On the Equivalence of Discrete Time Representations of Continuous Time ARMA Processes

Michael A. Thornton
University of Reading

and

Marcus J. Chambers
University of Essex

July 2011

Abstract

This paper explores techniques to derive the exact discrete time representation for data generated by a continuous time ARMA (autoregressive moving average) process. It compares the results of Bergstrom's (1983) widely used method, augmented by a stochastic integration-by-parts formula, with recently proposed techniques that use a different state space form. The continuous time ARMA$(2, 1)$ system is considered in detail and the equivalence of the discrete time representations in the two methods is demonstrated for models with stock variables and with flow variables. The technique is then used to derive a mapping from a univariate discrete time ARMA $(2, 1)$ process to a univariate continuous time ARMA $(2, 1)$ process observed at discrete intervals. This is used to derive conditions for the embeddability of such processes.

Keywords. Continuous time; ARMA process; state space; discrete time representation; embedding.

Address for Correspondence: Michael A. Thornton, School of Economics, University of Reading, Whiteknights, Reading, RG6 6AW, England.

Tel: +44 118 3785344; fax: +44 118 3786533; e-mail: m.a.thornton@reading.ac.uk.

1. INTRODUCTION

Interest continues to grow in time series models with dynamics that operate continuously through real time but which can be estimated from data collected at discrete points in time and/or over discrete intervals. A range of techniques has been developed to estimate the parameters of linear continuous time systems. In econometric applications, the most heavily deployed have been those involving: spectral representations, Robinson (1976); Kalman filtering of state space forms, Harvey and Stock (1985) and Zadrozny (1988); and, Gaussian estimation using the exact discrete time representation, Bergstrom (1983). Bergstrom (1990) eloquently conveyed the advantages of using exact discrete time representations of stochastic differential equation systems, which include efficiency of estimators and ease of computation. Analysts using this method incur some initial set-up costs in deriving the exact discrete representation, but these can also be seen as an opportunity to impose *a priori* restrictions on parameters. Techniques have been developed to derive the exact discrete representation of higher order continuous time systems with mixed stock-flow data, exogenous variables, stochastic trends, cointegrating relationships and mixed differential-difference equations; see, for example, Bergstrom (1986, 1997), McCrorie (2001) and Chambers (1999, 2009). The exact discrete representation of models with moving average errors has been slower to develop, in part because the vast majority of the literature has centred on models with errors defined as random measures, which are not mean-square differentiable[1], but this has been addressed recently by Chambers and Thornton (2011).

In the fundamental paper in the literature, Bergstrom (1983) considers the continuous time AR(p) process, for the $n \times 1$ vector $x(t)$

$$d[D^{p-1}x(t)] = [a_0 + A_{p-1}D^{p-1}x(t) + \ldots + A_0 x(t)]dt + \zeta(dt),$$

where D denotes the mean square differential operator, a_0 is an $n \times 1$ vector, $A_0, \ldots, A_{p-1}$ are $n \times n$ matrices of unknown coefficients, and $\zeta(dt)$ is a random measure. He shows that the unique, in the mean square sense, solution to this system, subject to initial conditions, takes the form $x(t) = y_1(t)$ where $y_1(t)$ is obtained from partitioning the $np \times 1$ vector, $y(t)$ into p $n \times 1$ vectors, $y_1(t), \ldots, y_k(t)$, and

$$y(t) = \int_0^t e^{(t-r)\widetilde{A}}\bar{\zeta}(dr) + e^{\widetilde{A}}y(0) + (e^{\widetilde{A}} - I)a_0,$$

[1] See Bergstrom (1984) for extensive discussion of this point.

where

$$a = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_0 \end{bmatrix}, \quad \widetilde{A} = \begin{bmatrix} 0 & I & 0 & \dots & 0 \\ 0 & 0 & I & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & I \\ A_0 & A_1 & A_2 & \dots & A_{p-1} \end{bmatrix}, \bar{\zeta} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \zeta \end{bmatrix},$$

and where the matrix exponential is defined by $e^{At} = I + \sum_{j=1}^{\infty}(At)^j/j!$.

