



# CARMA Approximations and Estimation

Silvia Lavagnini\*

Department of Mathematics, University of Oslo, Blindern, Norway

CARMA( $p, q$ ) processes are compactly defined through a stochastic differential equation (SDE) involving  $q + 1$  derivatives of the Lévy process driving the noise, despite this latter having in general no differentiable paths. We replace the Lévy noise with a continuously differentiable process obtained via stochastic convolution. The solution of the new SDE is then a stochastic process that converges to the original CARMA in an  $L^2$  sense. CARMA processes are largely applied in finance, and this article mathematically justifies their representation via SDE. Consequently, the use of numerical schemes in simulation and estimation by approximating derivatives with discrete increments is also justified for every time step. This provides a link between discrete-time ARMA models and CARMA models in continuous-time. We then analyze an Euler method and find a rate of convergence proportional to the square of the time step for discretization and to the inverse of the convolution parameter. These must then be adjusted to get accurate approximations.

## OPEN ACCESS

### Edited by:

Ehsan Azmoodeh,  
Ruhr University Bochum, Germany

### Reviewed by:

Marina Resta,  
University of Genoa, Italy  
Tommi Sottinen,  
University of Vaasa, Finland

### \*Correspondence:

Silvia Lavagnini  
silval@math.uio.no

### Specialty section:

This article was submitted to  
Mathematical Finance,  
a section of the journal  
Frontiers in Applied Mathematics and  
Statistics

**Received:** 30 April 2020

**Accepted:** 27 July 2020

**Published:** 11 November 2020

### Citation:

Lavagnini S (2020) CARMA  
Approximations and Estimation.  
Front. Appl. Math. Stat. 6:37.  
doi: 10.3389/fams.2020.00037

**Keywords:** CARMA processes, stochastic differential equations, Lévy processes, Euler approximation, stochastic convolution, normal inverse Gaussian process

## 1. INTRODUCTION

Continuous-time autoregressive moving-average processes, CARMA in short, represent the continuous-time version of the well-known ARMA models. Because of their linear structure, they are convenient for modeling empirical data, hence they find application in several fields, particularly in finance. Indeed, with a Lévy process as a driver for the noise, one obtains a rich class of possibly heavy-tailed continuous-time stationary processes, and such heavy tails are frequently observed in finance [1]. In Barndorff-Nielsen and Shepard [2], for example, a Lévy-driven Ornstein-Uhlenbeck process (corresponding to a CAR(1)) is employed in a stochastic volatility model, and [3] extend the idea to a wider class of Lévy-driven CARMA processes with non-negative kernel, allowing for non-negative, heavy-tailed processes with a larger range of autocovariance functions.

In view of financial applications, Brockwell et al. [4] derive efficient estimates for the parameters of a non-negative Lévy-driven CAR(1) process, and Brockwell et al. [5] generalize the technique to higher-order CARMA processes with a non-negative kernel. In Benth et al. [6] and Benth and Benth [7], the authors exploit the link between discrete-time ARMA and continuous-time ARMA to model the observed processes in a continuous manner, starting from discrete-time observations-based estimates, in the context of energy markets and weather derivatives. It is also worth mentioning that CARMA processes allow us to construct a wide range of long-memory processes, as shown, e.g., in Marquardt [8] and Marquardt and James [9].

Among the several financial applications, in Paschke and Prokopczuk [10, 11] a CARMA process is employed for the long-term equilibrium component in the two-factors spot price model introduced by Schwartz and Smith [12] (which uses an Ornstein-Uhlenbeck process), while Benth

and Koekebakker [13] use a CARMA process for the short-term factor instead. In Benth et al. [14] and García et al. [15], a CARMA process models the spot price in the electricity market, while in Benth et al. [6] and Benth and Benth [7] it models weather variables, such as temperature and wind speed. Finally, Andresen et al. [16] apply CARMA processes to interest rates modeling and [17] to the dynamics of regional ocean freight rates.

Formally, CARMA( $p, q$ ) processes are compactly defined through a stochastic differential equation (SDE) involving  $q + 1$  derivatives of the Lévy process driving the noise, despite the latter having in general no differentiable paths. Numerical schemes, such as the Euler scheme, are also applied both for simulations and estimation of CARMA processes. These schemes approximate derivatives with discrete increments with respect to a time step  $\Delta$ , and their convergence for  $\Delta$  approaching zero is indeed a well-studied topic. However, this opens up for a discrepancy since it is not mathematically clear what happens when the time step approaches zero, as, in general, the derivatives of a Lévy process are not well-defined.

We give a mathematical justification for the representation via SDE of a CARMA process, and, consequently, for the discrete approximations used for simulating and estimating the continuous-time dynamics of CARMA processes. This is done by replacing the Lévy process  $L$  driving the noise with a process smooth enough so that the SDE makes mathematical sense. For  $\varepsilon > 0$  we construct a continuously differentiable function  $L_\varepsilon$ , which converges in the  $L^2$  sense to the original Lévy process in the space  $[0, T] \times \Omega$ , for  $T < +\infty$ . More precisely, the process  $L_\varepsilon$  is constructed via stochastic convolution of  $L$  with a smooth sequence of functions  $\{\phi_\varepsilon\}_{\varepsilon>0}$  depending on the convolution parameter  $\varepsilon > 0$  and with first derivative converging to the delta-Dirac function when  $\varepsilon$  approaches zero. We show a rate of convergence of order 1.

For the  $L_\varepsilon$  continuous differentiable, the framework developed solves the differentiability issue for CAR( $p$ ) processes in which no more than the first derivative of the driving noise process is required. We can, however, make use of the result in Brockwell et al. [5, Proposition 2]. This states that, under some conditions on the coefficients of the SDE, a CARMA( $p, q$ ) process can be decomposed into a sum of  $p$  dependent CAR(1) processes. The approach can then be extended to CARMA( $p, q$ ) processes of every order  $q \geq 0$  by considering  $p$  different approximations based on the continuous differentiable  $L_\varepsilon$  and taking their sum. This justifies an  $L_\varepsilon$ -only continuous differentiable, as the general case  $q \geq 0$  is still covered.

We thus replace the Lévy process  $L$  with  $L_\varepsilon$  in the SDE defining the CARMA process  $X$ . This gives a different SDE, which depends on  $\varepsilon$ , whose solution is the new process  $X_\varepsilon$ . We show that  $X_\varepsilon$  converges to  $X$  in  $L^2$  ( $[0, T] \times \Omega$ ) and that the convergence is of order 2.

The  $\varepsilon$ -approximation of the Lévy process justifies the use of an Euler scheme for every time step  $\Delta$ . We present a possible approach for simulating  $L_\varepsilon$  starting from a simulated path of the Lévy process  $L$  and then  $X_\varepsilon$  starting from  $L_\varepsilon$ . This, for a fixed time interval  $\Delta$ , gives a new noise process  $L_\varepsilon^\Delta$  and a third process  $X_\varepsilon^\Delta$ . We show a rate of convergence for both  $L_\varepsilon^\Delta$  and  $X_\varepsilon^\Delta$  proportional to  $\Delta^2$  and inversely so to  $\varepsilon$ . This highlights that the

two parameters  $\Delta$  and  $\varepsilon$  must be chosen in a way to get accurate simulations and prevent them from exploding.

We finally consider a possible Euler scheme approach for estimating the CARMA dynamics starting from the simulated time series. It is worth mentioning that the process  $L_\varepsilon$  is not a Lévy process. Due to the convolution operator, it indeed fails on the crucial property of independent increments. We show in particular that the increments depend on the ratio  $\frac{\Delta}{\varepsilon}$ , and, as this ratio becomes smaller, the increments become more dependent on each other. This has a negative drawback when it comes to estimation. We find indeed that the Euler scheme recovers the true dynamics of the CARMA process only if  $\varepsilon$  is smaller than  $\Delta$ .

All the results are deeply investigated and are also confirmed by numerical experiments. In particular, we consider two different Lévy processes, a Brownian motion and a normal inverse Gaussian (NIG) process, and CARMA processes of two different orders, a Gaussian CAR(1) and a Gaussian CARMA(3, 2). Crucial to all the numerical simulations is the convolution property of both the Gaussian and the NIG distribution function, which allows us to compare path-wise different time grids by simulating only one path of the Lévy process.

Since CARMA processes have significant application within finance, from the modeling of the electricity spot price to the modeling of wind speed and temperature, with the present study, we justify its mathematical formulation and application in a rigorous way. Moreover, this may open for further applications in the context of sensitivity analysis with respect to paths of the driving noise process. Indeed, since the new process  $L_\varepsilon$  admits derivatives, one could potentially analyze the sensitivity of the CARMA process, and hence of the variable of interest (e.g., the spot price), with respect to the Lévy noise in a path-wise sense. It might also be that when observing a signal, the driving-noise process is more regular than, e.g., a Brownian motion. This might then be modeled by the regular process here introduced and studied.

The rest of the paper is organized as follows. In section 2 we construct the approximating process  $L_\varepsilon$  by stochastic convolution, and we study the rate of convergence. In section 3 we introduce CARMA processes and their spectral representation. We then show the convergence of the new process  $X_\varepsilon$  to the original CARMA process, and we briefly recover the general case  $q > 0$ . In section 4 we study an Euler scheme to simulate the process  $X_\varepsilon$  starting from the simulations of  $L_\varepsilon$ , and then to calibrate the parameters of the CARMA dynamics starting from the discrete time series. Appendix A in **Supplementary Material** contains the proofs of the main results.

## 2. CONSTRUCTION AND CONVERGENCE RESULTS

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. We consider a one-dimensional Lévy process  $L$ , which we assume to be a square-integrable martingale with characteristic triplet  $(\kappa, \Sigma, \nu)$ ,  $\nu$  being the Lévy measure. From Cont and Tankov [18, Proposition 3.13],

for every  $t > 0$  the variance of  $L(t)$  is of the form

$$\text{Var}(L(t)) = t\sigma^2 \quad \text{with} \quad \sigma^2 := \Sigma + \int_{-\infty}^{\infty} x^2\nu(dx). \quad (2.1)$$

Moreover, from Cont and Tankov [18, Proposition 3.18], we also know that  $L$  is a martingale if, and only if,

$$\int_{|x|\geq 1} |x|\nu(dx) < \infty \quad \text{and} \quad \kappa + \int_{|x|\geq 1} x\nu(dx) = 0. \quad (2.2)$$

We now consider a sequence of measurable functions  $\{\phi_\varepsilon\}_{\varepsilon>0}: [0, T] \rightarrow \mathbb{R}$  such that, for every  $\varepsilon > 0$ ,  $\phi_\varepsilon$  is twice continuously differentiable on the interval  $[0, T]$  and  $\phi_\varepsilon(0) = 0$ . We then construct the process  $\{L_\varepsilon(t), t \geq 0\}$  by stochastic convolution of the Lévy process  $L$  with  $\phi_\varepsilon$ :

$$L_\varepsilon(t) := \int_0^t \phi_\varepsilon(t-s)dL(s). \quad (2.3)$$

$L_\varepsilon$  is well-defined in  $[0, T]$  for every  $T \in ]0, +\infty[$ , and, in particular, the condition  $\phi_\varepsilon(0) = 0$  is used to prove the following result, where the symbol  $*$  denotes the convolution operator.

**Lemma 2.1.** *The process  $\{L_\varepsilon(t), t \geq 0\}$  in Equation (2.3) can be expressed by*

$$L_\varepsilon(t) = (\phi'_\varepsilon * L)(t)|_{[0,t]} = \int_0^t \phi'_\varepsilon(t-s)L(s)ds. \quad (2.4)$$

*Proof:* See Appendix A.1 in **Supplementary Material**.

