Sampling, embedding and inference for CARMA processes

Peter J. Brockwell* Alexander Lindner†

Abstract

A CARMA(p, q) process Y is a strictly stationary solution Y of the p^{th}-order formal stochastic differential equation $a(D)Y_t = b(D)DL_t$, where L is a two-sided Lévy process, $a(z)$ and $b(z)$ are polynomials of degrees p and q respectively, with $p > q$, and D denotes differentiation with respect to t. Since estimation of the coefficients of $a(z)$ and $b(z)$ is frequently based on observations of the Δ-sampled sequence $Y^\Delta := (Y_{n\Delta})_{n\in\mathbb{Z}}$, for some $\Delta > 0$, it is crucial to understand the relation betwen Y and Y^Δ. If $\mathbb{E}L_1^2 < \infty$ then Y^Δ is an ARMA sequence with coefficients depending on those of Y and the crucial problems for estimation are the determination of the coefficients of Y^Δ from those of Y (*the sampling problem*) and the determination of the coefficients of Y from those of Y^Δ (*the embedding problem*). In this paper we consider both questions and use the results to determine the asymptotic distribution, as $n \to \infty$, with Δ fixed, of $\sqrt{n\Delta}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, where $\hat{\boldsymbol{\beta}}$ is the quasi-maximum-likelihood estimator of the vector of coefficients of $a(z)$ and $b(z)$, based on n consecutive observations of Y^Δ.

MSC: 60G51; 60G10; 62M10

Keywords:CARMA process, sampling, embedding, complex-valued information matrix, quasi-maximum-likelihood estimation.

1 Introduction

Let $L = (L_t)_{t\in\mathbb{R}}$ be a Lévy process, i.e. a process with homogeneous independent increments, continuous in probability, with càdlàg sample paths and $L_0 = 0$. We shall make

*Statistics Department, Colorado State University, Fort Collins, Colorado 80523-1877, U.S.A. pjbrock@stat.colostate.edu, Tel: 970 412 3861, Fax: 970 491 7895.

†Institute of Mathematical Finance, Ulm University, Helmholtzstraße 18, 89081 Ulm, Germany alexander.lindner@uni-ulm.de

the additional assumption in this paper that $\mathbb{E}L_1^2 < \infty$. Then $\mathbb{E}L_t = mt$ for some $m \in \mathbb{R}$ and $\mathrm{Var}(L_t) = \nu^2 t$ for some $\nu \geq 0$.

For integers p and q such that $p > q$, let $a_1, \ldots, a_p, b_0, \ldots, b_{p-1}$ be real valued coefficients such that $b_j = 0$ for $j > q$, and $b_q = 1$. Define the polynomials $a(z)$ and $b(z)$ by

$$a(z) = z^p + a_1 z^{p-1} + \ldots + a_p = \prod_{k=1}^{p} (z - \lambda_k), \tag{1.1}$$

and

$$b(z) = b_0 + b_1 z + \ldots + b_q z^q = \prod_{k=1}^{q} (z - \mu_k) \tag{1.2}$$

Denote by D the differential operator with respect to t. A natural continuous-time analog of the difference equations defining an ARMA(p, q) sequence is the formal p^{th}-order stochastic differential equation

$$a(D)Y_t = b(D)DL_t. \tag{1.3}$$

Since the derivatives on the right of this equation do not exist in the usual sense we interpret this equation by means of a state-space formulation. More precisely, we define the $p \times p$ matrix and $p \times 1$ vectors,

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

with A defined to be $(-a_1)$ if $p = 1$. A CARMA(p, q) process Y driven by L and with characteristic polynomials $a(z)$ and $b(z)$ is then defined as a strictly stationary solution of the equations

$$Y_t = \mathbf{b}^T \mathbf{X}_t, \quad t \in \mathbb{R}, \tag{1.4}$$

and

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

i.e.

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t A\mathbf{X}_s \, ds + \mathbf{e}_p L_t, \tag{1.5}$$

(see Brockwell (2001)). We shall assume throughout that the zeroes of $a(z)$ are distinct and distinct from those of $b(z)$ and that the zeroes of both polynomials lie in the open left half of the complex plane, i.e. that the process is causal and invertible. Then the mean value of the process Y is $\mathbb{E}Y_t = b_0 m/a_p$ and the autocovariance function of Y is (see Brockwell (2001b)),

$$\kappa_Y(h) = \nu^2 \sum_{j=1}^{p} K_j e^{\lambda_j |h|},\ h \in \mathbb{R}, \tag{1.6}$$

where

$$K_j = \frac{b(\lambda_j)b(-\lambda_j)}{a'(\lambda_j)a(-\lambda_j)},\ j = 1, \ldots, p. \tag{1.7}$$

We shall also assume throughout the paper that $m = 0$ so that $\mathbb{E}L_t = \mathbb{E}Y_t = 0$ for all $t \in \mathbb{R}$. Results derived under this assumption can easily be translated to the case when $m \neq 0$ by adding $b_0 m/a_p$ to Y_t for each $t \in \mathbb{R}$.

Parameter estimation for a Lévy-driven CARMA process Y is often carried out (see, e.g. Garcia et al. (2011)) by quasi-maximum-likelihood (QML) estimation, which maximizes the Gaussian likelihood of the observations. These observations are frequently made at uniformly spaced times $n\Delta$, $n \in \mathbb{Z}$, for some fixed $\Delta > 0$. We shall refer to the sequence obtained by sampling Y at such times as the Δ-sampled sequence $Y^\Delta := (Y_{n\Delta})_{n\in\mathbb{Z}}$. QML estimation of the parameters of a (possibly multivariate) CARMA process based on uniformly spaced observations has been shown by Schlemm and Stelzer (2012) to be consistent and asymptotically normal under mild conditions.

The Gaussian likelihood under a specified CARMA model can be calculated, even when the observations are irregularly spaced in time, by using the Kalman recursions as described by Jones (1981) and maximizing numerically to obtain the QML estimate $\hat{\boldsymbol{\beta}}$ of the coefficient vector, $\boldsymbol{\beta} = (a_1\ \ldots\ a_p\ b_0\ \ldots\ b_{q-1})^T$.

An alternative approach, pioneered by Phillips (1959), Durbin (1961), Phillips (1973), Robinson (1977) and Bergstrom (1985) focuses on the Δ-sampled sequence which is known, when $\mathbb{E}L_1^2 < \infty$, to have an ARMA structure. If an ARMA model is fitted to the sampled data it is important to know whether or not there is a CARMA process of which it is the sampled sequence and if so to determine the parameters of the CARMA model. It is critical in this approach to have a clear understanding of the relations between the parameters of the sampled ARMA sequence and those of the underlying CARMA process. A clear specification of these relations also enables well-known asymptotic properties of the ARMA estimates to yield asymptotic properties of the CARMA estimates. In Section 2 we show how to determine the Δ-sampled ARMA parameters from the parameters of a CARMA(p, q) process by solving a polynomial equation of degree at most $p - 1$. In Section 3 we give conditions under which a specified ARMA(p, q) process has the autocovariances of a Δ-sampled CARMA process with autoregressive order p and show how to determine the parameters of the CARMA process, again by solving a polynomial equation of degree at most $p - 1$.

Based on the likelihood of a *complete* realization $(Y_t)_{t\in[0,T]}$ of a Gaussian CARMA process, Pham-Din-Tuan (1977) gave an explicit expression for the asymptotic distribution of $\sqrt{T}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ as $T\to\infty$. In Section 4 we use the results of Sections 2 and 3 together with the complex-valued information matrix (see Liu et al. (2014)) for the reciprocals of the zeros of the autoregressive and moving-average polynomials of an ARMA sequence (generalizing eq. (7.2.21) of Box et al. (2016) which applies only to real-valued autoregressive and moving-average zeros) to determine the asymptotic distribution of $\sqrt{n\Delta}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ as $n\to\infty$ with fixed $\Delta>0$, where $\hat{\boldsymbol{\beta}}$ here denotes the QML estimator of $\boldsymbol{\beta}$ based on observations of $Y_\Delta,\ldots,Y_{n\Delta}$. This is the relevant asymptotic distribution when estimation is based on observations of the Δ-sampled sequence rather than on a complete realization of Y. The results of Pham-Din-Tuan however provide a useful approximation to our results when Δ is small, as illustrated in the case of a particular CARMA$(2,1)$ process. For completeness we also include in Section 4 a brief account of the use of the Kalman recursions to compute the Gaussian likelihood and QML estimators of the parameters of a CARMA process.

Many numerical algorithms are available for the determination of the Δ-sampled sequence of a CARMA(p,q) process, including an asymptotic solution as $\Delta\to 0$ (see Brockwell et al. (2013)), however the algorithm provided by Theorem 2.1 reduces the problem of determining the exact ARMA parameters essentially to the factorization of a polynomial of order at most $p-1$. The embedding problem has been investigated by many authors, starting with Phillips (1959). More recently Huzii (2006) and Thornton and Chambers (2013) have given conditions for embeddability and applied them to low-order ARMA processes. However easily verifiable conditions under which an ARMA(p,q) sequence can be embedded in a CARMA(p,r) process for some $r<p$ have not previously been stated. Such conditions are provided in Theorem 3.5.

