

Prediction of Lévy-driven CARMA processes

Peter J. Brockwell* Alexander Lindner†

Abstract

The conditional expectations, $\mathbb{E}(Y(h)|Y(u), -\infty < u \leq 0)$ and $\mathbb{E}(Y(h)|Y(u), -M \leq u \leq 0)$ with $h > 0$ and $0 < M < \infty$ are determined for a continuous-time ARMA (CARMA) process $(Y(t))_{t \in \mathbb{R}}$ driven by a Lévy process L with $\mathbb{E}|L(1)| < \infty$. If $\mathbb{E}(L(1)^2) < \infty$ these are the minimum mean-squared error predictors of $Y(h)$ given $(Y(t))_{t \leq 0}$ and $(Y(t))_{-M \leq t \leq 0}$ respectively. Conditions are also established under which the sample-path of L can be recovered from that of Y , both when Y is causal and strictly stationary and (without these assumptions) when L is a pure-jump Lévy process. When $\mathbb{E}(L(1)^2) < \infty$ and Y is causal and strictly stationary the best *linear* predictors $P(Y(h)|Y(u), u \leq 0)$ and $P(Y(h)|Y(-n\Delta), n \in \mathbb{N})$ are also determined, the latter yielding a simple algorithm for determining the parameters of the ARMA process obtained by sampling the CARMA process at regular intervals.

JEL Classification: C02, G17

Keywords: Prediction, Lévy process, CARMA process, continuous time

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 $L(0) = 0$. For integers p and q such that $p > q$, we define a (complex valued) CARMA(p, q) process $Y = (Y(t))_{t \in \mathbb{R}}$, driven by L , by the equation

$$Y(t) = \mathbf{b}'\mathbf{X}(t), \quad t \in \mathbb{R}, \quad (1)$$

*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.

†Institut für Mathematische Stochastik, TU Braunschweig, Pockelsstraße 14, D-38106 Braunschweig, Germany a.lindner@tu-bs.de

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

$$d\mathbf{X}(t) = A\mathbf{X}(t) dt + \mathbf{e}_p dL(t), \quad (2)$$

or equivalently

$$\mathbf{X}(t) = e^{A(t-s)}\mathbf{X}(s) + \int_{(s,t]} e^{A(t-u)}\mathbf{e}_p dL(u), \quad \forall s \leq t \in \mathbb{R}, \quad (3)$$

with

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

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

Integration by parts allows us to rewrite (3) in the form,

$$\mathbf{X}(t) = e^{A(t-s)}\mathbf{X}(s) + \left[L(t)I_p - L(s)e^{A(t-s)} + \int_s^t L(u)Ae^{A(t-u)}du \right] \mathbf{e}_p, \quad (4)$$

where I_p denotes the $p \times p$ identity matrix. Thus we can interpret the integral in (3) as a pathwise integral and in fact all the integrals appearing in this paper can be interpreted pathwise. Under conditions specified in Sections 2 and 3 we shall show how to recover the sample-path of L from either \mathbf{X} or Y . In the Gaussian case, Pham (1977) used spectral analysis to express L in terms of \mathbf{X} and mean-square derivatives of Y , but his argument is restricted to square integrable CARMA processes and their second-order properties.

The equations (1) and (2) constitute the state-space representation of the formal p^{th} -order stochastic differential equation,

$$a(D)Y(t) = b(D)DL(t), \quad (5)$$

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

$$a(z) = z^p + a_1z^{p-1} + \cdots + a_p, \quad (6)$$

and

$$b(z) = b_0 + b_1z + \cdots + b_{p-1}z^{p-1}.$$

Equation (5) is the natural continuous-time analogue of the p^{th} -order linear difference equations used to define a discrete-time ARMA process (see e.g. Brockwell and Davis

(1991)). However, since the derivatives on the right-hand side of (5) do not exist as random functions, we base the definition on the state-space formulation (1) and (2).

Brockwell and Lindner (2009), generalizing results of Wolfe (1982) and Sato and Yamazato (1984) for the Lévy-driven Ornstein-Uhlenbeck equation, established necessary and sufficient conditions for the equations (1) and (2) to have a strictly stationary, not necessarily causal, solution $Y = (Y(t))_{t \in \mathbb{R}}$ and gave an explicit representation of the solution as an integral with respect to L . In Theorem 4.1 of that paper they also showed that there is no loss of generality in assuming that the polynomials $a(z)$ and $b(z)$ have no common zeroes. We shall therefore make this assumption whenever convenient.

The aim of this paper is to consider the prediction of $Y(h)$, $h \geq 0$, based on past observations. In Section 2, under causality, invertibility and stationarity conditions, we determine $\mathbb{E}(Y(h)|Y(u), -\infty < u \leq 0)$ and the best *linear* predictor $P(Y(h)|Y(u), -\infty < u \leq 0)$ when $\mathbb{E}(L(1))^2 < \infty$. In Section 3, under less restrictive conditions on $a(z)$ and $b(z)$ and without the stationarity assumption, but assuming that L has no Gaussian component, we determine $\mathbb{E}(Y(h)|Y(u), -M \leq u \leq 0)$. Finally in Section 4 we use the steady-state Kalman filter to determine the best linear predictor of $Y(h)$ based on the infinite discrete past $Y(-\Delta), Y(-2\Delta), Y(-3\Delta), \dots$, for $\Delta > 0$, assuming that $\mathbb{E}(L(1))^2 < \infty$ and that Y is strictly stationary and causal. With the aid of this predictor we establish a simple algorithm for calculating the parameters of the sampled ARMA process $(Y(n\Delta))_{n \in \mathbb{Z}}$.

2 Determination of $\mathbb{E}(Y(h)|Y(u), -\infty < u \leq 0)$ and $P(Y(h)|Y(u), -\infty < u \leq 0)$

2.1 Preliminaries

For $\mathbf{X}(t)$ satisfying (3) we shall also use the more explicit notation,

$$\mathbf{X}(t) = [X_0(t) \ X_1(t) \ \cdots \ X_{p-1}(t)]' \quad \text{and} \quad X(t) = X_0(t).$$

By (2), $X_j(t) - X_j(s) = \int_s^t X_{j+1}(u) du$ for $s \leq t$ and $j = 0, \dots, p-2$, and, since \mathbf{X} and hence X_{p-1} are càdlàg, it follows that X is $(p-2)$ -times differentiable with j^{th} derivative at time t , $X^{(j)}(t) = X_j(t)$, $j = 0, \dots, p-2$, while $X^{(p-2)} = X_{p-2}$ has right and left derivatives $X_{p-1}(t)$ and $X_{p-1}(t-)$, respectively. With the understanding that X_{p-1} will in general have jump discontinuities, we shall simply say that X is $(p-1)$ -times differentiable with j^{th} derivative

$$X^{(j)}(t) = X_j(t), \quad t \in \mathbb{R}, \quad j = 0, \dots, p-1. \quad (7)$$

To see the connection between the jumps of L and of Y we observe that by (4), $\Delta \mathbf{X}(t) := \mathbf{X}(t) - \mathbf{X}(t-) = L(t)\mathbf{e}_p$, so that

$$\Delta X^{(p-1)}(t) = \Delta X_{p-1}(t) = \Delta L(t).$$

On the other hand, since

$$Y(t) = b_0 X(t) + b_1 X^{(1)}(t) + \dots + b_{q-1} X^{(q-1)}(t) + X^{(q)}(t)$$

we see that Y is $(p-q-1)$ -times differentiable (again with the understanding that $Y^{(p-q-1)}$ will in general have jump discontinuities) and that

$$Y^{(p-q-1)}(t) = b_0 X^{(p-q-1)}(t) + b_1 X^{(p-q)}(t) + \dots + b_{q-1} X^{(p-2)}(t) + X^{(p-1)}(t).$$

Consequently,

$$\Delta Y^{(p-q-1)}(t) = \Delta X^{(p-1)}(t) = \Delta X_{p-1}(t) = \Delta L(t), \quad (8)$$

so that the jumps of L can be recovered from Y .

2.2 Recovery of the state vector

For the determination of the conditional expectation, the first step is the recovery of the realization of the state-vector process $(\mathbf{X}(t))_{t \leq 0}$ from that of $(Y(t))_{t \leq 0}$. As has been pointed out in Remark 4.3 in Brockwell and Lindner (2009), if $a(z)$ and $b(z)$ have common zeroes, then it may happen that a stationary CARMA process Y exists without a corresponding stationary state vector \mathbf{X} . Also, one can see from Theorem 4.2 in Brockwell and Lindner (2009) that in the case of common zeroes, there can be more than one state vector process leading to the same CARMA process Y . If $a(z)$ and $b(z)$ have no common zeroes, however this cannot happen as shown in the next proposition, which uses the concept of observability from linear systems theory (see e.g. Bernstein (2009), Section 12.3).

