Lévy process. A variety of choices for (Z_t) has been proposed in the literature: it can be a stable process, a variance Gamma (VG) process, a tempered stable (CGMY) process, a hypergeometric process or a normal inverse Gauss (NIG) process.

The motivation behind these models is that the returns of the stock process, say $(S_{t+h} - S_t)/S_t$ are assumed to be independent and stationary. While this is an attractive assumption, its consequences are less attractive. In particular it follows that the variance of the price process tends to infinity. A closer look at data in fact reveals that there is a weak correlation between, say, daily returns $(S_{t+1} - S_t)/S_t$.

We propose to introduce a new class of models allowing for decaying memory. For this purpose we first define a process Y_t which is the output of a finite dimensional stable linear SISO system, driven by a Lévy process:

$$dY_t = HdZ_t,$$

where H is a transfer function and Z is a Lévy process. Such a system will be called a Lévy system. Then we define the price process according to (1.2): $S_t = S_0 \exp Y_t$. In the statistical analysis of such systems, the fine characteristics of (Z_t) are also of prime importance. A very similar model was proposed in [13].

1.1. Problem Formulation. The difficulty of applying a maximum-likelihood (ML) method lies in the fact that there is no natural reference measure in the space of sample paths. To avoid this problem we focus on a simpler class of discrete time Lévy systems defined as

$$\Delta y_n = H(\theta^*)\Delta Z_n, \tag{1.3}$$

where ΔZ_n is the increment of a Lévy process Z over an interval $[(n-1)h, nh)$ with $\mathbb{E}[\Delta Z_n] = 0$, a property to be removed later. The Lévy-measure of Z will be denoted by $\nu(dx) = \nu(dx, \eta^*)$, where η^* denotes an unknown parameter-vector. The range of η^* is assumed to be known. The transfer function $H(\theta)$ is assumed to be stable and inverse stable for $\theta \in D \subset \mathbb{R}^p$, where D is a known open set. We suppose that

$$\int_{|x|\geq 1} |x|^q \nu(dx) < +\infty \tag{1.4}$$

for all $1 \leq q < Q$ with some constant Q.

Our aim is to identify both the system parameters, denoted by θ^*, both the noise parameters, denoted by η^*. The application of the ML method would solve the full identification problem along standard lines, assuming that the density function of ΔZ_n is known, which is unfortunately not the case. The objective of this paper is to present a combination of advanced techniques in systems identification with a specific statistical technique, widely used in the context in finance, called the ECF method, which can be interpreted as the Fourier transform of an ML method.

2. Lévy processes in finance. A wide range of geometric Lévy processes has been proposed by a variety of authors. A symmetric α-stable process, with $0 < \alpha < 2$, is defined via the Lévy measure

$$\nu(dx) = C|x|^{-1-\alpha}dx.$$

A recently widely studied class of Lévy processes is the CGMY process, due to Carr, Geman, Madan and Yor [5]. It is obtained by multiplying the Lévy-density of a stable process with a decreasing exponential on each half of the real axis. Its Lévy-measure, using standard parametrization, is of the form:

$$\nu(dx) = \frac{Ce^{-G|x|}}{|x|^{1+Y}}\mathbf{1}_{x<0}dx + \frac{Ce^{-Mx}}{|x|^{1+Y}}\mathbf{1}_{x>0}dx,$$

where $C, G, M > 0$, and $0 < Y < 2$. Intuitively, C controls the level of activity, G and M together control skewness. Y controls the density of small jumps, i.e. the fine structure. For $Y < 1$ the corresponding Lévy process is of finite variation. Allowing $Y = 0$ yields the Variance Gamma process that has been proposed by Madan, Carr and Chang [8].

The characteristic exponent of the CGMY process is given by

$$\psi(u) = C\Gamma(-Y)\left((M - iu)^Y - M^Y + (G + iu)^Y - G^Y\right), \tag{2.1}$$

where Γ denotes the gamma-function.Although the geometric CGMY model is widely used in finance, it cannot always be validated on real data. Surprisingly, even the assumed independence of daily log-returns may not be validated on historical data.

