

Identification of finite dimensional linear stochastic systems driven by Lévy processes

Máté Mánfay¹ and László Gerencsér²

Abstract—We study the problem of identifying a finite dimensional linear stochastic SISO system driven by a Lévy process. The latter are widely used in modelling financial time series. In a number of important examples the density function of the innovation term is unknown, but its characteristic function is explicitly known, possibly up to a few unknown parameters. In this paper we present and analyze a novel identification method that exploits the information on the characteristic function of the noise. It is obtained by adapting the empirical characteristic function method (ECF for short) developed for i.i.d. samples. We will show that the new method may be more efficient in estimating the system parameters than a plain prediction error method.

I. INTRODUCTION

In this paper we consider the problem of identifying a finite dimensional linear stochastic SISO system such that its innovation process is an i.i.d. sequence given by the increments of a Lévy process. This class of models is a natural extension of the class of finite dimensional Gaussian linear stochastic systems. However a novel technical challenge is that in a number of interesting applications the density function of the innovation terms is unknown, while its characteristic function is explicitly known, possibly up to a few unknown parameters, which are also to be identified. Surprisingly, this seemingly simple problem raises a number of non-trivial technical issues, some of which are still open.

The problem that we consider is of interest due to the fact that Lévy processes are widely used in modelling financial time series, see e.g. [12],[11],[5]. A common feature of the cited models is that the logarithm of the price process (or the log-return process) is a Lévy process with increments having explicitly known characteristic functions. A novelty of our class of models that we allow the increments of the the log-price process to have fading, but non-zero memory. This is best expressed by letting the increments of the original Lévy process pass through a finite dimensional linear stochastic system.

The main contribution of the paper is the development and analysis of a new system identification method exploiting the knowledge of the characteristic function of the noise, possibly up to a few unknown parameters. The main idea is the adaptation of the so-called ECF (empirical characteristic function) method developed for the statistical analysis of i.i.d. sequences, see [7],[6], with a characteristic function belonging to a known, parametric class of characteristic functions. The empirical characteristic function method has

been widely used in finance as an alternative to the maximum likelihood method (ML for short), assuming i.i.d. returns [6],[7],[13]. The ECF method for i.i.d. samples has the fascinating property that, by an appropriately chosen weighting function, the asymptotic covariance matrix of the estimator can be arbitrarily close to what is obtained by the ML method, which is known to be the inverse of the Fisher information matrix, see [10]. In fact, we present two possible extensions of the ECF method to solve our problem.

In the first method the estimation of system parameters and noise parameters are separately treated. This method is obtained by combining the prediction error method (PE for short) with the ECF method for i.i.d. sequences. In the second method we apply the ECF method directly by fitting the characteristic function of a single term of the innovation process to processed data.

We will focus on the second method that is capable to ensure better efficiency in estimating system parameters than a plain PE method. Here efficiency is measured by the asymptotic covariance matrix of the estimator. The possibility to outperform the PE method will be proven, and, in addition, empirical evidences will be given for the case of the so-called CGMY process, see [5].

II. DISCRETE-TIME LÉVY SYSTEMS

A Lévy process (Z_t) is a stochastic process with stationary and independent increments, but, aside from the Brownian motion, it has discontinuous sample paths. For an excellent introduction to Lévy processes see [3]. 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 t . In this case Z_t is a pure jump process, which means that the Lévy-Ito decomposition of Z_t does not have a Brownian motion component (but it may have a drift term). The intensity of $N(dt, dx)$ is defined and expressed as

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

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

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

¹MTA SZTAKI, Budapest, Hungary; Central European University, Budapest, Hungary

²MTA SZTAKI, Budapest, Hungary

The characteristic function of a Lévy process can be written as

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

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

A prime example for a Lévy process is the CGMY process, due to Carr, Geman, Madan and Yor [8], which is widely used process to model financial time series. It is also called a tempered stable process, because its Lévy-density is obtained by multiplying the Lévy-density of a stable process with a decreasing exponential. Its Lévy measure 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 of the process. The characteristic exponent of the CGMY process is

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

where Γ denotes the gamma-function.

To avoid technical difficulties concerning continuous time models we consider an alternative discrete-time model class, in which the output process (Δy_n) is defined via a discrete-time finite dimensional linear stochastic system

$$\Delta y_n = A \Delta Z_n, \quad (1)$$

where A , like before, represents the linear mapping from input to output, and the input terms ΔZ_n are defined as the increments of the Lévy process Z over interval $[(n-1)h, nh)$, with some fixed $h > 0$. For the sake of convenience we let $-\infty < n < +\infty$. We assume that a state space equation for this model is given by

$$\Delta X_{n+1} = H\Delta X_n + K\Delta Z_n \quad (2)$$

$$\Delta y_n = L\Delta X_n + \Delta Z_n. \quad (3)$$

We will call this model a discrete-time finite dimensional Lévy system.

