Ann Inst Stat Math (2014) 66:647–685
DOI 10.1007/s10463-014-0468-7

INVITED REVIEW ARTICLE

Recent results in the theory and applications of CARMA processes

P. J. Brockwell

Received: 9 October 2013 / Revised: 27 January 2014 / Published online: 1 June 2014

Abstract Just as ARMA processes play a central role in the representation of stationary time series with discrete time parameter, $(Y_n)_{n\in\mathbb{Z}}$, CARMA processes play an analogous role in the representation of stationary time series with continuous time parameter, $(Y(t))_{t\in\mathbb{R}}$. Lévy-driven CARMA processes permit the modelling of heavy-tailed and asymmetric time series and incorporate both distributional and sample-path information. In this article we provide a review of the basic theory and applications, emphasizing developments which have occurred since the earlier review in Brockwell (2001a, In D. N. Shanbhag and C. R. Rao (Eds.), *Handbook of Statistics 19*; *Stochastic Processes*: *Theory and Methods* (pp. 249–276), Amsterdam: Elsevier).

Keywords Time series · Stationary process · CARMA process · Sampled process · High-frequency sampling · Inference · Prediction

1 Introduction

Continuous-time autoregressive (CAR) processes have been of interest to physicists and engineers for many years [see e.g. Fowler (1936)]. Early papers dealing with the properties and statistical analysis of such processes, and of the more general continuous-time autoregressive moving average (CARMA) processes, include those

Support from the National Science Foundation Grant DMS-1107031 is gratefully acknowledged.

P. J. Brockwell (✉)
Statistics Department, Colorado State University, Fort Collins, CO 80523-1877, USA
e-mail: peter.brockwell@gmail.com

P. J. Brockwell
Statistics Department, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA

of Doob (1944), Bartlett (1946), Phillips (1959) and Durbin (1961). In the last 10 years there has been a resurgence of interest in these processes and in continuous-time processes more generally, partly as a result of the very successful application of stochastic differential equation models to problems in finance, particularly to the pricing of options. The proliferation of high-frequency data, especially in fields such as finance and turbulence, has stimulated interest also in the connections between continuous-time processes and the discrete-time processes obtained by sampling them at high frequencies and the possible use of continuous-time models to suggest inferential methods for such data. Numerous examples of econometric applications of continuous-time models are contained in the book of Bergstrom (1990). Continuous-time models have also been utilized very successfully for the modelling of irregularly-spaced data Jones (1981, 1985).

In this section we shall first outline some basic results from the theory of *weakly stationary* processes with index set $\mathbb{R}$. These depend heavily on integrals of the form $\int_{\mathbb{R}} g(t-u)\mathrm{d}\xi(u)$ with ξ an *orthogonal increment process*. In subsequent sections we shall make use of integrals of the form $\int_{\mathbb{R}} g(t-u)\mathrm{d}L(u)$ where L is a *Lévy process*, which has *independent* increments. Lévy processes are defined and discussed at the end of this section. The use of Lévy processes rather than orthogonal increment processes permits the study of strictly stationary processes which do not necessarily have finite second moments and also allows investigation of their joint distributions and sample-path properties. We begin with some definitions.

A complex-valued time series $(Y(t))_{t\in\mathbb{R}}$ is said to be *weakly stationary* if $E|Y(t)|^2 < \infty$ for all $t \in \mathbb{R}$, $E(Y(t))$ is independent of t, and $E\left[Y(t+h)\overline{Y(t)}\right]$ is independent of t for all $h \in \mathbb{R}$. ($\overline{Y(t)}$ denotes the complex conjugate of $Y(t)$). It is *strictly stationary* if the joint distribution of $(Y(t_1), Y(t_2), \ldots, Y(t_n))$ is the same as that of $(Y(t_1+h), Y(t_2+h), \ldots, Y(t_n+h))$ for all $n \in \mathbb{N}$, $t_i \in \mathbb{R}$, $i = 1, \ldots, n$, and $h \in \mathbb{R}$. We shall use the notation $\tilde{f}$ to denote the Fourier transform, $\tilde{f}(\omega) := \int_{\mathbb{R}} \mathrm{e}^{-i\omega t} f(t)\mathrm{d}t$, $\omega \in \mathbb{R}$, of a square integrable function f on $\mathbb{R}$. Then $f(t) = (2\pi)^{-1}\int_{\mathbb{R}} \mathrm{e}^{i\omega t}\tilde{f}(\omega)\mathrm{d}\omega$.

We shall consider processes defined on a probability space $(\Omega, \mathcal{F}, P)$. The set of complex-valued square-integrable random variables will be denoted by $L^2(\Omega, \mathcal{F}, P)$, or more concisely by $L^2(P)$, a Hilbert space if we define the inner product of X and Y as $\langle X, Y\rangle := E(X\overline{Y})$. The squared norm corresponding to this inner product is $||X||^2 := \langle X, X\rangle = E|X|^2$ and the random variables X and Y are said to be orthogonal if $\langle X, Y\rangle = 0$.

An orthogonal increment process (OIP) on $\mathbb{R}$ is a complex-valued stochastic process $(\xi(u))_{u\in\mathbb{R}}$ such that

$$\langle \xi(u), \xi(u)\rangle < \infty, \quad u \in \mathbb{R}, \tag{1}$$

$$\langle \xi(u), 1\rangle = 0, \quad u \in \mathbb{R}, \tag{2}$$

$$\langle \xi(u_4) - \xi(u_3), \xi(u_2) - \xi(u_1)\rangle = 0, \quad \text{if } (u_1, u_2] \cap (u_3, u_4] = \emptyset. \tag{3}$$

and

$$||\xi(u+\delta) - \xi(u)||^2 = E|\xi(u+\delta) - \xi(u)|^2 \to 0 \text{ as } \delta \downarrow 0. \tag{4}$$

For each OIP $(\xi(u))_{u\in\mathbb{R}}$ there is a unique non-decreasing right-continuous function F on $\mathbb{R}$, the associated distribution function (ADF), such that

$$F(0) = 0,$$

and

$$F(v) - F(u) = ||\xi(v) - \xi(u)||^2, \quad -\infty < u \le v < \infty. \tag{5}$$

If ξ is an orthogonal increment process with weakly stationary increments, i.e. if $E|\xi(t+h) - \xi(t)|^2$ is independent of t, then

$$F(t) = \sigma^2 t, \quad \text{for some } \sigma > 0. \tag{6}$$

The mean-square stochastic integral

The integral,

$$I(f) = \int_{\mathbb{R}} f(u)\mathrm{d}\xi(u),$$

where $(\xi(u))_{u\in\mathbb{R}}$ is an OIP with associated distribution function F and $f \in L^2(F)$, the space of complex-valued Borel-measurable functions on $\mathbb{R}$ which are square integrable with respect to F, is a linear, inner-product preserving mapping of $L^2(F)$ into $L^2(P)$. Properties:

If f, g, f_n, g_n are in $L^2(F)$, $f = \text{l.i.m.}_{n\to\infty} f_n$ and $g = \text{l.i.m.}_{n\to\infty} g_n$, then

(i)

$$E(I(f)) = 0. \tag{7}$$

(ii)

$$I(a_1 f + a_2 g) = a_1 I(f) + a_2 I(g) \quad \text{for all } a_1, a_2 \in \mathbb{C}. \tag{8}$$

(iii)

$$E(I(f)\overline{I(g)}) = \int_{\mathbb{R}} f(u)\overline{g(u)}\mathrm{d}F(u). \tag{9}$$

(iv)

$$E(I(f_n)\overline{I(g_n)}) \to E(I(f)\overline{I(g)}) = \int_{\mathbb{R}} f(u)\overline{g(u)}\mathrm{d}F(u). \tag{10}$$

The spectral representation of a weakly stationary process

If $(\xi(u))_{u\in\mathbb{R}}$ is an OIP with associated distribution function F such that $\int_{\mathbb{R}} \mathrm{d}F(u) < \infty$, then

$$Y(t) := \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tu}\mathrm{d}\xi(u), \quad t \in \mathbb{R}, \tag{11}$$

is a weakly stationary process which by (7) has zero mean and by (9) has autocovariance function,

$$\gamma(h) = E(Y(t+h)\overline{Y}(t)) = \int_{\mathbb{R}} e^{ihu} dF(u), \quad h \in \mathbb{R}. \tag{12}$$

A proof that *every* zero-mean weakly stationary process $(Y(t))_{t\in\mathbb{R}}$ with γ continuous at 0 has a representation (11) in terms of an OIP ξ whose associated distribution function F satisfies the condition $\int_{\mathbb{R}} dF(u) < \infty$ is given by Doob (1953), Theorem 4.1. F is known as the *spectral distribution function* of the weakly stationary process X. If F is absolutely continuous with density f, i.e. if $F(t) = \int_0^t f(u)du$, $t \in \mathbb{R}$, (where $\int_0^t := -\int_t^0$ if $t < 0$), then (12) becomes

$$\gamma(h) = \int_{\mathbb{R}} e^{ihu} f(u) du. \tag{13}$$

Continuous-time moving averages

If $(\xi(u))_{u\in\mathbb{R}}$ is an OIP with weakly stationary increments [see (6)] with ADF, $F(t) = \sigma^2 t,\ t \in \mathbb{R}$, and if $g \in L^2(F)$ then, by (7) and (9), the process,

$$Y(t) := \int_{\mathbb{R}} g(t-u) d\xi(u), \quad t \in \mathbb{R}, \tag{14}$$

is a weakly stationary process with zero mean and autocovariance function,

$$\gamma(h) = E(Y(t+h)\overline{Y}(t)) = \sigma^2 \int_{\mathbb{R}} g(t+h)\overline{g(t)} dt. \tag{15}$$

Processes of the form (14) are known as continuous-time moving average (or CMA) processes. By appropriate choice of the kernel function g they can be made to exhibit a very large range of autocorrelation functions for the representation of dependent data. From (15), the Fourier transform, $\tilde{\gamma}(\omega) := \int_{\mathbb{R}} e^{-i\omega h}\gamma(h)dh$ is clearly

$$\tilde{\gamma}(\omega) = \sigma^2 |\tilde{g}(\omega)|^2, \quad \omega \in \mathbb{R}. \tag{16}$$

Inversion of this Fourier transform shows that $\gamma(h)$ can also be expressed as

$$\gamma(h) := \frac{\sigma^2}{2\pi} \int_{\mathbb{R}} |\tilde{g}(\omega)|^2 e^{i\omega h} d\omega, \tag{17}$$

and consequently the continuous-time moving average (14) has the spectral density,

$$f(\omega) = \frac{\sigma^2}{2\pi} |\tilde{g}(\omega)|^2. \tag{18}$$

Conversely [see Yaglom (1987), p. 453], given a zero-mean weakly stationary process Y with spectral density f, then any function h such that $|h(\omega)|^2 = 2\pi\sigma^{-2} f(\omega)$ is the Fourier transform of a function g for which there exists an OIP ξ with ADF

(6) such that (14) holds. The function h is determined only to within multiplication by a factor $e^{i\Theta(\omega)}$ where Θ is a real-valued function. Moreover the class of zero-mean weakly stationary processes which can be represented as a *one-sided* moving average of the form (14) with $\mathbb{R}$ replaced by $\mathbb{R}_+ = [0, \infty)$, coincides with the class of zero-mean weakly stationary processes which have a spectral density $f(\omega) = F'(\omega)$ satisfying the Paley–Wiener condition,

$$\int_{\mathbb{R}} \frac{\log F'(\omega)}{1+\omega^2} \mathrm{d}\omega > -\infty. \tag{19}$$

Thus the class of zero-mean weakly stationary processes representable as (14) consists of those which have a spectral density, and the class representable as a one-sided moving average, i.e. as (14) with $\mathbb{R}$ replaced by $\mathbb{R}_+$, consists of those which have a spectral density F' satisfying (19).

The Wold–Karhunen decomposition

The general structure of continuous-time weakly stationary processes with mean-square continuous sample-paths was treated comprehensively by Doob (1953). Suppose that Y is such a process and assume (with no loss of generality) that $E(Y(t)) = 0$. Then Y is said to be *regular* if the best (i.e. minimum mean-squared error) linear predictor of $Y(t+h),\ h > 0$, based on observations of $Y(u), u \leq t$, i.e. the orthogonal projection of $Y(t+h)$ on the closed linear span $\mathcal{N}_t$ of $\{Y(u),\ u \leq t\}$ in $L^2(P)$ has strictly positive mean squared error for some (and then necessarily for all) $h > 0$. This is equivalent to the derivative F' of its spectral distribution function F satisfying the Paley–Wiener condition (19). Assuming regularity of Y, Doob derives the best linear predictor of $Y(t+h),\ h > 0$, based on observations of $Y(u), u \leq t$. He also establishes the Wold–Karhunen continuous-time analogue [Karhunen (1950)] of the Wold decomposition in discrete time, namely

$$Y(t) = U(t) + V(t) = \int_0^\infty g(t-u)\mathrm{d}\xi(u) + V(t), \tag{20}$$

where g is Lebesgue measurable, $g(t) = 0$ for $t \leq 0$, $\int_{\mathbb{R}} |g(t)|^2 \mathrm{d}t < \infty$,

$$\int_{\mathbb{R}} g(t)e^{i\omega t}\mathrm{d}t \neq 0, \quad \mathrm{Im}(\omega) > 0, \tag{21}$$

ξ is an orthogonal increment process with $E|\mathrm{d}\xi(t)|^2 = \mathrm{d}t$ such that $\xi(t_2) - \xi(t_1) \in \mathcal{N}_t$ if $t_1, t_2 \leq t$, the process V is deterministic, i.e. $V(t) \in \mathcal{N}_{-\infty} (:= \cap_{t\in\mathbb{R}}\mathcal{N}_t)$ for all $t \in \mathbb{R}$, and every ξ increment is orthogonal to every $V(t)$. Condition (21) is the continuous-time analogue of the discrete-time property of *invertibility*, often referred to in the spectral domain as the *minimum phase* property. The best linear predictor of $Y(t+h)$ in terms of $Y(u), u \leq t$, is then

$$P(Y(t+h)|\mathcal{N}_t) = \int_0^{\infty} g(t-u)\mathrm{d}\xi(u) + V(t), \tag{22}$$

and the mean-squared error is

$$\sigma_h^2 = \int_0^h |g(t)|^2 \mathrm{d}t. \tag{23}$$

The spectral distribution function of U is $F_U(t) = \int_0^t F'(u)\mathrm{d}u$ (where $\int_0^t := -\int_t^0$ if $t < 0$) and that of V is $F - F_U$.

