High frequency sampling of a continuous-time ARMA process

Peter J. Brockwell [*] Vincenzo Ferrazzano [†] Claudia Klüppelberg[‡]

May 2011

Abstract

Continuous-time autoregressive moving average (CARMA) processes have recently been used widely in the modeling of non-uniformly spaced data and as a tool for dealing with high-frequency data of the form $Y_{n\Delta}, n = 0, 1, 2, \ldots$, where Δ is small and positive. Such data occur in many fields of application, particularly in finance and the study of turbulence. This paper is concerned with the characteristics of the process $(Y_{n\Delta})_{n\in\mathbb{Z}}$, when Δ is small and the underlying continuous-time process $(Y_t)_{t\in\mathbb{R}}$ is a specified CARMA process.

AMS 2000 Subject Classifications: 60G51, 62M10.

Keywords: CARMA process, high frequency data, discretely sampled process

1 Introduction

Throughout this paper we shall be concerned with a CARMA process driven by a second-order zero-mean Lévy process L with $EL_1 = 0$ and $EL_1^2 = \sigma^2$. The process is defined as follows.

For non-negative integers p and q such that $q < p$, a CARMA(p, q) process $Y = (Y_t)_{t\in\mathbb{R}}$, with coefficients $a_1, \ldots, a_p$, $b_0, \ldots, b_q \in \mathbb{R}$, and driving Lévy process L, is defined to be a strictly stationary solution of the suitably interpreted formal equation,

$$a(D)Y_t = b(D)DL_t, \quad t \in \mathbb{R}, \tag{1.1}$$

where D denotes differentiation with respect to t, $a(\cdot)$ and $b(\cdot)$ are the polynomials,

$$a(z) := z^p + a_1 z^{p-1} + \cdots + a_p \quad \text{and} \quad b(z) := b_0 + b_1 z + \cdots + b_{p-1} z^{p-1},$$

and the coefficients b_j satisfy $b_q = 1$ and $b_j = 0$ for $q < j < p$. The polynomials $a(\cdot)$ and $b(\cdot)$ are assumed to have no common zeroes, and it will be assumed that the zeroes of the polynomial a all lie in the interior of the left half of the complex plane.

[*]Departments of Statistics, Colorado State University, Fort Collins, CO 80527, and Columbia University, New York, NY 10027, USA, email: pb2141@columbia.edu

[†]Center for Mathematical Sciences, Technische Universität München, 85748 Garching b. München, Germany, email: ferrazzano@ma.tum.de, http://www-m4.ma.tum.de/pers/ferrazzano/

[‡]Center for Mathematical Sciences, and Institute for Advanced Study, Technische Universität München, 85748 Garching b. München, Germany, email: cklu@ma.tum.de, http://www-m4.ma.tum.de

Since the derivative DL_t does not exist in the usual sense, we interpret (1.1) as being equivalent to the observation and state equations

$$Y_t = \mathbf{b}^T \mathbf{X}_t \,, \tag{1.2}$$

$$d\mathbf{X}_t = A\mathbf{X}_t dt + \mathbf{e}_p dL_t \,, \tag{1.3}$$

where

$$\mathbf{X}_t = \begin{pmatrix} X(t) \\ X^{(1)}(t) \\ \vdots \\ X^{(p-2)}(t) \\ X^{(p-1)}(t) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}, \quad \mathbf{e}_p = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \dots & -a_1 \end{pmatrix} \quad \text{and } A = -a_1 \text{ for } p = 1.$$

It is easy to check that the eigenvalues of the matrix A, which we shall denote by $\lambda_1, ..., \lambda_p$, are the same as the zeroes of the autoregressive polynomial $a(\cdot)$.

Under the conditions specified it has been shown (Brockwell and Lindner (2009), Lemma 2.3) that these equations have the unique strictly stationary solution,

$$Y_t = \int_{-\infty}^{\infty} g(t-u) dL_u, \tag{1.4}$$

where

$$g(t) = \begin{cases} \dfrac{1}{2\pi i} \displaystyle\int_\rho \frac{b(z)}{a(z)} e^{tz} dz = \sum_\lambda Res_{z=\lambda} \left(e^{zt} \frac{b(z)}{a(z)} \right), & \text{if } t > 0, \\ 0, & \text{if } t \le 0. \end{cases} \tag{1.5}$$

and ρ is any simple closed curve in the open left half of the complex plane encircling the zeroes of $a(\cdot)$. The sum is over the distinct zeroes λ of $a(\cdot)$ and $Res_{z=\lambda}(\cdot)$ denotes the residue at λ of the function in parentheses. Evaluating these residues, we can write g more explicitly as

$$g(t) = \sum_\lambda \frac{1}{(m(\lambda)-1)!} \left[D_z^{m(\lambda)-1} \left((z-\lambda)^{m(\lambda)} e^{zt} b(z)/a(z) \right) \right]_{z=\lambda} \mathbf{1}_{(0,\infty)}(t), \tag{1.6}$$

where $m(\lambda)$ denotes the multiplicity of the zero λ and D_z denotes differentiation with respect to z. The kernel g can also be expressed (Brockwell and Lindner (2009), equations (2.10) and (3.7)) as

$$g(t) = \mathbf{b}^\top e^{At} \mathbf{e}_p \mathbf{1}_{(0,\infty)}(t). \tag{1.7}$$

From this equation we see at once that g is infinitely differentiable on $(0, \infty)$ with k^{th} derivative,

$$g^{(k)}(t) = \mathbf{b}^\top e^{At} A^k \mathbf{e}_p, \ 0 < t < \infty.$$

Since $b_q = 1$ and $b_j = 0$ for $j > q$, the right derivatives $g^{(k)}(0+)$ satisfy

$$g^{(k)}(0+) = \mathbf{b}^\top A^k \mathbf{e}_p = \begin{cases} 0 & \text{if } k < p-q-1, \\ 1 & \text{if } k = p-q-1, \end{cases} \tag{1.8}$$

and in particular $g(0+) = 1$ if $p - q = 1$ and $g(0+) = 0$ if $p - q > 1$.