Assume now that $A = A(\theta^*)$, where θ^* is an unknown parameter-vector in \mathbf{R}^p . Similarly, assume that the Lévy measure of (Z_t) is given by $\nu(dx) = \nu(dx, \eta^*)$, where η^* denotes an unknown parameter-vector. The ranges of θ^* and η^* are assumed to be known. The fundamental problem to be discussed in this paper is to identify both the system dynamics and the noise parameters. We will establish sharp results for the error of the estimator, and compute their asymptotic covariance matrices.

III. THREE IDENTIFICATION PROBLEMS

In this section we formulate three identification problems related to discrete-time, finite dimensional Lévy systems, and sketch a possible path to their solution. The first, simplest problem is of mere technical interest:

Known system parameters, unknown noise parameters. In

this case define and compute

$$\varepsilon_n(\theta^*) = A^{-1}(\theta^*)\Delta y_n = A^{-1}(\theta^*)A(\theta^*)\Delta Z_n = \Delta Z_n,$$

assuming, for the sake of simplicity, that $\Delta Z_n = \varepsilon_n(\theta^*) = 0$ for $n \leq 0$. After that we can apply the ECF method for i.i.d. samples to obtain the estimation of η^* .

Known noise parameters, unknown system parameters. This is the simplest, technically interesting and non-trivial problem. If we knew the probability density function of the noise, say f , we could obtain the maximum likelihood estimate of θ^* via solving

$$\sum_{n=1}^N f_\theta(\varepsilon_n(\theta), \eta^*) = 0, \quad (4)$$

where f_θ is the derivative of f w.r.t. θ .

$$\varepsilon_n(\theta) = A^{-1}(\theta)\Delta y_n \quad (5)$$

is the estimated innovation process, see [17].

Under certain conditions the asymptotic covariance matrix of the ML estimate $\hat{\theta}_N$ is

$$\Sigma_{ML} = (R^*)^{-1} \mu,$$

where

$$R^* = \lim_{n \rightarrow \infty} \mathbb{E} [\varepsilon_{n\theta}(\theta^*) \varepsilon_{n\theta}^T(\theta^*)]$$

and

$$\mu = \lim_{n \rightarrow \infty} \mathbb{E} \left[\left(\frac{f'(\Delta Z_n)}{f(\Delta Z_n)} \right)^2 \right].$$

In our case, the p.d.f. of the noise distribution is not known. One might apply the prediction error method to estimate the system dynamics, i.e. θ^* . However, we will show, in the case of CGMY noise, that we may estimate θ^* in a more efficient way using an appropriate adaptation of the ECF method. In fact, this result is a special case of a more general result obtained for the general problem to be described in the next paragraph.

Both the system parameters and the noise parameters are unknown. The first method that we propose is quite straightforward: we estimate the system parameters using a PE method, then, using a certainty equivalence argument, we estimate the innovation process by inverting the system using the estimated parameters. Then, we estimate the noise parameters using ECF method for i.i.d. sequences. This method will be studied in section IV.

The second method, which is the main subject of this paper, estimates both the system parameters and noise parameters using an ECF method. First, an parameter-dependent, estimated innovation process $\varepsilon_n(\theta)$ is defined, then the characteristic function of the noise is fitted to empirical data defined in terms of $\varepsilon_n(\theta)$. Thus we get a score function that depends on both θ and η .

IV. COMBINING PE AND THE CLASSIC ECF METHOD

In this section we estimate the dynamics and the noise simultaneously in a straightforward way. We identify θ^* using only the orthogonality of ΔZ by applying a prediction

error method. This way we get an estimation $\hat{\theta}_N$ of θ^* , without using the Lévy measure of ΔZ . Then we apply an ECF method with the score function

$$h_n(u, \eta) = e^{iu\varepsilon_n(\hat{\theta}_N)} - \varphi(u, \eta)$$

to estimate η^* .