2 The sampling problem

As stated in Section 1 we assume (without loss of generality from a second-order point of view) that the p zeroes, $\lambda_1,\ldots,\lambda_p$, of $a(z)$ all lie in the open left half-plane of $\mathbb{C}$ (i.e. that the process is causal) and that $m=\mathbb{E}L_1=0$.

The Δ-sampled sequence $(Y_n^\Delta)_{n\in\mathbb{Z}}$ of the CARMA process Y is defined to be the sequence,

$$Y_n^\Delta=Y_{n\Delta},\ n\in\mathbb{Z}.$$

If we define

$$\phi(B):=\prod_{i=1}^{p}(1-e^{\lambda_i\Delta}B)=(1-d_1B-\ldots-d_pB^p),$$

then the Δ-sampled sequence $(Y_n^\Delta)_{n\in\mathbb{Z}}$ of Y satisfies (see Brockwell and Lindner (2009))

$$\phi(B)Y_n^\Delta = Y_n^\Delta - d_1 Y_{n-1}^\Delta - \ldots - d_p Y_{n-p}^\Delta = Z_n^1 + Z_{n-1}^2 + \ldots + Z_{n-p+1}^p,$$

where

$$Z_n^r := \int_{(n-1)\Delta}^{n\Delta} \mathbf{b}^T \left(e^{(r-1)A\Delta} - \sum_{j=1}^{r-1} d_j e^{(r-1-j)A\Delta} \right) e^{A(n-u)\Delta} \mathbf{e}_p \, dL_u, \ \ r = 1, \ldots, p.$$

This means that the Δ-sampled sequence satisfies an autoregressive equation of order p driven by a $(p-1)$-dependent sequence. Since we are assuming throughout that $\mathbb{E}(L_1^2) < \infty$ this implies (Brockwell and Davis (1991)) that

$$\prod_{j=1}^{p}(1 - e^{\lambda_j \Delta} B) Y_n^\Delta = \prod_{j=1}^{p-1}(1 - \eta_j B) Z_n, \ \ n \in \mathbb{Z}, \tag{2.1}$$

where $(Z_n)_{n\in\mathbb{Z}}$ is zero-mean white noise with variance σ^2 and $\eta_1, \ldots, \eta_{p-1}$ are moving-average coefficients, some of which may be zero. The sampling problem is to determine the coefficients η_j and the white noise variance σ^2. The following theorem reduces this problem to finding the zeroes of a polynomial of order at most $p-1$. The asymptotic form of the coefficients as $\Delta \to 0$ was determined by Brockwell et al. (2012) and (2013).

Theorem 2.1. Regular sampling of a CARMA process

Suppose that L is a zero-mean Lévy process with $\mathrm{Var}(L_1) = \nu^2 < \infty$ and let Y be the strictly stationary causal invertible CARMA(p, q) process defined by the formal equation,

$$\prod_{j=1}^{p}(D - \lambda_j) Y_t = \prod_{j=1}^{q}(D - \mu_j) D L_t, \tag{2.2}$$

where $e^{\lambda_i \Delta} \neq e^{\lambda_j \Delta}$ if $i \neq j$ and $\lambda_i \neq \mu_j$ for all i and j. As usual, D denotes differentiation with respect to t and the autoregressive and moving-average polynomials, $a(z) = \prod_{j=1}^{p}(z - \lambda_j)$ and $b(z) = \prod_{j=1}^{q}(z - \mu_j)$, are assumed to have real-valued coefficients.

Define, as in (1.7),

$$K_j := \frac{b(\lambda_j) b(-\lambda_j)}{a'(\lambda_j) a(-\lambda_j)}, \ \ j = 1, \ldots, p,$$

and the polynomial of degree $r \leq p-1$,

$$k(z) := \sum_{j=1}^{p} K_j \sinh(\lambda_j \Delta) \prod_{i \neq j}(z - \cosh(\lambda_i \Delta)).$$

Denoting by $k_1, \ldots, k_r$ the zeroes of $k(z)$, we define η_j, for $j = 1, \ldots, r$, to be the value of $k_j \pm \sqrt{k_j^2 - 1}$ with absolute value less than 1 and η_j, for $j > r$, to be zero.

Then $Y_n^\Delta := Y_{n\Delta},\ n \in \mathbb{Z}$, satisfies the causal invertible ARMA(p,r) equations,

$$\prod_{j=1}^{p}(1 - e^{\lambda_j \Delta}B)Y_n^\Delta = \prod_{j=1}^{r}(1 - \eta_j B)Z_n,\ n \in \mathbb{Z}, \tag{2.3}$$

where the sequence $(Z_n)_{n\in\mathbb{Z}}$ is zero mean white noise with variance

$$\sigma^2 = \frac{\nu^2 \prod_{j=1}^{p} e^{\lambda_j \Delta}}{\prod_{j=1}^{r} \eta_j}(-2)^{p-r} c_r, \tag{2.4}$$

and c_r is the coefficient of z^r in *$k(z)$. (If $c_p \neq 0$, then $r = p - 1$.)*

Proof. The autocovariance function of the CARMA process Y (see (1.6)) is

$$\kappa_Y(h) = \nu^2 \sum_{j=1}^{p} K_j e^{\lambda_j |h|},\ h \in \mathbb{R}.$$

The autocovariance function γ_Δ of the sampled sequence Y^Δ coincides at lag n, $n \in \mathbb{Z}$, with the autocovariance function of Y at lag $n\Delta$. Hence

$$\gamma_\Delta(n) = \nu^2 \sum_{j=1}^{p} K_j e^{\lambda_j \Delta |n|},\ n \in \mathbb{Z}. \tag{2.5}$$

The function γ_Δ is the autocovariance function of an ARMA sequence satisfying the difference equations (2.1), with some of the moving average parameters η_j possibly zero. Since γ_Δ is the absolutely summable autocovariance function of a covariance stationary sequence, it has the corresponding spectral density,

$$f_\Delta(\omega) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \gamma_\Delta(n) e^{-in\omega},\quad -\pi \le \omega \le \pi. \tag{2.6}$$

Substituting from (2.5) into (2.6) we find that

$$f_\Delta(\omega) = \frac{\nu^2}{2\pi} \sum_{j=1}^{p} \frac{K_j \sinh(\lambda_j \Delta)}{\cos\omega - \cosh(\lambda_j \Delta)} = \frac{\nu^2}{2\pi} \frac{k(\cos\omega)}{\prod_{j=1}^{p}(\cos\omega - \cosh(\lambda_j \Delta))}. \tag{2.7}$$

Since Y^Δ satisfies (2.1) for some $\eta_1, \ldots, \eta_{p-1}$ and σ^2, we can write

$$f_\Delta(\omega) = \frac{\sigma^2 \prod_{j=1}^{p-1}(1 + \eta_j^2 - 2\eta_j \cos\omega)}{2\pi \prod_{j=1}^{p}(-2e^{\lambda_j \Delta}) \prod_{j=1}^{p}(\cos\omega - \cosh(\lambda_j \Delta))} \tag{2.8}$$

(where some of the parameters η_j may be zero). Comparing (2.7) and (2.8) we see at once that the polynomial in $\cos\omega$ in the numerator of (2.8) is zero if and only if $\cos\omega$ is equal to one of the zeros k_j of the polynomial $k(z)$.

If $\sum_{j=1}^{p} K_j \sinh(\lambda_j \Delta) \neq 0$, then $k(z)$ has $p-1$ zeros, $k_1, \ldots, k_{p-1}$, so the parameters η_j, $j = 1, \ldots, p-1$ satisfy the equations,

$$\eta_j^2 - 2k_j\eta_j + 1 = 0, \ j = 1, \ldots, p-1,$$

and the invertible version of the sampled process is obtained by choosing each η_j to have absolute value less than 1. The white-noise variance σ^2 is then found by equating the coefficients of $(\cos\omega)^{p-1}$ in the two expressions for $f_\Delta(\omega)\prod_{j=1}^{p}(\cos\omega - \cosh(\lambda_j\Delta))$ obtained from (2.7) and (2.8).