Proposition 2.1. *Let Y be a CARMA(p, q) process, driven by the Lévy process L , such that $a(z)$ and $b(z)$ have no common zeroes.*

- (i) *Then the pair (A, \mathbf{b}') is observable, i.e. $[\mathbf{b}' A' \mathbf{b} \ \dots \ (A')^{p-1} \mathbf{b}] \in \mathbb{C}^{p \times p}$ has rank p .*
- (ii) *Let $I = [\alpha, \beta)$ be a non-empty real interval. Let $(\mathbf{X}(t))_{t \in I}$ and $(\tilde{\mathbf{X}}(t))_{t \in I}$ be two \mathbb{C}^p -valued state vector processes leading both to the same $(Y_t)_{t \in I}$, i.e. both $(\mathbf{X}(t))_{t \in I}$ and $(\tilde{\mathbf{X}}(t))_{t \in I}$ satisfy (3) for $s, t \in I$, $s \leq t$, and $Y(t) = \mathbf{b}' \mathbf{X}(t) = \mathbf{b}' \tilde{\mathbf{X}}(t)$ for $t \in I$. Then $(\mathbf{X}(t))_{t \in I} = (\tilde{\mathbf{X}}(t))_{t \in I}$.*
- (iii) *If Y is strictly stationary, then so is \mathbf{X} .*

Proof. (i) By Corollary 12.3.3 in Bernstein (2009), the given definition of observability is equivalent to the one used there. Further, by the PBH-test (see Bernstein (2009),

Corollary 12.3.19), (A, \mathbf{b}') is observable if and only if $\begin{bmatrix} \lambda I_p - A \\ \mathbf{b}' \end{bmatrix} \in \mathbb{C}^{(p+1) \times p}$ has rank p for all $\lambda \in \mathbb{C}$. Hence, if (A, \mathbf{b}') is not observable, there must be $\lambda \in \mathbb{C}$ and $\mathbf{v} \in \mathbb{C}^p \setminus \{\mathbf{0}\}$ such that $(\lambda I_p - A)\mathbf{v} = \mathbf{0}$ and $\mathbf{b}'\mathbf{v} = 0$. The first equation means that λ is an eigenvalue of A , hence a zero of $a(z)$, and that \mathbf{v} is a right eigenvector of A . Hence \mathbf{v} must be of the form $\mathbf{v} = \gamma[1 \ \lambda \ \lambda^2 \ \cdots \ \lambda^{p-1}]'$ for some $\gamma \in \mathbb{C} \setminus \{0\}$. This and the second equation imply that $\gamma b(\lambda) = \mathbf{b}'\mathbf{v} = 0$, contradicting the assumption that $a(z)$ and $b(z)$ have no common zeroes.

(ii) Suppose that \mathbf{X} and $\tilde{\mathbf{X}}$ are two state vectors in \mathbb{C}^p with the same corresponding CARMA process driven by the same Lévy process, i.e. that $\mathbf{b}'\mathbf{X}(t) = \mathbf{b}'\tilde{\mathbf{X}}(t) = Y(t)$ for all $t \in I$. Since both \mathbf{X} and $\tilde{\mathbf{X}}$ satisfy (2), their difference satisfies $\frac{d}{dt}(\mathbf{X}(t) - \tilde{\mathbf{X}}(t)) = A(\mathbf{X}(t) - \tilde{\mathbf{X}}(t))$ and $\mathbf{b}'(\mathbf{X}(t) - \tilde{\mathbf{X}}(t)) = 0$ for $t \in I$. Since (A, \mathbf{b}') is observable, this implies that $\mathbf{X}(t) - \tilde{\mathbf{X}}(t) = \mathbf{0}$ for all $t \in I$ (see Bernstein (2009), Corollary 12.3.3).

(iii) If Y is strictly stationary, it follows from Theorem 3.3 in Brockwell and Lindner (2009) that there is a strictly stationary \mathbf{X} such that (1) holds, since $a(z)$ and $b(z)$ have no common zeroes. Since the state vector is unique by (ii), the result follows. \square

In general we would like to determine the state vector in an explicit form using only Y . This can be done when L is a pure jump Lévy process by using (8) to recover the jumps (and hence the increments of L on finite intervals) from Y . This will be demonstrated in Section 3. It can also be done if Y is stationary and the zeroes of $b(z)$ all have strictly negative real parts, as we shall now show.

If $q = 0$ then $\mathbf{b}' = [1 \ 0 \ \cdots \ 0]$ so, from (1), the first component of $\mathbf{X}(t)$ is $X(t) = Y(t)$ and the remaining $p - 1$ components are obtained by successive differentiation of $Y(t)$ with respect to t . It therefore suffices, for the remainder of this section, to deal with the case $q \geq 1$.

Before stating the result (Theorem 2.2 below), we need the preliminary definitions,

$$B := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -b_0 & -b_1 & -b_2 & \cdots & -b_{q-1} \end{bmatrix}, \quad \mathbf{X}_q(t) := \begin{bmatrix} X(t) \\ X_1(t) \\ \vdots \\ X_{q-2}(t) \\ X_{q-1}(t) \end{bmatrix}, \quad \text{and } \mathbf{e}_q := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{C}^q.$$

If $q = 1$, then $B := -b_0 = -1$. Observe that from (2) we have

$$dX_j(t) = X_{j+1}(t) dt, \quad j = 0, \dots, q - 1.$$

Using (1) we further see that

$$dX_{q-1}(t) = X_q(t) dt = (Y(t) - b_0 X_0(t) - \dots - b_{q-1} X_{q-1}(t)) dt.$$

Hence \mathbf{X}_q satisfies the differential equation

$$d\mathbf{X}_q(t) = B\mathbf{X}_q(t) dt + \mathbf{e}_q Y(t) dt.$$

The solution satisfies (cf. (3))

$$\mathbf{X}_q(t) = e^{B(t-s)}\mathbf{X}_q(s) + \int_s^t e^{B(t-u)}\mathbf{e}_q Y(u)du, \quad -\infty < s \leq t < \infty.$$

Multiplying this equation by $e^{B(s-t)}$, and using the convention $\int_s^t = -\int_t^s$ for $s > t$, we see that this equation also holds for $s > t$, i.e.

$$\mathbf{X}_q(t) = e^{B(t-s)}\mathbf{X}_q(s) + \int_s^t e^{B(t-u)}\mathbf{e}_q Y(u)du, \quad \forall s, t \in \mathbb{R}. \quad (9)$$

Suppose now that the distinct eigenvalues of B (or equivalently the zeroes of the polynomial $b(z)$) are $\rho_1, \rho_2, \dots, \rho_m$, with multiplicities μ_1, \dots, μ_m , respectively. Then clearly $\sum_{i=1}^m \mu_i = q$. Define $T(j)$ to be the $q \times \mu_j$ matrix of right eigenvectors and generalized eigenvectors associated with the eigenvalue ρ_j , specifically,

$$T(j) = [\mathbf{v}_1(\rho_j) \ \mathbf{v}_2(\rho_j) \ \cdots \ \mathbf{v}_{\mu_j}(\rho_j)],$$

where $\mathbf{v}_1(\rho_j)$ is the right eigenvector,

$$\mathbf{v}_1(\rho_j) = [1 \ \rho_j \ \rho_j^2 \ \cdots \ \rho_j^{q-1}]',$$

and $\mathbf{v}_i(\rho_j)$, $1 < i \leq \mu_j$, are generalized eigenvectors satisfying

$$(\rho_j I_q - B)\mathbf{v}_i(\rho_j) = -\mathbf{v}_{i-1}(\rho_j), \quad 1 < i \leq \mu_j,$$

where I_q denotes the $q \times q$ identity matrix.

Now define T to be the $q \times q$ matrix,

$$T := [T(1) \ T(2) \ \cdots \ T(m)]. \quad (10)$$

Then $T^{-1}BT = \Lambda$, where Λ has the Jordan canonical form,

$$\Lambda = \text{diag}\{\Lambda_1, \dots, \Lambda_m\}, \quad (11)$$

in which Λ_j is the $\mu_j \times \mu_j$ matrix, ρ_j if $\mu_j = 1$, and

$$\Lambda_j = \begin{bmatrix} \rho_j & 1 & 0 & \cdots & 0 \\ 0 & \rho_j & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \rho_j \end{bmatrix} \text{ if } \mu_j > 1.$$

Furthermore

$$e^{Bt} = Te^{\Lambda t}T^{-1} = T\text{diag}\{e^{\Lambda_1 t}, \dots, e^{\Lambda_m t}\}T^{-1}, \quad (12)$$

where $e^{\Lambda_j t} = e^{\rho_j t}$ if $\mu_j = 1$ and

$$e^{\Lambda_j t} = e^{\rho_j t} \begin{bmatrix} 1 & t & \cdots & t^{\mu_j-1}/(\mu_j-1)! \\ 0 & 1 & t & \cdots & t^{\mu_j-2}/(\mu_j-2)! \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad \text{if } \mu_j > 1.$$

Theorem 2.2. *Let $q \geq 1$. Suppose that the polynomials $a(z)$ and $b(z)$ have no zeroes on the imaginary axis and no common zeroes, that $\mathbb{E}(\log^+ |L(1)|) < \infty$, and that Y is the unique strictly stationary solution of (3) (it exists and is unique by Theorem 3.3 of Brockwell and Lindner (2009)). Let ρ_j denote the distinct zeroes of the polynomial $b(z)$ and let $\mu(\rho_j)$ be the multiplicity of ρ_j (so that $\sum_{j=1}^m \mu(\rho_j) = q$). Define the matrices T and Λ as in (10) and (11) and define*

$$C_j := \text{diag}\{M_{j1} \ M_{j2} \ \cdots \ M_{jm}\}, \quad j = 1, \dots, m,$$

where $M_{ji} = I_{\mu_j}$, the $\mu_j \times \mu_j$ identity matrix, if $i = j$, and $M_{ji} = 0$ otherwise. Then \mathbf{X}_q has the representation

$$\begin{aligned} \mathbf{X}_q(t) &= \sum_{j: \Re(\rho_j) < 0} \int_{-\infty}^t TC_j e^{\Lambda(t-u)} T^{-1} \mathbf{e}_q Y(u) du \\ &\quad - \sum_{j: \Re(\rho_j) > 0} \int_t^{\infty} TC_j e^{\Lambda(t-u)} T^{-1} \mathbf{e}_q Y(u) du, \quad t \in \mathbb{R}, \end{aligned} \quad (13)$$

where the improper integrals over $(-\infty, t]$ and $[t, \infty)$ defining $\mathbf{X}_q(t)$ exist as almost sure limits $\lim_{T \rightarrow \infty} \int_{-T}^t$ and $\lim_{T \rightarrow \infty} \int_t^T$, respectively. In particular, $\mathbf{X}_q(t)$ is uniquely determined by $(Y(u))_{u \in \mathbb{R}}$, and uniquely determined by $(Y(u))_{u \leq t}$ if the eigenvalues of $b(z)$ all have strictly negative real parts.