The second order properties of mean-square continuous weakly stationary processes Y are very well understood. In particular [see Yaglom (1987), pp. 163, 164] the sample-paths have mean-square derivatives of order n if and only if the autocovariance function, $\gamma(h) := \mathrm{Cov}(Y(t+h), Y(t))$, has a derivative of order $2n$ at $h = 0$. However the second-order properties tell us little about the distributional and sample-path properties of Y. They are also restricted of course to processes for which $E|Y(t)|^2 < \infty$ for all t. In order to deal with observed time series with marginal distributions exhibiting asymmetry or heavy tails it is natural to consider processes of the form (14) in which the process ξ is replaced by a process with more closely specified stucture in order to generate continuous-time models with desired sample-path properties and marginal distributions. The natural candidates to replace ξ are the Lévy processes, whose path structure and distributional structure have been intensively studied in recent years. See, for example, Applebaum (2004), Bertoin (1996), Kyprianou (2006), Protter (2004), Sato (1999).

Lévy processes

A *Lévy process*, $L = (L(t))_{t\in\mathbb{R}}$ is a process with independent and stationary increments, continuous in probability, with sample-paths which are right-continuous with left limits (càdlàg) and $L(0) = L(0-) = 0$. The most celebrated examples are Brownian motion, B, for which $B(t), t \geq 0$, is normally distributed with mean μt and variance $\sigma^2 t$ for some fixed $\mu \in \mathbb{R}$ and $\sigma > 0$, and the Poisson process, N, for which $N(t), t \geq 0$, has the Poisson distribution with mean λt for some fixed $\lambda > 0$. Notice that the processes $B(t)$ and the centred Poisson process, $N(t) - \lambda t$, like all Lévy processes with finite second moments and zero mean, are also processes with stationary orthogonal increments. In fact if $\lambda = \sigma^2$ they have exactly the same second-order properties, even though their sample-paths are quite different. Those of Brownian motion are almost surely continuous while those of the process $N(t) - \lambda t$ are almost surely linear with added jumps of size 1.

If L is a Lévy process the distribution of $L(t),\ t \geq 0$, is characterized by a unique triplet (σ^2, ν, γ) where $\sigma \geq 0$, $\gamma \in \mathbb{R}$ and ν is a measure on the Borel subsets of $\mathbb{R}$ such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}} \min\{|x|^2, 1\}\, \nu(dx) < \infty$. This triplet determines the characteristic function of $L(t)$ via the Lévy–Khintchine formula,

$$E\mathrm{e}^{i\theta L(t)} = \exp(t\zeta(\theta)), \quad \theta \in \mathbb{R}, \tag{24}$$

where

$$\zeta(\theta) = i\theta\gamma - \frac{1}{2}\theta^2\sigma^2 + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{\{|x|\le 1\}})\, \nu(dx). \tag{25}$$

The measure ν is called the *Lévy measure* of L and σ^2 the *Gaussian variance*. Conversely, if $\gamma \in \mathbb{R}$, $\sigma \ge 0$ and ν is a Lévy measure, then there exists a Lévy process L, unique up to identity in law, such that (24) and (25) hold. The triplet (σ^2, ν, γ) is called the *characteristic triplet* of the Lévy process L. The class of possible distributions for $L(1)$ is precisely the class of infinitely divisible distributions. For Brownian motion B with $E(B(t)) = \mu t$ and $\mathrm{Var}(B(t)) = \sigma^2|t|$, the characteristic triplet is $(\sigma^2, 0, \mu)$ and for a compound Poisson process with jump-rate λ and jump-size distribution function F, the characteristic triplet is $(0, \lambda dF(\cdot), \int_{[-1,1]} \lambda x dF(x))$.

A Lévy process L is called a *subordinator* if it has non-decreasing sample paths. This happens if and only if $\sigma^2 = 0$, $\nu((-\infty, 0)) = 0$ and $\int_0^1 x\, \nu(dx) < \infty$. Examples of subordinators include compound Poisson processes with jump distribution concentrated on $[0, \infty)$, the gamma process and the inverse Gaussian process. Subordinators play a key role in the modelling of non-negative time series (Sect. 6).

The *jump* of a Lévy process L at time t is defined as

$$\Delta L(t) := L(t) - L(t-).$$

Apart from Brownian motion with drift, every Lévy process has jumps. The Lévy measure $\nu(B)$ of a Borel set B is the expected number of jumps of L in the time interval $[0, 1]$ with jump (possibly negative) in the set B, i.e.

$$\nu(B) = E \sum_{0<s\le 1} \mathbf{1}_B(\Delta L(s)).$$

A Lévy process L has finitely many jumps in every bounded time-interval if and only if the Lévy measure of L is finite. Every Lévy process is a semimartingale and its quadratic variation is given by $[L, L]_t = \sigma^2 t + \sum_{0<s\le t} \Delta L(s)^2$. See Applebaum (2004) and Protter (2004) for further information regarding integration with respect to semimartingales.

2 Lévy-driven CARMA processes

From now on we shall restrict attention to processes with values in $\mathbb{R}$ or $\mathbb{R}^m$ for some positive integer m. The correlation (or second-order) theory of weakly stationary time series with index $t \in \mathbb{R}$ was outlined in Sect. 1. In order to deal with processes which may not have finite second moments and to incorporate distributional and sample-path properties in the modelling of such processes, we shall replace the driving orthogonal-increment processes ξ which appear in the continuous time moving averages (14) by Lévy processes. The distinction between OIP-driven and Lévy-driven stationary processes is analogous to the distinction in discrete time between *weak* moving averages of uncorrelated white noise and *strong* moving averages of independent white noise.

A weak ARMA(p, q) process in discrete time [see Brockwell and Davis (1991) for details] is a weakly stationary solution of difference equations of the form,

$$\phi(B)Y_n = \theta(B)Z_n, \quad n \in \mathbb{Z}, \tag{26}$$

where B is the backward shift operator (i.e. for any sequence X, $B^k X_n = X_{n-k}$), $(Z_n)_{n\in\mathbb{Z}}$ is a sequence of uncorrelated random variables, $\phi(z)$ is a polynomial of degree p, $\theta(z)$ is a polynomial of degree q, and the polynomials have no common zeroes. A strong ARMA(p, q) process is defined in the same way except that the random variables Z_n are required to be independent and identically distributed (i.i.d.) rather than simply uncorrelated. There is then a unique strictly stationary solution Y [see Brockwell and Lindner (2010)], if and only if $\phi(z) \neq 0$ when $|z| = 1$ and $E \log^+ |Z_1| < \infty$.

A natural continuous-time analogue of the difference equation (26) with i.i.d. noise is the *formal* differential equation,

$$a(D)Y(t) = b(D)DL(t), \tag{27}$$

where L is a Lévy process, D denotes differentiation with respect to t, and $a(z)$ and $b(z)$ are polynomials of the form,

$$\begin{aligned} a(z) &= z^p + a_1 z^{p-1} + \cdots + a_p,\\ b(z) &= b_0 + b_1 z + \cdots + b_q z^q, \end{aligned}$$

with $b_q = 1$ and $q < p$. Since the derivatives on the right-hand side of (27) do not exist in the usual sense, we give the equation a meaningful interpretation by rewriting it in the *state-space form*,

$$Y(t) = \mathbf{b}^T \mathbf{X}(t), \quad t \in \mathbb{R}, \tag{28}$$

where $\mathbf{X} = (\mathbf{X}(t))_{t\in\mathbb{R}}$ is a process with values in $\mathbb{R}^p$, satisfying the stochastic differential equation,

$$\mathrm{d}\mathbf{X}(t) = A\mathbf{X}(t)\,\mathrm{d}t + \mathbf{e}_p\,\mathrm{d}L(t), \tag{29}$$

or equivalently

$$\mathbf{X}(t) = \mathrm{e}^{A(t-s)}\mathbf{X}(s) + \int_{(s,t]} \mathrm{e}^{A(t-u)}\mathbf{e}_p\,\mathrm{d}L(u), \quad \forall s \le t \in \mathbb{R}, \tag{30}$$

with

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{bmatrix}, \quad \mathbf{e}_p = \begin{bmatrix} 0\\ 0\\ \vdots\\ 0\\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_0\\ b_1\\ \vdots\\ b_{p-2}\\ b_{p-1} \end{bmatrix},$$

where $a_1, \ldots, a_p, b_0, \ldots, b_{p-1}$ are real-valued coefficients satisfying $b_q = 1$ and $b_j = 0$ for $j > q$. For $p = 1$ the matrix A is to be understood as $A = (-a_1)$.

Remark 1 It is easy to check that the eigenvalues of the matrix A are the same as the zeroes of the autoregressive polynomial $a(z)$. We shall refer throughout to a typical zero as λ and denote its multiplicity by $m(\lambda)$. Thus $\sum_{\lambda} m(\lambda) = p$.

Path properties

The integral in (30) is a special case of integration with respect to a semimartingale as defined in Protter (2004), Chapter 2. Since the integrand is deterministic, continuous and of bounded variation on $[s, t]$, we can integrate by parts to obtain the pathwise interpretation of (30),

$$\mathbf{X}(t) = \mathrm{e}^{A(t-s)}\mathbf{X}(s) + \left[L(t)I_p - L(s)\mathrm{e}^{A(t-s)} + \int_s^t L(u)A\mathrm{e}^{A(t-u)}\mathrm{d}u\right]\mathbf{e}_p, \quad (31)$$

where I_p is the $p \times p$ identity matrix. This shows, in particular, that the components of $\mathbf{X}$ are continuous except possibly for the last (i.e. pth) component, which is càdlàg, with jumps coinciding with those of the driving Lévy process L. It shows also that the derivative of the jth component of $\mathbf{X}$ is the $(j + 1)$th component, $j = 1, \ldots, p - 1$. Note however that the $(p-1)$th component is in general differentiable only in the sense of having right and left derivatives since the pth component has jump discontinuities which are the same as those of L. The general CARMA(p, q) process defined by (32) is the linear combination, $\mathbf{b}^T\mathbf{X}$, of components of $\mathbf{X}$ and therefore has non-differentiable sample-paths if $q = p - 1$, since in that case the linear combination includes the last component of $\mathbf{X}$. (CARMA(p, q) processes with $q \geq p$ can be defined as generalized random functions [Brockwell and Hannig (2010)] but we shall not consider them here).

Existence and uniqueness

Sufficient conditions for the existence of a strictly stationary solution of (28) and (30) were obtained by Brockwell (2001b). Subsequently Brockwell and Lindner (2009), generalizing results of Wolfe (1982) and Sato and Yamazato (1984) for the Lévy-driven Ornstein–Uhlenbeck equation, established necessary and sufficient conditions for the existence of a strictly stationary, not necessarily causal, solution $Y = (Y(t))_{t\in\mathbb{R}}$. In Theorem 4.1 of that paper they also showed that there is no loss of generality in assuming that the polynomials $a(z)$ and $b(z)$ have no common zeroes since if that is the case and if Y is a solution of (28) and (30), then it also satisfies the analogous equations corresponding to the polynomials obtained by cancelling the common factors of $a(z)$ and $b(z)$. We shall therefore assume throughout that $a(z)$ and $b(z)$ have no common zeroes. We shall also assume that the driving processs L is not deterministic. Under these assumptions the strictly stationary solution of (28) and (30) is given by Theorem 3.3 of Brockwell and Lindner (2009), restated below.

Theorem 1 *Let L be a Lévy process which is not deterministic and suppose that $a(\cdot)$ and $b(\cdot)$ have no common zeroes. Then the CARMA equations (28) and (29) have a strictly stationary solution $(Y(t))_{t\in\mathbb{R}}$ if and only if $E\log^+|L(1)| < \infty$ and $a(\cdot)$ is non-zero on the imaginary axis. In this case the solution Y is unique and is given by*

$$Y(t) = \int_{-\infty}^{\infty} g(t-u)\,\mathrm{d}L(u), \quad t \in \mathbb{R}, \tag{32}$$

where

$$g(t) = g_1(t)\mathbf{1}_{[0,\infty)}(t) - g_2(t)\mathbf{1}_{(-\infty,0)}(t), \tag{33}$$

and the functions $g_1(t)$ and $g_2(t)$ are the sums of the residues of the function $z \mapsto \mathrm{e}^{zt}b(z)/a(z)$ in the left and right halves of the complex plane respectively. Equivalently,

$$g_k(t) = \frac{1}{2\pi i}\int_{\rho_k}\frac{b(z)}{a(z)}\mathrm{e}^{zt}\mathrm{d}z,\ k = 1, 2, \tag{34}$$

where integration is anticlockwise around the simple closed curves ρ_1 and ρ_2 in the open left and right halves of the complex plane respectively, encircling the zeroes of $a(z)$.

Remark 2 The integrals $\int_{(-\infty,t]}$ and $\int_{(t,\infty)}$ appearing when (33) is substituted in (32) are defined as almost sure limits as $T \to \infty$ of $\int_{(-T,t]}$ and $\int_{(t,T)}$ respectively.

Remark 3 If $a(z)$ has no zeroes in the right half-plane then g_2 is zero and Y is said to be causal, or more precisely a causal function of L.

Remark 4 The kernel function g can also be written as the inverse Fourier transform in L^2 of the square integrable function $b(i\omega)/a(i\omega)$, i.e.

$$g(t) = \frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{e}^{i\omega t}\frac{b(i\omega)}{a(i\omega)}\mathrm{d}\omega, \tag{35}$$

taken to be right continuous at $t = 0$. The sums of residues $g_1(t)$ and $g_2(t)$ in (33) can also be evaluated more explicitly as

$$g(t) = \sum_{\lambda:\Re\lambda<0}\sum_{k=0}^{m(\lambda)-1} c_{\lambda k}t^k\mathrm{e}^{\lambda t}\mathbf{1}_{[0,\infty)}(t) - \sum_{\lambda:\Re\lambda>0}\sum_{k=0}^{m(\lambda)-1} c_{\lambda k}t^k\mathrm{e}^{\lambda t}\mathbf{1}_{(-\infty,0)}(t), \tag{36}$$

where the sums are over the distinct zeroes λ of $a(\cdot)$, $m(\lambda)$ denotes the multiplicity of λ, $\Re\lambda$ denotes the real part of λ, and the residue, $\sum_{k=0}^{m(\lambda)-1} c_{\lambda k}t^k\mathrm{e}^{\lambda t}$, of the function $z \mapsto \mathrm{e}^{zt}b(z)/a(z)$ at λ is

$$\sum_{k=0}^{m(\lambda)-1} c_{\lambda k}t^k\mathrm{e}^{\lambda t} = \frac{1}{(m(\lambda)-1)!}\left[D_z^{m(\lambda)-1}\left((z-\lambda)^{m(\lambda)}\mathrm{e}^{zt}b(z)/a(z)\right)\right]_{z=\lambda},$$

where D_z denotes differentiation with respect to z.