If $\sum_{j=1}^{p} K_j \sinh(\lambda_j \Delta) = 0$, then $k(z)$ has order $r < p-1$, the polynomial in $\cos\omega$ in the numerator of (2.8) is of order r and the parameters $\eta_1, \ldots, \eta_r$ satisfy the equations,

$$\eta_j^2 - 2k_j\eta_j + 1 = 0, \ j = 1, \ldots, r,$$

where $k_1, \ldots, k_r$ are the zeros of $k(z)$. For invertibility, the values of η_j with absolute values less than 1 are chosen and the value of the white-noise variance σ^2 is then obtained by equating the coefficients of $(\cos\omega)^r$ in the two expressions for $f_\Delta(\omega)\prod_{j=1}^{p}(\cos\omega - \cosh(\lambda_j\Delta))$ obtained from (2.7) and (2.8). □

Remark 2.2. *Repeated autoregressive zeros.* If the *distinct* zeros λ of $a(z)$ have multiplicities $m(\lambda)$, where $\sum_\lambda m(\lambda) = p$, then

$$\kappa(h) = \nu^2 \sum_\lambda e^{\lambda|h|} \sum_{j=0}^{m(\lambda)-1} K_{\lambda,j}|h|^j, \quad h \in \mathbb{R},$$

where the coefficients $K_{\lambda,j}$ can be found from equation (2.14) of Brockwell (2001b). Then γ_Δ in (2.5) must be replaced by

$$\gamma_\Delta(n) = \nu^2 \sum_\lambda e^{\lambda|\Delta n|} \sum_{j=0}^{m(\lambda)-1} K_{\lambda,j}|n\Delta|^j, \quad n \in \mathbb{Z}.$$

Equations (2.6) and (2.8) are unchanged but, on the right-hand side of (2.7), $k(\cos\omega)$ becomes a more complicated polynomial in $\cos\omega$ of degree less than or equal to $p-1$, specifically,

$$k(\cos\omega) = \prod_{j=1}^{p}(\cos\omega - \cosh(\lambda_j\Delta)) \sum_{n=-\infty}^{\infty} \nu^{-2}\gamma_\Delta(n)e^{-in\omega}.$$

If we replace k in the statement of Theorem 2.1 by this more general polynomial, the theorem is valid also when $a(z)$ has zeros of multiplicity possibly greater than 1.

Remark 2.3. If $\lambda_i \neq \lambda_j$ for all $i \neq j$ but for some $i \neq j$, $e^{\lambda_i\Delta} = e^{\lambda_j\Delta}$, as is the case if λ_i and λ_j differ by an integer multiple of $2\pi i/\Delta$, then Y^Δ is an ARMA(d, s) process with d equal to the number of distinct values of $e^{\lambda_j \Delta}$ and $s < d$. For example, if Y is a CARMA$(2,1)$ process with $\lambda_1 = -1 + i\pi$ and $\lambda_2 = -1 - i\pi$, then the sequence obtained by sampling at integer times is an AR(1) sequence. Such cases are rather special and, since they can be treated by the same arguments as used in the proof of the preceding theorem, we shall omit the details. □

Example 2.4. For the causal invertible CARMA(2,1) process with $a(z) = (z-\lambda_1)(z-\lambda_2)$ and $b(z) = z-\mu$, where $\lambda_1 \neq \lambda_2$, $\mu \neq \lambda_1$ and $\mu \neq \lambda_2$, we have, $K_1 = [\mu^2-\lambda_1^2]/[2\lambda_1(\lambda_1^2-\lambda_2^2)]$ and $K_2 = [\mu^2-\lambda_2^2]/[2\lambda_2(\lambda_2^2-\lambda_1^2)]$. The parameter η_1 in the invertible representation (2.3) is therefore the root with absolute value less than 1 of the equation,

$$z^2 - 2k_1 z + 1 = 0,$$

where

$$k_1 = \frac{K_1 \sinh(\lambda_1\Delta)\cosh(\lambda_2\Delta) + K_2 \sinh(\lambda_2\Delta)\cosh(\lambda_1\Delta)}{K_1 \sinh(\lambda_1\Delta) + K_2 \sinh(\lambda_2\Delta)}.$$

From η_1 the white-noise variance σ^2 is found directly from (2.4). As $\Delta \to 0$ we easily find that $\eta_1 = 1 + \mu\Delta + o(\Delta)$ and $\sigma^2 \sim \nu^2\Delta$. □

Remark 2.5. High frequency approximations.

The factorization of $k(z)$ required in the application of Theorem 2.1 cannot be carried out in an algebraically explicit manner for CARMA(p,q) processes with $p > 5$ since then the polynomial $k(z)$ is of order greater than 4. However general asymptotic expressions (as $\Delta \to 0$) for the moving-average coefficients and white-noise variance in the representation (2.3) were found by Brockwell et al. (2012) and (2013) and applied to inference for the kernel of the underlying CARMA process when observations are made at times $\Delta, 2\Delta, \ldots, n\Delta$, with Δ small and n large. Applying Theorem 1 of Brockwell et al. (2013) to Example 2.4 we find that $\eta_1 = 1 + \mu\Delta + o(\Delta)$ and $\sigma^2 \sim \nu^2\Delta$, as in the example above. □

3 The embedding problem

The second-order properties of a CARMA process (see Section 1) depend only on the coefficient vector $\boldsymbol{\beta} = (a_1, \ldots, a_p, b_0, \ldots, b_{q-1})^T$ and the variance ν^2 of the driving Lévy process at time 1.

The problem of determining parameters $\boldsymbol{\beta}$ and ν^2 of a real-valued CARMA process Y such that the autocovariances of Y at $n\Delta,\ n \in \mathbb{Z}$, coincide with the autocovariances at lags $n,\ n \in \mathbb{Z}$, of a given causal ARMA(p,q) sequence U with $p > q$ is more complicated. If such a $\boldsymbol{\beta}$ and ν^2 exist we shall say that U is Δ-*embeddable in* Y. Depending on U, there may be more than one such set of coefficients or possibly none at all.

The embedding theorem in this section gives readily verifiable conditions under which an ARMA(p,q) process is Δ-embeddable in a CARMA(p,r) process for some $r \leq p-1$ and specifies the parameters of such a CARMA process. The result depends on the following characterization of CARMA autocovariance functions. From (1.6) every real-valued causal CARMA(p,q) process with distinct $\lambda_1, \ldots, \lambda_p$, has an autocovariance function of the form $\kappa(h) := \sum_{j=1}^{p} c_j e^{\lambda_j |h|}$, $h \in \mathbb{R}$, for some $c_1, \ldots, c_p \in \mathbb{C}$. The following theorem establishes conditions under which such a linear combination of exponentials *is* in fact the autocovariance function of a CARMA process and specifies the parameters of a CARMA process which has κ as its autocovariance function.

Theorem 3.1. Characterization of CARMA autocovariance functions

Suppose that κ is a real-valued function of the form,

$$\kappa(h) := \sum_{j=1}^{p} c_j e^{\lambda_j |h|}, \quad h \in \mathbb{R},$$

where $c_j \in \mathbb{C}$, $\mathrm{Re}(\lambda_j) < 0$, $j = 1, \ldots, p$ and $\lambda_i \neq \lambda_j$ if $i \neq j$. Then the following statements are equivalent:

(i) κ is the autocovariance function of a mean-square continuous covariance stationary process, $(Y_t)_{t \in \mathbb{R}}$.

(ii) κ is non-negative definite.

(iii) $g(z) := \sum_{j=1}^{p} c_j \lambda_j \prod_{m \neq j} (z + \lambda_m^2) \leq 0$ for all $z \geq 0$.

(iv) κ is the autocovariance function of a real-valued CARMA(p,q) process Y satisfying the formal differential equation,

$$\prod_{j=1}^{p} (D - \lambda_j) Y_t = \prod_{j=1}^{q} (D + r_j e^{i\theta_j}) D L_t,$$

where q is the order of the polynomial g, $\mathbb{E}(L_1) = 0$, $\mathbb{E}(L_1^2) = -2K$, and K and $r_j e^{i\theta_j}$, $j = 1, \ldots q$, are found by writing $g(z) = K \prod_{j=1}^{q} (z + r_j^2 e^{2\theta_j i})$, with $r_j > 0$ and $-\pi/2 < \theta_j \leq \pi/2$.

Proof. If (i) holds then (ii) is an immediate consequence of Bochner's Theorem since κ is continuous at zero.

If (ii) holds, then, since κ is continuous and $f(\omega) := \frac{1}{2\pi} \int_{\mathbb{R}} \kappa(h) e^{-i\omega h} \, dh$, $\omega \in \mathbb{R}$, is integrable, we can write

$$\kappa(h) = \int_{\mathbb{R}} e^{i\omega h} f(\omega) \, d\omega.$$

In other words f is the spectral density of κ, so $f(\omega) \geq 0$ for all $\omega \in \mathbb{R}$. Direct evaluation of the integral defining f gives

$$f(\omega) = -\frac{\sum_{j=1}^{p} c_j \lambda_j \prod_{m \neq j} (\lambda_m^2 + \omega^2)}{\pi \prod_{j=1}^{p} (\lambda_j^2 + \omega^2)}, \quad \omega \in \mathbb{R}.$$

and non-negativity of f implies condition (iii).