The uniqueness of this solution is particularly interesting, since the state space representation of a continuous time AR(p) underlying $\widetilde{A}$ is not itself unique. A solution replacing $\widetilde{A}$ with the matrix

$$A = \begin{bmatrix} A_{p-1} & I & 0 & \dots & 0 \\ A_{p-2} & 0 & I & \dots & 0 \\ \vdots & & & & \vdots \\ A_1 & 0 & 0 & \dots & I \\ A_0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

is equally valid. Although the form $\widetilde{A}$ has been widely used in the exact discrete representation literature, the form A offers advantages in models with moving average errors. It has been used to estimate continuous time ARMA models by, among others, Zadrozny (1988) and Brockwell[2] (2004) using Kalman filter techniques and by Chambers and Thornton (2011) using the exact discrete representation.

It is clear that the matrices A and $\widetilde{A}$ are not equivalent. Moreover, in general, $e^{A} \neq e^{\widetilde{A}}$. Nevertheless, this paper demonstrates that this apparent conflict is overcome when the two solutions are transformed to describe the relationship between the observable $y_1(t) = x(t)$ and its lags, using methods given in Chambers (1999). We consider in detail the most tractable form of the model, a continuous time ARMA$(2,1)$, comparing the results of Chambers and Thornton (2011) with those derived using Chambers (1999). Using a stochastic integration-by-parts formula, established by McCrorie (2000), we demonstrate that equivalence also holds for the representation when $\zeta(dt)$ is replaced with a moving average stochastic process, even though this does not follow necessarily from the equivalence of autoregressive parameters. Not only does this address an interesting puzzle, it also gives

[2]Brockwell (2004) uses A for the state equation but captures the moving average error through an observation equation.

further insight into the impact of a moving average error in a continuous time model in a widely applicable form. In common with aggregation of discrete time models[3] the inclusion of a moving average error does not change the order of its exact discrete representation. Nevertheless we show it offers more flexibility in fitting moments in the data. Our results depend upon the development of some general mappings between the exponents of the coefficient matrices in the two state space forms, which are of wider interest.

The paper is organised as follows. Section 2 defines the continuous time ARMA system, and discusses the mappings between the two possible state space forms and between their exponents. Section 3 examines the continuous time ARMA$(2, 1)$ system in some detail. The exact discrete representations for the model in the two state space forms are then shown to be equivalent. Section 4 uses the exact discrete representation of this model to discuss the range of discrete time process which can be embedded in it. Section 5 concludes, while proofs and derivations are contained in an Appendix.

2. STATE SPACE REPRESENTATIONS OF CONTINUOUS TIME ARMA MODELS AND THEIR SOLUTIONS

The continuous time ARMA(p, q) model with intercept for the $n \times 1$ vector $x(t),\ t > 0$ is given by

$$D^p x(t) = a_0 + A_{p-1} D^{p-1} x(t) + \ldots + A_0 x(t) + u(t) + \Theta_1 Du(t) + \ldots + \Theta_q D^q u(t), \quad (1)$$

where the $n \times 1$ vector a_0 and the $n \times n$ matrices $A_0, \ldots, A_{p-1}, \Theta_1, \ldots, \Theta_q$ contain unknown coefficients. The process $u(t)$ is an $n \times 1$ continuous time white noise vector with variance matrix Σ. Systems such as (1) are of widespread interest. In this form they are not physically realisable, but become so when both sides are integrated p times. In a continuous time ARMA model we must impose the condition that $p > q$ so that $x(t)$ itself has an integrable spectral density matrix and, hence, finite variance.

One representation of (1), close to that widely used in the exact discrete representation literature, defines the $np \times 1$ state vector $\widetilde{y}(t) = [x(t)', Dx(t)', \ldots, D^{p-1}x(t)']'$, the error vector $\widetilde{u}'(t) = [0', 0', \ldots, 0', v(t)']$, with a moving average error

$$v(t) = u(t) + \Theta_1 Du(t) + \ldots + \Theta_q D^q u(t),$$

[3]Once the order of aggregation is reasonably large.

and matrices a and $\widetilde{A}$ as defined in section 1. A simple extension of equation (3) of Chambers (1999) enables the system to be written

$$D\widetilde{y}(t) = a + \widetilde{A}\widetilde{y}(t) + \widetilde{u}(t), \quad t > 0. \tag{2}$$

The formulae in Chambers (1999) remain exactly valid for (2), but the derivation of the autocovariance properties of the resulting discrete time disturbance vector is complicated by the presence of the derivatives of $u(t)$ in $\widetilde{u}(t)$. In the next section we outline how these should be evaluated.