Proof. Multiplying both sides of (9) on the left by $C_j T^{-1}$ and using (12), we obtain the relations,

$$C_j T^{-1} \mathbf{X}_q(t) = C_j e^{\Lambda(t-s)} T^{-1} \mathbf{X}_q(s) + \int_s^t C_j e^{\Lambda(t-u)} T^{-1} \mathbf{e}_q Y(u) du. \quad (14)$$

Since Y is stationary, so is the state vector \mathbf{X} by Proposition 2.1 (iii), so that in particular $\mathbf{X}_q(s)$ and $\mathbf{X}_q(0)$ have the same distribution for all $s \in \mathbb{R}$. Hence, by Slutsky's lemma,

if $\Re(\rho_j) < 0$, then $C_j e^{\Lambda(t-s)} T^{-1} \mathbf{X}_q(s) \rightarrow 0$ in probability as $s \rightarrow -\infty$. In particular, $\int_{-\infty}^t C_j e^{\Lambda(t-u)} T^{-1} \mathbf{e}_q Y(u) du$ exists as a limit in probability and we have

$$C_j T^{-1} \mathbf{X}_q(t) = \int_{-\infty}^t C_j e^{\Lambda(t-u)} T^{-1} \mathbf{e}_q Y(u) du. \quad (15)$$

If $\Re(\rho_j) > 0$, we let $s \rightarrow \infty$ and obtain by the same argument

$$C_j T^{-1} \mathbf{X}_q(t) = - \int_t^{\infty} C_j e^{\Lambda(t-u)} T^{-1} \mathbf{e}_q Y(u) du, \quad (16)$$

where the limit again exists as a limit in probability. That the limits in (15) and (16) are actually almost sure follows from Proposition 2.6 below and the dominated convergence theorem. Using equations (15) and (16) and observing that $\mathbf{X}_q(t) = \sum_{j=1}^m T C_j T^{-1} \mathbf{X}_q(t)$, we get (13). \square

An alternative expression for the first component of the state-vector can be obtained by an argument analogous to that which was used by Brockwell and Lindner (2009), Proposition 3.2, to express the strictly stationary solution $Y(t)$ in terms of the Lévy process L . Thus, the unique solution for $\mathbf{X}_q(t)$ in terms of Y is given by the following Corollary. This in turn determines the complete p -dimensional state-vector, $\mathbf{X}(t)$, by differentiation via (7).

Corollary 2.3. *Under the conditions of Theorem 2.2, let $\mathbf{l}_b(t)$ and $\mathbf{r}_b(t)$ be the sums of the residues of the column vector $e^{zt} b^{-1}(z) [1 \ z \ \dots \ z^{q-1}]'$ at the zeroes of $b(\cdot)$ with strictly negative and strictly positive real parts respectively. Then*

$$\mathbf{l}_b(t) + \mathbf{r}_b(t) = e^{Bt} \mathbf{e}_q, \quad t \in \mathbb{R}, \quad (17)$$

and

$$\mathbf{X}_q(t) = \int_{-\infty}^t \mathbf{l}_b(t-u) Y(u) du - \int_t^{\infty} \mathbf{r}_b(t-u) Y(u) du, \quad t \in \mathbb{R}, \quad (18)$$

where the improper integrals over $(-\infty, t]$ and $[t, \infty)$ are the almost sure limits $\lim_{T \rightarrow \infty} \int_{-T}^t$ and $\lim_{T \rightarrow \infty} \int_t^T$ respectively. In particular the first component, $X(t) = [1 \ 0 \ \dots \ 0] \mathbf{X}_q(t)$, $t \in \mathbb{R}$, can be written as

$$X(t) = \int_{-\infty}^{\infty} \kappa(t-u) Y(u) du, \quad t \in \mathbb{R}, \quad (19)$$

where

$$\kappa(t) = \left(\sum_{\rho: \Re \rho < 0} \sum_{k=0}^{\mu(\rho)-1} d_{\rho k} t^k e^{\rho t} \mathbf{1}_{(0, \infty)}(t) - \sum_{\rho: \Re \rho > 0} \sum_{k=0}^{\mu(\rho)-1} d_{\rho k} t^k e^{\rho t} \mathbf{1}_{(-\infty, 0)}(t) \right), \quad t \in \mathbb{R}, \quad (20)$$

and $\sum_{k=0}^{\mu(\rho)-1} d_{\rho k} t^k e^{\rho t}$ is the residue of $z \mapsto e^{zt}/b(z)$ at the zero ρ of $b(\cdot)$, i.e.

$$\sum_{k=0}^{\mu(\rho)-1} d_{\rho k} t^k e^{\rho t} = \frac{1}{(\mu(\rho) - 1)!} [D_z^{\mu(\rho)-1} ((z - \rho)^{\mu(\rho)} e^{zt}/b(z))]_{z=\rho},$$

where D_z denotes differentiation with respect to z . (For a zero ρ with $\mu(\rho) = 1$ the last sum reduces to $e^{\rho t}/b'(\rho)$.) The $(j+1)$ st component $X^{(j)}(t)$, $j = 1, \dots, p-1$, of the complete p -component state-vector $\mathbf{X}(t)$ is obtained from the first component $X(t)$ by differentiation.

Proof. Define $\tilde{\mathbf{X}}_q(t)$ to be the right-hand side of (18). Again, the integral converges almost surely by Proposition 2.6 below. It is easy to see from the definition that $\tilde{\mathbf{X}}_q$ is strictly stationary. The proofs of (17) and the fact that $\tilde{\mathbf{X}}_q(t)$ satisfies (9) are the same as the corresponding proofs in Proposition 3.2 of Brockwell and Lindner (2009). Since it follows from the proof of Theorem 2.2 that (9) can have at most one strictly stationary solution, namely \mathbf{X}_q , we obtain $\mathbf{X}_q = \tilde{\mathbf{X}}_q$ and hence the desired representation of \mathbf{X}_q . \square

Remark 2.4. By determining the state-vector $\mathbf{X}(t)$ explicitly in terms of the process Y , Theorem 2.2 also explicitly specifies the increments of the background driving Lévy process L . Thus, multiplying each side of (2) on the left by \mathbf{e}'_p , we obtain

$$L(t) - L(s) = \mathbf{e}'_p \left(\mathbf{X}(t) - \mathbf{X}(s) - \int_s^t A\mathbf{X}(u) du \right), \quad t \geq s. \quad (21)$$

With $L(0) = 0$, this gives the complete Lévy process.

Remark 2.5. If all the eigenvalues of B have strictly negative real parts, then the second term in (18) is zero, and the state vector is $\mathbf{X}(t) = [X(t) \ X^{(1)}(t) \ \dots \ X^{(p-1)}(t)]'$, where

$$X^{(j)}(t) = \sum_{k=0}^{j-q} \kappa^{(j-1-k)}(0+) Y^{(k)}(t) + \int_{(-\infty, t)} \kappa^{(j)}(t-u) Y(u) du, \quad 0 \leq j \leq p-1. \quad (22)$$

(Here we have used the identities $\kappa^{(j)}(0+) = 0$ for $0 \leq j < q-1$ which follow from (17) since $\kappa(t) = [1 \ 0 \ \dots \ 0] e^{Bt} \mathbf{e}_q$ for $t > 0$.) Similarly,

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

At any specified time t , L (and hence \mathbf{X}) almost surely has no jump and hence $X^{(p-1)}(t) = X^{(p-1)}(t-)$ and $Y^{(p-q-1)}(t) = Y^{(p-q-1)}(t-)$ almost surely. Identifying random variables which are almost surely equal, we see that $\mathbf{X}(t)$ can be represented as a function of $\{Y(u), -\infty < u \leq t\}$, without the need for direct evaluation of $Y^{(p-q-1)}(t)$ which, as a right-derivative, is defined in terms of $\{Y(s), s > t\}$. Moreover, from (19) – (21), if $\mathbb{E}L(1)^2 < \infty$ the process Y is *invertible* in the sense that $L(t) - L(s)$ is in the closed linear span of $\{Y(u), -\infty < u \leq t\}$ with respect to L^2 -convergence. (A detailed proof of this fact is completely analogous to the proof of Theorem 2.8 below.)

We conclude this subsection by showing that strictly stationary CARMA processes grow more slowly than any exponentially increasing function, a fact which was used in the proofs of Theorem 2.2 and Corollary 2.3.