If (iii) holds then g(z) is factorizable as

$$g(z) = K\prod_{j=1}^{q}(z + r_j^2 e^{2\theta_j i}),$$

where $r_j > 0$, $-\pi/2 < \theta_j \leq \pi/2$, $q < p$ is the order of the polynomial $g(z)$ and $K < 0$ is the coefficient of z^q. This implies that f is the spectral density and κ is the autocovariance function of the stationary CARMA(p, q) process satisfying the formal differential equation,

$$\prod_{j=1}^{p}(D - \lambda_j)Y_t = \prod_{j=1}^{q}(D + r_j e^{i\theta_j})DL_t,$$

with $\mathbb{E}(L_1) = 0$ and $\mathbb{E}(L_1^2) = -2K$. (If $\sum_{j=1}^{p} c_j\lambda_j \neq 0$ then $q = p - 1$.)

If (iv) holds then clearly (i) holds so we have established the required equivalence. □

Remark 3.2. The same theorem holds for the real-valued function

$$\kappa(h) = \sum_{\lambda} e^{\lambda|h|} \sum_{j=0}^{m(\lambda)-1} c_{\lambda,j}|h|^j, \quad h \in \mathbb{R},$$

where $c_{\lambda,j} \in \mathbb{C}$, $\text{Re}(\lambda) < 0$, $m(\lambda)$ is the multiplicity of λ, $\sum_\lambda$ denotes summation over *distinct* values of λ, $\sum_\lambda m(\lambda) = p$, and we replace the polynomial $g(z)$ in (iii) by the more complicated polynomial of degree at most $p - 1$, namely,

$$g(z) = -\frac{1}{2}\left[\prod_{\lambda}(z + \lambda^2)^{m(\lambda)}\right]\int_{\mathbb{R}} \kappa(h)e^{-ih\sqrt{z}}dh, \quad z \geq 0,$$

where $\prod_\lambda$ denotes multiplication over the distinct values of λ. □

A similar argument leads to a discrete-time analogue of Theorem 3.1 which we state without proof.

Proposition 3.3. *Suppose that γ is a real-valued function of the form,*

$$\gamma(h) := \sum_{j=1}^{p} c_j e^{\lambda_j|h|}, \quad h \in \mathbb{Z},$$

where $c_j \in \mathbb{C}$, $\text{Re}(\lambda_j) < 0$, $j = 1, \ldots, p$, and $e^{\lambda_i} \neq e^{\lambda_j}$ if $i \neq j$. Then the following statements are equivalent:

(i) γ is the autocovariance function of a covariance stationary sequence $(U_n)_{n\in\mathbb{Z}}$.

(ii) γ is non-negative definite.

(iii) $h(z) := \sum_{j=1}^{p} c_j \sinh(\lambda_j)\prod_{m\neq j}(\cosh(\lambda_j) - z) \leq 0$ for all $z \in [-1, 1]$.

(iv) γ is the autocovariance function of a real-valued ARMA(p, q) process where q is the order of the polynomial $h(z)$.

We need one further result for ARMA sequences.

Proposition 3.4. *Suppose that $(U_n)_{n\in\mathbb{Z}}$ is the ARMA(p,q) sequence satisfying the equations, with $p > q$,*

$$\phi(B)U_n = \theta(B)Z_n,\ \ n \in \mathbb{Z}, \tag{3.1}$$

where $(Z_n)_{n\in\mathbb{Z}}$ is zero-mean white noise with variance $\sigma^2 > 0$, B is the backward shift operator, $\phi(z) := \prod_{j=1}^{p}(1-\xi_j z)$ and $\theta(z) := \prod_{j=1}^{q}(1-\eta_j z)$. Suppose also that $\xi_i \neq \xi_j$, for $i \neq j$, that $\xi_i \neq \eta_j$ for all i and j, that $0 < |\xi_i| < 1$ and $0 < |\eta_j| < 1$ for all i and j, and that the coefficients of the polynomials $\phi(z)$ and $\theta(z)$ are real-valued. Define

$$G_j = -\xi_j \frac{\theta(\xi_j)\theta(\xi_j^{-1})}{\phi(\xi_j)\phi'(\xi_j^{-1})}. \tag{3.2}$$

Then the autocovariance function γ of $(U_n)_{n\in\mathbb{Z}}$ can be expressed as

$$\gamma(h) = \sigma^2 \sum_{j=1}^{p} G_j e^{|h|\log\xi_j},\ \ h \in \mathbb{Z}. \tag{3.3}$$

Proof. See Brockwell (2001b). □ □

For the ARMA sequence in Proposition 3.4 to be Δ-embeddable in a CARMA(p,r) process for some $r < p$ there must exist parameters ν^2 and $\boldsymbol{\beta} = (a_1,\ldots,a_p,b_0,\ldots,b_{r-1})^T$, or equivalently $\boldsymbol{\theta} = (\lambda_1,\ldots,\lambda_p,\mu_1,\ldots,\mu_r)^T$, such that the autocovariance at lag $n\Delta$ of a CARMA process with these parameters coincides with that of U at lag n. The following theorem gives conditions under which this is the case and specifies a set of parameters $(\lambda_1,\ldots,\lambda_p,\mu_1,\ldots,\mu_r)^T$ and ν^2 of a CARMA process into which the ARMA process is Δ-embeddable.

Theorem 3.5. Embedding an ARMA sequence

Suppose that $(U_n)_{n\in\mathbb{Z}}$, is the ARMA(p,q) sequence specified in Proposition 3.4. With G_j as in (3.2), define the (real-valued) polynomial of degree $r \le p-1$,

$$g(z) := \sigma^2\Delta^{-1}\sum_{j=1}^{p} G_j \log(\xi_j) \prod_{m\neq j}(z + \Delta^{-2}(\log\xi_m)^2), \tag{3.4}$$

where $\log(\xi_j)$, $j = 1,\ldots,p$, are any specified values of the logarithms such that if $\xi_j = \overline{\xi_k}$ then $\log\xi_j = \overline{\log\xi_k}$. (If $\sum_{j=1}^{p} G_j\log(\xi_j) \neq 0$ then $r = p-1$.)

If $g(z) \le 0$ for all $z \ge 0$, *then $g(z)$ can be factorized as*

$$g(z) = S\prod_{j=1}^{r}(z + r_j^2 e^{2i\theta_j}),$$

where $r \leq p-1$, $r_j > 0$, $\theta_j \in (-\pi/2, \pi/2]$ and $S < 0$. Under these conditions U has the autocovariance function of the Δ-sampled sequence obtained from the stationary CARMA process Y satisfying the causal invertible equation,

$$\prod_{j=1}^{p}(D - \Delta^{-1}\log\xi_j)Y_t = \prod_{j=1}^{r}(D + r_j e^{i\theta_j})DL_t. \tag{3.5}$$

where L is any driving Lévy process such that $\mathbb{E}L_1 = 0$, $\mathbb{E}(L_1)^2 = \nu^2$ and

$$\nu^2 = -2\sigma^2 S. \tag{3.6}$$

Proof. The autocovariance function of the ARMA process defined by (3.1) is

$$\gamma(h) = \sigma^2 \sum_{j=1}^{p} G_j \xi_j^{|h|}, \ h \in \mathbb{Z}.$$

Consider the function,

$$\kappa(h) = \sigma^2 \sum_{j=1}^{p} G_j e^{|h|\Delta^{-1}\log\xi_j}, \ h \in \mathbb{R}. \tag{3.7}$$

The values $\kappa(h\Delta)$ clearly coincide with $\gamma(h), \ h \in \mathbb{Z}$, as required.

By Theorem 3.1 with $c_j = \sigma^2 G_j$ and $\lambda_j = \Delta^{-1}\log\xi_j$, $j = 1, \ldots, p$, κ is the autocovariance function of a CARMA process with autoregressive order p if and only if $g(z) \leq 0$ for all $z \geq 0$, and in this case it is the autocovariance function of the process specified by (3.5) and (3.6). □

Corollary 3.6. *Define $\boldsymbol{\theta} := (\boldsymbol{\lambda}^T, \boldsymbol{\mu}^T)^T$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)^T$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_q)^T$ are respectively the vectors of autoregressive and moving-average zeros, in any prescribed order, of a causal invertible CARMA process satisfying the conditions of Theorem 2.1, let Θ denote the set of all such parameter vectors, and let T be the mapping $T : \boldsymbol{\theta} \mapsto \boldsymbol{\psi}$ where $\boldsymbol{\psi} = (\boldsymbol{\xi}^T, \boldsymbol{\eta}^T)^T$ with $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)^T$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{p-1})^T$, the autoregressive and moving-average parameter vectors of the Δ-sampled ARMA sequence respectively.*

Then, for any given $\boldsymbol{\theta}_0 = (\boldsymbol{\lambda}_0^T, \boldsymbol{\mu}_0^T)^T \in \Theta$, the restriction T_0 of T to $\Theta_0 := \{\boldsymbol{\theta} \in \Theta : |\mathrm{Im}(\lambda_j - \lambda_{0j})| < \pi, \ j = 1, \ldots, p\}$ is a bijective mapping onto $T_0\Theta_0$. □