Gaussian CARMA processes, of which the Gaussian Ornstein-Uhlenbeck process is an early example, were first studied in detail by Doob (1944) (see also Doob (1990)). The state-space formulation, (1.2) and (1.3) (with $\mathbf{b}^\top = [1\ 0\ \cdots\ 0]$) was used by Jones (1981) to carry out inference for time series with irregularly-spaced observations. This formulation leads naturally to the definition of Lévy-driven and non-linear CARMA processes (see Brockwell (2001) and the references therein). Fractionally integrated Lévy-driven CARMA processes were studied by Brockwell and Marquardt (2005).

Lévy-driven CARMA processes have been applied successfully to the modelling of stochastic volatility in finance (see Todorov and Tauchen (2006), Brockwell et al. (2006) and Haug and Czado (2007)), extending the celebrated Ornstein-Uhlenbeck model of Barndorff-Nielsen and Shephard (2001). The results presented here were motivated by preliminary studies of high-frequency turbulence data (see Ferrazzano (2010)) which appear to be well-fitted by a continuous time moving average process sampled at times $0, \Delta, 2\Delta, \ldots$, where Δ is small and positive. We return on this topic in Brockwell, Ferrazzano and Klüppelberg (2011). The application to turbulence data will be investigated in detail in Ferrazzano and Klüppelberg (2011).

Our paper is organised as follows. In Section 2 we derive an expression for the spectral density of the sampled sequence $Y^\Delta := (Y_{n\Delta})_{n\in\mathbb{Z}}$. It is known that the filtered process $(\phi(B)Y_n^\Delta)_{n\in\mathbb{Z}}$, where $\phi(B)$ is the filter defined in (3.1), is a moving average of order at most $p-1$. In Section 3, we determine the asymptotic behaviour of the spectral density and autocovariance function of $(\phi(B)Y_n^\Delta)_{n\in\mathbb{Z}}$ as $\Delta \downarrow 0$ and the asymptotic moving average coefficients and white noise variance in the cases $p - q = 1, 2$ and 3. In general we show that for small enough Δ the order of the moving average $(\phi(B)Y_n^\Delta)_{n\in\mathbb{Z}}$ is $p - 1$.

2 The spectral density of $Y^\Delta := (Y_{n\Delta})_{n\in\mathbb{Z}}$

From (1.5) we immediately see, since $g(t) = 0$ for $t < 0$, that the Fourier transform of g is

$$\tilde{g}(\omega) := \int_{\mathbb{R}} g(t)e^{i\omega t}dt = -\frac{1}{2\pi i}\int_\rho \frac{b(z)}{a(z)}\frac{1}{z + i\omega}dz = \frac{b(-i\omega)}{a(-i\omega)}, \quad \omega \in \mathbb{R}. \tag{2.1}$$

Since the autocovariance function $\gamma_Y(\cdot)$ is the convolution of $\sigma g(\cdot)$ and $\sigma g(-\cdot)$, its Fourier transform is given by

$$\tilde{\gamma}_Y(\omega) = \sigma^2\tilde{g}(\omega)\tilde{g}(-\omega) = \sigma^2\left|\frac{b(i\omega)}{a(i\omega)}\right|^2, \quad \omega \in \mathbb{R}.$$

The spectral density of Y is the inverse Fourier transform of γ_Y. Thus

$$f_Y(\omega) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\omega h}\gamma_Y(h)dh = \frac{1}{2\pi}\tilde{\gamma}_Y(-\omega) = \frac{\sigma^2}{2\pi}\left|\frac{b(i\omega)}{a(i\omega)}\right|^2, \quad \omega \in \mathbb{R}.$$

Substituting this expression into the relation

$$\gamma_Y(h) = \int_{\mathbb{R}} e^{i\omega h}f_Y(\omega)d\omega, \quad h \in \mathbb{R},$$

and changing the variable of integration from ω to $z = i\omega$ gives,

$$\gamma_Y(h) = \frac{\sigma^2}{2\pi i}\int_\rho \frac{b(z)b(-z)}{a(z)a(-z)}e^{|h|z}dz = \sigma^2\sum_\lambda Res_{z=\lambda}\left(\frac{b(z)b(-z)}{a(z)a(-z)}e^{z|h|}\right), \tag{2.2}$$

where the sum is again over the distinct zeroes of $a(\cdot)$.