Proposition 2.6. *Let Y be a stationary CARMA process such that $a(z)$ has no zeroes on the imaginary axis. Then $\lim_{|t| \rightarrow \infty} e^{-\varepsilon|t|} Y(t) = 0$ almost surely for every $\varepsilon > 0$.*

Proof. Since Y is stationary it follows from Theorem 4.2 in Brockwell and Lindner (2009) that $\mathbb{E} \log^+ |L(1)| < \infty$. By the Lévy-Itô-decomposition, we may write $L(t) = L_1(t) - L_2(t) + L_3(t)$, where L_1 and L_2 are subordinators, $(L_3(t))_{t \geq 0}$ is a square integrable martingale with expectation zero (see Sato (1999), Theorem 19.2), and L_1, L_2 and L_3 are independent. Then $\lim_{t \rightarrow \infty} t^{-1} L_3(t) = 0$ a.s. by the strong law of large numbers. Furthermore $\mathbb{E} \log^+ |L_j(1)| < \infty$, $j = 1, 2$. Then for every $\varepsilon > 0$, $\int_0^t e^{-\varepsilon s} dL_1(s)$ and $\int_{-t}^0 e^{-\varepsilon|s|} dL_1(s)$ converge almost surely as $t \rightarrow \infty$ (see Jurek and Mason (1993), Theorem 3.6.6). Lemma 17.6. in Sato (1999) then shows that $\lim_{|t| \rightarrow \infty} e^{-\varepsilon|t|} L_1(t) = 0$ almost surely. The same reasoning applies to L_2 and consequently

$$\lim_{|t| \rightarrow \infty} e^{-\varepsilon|t|} L(t) = 0 \text{ a.s.}, \quad \varepsilon > 0. \quad (23)$$

By Theorem 4.2 in Brockwell and Lindner (2009), $Y(t)$ can be written as a (finite) linear combination of terms of the form $\int_{-\infty}^t (t-u)^k e^{\lambda(t-u)} dL(u)$, with $k \in \mathbb{N}_0$ and $\Re(\lambda) < 0$, and $\int_t^\infty (t-u)^k e^{\lambda(t-u)} dL(u)$, with $k \in \mathbb{N}_0$ and $\Re(\lambda) > 0$, and these integrals converge almost surely. Using partial integration, we may write for $s \leq t$ and $\Re(\lambda) < 0$

$$\begin{aligned} \int_{(s,t]} (t-u)^k e^{\lambda(t-u)} dL(u) &= \mathbf{1}_{\{k=0\}} L(t) - (t-s)^k e^{\lambda(t-s)} L(s) \\ &\quad + \int_s^t L(u) e^{\lambda(t-u)} (k(t-u)^{k-1} \mathbf{1}_{\{k \neq 0\}} + \lambda(t-u)^k) du. \end{aligned}$$

Letting $s \rightarrow -\infty$, we conclude from (23) and the dominated convergence theorem that

$$\lim_{t \rightarrow -\infty} e^{-\varepsilon|t|} \int_{-\infty}^t (t-u)^k e^{\lambda(t-u)} dL(u) = \lim_{t \rightarrow \infty} e^{-\varepsilon|t|} \int_{-\infty}^t (t-u)^k e^{\lambda(t-u)} dL(u) = 0.$$

The analogous result for the integrals \int_t^∞ with $\Re(\lambda) > 0$, completes the proof. \square

2.3 Prediction of Y

We return now to the problem of determining $\mathbb{E}(Y(h)|Y(u), -\infty < u \leq 0)$. Assuming that the zeroes of the polynomial $a(z)$ have strictly negative real parts, we can express the strictly stationary $\mathbf{X}(t)$ (see Brockwell and Lindner (2009), equation (3.5)) as

$$\mathbf{X}(t) = \int_{-\infty}^t \mathbf{l}_a(t-u) dL(u) = \int_{-\infty}^t e^{A(t-u)} \mathbf{e}_p dL(u), \quad t \in \mathbb{R}, \quad (24)$$

where $\mathbf{l}_a(t)$ is the sum of the residues of the column vector $e^{zt}a^{-1}(z)[1 \ z \ \cdots \ z^{p-1}]$ at the zeroes of the polynomial $a(z)$ (cf. (18)). From (24) and (3) we see at once that \mathbf{X} is a Markov process and that $(\mathbf{X}(t))_{t \leq 0}$ is independent of $(L(t))_{t \geq 0}$. These facts plays a crucial role in the proof of the following theorem. From (24) it also follows that the strictly stationary state vector \mathbf{X} and the strictly stationary process Y satisfy $\mathbb{E}|\mathbf{X}(t)| < \infty$ and $\mathbb{E}|Y(t)| < \infty$ for all $t \in \mathbb{R}$ if $\mathbb{E}|L(1)| < \infty$, and $\mathbb{E}|\mathbf{X}(t)|^2 < \infty$ and $\mathbb{E}|Y(t)|^2 < \infty$ for all $t \in \mathbb{R}$ if $\mathbb{E}|L(1)|^2 < \infty$.

Theorem 2.7. *Suppose that the zeroes of the polynomials $a(z)$ and $b(z)$ all have strictly negative real parts, that $a(z)$ and $b(z)$ have no common zeroes, that $\mathbb{E}|L(1)| < \infty$ and that Y is the strictly stationary solution of (1) and (3). Then, for $h \geq 0$,*

$$\begin{aligned} \mathbb{E}(Y(h)|Y(s), s \leq 0) &= \mathbf{b}'e^{Ah}\mathbf{X}(0) + \mathbb{E}L(1) \int_0^h \mathbf{b}'e^{Au}\mathbf{e}_p du \\ &= \mathbf{b}'e^{Ah}\mathbf{X}(0) + \mathbb{E}L(1) \sum_{\lambda} \sum_{k=0}^{\nu(\lambda)-1} \alpha_{\lambda k} h^k e^{\lambda k}, \end{aligned} \quad (25)$$

where the components of $\mathbf{X}(0)$ ($= \mathbf{X}(0-)$ a.s.) are given by (22) with $t = 0$ if $q \neq 0$ and by $X_j(0) = Y^{(j)}(0)$ ($= Y^{(j)}(0-)$ a.s.) if $q = 0$. The sum \sum_{λ} is over the distinct zeroes λ of $a(\cdot)$ and $\nu(\lambda)$ denotes the multiplicity of λ . $\sum_{k=0}^{\nu(\lambda)-1} \alpha_{\lambda k} h^k e^{\lambda k}$ is the residue of $z \mapsto (e^{zh} - 1)b(z)/[a(z)z]$ at λ , i.e.

$$\sum_{k=0}^{\nu(\lambda)-1} \alpha_{\lambda k} h^k e^{\lambda k} = \frac{1}{(\nu(\lambda) - 1)!} [D_z^{\nu(\lambda)-1} ((z - \lambda)^{\nu(\lambda)}(e^{zh} - 1)b(z)/[a(z)z])]_{z=\lambda},$$

where D_z denotes differentiation with respect to z . (If $\nu(\lambda) = 1$ the last sum reduces to $b(\lambda)(e^{\lambda h} - 1)/[\lambda a'(\lambda)]$.)

If in addition $\mathbb{E}L(1)^2 < \infty$ then (25) is the minimum mean-squared error predictor of $Y(h)$ based on $\{Y(s), s \leq 0\}$ and its mean-squared error is

$$\begin{aligned} \text{Var} \left(\mathbf{b}' \int_0^h e^{A(h-u)} \mathbf{e}_p dL(u) \right) &= \text{Var}(L(1)) \int_0^h |\mathbf{b}'e^{Au}\mathbf{e}_p|^2 du \\ &= \text{Var}(L(1)) \int_0^h \left| \sum_{\lambda} \sum_{k=0}^{\mu(\lambda)-1} c_{\lambda k} u^k e^{\lambda u} \right|^2 du, \end{aligned} \quad (26)$$

where $\sum_{k=0}^{\mu(\lambda)-1} c_{\lambda k} u^k e^{\lambda u}$ is the residue of $z \mapsto e^{zu}b(z)/a(z)$ at λ .

Proof. From (1), (3) and the independence of $(\mathbf{X}(t))_{t \leq 0}$ and $(L_t)_{t \geq 0}$ we see at once that

$$\mathbb{E}(Y(h)|Y(s), s \leq 0) = \mathbf{b}'e^{Ah}\mathbb{E}(\mathbf{X}(0)|Y(s), s \leq 0) + \mathbb{E}L(1) \int_0^h \mathbf{b}'e^{Au}\mathbf{e}_p du, \quad h \geq 0.$$

Since, by Remark 2.5 (identifying almost surely equal random variables), $\mathbf{X}(0)$ is measurable with respect to the σ -algebra \mathcal{F}_0 generated by $\{Y(s), s \leq 0\}$,

$$\mathbb{E}(Y(h)|Y(s), s \leq 0) = \mathbf{b}'e^{Ah}\mathbf{X}(0) + \mathbb{E}L(1) \int_0^h \mathbf{b}'e^{Au}\mathbf{e}_p du, \quad h \geq 0,$$

where the components of the vector $\mathbf{X}(0)$ are found from (22) if $q \neq 0$ and from $X_j(0) = Y^{(j)}(0)$ if $q = 0$. The last term can be rewritten as in (25) by applying equation (2.10) of Brockwell and Lindner (2009) and Fubini's theorem, as in the proof of Proposition 5.1(b) of Brockwell and Lindner (2009).

The prediction error of the conditional expectation (25) is

$$Y(h) - \mathbb{E}(Y(h)|Y(s), s \leq 0) = \int_0^h \mathbf{b}'e^{A(h-u)}\mathbf{e}_p d(L(u) - u\mathbb{E}L(1)). \quad (27)$$

If $\mathbb{E}L(1)^2 < \infty$ the conditional expectation minimizes the expected squared prediction error over all measurable functions of $\{Y(u), u \leq 0\}$. This minimum value of the expected squared error is therefore

$$\text{Var} \left(\mathbf{b}' \int_0^h e^{A(h-u)}\mathbf{e}_p dL(u) \right) = \text{Var}(L(1)) \int_0^h \mathbf{b}'e^{Au}\mathbf{e}_p\mathbf{e}_p'e^{\bar{A}'u}\bar{\mathbf{b}} du$$

(here \bar{A} and $\bar{\mathbf{b}}$ denote the complex conjugates of A and \mathbf{b} , respectively). Using the expression for $\mathbf{b}'e^{Ah}\mathbf{e}_p$ in equation (2.10) of Brockwell and Lindner (2009), we obtain (26) as required. \square

The importance of the expression (26) for the mean-squared error when $\mathbb{E}L(1)^2 < \infty$ is that it provides a benchmark against which less efficient predictors can be measured. If, for example, prediction is based on $\{Y(n\Delta), n = -1, -2, \dots\}$ or on some bounded subset of $\{Y(u), u \leq 0\}$ then the corresponding mean-squared errors can be compared with (26) to check on their performance relative to the best predictor (25) based on $\{Y(u), u \leq 0\}$. We conclude this section by showing that the best *linear* predictor, $P(Y(h)|Y(s), s \leq 0)$, of $Y(h)$ based on $\{Y(u), u \leq 0\}$, i.e. the orthogonal projection of $Y(h)$ onto the closed linear span, $\overline{\text{sp}}\{Y(s), s \leq 0\}$, and the orthogonal projection $P(Y(h)|1, Y(s), s \leq 0)$ are both *equal* to the best predictor (25).