Proof. The inverse mapping T_0^{-1} is specified by Theorem 3.5 with the logarithm of each ξ_j chosen so that

$$|\mathrm{Im}(\Delta^{-1}\log\xi_j - \lambda_{0j})| < \pi, \ j = 1, \ldots, p.$$ □

Remark 3.7. If $g(z) > 0$ for some $z \geq 0$ then the function κ in (3.7) is not non-negative definite and is therefore not the autocovariance function of any CARMA process. There may however be alternative choices of the logarithms, equal modulo$(2i\pi)$ to those specified in the theorem, for which the condition $g(z) \leq 0$ for all $z \geq 0$ is satisfied. In this case the conclusions hold with the revised values of the logarithms. □

Remark 3.8. Since $g(z)$ is a polynomial, the condition that $g(z) \leq 0$ for all $z \geq 0$ is easily checked by evaluating $g(0)$, $\lim_{z\to\infty} g(z)$ and the values of g at each of the positive values of z where $g'(z) = 0$. □

Remark 3.9. If exactly m of the parameters ξ_i, say $\xi_1, \ldots, \xi_m$, are strictly negative real numbers then the theorem continues to hold with p replaced by $p+m$, $\log \xi_j := \log|\xi_j| + i\pi$, $\log \xi_{p+j} := \log|\xi_j| - i\pi$, $j = 1, \ldots, m$, and $G_{p+j} = G_j = -\frac{\xi_j \theta(\xi_j)\theta(\xi_j^{-1})}{2\phi(\xi_j)\phi'(\xi_j^{-1})}$, $j = 1, \ldots, m$. □

Remark 3.10. *Repeated autoregressive parameters.* An ARMA(p, q) sequence with $q < p$ and autoregressive parameters ξ having multiplicities $m(\xi)$, where $\sum_\xi m(\xi) = p$, has autocovariance function of the form,

$$\gamma(h) = \sigma^2 \sum_\xi \xi^{|h|} \sum_{j=0}^{m(\xi)-1} G_{\xi,j} |h|^j, \quad h \in \mathbb{Z},$$

where the coefficients $G_{\xi,j}$ can be found from Brockwell(2001b), equation (3.12). Theorem 3.5 holds for such sequences if $g(z)$ is redefined as in Remark 3.2 with λ replaced by $\Delta^{-1} \log \xi$ and $c_{\lambda,j}$ replaced by $\sigma^2 \Delta^{-j} G_{\xi,j}$.

□

Example 3.11. Sampling and recovering a CARMA(2,1) process

(i) Consider the Δ-sampled sequence Y^Δ, with $\Delta = .1$, of the stationary CARMA(2,1) process defined by the formal differential equation,

$$(D + .5)(D + 1)Y_t = (D + .25)DL_t, \tag{3.8}$$

where D as usual denotes differentiation with respect to t and the driving Lévy process L has moments $\mathbb{E}L_1 = 0$ and $\mathbb{E}(L_1^2) = 1$. Taking $\lambda_1 = -.5$ and $\lambda_2 = -1$, we find immediately that the constants K_1 and K_2 in Theorem 2.1 are given by

$$K_1 = -.25 \quad \text{and} \quad K_2 = .625.$$

Since $K_1 \sinh(-1) + K_2 \sinh(-2) \neq 0$, we deduce from the theorem that the Δ-sampled sequence Y^Δ with $\Delta = .1$ satisfies the ARMA(2,1) equation,

$$(1 - e^{-.05}B)(1 - e^{-.1}B)Y_n^\Delta = (1 - \eta B)Z_n, \ \{Z_n\} \sim \text{WN}(0, \sigma^2), \tag{3.9}$$

where η and σ^2 are to be determined. From Example 2.4, $k_1 = 1.0003132479211$. Hence, by Theorem 2.1,

$$\eta = k_1 - \sqrt{k_1^2 - 1} = 0.9752813889$$

and

$$\sigma^2 = 2(K_1 \sinh(.05) + K_2 \sinh(.1))e^{-.15}/\eta = .08842703.$$

(ii) Suppose on the other hand that we start with an ARMA sequence satisfying (3.9) with η and σ^2 as specified, and wish to find a CARMA$(2,r)$ process with $r \leq 1$ of which it is the Δ-sampled sequence with $\Delta = 0.1$. To do so we apply Theorem 3.5.

With $\xi_1 = e^{-.05}$, $\xi_2 = e^{-.1}$, $\log \xi_1 = -.05$, $\log \xi_2 = -.1$ and $\eta = .9752813889$, we readily find that G_1 and G_2 as defined in (3.2) are

$$G_1 = -2.8271899 \quad \text{and} \quad G_2 = 7.0679747.$$

Hence $\sum_{j=1}^{2} G_j \log \xi_j = -.56543798$ and $g(z) = -z/2 - 1/32$. From $g(z)$ and Theorem 3.5 we see at once that $\lambda_1 = -.5$, $\lambda_2 = -1$, $\mu = -.25$ and $\nu^2 = 1$. In other words we have recovered the CARMA model defined by (3.8). □

4 Inference for CARMA(p,q) processes

In this section we consider quasi-maximum-likelihood (QML) estimation for a zero-mean causal CARMA(p,q) process Y defined by the formal stochastic differential equation,

$$a(D)Y_t = b(D)DL_t,$$

based on observations $Y_1^\Delta, \ldots, Y_n^\Delta$ made at times $\Delta, \ldots, n\Delta$. Recall that causality implies that the zeros $\lambda_1, \ldots, \lambda_p$, of the polynomial $a(z)$ (which are the same as the eigenvalues of the matrix A in Section 1) have strictly negative real parts. QML estimation is estimation based on the (not necessarily valid) pretense that the process Y is Gaussian. We shall therefore take L, in the analysis of this section, to be νB where $0 < \nu < \infty$ and B is standard Brownian motion.

From (1.4) and (1.5) we immediately obtain the following discrete state-space representation of $Y_1^\Delta, \ldots, Y_n^\Delta$.

$$Y_i^\Delta = \mathbf{b}^T \mathbf{S}_i, i = 1, \ldots, n, \tag{4.10}$$

and

$$\mathbf{S}_{i+1} = e^{A\Delta}\mathbf{S}_i + \nu \int_{i\Delta}^{(i+1)\Delta} e^{A(u-i\Delta)} \mathbf{e}_p dB_u, \quad i = 1, 2, \ldots, \tag{4.11}$$

where $\mathbf{S}_i := \mathbf{X}_{i\Delta}$ and $\mathbf{X}_t$ is the state-vector in (1.4). The causality assumption implies that $\mathbf{X}_t = \int_0^\infty e^{A(t-u)}\mathbf{e}\, dB_u$, so that $\mathbf{S}_i$ has the multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\nu^2 \int_0^\infty e^{At}\mathbf{e}_p\mathbf{e}_p^T e^{A^T t}\, dt$.

The *observation equation* (4.10) and *state equation* (4.11) are in precisely the form required for application of the discrete-time Kalman recursions for calculation of the best one-step linear predictors of $Y_1^\Delta, \ldots, Y_n^\Delta$, their mean-squared errors, and hence the Gaussian likelihood of the observations. This is a special case of the state-space representation used by Jones (1981) for fitting Gaussian CARMA processes to irregularly spaced observations. A similar approach was also used in a more general setting by Bergstrom (1985).

To apply the Kalman recursions to (4.10) and (4.11) we rewrite them in the notation of Brockwell and Davis (1991; Proposition 12.2.2) as

$$Y_i^{\Delta} = \mathbf{b}^T \mathbf{S}_i, \quad i = 1, 2, \ldots, \tag{4.12}$$

and

$$\mathbf{S}_{i+1} = F\mathbf{S}_i + \mathbf{V}_i, i = 1, 2, \ldots, \tag{4.13}$$

where

$$F = e^{A\Delta} \quad \text{ and } \quad \mathbf{V}_i = \nu \int_{i\Delta}^{(i+1)\Delta} e^{A(u-i\Delta)} \mathbf{e}_p \, dB_u.$$

The random vectors $\mathbf{V}_i$ are iid multivariate Gaussian with mean $\mathbf{0}$ and covariance matrix

$$Q = \nu^2 \int_0^{\Delta} e^{Au} \mathbf{e}_p \mathbf{e}_p^T e^{A^T u} \, du.$$