Remark 5 Canonical decomposition. If $m(\lambda) = 1$ for every λ, the kernel reduces to

$$g(t) = \sum_{\lambda:\Re\lambda<0} e^{\lambda t} \frac{b(\lambda)}{a'(\lambda)} \mathbf{1}_{[0,\infty)}(t) - \sum_{\lambda:\Re\lambda>0} e^{\lambda t} \frac{b(\lambda)}{a'(\lambda)} \mathbf{1}_{(-\infty,0)}(t), \tag{37}$$

where $a'(\lambda)$ is the derivative of a at λ.

The simplified form of the kernel, (37), when the zeroes, $\lambda_1, \ldots, \lambda_p$ of the autoregressive polynomial $a(z)$ all have multiplicity one implies that in this case Y is a linear combination of (possibly complex-valued) CARMA$(1, 0)$ processes, $Y_1, \ldots, Y_p$, i.e.

$$Y(t) = \sum_{r=1}^{p} \alpha(\lambda_r) Y_r(t), \tag{38}$$

where

$$Y_r(t) = \begin{cases} \int_{(-\infty,t]} e^{\lambda_r(t-u)} \mathrm{d}L(u), & \text{if} \quad \Re\lambda_r < 0, \\ -\int_{(t,\infty)} e^{\lambda_r(t-u)} \mathrm{d}L(u), & \text{if} \quad \Re\lambda_r > 0, \end{cases}$$

and $\alpha(\lambda_r) = b(\lambda_r)/a'(\lambda_r),\ r = 1, \ldots, p$.

Remark 6 Matrix form of the kernel. The kernel g can also be expressed in matrix form as follows. For the zero λ of $a(z)$ with multiplicity $m(\lambda)$, define the p-component column-vector,

$$\boldsymbol{\phi}(\lambda) := [1\ \lambda\ \cdots\ \lambda^{p-1}]^T,$$

and the $p \times m(\lambda)$ matrix,

$$R(\lambda) = [\boldsymbol{\phi}(\lambda)\boldsymbol{\phi}^{(1)}(\lambda)\ \cdots\ \boldsymbol{\phi}^{(m(\lambda)-1)}(\lambda)], \tag{39}$$

where $\boldsymbol{\phi}^{(j)}$ denotes the jth derivative of $\boldsymbol{\phi}$. Labelling the distinct zeroes of $a(z)$ as $\lambda_1, \ldots, , \lambda_r$ (where $\sum_{i=1}^{r} m(\lambda_i) = p$), we next introduce the matrices,

$$T = [R(\lambda_1)\ R(\lambda_2)\ \cdots\ R(\lambda_r)]. \tag{40}$$

and

$$T^{-1} = \begin{bmatrix} L(\lambda_1) \\ L(\lambda_2) \\ \vdots \\ L(\lambda_r) \end{bmatrix}, \tag{41}$$

where, for each j, $L(\lambda_j)$ is an $m(\lambda_j) \times p$ matrix. Then

$$T^{-1}AT = \begin{bmatrix} J(\lambda_1) & 0 & \cdots & 0 \\ 0 & J(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J(\lambda_r) \end{bmatrix}, \tag{42}$$

where $J(\lambda_j)$ is a Jordan $m(\lambda_j) \times m(\lambda_j)$ block matrix with each diagonal component equal to λ_j, each superdiagonal component equal to one, and all other components equal to zero. Noting that $L(\lambda_i)R(\lambda_j)$ is the $m(\lambda_j) \times m(\lambda_j)$ identity matrix if $i = j$, and zero otherwise, we see that

$$\mathrm{e}^{At} = \sum_{i=1}^{r} R(\lambda_i)\mathrm{e}^{tJ(\lambda_i)}L(\lambda_i) = \sum_{i:\Re\lambda_i<0} + \sum_{i:\Re\lambda_i>0} R(\lambda_i)\mathrm{e}^{tJ(\lambda_i)}L(\lambda_i), \tag{43}$$

where $\mathrm{e}^{tJ(\lambda_i)}$ is the $m(\lambda_i) \times m(\lambda_i)$ upper triangular matrix with (j,k)-element $\mathrm{e}^{t\lambda_j}t^{k-j}/(k-j)!,\ k = j\ldots, m(\lambda_i),\ j = 1,\ldots, m(\lambda_i)$. Correspondingly,

$$\mathbf{b}^T\mathrm{e}^{At}\mathbf{e}_p = \sum_{i:\Re\lambda_i<0} + \sum_{i:\Re\lambda_i>0} \mathbf{b}^T R(\lambda_i)\mathrm{e}^{tJ(\lambda_i)}L(\lambda_i)\mathbf{e}_p,$$

where the first and second sums are the sums of the residues of the mapping $z \mapsto \mathrm{e}^{zt}b(z)/a(z)$ at the zeroes of $a(z)$ with negative and positive real parts respectively. Corresponding to (36) we thus have the representation of the kernel g as

$$g(t) = \sum_{\lambda:\Re\lambda<0} \mathbf{b}^T M(\lambda)\mathbf{e}_p \mathbf{1}_{[0,\infty)}(t) - \sum_{\lambda:\Re\lambda>0} \mathbf{b}^T M(\lambda)\mathbf{e}_p\mathbf{1}_{(-\infty,0)}, \tag{44}$$

where

$$M(\lambda) := R(\lambda)\mathrm{e}^{tJ(\lambda)}L(\lambda). \tag{45}$$

and the matrices $R(\lambda)$, $L(\lambda)$ and $\mathrm{e}^{tJ(\lambda)}$ are readily calculated for each of the distinct zeroes λ as described above. If all of the zeroes of $a(z)$ have negative real parts then the last term in (44) vanishes and, by (43), g reduces to

$$g(t) = \mathbf{b}^T\mathrm{e}^{At}\mathbf{e}_p\mathbf{1}_{[0,\infty)}(t). \tag{46}$$

Example 1 The stationary Ornstein–Uhlenbeck process. This is the special case of a CARMA(1, 0) process with $a(z) = z - \lambda$, $\lambda \in \mathbb{R}$, and $b(z) = 1$. Theorem 1 implies that the process exists if and only if $\lambda \neq 0$ and $E \log^+ |L_1| < \infty$. Under these conditions the process is unique and specified by (32) with

$$g(t) = \begin{cases} \mathrm{e}^{\lambda t}\mathbf{1}_{[0,\infty)}(t), & \text{if } \lambda < 0, \\ -\mathrm{e}^{\lambda t}\mathbf{1}_{(-\infty,0)}(t), & \text{if } \lambda > 0. \end{cases} \tag{47}$$

i.e.

$$Y(t) = \begin{cases} \int_{(-\infty,t]} \mathrm{e}^{\lambda(t-u)}\mathrm{d}L(u), & \text{if } \lambda < 0, \\ -\int_{(t,\infty)} \mathrm{e}^{\lambda(t-u)}\mathrm{d}L(u), & \text{if } \lambda > 0, \end{cases} \tag{48}$$

If $\lambda < 0$ the CARMA$(1, 0)$ process Y is causal and

$$Y(t) = \int_{(-\infty,t]} e^{\lambda(t-u)} dL(u), \text{ with } \lambda < 0. \tag{49}$$

The CARMA$(1, 0)$ process is also invertible in the sense that the increments $(L(t) - L(s))_{s \le t \le u}$ are in the closed linear span of $(Y(v))_{v \le u}$. This follows from the observation that in this case Y and X are the same so that (29) can be rewritten as

$$dY(t) = \lambda Y(t)dt + dL(t),$$

or equivalently,

$$L(t) - L(s) = Y(t) - Y(s) - \lambda \int_s^t Y(v)dv. \tag{50}$$

Recovery of the Lévy increments from the sample-path of Y

For inference concerning the nature of the driving process L of the CARMA process defined by (32) it is important to be able to recover the increments of L from a realization of Y as in the special case illustrated in Example 1. Provided the zeros of the polynomial $b(z)$ are all strictly less than zero, this can be done as indicated in the following theorem, for a proof of which we refer to Brockwell and Lindner (2014).

Theorem 2 *If the zeroes of the moving average polynomial $b(z)$ in (27) have strictly negative real parts then the increments of the driving Lévy process L satisfy*

$$L(t) - L(s) = \mathbf{e}_p^T \left(\mathbf{X}(t) - \mathbf{X}(s) - \int_s^t A\mathbf{X}(u)du \right), \ t \ge s, \tag{51}$$

where the components $X^{(0)}(t), X^{(1)}(t), \ldots, X^{(p-1)}(t)$, of the state-vector $\mathbf{X}(t)$are given by

$$X^{(j)}(t) = \sum_{k=0}^{j-q} \kappa^{(j-1-k)}(0+) Y^{(k)}(t) + \int_{(-\infty,t)} \kappa^{(j)}(t-u)Y(u)du, \quad 0 \le j \le p-1, \tag{52}$$

the superscripts denote order of differentiation, the sum is defined to be zero if $j - q < 0$, and $\kappa(t)$ is the sum of the residues of the mapping $z \mapsto e^{zt}/b(z)$ at the zeroes of $b(z)$. (If μ is a zero of $b(z)$ with multiplicity 1 then the residue at μ is $e^{\mu t}/b'(\mu)$).

If the zeroes of the autoregressive polynomial, $\lambda_1, \ldots, \lambda_p$ all have multiplicity 1, the increments of the driving Lévy process can also be recovered from any one of the component processes, $Y_r(t)$ in the canonical decomposition (38). Thus, cf. (50),

$$L(t) - L(s) = Y_r(t) - Y_r(s) - \lambda_r \int_s^t Y_r(v)dv, \quad r = 1, \ldots, p. \tag{53}$$

Example 2 The CARMA (2,1) process. For the L-driven CARMA process Y with autoregressive polynomial $a(z) = (z - \lambda_1)(z - \lambda_2) = z^2 + a_1 z + a_2$, and moving average polynomial $b(z) = z - \mu = b_0 + z$, with $\mu = -b_0 < 0$, the sum of the residues of $z \mapsto \mathrm{e}^{zt}/b(z)$ at the zeroes of $b(z)$ is just $\kappa(t) = \mathrm{e}^{-b_0 t}$. From (52) we therefore obtain, for $t \geq 0$,

$$X^{(0)}(t) = \int_{-\infty}^{t} \mathrm{e}^{-b_0(t-u)} Y(u)\mathrm{d}u = \mathrm{e}^{-b_0 t} X^{(0)}(0) + \int_0^t \mathrm{e}^{-b_0(t-u)} Y(u)\mathrm{d}u$$

and $X^{(1)}(t) = Y(t) - b_0 X^{(0)}(t)$. By (51) we then have, if $0 \leq s \leq t$,

$$L(t) - L(s) = Y(t) - Y(s) + (a_1 - b_0)(X^{(0)}(t) - X^{(0)}(s)) + a_2 \int_s^t X^{(0)}(u)\mathrm{d}u$$

where $X^{(0)}$ is given by (52).

Remark 7 Under the conditions of Theorem 1, the CARMA process Y defined by (32) can be written as $Y(t) = \mathbf{b}^T \mathbf{X}(t)$, where $\mathbf{X}$ is the unique strictly stationary solution of (30), namely

$$\mathbf{X}(t) = \int_{-\infty}^{\infty} \mathbf{f}(t-u)\,\mathrm{d}L(u), \quad t \in \mathbb{R}, \tag{54}$$

where

$$\mathbf{f}(t) = \mathbf{f}_1(t)\mathbf{1}_{[0,\infty)}(t) - \mathbf{f}_2(t)\mathbf{1}_{(-\infty,0)}(t), \tag{55}$$

and $\mathbf{f}_1(t)$ and $\mathbf{f}_2(t)$ are equal to the sums of the residues of the mapping $z \mapsto [1\ z\ \cdots\ z^{p-1}]^T \mathrm{e}^{zt}/a(z)$ in the left and right halves of the complex plane respectively. They simplify, when the zeroes of $a(\cdot)$ have multiplicity 1, to

$$\mathbf{f}_1(t) = \sum_{\lambda:\Re\lambda<0} \mathrm{e}^{\lambda t} \frac{[1\ \lambda\ \cdots\ \lambda^{p-1}]^T}{a'(\lambda)} \quad \text{and} \quad \mathbf{f}_2(t) = \sum_{\lambda:\Re\lambda>0} \mathrm{e}^{\lambda t} \frac{[1\ \lambda\ \cdots\ \lambda^{p-1}]^T}{a'(\lambda)}, \tag{56}$$

where $a'(\lambda)$ is the derivative of a at λ.

As in Remark 3, we can also write (regardless of the multiplicities of the zeroes of $a(z)$),

$$\mathbf{f}_1(t) = \sum_{\lambda:\Re\lambda<0} M(\lambda)\mathbf{e}_p \text{ and } \mathbf{f}_2(t) = \sum_{\lambda:\Re\lambda>0} M(\lambda)\mathbf{e}_p, \tag{57}$$

where $M(\lambda)$ was defined, for each distinct zero of $a(z)$, in (45). If the zeroes of $a(\cdot)$ all have negative real parts then $\mathbf{f}(t)$ can be expressed as

$$\mathbf{f}(t) = \mathrm{e}^{At}\mathbf{e}_p \mathbf{1}_{[0,\infty}(t). \tag{58}$$

In this case we can therefore write $Y(t) = \mathbf{b}^T \mathbf{X}(t)$ where

$$\mathbf{X}(t) = \int_{(-\infty,t]} \mathrm{e}^{A(t-u)}\mathbf{e}_p \mathrm{d}L(u). \tag{59}$$

Joint distributions

If Y is a causal CARMA process defined by (32) and the Lévy process L has the characteristic function specified in (24) and (25) then the the cumulant generating function of $Y(t_1), Y(t_2), \ldots, Y(t_n)$, $(0 \leq t_1 < t_2 < \cdots < t_n)$ is

$$\begin{aligned}
&\ln E[\exp(i\theta_1 Y(t_1) + \cdots + i\theta_n Y(t_n))] = \\
&\int_0^\infty \zeta\left(\sum_{i=1}^n \theta_i \mathbf{b}' \mathrm{e}^{A(t_i+u)}\right)\mathbf{e}_p \mathrm{d}u + \int_0^{t_1} \zeta\left(\sum_{i=1}^n \theta_i \mathbf{b}' \mathrm{e}^{A(t_i-u)}\right)\mathbf{e}_p \mathrm{d}u \\
&+ \int_{t_1}^{t_2} \zeta\left(\sum_{i=2}^n \theta_i \mathbf{b}' \mathrm{e}^{A(t_i-u)}\right)\mathbf{e}_p \mathrm{d}u + \cdots + \int_{t_{n-1}}^{t_n} \zeta\left(\theta_n \mathbf{b}' \mathrm{e}^{A(t_n-u)}\right)\mathbf{e}_p \mathrm{d}u. \qquad (60)
\end{aligned}$$

In particular, the marginal distribution of $Y(t)$ has cumulant generating function,

$$\ln E[\exp(i\theta Y(t))] = \int_0^\infty \zeta(\theta \mathbf{b}' \mathrm{e}^{Au} \mathbf{e}_p)\mathrm{d}u. \qquad (61)$$

A derivation of these results can be found in Brockwell (2001b).