An alternative state space representation is used in Chambers and Thornton (2011), in which the $np \times 1$ state vector is defined as $y(t) = [y_1(t)', \ldots, y_p(t)']'$, with $y_1(t) = x(t)$. The system may then be written as

$$Dy(t) = a + Ay(t) + \Theta u(t), \tag{3}$$

where a and A as defined in section 1 and $\Theta = \left(\Theta'_{p-1}, \Theta'_{p-2}, \ldots, \Theta'_1, I\right)'$, with $\Theta_j = 0$ for $j > q$.

The solution to (3), conditional on $y(0)$, can be written

$$y(t) = e^{At}y(0) + \int_0^t [a + e^{A(t-s)}\Theta u(s)]ds, \quad t > 0, \tag{4}$$

while the solution to (2), conditional on $\widetilde{y}(0)$, can be written

$$\widetilde{y}(t) = e^{\widetilde{A}t}\widetilde{y}(0) + \int_0^t [a + e^{\widetilde{A}(t-s)}\widetilde{u}(s)]ds, \quad t > 0. \tag{5}$$

It is clear that the two solutions will not, in general, be identical. There is, however, a mapping between the matrices A and $\widetilde{A}$, which holds at the level of the $n \times n$ sub-matrices of autoregressive parameters. We use $A_{[i,j]}$ to denote the $n \times n$ sub-matrix in the i'th block row of the j'th block column of A. Let the operator $(^{\triangle})$ denote translation of these sub-matrices across the transverse diagonal, so that when A is partitioned into these p^2 square sub-matrices, the block $A^{\triangle}_{[i,j]} \equiv A_{[p+1-j,p+1-i]}$. Clearly $A^{\triangle} = \widetilde{A}$ and $\widetilde{A}^{\triangle} = A$. The following proposition identifies a useful mapping.

PROPOSITION 1. *Let F and $\widetilde{F}$ denote, respectively, the partitioned exponents of the matrices A and $\widetilde{A}$ defined above and assume that $A_0 \neq 0$. Then*

(i) $F = \widetilde{F}^{\triangle}$ if and only if $A_iA_j = A_jA_i \quad \forall i,j = 0,...,p-1$,

otherwise

(ii) $F_{[i,j]} = \widetilde{F}^{\triangle}_{[i,j]}$ *if and only if* $i = 1$.

Proposition 1 shows that $F = \widetilde{F}^{\triangle}$ in the case of scalar processes or vector process with only one non-zero, non-diagonal sub-matrix. This is because in this near scalar case, the transverse block transpose operator shares the properties of the more familiar transpose operator. Under the more general case, it is still possible to map the top block row of F to the last block column of $\widetilde{F}$. Two further corollaries are worth highlighting. The first is that the relations used in the proof of proposition 1 also show that in the specific case of an integrated process, which has a zero root with $A_0 = 0$, then $F_{[p,j]} = \widetilde{F}^{\triangle}_{[p,j]} = 0, \quad \forall j$. The second is that Proposition 1 applies equally well when F and $\widetilde{F}$ are redefined as A^m and $\widetilde{A}^m$ for integer m. Such matrices are common in state-space solutions to temporally aggregated systems in discrete time, and our results can easily be translated to rival state space representations of such models.