Now let $\hat{\mathbf{S}}_i$ and $\hat{Y_i^{\Delta}}$ denote the orthogonal projections in L^2 of $\mathbf{S}_i$ and Y_i^{Δ} onto the closed linear span of $\{1, Y_j^{\Delta},\ j < i\}$. Then by the Kalman recursions (see Brockwell and Davis (1991; Proposition 12.2.2), the predictors $\hat{\mathbf{S}}_i$ and their error covariance matrices $\Omega_i = \mathbb{E}[(\mathbf{S}_i - \hat{\mathbf{S}}_i)(\mathbf{S}_i - \hat{\mathbf{S}}_i)^T]$ are uniquely determined by the initial conditions,

$$\hat{\mathbf{S}}_1 = 0, \quad \Pi_1 = \sigma^2 \int_0^{\infty} e^{At} \mathbf{e}_p \mathbf{e}_p^T e^{A^T t} \, dt, \quad \Psi_1 = 0_{p \times p}, \quad \Omega_1 = \Pi_1 - \Psi_1,$$

and the recursions,

$$\begin{cases} \Delta_i & = \mathbf{b}^T \Omega_i \mathbf{b} \\ \Theta_i & = F \Omega_i \mathbf{b} \\ \Pi_{i+1} & = F \Pi_i F^T + Q \\ \Psi_{i+1} & = F \Psi_i F^T + \Theta_i \Delta_i^{-1} \Theta_i^T \\ \Omega_{i+1} & = \Pi_{i+1} - \Psi_{i+1} \end{cases}$$

and

$$\hat{\mathbf{S}}_{i+1} = F_i \hat{\mathbf{S}}_i + \Theta_i \Delta_i^{-1} (Y_i^{\Delta} - \mathbf{b}^T \hat{\mathbf{S}}_i).$$

The best linear predictor of Y_i^{Δ} in terms of $\{1, Y_k, k < i\}$ is

$$\hat{Y_i^{\Delta}} = \mathbf{b}^T \hat{\mathbf{S}}_i$$

and its mean squared error is

$$\mathbb{E}[(Y_i^{\Delta} - \hat{Y_i^{\Delta}})^2] = \mathbf{b}^T \Omega_i \mathbf{b}.$$

Notice that we can write

$$\mathbb{E}[(Y_i^{\Delta} - \hat{Y_i^{\Delta}})^2] = r_{i-1}\nu^2,\ i = 1, 2, \ldots, \tag{4.14}$$

where the coefficients $r_0, r_1, \ldots,$ are equal to the corresponding mean squared errors when ν^2 is replaced in the Kalman recursions by 1.

The best linear predictors $\hat{Y_i^{\Delta}}$ are not affected by the value of ν so we can use the Kalman recursions with ν set equal to 1 to determine both the predictors and the constants r_i in (4.14).

The likelihood of the observations $Y_1^{\Delta}, \ldots, Y_n^{\Delta}$ under the Gaussian CARMA model with parameters $\mathbf{a} := (a_1, \ldots, a_p)^T$, $\mathbf{b} := (b_0, \ldots, b_{q-1})^T$ and ν^2 can then be written as (cf. Brockwell and Davis (1991; equ. (8.7.4)),

$$L\left(\mathbf{a}, \mathbf{b}, \nu^2\right) = \frac{1}{\sqrt{\left(2\pi\nu^2\right)^n r_0 \cdots r_{n-1}}} \exp\left\{ -\frac{1}{2\nu^2} \sum_{j=1}^{n} \frac{\left(Y_j^{\Delta} - \hat{Y_j^{\Delta}}\right)^2}{r_{j-1}} \right\}. \tag{4.15}$$

Differentiating $\log L\left(\mathbf{a}, \mathbf{b}, \nu^2\right)$ partially with respect to ν^2 and noting that $\hat{Y_j^{\Delta}}$ and r_j are independent of ν^2, we find that the QML estimators $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, and $\hat{\nu}^2$ satisfy the following equations:

$$\hat{\nu}^2 = n^{-1} S\left(\hat{\mathbf{a}}, \hat{\mathbf{b}}\right), \tag{4.16}$$

where

$$S\left(\mathbf{a}, \mathbf{b}\right) = \sum_{j=1}^{n} \left(Y_j^{\Delta} - \hat{Y_j^{\Delta}}\right)^2 / r_{j-1}, \tag{4.17}$$

and $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$ are the values of $\mathbf{a}$, $\mathbf{b}$ that minimize

$$\ell(\mathbf{a}, \mathbf{b}) = \log\left(n^{-1} S(\mathbf{a}, \mathbf{b})\right) + n^{-1} \sum_{j=1}^{n} \log r_{j-1}. \tag{4.18}$$

For any given parameter vectors $\mathbf{a}$ and $\mathbf{b}$, the reduced likelihood, $\ell(\mathbf{a}, \mathbf{b})$, can be computed from the data and the Kalman recursions as described above. A numerical search algorithm is then used to determine the values of $\mathbf{a}$ and $\mathbf{b}$ which minimize $\ell(\mathbf{a}, \mathbf{b})$. These are the QML estimators of $\mathbf{a}$ and $\mathbf{b}$ respectively. The corresponding QML estimator of ν^2 is obtained by substituting these values in (4.16).

Remark 4.1. Notice that the reduced likelihood of the data depends only on $\mathbf{a}$, $\mathbf{b}$ and the observed data, so that the QML estimation of $\mathbf{a}$ and $\mathbf{b}$ can be carried out regardless of the nature of the driving Lévy process L, and in particular when $\mathbb{E}(L_1^2) = \infty$ as in Garcia et al. (2011). □

4.1 Asymptotic behaviour of the estimators

As in Corollary 3.6 we shall denote by Θ the set of parameter vectors $\boldsymbol{\theta} = (\boldsymbol{\lambda}^T, \boldsymbol{\mu}^T)^T$ such that the components $\lambda_1, \ldots, \lambda_p$ and $\mu_1, \ldots, \mu_q$ have strictly negative real parts, $e^{\lambda_i} \neq e^{\lambda_j}$ for $i \neq j$ and $\lambda_i \neq \mu_j$ for all pairs i and j. We also introduce the parameters

$$\xi_i := e^{\lambda_i \Delta}, i = 1, \ldots, p, \tag{4.19}$$

and

$$\eta_i, i = 1, \ldots, p-1, \tag{4.20}$$

where the parameters η_i, $i = 1, \ldots, p-1$, were defined in Theorem 2.1. The sampled sequence Y^Δ then satisfies the causal invertible ARMA equations,

$$\prod_{i=1}^{p}(1 - \xi_i B)Y_n^\Delta = \prod_{i=1}^{p-1}(1 - \eta_i B)Z_n, \tag{4.21}$$

where the sequence Z is iid noise with mean zero and variance as specified in Theorem 2.1.

Introducing the parameter vectors,

$$\boldsymbol{\psi} := (\boldsymbol{\xi}^T, \boldsymbol{\eta}^T)^T = T\boldsymbol{\theta},\ \boldsymbol{\theta} \in \Theta,$$

where

$$\boldsymbol{\xi} := (\xi_1 \ \cdots \ \xi_p)^T \ \text{ and } \ \boldsymbol{\eta} := (\eta_1 \ \cdots \ \eta_{p-1})^T,$$

we recall from Corollary 3.6 that for any $\boldsymbol{\theta}_0 \in \Theta$ the restriction T_0 of the mapping T to $\Theta_0 := \{\boldsymbol{\theta} \in \Theta : |\mathrm{Im}(\Delta^{-1}\log \xi_j - \lambda_{0j})| < \pi,\ j = 1, \ldots, p\}$ maps Θ_0 bijectively onto $T_0\Theta_0 \subset \{\boldsymbol{\psi} : \max_j |\psi_j| < 1\}$. Values of $\boldsymbol{\theta}$ (uniquely related in the obvious way to the coefficient vectors $\mathbf{a}$ and $\mathbf{b}$) which minimize the reduced likelihood of the sampled observations will be mapped into values of $\boldsymbol{\psi}$ which minimize the reduced likelihood of the observations of Y^Δ under the model (4.21).

Let $\boldsymbol{\theta}_0$ be the true parameter vector of the CARMA process and let $\boldsymbol{\psi}_0$ denote $T_0\boldsymbol{\theta}_0$.

By the results of Schlemm and Stelzer (2012) we know that the QML estimator $\hat{\boldsymbol{\psi}}$ of $\boldsymbol{\psi}_0$ is consistent and that

$$\sqrt{n}(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0) \Rightarrow \mathrm{N}(\mathbf{0}, I(\boldsymbol{\psi}_0)^{-1}),\ \text{as } n \to \infty, \tag{4.22}$$

where $I(\boldsymbol{\psi})$ is the information matrix for $\boldsymbol{\psi}$ per observation of the Δ-sampled sequence. Knowing the mapping T_0 and its inverse, we can use this result to deduce the asymptotic distribution of the QML estimator of $\boldsymbol{\theta}_0$ and hence that of of $(a_1, \ldots, a_p, b_0, \ldots, b_{q-1})^T$.