We can now compute the spectral density of the sampled sequence $Y^\Delta := (Y_{n\Delta})_{n\in\mathbb{Z}}$. This spectral density f_Δ will play a key role in the subsequent analysis. We have, from Corollary 4.3.2 in Brockwell and Davis (1991),

$$f_\Delta(\omega) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma_Y(h\Delta)e^{-ih\omega}, \quad -\pi \le \omega \le \pi,$$

and, substituting for γ_Y from (2.2),

$$f_\Delta(\omega) = \frac{-\sigma^2}{4\pi^2 i}\int_\rho \frac{b(z)b(-z)}{a(z)a(-z)}\frac{\sinh(\Delta z)}{\cosh(\Delta z)-\cos(\omega)}dz, \quad -\pi \le \omega \le \pi. \tag{2.3}$$

3 The filtered sequence, $(\phi(B)Y_n^\Delta)_{n\in\mathbb{Z}}$

If $\lambda_1,\ldots,\lambda_p$ are the (not necessarily distinct) zeroes of $a(\cdot)$, then we know from Brockwell and Lindner (2009), Lemma 2.1, that if we apply the filter

$$\phi(B) := \prod_{j=1}^{p}(1-e^{\lambda_j\Delta}B) \tag{3.1}$$

to the sampled sequence, Y^Δ, we obtain a strictly stationary sequence which is $(p-1)$-correlated and is hence, by Lemma 3.2.1 of Brockwell and Davis (1991), a moving average process of order $p-1$ or less.

Our goal in this section is to study the asymptotic properties, as $\Delta \downarrow 0$, of the moving average $\theta(B)Z_n$ in the ARMA representation,

$$\phi(B)Y_n^\Delta = \theta(B)Z_n, \quad n\in\mathbb{Z}, \tag{3.2}$$

of the high-frequency sequence Y^Δ. Here B denotes the backward shift operator and $(Z_n)_{n\in\mathbb{Z}}$ is an uncorrelated sequence of zero-mean random variables with constant variance which we shall denote by τ^2.

We shall denote by f_{MA} the spectral density of $(\theta(B)Z_n)_{n\in\mathbb{Z}}$. Then, observing that the power transfer function of the filter (3.1) is

$$\psi(\omega) = |\prod_{j=1}^{p}(1-e^{\lambda_j\Delta+i\omega})|^2 = 2^p e^{-a_1\Delta}\prod_{i=1}^{p}(\cosh(\lambda_i\Delta)-\cos(\omega)), \quad -\pi\le\omega\le\pi, \tag{3.3}$$

we have

$$f_{MA}(\omega) = \psi(\omega)f_\Delta(\omega), \quad -\pi\le\omega\le\pi, \tag{3.4}$$

where $\psi(\omega)$ and $f_\Delta(\omega)$ are given by (3.3) and (2.3) respectively.

In principle the expression (3.4) determines the second order properties of $(\theta(B)Z_n)_{n\in\mathbb{Z}}$ and in particular the autocovariances $\gamma_{MA}(h)$ for $h=0,\ldots,p-1$. Ideally we would like to use these autocovariances to find the coefficients $\theta_1,\ldots,\theta_{p-1}$ and white noise variance τ^2, all of which are uniquely determined by the autocovariances, if we impose the condition that $\theta(\cdot)$ has no zeros in the interior of the unit circle. Determination of these quantities is equivalent to finding the corresponding factorization of the spectral density f_{MA} (see Sayed and Kailath (2001) for a recent paper on spectral factorization).

From (2.3), (3.3) and (3.4) we can calculate the spectral density $f_{MA}(\omega)$ as $-\sigma^2\psi(\omega)/(2\pi)$ times the sum of the residues in the left half plane of the integrand in (2.3) , i.e.

$$f_{MA}(\omega) = -\frac{\sigma^2}{2\pi}\psi(\omega)\sum_{\lambda} D_z^{m(\lambda)-1}\left(\frac{\sinh(\Delta z)b(z)b(-z)}{(\cosh(\lambda\Delta)-\cos(\omega))a(-z)\prod_{\mu\neq\lambda}(z-\mu)^{m(\mu)}}\right)_{z=\lambda}, \quad (3.5)$$

where the sum is over the distinct zeroes λ of $a(\cdot)$ and the product in the denominator is over the distinct zeroes μ of $a(\cdot)$, which are different from λ. The multiplicities of the zeroes λ and μ are denoted by $m(\lambda)$ and $m(\mu)$ respectively. When the zeroes $\lambda_1,\dots,\lambda_p$ each have multiplicity 1, the expression for $f_{MA}(\omega)$ simplifies to

$$f_{MA}(\omega) = \frac{(-2)^p e^{-a_1\Delta}\sigma^2}{2\pi}\sum_{i=1}^{p}\frac{b(\lambda_i)b(-\lambda_i)}{a'(\lambda_i)a(-\lambda_i)}\sinh(\lambda_i\Delta)\prod_{j\neq i}(\cos\omega-\cosh(\lambda_j\Delta)), \quad -\pi\le\omega\le\pi.$$

Although in principle the corresponding autocovariances $\gamma_{MA}(j)$ could be derived from f_{MA}, we derive a more direct explicit expression later as Proposition 3.6. The asymptotic behaviour of f_{MA} as $\Delta\downarrow 0$ is derived in the following theorem by expanding (2.3) in powers of Δ and evaluating the corresponding coefficients. Here and in all that follows we shall use the notation, $a(\Delta)\sim b(\Delta)$, to mean that $\lim_{\Delta\downarrow 0} a(\Delta)/b(\Delta) = 1$.