First, we define the estimated innovation process as in the previous sections and we use Condition 1 and 2 from Section IV. The prediction error method is obtained by minimizing the cost function

$$V_{P,N}(\theta) = \frac{1}{2} \sum_{n=1}^N \varepsilon_n^2(\theta).$$

In practice the estimated $\hat{\theta}_N$ is defined as the solution of

$$V_{P,N\theta}(\theta) = \sum_{n=1}^N \varepsilon_n(\theta) \varepsilon_{n\theta}(\theta) = 0.$$

The asymptotic cost function associated with the PE method is defined as

$$W_P(\theta) = \frac{1}{2} \lim_{n \rightarrow \infty} \mathbb{E} \varepsilon_n^2(\theta) = \frac{1}{2} \mathbb{E} \varepsilon_n^{*2}(\theta),$$

where $\varepsilon_n^*(\theta)$ is the innovation process that is calculated with stationary initial values. In general, the notation $(\cdot)^*$ will be used throughout this paper if the corresponding stochastic process is obtained by passing through a stationary process through an exponentially stable linear filter starting at $-\infty$, as opposed to initializing the filter at time 0 with some arbitrary initial condition, which is typically zero. We have

$$W_{P,\theta}(\theta^*) = 0 \text{ and } R_P^* := W_{P,\theta\theta}(\theta^*) = \mathbb{E} [\varepsilon_{n\theta}^*(\theta^*) \varepsilon_{n\theta}^{T*}(\theta^*)].$$

The asymptotic covariance matrix of the PE estimate of θ^* is given by

$$\Sigma_P = (\mathbb{E} [\varepsilon_{n\theta}^*(\theta^*) \varepsilon_{n\theta}^{T*}(\theta^*)])^{-1}. \quad (6)$$

We will use the following notation:

Definition IV.1. For a stochastic process X_n , and a function $f : \mathbb{Z} \rightarrow \mathbb{R}^+$ we say that $X_n = O_M^Q(f(n))$ if $\sup_n \frac{\mathbb{E}^{1/q} |X_n|^q}{f(n)} < \infty$ holds for $1 \leq q \leq Q$.

The proposition lemma, with minor variation, can be found in [15].

Proposition IV.1. Under Condition 1,2,3 below we have $\hat{\theta}_N - \theta^* = O_M^{Q/2}(N^{-1/2})$.

An ideal score function for the ECF method to estimate η^* would be defined by

$$h_{*,n}(u, \eta) = e^{iu\varepsilon_n^*(\theta^*)} - \varphi(u, \eta). \quad (7)$$

Since we are not given θ^* we define an alternative, θ -dependent score function via

$$h_n(u, \theta, \eta) = e^{iu\varepsilon_n(\theta)} - \varphi(u, \eta).$$

These are appropriate score functions since $\mathbb{E} h_n^*(u, \theta^*, \eta^*) = 0$.

Fix a set of real numbers u_1, \dots, u_k , with $k \geq \dim \eta$ and define

$$h_n(\theta, \eta) = (h_n(u_1, \theta, \eta), \dots, h_n(u_k, \theta, \eta))^T.$$

Then we obtain the estimate $\hat{\eta}_N$ of η^* by finding a least squares solution to the over-determined system of equations

$$h_n(\hat{\theta}_N, \eta) = 0.$$

More precisely, define the θ -dependent cost function

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

where K is a symmetric, positive definite weighting matrix. Then we obtain the estimate $\hat{\eta}_N$ of η^* by minimizing $V_{E,N}(\hat{\theta}_N, \eta)$.

Define the (θ -dependent) asymptotic cost function as

$$W_E(\theta, \eta) = \mathbb{E} |K^{-1/2} h_n^*(\theta, \eta)|^2.$$

Let its Hessian w.r.t. η at $\eta = \eta^*$ be denoted by

$$R_E^* = W_{E,\eta\eta}(\theta^*, \eta^*).$$

To formulate our result we need some technical conditions.

Condition 1 We assume that

$$\int_{|x| \geq 1} |x|^q \nu(dx) < +\infty \quad (8)$$

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

Note that Condition 1 holds with $Q = \infty$ in our benchmark examples, for stable and CGMY processes.

Condition 2 $A(\theta)$ is assumed to be exponentially stable and exponentially inverse stable for $\theta \in G_\theta \subset \mathbb{R}^p$, where G_θ is a known open set.

Let ρ be the joint parameter i.e. $\rho = (\theta, \eta)$. Let D_ρ and D_ρ^* be compact domains such that $\rho^* \in D_\rho^* \subset \text{int } D_\rho$ and $D_\rho \subset G_\rho$.

Condition 3 The equations $W_{P,\theta}(\theta) = 0$, and $W_{E,\eta}(\eta) = 0$ have a unique solution in D_ρ^* .