Example 3 Symmetric stable L. If L is a symmetric stable process, then

$$\ln E\mathrm{e}^{i\theta L(t)} = -ct|\theta|^\alpha, \quad c > 0,\ 0 < \alpha \leq 2,$$

and by (61), $Y(t)$ has the symmetric stable marginal distribution with

$$\ln E\mathrm{e}^{i\theta Y(t)} = -c|\theta|^\alpha \int_0^\infty |\mathbf{b}' \mathrm{e}^{Au} \mathbf{e}_p|^\alpha \mathrm{d}u.$$

Example 4 CARMA(1,0). If Y is the causal CARMA(1,0) process (49) and $\ln E\mathrm{e}^{i\theta L(t)} = t\zeta(\theta)$, then by (61),

$$\kappa(\theta) := \ln E[\exp(i\theta Y(t))] = \int_0^\infty \zeta(\theta \mathrm{e}^{\lambda u})\mathrm{d}u,$$

from which, by a change of variable to $y = \theta \mathrm{e}^{\lambda u}$ in the last integral, we find that

$$\kappa(\theta) = |\lambda|^{-1} \int_0^\theta y^{-1} \zeta(y)\mathrm{d}y, \qquad (62)$$

which, in the special case of Example 3 reduces to $\kappa(\theta) = -c|\theta|^\alpha/(\alpha|\lambda|)$. The result (62), or equivalently $\zeta(\theta) = |\lambda|\theta\kappa'(\theta)$, is essentially the same as Eq. (12) of Barndorff-Nielsen and Shephard (2001) who used it to choose the Lévy process (i.e. $\zeta(\theta)$) required to generate a specified marginal distribution for the corresponding CARMA(1,0) process.

3 Second-order CARMA processes

If the driving Lévy process in (29) has the property $EL(1)^2 < \infty$ then $EY(t)^2 < \infty$ and the process Y is not only strictly stationary but also covariance stationary, i.e. $E(Y(t))$ is independent of t and $\mathrm{Cov}(Y(t+h), Y(t))$ is independent of t for all $h \in \mathbb{R}$ and denoted by $\gamma(h)$. In this section we shall examine these second order properties and the corresponding frequency-domain properties of the process Y.

For a second-order non-deterministic Lévy process the mean and variance of $L(t)$ are necessarily of the form $EL(t) = \xi t$ and $\mathrm{Var}(L(t)) = \sigma^2 t$, where $\xi = EL(1) \in \mathbb{R}$ and $\sigma = \sqrt{\mathrm{Var}(L(1))} > 0$. We shall use ξ and σ to denote these quantities whenever they exist. The process $(L(t) - \xi t)_{t\in\mathbb{R}}$ is an OIP as defined in Sect. 1, with associated distribution function, $F(t) = \sigma^2 t$.

The expected value of $Y(t)$ can be computed from (32) as $E(Y(t)) = \xi \int_{-\infty}^{\infty} g(t-u)\mathrm{d}u$. Substituting for g from (33) and (34) and integrating we obtain

$$EY(t) = -\frac{\xi}{2\pi i}\left[\int_{\rho_1} + \int_{\rho_2} \frac{b(z)}{za(z)}\mathrm{d}z\right] = \xi \mathrm{Res}_{z=0}\left[\frac{b(z)}{za(z)}\right] = \frac{\xi b_0}{a_p}. \qquad (63)$$

(Henceforth we shall use the notation $\mathrm{Res}_{z=\zeta} f(z)$ to denote the residue of the function f at ζ).

From (35) we see at once that the Fourier transform of the kernel g of Y is

$$\tilde{g}(\omega) = \frac{b(i\omega)}{a(i\omega)}.$$

Hence, by (18), Y has spectral density,

$$f(\omega) = \frac{\sigma^2}{2\pi}\left|\frac{b(i\omega)}{a(i\omega)}\right|^2.$$

Substituting this expression into (13) and changing the variable of integration from ω to $z = i\omega$ gives,

$$\gamma(h) = \frac{\sigma^2}{2\pi i}\int_{\rho} \frac{b(z)b(-z)}{a(z)a(-z)} \mathrm{e}^{|h|z}\mathrm{d}z = \sigma^2 \sum_{\tau:\Re\tau<0} \mathrm{Res}_{z=\tau}\left(\frac{b(z)b(-z)}{a(z)a(-z)}\mathrm{e}^{z|h|}\right), \qquad (64)$$

where the integration is clockwise around any simple closed curve ρ in the open left half-plane enclosing the distinct zeroes τ of $a(z)a(-z)$ with negative real parts.

If the process is causal the zeroes, τ, of $a(z)a(-z)$ in the left half-plane are the same as the zeroes, λ, of $a(z)$ and if in addition the multiplicity, $m(\lambda)$, of each of the zeroes is one, then $\gamma(h)$ takes the especially simple form,

$$\gamma(h) = \sigma^2 \sum_{\lambda} \frac{b(\lambda)b(-\lambda)}{a'(\lambda)a(-\lambda)}\mathrm{e}^{\lambda|h|}, \quad h \in \mathbb{R}. \qquad (65)$$

If we are concerned only with second-order properties of Y, i.e. properties which depend only on $EY(t)$ and $\gamma(\cdot)$, then there is no loss of generality in assuming that the zeroes of $a(z)$ and $b(z)$ all have negative real parts since, if some of them have positive real parts, a CARMA process of the same order and with the same mean and autocovariance function (but not in general with the same joint distributions) is obtained by changing the sign of each zero with strictly positive real part and multiplying L by $(-1)^m$ where m is the sum of the multiplicities of all of the zeroes whose signs are reversed. If the driving Lévy process is Brownian motion then Y is Gaussian and the distributional properties of Y are completely determined by its second-order properties. More generally it is very appealing, particularly from the point of view of simulating Y, that $Y(t)$ should depend only on the sample-path of L up to and including time t. Moreover, by Theorem 2, if the zeroes of $b(z)$ all have negative real parts the process is invertible, i.e. the increments $(L(t) - L(s))_{s \le t \le u}$ are in the closed linear span of $(Y(v))_{v \le u}$. For these reasons we shall assume for the remainder of this section that the zeroes of $a(z)$ and $b(z)$ all have negative real parts.

Under these conditions the general expression (64) for the autocovariance function of Y can also be expressed neatly in the matrix form,

$$\gamma(h) = \mathbf{b}'\mathrm{Cov}(\mathbf{X}(\mathbf{t}+\mathbf{h}), \mathbf{X}(\mathbf{t}))\mathbf{b}, \tag{66}$$

where, from (59), we see that

$$\mathrm{Cov}(\mathbf{X}(t+h), \mathbf{X}(t)) = \sigma^2 \mathrm{e}^{Ah}\Sigma, \quad h \ge 0, \tag{67}$$

with

$$\Sigma = \int_0^\infty \mathrm{e}^{At}\mathbf{e}_p\mathbf{e}_p'\mathrm{e}^{A't}\mathrm{d}t.$$

The matrix $\sigma^2\Sigma$ is clearly the covariance matrix of $\mathbf{X}(t)$. Since the jth component of $\mathbf{X}(t)$ is the $(j-1)$th mean square derivative of the first component, and since the first component is the CARMA$(p, 0)$ process obtained by setting $b(z) = 1$ in the definition of Y, the components $\sigma^2\Sigma_{i,j}$ have the form

$$\sigma^2\Sigma_{i,j} = \begin{cases} (-1)^{i+1}\kappa^{(i+j-2)}(0), & \text{if} \quad (i+j) \text{ is even,} \\ 0, & \text{otherwise,} \end{cases} \tag{68}$$

where $\kappa^{(j)}(0)$ denotes the jth derivative at 0 of the autocovariance function of the first component of $\mathbf{X}$ which, from (64), is

$$\kappa(h) = \sigma^2 \sum_\lambda \mathrm{Res}_{z=\lambda} \frac{\mathrm{e}^{z|h|}}{a(z)a(-z)},$$

and the sum, as usual, is over the distinct zeroes of $a(z)$ (which, as we are assuming, all have negative real parts). Thus

$$\kappa^{(j)}(0) = \sigma^2 \sum_{\lambda} \mathrm{Res}_{z=\lambda} \frac{z^j}{a(z)a(-z)}, \tag{69}$$

and, if the zeroes of $a(z)$ all have multiplicity one,

$$\kappa^{(j)}(0) = \sigma^2 \sum_{\lambda} \frac{\lambda^j}{a'(\lambda)a(-\lambda)}.$$

Calculation of the matrix $\sigma^2\Sigma$ can easily be carried out using (68) and (69) or by observing that the matrix $M := \Sigma^{-1} = [m_{i,j}]_{i,j=1}^p$ has the very simple representation [Arato (1982)],

$$m_{i,j} = \begin{cases} 2\sum_{k=0}^{\infty}(-1)^k a_{p-i-k}a_{p-j+1+k}, & \text{if } (i+j) \text{ is even,} \\ 0, & \text{otherwise,} \end{cases} \tag{70}$$

where $a_0 := 1$ and $a_j := 0$ if $j < 0$ or $j > p$. Finally from (66) and (67) we obtain the matrix representation of the autocovariance function γ of Y,

$$\gamma(h) = \sigma^2 \mathbf{b}' \mathrm{e}^{A|h|} \Sigma \mathbf{b}, \quad h \in \mathbb{R}. \tag{71}$$

Remark 8 Although, *from a second-order point of view*, one cannot distinguish between a causal and a non-causal process with parameters such that the second-order properties are the same, the distributional and sample-path properties of the causal and non-causal processes can be quite different. We return to this issue in Sect. 10 where we discuss non-causal processes in more detail.

4 Relating continuous-time and discrete-time ARMA processes

The discrete-time analogue of a Lévy-driven CARMA process is the ARMA process driven by i.i.d. noise. The ARMA(p, q) process with autoregressive coefficients $\phi_1, \ldots, \phi_p$ and moving average coefficients $\theta_1, \ldots, \theta_q$, driven by the i.i.d. sequence $(Z_n)_{n\in\mathbb{Z}}$, is a strictly stationary solution of the *difference* equations,

$$\phi(B)Y_n = \theta(B)Z_n, \tag{72}$$

where B is the backward shift operator, i.e. $B^jY_n = Y_{n-j}$, and $\phi(z)$, $\theta(z)$ are the autoregressive and moving average polynomials,

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q,$$

respectively. Brockwell and Lindner (2010) established the following existence and uniqueness theorem for strictly stationary solutions of (72).

Theorem 3 *Suppose that $(Z_n)_{n\in\mathbb{Z}}$ is a non-deterministic independent white noise sequence. Then the ARMA equation (72) admits a strictly stationary solution $(Y_n)_{n\in\mathbb{Z}}$ if and only if*

(i) $E\log^+|Z_1| < \infty$ and all singularities of $\theta(z)/\phi(z)$ on the unit circle are removable, i.e. if $\phi(z)$ has a zero ζ of multiplicity m on the unit circle, then ζ is a zero of $\theta(z)$ of multiplicity at least m,
or

(ii) all singularities of $\theta(z)/\phi(z)$ in $\mathbb{C}$ are removable.
If (i) or (ii) above holds, then a strictly stationary solution of (72) is given by

$$Y_n = \sum_{k=-\infty}^{\infty} \psi_k Z_{n-k}, \quad n \in \mathbb{Z}, \tag{73}$$

where

$$\sum_{k=-\infty}^{\infty} \psi_k z^k = \frac{\Theta(z)}{\Phi(z)}, \quad 1-\delta < |z| < 1+\delta \ \textit{for some}\ \delta \in (0,1),$$

is the Laurent expansion of $\Theta(z)/\Phi(z)$. The sum in (73) converges absolutely almost surely.
If Φ does not have a zero on the unit circle, then (73) is the unique strictly stationary solution of (72).

Remark 9 The sequence $(\psi_n)_{n\in\mathbb{Z}}$ of Laurent coefficients is the discrete-time analogue of the kernel $(g(t))_{t\in\mathbb{R}}$ in Theorem 1. The coefficient ψ_n can be expressed as

$$\psi_n = \frac{1}{2\pi i}\oint z^{-n-1}\frac{\theta(z)}{\phi(z)}\mathrm{d}z, \tag{74}$$

where the integral is anticlockwise around the unit circle. Equivalently [cf. (33)]

$$\psi_n = \psi_n^{(1)}\mathbf{1}_{[0,\infty)}(n) - \psi_n^{(2)}\mathbf{1}_{(-\infty,0)}(n), \tag{75}$$

where $\psi_n^{(1)}$ and $\psi_n^{(2)}$ are the sums of the residues of $z \mapsto z^{-n-1}\theta(z)/\phi(z)$ in the exterior and interior of the unit disc respectively, the former including the residue at infinity which is non-zero if $0 \le n \le q-p$ and zero for all non-negative n if $p > q$. (Recall that the residue at infinity of a meromorphic function f is defined as $\mathrm{Res}_{z=0}\left[-z^{-2}f(z^{-1})\right]$).

If Y is the CARMA(p,q) process (32) with $q = p-1$, $a(z) = \prod_{j=1}^{p}(z-\lambda_j)$, $b(z) = \prod_{j=1}^{q}(z-\mu_j)$ and kernel g, and if $(U_{n\Delta})_{n\in\mathbb{Z}}$ is the ARMA(p,q) process on the discrete time grid $\{n\Delta, n \in \mathbb{Z}\}$ with $\phi(z) = \prod_{j=1}^{p}(1-e^{\lambda_j\Delta}z)$, $\theta(z) = \prod_{j=1}^{q}(1-e^{\mu_j\Delta}z)$ and some driving i.i.d. sequence $(Z_n)_{n\in\mathbb{Z}}$ such that $E\log^+|Z_1| < \infty$, then it is straightforward to check, by making the change of variable $w = (1-z)/\Delta$ in (74) that

$$\lim_{\Delta\to 0}\psi_{[t/\Delta]} = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{zt}\frac{b(z)}{a(z)}dz = g(t), \tag{76}$$

i.e. that the kernel of $(U_{n\Delta})_{n\in\mathbb{Z}}$ at lag $[t/\Delta]$, the integer part of t/Δ, converges, as the grid spacing $\Delta \to 0$, to the kernel of Y at lag t. One can thus approximate the kernel of Y, for small Δ, using that of the discrete-time process $(U_{n\Delta})_{n\in\mathbb{Z}}$ as indicated in (76). We shall return to this issue in connection with the sampling of CARMA processes discussed in Sect. 7. The ARMA process $(U_{n\Delta})_{n\in\mathbb{Z}}$ introduced in this paragraph should not however be confused with the sampled process $(Y_n^\Delta := Y(n\Delta))_{n\in\mathbb{Z}}$ obtained by sampling Y at the grid points $n\Delta$, $n \in \mathbb{Z}$.