Theorem 3.1. *The spectral density f_{MA} of $(\theta(B)Z_n)_{n\in\mathbb{Z}}$ in the ARMA representation* (3.2) *of the sampled process Y^Δ has the asymptotic form, as $\Delta\downarrow 0$,*

$$f_{MA}(\omega)\sim\frac{\sigma^2}{2\pi}(-1)^{p-q-1}\Delta^{2(p-q)-1}c_{p-q-1}(\omega)2^{p-1}(1-\cos\omega)^p, \quad -\pi\le\omega\le\pi, \quad (3.6)$$

where $c_k(\omega)$ is the coefficient of x^{2k+1} in the power series expansion

$$\frac{\sinh x}{\cosh x-\cos\omega}=\sum_{k=0}^{\infty}c_k(\omega)x^{2k+1}. \quad (3.7)$$

In particular, $c_0(\omega)=\frac{1}{1-\cos\omega}$, $c_1(\omega)=-\frac{2+\cos\omega}{6(1-\cos\omega)^2}$, $c_2(\omega)=\frac{33+26\cos\omega+\cos(2\omega)}{240(1-\cos\omega)^3}$, $\dots$.

Proof. The integrand in (2.3) can be expanded as a power series in Δ using (3.7). The integral can then be evaluated term by term using the identities, (see Example 3.1.2.3. of Mitrinović and Kečkić (1984))

$$\frac{1}{2\pi i}\int_\rho z^{2k+1}\frac{b(z)b(-z)}{a(z)a(-z)}dz = -\frac{1}{2}Res_{z=\infty}\left(\frac{z^{2k+1}b(z)b(-z)}{a(z)a(-z)}\right), \quad k\in\{0,1,2,\dots\},$$

from which we obtain, in particular,

$$\frac{1}{2\pi i}\int_\rho z^{2k+1}\frac{b(z)b(-z)}{a(z)a(-z)}dz = \begin{cases}0 & \text{if } 0\le k<p-q-1,\\ \frac{(-1)^{p-q}}{2} & \text{if } k=p-q-1.\end{cases}$$

Substituting the resulting expansion of the integral (2.3) and the asymptotic expression $\psi(\omega)\sim 2^p(1-\cos\omega)^p$ into (3.4) and retaining only the dominant power of Δ as $\Delta\to 0$, we arrive at (3.6). $\square$

Corollary 3.2. *The following special cases are of particular interest.*

$$p - q = 1: \quad f_{MA}(\omega) \sim \frac{\sigma^2\Delta}{2\pi} 2^q(1-\cos\omega)^q. \tag{3.8}$$

$$p - q = 2: \quad f_{MA}(\omega) \sim \frac{\sigma^2\Delta^3}{2\pi}\left(\frac{2}{3} + \frac{\cos\omega}{3}\right) 2^q(1-\cos\omega)^q. \tag{3.9}$$

$$p - q = 3: \quad f_{MA}(\omega) \sim \frac{\sigma^2\Delta^5}{2\pi}\left(\frac{11}{20} + \frac{13\cos\omega}{30} + \frac{\cos(2\omega)}{60}\right) 2^q(1-\cos\omega)^q. \tag{3.10}$$

Proof. These expressions are obtained from (3.6) using the values of $c_0(\omega), c_1(\omega)$ and $c_2(\omega)$ given in the statement of the theorem. □

Remark 3.3. (i) The right-hand side of (3.8) is the spectral density of a q-times differenced white noise with variance $\sigma^2\Delta$. It follows that, if $q = p-1$, then the moving average polynomial $\theta(B)$ in (3.2) is asymptotically $(1-B)^q$ and the white noise variance τ^2 is asymptotically $\sigma^2\Delta$ as $\Delta \to 0$. This result is stated with the corresponding results for $p-q=2$ and $p-q=3$ in the following corollary.

(ii) By Proposition 3.32 of Marquardt and Stelzer (2007) a CARMA(p,q)-process has sample paths which are $(p-q-1)$-times differentiable. Consequently to represent processes with non-differentiable sample-paths it is necessary to restrict attention to the case $p-q=1$. It is widely believed that sample-paths with more than two derivatives are too smooth to represent the processes observed empirically in finance and turbulence (see e.g. Jacod and Todorov (2010), Jacod et al. (2010)) so we are not concerned with the cases when $p-q>3$. □

Corollary 3.4. *The moving average process $X_n := \theta(B)Z_n$ in (3.2) has for $\Delta \downarrow 0$ the following asymptotic form.*

(a) If $p-q=1$, then

$$X_n = (1-B)^q Z_n, \quad n \in \mathbb{Z},$$

where $\tau^2 := Var(Z_n) = \sigma^2\Delta$.

(b) If $p-q=2$, then

$$X_n = (1+\theta B)(1-B)^q Z_n, \quad n \in \mathbb{Z},$$

where $\theta = 2-\sqrt{3}$ and $\tau^2 := Var(Z_n) = \sigma^2\Delta^3(2+\sqrt{3})/6$.

(c) If $p-q=3$, then

$$X_n = (1+\theta_1 B + \theta_2 B^2)(1-B)^q Z_n, \quad n \in \mathbb{Z},$$

where $\theta_2 = 2\left(8+\sqrt{30}\right) - \sqrt{375+64\sqrt{30}}$, $\theta_1 = 26\theta_2/(1+\theta_2) = 13 - \sqrt{135+4\sqrt{30}}$ and $\tau^2 = \left(2\left(8+\sqrt{30}\right) + \sqrt{375+64\sqrt{30}}\right)\Delta^5\sigma^2/120$.

Proof. (a) follows immediately from Theorem 4.4.2 of Brockwell and Davis (1991).