For  $\psi_\varepsilon(x) := \phi'_\varepsilon(x)$ , since  $\phi_\varepsilon(0) = 0$ , we write that  $\phi_\varepsilon(t) = \int_0^t \psi_\varepsilon(u)du$  so that  $\phi_\varepsilon$  is twice continuously differentiable for any continuously differentiable  $\psi_\varepsilon$ . We introduce the function  $\psi(x) := \frac{2}{\sqrt{2\pi}} \exp\{-\frac{1}{2}x^2\}$ , and for every  $\varepsilon > 0$  we define

$$\psi_\varepsilon(x) := \frac{1}{\varepsilon} \psi\left(\frac{x}{\varepsilon}\right) = \frac{2}{\sqrt{2\pi}\varepsilon} \exp\left\{-\frac{x^2}{2\varepsilon^2}\right\}. \quad (2.5)$$

Let us notice that  $\int_{\mathbb{R}} \psi(x)dx = \int_{\mathbb{R}} \psi_\varepsilon(x)dx = 2$ , while the total area of the mollifier function usually equals 1. However, we construct the process  $\{L_\varepsilon(t), t \geq 0\}$  by stochastic convolution restricted to the interval  $[0, t]$ , and, denoting by  $\Phi$  the standard Gaussian cumulative distribution function, the area of  $\psi_\varepsilon$  restricted to  $[0, t]$  is

$$(\phi_\varepsilon(t)) = \int_0^t \psi_\varepsilon(x)dx = 2 \left( \Phi\left(\frac{t}{\varepsilon}\right) - \frac{1}{2} \right) \xrightarrow{\varepsilon \downarrow 0} 1, \quad (2.6)$$

so that the unitary total integral property for  $\psi_\varepsilon$  is preserved in the limit.

By differentiating Equation (2.4) with respect to  $t$ , according to the Leibniz rule, we get

$$L'_\varepsilon(t) = \phi'_\varepsilon(0)L(t) + \int_0^t \phi''_\varepsilon(t-s)L(s)ds. \quad (2.7)$$

This shows that  $L_\varepsilon \in C^1([0, T])$ , though  $L_\varepsilon \notin C^n([0, T])$  for any  $n > 1$ , because of the term  $L(t)$  which is not differentiable, unless  $\phi'_\varepsilon(0) = \psi_\varepsilon(0) = 0$ , which is not our case. Hence, we start with the Lévy process  $L$ , which is not differentiable, and we end up with  $L_\varepsilon$ , which is continuously differentiable. For a CARMA( $p, q$ ) process, we need  $L_\varepsilon \in C^{q+1}([0, T])$ . From Equation (2.7), this means to have  $\phi_\varepsilon^{(m)}(0) = 0$  for every  $0 \leq m \leq q + 1$ ,  $\phi_\varepsilon^{(m)}$  being the  $m$ -th derivative of  $\phi_\varepsilon$ . However, for our particular choice of the convolution function  $\psi_\varepsilon$ , these conditions of null derivatives in zero are not satisfied. Another possible choice would have been, for example, to take a sequence of gamma densities with scale parameter  $\theta_\varepsilon$  converging to 0 as  $\varepsilon$  converges to 0 and shape parameter  $k > 0$ , such as  $d_\varepsilon(t) = \frac{1}{\Gamma(k)\theta_\varepsilon^k} t^{k-1} e^{-\frac{t}{\theta_\varepsilon}}$ . For  $k \geq q + 1$ , the first  $q$  derivatives of  $d_\varepsilon$  are indeed null in  $t = 0$ .

With  $L_\varepsilon$  constructed above, we thus address the differentiability issue in the stochastic differential equation defining the CARMA process only for  $q = 0$ , namely when the process is a CAR( $p$ ). However, Brockwell et al. [5, Proposition 2] shows that, under some conditions, we can express a CARMA( $p, q$ ) process as the sum of  $p$  dependent and possibly complex-valued CAR(1) processes. The idea is then to construct the  $\varepsilon$ -approximation with  $L_\varepsilon$  in Equation (2.3) for each of the  $p$  processes of CAR(1)-type and obtain the  $\varepsilon$ -CARMA( $p, q$ ) process by summing them up. We can thus deal with  $q > 0$ , but it must be done separately. With the gamma densities mentioned above, for example, the special treatment for  $q > 0$  would not be necessary. However, we started this analysis from CAR( $p$ ) processes, and in this setting the scaled Gaussian densities  $\psi_\varepsilon$  make the job. We shall analyze the case  $q > 0$  in section 3.2.

A first important fact about  $L_\varepsilon$  in Equation (2.3) is that the convolution operation does not preserve the properties of the Lévy process  $L$ .

**Proposition 2.2.** *The process  $\{L_\varepsilon(t), t \geq 0\}$  is not a Lévy process.*

*Proof:* See Appendix A.2 in **Supplementary Material**.

By Proposition 2.2, for  $L = B$ , a Brownian motion, the approximated process  $B_\varepsilon$  is no longer a Brownian motion. We can, however, show that the distribution of  $B$  is preserved.

**Lemma 2.3.** *For every  $\varepsilon > 0$ , the process  $\{B_\varepsilon(t), t \geq 0\}$  has Gaussian distribution with zero mean and variance*

$$\begin{aligned} \text{Var}(B_\varepsilon(t)) &= \int_0^t \phi_\varepsilon^2(s)ds \\ &= t + 4t\Phi\left(\frac{t}{\varepsilon}\right) \left( \Phi\left(\frac{t}{\varepsilon}\right) - 1 \right) + \\ &+ 4\varepsilon \left( 2\varphi\left(\frac{t}{\varepsilon}\right) \Phi\left(\frac{t}{\varepsilon}\right) - \frac{1}{\sqrt{2\pi}} \left( \Phi\left(\frac{\sqrt{2}t}{\varepsilon}\right) - \frac{1}{2} \right) - \varphi\left(\frac{t}{\varepsilon}\right) \right). \end{aligned}$$

*Proof:*  $B_\varepsilon$  is defined in Equation (2.3) as the stochastic integral of  $\phi_\varepsilon$  with respect to the Brownian motion  $B$ . Then  $B_\varepsilon$  has Gaussian distribution with zero mean and variance  $\int_0^t \phi_\varepsilon^2(s)ds$ , which can be explicitly calculated.  $\square$

Let us point out that Lemma 2.3 is valid for every choice of the functions  $\{\phi_\varepsilon\}_{\varepsilon>0}$ , with the obvious modification for the variance formula  $\int_0^t \phi_\varepsilon^2(s)ds$ . Moreover, from Lemma 2.3, we get that  $\lim_{\varepsilon \downarrow 0} \text{Var}(B_\varepsilon(t)) = t$ , as one would expect. **Figure 1** shows indeed the typical bell-shape for a Gaussian distribution. We sampled the process  $B_\varepsilon$  at time  $t = 1$  for 5,000 times and different values of  $\varepsilon$ . The empirical (black line) and the theoretical (red line) distributions (according to Lemma 2.3) are also shown. For  $\varepsilon = 1$ , there is a certain discrepancy between empirical and theoretical distributions, which may be due to the fact that for big values of  $\varepsilon$  the increments of  $B_\varepsilon$  are more dependent from each other than for smaller values of  $\varepsilon$ . The same phenomena will be encountered in section 4.3.

### 2.1. Convergence of the Lévy Approximation

We prove now the convergence of the process  $L_\varepsilon$  to the Lévy process  $L$ .

**Theorem 2.4.** *The process  $\{L_\varepsilon(t), t \geq 0\}$  converges to  $\{L(t), t \geq 0\}$  in  $L^2([0, T] \times \Omega)$ .*

*Proof:* See Appendix A.3 in **Supplementary Material**. The proof is based on some results from Benth et al. [19, Appendix A].

**Figure 2** shows graphically the convergence stated in Theorem 2.4 for a Brownian motion and for a normal inverse Gaussian process with parameters  $(\alpha, \beta, \delta, \mu) = (1, 0, \Delta^{true}, 0)$ ,  $\Delta^{true}$  being the time step used in the Euler discretization (see section 4). As  $\varepsilon$  decreases, the error between  $L_\varepsilon$  and  $L$  decreases. Even with a large  $\varepsilon$  ( $\varepsilon = 1$ ), the process  $L_\varepsilon$  is able to capture the shape of  $L$ . It seems however to be slightly shifted to the right.

We compute explicitly the squared error between  $L_\varepsilon$  and  $L$  at the terminal point  $t = T$ .

**Lemma 2.5.** *For every  $\varepsilon > 0$  and  $T < +\infty$ , the squared error between  $L_\varepsilon$  and  $L$  is given by*

$$\begin{aligned} \mathbb{E} \left[ (L_\varepsilon(T) - L(T))^2 \right] &= 4\sigma^2 T \left( \Phi \left( \frac{T}{\varepsilon} \right) - 1 \right)^2 + \\ &+ 4\sigma^2 \varepsilon \left( \frac{1}{\sqrt{2\pi}} + 2\varphi \left( \frac{T}{\varepsilon} \right) \left( \Phi \left( \frac{T}{\varepsilon} \right) - 1 \right) + \right. \\ &\quad \left. - \frac{1}{\sqrt{\pi}} \left( \Phi \left( \frac{\sqrt{2}T}{\varepsilon} \right) - \frac{1}{2} \right) \right), \end{aligned}$$

for  $\sigma^2$  in Equation (2.1) and  $\varphi$  and  $\Phi$  respectively the density and the cumulative distribution functions of a standard Gaussian random variable.

*Proof:* The proof applies the integration by parts formula several times. □

The following lemma concerning the cumulative Gaussian distribution function  $\Phi$  is useful.

**Lemma 2.6.** *The function  $\Phi$  can be expanded by integration by parts into the series*

$$\Phi(x) = \frac{1}{2} + \frac{e^{-x^2/2}}{\sqrt{2\pi}} \left( x + \sum_{n=1}^{\infty} \frac{x^{2n+1}}{(2n+1)!!} \right), \tag{2.8}$$

!! denoting the double factorial. Moreover, for large  $x$ , the asymptotic expansion is also valid

$$\Phi(x) = 1 - \frac{e^{-x^2/2}}{x\sqrt{2\pi}} \left( 1 + \sum_{n=1}^{\infty} (-1)^n \frac{(2n-1)!!}{x^{2n}} \right). \tag{2.9}$$

We now use the expansion (2.8) to calculate the rate of convergence of  $L_\varepsilon$  to the Lévy process  $L$ , starting from Lemma 2.5. The expansion (2.9) will be used further in the paper.

**Proposition 2.7.** *For every  $\varepsilon > 0$  and  $T < +\infty$ , there exists a constant  $K_1 > 0$  such that*

$$\mathbb{E} \left[ (L_\varepsilon(T) - L(T))^2 \right] \leq K_1 \varepsilon + O(\varepsilon^4).$$

*Proof:* See Appendix A.4 in **Supplementary Material**.

Similarly to Lemma 2.3, the rate of convergence found in Proposition 2.7 can be calculated explicitly for every choice of the functions  $\{\phi_\varepsilon\}_{\varepsilon>0}$ , possibly obtaining a different order. The  $L^2$  norm of the approximation error studied in Proposition 2.7 as a function of  $\varepsilon$  is in **Figure 3A** for the Brownian motion and in **Figure 3B** for the NIG Lévy process. In the next section we shall introduce CARMA processes in a more rigorous way and study their  $\varepsilon$ -approximation.