3. THE CONTINUOUS TIME ARMA(2, 1) PROCESS

We now discuss the continuous time ARMA$(2, 1)$ system in some detail. This is the simplest case to capture the essential features of continuous time ARMA models while remaining relatively tractable, the integration-by-parts technique becoming unwieldy for higher order models. The system is therefore

$$D^2x(t) = a_0 + A_1Dx(t) + A_0x(t) + u(t) + \Theta_1Du(t), \quad -\infty < t < \infty. \tag{6}$$

We proceed as if the underlying model is stationary, although the non-stationary model where A_0 equals 0, is nested as a special (and simpler) case; see the application to consumption data in Chambers and Thornton (2011). In the first instance we consider the case in which the variables are stocks, i.e. the observed sequence is $x_1, x_2, \ldots, x_T$ where $x_t = x(t)$ $(t = 1, \ldots, T)$ and T denotes sample size. For ease of notation, we define the partition of a $2n \times 2n$ matrix K into four $n \times n$ sub-matrices as $K_{ij} \equiv K_{[i,j]}$ $(i, j = 1, 2)$.

3.1. Chambers and Thornton (2011) approach

The solution for the state vector is

$$y(t) = \int_{-\infty}^{t} [a + e^{A(t-s)}\Theta u(s)]ds,$$

where

$$A = \begin{pmatrix} A_1 & I \\ A_0 & 0 \end{pmatrix}, \quad \Theta = \begin{pmatrix} \Theta_1 \\ I \end{pmatrix}.$$

From the solution of (4) it is straightforward to show that $y(t)$ satisfies the stochastic difference equation

$$y(t) = c + Cy(t-1) + \epsilon(t), \quad \epsilon(t) = \int_{t-1}^{t} C(t-s)\Theta u(s)ds, \quad t = 1, \ldots, T, \tag{7}$$

where $C(r) = e^{rA}$, $C = C(1) = e^{A}$ and $c = \Phi a$, $\Phi = \Phi(1)$ and where $\Phi(r) = \int_{s=0}^{r} e^{sA}ds$. The objective is then to derive a stochastic difference equation for x_t.

Application of Corollary 1 in Chambers and Thornton (2011) with $p = 2$ yields

$$x_t = f + F_1 x_{t-1} + F_2 x_{t-2} + \eta_t, \tag{8}$$

with, after simplification, $F_1 = C_{11} + C_{12}C_{22}C_{12}^{-1}$, $F_2 = C_{12}[C_{21} - C_{22}C_{12}^{-1}C_{11}]$ and $f = [\Phi_{12} + C_{12}(\Phi_{22} - C_{22}C_{12}^{-1}\Phi_{12})]a_0$. Assumptions 1–3 in Chambers and Thornton (2011), which relate to reconstructability and detectability in optimal control theory, ensure that C_{12}^{-1} and C_{22}^{-1} exist.

Theorem 3 of Chambers (1999) enables the disturbance vector to be written

$$\eta_t = \eta(t) = C_0\epsilon(t) + C_1\epsilon(t-1),$$

where $\epsilon(t)$ is defined in (7), $C_0 = S_1$ and $C_1 = C_{12}[S_2 - C_{22}C_{12}^{-1}S_1]$. Simplifying,

$$\begin{aligned} \eta_t &= \int_{t-1}^{t} \{C_{12}(t-r) + C_{11}(t-r)\Theta_1\} u(r)dr \\ &+ C_{12}\int_{t-2}^{t-1} \big\{ \big[C_{22}(t-1-r) - C_{22}C_{12}^{-1}C_{12}(t-1-r)\big] \\ &+ \big[C_{21}(t-1-r) - C_{22}C_{12}^{-1}C_{11}(t-1-r)\big]\,\Theta_1 \big\} u(r)dr. \end{aligned}$$

3.2. Augmented Bergstrom (1983) approach

An alternative way of deriving the exact discrete model in (8) is to use the approach of Bergstrom (1983) with his continuous time white noise disturbance replaced by the continuous time MA(1). Theorem 2 of Bergstrom (1983) delivers

$$x_t = g + G_1 x_{t-1} + G_2 x_{t-2} + h_t, \tag{9}$$

where $G_1 = H_{11} + H_{12}H_{22}H_{12}^{-1}$, $G_2 = H_{12}[H_{21} - H_{22}H_{12}^{-1}H_{11}]$ and

$$H = e^{\widetilde{A}} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \quad \widetilde{A} = \begin{bmatrix} 0 & I \\ A_0 & A_1 \end{bmatrix}.$$