To establish (b) we observe from (3.9) that the required moving average is the q times differenced MA(1) process with autocovariances at lags zero and one, $\gamma(0) = 2\sigma^2\Delta^3/3$ and $\gamma(1) = \sigma^2\Delta^3/6$. Expressing these covariances in terms of θ and τ^2 gives the equations,

$$(1+\theta^2)\tau^2 = 2\sigma^2\Delta^3/3,$$

$$\theta\tau^2 = \sigma^2\Delta^3/6,$$

from which we obtain a quadratic equation for θ. Choosing the unique solution which makes the MA(1) process invertible gives the required result.

The proof of (c) is analogous. The corresponding argument yields a quartic equation for θ_2. The particular solution given in the statement of (b) is the one which satisfies the condition that $\theta(z)$ is nonzero for all complex z such that $|z| < 1$. $\square$

Although the absence of the moving-average coefficients, b_j, from Corollary 3.4 suggests that they cannot be estimated from very closely-spaced observations, the coefficients do appear if the expansions are taken to higher order in Δ. The apparent weak dependence of the sampled sequence on the moving-average coefficients as $\Delta \downarrow 0$ is compensated by the increasing number of available observations.

In principle the autocovariance function γ_{MA} can be calculated, as indicated earlier, from the corresponding spectral density f_{MA} given by (3.4) and (2.1). Below we derive a more direct representation of γ_{MA} and use it to prove Theorem 3.7, which is the time-domain analogue of Theorem 3.1.

Define $B_\Delta g(t) = g(t-\Delta)$ for $t \in \mathbb{R}$. We show that $\phi(B_\Delta)g(\cdot) \equiv 0$ for $t > p\Delta$.

Lemma 3.5. *Let Y be the* CARMA(p,q) *process* (1.4) *and* $\Delta > 0$. *Define* $\phi(B)$ *as in* (3.1). *Then*

$$\phi(B_\Delta)g(t) := \prod_{j=1}^{p}(1 - e^{\lambda_j\Delta}B_\Delta)g(t) = 0, \quad t > p\Delta. \tag{3.11}$$

Proof. Rewriting the product in (3.11) as a sum we find $\phi(B_\Delta)g(t) = \sum_{j=0}^{p} A_j^p g(t - j\Delta)$, which has Fourier transform (invoking the shift property and the right hand side of (2.1))

$$\prod_{\lambda}(1 - e^{\Delta(\lambda+i\omega)})^{m(\lambda)}\frac{b(-i\omega)}{a(-i\omega)}, \quad \omega \in \mathbb{R},$$

where the product is taken over the distinct zeroes of $a(\cdot)$ having multiplicity $m(\lambda)$. Using the fact that the product of Fourier transforms corresponds to the convolution of functions, we obtain from (1.5)

$$\phi(B_\Delta)g(t) = -\frac{1}{2\pi i}\int_\rho \prod_\lambda (1-e^{\Delta(\lambda-z)})^{m(\lambda)}\frac{b(z)}{a(z)}e^{tz}dz = -\sum_\lambda Res_{z=\lambda}\left(e^{zt}b(z)\prod_\lambda \frac{(1-e^{\Delta(\lambda-z)})^{m(\lambda)}}{(z-\lambda)^{m(\lambda)}}\right).$$

Now note that, for every of the distinct zeroes λ_j,

$$\lim_{z\to\lambda_i}\frac{(1-e^{\Delta(\lambda_j-z)})^{m(\lambda_j)}}{(z-\lambda_j)^{m(\lambda_j)}} = \Delta^{m(\lambda_j)}.$$

The singularities at $z = \lambda_j$ are removable and, therefore, using Cauchy's residue theorem, Theorem 1 of Section 3.1.1, p. 25, and Theorem 2 of Section 2.1.2, p. 7, of Mitrinović and Kečkić (1984), the filtered kernel is zero for every $t \in \mathbb{R}$. $\square$

Proposition 3.6. *Let Y be the* CARMA(p,q) *process* (1.4) *and $\Delta > 0$. The autocovariance at lag n of $(\phi(B)Y_j^{\Delta})_{j\in\mathbb{Z}}$ is, for $n = 0, 1, \ldots, p-1$,*

$$\gamma_{MA}(n) = \sigma^2 \sum_{i=1}^{p-n} \sum_{k=0}^{n+i-1} \sum_{h=0}^{i-1} A_k^p A_h^p \int_{(i-1)\Delta}^{i\Delta} g(s - h\Delta) g(s - (k-n)\Delta) ds, \qquad (3.12)$$

with

$$A_k^p = (-1)^k \sum_{\{i_1,\ldots,i_k\}\in C_k^p} e^{\Delta(\lambda_{i_1}+\cdots+\lambda_{i_k})}, \quad k = 1, \ldots, p. \qquad (3.13)$$

The sum in (3.13) *is taken over the $\binom{p}{k}$ subsets of size k of $\{1, 2, \ldots, p\}$.*

Proof. We note that $\gamma_{MA}(n)$ is the same as $\mathbb{E}[(\phi(B_\Delta)Y)_t(\phi(B_\Delta)Y)_{t+\Delta n}]$ and use the same expansion as in the proof of Lemma 3.5, i.e.

$$\phi(B_\Delta) = \prod_{j=1}^{p} (1 - e^{\lambda_j \Delta} B_\Delta) = \sum_{k=0}^{p} A_k^p B_\Delta^k, \qquad (3.14)$$

which we apply to Y. Observe that for $t \in \mathbb{R}$, setting $t_k := t - k\Delta$ for $k = 0, \ldots, p$, and $t_{p+1} := -\infty$,

$$B_\Delta^k Y_t = B_\Delta^k \int_{-\infty}^{t} g(t-u) dL_u = \int_{-\infty}^{t_k} g(t_k - u) dL_u = \sum_{i=k}^{p} \int_{t_{i+1}}^{t_i} g(t_k - u) dL_u. \qquad (3.15)$$