5 Integrated CARMA processes

The fractionally integrated process, FICARMA(p,d,q), $d \in (0,1/2)$

The fractionally integrated discrete-time ARMA(p,d,q) process (with $d \in (0,0.5)$ driven by a finite variance, zero mean, uncorrelated sequence of random variables, $(Z_n)_{n\in\mathbb{Z}}$, was introduced by Granger and Joyeux (1980) and Hosking (1981) as a *weakly* stationary solution of the difference equations [cf. (72)]

$$\phi(B)Y_n = \theta(B)(1-B)^{-d}Z_n, \tag{77}$$

where $(1-B)^{-d} := \sum_{j=0}^{\infty}\beta_j B^j$ and $\beta_j = \Gamma(d+j)/[\Gamma(d)\Gamma(j+1)]$. The sequence $(\beta_n)_{n\in\mathbb{N}_0}$ is the fractional summation kernel of order d.

Hosking (1981) showed that a sufficient condition for the existence of a weakly stationary solution is that $d < 1/2$ and $\phi(z)$ has no zero on the unit circle. It is now well-known [see e.g. Brockwell and Davis (1991), Theorem 13.2.2] that if ϕ and θ have no common zeroes and if $\phi(z) \neq 0$ for all z such that $|z| = 1$, then there is a unique purely non-deterministic stationary solution given by

$$Y_n = \sum_{j=-\infty}^{\infty}\zeta_j(1-B)^{-d}Z_{n-j},$$

where $\sum_{j=-\infty}^{\infty} \zeta_j z^j$ is the Laurent expansion of $\theta(z)/\phi(z)$, valid for $|1-|z||$ sufficiently small. The solution can also be written as the mean-square convergent sum,

$$Y_n = \sum_{j=-\infty}^{\infty} \psi_j Z_{n-j}, \tag{78}$$

where the sequence $(\psi_n)_{n\in\mathbb{Z}}$ is the convolution of the ARMA kernel $(\zeta_n)_{n\in\mathbb{Z}}$ and the fractional summation kernel $(\beta_n)_{n\in\mathbb{N}_0}$. If we make the stronger assumption that the driving noise sequence $(Z_n)_{n\in\mathbb{Z}}$ is i.i.d. then the sum (78) is absolutely convergent with probability one if $E|Z_0|^{1/(1-d)} < \infty$. [See Kokoszka and Taqqu (1995) and Vollenbröker (2012)].

Motivated by the form of the kernel in (78), Brockwell (2004) and Brockwell and Marquardt (2005) defined a fractionally integrated continuous-time ARMA (denoted FICARMA) process, driven by the second-order zero-mean Lévy process L, as the strictly (and covariance) stationary process,

$$Y_d(t) = \int_{(-\infty,\infty)} g_d(t-u)\mathrm{d}L(u),\ d \in (0, 0.5), \tag{79}$$

where g_d is the convolution of the CARMA kernel g in (33) with the Riemann-Liouville fractional integration kernel,

$$h(t) = \frac{t^{d-1}}{\Gamma(d)}\mathbf{1}_{[0,\infty)}(t), \tag{80}$$

i.e.

$$g_{\mathrm{d}}(t) = \int_{[0,\infty)} g(t-u)\frac{u^{d-1}}{\Gamma(d)}\mathrm{d}u,\ t \in \mathbb{R}. \tag{81}$$

It is sufficient to assume in the definition (79) that $a(z)$ has no zeroes on the imaginary axis. Brockwell and Marquardt (2005) made the stronger assumption that the zeroes of $a(z)$ all have strictly negative real parts, in which case Y is a causal function of L. Under the additional assumption that the zeroes of $a(z)$ all have multiplicity one, they also derived explicit expressions, in terms of the confluent hypergeometric function of the first kind, for the kernel g_d and the autocovariance function γ_d of the process Y_d. They found interesting parallels between the asymptotic behaviour of both the kernel and the autocovariance functions of the continuous- and discrete-time fractionally integrated processes. For the CARMA(p, d, q) process,

$$g_d(t) \sim \frac{t^{d-1}}{\Gamma(d)}\frac{b(0)}{a(0)} \text{ as } t \to \infty,$$
$$\gamma_d(h) \sim h^{2d-1}\frac{\mathrm{Var}(L(1))\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}\left[\frac{b(0)}{a(0)}\right]^2 \quad \text{as } h \to \infty,$$

while for the ARMA(p, d, q) process (77) [see Beran (1994)],

$$\psi_j \sim \frac{j^{d-1}}{\Gamma(d)} \frac{\theta(1)}{\phi(1)} \quad \text{as } j \to \infty,$$
$$\gamma_Y(h) \sim h^{2d-1} \frac{\mathrm{Var}(Z_0)\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)} \left[\frac{\theta(1)}{\phi(1)}\right]^2 \quad \text{as } h \to \infty.$$

The power, as opposed to exponential, decay of their kernel and autocovariance functions is what distinguishes the fractionally integrated CARMA/ARMA processes from their CARMA/ARMA counterparts.

Remark 10 A fractionally integrated Gaussian CARMA process was introduced by Comte and Renault (1996) as a CARMA process driven by fractionally integrated Brownian motion. An analogous approach in which fractional Brownian motion was replaced by a fractional Lévy motion was taken by Marquardt (2006). This approach opens the door to a much larger class of long-memory continuous-time moving average processes.

Remark 11 The connection between the kernels of the CARMA process and the corresponding fractionally integrated process can be seen by comparing the CARMA kernel (35) with the FICARMA kernel,

$$g_{\mathrm{d}}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{i\omega t} (i\omega)^{-d} \frac{b(i\omega)}{a(i\omega)} \mathrm{d}\omega, \quad t \in \mathbb{R}. \tag{82}$$

Remark 12 It is crucial, in the definition (79) of the second-order fractionally integrated CARMA process, that $EL(1) = 0$. If $EL(1) = \mu \neq 0$ then the integral (79) is not defined.

Remark 13 A heavy-tailed fractionally integrated process can be defined by (79) provided $EL(1)^{1/1-d} < \infty$. This is the analogue of the corresponding condition $E|Z_0|^{1/1-d}$ for the existence of the heavy-tailed discrete-time fractionally-integrated ARMA process.

The integrated process ICARMA(p, d, q), $d \in \mathbb{N}$

The d-fold integrated CARMA process with $d \in \mathbb{N}$ is the non-stationary process defined by

$$\begin{aligned} Y_d(t) &= \int_0^t \int_0^{u_{d-1}} \cdots \int_0^{u_1} Y(u)\, \mathrm{d}u\, \mathrm{d}u_1 \ldots \mathrm{d}u_{d-1} \\ &= \int_0^t \frac{(t-u)^{d-1}}{\Gamma(d)} Y(u) \mathrm{d}u, \ d \in \mathbb{N}, \end{aligned} \tag{83}$$

where Y is a CARMA(p, q) process and $\frac{t^{d-1}}{\Gamma(d)}$ is the d-fold integration kernel introduced (for $d \in (0, 1/2)$) in (81).

The integral on the right of (83) can be evaluated explicitly as follows. From (29) we see at once that

$$\int_0^t \mathbf{X}(u)\mathrm{d}u = A^{-1}[\mathbf{X}(t) - \mathbf{X}(0) - \mathbf{e}_p L(t)]. \tag{84}$$

Iterating (84) $(d-1)$ times gives the d-fold integral,

$$\begin{aligned}\mathbf{X}_d(t) &= \int_0^t \frac{(t-u)^{d-1}}{\Gamma(d)}\mathbf{X}(u)\mathrm{d}u \\ &= A^{-d}\left[\mathbf{X}(t) - \sum_{j=0}^{d-1}\frac{A^j}{j!}\left(t^j\mathbf{X}(0) - \int_0^t (t-u)^j \,\mathrm{d}L(u)\,\mathbf{e}_p\right)\right],\end{aligned} \tag{85}$$

where, from (59), $\mathbf{X}(t) = \int_{(-\infty,t]} \mathrm{e}^{A(t-u)}\mathbf{e}_p \mathrm{d}L(u)$, and from (28),

$$Y_d(t) = \mathbf{b}^T \mathbf{X}_{\mathrm{d}}(t). \tag{86}$$

Example 5 The integrated stationary Ornstein–Uhlenbeck process. For the causal stationary Ornstein–Uhlenbeck process,

$$Y(t) = \int_{(-\infty,t]} \mathrm{e}^{\lambda(t-u)}\mathrm{d}L(u),\ \ \lambda < 0,$$

the process and the state vector are the same and so from (85) with $d = 1$ we find at once that the integrated process Y_1 is given by

$$Y_1(t) = \lambda^{-1}[Y(t) - Y(0) - L(t)],\ \ t \geq 0. \tag{87}$$

This process has been studied in detail by Barndorff-Nielsen and Shephard (2003). The canonical decomposition (38), in conjunction with (87), gives the following simple expression for the integrated process Y_1 when Y is a causal CARMA process with distinct autoregressive zeroes:

$$Y_1(t) = \sum_{\lambda} \lambda^{-1}\alpha(\lambda)[Y^{(\lambda)}(t) - Y^{(\lambda)}(0) - L(t)],$$

where the sum is over the zeroes of $a(z)$, the coefficients $\alpha(\lambda)$ are defined as in (38) and $Y^{(\lambda)}$ is the CARMA(1, 0) process,

$$Y^{(\lambda)} = \int_{(-\infty,t]} \mathrm{e}^{\lambda(t-u)}\mathrm{d}L(u).$$

6 Non-negative CARMA processes

CARMA processes with a non-negative kernel, driven by a non-decreasing Lévy process constitute a useful and very general class of stationary, non-negative continuous-time processes which have been used, in particular, for the modelling of stochastic volatility.

In financial econometrics a Lévy-driven CAR(1) (or stationary Ornstein–Uhlenbeck) process was used by Barndorff-Nielsen and Shephard (2001) to represent the spot volatility $V(t)$ in their celebrated model,

$$\mathrm{d}X^*(t) = (\mu + \beta V(t))\mathrm{d}t + \sqrt{V(t)}\mathrm{d}W(t), \tag{88}$$

for the logarithm, $X^*(t)$, of the price of an asset at time t. In this model μ and β are constants, W is standard Brownian motion and the volatility process V is a stationary causal non-negative Lévy-driven Ornstein–Uhlenbeck process, independent of W, satisfying

$$\mathrm{d}V(t) - \lambda V(t)\mathrm{d}t = \mathrm{d}L(t), \quad \lambda < 0,$$

i.e.

$$V(t) = \int_{-\infty}^{t} \exp(\lambda(t-u))\mathrm{d}L(u). \tag{89}$$

Since the kernel function $g(t) = \mathrm{e}^{\lambda t}$ is non-negative, the volatility process will be non-negative (as required) if the Lévy process L has non-decreasing sample-paths. Lévy processes with this property are known as *subordinators*. An example is the gamma process, whose increments over any interval of length t have a gamma distribution with probability density $(\Gamma(\alpha t))^{-1}\beta^{\alpha}x^{\alpha t-1}\mathrm{e}^{-\beta x}\mathbf{1}_{[0,\infty)}(t)$ for some $\alpha > 0$ and $\beta > 0$.

In financial econometrics it is normally assumed that V has finite second moments. The mean and autocovariance function of V are then, from (63) and (65), $EV(t) = \xi/|\lambda|$ and $\gamma(h) = \sigma^2\mathrm{e}^{\lambda|h|}/(2|\lambda|)$. The dependence structure implied by this autocovariance function is more restrictive than one would like for modelling purposes and so Barndorff-Nielsen and Shephard (2002) proposed the use of superpositions of Ornstein–Uhlenbeck processes in order to expand the class of achievable autocovariance functions. However a considerably larger class can be achieved by replacing the Ornstein–Uhlenbeck process, (89) by a subordinator-driven causal CARMA process with non-negative kernel. Conditions on the defining polynomials $a(z)$ and $b(z)$ under which the kernel is non-negative were given by Tsai and Chan (2005) and, in the special case of the CARMA(2,1) process by Brockwell and Davis (2001). Even with these restrictions, the class of achievable autocovariance functions is large and includes in particular non-monotone functions.

Simulation of Lévy-driven CARMA stochastic volatility models was considered by Todorov and Tauchen (2006). Brockwell et al. (2011) considered estimation for subordinator-driven non-negative CARMA(p, q) processes based on uniformly-spaced observations. They also considered the problem of recovering the increments of the driving subordinator from closely-spaced observations of the CARMA process and found that for the DM/US$ exchange rate, from December 1st, 1986, through June 30th, 1999, a gamma-driven CARMA(2,1) process fitted the *daily realized volatility* series reasonably well.

Brockwell and Lindner (2013) considered parameter estimation for a causal subordinator-driven CARMA(p, q) model for the *spot volatility process* V by showing that for any fixed $\Delta > 0$, the Δ-integrated volatility sequence,

$$I_n^\Delta := \int_{(n-1)\Delta}^{n\Delta} V(t)\mathrm{d}t,\ n \in \mathbb{Z}, \tag{90}$$

satisfies the discrete-time ARMA equations

$$\phi(B)I_n^\Delta = U_n,\ n \in \mathbb{Z}, \tag{91}$$

where B is the backward shift operator ($B^j Y_n := Y_{n-j}$), $\phi(z)$ is the polynomial,

$$\phi(z) = \prod_\lambda (1 - \mathrm{e}^{\lambda\Delta} z)^{m(\lambda)},$$

the product is over the distinct zeroes of the autoregressive polynomial $a(z)$, and $(U_n)_{n\in\mathbb{Z}}$ is a p-dependent sequence which can be expressed as a moving average,

$$U_n = EU_0 + \theta(B)Z_n,$$

where $\theta(z)$ is a polynomial of the form,

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_p z^p,$$

$(Z_n)_{n\in\mathbb{Z}}$ is an uncorrelated (but not necessarily independent) zero-mean white noise sequence, and $EU_0 = \Delta\phi(1)\xi b_0/a_p$. This implies that $(I_n^\Delta)_{n\in\mathbb{Z}}$ is a weak ARMA(p, q) process with $q \leq p$. In the case $p = 1$ this was already well-known [Barndorff-Nielsen and Shephard (2001)], however the autocorrelation function of the ARMA(1,1) model is restricted for lags h greater than zero to functions of the form $c\phi_1^h$, $c, \phi_1 > 0$. The purpose of introducing the finite variance CARMA(p, q) model for spot volatility was to escape from this restriction in order to obtain a more realistic representation of integrated volatility as estimated in practice by the so-called *realized volatility*, which we shall denote by R^Δ. The time-interval Δ is usually one day and the realized volatility is typically calculated from observations of the asset price at 30-min intervals. Andersen et al. (2003) consider the forecasting of realized volatility based on high-frequency data, indicating that for the daily realized volatility of exchange rates, 30-min data achieves a reasonable compromise between high-frequency and the interference effect of market microstructure. Excellent accounts of realized volatility with many references are given by Andersen and Benzoni (2009) and Mykland and Zhang (2012).