The intercept term is $g = (I - G_1 - G_2)A_0^{-1}a_0 = [\Psi_{12} + H_{12}(\Psi_{22} - H_{22}H_{12}^{-1}\Psi_{12})]a_0$ where

$$\Psi = \int_{r=0}^{1} e^{\widetilde{A}r} dr = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}.$$

Furthermore the discrete time disturbance in (9) is

$$h_t = h(t) = \int_{t-1}^{t} H_{12}(t-r)[u(r) + \Theta_1 Du(r)]dr + \int_{t-2}^{t-1} P(t-1-r)[u(r) + \Theta_1 Du(r)]dr,$$

where $H(r) = e^{r\widetilde{A}}$ and $P(r) = H_{12}[H_{22}(r) - H_{22}H_{12}^{-1}H_{12}(r)]$.

In the case of the continuous time ARMA $(2, 1)$, it is possible to evaluate these integrals of the derivative $Du(t)$ using the following integration by-parts formula, used by Simos (1996) and established formally by McCrorie (2000).

LEMMA *Let $x(t)$ be an $n \times 1$ continuous time mean square differentiable random process. Then the following integration-by-parts formula is valid:*

$$\int_{t-1}^{t} \psi(s)Dx(s)ds = \psi(t)x(t) - \psi(t-1)x(t-1) - \int_{t-1}^{t} \psi'(s)x(s)ds.$$

Applying the lemma to the first term in h_t involving $Du(r)$ gives

$$\begin{aligned}
& \int_{t-1}^{t} H_{12}(t-r)\Theta_1 Du(r)dr \\
= \; & H_{12}(0)\Theta_1 u(t) - H_{12}(1)\Theta_1 u(t-1) - \int_{t-1}^{t} H_{12}'(t-r)\Theta_1 u(r)dr \\
= \; & -H_{12}\Theta_1 u(t-1) + \int_{t-1}^{t} H_{22}(t-r)\Theta_1 u(r)dr.
\end{aligned}$$

The final equality results since $H'(t-r) = (d/dr)e^{\widetilde{A}(t-r)} = -\widetilde{A}H(t-r)$, $H(0) = I$ and $H(1) \equiv H$. In a similar fashion,

$$\begin{aligned}
& \int_{t-1}^{t} H_{22}(t-r)\Theta_1 Du(r)dr \\
= \; & H_{22}(0)\Theta_1 u(t) - H_{22}(1)\Theta_1 u(t-1) - \int_{t-1}^{t} H_{22}'(t-r)\Theta_1 u(r)dr \\
= \; & \Theta_1 u(t) - H_{22}\Theta_1 u(t-1) \\
& + A_0 \int_{t-1}^{t} H_{12}(t-r)\Theta_1 u(r)dr + A_1 \int_{t-1}^{t} H_{22}(t-r)\Theta_1 u(r)dr.
\end{aligned}$$

After substituting these expressions back, terms involving $u(t-1)$ and $u(t-2)$ cancel out, meaning that h_t, like η_t, comprises terms involving integrals with respect to $u(t)$,

$$\begin{aligned}
h_t &= \int_{t-1}^{t} \{H_{12}(t-r) + H_{22}(t-r)\Theta_1\} u(r)dr \\
&+ H_{12} \int_{t-2}^{t-1} \big\{[H_{22}(t-1-r) - H_{22}H_{12}^{-1}H_{12}(t-1-r)] \\
&+ [A_0 H_{12}(t-1-r) + A_1 H_{22}(t-1-r) - H_{22}H_{12}^{-1}H_{22}(t-1-r)]\Theta_1\big\} u(r)dr.
\end{aligned}$$

There are clear similarities in the relationship between the matrix exponentials and the discrete time autoregressive matrices in (8) and (9), although the underlying state space matrices are different (A and $\widetilde{A}$). The disturbance vectors do, however, appear to be rather different. The presence of the submatrices A_0 and A_1 in h_t means that even if one were to take for granted an equivalence of the autoregressive parts, the equivalence of the moving average errors does not necessarily follow. Nevertheless, this is established in the following proposition.