### 3. CARMA REPRESENTATION AND CONVERGENCE RESULTS

For  $p > q$  and  $t \geq 0$ , a second-order Lévy-driven continuous-time ARMA process of order  $(p, q)$  is formally defined by the stochastic differential equation

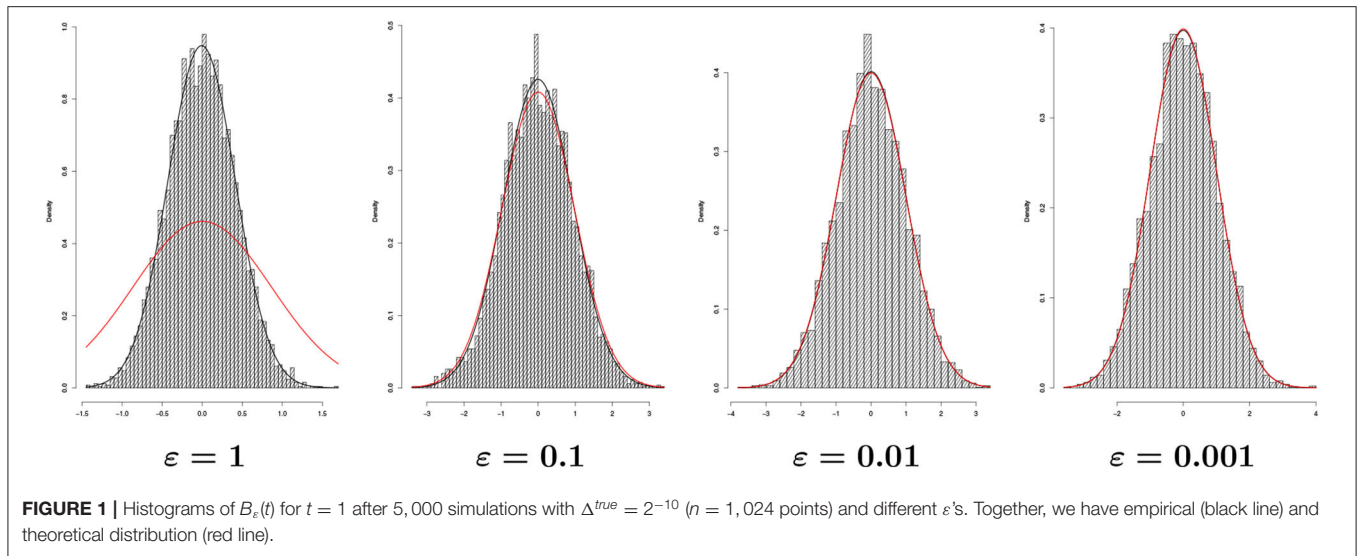
$$P \left( \frac{d}{dt} \right) X(t) = Q \left( \frac{d}{dt} \right) \frac{d}{dt} L(t), \tag{3.1}$$

where  $P$  and  $Q$  are the characteristic polynomials of the CARMA process of the form

$$\begin{aligned} P(z) &= z^p + a_1 z^{p-1} + \dots + a_p \\ Q(z) &= b_0 + b_1 z + \dots + b_{p-1} z^{p-1} \end{aligned} \tag{3.2}$$

with  $a_j \geq 0, j = 1, \dots, p$ , and  $a_p > 0, b_q = 1$  and  $b_j = 0$  for  $q < j < p$ . Equation (3.1) is, however, usually interpreted by means of its state-space representation

$$\begin{cases} d\mathbf{Y}(t) = \mathbf{A}\mathbf{Y}(t)dt + \mathbf{e}_p dL(t) \\ X(t) = \mathbf{b}^\top \mathbf{Y}(t) \end{cases}, \tag{3.3}$$



where  $\top$  denotes the transpose operator, and

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \dots & -a_1 \end{pmatrix}, \mathbf{e}_p = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}.$$

By Itô's rule, the solution of the first equation in (3.3) at time  $t \geq s$  is given by

$$\mathbf{Y}(t) = e^{A(t-s)}\mathbf{Y}(s) + \int_s^t e^{A(t-u)}\mathbf{e}_p dL(u).$$

Moreover, from Brockwell et al. [5, Proposition 1], if the eigenvalues  $\lambda_1, \dots, \lambda_p$  of  $A$  satisfy the condition

$$\text{Re}(\lambda_j) < 0 \quad \text{for } j = 1, \dots, p, \tag{3.4}$$

then the process  $\{\mathbf{Y}(t), t \geq 0\}$  is strictly stationary and the CARMA( $p, q$ ) process is given by

$$X(t) = \int_0^t g(t-s)dL(s), \tag{3.5}$$

$$g(x) := \mathbf{b}^\top e^{Ax} \mathbf{e}_p, \quad x \geq 0. \tag{3.6}$$

The function  $g$  is referred to as the kernel function associated to the CARMA process. We also notice that the process  $X$  is a causal function of  $L$ , in the sense that  $X(t)$  is independent of  $\{L(s) - L(t), s \geq t\}$  for all  $t \geq 0$ . We shall assume that condition (3.4) is satisfied, referring to it as the causality condition or stationarity condition.

In a similar way, we introduce the state-space representation for  $X_\varepsilon$ . This is the process obtained by replacing the Lévy process  $L$  with the smoother  $L_\varepsilon$ , namely

$$\begin{cases} d\mathbf{Y}_\varepsilon(t) = A\mathbf{Y}_\varepsilon(t)dt + \mathbf{e}_p dL_\varepsilon(t) \\ X_\varepsilon(t) = \mathbf{b}^\top \mathbf{Y}_\varepsilon(t) \end{cases}. \tag{3.7}$$

The solution to Equation (3.7) can be expressed in terms of the kernel function  $g$  by

$$X_\varepsilon(t) = \int_0^t g(t-s)dL_\varepsilon(s). \tag{3.8}$$

We point out that the kernel function  $g$  in Equation (3.6) is continuously differentiable. Its value at  $x = 0$  is given in the following lemma, where  $\delta$  denotes the Kronecker delta.

**Lemma 3.1.** For  $p > q$ , the value of the kernel function at  $x = 0$  is equal to  $g(0) = \delta_{p,q+1}$ .

*Proof:* The proof follows by direct verification. □

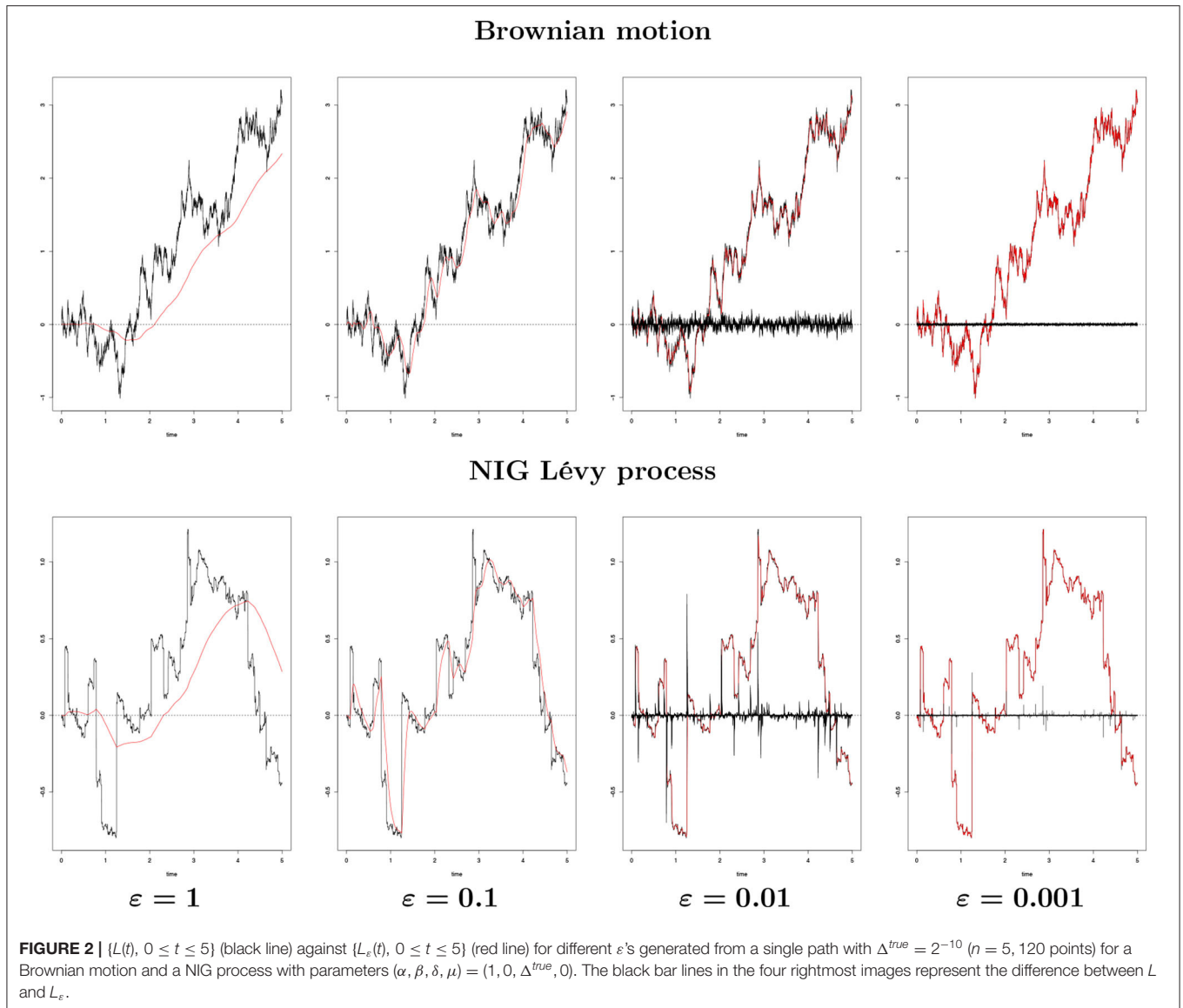
From the basic algebra analysis, if  $A$  has distinct eigenvalues, then the matrix exponential  $e^{Ax}$  can be decomposed into the matrix product

$$e^{Ax} = V e^{\Lambda x} V^{-1}, \tag{3.9}$$

$\Lambda$  and  $V$  being the eigenvalues and eigenvectors matrices. From Brockwell et al. [5, Remark 2], the eigenvalues  $\lambda_1, \dots, \lambda_p$  of the matrix  $A$  coincide with the roots of the polynomial  $P$  in Equation (3.2), and the  $i$ -th eigenvector is of the form  $\mathbf{v}_i = [1, \lambda_i, \lambda_i^2, \dots, \lambda_i^{p-1}]^\top$ . From Brockwell et al. [5, Proposition 2], we also get that  $V^{-1}\mathbf{e}_p = [\gamma_1, \gamma_2, \dots, \gamma_p]^\top$ , where

$$\gamma_i := \frac{1}{P'(\lambda_i)} = \frac{1}{\prod_{\substack{1 \leq m \leq p \\ m \neq i}} (\lambda_i - \lambda_m)} \quad \text{for } i = 1, \dots, p, \tag{3.10}$$

$P'$  being the derivative of the polynomial  $P$ . We then find the following spectral representations for the kernel function  $g$  and its first derivative  $h(x) := g'(x)$ .