Parameter estimation for V was carried out by searching numerically for the coefficients of a causal and invertible CARMA model which minimizes the sum of squares of the linear one-step prediction errors, when the corresponding weak ARMA model for I^Δ is applied to the data R^Δ. The sum of squares of the one-step linear prediction errors can be calculated directly as a function of the CARMA parameters and minimized numerically. Under mild conditions this procedure leads to strongly consistent estimators of the CARMA coefficients.

Using these estimates, we can simulate the spot volatility process V and the corresponding Δ-integrated volatility I^Δ using a variety of driving subordinators. Choosing the mean and variance of the driving subordinators so as to match the sample mean and variance of the realized volatility process V^Δ, we can then compare the empirical marginal distribution of the simulated integrated volatility series with that of V^Δ. Application of this technique to the DM/US$ daily realized volatility series gave a remarkably good fit to the empirical marginal distribution of daily realized volatility using the least squares CARMA(2,1) spot volatility model driven by a gamma subordinator with appropriately chosen $\mathrm{E}L(1)$ and $\mathrm{Var}L(1)$.

7 Sampling and embedding

Since continuous-time realizations are rarely if ever observed, it is important to understand the nature of the processes obtained when a continuous-time stationary process is sampled, in particular when it is sampled at uniformly-spaced times $\{n\Delta, n \in \mathbb{Z}\}$.

From Brockwell and Lindner (2009) (Lemma 2.1) we know that if Y is the causal CARMA process (32), then the sampled process $(Y_n^\Delta := Y(n\Delta))_{n\in\mathbb{Z}}$ satisfies the discrete-time equations,

$$\phi(B)Y_n^\Delta = U_n, \ n \in \mathbb{Z}, \tag{92}$$

where B is the backward shift operator ($B^j Y_n^\Delta := Y_{n-j}^\Delta$), $\phi(z)$ is the same polynomial as in (91), i.e.

$$\phi(z) = \prod_\lambda (1 - \mathrm{e}^{\lambda\Delta} z)^{m(\lambda)}, \tag{93}$$

and $(U_n)_{n\in\mathbb{Z}}$ is a $(p-1)$-dependent sequence which can be writen explicitly as

$$U_n = \mathbf{b}^T \sum_{r=0}^{p-1} \sum_{j=0}^{r} d_j \mathrm{e}^{(r-j)A\Delta} \mathbf{R}_{n-r}, \ n \in \mathbb{Z}, \tag{94}$$

where d_j is the coefficient of z^j in $\phi(z)$, $j = 0, \ldots, p$, and

$$\mathbf{R}_n = \int_{(n-1)\Delta}^{n\Delta} \mathrm{e}^{A(n\Delta - u)} \mathbf{e}_p \, dL(u), \ n \in \mathbb{Z}.$$

If $EL(1)^2 < \infty$ and $EL(1) = \xi$ then, since the sequence U in (92) is $(p-1)$-dependent, the mean-corrected sampled series $(Y_n^\Delta - \xi b_0/a_p)_{n\in\mathbb{Z}}$ satisfies the ARMA equations,

$$\phi(B)(Y_n^\Delta - \xi b_0/a_p) = \theta(B)Z_n, \tag{95}$$

where $\phi(\cdot)$ is the polynomial defined in (93) and $\theta(z)$ is a polynomial of the form,

$$\theta(z) = \prod_{j=1}^{p-1}(1 - \zeta_j z), \tag{96}$$

with $|\zeta_j| < 1$ for all j, and $(Z_n)_{n\in\mathbb{Z}}$ an uncorrelated (but not necessarily independent) zero-mean white noise sequence with finite variance σ_Δ^2. The Wold representation of $Y_n^\Delta - \xi b_0/a_p$ is then

$$Y_n^\Delta - \xi b_0/a_p = \psi(B)Z_n \tag{97}$$

where

$$\psi(z) = \theta(z)/\phi(z), \ |z| \leq 1.$$

Since the AR polynomial in (92) is known exactly, in order to study the second-order properties of the sampled process Y^Δ, we shall focus attention on the moving average component $\theta(B)Z_n$, and in particular on the polynomial $\theta(\cdot)$ and the white-noise variance σ_Δ^2. The spectral density f_Δ of Y^Δ is well-known [see e.g. Bloomfield (2000)] to be related to that of Y by

$$f_\Delta(\omega) = \frac{1}{\Delta}\sum_{k=-\infty}^{\infty} f_Y\left(\frac{\omega + 2k\pi}{\Delta}\right), \ \ \omega \in [-\pi, \pi],$$

but for our purposes it will be more convenient to express f_Δ as

$$f_\Delta(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \mathrm{e}^{-in\omega}\gamma(n\Delta), \quad -\pi \leq \omega \leq \pi. \tag{98}$$

where the function γ was specified in (64). Hence

$$f_\Delta(\omega) = \frac{\sigma^2}{2\pi}\sum_\lambda \mathrm{Res}_{z=\lambda}\left[\frac{b(z)b(-z)}{a(z)a(-z)} \cdot \frac{\sinh(\Delta z)}{\cos\omega - \cosh(\Delta z)}\right], \ \ \pi \leq \omega \leq \pi. \tag{99}$$

where the sum is over the distinct zeroes of $a(z)$ which, as we are assuming, all lie in the open left half of the complex plane.

From (95) and the definition of ϕ the spectral density of the moving average $\theta(B)Z_n$ is given by

$$f_{MA}(\omega) = \psi(\omega)f_\Delta(\omega), \quad -\pi \leq \omega \leq \pi, \tag{100}$$

where

$$\psi(\omega) = \left|\prod_\lambda(1 - \mathrm{e}^{i\omega+\lambda\Delta})\right|^2 = 2^p\mathrm{e}^{-a_1\Delta}\prod_{j=1}^{p}(\cosh(\lambda_j\Delta) - \cos(\omega)),$$

and $\lambda_1, \ldots, \lambda_p$ are the (not necessarily distinct) zeroes of $a(\cdot)$. But we know, from (95) and (96), that

$$f_{MA}(\omega) = \frac{\sigma_\Delta^2}{2\pi} \prod_{j=1}^{p-1} \left|1 - \zeta_j e^{-i\omega}\right|^2. \tag{101}$$

The determination of σ_Δ^2 and $\theta(\cdot)$, or equivalently $\zeta_1, \ldots, \zeta_{p-1}$, thus boils down to the factorization of the polynomial (100). Explicit solution of this problem for fixed positive Δ is not possible in general, however the asymptotic behaviour of σ_Δ^2 and the coefficients ζ_j as $\Delta \to 0$ can be derived as shown below.

High-frequency sampling when $EL(1)^2 < \infty$

Assuming that the process Y is both causal and invertible, we can write the autoregressive and moving average polynomials defining Y as

$$a(z) = \prod_{i=1}^{p}(z - \lambda_i) \text{ and } b(z) = \prod_{i=1}^{q}(z - \mu_i)$$

where $\Re\lambda_i < 0, i = 1, \ldots, p$ and $\Re\mu_i \leq 0, i = 1, \ldots, q$. Now let

$$\beta(x) := \text{Res}_{z=0} \frac{z^{-2(p-q)} \sinh(z)}{\cosh(z) - 1 + x} = \frac{1}{(2(p-q)-1)!x^{p-q}} \prod_{i=1}^{p-q-1} (x - \xi_i), \tag{102}$$

and define

$$\eta(\xi_i) = \xi_i - 1 \pm \sqrt{(\xi - 1)^2 - 1}, \quad i = 1, \ldots, p - q - 1, \tag{103}$$

where the sign is chosen so that $|\eta(\xi)| < 1$.

It was shown by Brockwell et al. (2013) that, as $\Delta \to 0$, the zeroes of the spectral density, $f_{MA}(\omega)$, $\pi \leq \omega \leq \pi$, of $\theta(B)Z_n$ in (95) occur where

$$\cos(\omega) = 1 - \xi_j(1 + o(1)), \quad j = 1, \ldots, p - q - 1,$$

and

$$\cos(\omega) = 1 + \frac{\mu_j^2 \Delta^2}{2}(1 + o(1)), \quad j = 1, \ldots, q.$$

The invertible moving average polynomial is therefore given by (96) with

$$\zeta_j = -\eta(\xi_j) + o(1), \quad j = 1, \ldots, p - q - 1,$$

and

$$\zeta_{p-q-1+j} = 1 + \mu_j \Delta + o(\Delta), \quad j = 1, \ldots, q.$$

From the asymptotic form of $f_{MA}(\omega)$ as $\Delta \to 0$ it was also found that the variance σ_Δ^2 of Z_n satisfies

$$\sigma_\Delta^2 = \frac{\Delta^{2(p-q)-1}\sigma^2}{[2(p-q)-1]! \prod_{i=1}^{p-q-1} \eta(\xi_i)}(1 + o(1)),$$

where σ^2 is the variance of $L(1)$. Applying these asymptotic results to the Wold representation (97) of the sampled process, Brockwell et al. (2013), assuming that the zeroes of $a(z)$ all have multiplicity one, derived the asymptotic behaviour of the Wold coefficients ψ_k^Δ of the sampled process Y^Δ and used them to show the pointwise convergence, as $\Delta \to 0$, of the function

$$g^\Delta(t) := \sum_{j=0}^{\infty} \frac{\sigma_\Delta}{\sqrt{\Delta}} \psi_j^\Delta \mathbf{1}_{[j\Delta,(j+1)\Delta)}(t) \tag{104}$$

to σg (or to g if L is standardized so that $\mathrm{Var}(L_1) = 1$), where g is the kernel of the CARMA process Y. They then used a non-parametric estimator of $g^\Delta(t)$ to estimate $g(t)$, showing for Gaussian CARMA processes that, if n is the number of observations and $\Delta(n)$ is the sampling interval, the estimator is consistent as $n \to \infty$ provided $\Delta(n) \to 0$ in such a way that $n\Delta(n) \to \infty$. Under stronger restrictions on the rates, the estimator is also asymptotically normal with variance depending on g. The estimator was used by Brockwell et al. (2013) in their analysis of the extremely high-frequency (5000 Hz) Brookhaven turbulent windspeed data. A spectral approach to estimation based on high-frequency samples from a CARMA process has been taken by Fasen and Fuchs (2013).

Embedding

We have seen at the beginning of this section that if Y is a causal Lévy-driven CARMA(p, q) process the sampled sequence $(Y_n^\Delta)_{n\in\mathbb{Z}}$ is a strictly stationary process satisfying (92) and that if $EL(1)^2 < \infty$ it satisfies the ARMA$(p, p-1)$ equations (95) driven by an uncorrelated (but not in general i.i.d.) white noise sequence. It may also satisfy ARMA equations of lower order if the polynomials $\phi(z)$ and $\theta(z)$ have common factors. Determination of the parameters of the moving average term $\theta(B)Z_n$ and the variance of Z_n in (95) requires factorization of the spectral density (100). This factorization cannot be done analytically except in very simple cases but algorithms exist for doing it numerically.

The fact that the regularly sampled process is an ARMA process suggests the possibility of estimating parameters for the continuous-time process by estimating the ARMA parameters of the sampled process and determining a corresponding continuous-time ARMA process. This technique was first used by Phillips (1959). It raises the question of whether, for a given ARMA process $(X_n)_{n\in\mathbb{Z}}$, there exists a CARMA process Y whose sampled sequence $(Y(n))_{n\in\mathbb{Z}}$ has the same autocovariance function as the sequence X and if so, is the process Y unique. The answer to both questions is no. It was shown by Brockwell and Brockwell (1999) that if the moving average polynomial of the ARMA process has a zero on the unit circle then its autocovariance function cannot be that of any regularly sampled CARMA process. When a suitable CARMA process can be found it may or may not be unique and the possible non-uniqueness gives rise to the so-called problem of aliassing. This is illustrated by Example 5 in Brockwell (2001a) which demonstrates that, depending on the ARMA process there may be zero, one, some finite number or infinitely many CARMA processes Y for which the autocovariance function of $(Y(n))_{n\in\mathbb{Z}}$ matches that of the specified ARMA process. There have been many studies of these and related questions but there is still no simple criterion, in terms of ARMA parameters, for deciding whether or not embedding is possible or determining the one or more CARMA processes in which the embedding can be made. A recent paper (Thornton and Chambers (2013)) provides a good list of references on this topic and characterizes the ARMA processes which can be embedded in a CARMA(2, 1) process.

8 Inference for CARMA processes

When observations of the second-order CARMA process defined by (28) and (30) with $E(L(1)) = \xi$ and $\mathrm{Var}(L(1)) = \sigma^2$ are available at the possibly irregularly-spaced times $t_1, \ldots, t_N$, the corresponding state-vectors and observations at those times are immediately found to satisfy the equations,

$$Y(t_i) = \mathbf{b}^T \mathbf{X}(t_i) \tag{105}$$

and

$$\mathbf{X}(t_i) = \mathrm{e}^{A(t_i - t_{i-1})}\mathbf{X}(t_{i-1}) + \int_{t_{i-1}}^{t_i} \mathrm{e}^{A(t-t_{i-1})}\mathbf{e}_p \mathrm{d}L(t), \tag{106}$$

where $\mathbf{X}(t_1)$ has the distribution of $\int_0^\infty \mathrm{e}^{Au}\mathbf{e}\mathrm{d}L(u)$. The *observation equation* (105) and *state equation* (106) are in precisely the form required for application of the discrete-time Kalman recursions [see e.g. Brockwell and Davis (1991)] in order to compute numerically the best one-step linear predictors of $Y(t_1), \ldots, Y(t_N)$, their mean-squared errors, and hence the Gaussian likelihood of the observations in terms of the coefficients $\{a_j, 1 \le j \le p;\ b_j, 0 \le j < q\}$ and the parameters ξ and σ of L. Jones (1981) used this representation, together with numerical maximization of the calculated Gaussian likelihood, to compute maximum Gaussian likelihood estimates of the parameters for time series with irregularly spaced data. A similar approach was used in a more general setting by Bergstrom (1985). If the observations are uniformly spaced, an alternative approach due to Phillips (1959) is to fit a discrete-time ARMA model to the observations and then to determine a CARMA process in which the discrete-time process can be embedded. (Recalling the results of Sect. 7 however, it may be the case that there is no CARMA process in which the fitted ARMA process can be embedded). In the general case it is also important to make inferences about the driving process L in the general case when it might not be Gaussian.