First we need a generalization of the expression for $I(\boldsymbol{\psi})$ given by Box et al. (2016; eq. (7.2.21) to allow for possibly complex components of $\boldsymbol{\psi}$. Noting that our definition of $I(\boldsymbol{\psi})$

is *per observation* and using the definition of Liu et al. (2014; eq. (3)), we obtain the Hermitian information matrix for the possibly complex-valued parameters $\xi_1, \ldots, \xi_p, \eta_1, \ldots, \eta_q$,

$$I(\boldsymbol{\psi}) = \begin{bmatrix} (1-\xi_1\overline{\xi}_1)^{-1} & \cdots & (1-\xi_p\overline{\xi}_1)^{-1} & -(1-\eta_1\overline{\xi}_1)^{-1} & \cdots & -(1-\eta_{p-1}\overline{\xi}_1)^{-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ (1-\xi_1\overline{\xi}_p)^{-1} & \cdots & (1-\xi_p\overline{\xi}_p)^{-1} & -(1-\eta_1\overline{\xi}_p)^{-1} & \cdots & -(1-\eta_{p-1}\overline{\xi}_p)^{-1} \\ -(1-\xi_1\overline{\eta}_1)^{-1} & \cdots & -(1-\xi_p\overline{\eta}_1)^{-1} & (1-\eta_1\overline{\eta}_1)^{-1} & \cdots & (1-\eta_{p-1}\overline{\eta}_1)^{-1} \\ \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ -(1-\xi_1\overline{\eta}_{p-1})^{-1} & \cdots & -(1-\xi_p\overline{\eta}_{p-1})^{-1} & (1-\eta_1\overline{\eta}_{p-1})^{-1} & \cdots & (1-\eta_{p-1}\overline{\eta}_{p-1})^{-1} \end{bmatrix}. \tag{4.23}$$

Since

$$\log \xi_j = \Delta\lambda_j, \; j = 1, \ldots, p,$$

and

$$\mu_j = f_j(\xi_1, \ldots, \xi_p, \eta_1, \ldots, \eta_{p-1}), \; j = 1, \ldots, p-1,$$

where the functions f_j are determined by Corollary 3.6, the Jacobian of T_0^{-1} is

$$J = \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\psi}} = \begin{bmatrix} \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{\xi}} & 0 \\ \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\xi}} & \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\eta}} \end{bmatrix}, \tag{4.24}$$

where

$$\frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{\xi}} = \Delta^{-1} \begin{bmatrix} e^{-\lambda_1\Delta} & 0 & \cdots & 0 \\ 0 & e^{-\lambda_2\Delta} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & e^{-\lambda_p\Delta} \end{bmatrix},$$

$$\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\xi}} = \left[\frac{\partial \mu_i}{\partial \xi_j}\right]_{q\times p} \quad \text{and} \quad \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\eta}} = \left[\frac{\partial \mu_i}{\partial \eta_j}\right]_{q\times(p-1)}.$$

It then follows from (4.22) that the QML estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}_0$ satisfies

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \Rightarrow \mathrm{N}(\mathbf{0}, JI(\boldsymbol{\psi}_0)^{-1}\overline{J}^T), \text{ as } n \to \infty, \tag{4.25}$$

with $I(\boldsymbol{\psi})$ as in (4.23).

The coefficients a_i, $i = 1, \ldots, p$ and b_j, $j = 1, \ldots, q$ are expressed in terms of the zeros $\lambda_1, \ldots, \lambda_p$ and $\mu_1, \ldots, \mu_q$ by the familiar relations,

$$a_i = (-1)^i \sum \lambda_{j_1} \ldots \lambda_{j_i}, \; i = 1, \ldots, p, \tag{4.26}$$

where the sum is over all distinct subsets $\{j_1, \ldots, j_i\}$ of $\{1, \ldots, p\}$, and

$$b_{q-i} = (-1)^i \sum \mu_{j_1} \ldots \mu_{j_i}, \; i = 1, \ldots, q, \tag{4.27}$$

where the sum is over all distinct subsets $\{j_1, \ldots, j_i\}$ of $\{1, \ldots, q\}$. Defining

$$\boldsymbol{\beta} := (a_1 \ \cdots \ a_p \ b_0 \ \cdots \ b_{q-1}),$$

we find from (4.25) that the QML estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}_0$ satisfies

$$\sqrt{n\Delta}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \Rightarrow \mathrm{N}(\mathbf{0}, \Delta K J I(\boldsymbol{\psi}_0)^{-1} \overline{J}^T \overline{K}^T), \text{ as } n \to \infty \text{ with } \Delta \text{ fixed}, \tag{4.28}$$

where

$$K := \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{\lambda}} & 0 \\ 0 & \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{\mu}} \end{bmatrix} \tag{4.29}$$

can be readily evaluated from (4.26) and (4.27).

Example 4.2. CARMA(2,1)

For the CARMA(2,1) process defined in Example 3.11 with parameters $\lambda_1 = -0.5$, $\lambda_2 = -1.0$ and $\mu = -0.25$, the parameters of the sequence sampled at intervals of length $\Delta = 1$ are $\xi_1 = e^{-.5}$, $\xi_2 = e^{-1}$ and, from Theorem 2.1, $\eta = 0.7595699$. The coefficient G_1 in Theorem 3.5 is given by

$$G_1 = \frac{\xi_1(1 - \eta\xi_1)(1 - \eta\xi_1^{-1})}{(1 - \xi_1^2)(1 - \xi_1\xi_2)(\xi_1 - \xi_2)}$$

and the coefficient G_2 is the same with ξ_1 and ξ_2 interchanged.

By Corollary 3.6 the moving average zero μ can be expressed in terms of ξ_1, ξ_2 and η as

$$\mu = \Delta^{-1} \left[\log \xi_1 \log \xi_2 \frac{G_1 \log \xi_2 + G_1 \log \xi_1}{G_1 \log \xi_1 + G_2 \log \xi_2} \right]^{1/2}.$$

From this expression it is easy to evaluate numerically the derivatives $\frac{\partial \mu}{\partial \xi_1}$, $\frac{\partial \mu}{\partial \xi_2}$ and $\frac{\partial \mu}{\partial \eta}$ at $\xi_1 = e^{-0.5}$, $\xi_2 = e^{-1.0}$ and $\eta = 0.7595699$, and hence the Jacobian matrix in (4.24), which takes the value

$$J = \begin{bmatrix} 1.64872 & 0 & 0 \\ 0 & 2.71828 & 0 \\ -0.24747 & -0.15210 & 1.57833 \end{bmatrix}.$$

The matrix K in (4.29) is easily evaluated from (4.26) and (4.27) as

$$K = \begin{bmatrix} -1 & -1 & 0 \\ \lambda_2 & \lambda_1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 \\ -1 & -.5 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

At $\boldsymbol{\psi} = (e^{-0.5} \ e^{-1.0} \ .9752813889)^T$ the matrix $I(\boldsymbol{\psi})$ is easily evaluated from (4.23) as

$$I(\boldsymbol{\psi}) = \begin{bmatrix} 1.58198 & 1.28722 & -1.85426 \\ 1.28722 & 1.15652 & -1.38779 \\ -1.85426 & -1.38779 & 2.36377 \end{bmatrix}.$$

The matrix $\Delta KJI(\boldsymbol{\psi})^{-1}\overline{J}^T\overline{K}^T$ in (4.28) is therefore

$$\Delta KJI(\boldsymbol{\psi})^{-1}\overline{J}^T\overline{K}^T = \begin{bmatrix} 22.0217 & 15.7619 & -9.27685 \\ 15.7619 & 66.7833 & -41.7082 \\ -9.27685 & -41.7082 & 27.1195 \end{bmatrix}, \text{ if } \Delta = 1.0 \tag{4.30}$$

For the same CARMA model sampled at intervals of lengths $\Delta = 0.1$ and $\Delta = 0.01$, analogous calculations give

$$\Delta KJI(\boldsymbol{\psi})^{-1}\overline{J}^T\overline{K}^T = \begin{bmatrix} 11.1168 & 15.8567 & -9.8981 \\ 15.8567 & 34.4869 & -20.5699 \\ -9.8981 & -20.5699 & 12.8352 \end{bmatrix}, \text{ if } \Delta = 0.1 \tag{4.31}$$

and

$$\Delta KJI(\boldsymbol{\psi})^{-1}\overline{J}^T\overline{K}^T = \begin{bmatrix} 11.0056 & 15.9804 & -9.98774 \\ 15.9804 & 33.5816 & -20.0510 \\ -9.98774 & -20.0510 & 12.5319 \end{bmatrix}, \text{ if } \Delta = 0.01. \tag{4.32}$$

Equation (4.30) shows that if observations of this CARMA process are made at unit intervals over a large time interval $T(= n\Delta)$, the approximate variances of the components $\hat{a_1}$, $\hat{a_2}$ and $\hat{b_0}$ of the QML estimator $\hat{\boldsymbol{\beta}}$ are $22.0/T, 66.8/T$ and $27.1/T$ respectively. If, for the same time interval, observations are made at intervals $\Delta = .1$, equation (4.31) shows that these variances are reduced substantially to $11.1/T$, $34.5/T$ and $12.8/T$ respectively. However equation (4.32) shows that there is only a very slight further improvement, to $11.0/T$, $33.6/T$ and $12.5/T$ respectively, gained by using observations made at intervals of length $\Delta = .01$. This indicates that there is little point in taking observations at time intervals less than 0.1 for this model. Lower bounds for these asymptotic variances can be found from the results of Pham-Din-Tuan (1977), as discussed below.