Lemma 3.2. The functions  $g$  and  $h$  can be expressed in terms of the spectral representations:

$$g(x) = \sum_{i,j=1}^p b_j \lambda_i^j \gamma_i e^{\lambda_i x} \quad \text{and} \quad h(x) = \sum_{i,j=1}^p \tilde{b}_j \lambda_i^j \gamma_i e^{\lambda_i x},$$

for  $\gamma_i$  in Equation (3.10), and  $\tilde{b}_j = [\mathbf{b}^\top A]_j, j = 1, \dots, p$ .

*Proof:* The proof comes from Equations (3.6) and (3.9) by writing explicitly the matrix product  $g(x) = \mathbf{b}^\top V e^{\Lambda x} V^{-1} \mathbf{e}_p$  and  $h(x) = g'(x) = \mathbf{b}^\top A V e^{\Lambda x} V^{-1} \mathbf{e}_p$ .  $\square$

These representations will be used for proving the convergence results in the next section.

### 3.1. Convergence of the CARMA Approximation

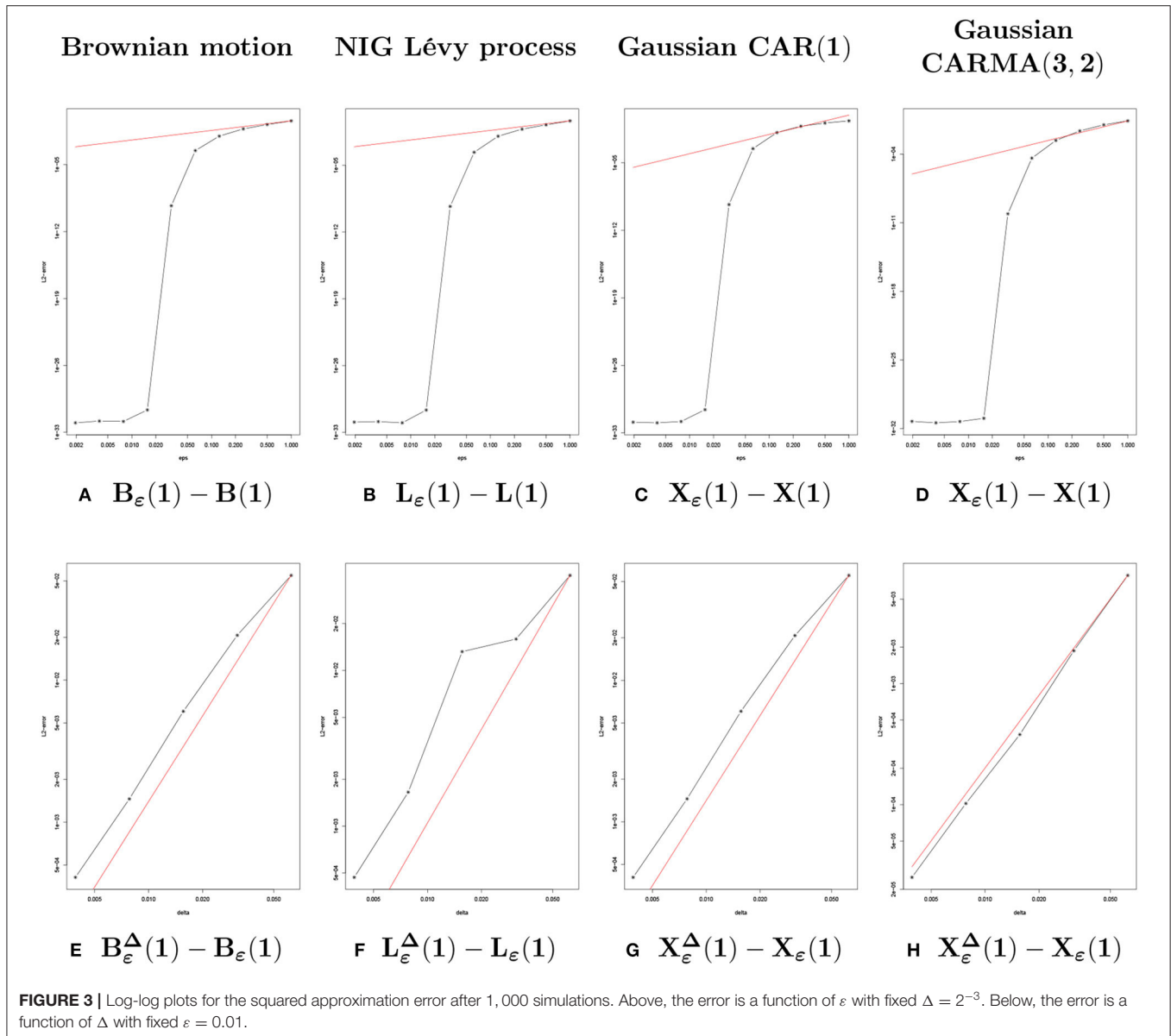
We prove the convergence of  $X_\epsilon$  to  $X$ . We first start with the following preliminary result.

Lemma 3.3. The processes  $\{X(t), t \geq 0\}$  and  $\{X_\epsilon(t), t \geq 0\}$  can be rewritten by

$$X(t) = \int_0^t h(t-s)L(s)ds \quad \text{and} \quad X_\epsilon(t) = \int_0^t h_\epsilon(t-s)L(s)ds.$$

Here  $h(x) = g'(x)$  and  $h_\epsilon(x) := g(0)\psi_\epsilon(x) + \int_0^x h(x-v)\psi_\epsilon(v)dv$ , for  $g(0)$  given in Lemma 3.1.

*Proof:* See Appendix A.5 in **Supplementary Material**.



From Lemma 3.2, we also derive the spectral representation

$$h_\epsilon(x) = g(0)\psi_\epsilon(x) + \sum_{i,j=1}^p \tilde{b}_j \lambda_i^j \gamma_i \int_0^x e^{\lambda_i(x-v)} \psi_\epsilon(v) dv, \quad (3.11)$$

and prove the first convergence result.

**Proposition 3.4.** *For every  $x > 0$ , the function  $h_\epsilon(x)$  converges pointwise to  $h(x)$ .*

*Proof:* See Appendix A.6 in **Supplementary Material**.

In Proposition 3.4, we do not require the limit to hold in  $x = 0$ . However, this does not affect the convergence in  $L^2$  since  $X$  and  $X_\epsilon$  are both integral functions, and  $\{x = 0\}$  is a set of measure 0. Indeed, we can prove the required convergence for  $X_\epsilon$ .

**Theorem 3.5.** *The process  $\{X_\epsilon(t), t \geq 0\}$  converges to  $\{X(t), t \geq 0\}$  in  $L^2([0, T] \times \Omega)$ .*

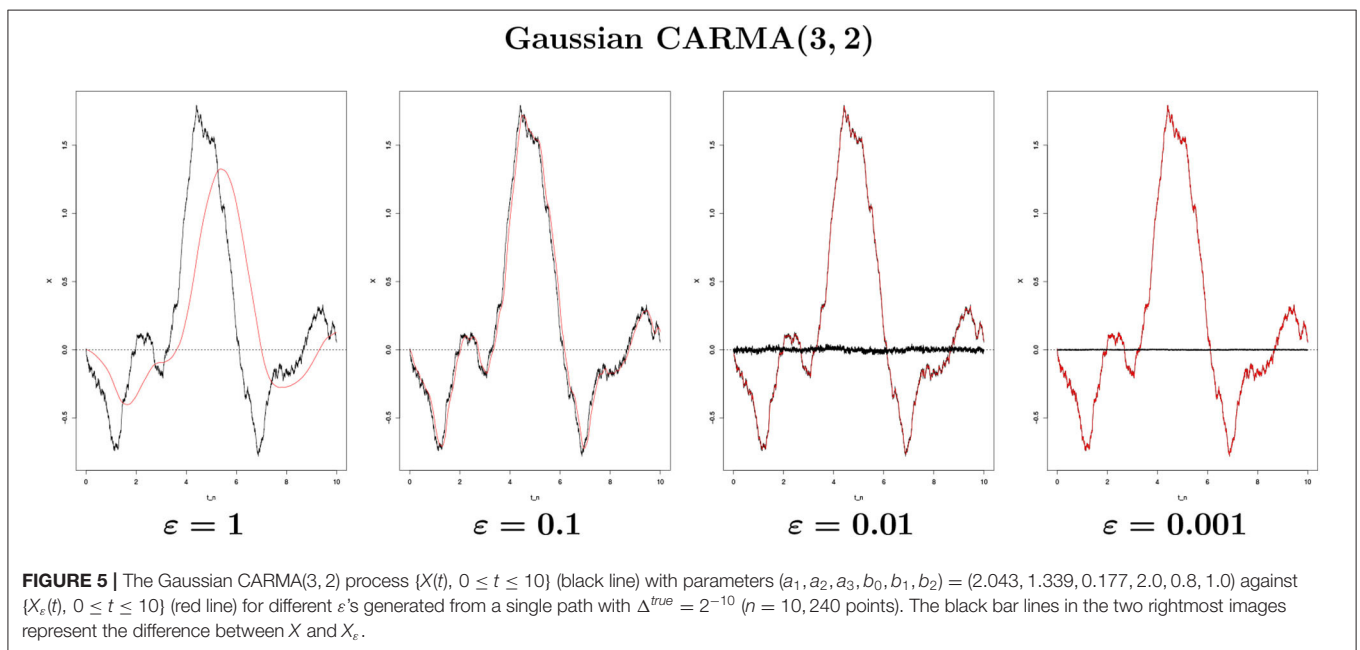
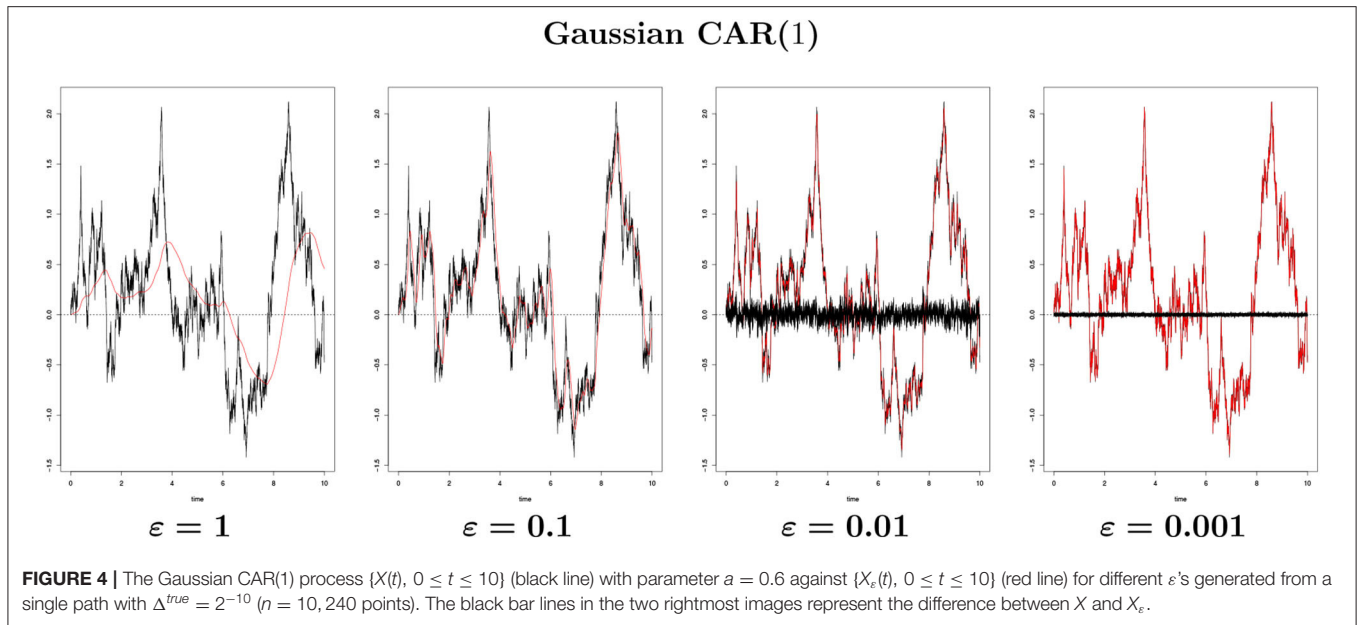
*Proof:* See Appendix A.7 in **Supplementary Material**.