For a CAR(p) process observed continuously on the time interval $[0, T]$, Hyndman (1993) derived continuous-time analogues of the discrete-time Yule-Walker equations for estimating the coefficients. For a Gaussian CARMA process observed continuously on $[0, T]$, the exact likelihood function was determined by Pham-Dinh (1977), who also gave a computational algorithm for computing approximate maximum likelihood estimators of the parameters which are asymptotically normal and efficient. The determination of the exact likelihood, conditional on the initial state vector $\mathbf{X}(0)$, can also be carried out for non-linear Gaussian CAR(p) processes and maximum conditional likelihood estimators expressed in terms of stochastic integrals [see Brockwell et al. (2007b), where this method of estimation is applied to threshold CAR processes observed at closely spaced times, using sums to approximate the stochastic integrals involved].

For general Lévy-driven CARMA processes, estimation procedures which provide information concerning the driving process L are less well-developed. One approach, when $E(L(1)^2) < \infty$, is to estimate the parameters $\{a_j, 1 \leq j \leq p;\ b_j, 0 \leq j < q\}$ by using the Kalman recursions applied to (105) and (106) either to maximize the Gaussian likelihood or to minimize the sum of squares of the one-step prediction errors. If the observations had been made continuously on $[0, T]$, these estimates could be used with Theorem 2 to estimate, for any observed or assumed $\mathbf{X}(0)$, a realization of L on $[0, T]$. The increments of this estimated realization could then be examined and a driving Lévy process chosen whose increments are compatible with the increments of the realization. In practice the CARMA process is not observed continuously but, if the observations are closely-spaced, a discretized version of this procedure can be used as in Brockwell et al. (2007a), where it was applied to subordinator-deriven stationary Ornstein–Uhlenbeck processes. Inference for such processes has also been investigated by Jongbloed et al. (2005). Inference for second-order non-negative subordinator-driven CARMA processes was carried out in Brockwell et al. (2011). For detailed analyses of the recovery of the Lévy increments from a CARMA process observed at discrete times see Brockwell and Schlemm (2013) and Ferrazzano and Fuchs (2013), and for the use of the increments in bootstrapping CAR processes see Brockwell et al. (2014). For stable CARMA processes with exponent $\alpha \in (0, 2)$, $E(L(1)^2) = \infty$, however it was found by Garcia et al. (2011) that estimation of the coefficients of a stable CARMA(2,1) process from observations at regular intervals could be satisfactorily achieved by treating the CARMA process as a second-order processs, estimating the coefficients of the sampled ARMA(2,1) process by standard second-order techniques, and transforming the estimated ARMA coefficents into corresponding CARMA coefficients.

9 Prediction of causal invertible CARMA processes

For regular weakly stationary processes Y with zero mean and zero deterministic component (this class includes all second-order zero-mean CARMA processes) the minimum mean-squared error linear predictor of $Y(h)$, $h > 0$, based on observation of $\{Y(s), s \leq 0\}$ is given by (22) where g is the causal invertible kernel in the Wold-Karhunen representation of Y.

If Y is the L-driven CARMA process defined by (28) and (29) with $EL(1) = 0$, $\mathrm{Var}(L(1)) = \sigma^2 < \infty$, and the zeroes of $a(z)$ and $b(z)$ all in the open left half-plane, then from (32) it has the representation

$$Y(t) = \int_{(-\infty,t]} g_1(t-u)\mathrm{d}L(u), \tag{107}$$

where the kernel g_1 was defined in Theorem 1. Noting, by (51), (52) and (107), that the closed linear span $\mathcal{N}_0$ of $\{Y(s), s \leq 0\}$ is the same as the closed linear span of the increments $\{L(v) - L(u), u < v \leq 0\}$, the minimum mean-squared error predictor of $Y(h)$ in $\mathcal{N}_0$ is the orthogonal projection,

$$P(Y(h)|\mathcal{N}_0) = \int_{(-\infty,0]} g_1(h-u)\mathrm{d}L(u), \tag{108}$$

with mean-squared error

$$\sigma_h^2 = \int_0^h g_1(t)^2 \mathrm{d}t. \tag{109}$$

To express (108) directly in terms of $\{Y(s),\ s \leq 0\}$ rather than $\{L(s),\ s \leq 0\}$ we multiply (29) on the left by $\mathbf{e}_p^T$ to obtain

$$\mathrm{d}L(u) = \mathbf{e}_p^T \left[\mathrm{d}\mathbf{X}(u) - A\mathbf{X}(u)\mathrm{d}u\right],$$

where $\mathbf{X}(t)$ was expressed in terms of Y in (52).

Example 6 CARMA(2,1). For the L-driven CARMA(2, 1) process with $EL(1) = 0$, $\mathrm{Var}(L(1)) = \sigma^2 < \infty$,

$$a(z) = z^2 + a_1 z + a_2 = (z - \lambda_1)(z - \lambda_2),\ \ \Re\lambda_i < 0,\ i = 1, 2,$$

and

$$b(z) = z + b_0,\ \ b_0 > 0,$$

we find, using the preceding arguments, that $P(Y(h)|\mathcal{N}_0),\ h > 0$, is given by (108) with

$$\mathrm{d}L(u) = \mathrm{d}Y(u) + (a_1 - b_0)Y(u)\mathrm{d}u + a(-b_0)\left[\int_{-\infty}^u \mathrm{e}^{-b_0(u-v)}Y(v)\mathrm{d}v\right]\mathrm{d}u.$$

In practice a complete realization of $\{Y(s),\ s \leq 0\}$ is never available, but nevertheless the predictor (108) and its mean squared error (109) provide a useful benchmark against which the performance of other estimators such as $P[Y(h)|Y(n\Delta), n =$

$0, -1, -2, \ldots]$, $\Delta > 0$, can be assessed. Brockwell and Lindner (2014) consider a variety of predictors including the latter and the conditional expectations $E[Y(h)|Y(s), -\infty < s \le 0]$, $E[Y(h)|Y(s), -M \le s \le 0]$ when they exist. In particular they show that, if $E(L(1)^2) < \infty$, the minimum mean-squared error predictor, $E[Y(h)|Y(s), -\infty < s \le 0]$, is the same as the minimum mean-squared error *linear* predictor (108).

10 Non-causal CARMA processes

The kernel function g of a CARMA(p, q) process Y whose autoregressive and moving average zeroes all fall in the open left half of the complex plane has the property that $g(t) = 0$ for $t < 0$. It also has a jump discontinuity at $t = 0$ if and only if $q = p - 1$, and in this case the jump discontinuity is of size one. This reflects the fact that the sample-paths of such a process are continuous if $q < p - 1$ and have the same discontinuities as the driving Lévy process L if $q = p - 1$. The kernel function is the continuous-time analogue of the impulse response function in discrete time in the sense that $g(t)$ is the contribution to $Y(t + T)$ from a jump in the driving process of size one at time T.

It has already been pointed out that the autocovariance structure of the CARMA process is unchanged by reversing the signs of the zeroes of $a(z)$ ad $b(z)$. However the sample-path properties of the original and modified processes will be different because the kernel g will change.

For example the causal CAR(2) process Y with $a(z) = (z - \lambda)^2$, $\lambda < 0$, has the kernel function [from (33)],

$$g(t) = t\mathrm{e}^{-|\lambda|t}\mathbf{1}_{[0,\infty)}(t), \tag{110}$$

and autocorrelation function,

$$\rho(h) = \mathrm{e}^{-|\lambda h|}(1 + |\lambda h|). \tag{111}$$

On the other hand the non-causal CAR(2) process Y_{nc}, with $a(z) = z^2 - \lambda^2$, has the same autocorrelation function but kernel

$$g_{nc}(t) = -\frac{1}{2|\lambda|}\mathrm{e}^{-|\lambda t|}, \quad t \in \mathbb{R},$$

indicating that the sample paths, unlike those of Y, exhibit an instantaneous response to a jump in the driving process since $g_{nc}(0-) = g_{nc}(0+) = -\frac{1}{2|\lambda|} \neq 0$.

Modelling with non-causal CARMA processes is a relatively unexplored area. Schnurr and Woerner (2011) have used the non-causal CAR(2) process $2\lambda Y_{nc}$, with kernel $\mathrm{e}^{-|\lambda t|}$, $t \in \mathbb{R}$, to obtain substantially improved fits to high-frequency financial data than are obtainable using the causal stationary Ornstein–Uhlenbeck process with $a(z) = z - \lambda$ and kernel $\mathrm{e}^{-|\lambda t|}\mathbf{1}_{[0,\infty)}(t)$. They refer to the process $2\lambda Y_{nc}$ as the *well-balanced Ornstein–Uhlenbeck process*. Its autocorrelation function at lag h

is given by (111), as compared with $e^{-|\lambda h|}$ for the corresponding causal stationary Ornstein–Uhlenbeck process.

11 Non-linear CARMA processes

A family of non-linear Gaussian continuous-time autoregressive models which includes CTAR(p) (continuous-time *threshold* autoregressive) processes of order p was discussed in Brockwell (2001a). For closely-spaced data Brockwell et al. (2007b) developed an estimation technique based on the exact likelihood of the continuous time process conditional on the initial state-vector, approximating the stochastic integrals which appear in the maximum likelihood parameter estimators by approximating sums.

12 Continuous-time GARCH processes

Another important class of non-linear models related to CARMA processes is the COGARCH family. The COGARCH(1, 1) process was introduced by Klüppelberg et al. (2004) as a continuous-time analogue of the celebrated discrete-time GARCH(1,1) model for stochastic volatility of Bollerslev (1986).

Given an i.i.d. sequence $(\varepsilon_n)_{n\in\mathbb{N}_0}$ and constants $\beta > 0$, $\lambda_1, \ldots, \lambda_q \geq 0$ and $\delta_1, \ldots, \delta_p \geq 0$ with $q \in \mathbb{N}$ and $p \in \mathbb{N}_0$ and $\lambda_q > 0$, a GARCH(q, p) process $(Y_n)_{n\in\mathbb{N}_0}$ with volatility process $(V_n)_{n\in\mathbb{N}_0}$ is defined by the equations,

$$Y_n = \sqrt{V_n}\,\varepsilon_n, \quad n \in \mathbb{N}_0, \tag{112}$$

$$V_n = \beta + \sum_{i=1}^{q} \lambda_i Y_{n-i}^2 + \sum_{j=1}^{p} \delta_j V_{n-j}, \quad n \geq \max\{p, q\}, \tag{113}$$

with V_n independent of $(\varepsilon_{n+h})_{h\in\mathbb{N}_0}$ and non-negative for every $n \in \mathbb{N}_0$. For $p = 0$ the process is the ARCH(q) process of Engle (1982).

The COGARCH(1,1) process was defined by Klüppelberg et al. (2004) as follows. Given a driving Lévy process $M = (M(t))_{t\geq 0}$ with nonzero Lévy measure, independent of a starting random variable $V(0) \geq 0$, and constants $\beta, \delta > 0$ and $\lambda \geq 0$, the *COGARCH(1,1) process* $(G(t))_{t\geq 0}$ with *volatility process* $(V(t))_{t\geq 0}$ is specified by the equations,

$$G(0) = 0, \quad \mathrm{d}G(t) = \sqrt{V(t-)}\,\mathrm{d}M(t), \quad t \geq 0,$$

where

$$V(t) = \left(\beta \int_0^t e^{\xi(s-)}\,\mathrm{d}s + V(0)\right) e^{-\xi(t)}, \quad t \geq 0,$$

and $\xi = (\xi(t))_{t\geq 0}$ is defined by

$$\xi(t) := -t \log \delta - \sum_{0<s\leq t} \log(1 + \lambda\delta^{-1}(\Delta M_s)^2), \quad t \geq 0,$$

where $\Delta M_s := M(s) - M(s-)$. The volatility process satisfies the stochastic differential equation,

$$\mathrm{d}V(t) = V(t-)\mathrm{d}(t \log \delta + \lambda\delta^{-1}[M, M]_t^{(d)}) + \beta \mathrm{d}t, \quad t \geq 0,$$

where $[M, M]_t^{(d)} = \sum_{0<s\leq t} (\Delta M_s)^2$ denotes the discrete part of the quadratic variation of M. A multivariate extension of the COGARCH(1,1) process has been obtained by Stelzer (2010).

Necessary and sufficient conditions for the existence of a strictly stationary volatility process were obtained by Klüppelberg et al. (2004) and, under certain assumptions, they showed that non-overlapping increments of the corresponding process G are uncorrelated, while the autocorrelation function of $((G_{rh} - G_{r(h-1)})^2)_{h\in\mathbb{N}}$ is that of an ARMA(1,1) process for any $r > 0$. Zero correlation between the increments of G and serial correlation between the squared increments are two of the so-called *stylized facts* of empirical financial time series.

Various techniques for estimation of COGARCH(1,1) processes have been developed. Haug et al. (2007) use a generalized method of moments based on observations $G(0), G(1), G(2), \ldots, G(n)$. They show that their estimators are strongly consistent and under further moment assumptions, asymptotically normal. Other proposed estimation methods include the pseudo-maximum likelihood estimator of Maller et al. (2008), which they use to fit a COGARCH(1,1) model to the ASX200 index of the Australian Stock exchange, and the Markov Chain Monte Carlo estimator of Müller (2010).