Theorem 2.8. *If the assumptions of Theorem 2.7 are satisfied and $\mathbb{E}L(1)^2 < \infty$ then, for $h \geq 0$, the best linear predictor $P(Y(h)|Y(s), s \leq 0)$ and the predictor $P(Y(h)|1, Y(s), s \leq 0)$ are both equal to the minimum mean-squared error predictor $\mathbb{E}(Y(h)|Y(s), s \leq 0)$ given by (25), with mean squared error (26).*

Proof. Since

$$\overline{\text{sp}}\{Y(s), s \leq 0\} \subset \overline{\text{sp}}\{1, Y(s), s \leq 0\} \subset L^2(\Omega, \sigma\{Y(s) : s \leq 0\}, \mathbb{P}) \subset L^2(\Omega, \mathcal{F}, \mathbb{P}),$$

where $(\Omega, \mathcal{F}, \mathbb{P})$ is the underlying probability space, it will be sufficient to show that $\mathbb{E}(Y(h)|Y(s), s \leq 0)$ as given by (25) is in $\mathcal{M} := \overline{\text{sp}}\{Y(s), s \leq 0\}$. To see this, observe first that $\mathbb{E}|\mathbf{X}(t)|^2 < \infty$ for all $t \in \mathbb{R}$ by (24), since \mathbf{l}_a decreases exponentially (and hence is in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$). Let $T > 0$. By (4), there is a constant C_T such that

$$\sup_{t \in [0, T]} |\mathbf{X}(t)| \leq C_T \left(|\mathbf{X}(0)| + \sup_{t \in [0, T]} |L(t)| \right),$$

from which we conclude that

$$\mathbb{E} \sup_{t \in [-T, 0]} |\mathbf{X}(t)|^2 < \infty \quad (28)$$

by stationarity and Theorem 25.18 of Sato (1999). Hence also $\mathbb{E} \sup_{t \in [-T, 0]} |Y(t)|^2 < \infty$ for all T . Let κ be given by (20). Since $N^{-1} \sum_{i=1}^N \kappa(iT/N) Y(-iT/N) \in \mathcal{M}$ for all $N \in \mathbb{N}$ and since it converges almost surely to $\int_{-T}^0 \kappa(-u) Y(u) du$ as $N \rightarrow \infty$, it converges also in L^2 by Lebesgue's dominated convergence theorem. Hence $\int_{-T}^0 \kappa(-u) Y(u) du \in \mathcal{M}$. Next, observe that

$$\left| \int_{-T}^0 \kappa(-u) Y(u) du \right| \leq \int_{-\infty}^0 |\kappa(-u) Y(u)| du \leq \sum_{n=0}^{\infty} \sup_{u \in [n, n+1]} |\kappa(u)| \sup_{u \in [-n-1, -n]} |Y(u)|,$$

and since $\mathbb{E} \sup_{u \in [-n-1, -n]} |Y(u)|^2 = \mathbb{E} \sup_{u \in [-1, 0]} |Y(u)|^2 < \infty$ by stationarity, the exponential decrease of κ and Minkowski's inequality show that $\mathbb{E} \left(\int_{-\infty}^0 |\kappa(-u) Y(u)| du \right)^2 < \infty$. Hence, by (19) and Lebesgue's dominated convergence theorem, $\int_{-T}^0 \kappa(-u) Y(u) du$ converges in L^2 to $X(0)$ as $T \rightarrow \infty$, from which we conclude that $X(0) \in \mathcal{M}$. The same reasoning can be applied to show that $X(t) \in \mathcal{M}$ for all $t < 0$. Since

$$\left| X_1(0) - \frac{X(0) - X(-\delta)}{\delta} \right| = \left| \frac{1}{\delta} \int_{-\delta}^0 (X_1(0) - X_1(u)) du \right| \leq 2 \sup_{u \in [-\delta, 0]} |X(u)|,$$

letting $\delta \downarrow 0$ we conclude from Lebesgue's dominated convergence theorem, using (28) and (7), that $\delta^{-1}(X(0) - X(-\delta))$ converges in L^2 to $X_1(0-) = X_1(0)$ a.s., and hence that $X_1(0) \in \mathcal{M}$. The same reasoning can be applied to show that $X_1(t) \in \mathcal{M}$ for all $t \leq 0$. Iterating this argument, we see that $X(t), X_1(t), \dots, X_{p-1}(t) \in \mathcal{M}$ for all $t \leq 0$ and hence $\mathbf{b}'e^{Ah}\mathbf{X}(0) \in \mathcal{M}$. Further, since $L(0) = 0$, from (21) we conclude that $L(s) \in \mathcal{M}$ for all $s \leq 0$. Since

$$\mathbb{E} |t^{-1}L(t) - \mathbb{E}L(1)|^2 = t^{-2} \text{Var}(L(t)) = |t|^{-1} \text{Var}(L(1)) \rightarrow 0, \quad t \rightarrow -\infty,$$

we conclude that $t^{-1}L(t)$ converges in L^2 to $\mathbb{E}L(1)$ as $t \rightarrow -\infty$. Hence $\mathbb{E}L(1) \in \mathcal{M}$ and from (25) we conclude that $\mathbb{E}(Y(h)|Y(s), s \leq 0) \in \mathcal{M}$, finishing the proof. \square

3 Determination of $\mathbb{E}(Y(h)|Y(u), -M \leq u \leq 0)$

In this section we shall determine $\mathbb{E}(Y(h)|Y(u), -M \leq u \leq 0)$ for $h \geq 0$ when L is a pure jump Lévy process. We do not need to assume that Y is strictly stationary, but require it to be causal or, more precisely, non-anticipative in the sense that $(L(t))_{t \geq 0}$ and $(Y(t))_{t \in [-M, 0]}$ are independent. The first step is the recovery of $(L(t))_{t \in [-M, 0]}$ from the realisation of Y on $[-M, 0]$. This is achieved using (8) and the Lévy-Itô decomposition as shown below in Proposition 3.1.

Throughout the section we shall suppose that $L = (L(t))_{t \in \mathbb{R}}$ is a two-sided Lévy process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with characteristic triplet $(0, \Pi, \gamma)$, i.e. that L is a pure jump Lévy process with Lévy measure Π and that $L(1)$ has characteristic function,

$$\mathbb{E}e^{iL(1)z} = \exp \left\{ i\gamma z + \int_{\mathbb{R}} (e^{ixz} - 1 - ixz \mathbf{1}_{|x| \leq 1}) \Pi(dx) \right\}, \quad z \in \mathbb{R}.$$

By the Lévy-Itô-decomposition (see e.g. Sato (1999), Theorem 19.2) there is a set $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that, for each $\omega \in \Omega_0$, the path $L(\cdot, \omega)$ is càdlàg on \mathbb{R} and

$$\lim_{\varepsilon \downarrow 0} \left(\gamma t - \sum_{t < s \leq 0, |\Delta L(s, \omega)| > \varepsilon} \Delta L(s, \omega) - t \int_{\varepsilon < |x| \leq 1} x \Pi(dx) \right) = L(t, \omega) \quad \forall t \in [-M, 0], \quad (29)$$

the limit being uniform in t on $[-M, 0]$. Then, for each $\omega \in \Omega_0$, $Z_t(\omega) := Y^{(p-q-1)}(t, \omega)$ exists in the sense that $Y^{(p-q-2)}$ is differentiable from the left and from the right on \mathbb{R} with càdlàg derivative $Z = Y^{(p-q-1)}$.

Proposition 3.1. *Let $Y = (Y(t))_{t \in \mathbb{R}}$ be a CARMA(p, q)-process with characteristic polynomials $a(z)$ and $b(z)$, driven by L as defined in the preceding paragraph, and with initial state vector $\mathbf{X}(0)$. (We do not require Y to be either strictly stationary or causal, nor do we exclude the possibility that $a(z)$ and $b(z)$ have common zeroes.) For any $M > 0$ we define $\mathcal{G}_M := \sigma\{Y(u) : u \in [-M, 0]\}$. Then*

(i) *$L(t)$ is almost surely equal to a \mathcal{G}_M -measurable random variable for each $t \in [-M, 0]$, more precisely*

(ii) *if $(x_m)_{m \in \mathbb{N}}$ is a strictly decreasing sequence of real numbers such that $\lim_{m \rightarrow \infty} x_m = 0$ and $\Pi(\{-x_m, x_m\}) = 0$ for all $m \in \mathbb{N}$, and if*

$$\begin{aligned} R_{n,m}(t, \omega) &:= \sum_{i=0}^{n-1} (Z_{-iM/n} - Z_{-(i+1)M/n}) \mathbf{1}_{\{|Z_{-iM/n} - Z_{-(i+1)M/n}| \geq x_m\}} \mathbf{1}_{[t, 0]}(-iM/n), \end{aligned}$$

for $n, m \in \mathbb{N}$, $t \in [-M, 0]$ and $\omega \in \Omega_0$, then there is a set $\Omega_1 \in \mathcal{F}$ such that $\Omega_1 \subset \Omega_0$, $\mathbb{P}(\Omega_1) = 1$, $\lim_{n \rightarrow \infty} R_{n,m}(t, \omega)$ exists for all $\omega \in \Omega_1$, $m \in \mathbb{N}$ and $t \in [-M, 0]$, and such

that

$$L(t, \omega) = \gamma t - \lim_{m \rightarrow \infty} \left(t \int_{|x| \in (x_m, 1]} x \Pi(dx) + \lim_{n \rightarrow \infty} R_{n,m}(t, \omega) \right), \quad t \in [-M, 0], \quad \omega \in \Omega_1, \quad (30)$$

where the limit in m is uniform in t on $[-M, 0]$.