3. ECF identification of Lévy systems. The ECF method has been widely used in finance as an alternative to the ML Method, assuming i.i.d. returns [3], [4], [9]. We adapt this technique to the problem of identifying the discrete-time Lévy system described in (1.3). Fix a realization of H in its innovation form, i.e. assume that H and its inverse are stable. Then the estimated innovation process $(\varepsilon_n(\theta))$ is defined via the inverse filter:

$$\varepsilon_n(\theta) = H^{-1}(\theta)\Delta y_n, \tag{3.1}$$

with zero initial conditions. Let $\varepsilon_n^*(\theta)$ denote the stationary solution of (3.1) when $-\infty < n < \infty$. Then we have for $k \geq 0$

$$\varepsilon_k^*(\theta) = \varepsilon_k(\theta) + r_k, \tag{3.2}$$

where $r_k = O_M^Q(\alpha^k)$ with some $0 < \alpha < 1$, meaning that for all $1 \leq q < Q$

$$\sup_k \alpha^{-k}\mathbb{E}^{1/q}|r_k|^q < \infty.$$

The score functions to be used following the basic idea of the ECF method are defined as

$$h_k(u;\theta,\eta) = e^{iu\varepsilon_k(\theta)} - \varphi(u,\eta) \tag{3.3}$$

$$h_k^*(u;\theta,\eta) = e^{iu\varepsilon_k^*(\theta)} - \varphi(u,\eta) \tag{3.4}$$

with $u \in \mathbb{R}$. These are indeed appropriate score functions, since we obviously have

$$\mathbb{E}\left[h_k^*(u;\theta^*,\eta^*)\right] = 0.$$

Following the philosophy of the ECF method take a fix set u_i-s, and define the k-dimensional vector

$$h_n(\theta,\eta) = \left(h_n(u_1;\theta,\eta),\ldots,h_n(u_k;\theta,\eta)\right)^T.$$

Let $K > 0$ be a fixed $k \times k$ weighting matrix. Then we define the cost function as

$$V_N = V_N(\theta, \eta) = \sum_{n=1}^{N} |K^{-1/2} h_n(\theta, \eta)|^2.$$

By minimizing V_N we obtain the estimation θ_N and η_N of θ^* and η^*, respectively.

4. Analysis. We show that the gradient of $V_N(\theta, \eta)$ serves as an alternative score function. Differentiating V_N w.r.t θ and η we get the equations

$$\sum_{n=1}^{N} \left(h_{n\theta}^T(\theta, \eta) K^{-1} \bar{h}_n(\theta, \eta) + h_n^T(\theta, \eta) K^{-1} \bar{h}_{n\theta}(\theta, \eta) \right) = 0, \tag{4.1}$$

$$\sum_{n=1}^{N} \left(h_{n\eta}^T(\theta, \eta) K^{-1} \bar{h}_n(\theta, \eta) + h_n^T(\theta, \eta) K^{-1} \bar{h}_{n\eta}(\theta, \eta) \right) = 0, \tag{4.2}$$

where $\bar{h}$ is the conjugate of h. Note that, setting $\theta = \theta^*$, the second equation is just the optimality condition of the ECF method for i.i.d. samples. The derivative of the score function is

$$h_{n\theta}(u, \theta, \eta) = e^{iu\varepsilon_n(\theta)} iu\varepsilon_{n\theta}(\theta). \tag{4.3}$$

Hence in the first equation $h_{n\theta}(\theta, \eta)$ and $h_n(\theta, \eta)$ are not independent. However, the next lemma shows that $h_{n\theta}(\theta, \eta)$ and $h_n(\theta, \eta)$ are essentially uncorrelated.

LEMMA 4.1. *For any η we have $\mathbb{E}\left[V_{N\theta}(\theta^*, \eta)\right] = O_M^Q(\alpha^n)$, and in addition $\mathbb{E}\left[V_{N\eta}(\theta^*, \eta^*)\right] = O_M^Q(\alpha^n)$.*

Proof. First note that (3.2) implies

$$\mathbb{E}\left[h_{n\theta}^T(\theta^*, \eta) K^{-1} \bar{h}_n(\theta^*, \eta)\right] = \mathbb{E}\left[h_{n\theta}^{*T}(\theta^*, \eta) K^{-1} \bar{h}_n^*(\theta^*, \eta)\right] + O_M^Q(\alpha^n).$$

Consider the n^{th} term in (4.1). We have

$$\begin{aligned} &\mathbb{E}\left(h_{n\theta}^{*T}(\theta^*, \eta) K^{-1} \bar{h}_n^*(\theta^*, \eta)\right) = \\ &= \sum_{l=1}^{k} \sum_{m=1}^{k} K_{l,m}^{-1} \mathbb{E}\left[\left(e^{iu_l\varepsilon_n^*(\theta^*)} iu_l\varepsilon_n^*(\theta^*)\right)\left(e^{-iu_m\varepsilon_n^*(\theta^*)} - \varphi(-u_m, \eta)\right)\right]. \end{aligned} \tag{4.4}$$