**Figure 4** shows the convergence proven in Theorem 3.5 for a Gaussian CAR(1) process with parameter  $a = 0.6$ . We notice that  $X_\epsilon$  captures the shape of  $X$  even with  $\epsilon$  large, and we observe a right-shifting of  $X_\epsilon$  with respect to  $X$ , as noticed already for the Lévy process approximation.

We now study explicitly the rate of convergence of  $X_\epsilon$  to the CARMA process  $X$ .

**Proposition 3.6.** *For every  $\epsilon > 0$  and  $T < +\infty$ , there exists a constant  $K_2 > 0$  such that*

$$\mathbb{E} \left[ (X_\epsilon(T) - X(T))^2 \right] \leq K_2 \epsilon^2 + O(\epsilon^3).$$



*Proof:* See Appendix A.8 in **Supplementary Material**.

The rate of convergence obtained in Proposition 3.6 is shown in **Figure 3C** for a CAR(1).

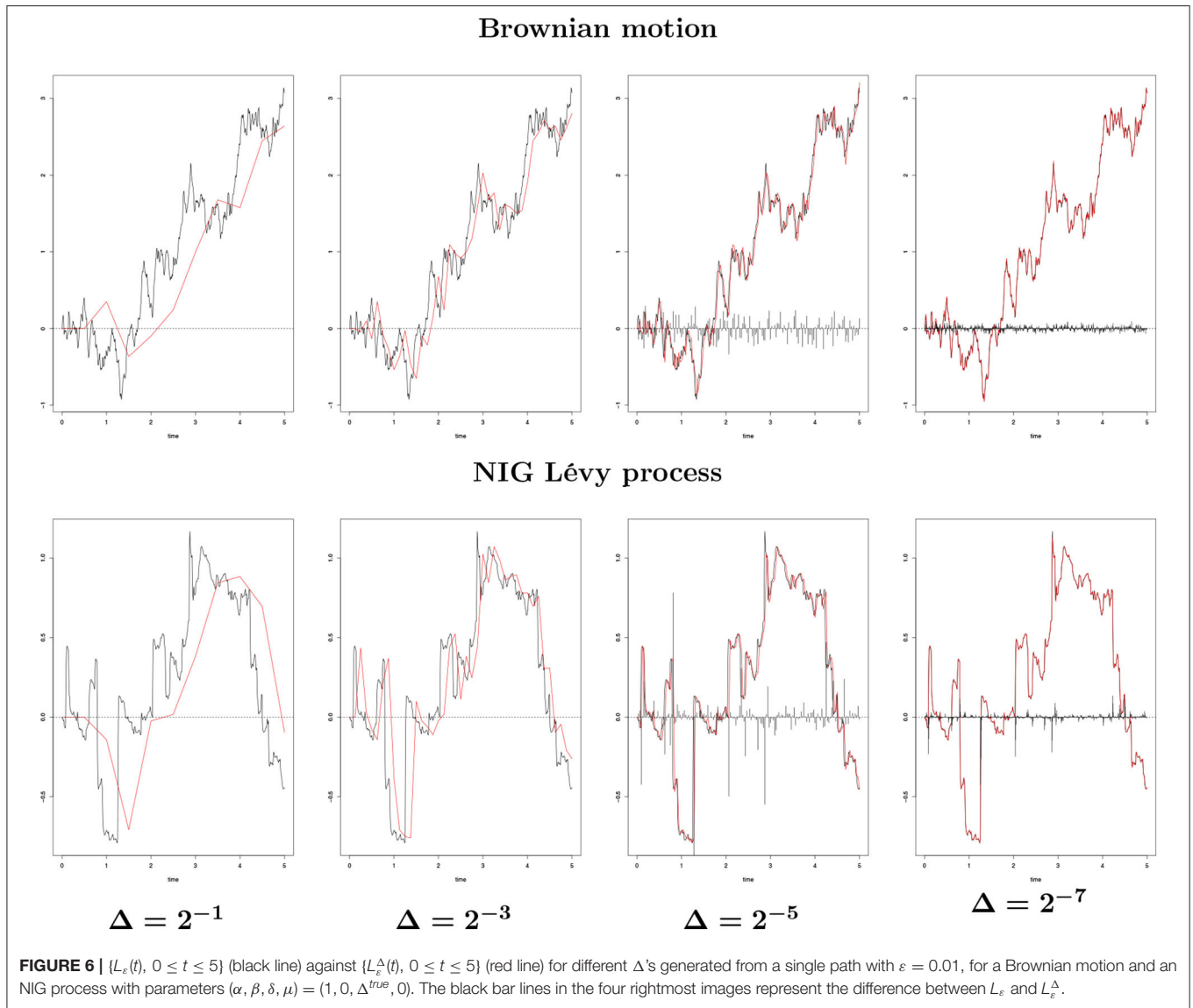
We point out that the results in this section have been stated and proved for a general CARMA( $p, q$ ) process. The convergence results are indeed valid for every  $q \geq 0$ , even when  $L_\epsilon$  is not smooth enough to make the SDE having mathematical sense. We then present them for the general case. Since the same ideas apply to other families of convolution functions, such as the Gamma densities previously mentioned, this is useful for possible further applications.

We end this section with a distribution property holding for  $X_\epsilon$  when the original noise is driven by a Brownian motion. If the CARMA process  $X$  is driven by a Brownian motion  $B$ , then  $X$  has Gaussian distribution. We show that the same holds for  $X_\epsilon$ . This is indeed not a straightforward result since the process  $B_\epsilon$  is no longer a Brownian motion by means of Proposition 2.2. A rigorous proof is thus necessary.

**Proposition 3.7.** *For every  $\epsilon > 0$ , the process  $\{X_\epsilon(t), t \geq 0\}$  has Gaussian distribution.*

*Proof:* See Appendix A.9 in **Supplementary Material**.





### 3.2. When $q > 0$

The process  $L_\varepsilon$  is continuously differentiable but is not twice continuously differentiable. However, at the right-hand side of the stochastic differential equation (3.1), we require  $q + 1$  derivatives of the driving noise process. For the general case  $q > 0$  we need the following result.

**Proposition 3.8.** *Under the causality condition (3.4) and if the matrix  $A$  has distinct eigenvalues, the CARMA( $p, q$ ) process  $X$  can be expressed by the sum*

$$X(t) = \sum_{r=1}^p X^{(r)}(t) \quad \text{with}$$

$$X^{(r)}(t) := \int_0^t \alpha_r e^{\lambda_r(t-u)} dL(u), \quad r = 1, \dots, p.$$

Here, the  $X^{(r)}$  are dependent and possibly complex-valued CAR(1) processes with  $\alpha_r := \frac{Q(\lambda_r)}{P'(\lambda_r)}$ .

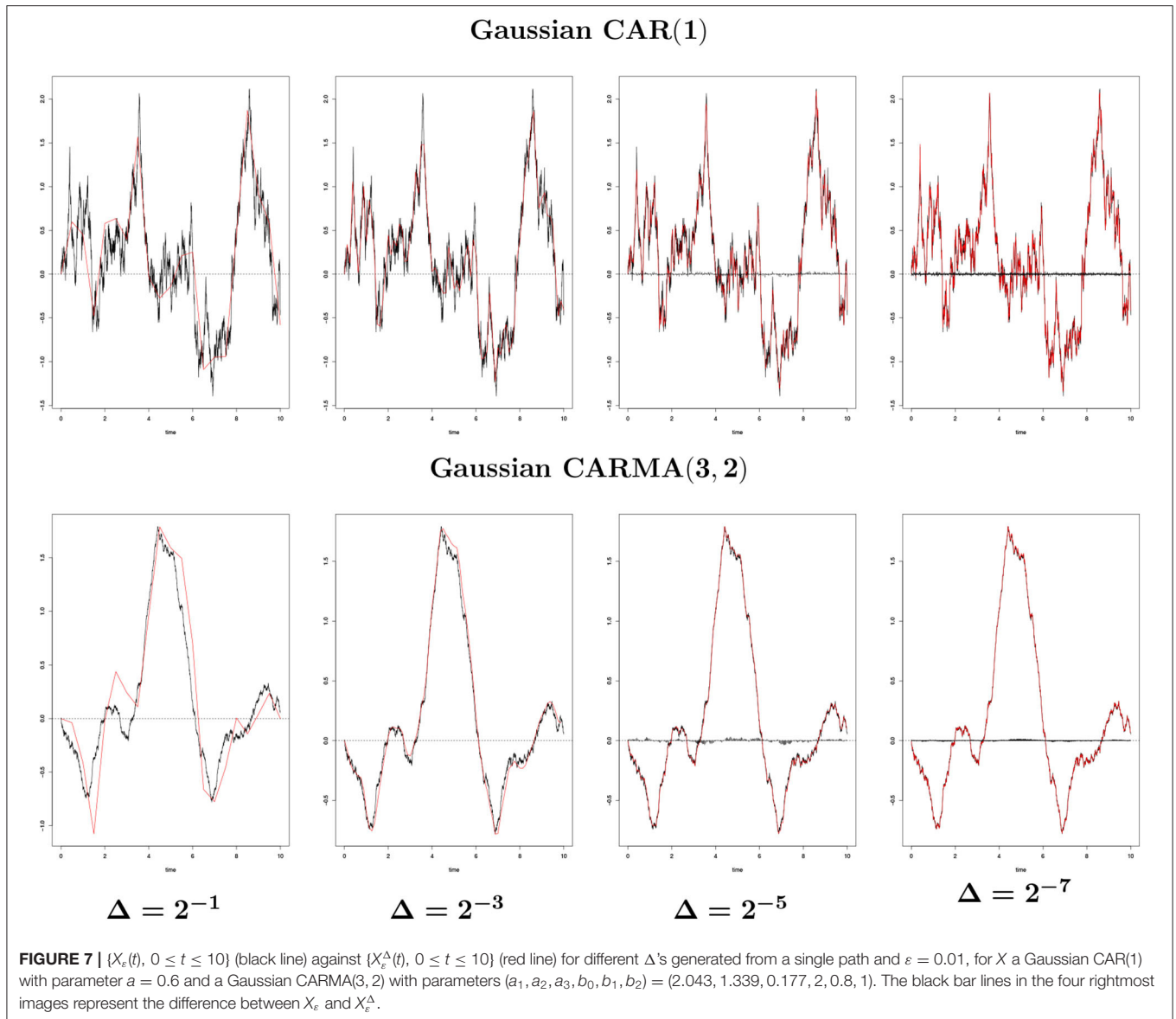
*Proof:* See [5, Proposition 2].