Theorem IV.1. Under Conditions 1,2 and 3 we have

$$\hat{\eta}_N - \eta^* = -(R_E^*)^{-1} \frac{1}{N} V_{E,N\eta}(\eta^*) + O_M^{Q/(2(p+q))}(N^{-1}).$$

The proof is obtained by the same methods that are used to prove Theorem VI.1. See the next section for more details.

V. SINGLE TERM ECF METHOD FOR LÉVY SYSTEMS

In this section we present an alternative extension of the ECF method in which the system dynamics and the noise parameters are estimated simultaneously. Consider a discrete-time, finite dimensional Lévy system described by the state-space equation (2).

The effect of non-stationary initial values decay exponentially, so that we have

$$\varepsilon_n(\theta^*) - \Delta Z_n = r_n$$

with some $0 < \alpha < 1$ and $r_n = O_M^Q(\alpha^n)$

We define the score functions via single terms of the processed data $\varepsilon_n(\theta)$ as

$$h_n(u; \theta, \eta) = e^{iu\varepsilon_n(\theta)} - \varphi(u, \eta). \quad (9)$$

These are indeed appropriate score functions, because

$$\mathbb{E}[h_n^*(u; \theta^*, \eta^*)] = 0$$

holds. Take a fix set $\{u_i : u_i \in \mathbb{R}, i = 1, \dots, k\}$, with $k \geq \dim\theta + \dim\eta$. We will fit the exact characteristic function to the empirical characteristic function at these values. Let

$$h_n(\theta, \eta) = (h_n(u_1; \theta, \eta), \dots, h_n(u_k; \theta, \eta))^T.$$

We will find the estimator of (θ^*, η^*) by finding a weighted least squares solution to the set of over-determined system of equations

$$h_n(\theta, \eta) = 0 \quad n = 1, \dots, N.$$

Fix a $k \times k$ weighting matrix K , and define the cost function

$$V_N = V_N(\theta, \eta) = \sum_{n=1}^N |K^{-1/2} h_n(\theta, \eta)|^2. \quad (10)$$

We obtain the estimates $\hat{\theta}_N$ and $\hat{\eta}_N$ of θ^* and η^* , respectively, by solving

$$\begin{aligned} V_{N\theta}(\theta, \eta) &= 0 \\ V_{N\eta}(\theta, \eta) &= 0. \end{aligned}$$

VI. ANALYSIS OF THE NEW ECF METHOD

In this section we analyze the single term ECF method presented in the previous section. Our first main result is Theorem VI.1 which gives a precise characterization of the estimation error. The second main results are Theorem VI.2 and Theorem VI.3 providing a formula for the asymptotic covariance matrix of the estimators of (θ^*, η^*) . Differentiating V_N w.r.t. θ and η we get the equations

$$V_{N\theta}^*(\theta, \eta) = \sum_{n=1}^N (h_{n\theta}^{T*}(\theta, \eta) K^{-1} \bar{h}_n^*(\theta, \eta) + h_{n\eta}^{T*}(\theta, \eta) K^{-1} \bar{h}_{n\theta}^*(\theta, \eta)) = 0, \quad (11)$$

$$V_{N\eta}^*(\theta, \eta) = \sum_{n=1}^N (h_{n\eta}^{T*}(\theta, \eta) K^{-1} \bar{h}_n^*(\theta, \eta) + h_{n\theta}^{T*}(\theta, \eta) K^{-1} \bar{h}_{n\eta}^*(\theta, \eta)) = 0. \quad (12)$$

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. Note that for the derivative of the score function we have

$$h_{n\theta}(u, \theta, \eta) = e^{iu\varepsilon_n(\theta)} iu\varepsilon_n(\theta). \quad (13)$$

Proposition VI.1. *For any η we have $\mathbb{E}[V_{N\theta}^*(\theta^*, \eta)] = 0$, and in addition $\mathbb{E}[V_{N\eta}^*(\theta^*, \eta^*)] = 0$.*

Let ρ be the joint parameter i.e. let $\rho = (\theta, \eta)$, and define

the asymptotic cost function

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

Condition 3' The equation $W_\rho(\rho) = 0$ has a unique solution in D_ρ^* .

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

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

It is easy to see that

$$R^* = \begin{pmatrix} W_{\theta\theta}(\theta^*) & 0 \\ 0 & W_{\eta\eta}(\eta^*) \end{pmatrix}$$

i.e. R^* is a block diagonal matrix. The following result gives a precise characterization of the estimation error by providing an upper bound for the rate of the convergence also for the residual terms.