Fix l, m and compute one term using the tower law:

$$\begin{aligned} &\mathbb{E}\left[\left(e^{iu_l\varepsilon_n^*(\theta^*)} iu_l\varepsilon_{n\theta}^*(\theta^*)\right)\left(e^{-iu_m\varepsilon_n^*(\theta^*)} - \varphi(-u_m, \eta)\right)\right] = \\ &= \mathbb{E}\left[\mathbb{E}\left[\left(e^{iu_l\varepsilon_n^*(\theta^*)} iu_l\varepsilon_{n\theta}^*(\theta^*)\right)\left(e^{-iu_m\varepsilon_n^*(\theta^*)} - \varphi(-u_m, \eta)\right) | F_{n-1}^{\Delta Z}\right]\right], \end{aligned} \tag{4.5}$$

where $F_{n-1}^{\Delta Z} = \sigma\left\{\Delta Z_k : k \leq n-1\right\}$. Here we used that $\bar{\varphi}(u, \eta) = \varphi(-u, \eta)$. Due to the fact that $\epsilon_{n\theta}^*(\theta^*)$ is $F_{n-1}^{\Delta Z}$ measurable, (4.5) can be written as

$$\begin{aligned} &\mathbb{E}\left[iu_l\epsilon_{n\theta}^*(\theta^*)\left(\varphi(u_l - u_m, \eta^*) - \varphi(u_l, \eta^*)\varphi(-u_m, \eta)\right)\right] = \\ &\left(\varphi(u_l - u_m, \eta^*) - \varphi(u_l, \eta^*)\varphi(-u_m, \eta)\right)\mathbb{E}\left(iu_l\epsilon_{n\theta}^*(\theta^*)\right) = 0. \end{aligned}$$

Similarly for the n^{th} term of (4.2) we have

$$h^*_{n\eta}(u,\theta,\eta) = -\varphi_\eta(u,\eta),$$

which is non-random implying that

$$\mathbb{E}h^*_{n\eta}(u,\theta^*,\eta^*)K^{-1}\bar{h}^*_n(u,\theta^*,\eta^*) = 0.$$

□

Define $\rho = (\theta,\eta)$, and define the asymptotic cost function by

$$W(\theta,\eta) = W(\rho) = \mathbb{E}\left|K^{-1/2}h^*_n(\rho)\right|^2.$$

A crucial object is the Hessian of W at $\rho = \rho^*$:

$$R^* = W_{\rho\rho}(\rho^*).$$

It is easy to see that R^* is block diagonal matrix, with diagonal blocks $R^*(\theta^*) = W_{\theta\theta}(\theta^*)$ and $R^*(\eta^*) = W_{\eta\eta}(\eta^*)$. The following result provides a precise characterization of the estimation error:

Theorem 4.2. *Under conditions seen in Section 1.1 we have*

$$\hat{\rho}_N - \rho^* = -(R^*)^{-1}\frac{1}{N}V^*_{N\rho}(\rho^*) + O^Q_M(N^{-1})$$

Idea of proof: The proof of Theorem 4.2 is obtained by adapting methods developed in [11]. From (1.4) it follows that $\mathbb{E}|\Delta Z_n|^q < +\infty$ for all $1 \le q < Q$. Thus we get that both y and $\varepsilon_n(\theta)$ are L-mixing in a restricted sense [12]. From here we obtain, following [12], that for any compact $D_0 \subset D$

$$\sup_{(\theta,\eta)\in D_0}\left|\frac{1}{N}\sum_{n=1}^{N}\left(|K^{-1/2}h_n(\theta,\eta)|^2 - W(\theta,\eta)\right)\right| = O^Q_M(N^{-1/2}).$$

Similar results hold for the derivatives of the terms in the bracket, whenever H and φ are sufficiently smooth in the parameters. From here Theorem 4.2 follows using arguments of [11].

Lemma 4.3. *Let $K = I$. Then*

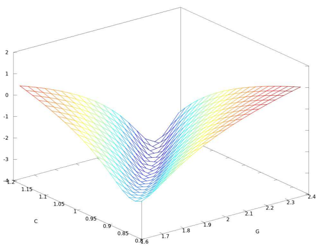
$$R^*(\theta^*) = 2\mathbb{E}\left[\varepsilon_{n\theta}(\theta^*)\varepsilon^T_{n\theta}(\theta^*)\right]\sum_{j=1}^{k}u_j^2\left|\varphi(u_j,\eta^*)\right|^2$$

and

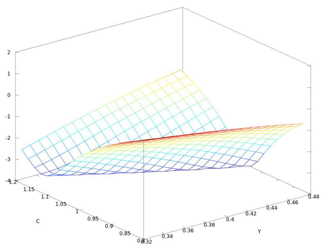
$$S^* := \frac{1}{N}\mathrm{Cov}(V^*_{N\theta}(\theta^*),V^*_{N\theta}(\theta^*)) =$$
$$2\mathbb{E}\left[\varepsilon_{n\theta}(\theta^*)\varepsilon^T_{n\theta}(\theta^*)\right]\sum_{j=1}^{k}u_j^2\left(|\varphi(u_j,\eta^*)|^2 - \Re\left(\varphi(2u_j,\eta^*)\varphi(-u_j,\eta^*)^2\right)\right).$$

The proof follows the line of arguments for Lemma 4.1. One can obtain similar results for arbitrary K, but calculations are considerably simplifed if we take $K = I$. Note that our theorem implies that the asymptotic covariance matrix of $\hat{\rho}_N - \rho^*$ is $(R^*)^{-1}S^*(R^*)^{-1}$.