Proof. (i) is an immediate consequence of (ii), since the right-hand-side of (30) is \mathcal{G}_M -measurable.

We shall now prove (ii). That Z exists on Ω_0 and is càdlàg and the fact that $\Delta L(s, \omega) = \Delta Z_s(\omega)$ for each $\omega \in \Omega_0$ has already been observed. Since $\Pi(\{-x_m, x_m\}) = 0$, there is some set $\Omega_1 \in \mathcal{F}$, $\Omega_1 \subset \Omega_0$ with $\mathbb{P}(\Omega_1) = 1$ and $\Delta Z_s(\omega) \neq |x_m|$ for all $m \in \mathbb{N}, \omega \in \Omega_1$ and $s \in \mathbb{R}$, and with $\Delta Z_{-M}(\omega) = 0$ for all $\omega \in \Omega_1$. Define

$$W_t := Z_t + \sum_{t < u \leq 0, |\Delta Z_u| \geq x_m} \Delta Z_u, \quad t \in [-M, 0],$$

for fixed $\omega \in \Omega_1$ and $m \in \mathbb{N}$. Then W is càdlàg on $[-M, 0]$ and satisfies $|\Delta W_s| < \varepsilon$ for each $s \in [-M, 0]$ for some $\varepsilon = \varepsilon(\omega) < x_m$. Hence, for each $s \in [-M, 0]$ there exists an open neighborhood $U_{s,m}$ in $[-M, 0]$ of s such that $|W_u - W_v| < \varepsilon$ for all $u, v \in U_{s,m}$. The compactness of $[-M, 0]$ implies that there exists $\delta(m, \omega) > 0$ such that $|W_u - W_v| < x_m$ whenever $u, v \in [-M, 0]$ with $|u - v| \leq \delta$. In particular,

$$\sum_{i=0}^{n-1} \mathbf{1}_{\{|W_{-iM/n} - W_{-(i+1)M/n}| \geq x_m\}} \mathbf{1}_{[t, 0]}(-(i+1)M/n) = 0$$

for large enough n , uniformly in t on $[-M, 0]$. If Z (equivalently L) has no jump of absolute size greater than or equal to x_m in $(-(i+1)M/n, -iM/n]$, then $Z_{-iM/n} - Z_{-(i+1)M/n}$ is equal to $W_{-iM/n} - W_{-(i+1)M/n}$. For $\omega \in \Omega_1$ denote by $S(m, \omega)$ the finite set of all $s \in (-M, 0]$ with $|\Delta Z_s(\omega)| = |\Delta L(s, \omega)| \geq x_m$. For $s \in S(m, \omega)$, define $i(s) \in \{0, \dots, n-1\}$ through the property that $s \in (-(i(s)+1)M/n, -i(s)M/n]$. Then for sufficiently large n , the indices $i(s)$ are distinct for different $s \in S(m, \omega)$, and

$$R_{n,m}(t, \omega) = \sum_{s \in S(m, \omega)} (Z_{-i(s)M/n} - Z_{-(i(s)+1)M/n}) \mathbf{1}_{\{|Z_{-i(s)M/n} - Z_{-(i(s)+1)M/n}| \geq x_m\}} \mathbf{1}_{[t, 0]}(-(i(s)+1)M/n).$$

Since Z has no jumps of absolute size x_m on Ω_1 , we obtain for $\omega \in \Omega_1$ and $s \in S(m, \omega)$

$$\begin{aligned} \lim_{n \rightarrow \infty} (Z_{-i(s)M/n} - Z_{-(i(s)+1)M/n}) &= \Delta Z_s \quad \text{and} \\ \lim_{n \rightarrow \infty} \mathbf{1}_{\{|Z_{-i(s)M/n} - Z_{-(i(s)+1)M/n}| \geq x_m\}} &= \mathbf{1}_{\{|\Delta Z_s| > x_m\}}. \end{aligned}$$

Hence $\lim_{n \rightarrow \infty} R_{n,m}(t, \omega)$ exists for each $\omega \in \Omega_1$ and is equal to $\sum_{t < s \leq 0, |\Delta Z_s| > x_m} \Delta Z_s$. Equation (30) then follows immediately from (29). \square

Remark 3.2. If the exceptional set $\Omega \setminus \Omega_1$ appearing in the statement of Proposition 3.1 is the empty set, which can without loss of generality be assumed by redefining L on $\Omega \setminus \Omega_1$ as $L(t, \omega) = 0$, then $L(t)$ is not only almost surely equal to a \mathcal{G}_M -measurable random variable, but is \mathcal{G}_M -measurable. A similar remark is true for the random variables $\mathbf{f}_M(t)$ and $\mathbf{X}(0)$ appearing in Theorem 3.3 below.

Theorem 3.3. *Let $L = (L(t))_{t \in \mathbb{R}}$ be a pure jump Lévy process, and let $Y = (Y(t))_{t \in \mathbb{R}}$ be a CARMA(p, q) process driven by L , with characteristic polynomials $a(z)$ and $b(z)$, such that $(Y(s))_{s \in [-M, 0]}$ is independent of $(L(s))_{s \geq 0}$ (as is the case if Y is strictly stationary and causal). Suppose that $a(z)$ and $b(z)$ have no common zeroes, and let $M > 0$. Define*

$$\mathbf{f}_M(t) := \left[L(t)I_p - L(-M)e^{A(t+M)} + \int_{-M}^t L(u)Ae^{A(t-u)} du \right] \mathbf{e}_p, \quad t \in [-M, 0].$$

Then $\mathbf{f}_M(t)$ is almost surely equal to a \mathcal{G}_M -measurable random variable for each $t \in [-M, 0]$ and can be calculated from (30). Further, the matrix $\int_0^M e^{A't} \mathbf{b} \mathbf{b}' e^{At} dt$ is invertible and $\mathbf{X}(0)$ is given by

$$\mathbf{X}(0) = e^{AM} \left(\int_0^M e^{A't} \mathbf{b} \mathbf{b}' e^{At} dt \right)^{-1} \int_0^M e^{A't} \mathbf{b} (Y(t-M) - \mathbf{b}' \mathbf{f}_M(t-M)) dt + \mathbf{f}_M(0). \quad (31)$$

In particular, $\mathbf{X}(0)$ is almost surely equal to a \mathcal{G}_M -measurable random variable. Further, if $\mathbb{E}|L(1)| < \infty$ and $\sup_{s \in [-M, 0]} \mathbb{E}|Y(s)| < \infty$ (the latter will hold if $\mathbb{E}|L(1)| < \infty$ and Y is strictly stationary and causal), then $\mathbb{E}|\mathbf{X}(0)| < \infty$, and for $h \geq 0$ we have $\mathbb{E}|Y(h)| < \infty$ and

$$\mathbb{E}(Y(h)|Y(s), s \in [-M, 0]) = \mathbf{b}' e^{Ah} \mathbf{X}(0) + \mathbb{E}L(1) \int_0^h \mathbf{b}' e^{Au} \mathbf{e}_p du, \quad (32)$$

which can be simplified as in (25). If in addition $\mathbb{E}|Y(h)|^2 < \infty$ (as is the case if $\mathbb{E}L(1)^2 < \infty$ and Y is strictly stationary and causal), then (32) is the minimum mean-squared error predictor of $Y(h)$ based on $\{Y(s), s \leq 0\}$. The corresponding prediction error and mean-squared error are given by (27) and (26) respectively.

Proof. By (4) we have

$$\mathbf{X}(t) = e^{A(t+M)} \mathbf{X}(-M) + \mathbf{f}_M(t), \quad t \in [-M, 0].$$

Define

$$\mathbf{Z}(t) := \mathbf{X}(t) - \mathbf{f}_M(t).$$

Then, since $\mathbf{f}_M(-M) = 0$,

$$\mathbf{Z}(t) = e^{A(t+M)} \mathbf{Z}(-M), \quad t \in [-M, 0].$$

Define

$$w(t) := \mathbf{b}'\mathbf{Z}(t) = (Y(t) - \mathbf{b}'\mathbf{f}_M(t)), \quad t \in [-M, 0].$$

Since $a(z)$ and $b(z)$ have no common zeroes, the pair (A, \mathbf{b}') is observable by Proposition 2.1. Hence $\int_0^M e^{A't} \mathbf{b}\mathbf{b}' e^{At} dt$ is invertible and

$$\mathbf{X}(-M) = \mathbf{Z}(-M) = \left(\int_0^M e^{A't} \mathbf{b}\mathbf{b}' e^{At} dt \right)^{-1} \int_0^M e^{A't} \mathbf{b}w(t-M) dt$$

(cf. Lemma 12.3.2 and Equation (12.3.18) in Bernstein (2009)). This gives (31) since $\mathbf{X}(0) = e^{AM}\mathbf{X}(-M) + \mathbf{f}_M(0)$. If $\mathbb{E}|L(1)| < \infty$ and $\sup_{s \in [-M, 0]} \mathbb{E}|Y(s)| < \infty$, then it follows that $\sup_{s \in [-M, 0]} \mathbb{E}|L(s)| < \infty$ (see Sato (1999), Theorem 25.18). Hence $\mathbb{E}|\mathbf{X}(0)| < \infty$ by (31) and $\mathbb{E}|Y(h)| < \infty$ by (3). The remaining assertions now follow as in the proof of Theorem 2.7. \square

Corollary 3.4. *Under the assumptions of Theorem 3.3, if $\mathbb{E}|L(1)| < \infty$ and $\sup_{s \in [-R, 0]} \mathbb{E}|Y(s)| < \infty$ for all $R > 0$, then*

$$\mathbb{E}(Y(h)|Y(s), s \in [-M_1, 0]) = \mathbb{E}(Y(h)|Y(s), s \in [-M_2, 0])$$

for all $h \geq 0$ and $0 < M_1, M_2 < \infty$. Under the additional assumptions that the zeroes of $a(z)$ and $b(z)$ have strictly negative real parts and Y is strictly stationary and causal,

$$\mathbb{E}(Y(h)|Y(s), s \leq 0) = \mathbb{E}(Y(h)|Y(s), s \in [-M, 0]).$$

Proof. This follows from the fact that the prediction errors are given by (27) and hence the same in each case. \square

We have not determined the best linear predictor $P(Y(h)|Y(s), s \in [-M, 0])$ in the setting of Theorem 3.3. Since the construction of $\mathbb{E}(Y(h)|Y(s), s \in [-M, 0])$ is nonlinear, it may well be that $P(Y(h)|Y(s), s \in [-M, 0])$ and $\mathbb{E}(Y(h)|Y(s), s \in [-M, 0])$ differ. We have not investigated this issue further.