Theorem VI.1. *Under Conditions 1,2 and 3' we have*

$$\hat{\rho}_N - \rho^* = -(R^*)^{-1} \frac{1}{N} V_{N\rho}(\rho^*) + O_M^{Q/(2(p+q))}(N^{-1}).$$

Sketch of the proof: The proof of Theorem (VI.1) is obtained by adapting the arguments given in [15], and other known techniques in system identification methods. First, we note that the assumption $\int_{|x| \geq 1} |x|^q \nu(dx) < +\infty$ for all $1 \leq q \leq Q$ implies that $\mathbb{E}|\Delta Z_n|^q < +\infty$ for all $1 \leq q \leq Q$. Thus we get that the processes $y, \varepsilon(\theta)$ and its derivative with respect to $\varepsilon_\theta(\theta)$ are L -mixing in a restricted sense. From here we obtain by the methods developed in [15] the key inequality

$$\sup_{\rho \in D} \left\| \frac{1}{N} V_{N\rho}(\rho) - W_\rho(\rho) \right\| = O_M^{Q/(2(p+q))}(N^{-1/2}).$$

Similar results hold for the second and third order derivatives, whenever A and φ are sufficiently smooth in the parameters.

Write

$$0 = V_{N\rho}(\hat{\rho}_N) = V_{N\rho}(\rho^*) + \bar{V}_{N\rho\rho}(\hat{\rho} - \rho^*),$$

where $\bar{V}_{N\rho\rho} = \int_0^1 V_{N\rho\rho}((1-\lambda)\rho^* + \lambda\hat{\rho}_N) d\lambda$. Thus

$$\hat{\rho}_N - \rho^* = -\bar{V}_{N\rho\rho}^{-1} V_{N\rho}(\rho^*). \quad (14)$$

Let Ω_N be the set where the solution of $V_{N\rho}(\rho) = 0$ is unique, for this set $P(\Omega_N) > 1 - O(N^{-s})$ holds with any $0 < s \leq Q/2$, where $P(\omega)$ denotes the probability of event ω . Following [15] once again we get the following inequality:

$$\chi_{\Omega_N} \left\| \bar{V}_{N\rho\rho}^{-1} - \frac{1}{N} W_{\rho\rho}^{-1}(\rho^*) \right\| = O_M^{Q/(2(p+q))}(N^{-3/2}),$$

with χ_ω being the indicator function of event ω . Finally using (14) and the last equation together concludes the proof.

The following lemma provides an explicit expression for the Hessian of W :

Theorem VI.2. *Under Conditions 1,2,3' we have*

$$R^* = \begin{pmatrix} W_{\theta\theta}(\theta^*) & 0 \\ 0 & W_{\eta\eta}(\eta^*) \end{pmatrix}$$

i.e. R^* is block diagonal, and here

$$W_{\theta\theta}(\theta^*) = w \mathbb{E} [\varepsilon_{n\theta}^*(\theta^*) \varepsilon_{n\theta}^{*T}(\theta^*)],$$

with

$$w = \sum_{l=1}^k \sum_{m=1}^k K_{l,m}^{-1} ((u_l^2 + u_m^2) \varphi(u_l, \eta^*) \varphi(-u_m, \eta^*) - (u_l - u_m)^2 \varphi(u_l - u_m, \eta^*)),$$

and

$$(W_{\eta\eta})_{j,j'}(\eta^*) = \sum_{l=1}^k \sum_{m=1}^k K_{l,m}^{-1} (\varphi_{\eta_j}(u_l, \eta^*) \varphi_{\eta_{j'}}(-u_m, \eta^*) + \varphi_{\eta_j'}(u_l, \eta^*) \varphi_{\eta_j}(-u_m, \eta^*)).$$

Remark 1: Note that the expression for $(W_{\eta\eta})_{j,j'}(\eta^*)$ is identical to what was obtained in [6] for i.i.d. samples.

Remark 2: Since we have $w \geq 0$, the expression for w yields a non-trivial inequality for characteristic functions.