By means of Proposition 3.8, starting from  $L_\varepsilon$  in Equation (2.3), the idea is to construct  $p$ -approximating processes and obtain  $X_\varepsilon$  as the sum of them, namely,

$$X_\varepsilon(t) = \sum_{r=1}^p X_\varepsilon^{(r)}(t) \quad \text{with}$$

$$X_\varepsilon^{(r)}(t) = \int_0^t \alpha_r e^{\lambda_r(t-u)} dL_\varepsilon(u), \quad r = 1, \dots, p.$$

In particular, from Itô's calculus, we can prove that, for any  $r = 1, \dots, p$  and  $t \geq 0$ , the process  $X^{(r)}$ , respectively  $X_\varepsilon^{(r)}$ , satisfies the



stochastic differential equation

$$dX_{(\varepsilon)}^{(r)}(t) = \lambda_r X_{(\varepsilon)}^{(r)}(t)dt + \alpha_r dL_{(\varepsilon)}(t), \tag{3.12}$$

that will be used for numerical purposes in section 4. In Equation (3.12), we write  $\varepsilon$  in parentheses since it applies to both  $X^{(r)}$  and  $X_\varepsilon^{(r)}$ .

In **Figure 5** we consider a Gaussian CARMA(3,2) process with parameters  $(a_1, a_2, a_3) = (2.043, 1.339, 0.177)$  taken from [6] where a CAR(3) is used to model the daily average temperature dynamics in Stockholm (Sweden) and  $(b_0, b_1, b_2) = (2.0, 0.8, 1.0)$ . In **Figure 3D**, we observe the rate of convergence as a function of  $\varepsilon$ .

## 4. NUMERICAL SCHEMES AND RESULTS

We consider a possible approach to simulate  $X$  and  $X_\varepsilon$  for some CARMA parameters, and we then estimate these parameters starting from the simulated time series. In particular, we need first to simulate  $L$ , then  $L_\varepsilon$ , and only afterwards can we simulate  $X_\varepsilon$ . For  $n > 0$ , we shall consider the partition  $\Pi(n) := \{0 = s_0^{(n)} < s_1^{(n)} < \dots < s_n^{(n)} = T\}$  such that  $\Delta s_j^{(n)} := s_j^{(n)} - s_{j-1}^{(n)}$  is constant, that is  $\Delta s_j^{(n)} = T/n = \Delta(n) =: \Delta$ , and  $s_j^{(n)} = j\Delta$  for each  $j = 0, \dots, n$ .

### 4.1. The Lévy Simulations

By definition, for a fixed time interval  $\Delta$ , the increments of the Lévy process  $\Delta L(s_j^{(n)}) := L(s_j^{(n)}) - L(s_{j-1}^{(n)})$ ,  $j = 1, \dots, n$ , are independent and identically distributed random variables

following a certain distribution, which depends only on the time increment  $\Delta$ . For example, for  $L = B$ , the increments  $\Delta B(s_j^{(n)})$  have all Gaussian distribution with zero mean and variance  $\Delta$ . A possible approach to simulate  $L_\varepsilon$  in an exact way is to find the distribution of  $L_\varepsilon(t)$ , starting from the distribution of  $L(t)$ . For example, for  $L = B$ , by means of Lemma 2.3, we can simulate  $B_\varepsilon$  sampling directly from a Gaussian distribution with zero mean and variance in Lemma 2.3. However, with this approach,  $L$  and  $L_\varepsilon$  are from two different  $\omega$ 's in  $\Omega$ , and comparing the two paths we would not get the error measure required. We then need an alternative approach that simulates  $L_\varepsilon$  and  $L$  from the same  $\omega \in \Omega$ .

From Equation (2.3),  $L_\varepsilon$  is the stochastic integral of  $\phi_\varepsilon$  with respect to the Lévy process  $L$ . From the definition of the stochastic integral as a limit of approximating Riemann-Stieltjes sums [20], for  $t \in \Pi(n)$  such that  $t = s_m^{(n)} = m\Delta$ , for  $m = m(t)$  and  $1 \leq m \leq n$ , we approximate  $L_\varepsilon(t)$  by approximating the integral with a sum, namely

$$L_\varepsilon(t) = \int_0^t \phi_\varepsilon(t-s)dL(s) \approx \sum_{j=1}^m \phi_\varepsilon(t-s_{j-1}^{(n)})\Delta L(s_j^{(n)}) = : L_\varepsilon^\Delta(t), \quad (4.1)$$

where  $t - s_{j-1}^{(n)} = (m + 1 - j)\Delta = s_{m+1-j}^{(n)}$ . The approximation scheme (4.1) adds an error to the simulation. It, however, constructs  $L_\varepsilon$  starting from the simulations of  $L$ , allowing to compare the two paths for different  $\Delta$ 's. In the next proposition, we study the approximation error.

**Proposition 4.1.** *For every  $\Delta > 0$ ,  $\varepsilon > 0$  and  $T < +\infty$ , there exists a constant  $K_3 > 0$  such that the simulation error is bounded by*

$$\mathbb{E} \left[ (L_\varepsilon^\Delta(T) - L_\varepsilon(T))^2 \right] \leq K_3 \frac{\Delta^2}{\varepsilon}.$$

*Proof:* See Appendix A.10 in **Supplementary Material**.

With Proposition 4.1, we estimated an upper bound for the error added by the Euler scheme for simulating the process  $L_\varepsilon$  by means of the process  $L_\varepsilon^\Delta$ . In particular, it shows that, for a fixed  $\varepsilon > 0$ , the simulation error decreases at the decreasing of the time step  $\Delta$ , as we would expect. In Kloeden et al. [21, section 9.3], the authors study the rate of convergence for the Euler scheme when applied to an Itô process, with respect to the absolute error criterion, finding a strong rate of convergence equal to 1/2. Here, in line with the rest of the paper, we consider the squared error in Proposition 4.1 instead. This gives a rate of convergence equal to 2, corresponding to a strong order 1. From Proposition 4.1, we also notice that the bound depends on the inverse of  $\varepsilon$ , meaning that, for  $\varepsilon$  going to zero, the simulation error might explode.

**Figure 5** shows the convergence of the Euler scheme (4.1) for a Brownian motion and for the normal inverse Gaussian process introduced in section 2.1. Since an exact simulation is not possible, to generate  $B_\varepsilon$ , we consider the scheme (4.1) with a small time step  $\Delta^{true} = 2^{-10}$ . Moreover, since we want to

compare  $B_\varepsilon$  and  $B_\varepsilon^\Delta$  path-wise for different values of  $\Delta$ ,  $B_\varepsilon$  and  $B_\varepsilon^\Delta$  must be related to the same  $\omega \in \Omega$ , that is, to the same realization of the Brownian motion  $B$ . To get this, we exploit the following well-known property: given  $X_i \sim \mathcal{N}(0, \sigma_i^2)$ , for  $i = 1, \dots, k$ , independent Gaussian random variables, their sum follows the distribution  $\sum_{i=1}^k X_i \sim \mathcal{N}(0, \sum_{i=1}^k \sigma_i^2)$ . Thus, we start by simulating the *true* process  $B_\varepsilon$  for a certain time step  $\Delta^{true} \ll 1$ . For this, we sample the increments  $\Delta B(s_j^{(n)})$  from a Gaussian distribution with zero mean and variance  $\Delta^{true}$  and obtain  $B_\varepsilon$  as in Equation (4.1). Then, we choose a time step  $\Delta$  multiple of  $\Delta^{true}$ , namely,  $\Delta = k\Delta^{true}$  for a certain  $k \in \mathbb{N}$ . The Brownian motion path to be used in Equation (4.1) for simulating  $B_\varepsilon^\Delta$  is then obtained by summing up the increments of  $B$  (above simulated) at groups of  $k$  elements. These increments have indeed Gaussian distribution with zero mean and variance  $\sum_{i=1}^k \Delta^{true} = k\Delta^{true} = \Delta$ , as required. In **Figure 6**, we considered  $\Delta^{true} = 2^{-10}$  and  $\Delta \in \{2^{-7}, 2^{-5}, 2^{-3}, 2^{-1}\}$ . We see that, for a fixed  $\varepsilon$ , as  $\Delta$  decreases, the error between  $B_\varepsilon^\Delta$  and  $B_\varepsilon$  decreases. Even with a large  $\Delta$  ( $\Delta = 2^{-1}$ ), the process  $B_\varepsilon^\Delta$  captures the shape of  $B_\varepsilon$ . In **Figure 3E** we show the rate of convergence of the Euler scheme proven in Proposition 4.1 as function of time step  $\Delta$  and for a fixed  $\varepsilon$  ( $\varepsilon = 0.01$ ).

Similarly, for the NIG Lévy process, to simulate and compare  $L_\varepsilon$  with  $L_\varepsilon^\Delta$  path-wise, we use the same idea as in the Brownian motion case. We can indeed exploit the convolution property of the NIG distribution (see [22]): given  $X_i \sim \text{NIG}(\alpha, \beta, \delta_i, \mu_i)$ , for  $i = 1, \dots, k$ , independent normal inverse Gaussian random variables with common parameters  $\alpha$  and  $\beta$ , but individual scale-location parameters  $\delta_i$  and  $\mu_i$ , their sum follows  $\sum_{i=1}^k X_i \sim \text{NIG}(\alpha, \beta, \sum_{i=1}^k \delta_i, \sum_{i=1}^k \mu_i)$ . In section 2.1, we introduced NIG distribution with parameters  $(1, 0, \Delta^{true}, 0)$ , for  $\Delta^{true}$  being the time step used in the Euler discretization. As above, we thus simulate the *true* process  $L_\varepsilon$  from Equation (4.1) by sampling the increments of  $L$  from a NIG distribution with parameters  $(1, 0, \Delta^{true}, 0)$  for a certain  $\Delta^{true} \ll 1$ . Then, we obtain  $L_\varepsilon^\Delta$  by choosing  $\Delta$  so that  $\Delta = k\Delta^{true}$  for a certain  $k \in \mathbb{N}$ , and summing up the increments of  $L$  (simulated above) at groups of  $k$  elements. These increments have indeed NIG distribution with parameters  $(1, 0, \sum_{i=1}^k \Delta^{true}, \sum_{i=1}^k 0) = (1, 0, \Delta, 0)$  as required. We consider the same  $\Delta$ 's as for the Brownian motion. We notice that, for a fixed  $\varepsilon$ ,  $L_\varepsilon^\Delta$  captures the shape of  $L_\varepsilon$  even with a large  $\Delta$ . In **Figure 3F** we show the rate of convergence for the Euler scheme proven in Proposition 4.1 as function of the time step  $\Delta$  and for a fixed  $\varepsilon$  ( $\varepsilon = 0.01$ ).

From Proposition 2.7 and 4.1, we calculate the squared error for  $L_\varepsilon^\Delta$  with respect to  $L$ .

**Proposition 4.2.** *For every  $\Delta > 0$  and  $\varepsilon > 0$  small enough, we get the following bound*

$$\mathbb{E} \left[ (L_\varepsilon^\Delta(T) - L(T))^2 \right] \leq 2K_1\varepsilon + 2K_3 \frac{\Delta^2}{\varepsilon}.$$

*Proof:* This is a direct consequence of Proposition 2.7 and 4.1 by the inequality  $(a + b)^2 \leq 2(a^2 + b^2)$  for  $a = L_\varepsilon^\Delta(T) - L_\varepsilon(T)$  and  $b = L_\varepsilon(T) - L(T)$ .  $\square$

Proposition 4.2 shows that the squared error between  $L_\varepsilon^\Delta$  and  $L$  decreases when  $\Delta$  decreases. However, because of the term proportional to the inverse of  $\varepsilon$ , it might also explode for small values of  $\varepsilon$ . We then need to tune the parameters  $\Delta$  and  $\varepsilon$  to find a combination that minimizes the error preventing it from taking large values.