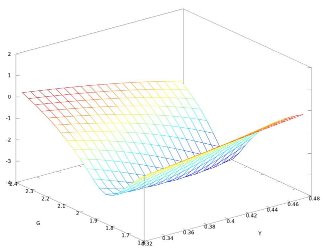
5. Statistical Estimation of Stock Prices. In this section we summarize some numerical results. Consider a process defined by $\Delta Y_n = a\Delta Y_{n-1} + \Delta Z_n$ which has been simulated with parameters $C = 1, G = M = 2, Y = 0.4, a = 0.1$. for $1 \leq n \leq 5000$. Define an error function as the difference of the empirical characteristic function and of the characteristic function. In Figure (5.1.) the logarithm of the error function is plotted against different pairs of arguments. These figures suggest that the proposed ECF estimation is numerically ill-conditioned and that is what we experienced in practice. To understand the causes of the numerical problem we also considered the estimation of a plain CGMY process (with $a = 0$), assuming that the parameters C and Y known. Surprisingly, the estimations of G and M are fairly accurate in this case. We got that $\hat{G} = 0.5089$ and $\hat{M} = 0.6928$, while $G^* = 0.6$ and $M^* = 0.7$



(a) C, G error



(b) C,Y error



(c) G, Y error

Fig. 5.1. *CGMY errors*

6. Discussion. One could solve our identification problem using the orthogonality of ΔZ only by applying a prediction error method to get an estimation $\hat{\theta}_N$ of θ^*, and then, applying an ECF method with score function $h_n(u, \eta) = e^{iu\varepsilon_n(\hat{\theta}_N)} - \varphi(u, \eta)$ to estimate η^*. Comparing the efficiency of these methods is the subject of further investigations. According to the literature the optimal choice of K is $K = \mathbb{E}\left[h(\theta^*, \eta^*)\bar{h}(\theta^*, \eta^*)^T\right]$ for i.i.d samples using a continuum of u-s. It is an open problem whether an approximately optimal choice of K would significantly improve the quality of the estimation of θ^*. Since the estimations of G and M were fairly accurate it might be a promising idea to mix the ECF method with other techniques.

REFERENCES

[1] J. Jacod, A.N. Shiryaev, *Limit theorems for stochastic processes (2. ed.)*, Springer, (2002)

[2] P. Carr, H. Geman, D. Madan and M. Yor, *The fine structure of asset returns: an empirical investigation*, in Journal of Business, 75 (2) (2002), pp. 305-332.

[3] M. Carrasco and J.-P. Florens, *Generalization of GMM to a continuum of moment conditions*, in Econometric Theory, 16 (06) (2000), pp. 797–834.

[4] M. Carrasco and J. P. Florens, *Efficient GMM estimation using the empirical characteristic function*, in Idei working papers, 140 (2002)

[5] P. Carr, H. Geman, D.B. Madan and M. Yor, *The Fine Structure of Asset Returns: An Empirical Investigation*, in The Journal of Business, 75 (2000), pp. 305–322.

[6] R. Cont, P. Tankov, *Financial Modelling with Jump Processes*, in Journal of the American Statistical Association, 101 (2006), pp. 1315-1316.

[7] A. Feuerverger and P. McDunnogh, *On the efficiency of empirical characteristic function procedures*, in J.R. Stat. Soc. B, 43 (1) (1981), pp. 20–47.

[8] B. Madan, P. Carr and C. Chang, *The Variance Gamma Process and Option Pricing*, in European Finance Review, 2 (1998), pp. 79–105.

[9] J. Yu, *Empirical characteristic function estimation and its applications*, in Econometric Reviews, 23 (2) (2004), pp. 93–123.

[10] Y. Miyahara and A. Novikov, *Geometric Lévy Process Pricing Model*, in Research Paper Series 66, Quantitative Finance Research Centre, University of Technology, Sydney, (2001)

[11] L. Gerencsér, *On the martingale approximation of the estimation error of ARMA parameters*, in System & Control Letters, 15 (1990), pp. 417–423.

[12] L. Gerencsér, *On a class of mixing processes*, in Stochastics, 26 (1989), pp. 165–191.

[13] P.J. Brockwell, *Lévy-Driven Carma Processes* , in Annals of the Institute of Statistical Mathematics, 53 (1) (2001), pp. 113–124.