Applying the operator (3.14) to Y_t, using (3.15) and interchanging the order of summation gives

$$(\phi(B_\Delta)Y)_t = \sum_{m=0}^{p} \int_{t_{m+1}}^{t_m} \sum_{k=0}^{m} A_k^p g(t_k - u) dL_u. \qquad (3.16)$$

From Lemma 3.5 we know that the contribution from the term corresponding to $m = p$ is zero. By stationarity, the autocovariance function is independent of t, hence we can choose $t = \Delta n$. Then we obtain

$$\begin{aligned}
(\phi(B_\Delta)Y)_{n\Delta} &= \sum_{j=0}^{p-1} \int_{\Delta(n-j-1)}^{\Delta(n-j)} \sum_{k=0}^{j} A_k^p g((n-k)\Delta - u) dL_u \\
&= \sum_{j=1}^{p} \int_{\Delta(n-j)}^{\Delta(n-j+1)} \sum_{k=0}^{j-1} A_k^p g((n-k)\Delta - u) dL_u.
\end{aligned}$$

For $t = 0$, we obtain analogously

$$(\phi(B_\Delta)Y)_0 = \sum_{i=1}^{p} \int_{-\Delta i}^{-\Delta(i-1)} \sum_{h=0}^{i-1} A_h^p g(-\Delta h - u) dL_u.$$

For the autocovariance function we obtain for $n = 0, \ldots, p-1$ by using the fact that L has orthogonal increments,

$$\begin{aligned}
\gamma_{MA}(n) &= \mathbb{E}[(\phi(B_\Delta)Y)_0(\phi(B_\Delta)Y)_{n\Delta}] \\
&= \sigma^2 \sum_{i=1}^{p-n} \sum_{k=0}^{i+n-1} \sum_{h=0}^{i-1} A_k^p A_h^p \int_{-i\Delta}^{-(i-1)\Delta} g((n-k)\Delta - u) g(-\Delta h - u) du.
\end{aligned}$$

Finally, (3.12) is obtained by changing the variable of integration from u to $s = -u$. $\square$

Theorem 3.7. *The autocovariance function $\gamma_{MA}(n)$ for $n = 1, \ldots, p-1$ has for $\Delta \downarrow 0$ the asymptotic form*

$$\gamma_{MA}(n) \sim \frac{\sigma^2 \Delta^{2(p-q)-1}}{((p-q-1)!)^2} \sum_{i=1}^{p-n} \sum_{k=0}^{n+i-1} \sum_{h=0}^{i-1} (-1)^{h+k} \binom{p}{k} \binom{p}{h} C(h,k,i,n;p-q-1), \tag{3.17}$$

where for $N \in \mathbb{N}_0$

$$C(h,k,i,n;N) \quad := \quad \int_0^1 (s+i-1-h)^N (s+i-1-k+n)^N ds.$$

Proof. We can rewrite the integral in (3.12) as

$$\Delta \int_0^1 g((s+i-1-h)\Delta) g((s+i-1-k+n)\Delta) ds. \tag{3.18}$$

Since g is infinitely differentiable on $(0,\infty)$ and the right derivatives at 0 exist, the integrand has one-sided Taylor expansions of all orders $M \in \mathbb{N}$,

$$\sum_{l=0}^{M} \frac{d^l \left[g((s+i-1-h)\Delta) g((s+i-1-k+n)\Delta) \right]}{d\Delta^l} \bigg|_{\Delta=0^+} \frac{\Delta^l}{l!} + o\left(\Delta^M\right)$$

$$= \sum_{l=0}^{M} \sum_{m=0}^{l} \binom{l}{m} (s+i-1-h)^{l-m} (s+i-1-k+n)^m g^{(l-m)}(0+) g^{(m)}(0+) \frac{\Delta^l}{l!} + o\left(\Delta^M\right),$$

as $\Delta \downarrow 0$. Choose $M = 2(p-q-1)$. Then by (1.8) there is only one term in the double sum which does not vanish, namely the term for which $m = p-q-1 = l-m$. Setting $N := p-q-1$ (so that $M = 2N$) the sum reduces to

$$\binom{2N}{N} (s+i-1-h)^N (s+i-1-k+n)^N \frac{1}{(2N)!} \Delta^{2N} + o\left(\Delta^{2N}\right).$$

Since $\binom{2N}{N}/(2N)! = (N!)^{-2}$, the integral in (3.18) is for $\Delta \downarrow 0$ asymptotically equal to

$$\frac{\Delta^{2N+1}}{(N!)^2} \int_0^1 (s+i-1-h)^N (s+i-1-k+n)^N ds + o(\Delta^{2N+1}), \tag{3.19}$$

and, since

$$\lim_{\Delta \downarrow 0} \sum_{\{i_1,\ldots,i_h\} \in C_h^p} e^{\Delta(\lambda_{i_1} + \cdots + \lambda_{i_h})} = \binom{p}{h},$$

we also have

$$A_k^p A_h^p = (-1)^{h+k} \binom{p}{k} \binom{p}{h} + o(1) \quad \text{as } \Delta \downarrow 0. \tag{3.20}$$