### 4.2. The CARMA Simulations

It is possible to construct  $X$  and  $X_\varepsilon$  starting from Equations (3.5) and (3.8) and considering the same kind of numerical approximation as done for  $L_\varepsilon$ , that is, by replacing the integral with a suitable sum. However, to not add a second simulation error to the one already due to  $L_\varepsilon$ , which is approximated by  $L_\varepsilon^\Delta$ , we start from the state-space representations instead, Equations (3.3) and (3.7), respectively. By using the Euler scheme, according to [21, Chapter 9], we get (here we write  $\varepsilon$  in parenthesis since the equation holds both for  $X$  and  $X_\varepsilon$ ):

$$\begin{aligned} \mathbf{Y}_{(\varepsilon)}^\Delta(s_j^{(n)}) &= \mathbf{Y}_{(\varepsilon)}^\Delta(s_{j-1}^{(n)}) + A\mathbf{Y}_{(\varepsilon)}^\Delta(s_{j-1}^{(n)})\Delta + \mathbf{e}_p\Delta L_{(\varepsilon)}(s_j^{(n)}), \\ X_{(\varepsilon)}^\Delta(s_j^{(n)}) &= \mathbf{b}^\top \mathbf{Y}_{(\varepsilon)}^\Delta(s_j^{(n)}), \end{aligned} \tag{4.2}$$

where  $\Delta L_{(\varepsilon)}(s_j^{(n)}) = L_{(\varepsilon)}(s_j^{(n)}) - L_{(\varepsilon)}(s_{j-1}^{(n)})$ , for  $j = 1, \dots, n$ . For the reasons already discussed, we apply the scheme (4.2) only for  $q = 0$ , that is, when  $\mathbf{b} = \mathbf{e}_1$ . For  $q > 0$ , by means of Proposition 3.8, we instead apply the Euler scheme to the stochastic differential equation (3.12) for each of the  $p$  components  $X_{(\varepsilon)}^{(r)}$  of  $X_{(\varepsilon)}$ , and we get  $X_{(\varepsilon)}$  as the sum of them, namely,

$$\begin{aligned} X_{(\varepsilon)}^\Delta(s_j^{(n)}) &= X_{(\varepsilon)}^\Delta(s_{j-1}^{(n)}) + \lambda_r X_{(\varepsilon)}^\Delta(s_{j-1}^{(n)})\Delta + \alpha_r \Delta L_{(\varepsilon)}(s_j^{(n)}), \\ X_{(\varepsilon)}^\Delta(s_j^{(n)}) &= \sum_{r=1}^p X_{(\varepsilon)}^{\Delta(r)}(s_j^{(n)}). \end{aligned}$$

When simulating  $X_\varepsilon^\Delta$ , we use the approximation  $L_\varepsilon^\Delta$  in Equation (4.1), and this adds an error. We study the rate of convergence for  $X_\varepsilon^\Delta$  to the true process  $X_\varepsilon$  in the following proposition.

Proposition 4.3. For every  $\varepsilon > 0$  and  $\Delta < \max_{j=1, \dots, p} \frac{-2\text{Re}(\lambda_j)}{|\lambda_j|^2}$ , there exist two constants  $K_4, K_5 > 0$  such that the simulation error for the Euler scheme (4.2) is bounded by

$$\mathbb{E} \left[ (X_\varepsilon^\Delta(T) - X_\varepsilon(T))^2 \right] \leq K_4 \frac{\Delta^2}{\varepsilon} + K_5 \frac{\Delta^4}{\varepsilon^4}.$$

Proof: See Appendix A.11 in Supplementary Material.

Let us notice that the bound on  $\Delta$  in Proposition 4.3 has been introduced to obtain the estimate for the squared error. The scheme (4.2) works however for every  $\Delta > 0$ .

In Figure 7 we show the convergence for the simulations of a Gaussian CAR(1) and a Gaussian CARMA(3,2). In both cases we notice that for  $\Delta$  decreasing, the Euler approximation  $X_\varepsilon^\Delta$  converges to  $X_\varepsilon$ , and, even with a large value of  $\Delta$  ( $\Delta = 2^{-1}$ ),  $X_\varepsilon^\Delta$  captures the shape of  $X_\varepsilon$ .

TABLE 1 | Parameters estimated and standard deviation (in parenthesis) for different  $\varepsilon$ 's and  $\Delta$ 's after 1,000 simulations for a Gaussian CAR(1) with  $a = 0.6$ .

Gaussian CAR(1) parameters estimated		
$\Delta = 2^0$		
$\varepsilon$	$\hat{a}$	$\hat{a}_\varepsilon$
$2^0$	0.6010288 (0.02889827)	0.3530898 (0.01973318)
$*2^{-1}$	0.6017733 (0.02791626)	0.5633918 (0.02691802)
$*2^{-2}$	0.6005038 (0.02813515)	0.6004506 (0.02813376)
$*2^{-3}$	0.6008589 (0.0300782)	0.6008589 (0.0300782)
$*2^{-4}$	0.6013792 (0.02890727)	0.6013792 (0.02890727)
$\Delta = 2^{-1}$		
$\varepsilon$	$\hat{a}$	$\hat{a}_\varepsilon$
$2^0$	0.6041548 (0.03362481)	0.1694637 (0.01097814)
$2^{-1}$	0.6011999 (0.0325313)	0.3378004 (0.01978348)
$*2^{-2}$	0.6036319 (0.03193889)	0.558025 (0.02990832)
$*2^{-3}$	0.6002732 (0.03133392)	0.6002087 (0.03133112)
$*2^{-4}$	0.6025603 (0.03244405)	0.6025603 (0.03244405)
$\Delta = 2^{-2}$		
$\varepsilon$	$\hat{a}$	$\hat{a}_\varepsilon$
$2^0$	0.6037737 (0.03312141)	0.0809573 (0.005357486)
$2^{-1}$	0.6016175 (0.03297326)	0.1679081 (0.01021751)
$2^{-2}$	0.6026758 (0.03342669)	0.3321287 (0.01916501)
$*2^{-3}$	0.6032027 (0.03258187)	0.5541501 (0.03015163)
$*2^{-4}$	0.6041056 (0.03402852)	0.6040349 (0.03402487)

The underlined digits are the common digits in  $\hat{a}$  and  $\hat{a}_\varepsilon$ . The symbol \* marks when condition  $\varepsilon < \Delta$  is satisfied.

In Figures 3G,H, we show the rate of convergence for the two processes as function of  $\Delta^2$  and for a fixed value of  $\varepsilon$  ( $\varepsilon = 0.01$ ).

From Proposition 3.6 and 4.3, we estimate the squared error for  $X_\varepsilon^\Delta$  with respect to  $X$ .

Proposition 4.4. For every  $\Delta > 0$  and  $\varepsilon > 0$  small enough, the following bound applies:

$$\mathbb{E} \left[ (X_\varepsilon^\Delta(T) - X(T))^2 \right] \leq 2K_2 \varepsilon^2 + 2K_4 \frac{\Delta^2}{\varepsilon} + 2K_5 \frac{\Delta^4}{\varepsilon^4}.$$

Proof: This is a direct consequence of Proposition 3.6 and 4.3 by the inequality  $(a + b)^2 \leq 2(a^2 + b^2)$  for  $a = X_\varepsilon^\Delta(T) - X_\varepsilon(T)$  and  $b = X_\varepsilon(T) - X(T)$ .  $\square$

### 4.3. CARMA Estimation

We introduce an approach to estimate the parameters  $(a_1, \dots, a_p, b_0, \dots, b_q)$  starting from the time series for  $\{X(t), 0 \leq t \leq T\}$  and  $\{X_\varepsilon(t), 0 \leq t \leq T\}$  by applying the Euler scheme. Following Benth et al. [7, section 4.3], we focus on CAR models only, that is, on the process  $X(t) = Y_1(t)$ , for  $Y_1(t)$  being the first coordinate of the vector  $\mathbf{Y}(t)$  in Equation

**TABLE 2** | Parameters estimated and standard deviation (every second row) for different  $\varepsilon$ 's and  $\Delta$ 's after 1,000 simulations for a Gaussian CAR(3) with  $a = (2.043, 1.339, 0.177)$ .

Gaussian CAR(3) parameters estimated		
$\Delta = 2^0$		
$\varepsilon$	$\hat{a}$	$\hat{a}_\varepsilon$
$2^0$	( <u>2.0465438</u> , <u>1.3448458</u> , <u>0.1795313</u> )	(1.6656158, <u>1.0061226</u> , <u>0.1163607</u> )
	(0.03182854, 0.04668409, 0.02019277)	(0.03093857, 0.03861325, 0.01357942)
* $2^{-1}$	( <u>2.0448233</u> , <u>1.3424365</u> , <u>0.1793535</u> )	(1.9980868, <u>1.2946571</u> , <u>0.1702482</u> )
	(0.03255849, 0.04568522, 0.01961278)	(0.03248752, 0.04472464, 0.01870831)
* $2^{-2}$	( <u>2.0444178</u> , <u>1.3425449</u> , <u>0.1799632</u> )	( <u>2.0443555</u> , <u>1.3424787</u> , <u>0.1799505</u> )
	(0.03155042, 0.04637146, 0.01957948)	(0.03155036, 0.04637017, 0.01957821)
* $2^{-3}$	( <u>2.0441555</u> , <u>1.3413174</u> , <u>0.1784795</u> )	( <u>2.0441555</u> , <u>1.3413174</u> , <u>0.1784795</u> )
	(0.03174612, 0.04494131, 0.02049703)	(0.03174612, 0.04494131, 0.02049703)
* $2^{-4}$	( <u>2.0455070</u> , <u>1.3438790</u> , <u>0.1805164</u> )	( <u>2.0455070</u> , <u>1.3438790</u> , <u>0.1805164</u> )
	(0.03113649, 0.04640021, 0.02003383)	(0.03113649, 0.04640021, 0.02003383)
$\Delta = 2^{-1}$		
$\varepsilon$	$\hat{a}$	$\hat{a}_\varepsilon$
$2^0$	(2.047198, 1.344582, 0.1179701)	(0.93101668, 0.74900497, 0.054728)
	(0.04400026, 0.05788283, 0.02091601)	(0.024569478, 0.036776509, 0.006575435)
$2^{-1}$	(2.0468192, <u>1.3450024</u> , <u>0.1795957</u> )	(1.4472312, <u>1.0800373</u> , <u>0.1083792</u> )
	(0.04272407, 0.05713749, 0.02003501)	(0.03343550, 0.04763322, 0.01227647)
* $2^{-2}$	(2.0451238, <u>1.3438020</u> , <u>0.1798912</u> )	(1.9600235, <u>1.3032478</u> , <u>0.1692412</u> )
	(0.04093678, 0.05699320, 0.02153571)	(0.04002300, 0.05576660, 0.02030978)
* $2^{-3}$	( <u>2.0466145</u> , <u>1.3420588</u> , <u>0.1787042</u> )	( <u>2.0464984</u> , <u>1.3420004</u> , <u>0.1786897</u> )
	(0.04360794, 0.05849717, 0.02093359)	(0.04360660, 0.05849502, 0.02093195)
* $2^{-4}$	( <u>2.0475337</u> , <u>1.3460992</u> , <u>0.1809728</u> )	( <u>2.0475337</u> , <u>1.3460992</u> , <u>0.1809728</u> )
	(0.04374334, 0.05753167, 0.02091458)	(0.04374334, 0.05753167, 0.02091458)
$\Delta = 2^{-2}$		
$\varepsilon$	$\hat{a}$	$\hat{a}_\varepsilon$
$2^0$	(2.0488046, 1.3472925, 0.1797796)	(0.44880714, 0.64200476, 0.02334978)
	(0.05145854, 0.06480728, 0.02143492)	(0.013049683, 0.034215080, 0.002944983)
$2^{-1}$	(2.0464197, 1.3458156, 0.1800446)	(0.76315132, 0.93934128, 0.05071483)
	(0.04978053, 0.06403133, 0.02119095)	(0.020194320, 0.046465195, 0.006227828)
$2^{-2}$	(2.0496128, <u>1.3512352</u> , <u>0.1802224</u> )	(1.2995395, <u>1.1838183</u> , <u>0.1029504</u> )
	(0.05054903, 0.06626747, 0.02097240)	(0.03354603, 0.05883910, 0.01220843)
* $2^{-3}$	(2.0499813, <u>1.3469554</u> , <u>0.1801088</u> )	(1.9279250, <u>1.3210808</u> , <u>0.1673256</u> )
	(0.05507782, 0.06371013, 0.02150337)	(0.05230973, 0.06260054, 0.02005562)
* $2^{-4}$	( <u>2.0486265</u> , <u>1.3471175</u> , <u>0.1793858</u> )	( <u>2.0484564</u> , <u>1.3470775</u> , <u>0.1793683</u> )
	(0.05069210, 0.06491714, 0.02095847)	(0.05068899, 0.06491685, 0.02095691)