4 The sampled sequence $(Y_n^\Delta)_{n \in \mathbb{Z}}$ and $P[Y(t)|Y_n^\Delta, n < 0]$

In this section we are concerned with linear minimum-mean-squared error prediction of the stationary CARMA(p, q) process Y in (1) at times $h \in [0, \infty)$, based on observations at the discrete times $\{n\Delta, n = -1, -2, \dots\}$ when it is assumed, as before, that the parameters of the CARMA process are known. We shall assume in this section that $\mathbb{E}L(1) = 0$ so that $\mathbb{E}Y(t) = 0$ for all t and that $\mathbb{E}L(1)^2 =: \sigma^2 \in (0, \infty)$. We shall also assume that the zeroes, $\lambda_1, \dots, \lambda_p$, of the autoregressive polynomial $a(z)$ in (6) all have strictly negative

real parts, and that the state vector \mathbf{X} is strictly stationary (by Proposition 2.1, this is necessarily the case if Y is strictly stationary and $a(z)$ and $b(z)$ have no common zeroes). Finally, we assume that $a_1, \dots, a_p, b_0, \dots, b_q \in \mathbb{R}$. Prediction of $Y(h)$ when h is an integer multiple of Δ can be done using standard discrete-time techniques applied to the sampled process $(Y_n^\Delta := Y(n\Delta))_{n \in \mathbb{Z}}$, which satisfies the ARMA(p, q) equations,

$$\prod_{i=1}^p (1 - e^{\lambda_i \Delta} B) Y_n^\Delta = \theta^\Delta(B) Z_n^\Delta, \quad \{Z_n^\Delta\} \sim \text{WN}(\sigma_\Delta^2), \quad (33)$$

where $\theta^\Delta(z) = 1 + \theta_1^\Delta z + \dots + \theta_{p-1}^\Delta z^{p-1}$ is a polynomial of degree less than or equal to $p - 1$, all of whose zeroes lie in the exterior of the closed unit disc. (There can be no zero on the unit circle by a result of Brockwell and Brockwell (1999).) Its coefficients and the white-noise variance σ_Δ^2 are determined by the parameters of the process Y . The calculation of the moving average coefficients and of σ_Δ^2 can in principle be carried out by computing the autocovariance function of $\{\prod_{i=1}^p (1 - e^{\lambda_i \Delta} B) Y_n^\Delta\}$ and determining the coefficients and white-noise variance of the corresponding invertible moving average using Wilson's algorithm. This however is a complicated and tedious procedure. In this section we develop a simpler approach using the Kalman recursions which not only determines all of the parameters appearing in (33) but also provides best linear forecasts of $Y(h)$ for *all* non-negative h , not just for values which are integer-multiples of Δ .

From (3) we have

$$\mathbf{X}(t) = e^{At} \mathbf{X}(0) + \int_0^t e^{A(t-u)} \mathbf{e}_p dL(u),$$

whence $P[\mathbf{X}(t)|Y_n^\Delta, n < 0] = e^{At} P[\mathbf{X}(0)|Y_n^\Delta, n < 0]$ by our assumption that the zeroes of $a(z)$ all have strictly negative real parts so that $\{\mathbf{X}(t) : t \leq 0\}$ and $\{L(t) : t \geq 0\}$ are independent. From (1) and (3) we therefore have

$$P[Y(t)|Y_n^\Delta, n < 0] = \mathbf{b}' e^{At} P[\mathbf{X}(0)|Y_n^\Delta, n < 0], \quad (34)$$

where $P[\mathbf{X}(0)|Y_n^\Delta, n < 0]$ can be found by applying the one-step steady-state Kalman state predictor to the *discrete* state-space model,

$$\mathbf{X}_{n+1} = F \mathbf{X}_n + \mathbf{V}_n, \quad \{\mathbf{V}_n\} \sim \text{IID}(\mathbf{0}, Q), \quad (35)$$

$$Y_n^\Delta = \mathbf{b}' \mathbf{X}_n, \quad (36)$$

where $\mathbf{X}_n := \mathbf{X}(n\Delta)$, $F = e^{A\Delta}$, $\mathbf{V}_n = \int_{n\Delta}^{(n+1)\Delta} e^{A((n+1)\Delta-u)} \mathbf{e}_p dL(u)$ and $Q = \mathbb{E}(\mathbf{V}_n \mathbf{V}_n') = \sigma^2 \int_0^\Delta e^{At} \mathbf{e}_p \mathbf{e}_p' e^{A't} dt$. If the eigenvalues $\lambda_1, \dots, \lambda_p$ of the matrix A each have multiplicity one, we can also write

$$Q = \sigma^2 \sum_{r=1}^p \sum_{s=1}^p \frac{e^{\Delta(\lambda_r + \lambda_s)} - 1}{(\lambda_r + \lambda_s) \prod_{k \neq r} (\lambda_r - \lambda_k) \prod_{m \neq s} (\lambda_s - \lambda_m)} [\lambda_r^{i-1} \lambda_s^{j-1}]_{i,j=1}^p.$$

Since obviously $[\lambda I_p - A, \mathbf{e}_p]$ has full rank p for any $\lambda \in \mathbb{C}$, the pair (A, \mathbf{e}_p) is controllable by the PBH test for controllability (see Bernstein (2009), Corollary 12.6.19), implying that Q is invertible by Lemma 12.6.2 in Bernstein (2009).

Our goal now is to determine $\hat{\mathbf{X}}_0 := P[\mathbf{X}(0)|Y_n^\Delta, n < 0]$ from the steady-state Kalman recursions applied to the discrete-time system (35) and (36). Then (34) gives the required predictor of $Y(t)$.

Since $(\mathbf{X}_n)_{n \in \mathbb{Z}}$ is strictly stationary, since all eigenvalues of F have modulus strictly less than 1 and since Q is (strictly) positive definite, it follows from Proposition 13.2 in Hamilton (1994) that the steady-state mean-square error covariance matrix, $\Omega := \mathbb{E}[(\mathbf{X}_0 - \hat{\mathbf{X}}_0)(\mathbf{X}_0 - \hat{\mathbf{X}}_0)']$ is the unique positive semidefinite solution of the Riccati equation,

$$\Omega = F\Omega F' + Q - F\Omega \mathbf{b}(\mathbf{b}'\Omega \mathbf{b})^{-1} \mathbf{b}'\Omega F',$$

and that the sequence $\{\Omega_n\}$ defined by the recursions,

$$\Omega_{n+1} = F\Omega_n F' + Q - F\Omega_n \mathbf{b}(\mathbf{b}'\Omega_n \mathbf{b})^{-1} \mathbf{b}'\Omega_n F', \quad (37)$$

with $\Omega_0 = I_p$, the $p \times p$ identity matrix, converges to Ω as $n \rightarrow \infty$ (here, $(\mathbf{b}'\Omega_n \mathbf{b})^{-1}$ needs to be interpreted as a generalized inverse if $\mathbf{b}'\Omega_n \mathbf{b} = 0$, e.g. as being 0, but it is easy to see that the limit satisfies $\mathbf{b}'\Omega \mathbf{b} > 0$ and hence $\mathbf{b}'\Omega_n \mathbf{b} > 0$ for large n , since otherwise $Y(0)$ could be exactly predicted in terms of $(Y_n^\Delta : -n \in \mathbb{N})$, contradicting equation (26)).

The steady-state Kalman gain is

$$K = F\Omega \mathbf{b}(\mathbf{b}'\Omega \mathbf{b})^{-1}$$

and the eigenvalues of the matrix $F - K\mathbf{b}'$ are all strictly less than one in absolute value (cf. Hamilton (1994), Proposition 13.2).