The COGARCH(q, p) process of Brockwell et al. (2006) was introduced with the aim of allowing a wider range of autocorrelation structures for the volatility process than permitted by the COGARCH(1,1) model. From (112) and (113) we see that the volatility (V_n) of a GARCH(q, p) process can be regarded as a "self-exciting" ARMA$(p, q-1)$ process driven by $(V_{n-1}\varepsilon_{n-1}^2)$ together with the "mean correction" β. This motivated the definition of the volatility process $(V(t))_{t\geq 0}$ of a continuous-time GARCH(q, p) process as a "self-exciting mean corrected" CARMA$(p, q-1)$ process driven by an appropriate noise term. More precisely, let $M = (M(t))_{t\geq 0}$ be a Lévy process with nonzero Lévy measure. With $p, q \in \mathbb{N}$ such that $q \leq p$, $a_1, \ldots, a_p, b_0, \ldots, b_{p-1} \in \mathbb{R}$, $\beta > 0$, $a_p \neq 0$, $b_{q-1} = 1$ and $b_q = \ldots = b_{p-1} = 0$, define the $p \times p$-matrix $\mathbf{A}$ and the vectors $\mathbf{b}, \mathbf{e}_p \in \mathbb{R}^p$ as in (28) and (29). Define the volatility process $(V(t))_{t\geq 0}$ with parameters $\mathbf{A}$, $\mathbf{b}$, β and driving Lévy process M by

$$V(t) = \beta + \mathbf{b}'\mathbf{X}(t), \quad t \geq 0,$$

where the state process $\mathbf{X} = (\mathbf{X}(t))_{t\geq 0}$ is the unique solution of the stochastic differential equation

$$\begin{aligned}\mathrm{d}\mathbf{X}(t) &= \mathbf{A}\mathbf{X}(t-)\,\mathrm{d}t + \mathbf{e}_p V(t-)\, d[M,M]_t^{(d)}\\ &= \mathbf{A}\mathbf{X}(t-)\,\mathrm{d}t + \mathbf{e}_p(\beta + \mathbf{b}'\mathbf{X}(t-))\, d[M,M]_t^{(d)},\end{aligned}$$

with initial value $\mathbf{X}(0)$, independent of M. If the process $(V(t))_{t\geq 0}$ is non-negative almost surely, then $G = (G(t))_{t\geq 0}$, defined by

$$G(0) = 0, \quad \mathrm{d}G(t) = \sqrt{V(t-)}\,\mathrm{d}M(t),$$

is a *COGARCH* (q, p) *process* with parameters $\mathbf{A}$, $\mathbf{b}$, β and driving Lévy process M.

It can be shown that for $p = q = 1$ this definition is equivalent to the earlier definition of the COGARCH(1,1) process. Brockwell et al. (2006) give sufficient conditions for the existence of a strictly stationary solution $(V_t)_{t\geq 0}$ and its positivity, and show that $(V_t)_{t\geq 0}$ has the same autocorrelation structure as a CARMA$(p, q-1)$ process. Under suitable conditions it is further shown that non-overlapping increments of G are uncorrelated, while their squares are not. More precisely,

$$\mathrm{Cov}((G_t - G_{t-r})^2, (G_{t+h} - G_{t+h-r})^2) = \mathbf{b}'\mathrm{e}^{(A+EM_1^2\mathbf{e}\mathbf{b}')h} H_r, \quad h \geq r > 0,$$

where $H_r \in \mathbb{C}^p$ is independent of h.

13 Conclusions

The theory and applications of CARMA processes have expanded rapidly in the 12 years since the overview given in Brockwell (2001a) and even since the more recent financially oriented review in Brockwell (2009). The present review has attempted to provide the basic theory of Lévy-driven CARMA and related processes, taking into account the results which have appeared in the last few years. These include the basic existence and uniqueness theorem, inference based on recovery of the driving Lévy process from high-frequency data, the use of non-causal models and a number of applications in the study of financial time series and turbulence. To these applications should also be added reference to the CARMA interest rate model (Andresen et al. 2012), the application of stable CARMA processes to futures pricing in electricity markets (Benth et al. 2013) and applications to signal extraction (McElroy 2013). The potential for further applications, with the proliferation of high-frequency data in so many fields, and further theoretical developments, particularly with respect to multivariate models (Marquardt and Stelzer 2007), nonlinear models, non-causal modelling, sampling and embedding remains broad and challenging.

Acknowledgments Much of the recent material reviewed in this article is joint work with Alexander Lindner, to whom I am indebted for many valuable discussions. I am also indebted to two referees for their careful reading of the manuscript and many helpful comments.

References

Andersen, T. G., Benzoni, L. (2009). Realized volatility. In T. G. Andersen, R. A. Davis, J.-P. Krei, T. H. Mikosch (Eds.), *Handbook of Financial Time Series* (pp. 555–575). Berlin: Springer.

Andersen, T.G., Bollerslev, T., Diebold, F. X., Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*, 579–625.

Andresen, A., Benth, F.E., Koekebakker, S., Zakamulin, V. (2012). The CARMA interest rate model. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1138632

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge: Cambridge University Press.

Arato, M. (1982). *Linear Stochastic Systems with Constant Coefficients*. Berlin: Springer.

Barndorff-Nielsen, O. E., Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society Series B*, *63*, 167–241.

Barndorff-Nielsen, O. E., Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B*, *64*, 253–280.

Barndorff-Nielsen, O.E., Shephard, N. (2003). Integrated OU processes and non-Gaussian OU-based stochastic volatility models. *Scandinavian Journal of Statistics*, *30*, 277–295.

Bartlett, M.S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society*, *7*, 27–41 (Supplement).

Benth, F. E., Klüppelberg, C., Müller, G., Vos, L. (2013). Futures pricing in electricity markets based on stable CARMA spot models. http://arxiv.org/pdf/1201.1151

Beran, J. (1994). *Statistics for Long Memory Processes. vol. 61 of Monographs on Statistics and Applied Probability*. New York: Chapman and Hall.

Bergstrom, A. R. (1985). The estimation of parameters in non-stationary higher-order continuous-time dynamic models. *Econometric Theory*, *1*, 369–385.

Bergstrom, A. R. (1990). *Continuous Time Econometric Modelling*. Oxford: Oxford University Press.

Bertoin, J. (1996). *Lévy processes*. Cambridge: Cambridge University Press.

Bloomfield, P. (2000). *Fourier Analysis of Time Series, an Introduction* (2nd ed.). New York: Wiley.

Bollerslev, T. (1986). Generalised autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *51*, 307–327.

Brockwell, A. E., Brockwell, P.J. (1999). A class of non-embeddable ARMA processes. *Journal of Time Series Analysis*, *20*, 483–486.

Brockwell, P. J. (2001a). Continuous-time ARMA Processes. In D.N. Shanbhag, C.R. Rao (Eds.), *Handbook of Statistics 19*; *Stochastic Processes*: *Theory and Methods* (pp. 249–276). Amsterdam: Elsevier.

Brockwell, P. J. (2001b). Levy-driven CARMA processes. *Annals of the Institute of Statistical Mathematics*, *53*, 113–124.

Brockwell, P. J. (2004). Representations of continuous-time ARMA processes. *Journal of Applied Probability*, *41A*, 375–382.

Brockwell, P. J. (2009). Lévy-driven continuous-time ARMA processes. In T.G. Andersen, R.A. Davis, J.-P. Krei, Th Mikosch (Eds.), *Handbook of Financial Time Series* (pp. 457–479). Berlin: Springer.

Brockwell, P. J., Davis, R.A. (1991). *Time Series*: *Theory and Methods* (2nd ed.). New York: Springer.

Brockwell, P. J., Davis, R.A. (2001). Discussion of Levy-driven Ornstein–Uhlenbeck processes and some of their applications in financial economics by O. Barndorff-Nielsen and N. Shephard. *Journal of the Royal Statistical Society Series B*, *63*, 218–219.

Brockwell, P. J., Hannig, J. (2010). CARMA(p, q) generalized random processes. *Journal of Statistical Planning and Inference*, *140*, 3613–3618.

Brockwell, P. J., Lindner, A. (2009). Existence and uniqueness of stationary Lévy-driven CARMA processes. *Stochastic Processes and their Applications*, *119*, 2660–2681.

Brockwell, P. J., Lindner, A. (2010). Strictly stationary solutions of autoregressive moving average equations. *Biometrika*, *97*, 765–772.

Brockwell, P. J., Lindner, A. (2013). Integration of CARMA processes and spot volatility modelling. *Journal of Time Series Analysis*, *34*, 156–167.

Brockwell, P. J., Lindner, A. (2014). Prediction of stationary Lévy-driven CARMA processes. *Journal of Econometrics* (to appear)

Brockwell, P. J., Marquardt, T. (2005). Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statistica Sinica*, *15*, 477–494.

Brockwell, P. J., Schlemm, E. (2013). Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete obsservations. *Journal of Multivariate Analysis*, *115*, 217–251.
Brockwell, P. J., Chadraa, E., Lindner, A. (2006). Continuous-time GARCH processes. *Annals of Applied Probabiity*, *16*, 790–826.
Brockwell, P. J., Davis, R. A., Yang, Y. (2007a). Estimation for non-negative Lévy-driven Ornstein–Uhlenbeck processes. *Journal of Applied Probability*, *44*, 977–989.
Brockwell, P. J., Davis, R. A., Yang, Y. (2007b). Continuous-time Gaussian autoregression. *Statistica Sinica*, *17*, 63–80.
Brockwell, P. J., Davis, R. A., Yang, Y. (2011). Estimation for non-negative Lévy-driven CARMA processes. *Journal of Business and Economic Statistics*, *29*, 250–259.
Brockwell, P. J., Ferrazzano, V., Kluppelberg, C. (2013). High frequency sampling and kernel estimation for continuous-time moving average processes. *Journal of Time Series Analysis*, *34*, 385–404.
Brockwell, P. J., Kreiss, J. -P., Niebuhr, T. (2014). Bootstrapping continuous-time autoregressive processes. *Annals of the Institute of Statistical Mathematics*, *66*, 75–92.
Comte, F., Renault, E. (1996). Long memory continuous time models. *Journal of Econometrics*, *73*, 101–149.
Doob, J.L. (1944). The elementary Gaussian processes. *Annals of Mathematical Statistics*, *25*, 229–282.
Doob, J.L. (1953). *Stochastic Processes*. New York: Wiley.
Durbin, J. (1961). Efficient fitting of linear models for continuous stationary time series from discrete data. *Bulletin of the International Statistical Institute*, *38*, 273–281.
Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*, 987–1008.
Fasen, V., Fuchs, F. (2013). Spectral estimates for high-frequency sampled continuous-time autoregressive moving average processes. *Journal of Time Series Analysis*, *34*, 532–551.
Ferrazzano, V., Fuchs, F. (2013). Noise recovery for Lvy-driven CARMA processes and high-frequency behaviour of approximating Riemann sums. *Electronic Journal of Statistics*, *7*, 533–561.
Fowler, R.H. (1936). *Statistical Mechanics*. Cambridge: Cambridge University Press.
Garcia, I., Klüppelberg, C., Müller, G. (2011). Estimation of stable CARMA models with an application to electricity spot prices. *Statistical Modelling*, *11*, 447–470.
Granger, C.W.J., Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Anaysis*, *1*, 15–29.
Haug, S., Klüppelberg, C., Lindner, A., Zapp, M. (2007). Method of moment estimation in the COGARCH(1,1) model. *The Econometrics Journal*, *10*, 320–341.
Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, *68*, 165–176.
Hyndman, R.J. (1993). Yule-Walker estimates for continuous-time autoregressive models. *Journal of Time Series Analysis*, *14*, 281–296.
Jones, R.H. (1981). Fitting a continuous time autoregression to discrete data. In D.F. Findley (Ed.), *Applied Time Series Analysis II* (pp. 651–682). New York: Academic Press.
Jones, R.H. (1985). Time series analysis with unequally spaced data. In E.J. Hannan, P.R. Krishnaiah, M.M. Rao (Eds.), *Time Series in the Time Domain, Handbook of Statistics 5* (pp. 157–178). Amsterdam: North Holland.
Jongbloed, G., van der Meulen, F. H., van der Waart, A.W. (2005). Non-parametric inference for Lévy-driven Ornstein–Uhlenbeck processes. *Bernoulli*, *11*, 759–791.
Karhunen, K. (1950). Über die Struktur stationärer zufälliger Funktionen. *Arkiv för Matematik*, *1*, 141–160.
Klüppelberg, C., Lindner, A., Maller, R. (2004). Stationarity and second order behaviour of discrete and continuous-time GARCH(1,1) processes. *Journal of Applied Probability*, *41*, 601–622.
Kokoszka, P. S., Taqqu, M.S. (1995). Fractional ARIMA with stable innovations. *Stochastic Processes and their Applications*, *60*, 19–47.
Kyprianou, A. E. (2006). *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Berlin: Springer.
Maller, R. A., Müller, G., Szimayer, A. (2008). GARCH modelling in continuous-time for irregularly spaced time series data. *Bernoulli*, *14*, 519–542.
Marquardt, T. (2006). Fractional Lévy processes with an application to long memory moving average processes. *Bernoulli*, *12*, 1099–1126.
Marquardt, T., Stelzer, R. (2007). Multivariate CARMA processes. *Stochastic Processes and their Applications*, *117*, 96–120.

McElroy, T. (2013). Forecasting continuous-time processes with applications to signal extraction. *Annals of the Institute of Statistical Mathematics*, *65*, 439–456.

Müller, G. (2010). MCMC estimation of the COGARCH(1,1) model. *Journal of Financial Econometrics*, *8*, 481–510.

Mykland, P. A., Zhang, L. (2012). The econometrics of high-frequency data. In M. Kessler, A. Lindner, M. Srensen (Eds.), *Statistical Methods for Stochastic Differential Equations* (pp. 109–190). Boca Raton: Chapman and Hall/CRC.

Pham-Dinh, T. (1977). Estimation of parameters of a continuous time Gaussian stationary process with rational spectral density. *Biometrika*, *64*, 385–389.

Phillips, A.W. (1959). The estimation of parameters in systems of stochastic differential equations. *Biometrika*, *46*, 67–76.

Protter, P. (2004). *Stochastic Integration and Differential Equations* (2nd ed.). New York: Springer.

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

Sato, K., Yamazato, M. (1984). Operator-selfdecomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stochastic Processes and their Applications*, *17*, 73–100.

Schnurr, A., Woerner, J.H.C. (2011). Well-balanced Lévy-driven Ornstein–Uhlenbeck processes. *Statistics and Risk Modelling*, *28*, 343–357.

Stelzer, R. (2010). Multivariate COGARCH(1,1) processes. *Bernoulli*, *16*, 80–115.

Thornton, M. A., Chambers, M.J. (2013). Continuous-time autoregressive moving average processes in discrete time: representation and embeddability. *Journal of Time Series Analysis*, *34*, 552–561.

Todorov, V., Tauchen, G. (2006). Simulation methods for Lévy-driven CARMA stochastic volatility models. *Journal of Business and Economic Statistics*, *24*, 455–469.

Tsai, H., Chan, K. S. (2005). A note on non-negative continuous-time processes. *Journal of the Royal Statistical Society Series B*, *67*, 589–597.

Vollenbröker, B. (2012). Strictly stationary solutions of ARMA equations with fractional noise. *Journal of Time Series Analysis*, *33*, 570–582.

Wolfe, S.J. (1982). On a continuous analog of the stochastic difference equation $x_n = \rho x_{n-1} + b_n$. *Stochastic Processes and their Applications*, *12*, 301–312.

Yaglom, I.M. (1987). *Correlation Theory of Stationary Random Processes*. New York: Springer.