The next step in calculating the asymptotic covariance matrix of $\hat{\theta}_N$ is the computation of $S^* = \text{Cov}(V_{N\theta}^*(\rho^*), V_{N\theta}^*(\rho^*))$. For this we need to introduce the following auxiliary function: $F(a, b, c, d, \eta) =$

$$ab[\varphi(a+b+c+d, \eta) - \varphi(a+b+c, \eta)\varphi(d, \eta) - \varphi(a+b+d, \eta)\varphi(c, \eta) + \varphi(a+b, \eta)\varphi(c, \eta)\varphi(d, \eta)].$$

Theorem VI.3. Under Conditions 1,2,3' we have for

$$S^* = \text{Cov}(V_{N\theta}^*(\rho^*), V_{N\theta}^*(\rho^*)) = s \mathbb{E} [\varepsilon_{n\theta}^*(\theta^*) \varepsilon_{n\theta}^{*T}(\theta^*)],$$

where $s = \sum_{l,m,s,t=1}^N K_{l,m}^{-1} K_{s,t}^{-1} \times$

$$(F(u_l, u_s, -u_m, -u_t, \eta^*) + F(u_l, -u_t, -u_m, u_s, \eta^*) + F(-u_m, u_s, u_l, -u_t, \eta^*) + F(-u_m, -u_t, u_l, u_s, \eta^*)).$$

Note that both R^* and S^* are of the form $c \Sigma_P$, where Σ_P is the asymptotic covariance matrix for the prediction error method, see above (6), and c is a constant. The last two theorems and Theorem VI.1 together gives an exact formula for the asymptotic covariance matrix of the estimator.

Theorem VI.4. Under Conditions 1,2 and 3' the asymptotic covariance matrix of the ECF estimator for θ^* can be written as

$$\Sigma_E = (R^*)^{-1} S^* (R^*)^{-1} = \frac{s}{w^2} \Sigma_P, \quad (15)$$

where the s and w are given in Theorems VI.3 and VI.2.

VII. EFFICIENCY OF THE ECF METHOD

In view of the efficiency of the ECF method for i.i.d samples the question arises what can be achieved by the proposed adaptation of the ECF method when identifying the dynamics of a linear stochastic system. We do not have an answer to this general question, but we will show that the commonly used PE method can be outperformed by an appropriately calibrated ECF method when the noise is CGMY. Surprisingly, we will see that the ECF method may outperform the PE method by using a single u sufficiently close to 0. Letting u tend to 0 the asymptotic variance of

the ECF estimate tends to asymptotic variance of the PE estimate. On the other hand, numerical investigations show that increasing the number of u -s used in the ECF method may not improve the efficiency significantly. For $k = 1$ the asymptotic variance of $\hat{\theta}_N$ obtained by the ECF method is, using Theorems VI.3 and VI.2,

$$\left(\mathbb{E} [\varepsilon_{n\theta}^*(\theta^*) \varepsilon_{n\theta}^{*T}(\theta^*)] \right)^{-1} \times -\frac{1}{4u^2} \left(\frac{\varphi(2u)}{\varphi^2(u)} + \frac{\varphi(-2u)}{\varphi^2(-u)} - \frac{2}{\varphi(u)\varphi(-u)} \right) := -\frac{1}{4u^2} g(u).$$

Recall that the asymptotic variance of $\hat{\theta}_N$ obtained by the PE method is $\Sigma_P = \left(\mathbb{E} [\varepsilon_{n\theta}(\theta^*) \varepsilon_{n\theta}^T(\theta^*)] \right)^{-1}$. Thus the ECF estimator outperforms the PE estimator if

$$\frac{s}{w^2} = -\frac{1}{4u^2} g(u) < 1$$

for some u .

Theorem VII.1. For all $u \neq 0$, sufficiently close to 0 we have $s/w^2 < 1$, and thus the corresponding single-term ECF estimator of the system parameter θ^* , with $k = 1$, outperforms the PE estimator.

Proof: Let us compute the Taylor expansion of $g(u)$ around 0. The first three derivatives of $\varphi(u)$ for a CGMY process with zero expectation are given by

$$\begin{aligned} \varphi(0) &= 1, \\ \varphi_u(0) &= i\mathbb{E}[\Delta Z_n] = 0, \\ \varphi_{uu}(0) &= -\mathbb{E}[(\Delta Z_n)^2] := -1, \\ \varphi_{uuu}(0) &= -i\mathbb{E}[(\Delta Z_n)^3] = 0. \end{aligned} \quad (16)$$

After a lengthy computation, that we omit, we get that

$$g(u) = -4u^2 + \frac{4}{3}G^{-2}(Y-2)(Y-3)u^2 + \mathcal{O}(u^6). \quad (17)$$

Thus

$$\begin{aligned} \frac{s}{w^2} &= -\frac{1}{4u^2} \left(\frac{\varphi(2u)}{\varphi^2(u)} + \frac{\varphi(-2u)}{\varphi^2(-u)} - \frac{2}{\varphi(u)\varphi(-u)} \right) = \\ &= 1 - \frac{1}{3}G^{-2}(Y-2)(Y-3)u^2 + \mathcal{O}(u^4). \end{aligned}$$