The common digits in  $\hat{a}$  and  $\hat{a}_\varepsilon$  are underlined. The symbol \* marks when condition  $\varepsilon < \Delta$  is satisfied.

(3.3). Then, for a uniform time step  $\Delta$ , we get the relation

$$\begin{aligned}
 & \frac{1}{\Delta^{p-1}} \sum_{j=0}^p (-1)^j z_j^p Y_1(t + (p-j)\Delta) = \\
 & = - \sum_{i=1}^p a_{p-i+1} \frac{1}{\Delta^{i-2}} \sum_{j=0}^{i-1} (-1)^j z_j^{i-1} Y_1(t + (i-1-j)\Delta) + \varepsilon(t),
 \end{aligned}
 \tag{4.3}$$

$\varepsilon(t)$  being the Lévy noise. The coefficients  $z_j^i$  are defined as  $z_0^i = z_i^i = 1$  for  $i = 1, \dots, p$ , and, through the recursion  $z_j^i = z_{j-1}^{i-1} + z_j^{i-1}$ , for  $j = 1, \dots, p-1$  and  $i \geq 2$ . We refer to Benth et al. [7, section 4.3] for more details. In particular, Equation (4.3) is a linear combination of  $Y_1(t + p\Delta), \dots, Y_1(t)$  and the error term  $\varepsilon(t)$ . Thus, we have an AR( $p$ ) representation model for  $Y_1(t)$ , and the parameters  $a_1, \dots, a_p$  can be estimated by means of relation (4.3) starting from the AR estimated parameters. We consider as examples  $p = 1$  and  $p = 3$ . For  $p = 1$ ,  $X(t) = Y_1(t)$  is

an Ornstein-Uhlenbeck process and Equation (4.3) becomes

$$Y_1(t + \Delta) = \{1 - a_1 \Delta\} Y_1(t) + \epsilon(t),$$

which coincides with the Euler scheme (4.2) when  $p = 1$ . For  $p = 3$ , from Equation (4.3) we get

$$\begin{aligned} & \frac{1}{\Delta^2} \{z_0^3 Y_1(t + 3\Delta) - z_1^3 Y_1(t + 2\Delta) + \\ & \quad + z_2^3 Y_1(t + \Delta) - z_3^3 Y_1(t)\} \\ & = -a_3 \Delta z_0^0 Y_1(t) - a_2 \{z_1^1 Y_1(t + \Delta) - z_1^1 Y_1(t)\} + \\ & \quad - \frac{a_1}{\Delta} \{z_0^2 Y_1(t + 2\Delta) - z_1^2 Y_1(t + \Delta) + z_2^2 Y_1(t)\} + \epsilon(t). \end{aligned}$$

From the definition for  $z_j^i$ , rearranging the terms, the last equation can be rewritten as

$$\begin{aligned} Y_1(t + 3\Delta) &= \theta_1 Y_1(t + 2\Delta) + \theta_2 Y_1(t + \Delta) + \theta_3 Y_1(t) + \Delta^2 \epsilon(t), \\ \text{with } \theta_1 &:= 3 - a_1 \Delta, \quad \theta_2 := 2a_1 \Delta - a_2 \Delta^2 - 3, \\ \theta_3 &:= 1 - a_1 \Delta + a_2 \Delta^2 - a_3 \Delta^3. \end{aligned}$$

Then, one first estimates the AR coefficients  $\theta_1, \theta_2, \theta_3$  and afterwards recover the CAR coefficients  $a_1, a_2, a_3$  by

$$a_1 = \frac{3 - \theta_1}{\Delta}, \quad a_2 = \frac{3 - 2\theta_1 - \theta_2}{\Delta^2}, \quad a_3 = \frac{1 - \theta_1 - \theta_2 - \theta_3}{\Delta^3}.$$

This scheme holds for every Lévy process. To recover the case  $q > 0$  different approaches have to be considered, see, e.g., [23] and [24].

In **Tables 1, 2** we report the parameters estimated for a Gaussian CAR(1) and a Gaussian CAR(3) for different  $\epsilon$ 's and different  $\Delta$ 's. In particular,  $\hat{a}$  is the vector of parameters estimated starting from the time series for  $X$ , and  $\hat{a}_\epsilon$  is the vector of parameters estimated starting from the time series for  $X_\epsilon$ . The underlined digits are the common digits between  $\hat{a}$  and  $\hat{a}_\epsilon$ , with respect to the same  $\epsilon$  and the same  $\Delta$ . The estimates get more

accurate for a small  $\epsilon$ , but they are not accurate if  $\epsilon$  is bigger than  $\Delta$ . We marked the cases corresponding to  $\epsilon < \Delta$  with the symbol \*, and we notice that as soon as  $\epsilon$  gets smaller than  $\Delta$ , the accuracy improves and gets better for smaller values of  $\epsilon$ . We summarize this empirical result with a proposition.

**Proposition 4.5.** *The Euler scheme (4.3) converges to the true CARMA parameters if  $\epsilon < \Delta$ .*

*Proof:* See Appendix A.12 in **Supplementary Material**.

A more rigorous study should be performed regarding the condition in Proposition 4.5, analysing in details the estimation technique for ARMA processes, which is the starting point to estimate the CARMA parameters. This is, however, beyond the purposes of the current article.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## ACKNOWLEDGMENTS

The author acknowledges Fred Espen Benth for the several fruitful discussions that let to this paper and two referees for additional interesting comments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fams.2020.00037/full#supplementary-material>

## REFERENCES

- Brockwell PJ. Lévy-driven CARMA processes. *Ann Inst Stat Math.* (2001) 53:113–24. doi: 10.1023/A:1017972605872
- Barndorff-Nielsen OE, Shepard N. Non-Gaussian OU based models and some of their uses in financial economics. *J R Stat Soc B Stat Methodol.* (2001) 63:167–241. doi: 10.1111/1467-9868.00282
- Brockwell PJ, Marquardt T. Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Stat Sin.* (2005) 15:477–94. Available online at: <https://www.jstor.org/stable/24307365>
- Brockwell PJ, Davis RA, Yang Y. Estimation for nonnegative Lévy-driven Ornstein-Uhlenbeck processes. *J Appl Probab.* (2007) 44:977–89. doi: 10.1239/jap/1197908818
- Brockwell PJ, Davis RA, Yang Y. Estimation for non-negative Lévy-driven CARMA processes. *J Bus Econ Stat.* (2011) 29:250–9. doi: 10.1198/jbes.2010.08165
- Benth FE, Benth JS, Koekebakker S. *Stochastic Modelling of Electricity and Related Markets*. Vol. 11. Singapore: World Scientific (2008).
- Benth FE, Benth JS. *Modelling and Pricing in Financial Markets for Weather Derivatives*. Vol. 17. Singapore: World Scientific (2012).
- Marquardt T. Fractional Lévy processes with an application to long memory moving average processes. *Bernoulli.* (2006) 12:1099–126. doi: 10.3150/bj/1165269152
- Marquardt T, James LF. *Generating Long Memory Models Based on CARMA Processes*. Technical report (2007).
- Paschke R, Prokopczuk M. Integrating multiple commodities in a model of stochastic price dynamics. *J Energy Markets.* (2009) 2:47–82. doi: 10.21314/JEM.2009.025
- Paschke R, Prokopczuk M. Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. *J Bank Finance.* (2010) 34:2742–52. doi: 10.1016/j.jbankfin.2010.05.010
- Schwartz E, Smith JE. Short-term variations and long-term dynamics in commodity prices. *Manage Sci.* (2000) 46:893–911. doi: 10.1287/mnsc.46.7.893.12034
- Benth FE, Koekebakker S. Pricing of forwards and other derivatives in cointegrated commodity markets. *Energy Econ.* (2015) 52:104–17. doi: 10.1016/j.eneco.2015.09.009
- Benth FE, Klüppelberg C, Müller G, Vos L. Futures pricing in electricity markets based on stable CARMA spot models.

- Energy Econ.* (2014) 44:392–406. doi: 10.1016/j.eneco.2014.03.020
15. García I, Klüppelberg C, Müller G. Estimation of stable CARMA models with an application to electricity spot prices. *Stat Model.* (2011) 11:447–70. doi: 10.1177/1471082X1001100504
  16. Andresen A, Benth FE, Koekebakker S, Zakamulin V. The CARMA interest rate model. *Int J Theor Appl Finance.* (2014) 17:1450008. doi: 10.1142/S0219024914500083
  17. Adland R, Benth FE, Koekebakker S. Multivariate modeling and analysis of regional ocean freight rates. *Transport Res E.* (2018) 113:194–221. doi: 10.1016/j.tre.2017.10.014
  18. Cont R, Tankov P. *Financial Modelling With Jump Processes.* Chapman and Hall/CRC (2004).
  19. Benth FE, Di Persio L, Lavagnini S. Stochastic modelling of wind derivatives in energy markets. *Risks.* (2018) 6:56. doi: 10.3390/risks6020056
  20. Øksendal BK. *Stochastic Differential Equations: An Introduction With Applications.* 5th ed. Heidelberg; New York, NY: Springer-Verlag (2003).
  21. Kloeden PE, Platen E. *Numerical Solution of Stochastic Differential Equation.* Berlin; Heidelberg: Springer-Verlag (1992).
  22. Hanssen A, Øigård TA. The normal inverse Gaussian distribution: a versatile model for heavy-tailed stochastic processes. In: *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings.* Vol. 6. Salt Lake City, UT (2001). doi: 10.1109/ICASSP.2001.940717
  23. Brockwell PJ, Davis RA. *Introduction to Time Series and Forecasting.* 2nd ed. New York, NY: Springer (2002). doi: 10.1007/b97391
  24. Stelzer R. CARMA processes driven by non-Gaussian noise. *arXiv preprint arXiv:1201.0155* (2011).

**Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Lavagnini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.