PROPOSITION 2. *Let $x(t)$ be generated by (6) and let the observed sequence be $x_t = x(t)$ $(t = 1, \ldots, T)$. Then the discrete time representations (8) and (9) are equivalent in the sense that: (i) $F_1 = G_1$; (ii) $F_2 = G_2$; (iii) $f = g$; and, (iv) $\eta_t = h_t$.*

The proof of the above result extends the method of Bergstrom (1983) to account for the continuous time MA disturbances. Although it is, in principle, applicable to higher-order models, doing so would require repeated sequential application of Lemma 1 which renders it the less attractive approach.

The various expressions for η_t and h_t do, however, help to highlight the impact of the moving average error in a continuous time model on the structure of its exact discrete representation. The Θ_1 matrix does not just scale all of the moments of η_t proportionally. While these moments are still subject to constraint by the C matrix, they are not rigidly determined by it and so the analyst has more freedom to produce a statistically acceptable model.

Although the recursive 'Gaussian' techniques in which the exact discrete representation is employed offer computational advantages over rival methods, they still involve finding the Choleski decomposition of the covariance matrix. This matrix has a block Toeplitz structure with non-zero blocks around the principal diagonal corresponding to the order of moving average in the error term in the exact discrete representation, which is of order $p-1$ if the data are stocks and p if the data contain flows, see Bergstom (1983). Because they do not increase this order, including a moving average error in the underlying continuous time model is the computationally far easier way of adding flexibility than increasing the order of the autoregressive part. The empirical relevance of continuous time models with moving average components was demonstrated by Chambers and Thornton (2011) in applications looking at sunspot, interest rate and household consumption data.

It is straightforward to extend the mapping to the case when x_t is a flow variable, i.e. when $x_t = \int_{t-1}^{t} x(\tau)d\tau$ $(t = 1, \ldots, T)$ with the consequence that $\eta_t = \int_{t-1}^{t} \eta(\tau)d\tau$ and $h_t = \int_{t-1}^{t} h(\tau)d\tau$. Demonstrating equivalence in models involving both stocks and flows is considerably more complicated and we do not attempt it here. The complications arise because the relevant selection matrices, as set out in Chambers (1999), no longer translate to the partitions of C and H used above. The same is largely true of higher order models where, for example, C_{11} no longer has the same dimension as C_{22} and the matrix C_{12} extends beyond the top block row of C. Establishing equivalence for scalar processes is more straightforward. Doing so enables us to invert the mapping between the continuous time and discrete time models and thus consider when it is possible to 'embed' a discrete time process in a continuous time one.

4. EMBEDDING A DISCRETE TIME ARMA IN A CONTINUOUS TIME ARMA

The question of whether it is possible for a real valued discrete time scalar ARMA process to have been generated by observing a continuous time ARMA process at regular discrete intervals, known as 'embedding' the discrete time process in a continuous time process, has been considered by, among others, Phillips (1959). More recently, He and Wang (1989) develop arguments that are closely aligned to the exact discrete representation based on expressing a discrete time ARMA (p, q), with $p > q$ as a $p + r$ dimensional AR(1), where r is the number of negative real roots in the autoregressive polynomial. Their decision not to discuss the impact of a moving average error is shown to be an oversight by Brockwell (1995) and by Brockwell and Brockwell (1998), who show that discrete time processes with zeros on the unit circle cannot be embedded in stationary continuous time processes. These analyses and that by Huzii (2007), who gives some general conditions and discusses the discrete time ARMA $(2,1)$ at length, are based on frequency domain representations of the processes. The exact discrete representation remains a valid analysis provided due care is taken over the moving average terms. One advantage is that, subject to initial conditions, it does not depend on stationarity. In addition, it naturally provides an explicit mapping from the four dimensional parameter space defining a discrete time ARMA $(2,1)$ to the four dimensional parameter space defining a continuous time process from which it may have been sampled, subject to consideration of aliasing, see Priestley (1981, pp 226). The following mapping also has practical uses enabling an analyst to use estimates from a discrete time model as initial values in a continuous time maximisation procedure.