Since $G < 0$ and $0 < Y < 2$, the coefficient of u^2 is negative. Hence, by choosing u sufficiently small $\frac{s}{w^2} < 1$ can be achieved. \square

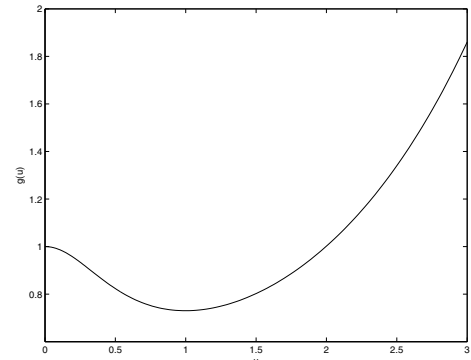


Fig. 1: Efficiency as a function of a single u

The function $g(u)$ is plotted against u for a CGMY process with parameters $C = 0.564, G = M = 1, Y = 0.5$ in Figure I. The choice $G = M$ ensures that the process has zero expectation. The minimal value of g is approximately 0.73. In Table I. we present the values of the asymptotic variances of $\hat{\theta}_N$ as we increase the number of u -s used for the ECF method with $K = I$. In each box of this table the top number is the number of u -s that are used, and below it is the corresponding value of s/w^2 . If the number of u -s is k , then we choose $(u_1, \dots, u_k) = (0.1, 0.2, \dots, 0.1k)$.

1	2	3	4	5
0.988	0.962	0.923	0.893	0.859
6	7	8	9	10
0.827	0.799	0.775	0.755	0.738
11	12	13	14	15
0.725	0.714	0.705	0.699	0.694
16	17	18	19	20
0.690	0.688	0.687	0.687	0.688

TABLE I: Increasing the number of u -s

We can see that by using several u -s we could reduce the variance, but this is not necessarily so, see the last two boxes for $k = 19$ and $k = 20$.

APPENDIX

Let $(x_n(\theta))$ be a parameter-dependent stochastic process with $n \geq 0$ and θ being a p -dimensional parameter vector taking its value in a set $D \subset \mathbb{R}^p$.

Definition .1. We say that $x_n(\theta)$ is M -bounded of order Q if for all $1 \leq q \leq Q$,

$$M_q^Q(x) = \sup_{n \geq 0, \theta \in D} \mathbb{E}^{1/q} |x_n(\theta)|^q < \infty$$

Let $e_i, i \geq 0$ be i.i.d. random variables. Define

$$\mathcal{F}_n = \sigma \{e_i : i \leq n\} \quad \text{and} \quad \mathcal{F}_n^+ = \sigma \{e_i : i > n\}.$$

Definition .2. We say that a stochastic process $(x_n(\theta))$ is L -mixing of order Q with respect to $(\mathcal{F}_n, \mathcal{F}_n^+)$, uniformly in θ , if it is \mathcal{F}_n -adapted, M -bounded of order Q , and defining for any non-negative integer r

$$\gamma_q(r, x) = \sup_{n \geq r, \theta \in D} \mathbb{E}^{1/q} |x_n(\theta) - \mathbb{E}[x_n(\theta) | \mathcal{F}_{n-r}^+]|^q,$$

we have for all $1 \leq q \leq Q$,

$$\Gamma_q(x) = \sum_{r=0}^{\infty} \gamma_q(r, x) < \infty.$$

Proof of Lemma VI.1 *Proof:* For the n -th term in (12) we have

$$\sum_{l=1}^k \sum_{m=1}^k K_{l,m}^{-1} \mathbb{E} \left[\left(e^{iu_l \varepsilon_n^*(\theta^*)} iu_l \varepsilon_n^*(\theta^*) \right) \times \left(e^{-iu_m \varepsilon_n^*(\theta^*)} - \varphi(-u_m, \eta) \right) \right].$$