Combining (3.19) and (3.20), we obtain (3.17). □

Remark 3.8. (i) For computations the following expansion may be useful (as usual we set $0^0 = 1$)

$$\begin{aligned} C(h,k,i,n;N) &:= \int_0^1 (s+i-1-h)^N (s+i-1-k+n)^N ds \\ &= \sum_{l_1,l_2=0}^{N} \binom{N}{l_1}\binom{N}{l_2} (i-1-h)^{N-l_1}(i-1-k+n)^{N-l_2} \int_0^1 s^{l_1+l_2} ds \\ &= \sum_{l_1,l_2=0}^{N} \binom{N}{l_1}\binom{N}{l_2} \frac{1}{l_1+l_2+1}(i-1-h)^{N-l_1}(i-1-k+n)^{N-l_2}. \end{aligned}$$

Furthermore, we observe that C depends on p and q only through $p-q$.
(ii) Note that the right hand sides of (3.5) and (3.6) are the discrete Fourier transforms of (3.12) and (3.17), respectively. Note also the symmetry between (3.5) and (3.12) in the dependence on Δ and $p-q-1$. $\square$

So far we know that the moving average process $X_n = \theta(B)Z_n$ from (3.2) is of order not greater than $p-1$ but possibly lower. Our next result presents an asymptotic formula for $\gamma_{MA}(p-1)$, which shows clearly that this term is not 0.

Corollary 3.9. *For lag $n = p-1$ the autocovariance formula* (3.17) *reduces to*

$$\gamma_{MA}(p-1) \sim (-1)^q \frac{\sigma^2 \Delta^{2(p-q)-1}}{(2(p-q-1))!} \tag{3.21}$$

and $\gamma_{MA}(p-1)$ is therefore non-zero for all sufficiently small $\Delta > 0$.

Proof. From the expansion (3.17) we find

$$\gamma_{MA}(p-1) \sim \frac{\sigma^2\Delta^{2(p-q)-1}}{((p-q-1)!)^2} \sum_{k=0}^{p-1} (-1)^k \binom{p}{k} C(0,k,1,p-1;p-q-1). \tag{3.22}$$

Set $d := p-q \geq 1$, then

$$C(0,k,1,p-1;d-1) = \int_0^1 s^{d-1}(s-k+p-1)^{d-1} ds,$$

and, from Remark 3.8, this is a polynomial of order $d-1$. In order to apply known results on the difference operator, we define the polynomial $f(x) = \int_0^1 s^{d-1}(x+s+p-1)^{d-1}ds$. Then, using Eq. (5.40), p. 188, and the last formula on p. 189 in Graham et al. (1994), the sum in (3.22) can be written as

$$\begin{aligned} &\sum_{k=0}^{p-1} (-1)^k \binom{p}{k} C(0,k,1,p-1;d-1) \\ =\ & \sum_{k=0}^{p} (-1)^k \binom{p}{k} f(x-k)|_{x=0} - (-1)^p \binom{p}{p} C(0,k,1,p-1;d-1) \\ =\ & 0 + (-1)^{p+1} \int_0^1 s^{d-1}(s-1)^{d-1} ds \ = \ (-1)^{p+d} \int_0^1 s^{d-1}(1-s)^{d-1} ds, \end{aligned} \tag{3.23}$$

$p-q$	1	2	3	4
$\gamma_{MA}(p-1)$	$\Delta(-1)^{p-1}\sigma^2$	$6^{-1}\Delta^3(-1)^{p-2}\sigma^2$	$120^{-1}\Delta^5(-1)^{p-3}\sigma^2$	$5040^{-1}\Delta^7(-1)^{p-4}\sigma^2$

Table 1: Values of $\gamma_{MA}(p-1)$ for $p-q=1,\ldots,4$.

where we have used the fact that $d-1=p-q-1<p$. To obtain Eq. (3.21) it suffices to note that $(-1)^{p+d}=(-1)^{2p-q}=(-1)^q$ and that the integral in (3.23) is a beta function. Hence

$$\int_0^1 s^{d-1}(1-s)^{d-1}ds=\frac{(\Gamma(d))^2}{\Gamma(2d)}=\frac{((d-1)!)^2}{(2d-1)!}>0, \quad d\in\mathbb{N}.$$

□

Remark 3.10. If Y is the CARMA(p,q) process (1.4) then, from Theorem 3.1, the spectral density of $(1-B)^{p-q}Y^\Delta$ is asymptotically, as $\Delta\downarrow 0$,

$$\frac{\sigma^2}{2\pi}\Delta(-2\Delta^2)^{p-q-1}c_{p-q-1}(\omega)(1-\cos\omega)^{p-q}, \quad \pi\le\omega\le\pi.$$

If $p-q=1$, 2 or 3 this reduces to the corresponding spectral densities in Corollary 3.2, each divided by $2^q(1-\cos\omega)^q$. The corresponding moving average representations are as in Corollary 3.4 without the factors $(1-B)^q$.

In particular, for the CAR(1) process, $(1-B)Y^\Delta$ has a spectral density which is asymptotically $\sigma^2\Delta/(2\pi)$ so that, in the Gaussian case, the increments of Y^Δ for small Δ approximate those of Brownian motion with variance $\sigma^2 t$. □

In this paper we have considered only second-order properties of Y^Δ. It is possible (see Brockwell (2001), Theorem 2.2) to express the joint characteristic functions, $\mathbb{E}\exp(i\sum_{k=1}^m\theta_k Y_k^\Delta)$, for $m\in\mathbb{N}$, in terms of the coefficients a_j and b_j and the function $\xi(\cdot)$, where $\xi(\theta)$, for $\theta\in\mathbb{R}$, is the exponent in the characteristic function, $\mathbb{E}e^{i\theta L_1}=e^{\xi(\theta)}$, of L_1. In particular the marginal characteristic function is given by $\mathbb{E}\exp(i\theta Y_k^\Delta)=\exp\int_0^\infty\xi(\theta\mathbf{b}'e^{Au}\mathbf{e})du$, where $\mathbf{b}$, A and $\mathbf{e}$ are defined as in (1.2)-(1.3).