4.1.1 High frequency observations

For the causal invertible CARMA(p, q) process defined by (1.3), the information matrix $I(\boldsymbol{\psi})$ in (4.23) for the parameters of the sampled sequence can be re-expressed in terms of the parameters λ_i and μ_i of the underlying CARMA process. By appropriately ordering the rows and columns, we then readily find that

$$\Delta I(\boldsymbol{\psi}) \to M = [M_{ij}]_{i,j=1}^{2p-1}, \text{ as } \Delta \to 0,$$

where

$$M_{ij} = \begin{cases} -(\overline{\lambda_i} + \lambda_j)^{-1}, & 1 \le i \le p,\ 1 \le j \le p, \\ (\overline{\lambda_i} + \mu_j)^{-1}, & 1 \le i \le p,\ p+1 \le j \le p+q, \\ (\overline{\mu_i} + \lambda_j)^{-1}, & p+1 \le i \le p+q,\ 1 \le j \le p, \\ -(\overline{\mu_i} + \mu_j)^{-1}, & p+1 \le i \le p+q,\ p+1 \le j \le p+q, \\ 0, & \text{otherwise}. \end{cases} \qquad (4.33)$$

With corresponding ordering of rows and columns,

$$\Delta J \to L = [L_{ij}]_{i=1,\ j=1}^{p+q,\ 2p-1},\ \text{as } \Delta \to 0, \qquad (4.34)$$

where

$$L_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise}. \end{cases}$$

Using the inversion formula for block matrices, we find that the upper-left $(p+q) \times (p+q)$ truncation of $(\Delta I(\boldsymbol{\psi}))^{-1}$ converges as $\Delta \to 0$ to M_{p+q}^{-1}, the inverse of the upper-left $(p+q) \times (p+q)$ truncation of the matrix M defined in (4.33).

From these limits we find that

$$\lim_{\Delta \to 0} \Delta J I(\boldsymbol{\psi})^{-1} \overline{J}^T = M_{p+q}^{-1}.$$

and recalling from (4.25) that for any fixed $\Delta > 0$,

$$\sqrt{n\Delta}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \Rightarrow \mathrm{N}(\mathbf{0}, \Delta J I(\boldsymbol{\psi})^{-1} \overline{J}^T),\ \text{as } n \to \infty, \qquad (4.35)$$

we see that the asymptotic covariance matrix on the right of (4.35) converges as $\Delta \to 0$ to M_{p+q}^{-1}. Correspondingly, for the QML estimator of $\boldsymbol{\beta} = (a_1\ \dots\ a_p\ b_0\ \dots\ b_{q-1})^T$ we have

$$\sqrt{n\Delta}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow \mathrm{N}(\mathbf{0}, \Delta K J I(\boldsymbol{\psi})^{-1} \overline{J}^T \overline{K}^T),\ \text{as } n \to \infty, \qquad (4.36)$$

and the asymptotic covariance in (4.36) converges as $\Delta \to 0$ to $K M_{p+q}^{-1} \overline{K}^T$.

This result gives an easily calculated approximation, $K M_{p+q}^{-1} \overline{K}^T$, to the variance of the asymptotic normal distribution of $\sqrt{n\Delta}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ in(4.36) when Δ is small. Applying this approximation to Example 4.2, we find that

$$K M_{p+q}^{-1} \overline{K}^T = \begin{bmatrix} 11.00 & 16.00 & -10.00 \\ 16.00 & 33.50 & -20.00 \\ -10.00 & -20.00 & 12.50 \end{bmatrix},$$

which is a good approximation to the corresponding exact matrices in (4.31) and (4.32).

Pham-Dinh-Tuan (1977) showed that if the CARMA process Y is Gaussian, the exact maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta} = (\lambda_1 \ \ldots \ \lambda_p \ \mu_1 \ \ldots \ \mu_q)^T$ based on $(Y_t)_{t\in[0,T]}$ satisfies

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \Rightarrow \mathrm{N}(\mathbf{0}, R^{-1}), \text{ as } T \to \infty, \tag{4.37}$$

where

$$R_{jk} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta_j} \left[\frac{a(i\omega)}{b(i\omega)} \right] \frac{\partial}{\partial \theta_k} \left[\frac{a(-i\omega)}{b(-i\omega)} \right] \left| \frac{b(i\omega)}{a(i\omega)} \right|^2 d\omega, \ j,k = 1, \ldots, p+q.$$

Using contour integration it can be shown that, if $\lambda_1, \ldots, \lambda_p$, are distinct,

$$R = M_{p+q}.$$

Hence the asymptotic covariance in (4.35) converges as $\Delta \to 0$ to the asymptotic covariance in (4.37) based on *continuous* observation of the corresponding Gaussian CARMA process on the time interval $[0, T]$. Similarly the asymptotic covariance in (4.36) converges as $\Delta \to 0$ to the asymptotic covariance $KR^{-1}\overline{K}^T$ of $\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, where $\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimator of $\boldsymbol{\beta}$ based on $(Y_t)_{t\in[0,T]}$ when Y is Gaussian. Pham-Dinh-Tuan (1977) also gives an algorithm for approximating the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ in the Gaussian case which can be used as an initial approximation for numerical QML estimation via the Kalman recursions when Δ is small.

5 Conclusions

We have established a simple algorithm for determining the parameters of the Δ-sampled ARMA sequence of a second-order CARMA process and derived conditions under which an ARMA(p, q) sequence with $q < p$ has the autocovariance function of a Δ-sampled CARMA(p, r) process for some $r < p$, determining, under these conditions, the parameters of such a CARMA process. The results were used to determine the asymptotic behaviour of the QML estimators of the coefficients of a CARMA process based on observations made at time intervals of length Δ. These results provide information, for any given parameter vectors $\mathbf{a}$ and $\mathbf{b}$, on the frequency of observations on $[0, T]$ beyond which only marginal improvement in estimation accuracy is possible. The asymptotic results as $\Delta \to 0$ were also related to the corresponding results of Pham-Dinh-Tuan (1977) on maximum likelihood estimation for a Gaussian CARMA process based on continuous observation of the process on $[0, T]$.

Acknowledgments

We are indebted to two referees whose comments led to considerable improvements on the original manuscript.

References

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

Box, G.E.P, Jenkins, G.M., Reinsel, G.C. and Ljung, G.M., 2016, Time Series Analysis, Forecasting and Control, 5th edition. Prentice-Hall, Englewood Cliffs, New Jersey.

Brockwell, P.J., 2001, Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.* **52**, 1–18.

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

Brockwell, P.J. and Davis, R.A., 1991, Time Series: Theory and Methods, 2nd edition. New York, Springer.-Verlag.

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

Brockwell, P.J., Ferrazzano, V. and C. Klüppelberg, 2012, High frequency sampling of a continuous-time ARMA process. *J. Time Series Analysis* **33**, 152–160.

Brockwell, P.J., Ferrazzano, V. and Klüppelberg C. 2013, High-frequency sampling and kernel estimation for continuous-time moving average processes. *J.Time Series Analysis* **35**, 385–404.

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.

Garcia, I., Klüppelberg, C. and Müller, G., 2011, Estimation of stable CARMA models with an application to electricity spot prices. Statistical Modelling 11, 2011, 447–470.

Huzii, M., 2006, Embedding a Gaussian discrete-time autoregressive moving average process in a Gaussian continuous-time autoregressive moving average process. *Journal of Time Series Analysis*, 28, 498-520.

Jones, R.H., 1981, Fitting a continuous time autoregression to discrete data. In *Applied Time Series Analysis II* (ed. D.F. Findley). Academic Press, New York, 651–682.

Liu, W., Wang, Y. and Xie, W., 2014, Fisher information matrix, Rao test, and Wald test for complex-valued signals and their applications. *Signal Processing*, 94, 1–5.

Pham-Dinh-Tuan, 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.

Phillips, P.C.B., 1973, The problem of identification in finite parameter continuous time models. *J. of Econometrics* **1**, 351–362.

Robinson, P.M., 1977, The construction and estimation of continuous time models and discrete approximations in econometrics, *J. Econometrics*, **6**, 173–197.

Schlemm, E. and Stelzer, R., 2012, Quasi maximum likelihood estimation for strongly mixing state space models and multivariate Lvy-driven CARMA processes, *Electron. J. Stat.*, **6**, 21852234.

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.