Compute the (l, m) -th term using the tower law:

$$\begin{aligned} \mathbb{E} \left[\left(e^{iu_l \varepsilon_n^*(\theta^*)} iu_l \varepsilon_n^*(\theta^*) \right) \left(e^{-iu_m \varepsilon_n^*(\theta^*)} - \varphi(-u_m, \eta) \right) \right] &= \\ &= \mathbb{E} \left[\mathbb{E} \left[\left(e^{iu_l \varepsilon_n^*(\theta^*)} iu_l \varepsilon_n^*(\theta^*) \right) \times \right. \right. \\ &\quad \left. \left. \left(e^{-iu_m \varepsilon_n^*(\theta^*)} - \varphi(-u_m, \eta) \right) \middle| \mathcal{F}_{n-1}^{\Delta Z} \right] \right], \end{aligned}$$

where $\mathcal{F}_{n-1}^{\Delta Z} = \sigma \{ \Delta Z_k : k \leq n-1 \}$. Here we used that $\bar{\varphi}(u, \eta) = \varphi(-u, \eta)$. Since $\varepsilon_n^*(\theta^*)$ is $\mathcal{F}_{n-1}^{\Delta Z}$ measurable, (18) reads as

$$\begin{aligned} &\mathbb{E} \left[iu_l \varepsilon_n^*(\theta^*) \times \right. \\ &\quad \left. \mathbb{E} \left[e^{i(u_l - u_m) \varepsilon_n^*(\theta^*)} - e^{iu_l \varepsilon_n^*(\theta^*)} \varphi(-u_m, \eta) \middle| \mathcal{F}_{n-1}^{\Delta Z} \right] \right] = \\ &\quad \mathbb{E} [iu_l \varepsilon_n^*(\theta^*) (\varphi(u_l - u_m, \eta^*) - \varphi(u_l, \eta^*) \varphi(-u_m, \eta))] = \\ &\quad (\varphi(u_l - u_m, \eta^*) - \varphi(u_l, \eta^*) \varphi(-u_m, \eta)) \mathbb{E} [iu_l \varepsilon_n^*(\theta^*)] = 0. \end{aligned}$$

In the last equation we used that $\mathbb{E}[\Delta Z_n] = 0$.

Similarly for the n -th term of (12) we have $h_{n\eta}^*(u, \theta, \eta) = -\varphi_\eta(u, \eta)$, which is non-random, so it follows that $\mathbb{E} [h_{n\eta}^*(u, \theta^*, \eta^*) K^{-1} \bar{h}_n^*(u, \theta^*, \eta^*)] = 0$. \square

REFERENCES

- [1] L. BACHELIER *Théorie de la spéculation*, in Annales Scientifiques de l'École Normale Supérieure, 3 (17) (1900), pp. 2186.
- [2] F. BLACK, M. SCHOLES *The pricing of options and corporate liabilities*, in Journal of Political Economy, 81 (1973), pp. 345-637.
- [3] J. JACOD, A.N. SHIRYAEV, *Limit theorems for stochastic processes (2. ed.)*, Springer, (2002)
- [4] K. SATO *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, (1999)
- [5] 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.
- [6] M. CARRASCO AND J.-P. FLORENS, *Generalization of GMM to a continuum of moment conditions*, in Econometric Theory, 16 (06) (2000), pp. 797-834.
- [7] M. CARRASCO AND J. P. FLORENS, *Efficient GMM estimation using the empirical characteristic function*, in Idei working papers, 140 (2002)
- [8] 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.
- [9] R. CONT, P. TANKOV, *Financial Modelling with Jump Processes*, in Journal of the American Statistical Association, 101 (2006), pp. 1315-1316.
- [10] 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.
- [11] B. MADAN, P. CARR AND C. CHANG, *The Variance Gamma Process and Option Pricing*, in European Finance Review, 2 (1998), pp. 79-105.
- [12] B. MANDELBROT, *The Variation of Certain Speculative Prices*, in Journal of Business, 35, (1963)
- [13] J. YU, *Empirical characteristic function estimation and its applications*, in Econometric Reviews, 23 (2) (2004), pp. 93-123.
- [14] Y. MIYAHARA AND A. NOVIKOV, *Geometric Lévy Process Pricing Model*, in Research Paper Series 66, Quantitative Finance Research Centre, University of Technology, Sydney, (2001)
- [15] L. GERENCSÉR, *On the martingale approximation of the estimation error of ARMA parameters*, in System & Control Letters, 15 (1990), pp. 417-423.
- [16] L. GERENCSÉR, *On a class of mixing processes*, in Stochastics, 26 (1989), pp. 165-191.
- [17] L. GERENCSÉR, GY. MICHALETZKY, Z. REPPA *A two-step maximum-likelihood identification of non-Gaussian systems*, in Proceedings of the 15th IFAC World Congress 15 (2002)