These expressions are awkward to use in practice, however Brockwell, Davis and Yang (2011) have found that least squares estimation (which depends only on second-order properties) for closely and uniformly spaced observations of a CARMA(2,1) process on a fixed interval $[0,T]$ gives good results. They find in simulations that for large T the empirically-determined sample covariance matrix of the estimators of a_1,a_2 and b_0 is close to the matrix calculated from the asymptotic (as $T\to\infty$) covariance matrix of the maximum likelihood estimators based on continuous observation on $[0,T]$ of the corresponding Gaussian CARMA process.

4 Conclusions

When a CARMA(p,q) process Y is sampled at times $n\Delta$ for $n\in\mathbb{Z}$, it is well-known that the sampled process Y^Δ satisfies discrete-time ARMA equations of the form (3.2). The determination of the moving average coefficients and white noise variance for given grid size Δ, however, is a non-trivial procedure. In this paper we have focussed on *high frequency sampling* of Y. We have determined the relevant second order quantities, the spectral density f_{MA} of the moving average on the right-hand side of (3.2) and its asymptotic representation as $\Delta\downarrow 0$. This includes the moving average coefficients as well as the variance of the innovations. We also derived an explicit expression for the autocovariance function γ_{MA} and its asymptotic representation as $\Delta \downarrow 0$. This shows, in particular, that the moving average is of order $p-1$ for Δ sufficiently small.

5 Acknowledgments

PJB gratefully acknowledges the support of this work by NSF Grant DMS-1107031 and the Institute of Advanced Studies at Technische Universität München where this work was initiated. The work of Vincenzo Ferrazzano was supported by the International Graduate School of Science and Engineering (IGSSE) of Technische Universität München. We are also indebted to a referee for valuable comments.

References

Barndorff-Nielsen, O. E. and Shephard, N.: 2001, Non-Gaussian OU based models and some of their uses in financial economics, *J. Roy. Statist. Soc. Ser. B* **63**, 167–241.

Brockwell, P. J.: 2001, Lévy-driven CARMA processes, *Ann. Inst. Statist. Math.* **53**(1), 113–123.

Brockwell, P. J., Chadraa, E. and Lindner, A. M.: 2006, Continuous time GARCH processes of higher order, *Ann. Appl. Probab.* **16**, 790–826.

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

Brockwell, P. J., Davis, R. A. and Yang, Y.: 2011, Estimation for non-negative Lévy-driven CARMA processes, *Journal of Business and Economic Statistics* **29**(2), 250–259.

Brockwell, P. J., Ferrazzano, V. and Klüppelberg, C.: 2011, Kernel estimation in continuous-time moving average processes. Submitted for publication.

Brockwell, P. J. and Marquardt, T.: 2005, Lévy-driven and fractionally integrated ARMA processes with continuous time parameter, *Statistica Sinica* **15**(2), 477–494.

Brockwell, P. and Lindner, A.: 2009, Existence and uniqueness of stationary Lévy-driven CARMA processes, *Stoch. Proc. Appl.* **119**, 2625–2644.

Doob, J. L.: 1944, The elementary Gaussian processes, *Ann. Math. Stat.* **15**, 229–282.

Doob, J. L.: 1990, *Stochastic Processes*, 2nd edn, Wiley.

Ferrazzano, V.: 2010, Windspeed recording process and related issues, *Technical report*, Technische Universität München, Munich. *http://www-m4.ma.tum.de/Papers/*.

Ferrazzano, V. and Klüppelberg, C.: 2011, Statistical estimation of turbulence. In preparation.

Graham, R. L., Knuth, D. E. and Patashnik, O.: 1994, *Concrete Mathematics*, 2nd edn, Addison-Wesley, Reading, MA.

Haug, S. and Czado, C.: 2007, An exponential continuous time GARCH process, *J. Appl. Probab.* **44**(4), 960–976.

Jacod, J., Klüppelberg, C. and Müller, G.: 2010, Testing for zero correlation or for a functional relationship between price and volatility jumps. Submitted for publication.

Jacod, J. and Todorov, V.: 2010, Do price and volatility jump together?, *Ann. Appl. Probab.* **20**(4), 1425–1469.

Jones, R.: 1981, Fitting a continuous time autoregression to discrete data, *Applied time series analysis II, Proc. Symp., Tulsa/USA* pp. 651–675.

Marquardt, T. and Stelzer, R.: 2007, Multivariate CARMA processes, *Stoch. Proc. Appl.* **117**, 96–120.

Mitrinović, D. and Kečkić, J. D.: 1984, *The Cauchy Method of Residues: Theory and Application*, Reidel, Dordrecht, Holland.

Sayed, A. H. and Kailath, T.: 2001, A survey of spectral factorization methods, *Numer. Linear Algebra Appl.* **8**, 467–496.

Todorov, V. and Tauchen, G.: 2006, Simulation methods for Lévy driven continuous-time autoregressive moving average CARMA stochastic volatility models, *J. Bus. Econom. Statist.* **24**(4), 455–469.