Since we have assumed that the eigenvalues of the matrix A all have negative real parts, the stationary solution of (35) is

$$\mathbf{X}_n = \sum_{j=0}^{\infty} F^j \mathbf{V}_{n-1-j}$$

and the corresponding best linear predictor $\hat{\mathbf{X}}_n := P[\mathbf{X}_n|Y_m^\Delta, m \in \mathbb{Z}, m < n]$ satisfies

$$\begin{aligned} \hat{\mathbf{X}}_n &= F\hat{\mathbf{X}}_{n-1} + K\mathbf{b}'(\mathbf{X}_{n-1} - \hat{\mathbf{X}}_{n-1}) \\ &= (F - K\mathbf{b}')\hat{\mathbf{X}}_{n-1} + KY_{n-1}^\Delta \\ &= \sum_{j=1}^{\infty} (F - K\mathbf{b}')^{j-1} KY_{n-j}^\Delta \end{aligned}$$

(cf. equations (13.5.8) and (13.5.12) in Hamilton (1994)). Hence, from (34), we have the required predictor,

$$P[Y(h)|Y_n^\Delta, n < 0] = \mathbf{b}'e^{Ah} \sum_{j=1}^{\infty} (F - K\mathbf{b}')^{j-1} KY_{-j}^\Delta, \quad (38)$$

where $F = e^{A\Delta}$ and $K = e^{A\Delta}\Omega\mathbf{b}(\mathbf{b}'\Omega\mathbf{b})^{-1}$. The prediction error is

$$Y(h) - P[Y(h)|Y_n^\Delta, n < 0] = \mathbf{b}'e^{Ah}(\mathbf{X}_0 - \hat{\mathbf{X}}_0) + \int_0^h \mathbf{b}'e^{A(h-u)}\mathbf{e}_p dL(u),$$

with variance,

$$v(h) = \mathbf{b}'e^{Ah}\Omega e^{A'h}\mathbf{b} + \sigma^2 \int_0^h \mathbf{b}'e^{Au}\mathbf{e}_p\mathbf{e}_p'e^{A'u}\mathbf{b} du. \quad (39)$$

If the eigenvalues of A all have multiplicity one, the second term on the right of (39) can be expressed as

$$\sigma^2 \sum_{r=1}^p \sum_{s=1}^p \frac{b(\lambda_r)b(\lambda_s)(e^{(\lambda_r+\lambda_s)h} - 1)}{a'(\lambda_r)a'(\lambda_s)(\lambda_r + \lambda_s)}.$$

It is the mean squared error of the best predictor $\mathbb{E}[Y(t)|Y(u), u < 0] = P[Y(t)|Y(u), u < 0]$ based on continuous observation of Y on $(-\infty, 0)$ (see Theorems 2.7 and 2.8). The first term on the right of (39) thus quantifies the loss of accuracy incurred by using only the best linear predictor based on the discrete past observations rather than the minimum mean squared error predictor based on the entire past $\{Y(t), t < 0\}$.

To conclude this section we use (38) with $t = 0$ to determine the parameters of the moving average parameters and white noise variance in the representation (33) of the sampled process. From (33) we obtain the representation of the best linear predictor \hat{Y}_0^Δ of Y_0^Δ in terms of $Y_{-1}^\Delta, Y_{-2}^\Delta, \dots$,

$$\hat{Y}_0^\Delta = \sum_{j=1}^{\infty} -\pi_j Y_{-j}^\Delta,$$

where π_j is the coefficient of z^j in the power-series expansion of

$$\pi(z) = \frac{\phi(z)}{\theta^\Delta(z)}, |z| < 1, \quad (40)$$

and $\phi(z) = \prod_{i=1}^p (1 - e^{\lambda_i \Delta} z) = \sum_{j=0}^p \phi_j z^j$ (see e.g. Brockwell and Davis (1991), Theorem 5.5.1)). But from (38)

$$\pi_j = -\mathbf{b}'(F - K\mathbf{b}')^{j-1}K, \quad j = 1, 2, \dots$$

Since $\pi(z)$ and $\phi(z)$ are both known, it is an elementary matter to compute the coefficients of $\theta^\Delta(z)$ recursively from (40). Thus $\theta_0^\Delta = 1, \pi_0 = 1$ and

$$\theta_k^\Delta = \phi_k - \sum_{j=0}^{k-1} \theta_j^\Delta \pi_{k-j}, \quad k = 1, \dots, p-1. \quad (41)$$

The corresponding white-noise variance in (33) is

$$\sigma_{\Delta}^2 = \mathbb{E}(Y_0^{\Delta} - \hat{Y}_0^{\Delta})^2 = \mathbf{b}'\Omega\mathbf{b}. \quad (42)$$

Remark 4.1. Having determined the polynomial $\theta^{\Delta}(z)$ as described above, it is a trivial matter to derive the coefficients ψ_j in the Wold representation,

$$Y_n^{\Delta} = \sum_{j=0}^{\infty} \psi_j Z_{n-j}^{\Delta}, \quad \{Z_n\} \sim \text{WN}(0, \sigma_{\Delta}^2),$$

from the power series expansion,

$$\psi(z) := \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta^{\Delta}(z)}{\phi(z)}, \quad |z| \leq 1.$$

The white-noise variance σ_{Δ}^2 in (42) can also be derived from Kolmogorov's formula,

$$\sigma_{\Delta}^2 = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f_{\Delta}(\omega) d\omega \right\},$$

where f_{Δ} is the spectral density of the sampled process $(Y_n^{\Delta})_{n \in \mathbb{Z}}$, which can be expressed as a sum of residues

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

and the sum is over the distinct zeroes λ of $a(z)$ (this follows from Brockwell et al. (2012)).

Example 4.2. [CARMA(2,1)]

Consider the CARMA(2,1) process defined by the (formal) stochastic differential equation,

$$(D^2 + 1.5D + .5)Y(t) = (D + 2)DL(t),$$

with $\mathbb{E}L(1) = 0$ and $\sigma^2 = \mathbb{E}L(1)^2 = 1$. In this case we have $\lambda_1 = -.5$, $\lambda_2 = -1$,

$$A = \begin{bmatrix} 0 & 1 \\ -.5 & -1.5 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and, taking } \Delta = 1, \quad Q = \begin{bmatrix} .114506 & .113909 \\ .113909 & .289797 \end{bmatrix}.$$

Iterating (37) with $F = e^A$, we find that Ω_n is constant to six decimal places for $n \geq 4$ and so, to this accuracy,

$$\Omega = \begin{bmatrix} .114709 & .114831 \\ .114831 & .293984 \end{bmatrix}, \quad K = \begin{bmatrix} .446226 \\ -.011950 \end{bmatrix} \quad \text{and } F - K\mathbf{b}' = \begin{bmatrix} -.047271 & .031076 \\ -.214751 & .141179 \end{bmatrix}.$$

From these matrices we immediately obtain

$$\hat{\mathbf{X}}_0^1 = \sum_{j=1}^{\infty} (F - K\mathbf{b}')^{j-1} K Y_{-j}^1$$

and the required predictor $P[Y(t)|Y_n^1, n < 0], t \in [0, \infty)$, is then found at once from (38). In particular

$$\hat{Y}_0^1 = \mathbf{b}'\hat{\mathbf{X}}_0^1 = .880502Y_{-1}^1 - .140444Y_{-2}^1 + \dots$$

Since $\phi_1 = -e^{-.5} - e^{-1} = -.974410$ and $\pi_1 = -.880502$ we also find immediately from (41) that the moving average coefficient of the sampled process appearing in (33) is

$$\theta_1^1 = \phi_1 - \pi_1 = -.093908.$$

The white-noise variance σ_1^2 is found from (42) to be

$$\sigma_1^2 = 1.21214.$$

It is interesting in this example to compare the mean squared errors of the predictors $\mathbb{E}[Y(0)|Y(u), u \leq -1]$ and $P[Y(0)|Y(-1), Y(-2), \dots]$. The former, by Theorem 2.7, is

$$\sigma^2 \int_0^1 \mathbf{b}'e^{At}\mathbf{e}_p\mathbf{e}_p'e^{A't}\mathbf{b} dt = \mathbf{b}'Q\mathbf{b} = 1.20346$$

and the latter is

$$\mathbb{E}(Y(0) - \hat{Y}_0^1)^2 = \mathbf{b}'\Omega\mathbf{b} = 1.21214.$$

This shows that in this case, from the point of view of one-step minimum mean squared error prediction, there is very little to be gained by attempting to calculate the conditional expectation, $\mathbb{E}[Y(0)|Y(-1), Y(-2), \dots]$ since we know that its mean squared error cannot be smaller than 1.20346, a value only barely smaller than the mean squared error of the linear predictor, $P[Y(0)|Y(-1), Y(-2), \dots]$.

The steady-state Kalman recursions thus permit very simple determination of the best linear predictors of Y at all non-negative times, based on observations at times $\{-n\Delta, n \in \mathbb{N}\}$, and provide also a very simple method for computing all the parameters in the ARMA equations for the process obtained by sampling the CARMA process on any uniform lattice with spacing Δ .

Acknowledgments

PB gratefully acknowledges the support of NSF Grant DMS-1107031. We are also indebted to two referees for their valuable comments on the manuscript.

References

Bernstein, D.S., 2009, Matrix Mathematics. Theory, Facts, and Formulas, 2nd edition, Princeton University Press, Princeton.

- Brockwell, A.E. and P.J. Brockwell, 1999, A class of non-embeddable ARMA processes. *Journal of Time Series Analysis* 20, 483–486.
- Brockwell, P.J. and R.A. Davis, 1991, *Time Series: Theory and Methods*, 2nd edition. New York, Springer.-Verlag.
- Brockwell, P.J. and A. Lindner, 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. *Journal of Time Series Analysis* 33, 152–160.
- Hamilton, J.D., 1994, *Time Series Analysis*. Princeton University Press, Princeton.
- Jurek, Z. J. and J.D. Mason, 1993, *Operator Limit Distributions in Probability Theory*. Wiley, New York.
- Pham, D.-T., 1977, Estimation of parameters of a continuous time Gaussian stationary process with rational spectral density. *Biometrika* 64, 385–389.
- Sato, K., 1999, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.
- Sato, K. and M. Yamazato, 1984, Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Processes and their Applications* 17, 73–100.
- Wolfe, S.J., 1982, On a continuous analogue of the stochastic difference equation $x_n = \rho x_{n-1} + b_n$. *Stochastic Processes and their Applications* 12, 301–312.