Consider the discrete time ARMA $(2,1)$ process

$$y_t - b_1 y_{t-1} - b_2 y_{t-2} = (1-\mu L)(1-\nu L)y_t = (1-\phi L)e_t, \quad t = 1,\ldots,T, \tag{10}$$

where e_t is a zero mean white noise process with variance σ_e^2 and the continuous time process

$$D^2x(t) - a_1 Dx(t) - a_0 x(t) = (D-\alpha)(D-\beta)x(t) = u(t) + \theta Du(t), \quad -\infty < t < \infty, \tag{11}$$

where $u(t)$ is a zero mean white noise process with variance σ_u^2.

PROPOSITION 3. *Let y_t be generated by (10). Let $x(t)$ be generated by (11) and let the observed sequence be $x_t = x(t)$ $(t = 1,\ldots,T)$. Then the two processes are equivalent if:*

(i) $e^{\alpha} = \mu,\ e^{\beta} = \nu;$

(ii) $\theta = \pm\sqrt{\dfrac{\beta g(\mu,\nu;\phi) - \alpha g(\nu,\mu;\phi)}{\alpha\beta[\alpha g(\mu,\nu;\phi) - \beta g(\nu,\mu;\phi)]}};$ *and,*

(iii) $\sigma_u^2 = \dfrac{2\alpha\beta(\alpha^2-\beta^2)(1+\phi^2)\sigma_e^2}{\beta(1-\alpha^2\theta^2)(1-\mu^2)(1+\nu^2) - \alpha(1-\beta^2\theta^2)(1+\mu^2)(1-\nu^2)},$

where $g(x,y;\phi) = (1-x^2)[\phi(1+y^2) - (1+\phi^2)y] = (1-x^2)(\phi-y)(1-\phi y).$

Note that terms in odd powers of θ have all cancelled, meaning that moving average processes with the same absolute value for θ are observationally equivalent. Taking the positive square root ensures that the process is 'minimum phase', which corresponds to invertibility in a discrete time process.

The question of embeddability is equivalent to asking under what circumstances will the processes y_t and $x(t)$ be real. That is to say for which real parameters $(b_1, b_2, \phi, \sigma_e)$ does the mapping present parameters $(a_0, a_1, \theta, \sigma_u)$ that are also real numbers. For (b_1, b_2) to be real then μ and ν must either be conjugates or real valued. When μ and ν are a conjugate pair (including identical real roots), so are the parameters $\alpha = \ln(\mu\nu)/2 + i\lambda$ and $\beta = \ln(\mu\nu)/2 - i\lambda$, with $\lambda = \cos^{-1}(\frac{\mu+\nu}{2\sqrt{\mu\nu}}) + 2\pi j,\ j = 0, \pm 1, \pm 2, \ldots$ and so a_0 and a_1 are real. In this case the autoregressive parameters are subject to aliasing, with α and β identified only up to the addition of $2\pi ij$, with consequences for the identification of a_0, see the discussion in Brockwell (1995). When μ and ν are not conjugates, neither are α and β. The parameters a_0, a_1 will only be real if α and β are. This will be the case if μ and ν are positive and so $\alpha = \ln(\mu)$ and $\beta = \ln(\nu)$, but cannot be if either μ or ν are negative. Note that the need for both y_t and x_t to be real valued rules out aliasing in this case.

Given that ϕ is real, σ_u and the values of $x(t)$ are real if θ is. In the special case of conjugate roots where μ and ν are equal, real and negative, the expression in Proposition 3 (ii) simplifies to $\theta = i/\sqrt{a_0}$, which must be imaginary. It is not possible, therefore, to embed such discrete time processes in continuous time ARMA$(2,1)$ processes, although it may be possible in higher order continuous time ARMA processes, see Huzii(2007). For a given triple, (μ, ν, ϕ) there needs to exist (α, β), with $e^\alpha = \mu$ and $e^\beta = \nu$ such that the expression in Proposition 3(ii) is real. This is considered below.

PROPOSITION 4. *The real valued process y_t generated by (10) can be embedded in a continuous time process generated by (11) provided:*

(i) for real α, β:

$$\alpha\beta[g(\mu,\nu;\phi) + g(\nu,\mu;\phi)]^2 > (\alpha+\beta)^2[g(\mu,\nu;\phi)g(\nu,\mu;\phi)];$$

(ii)whereas for complex α, β:

$$\alpha\beta[g(\mu,\nu;\phi) + g(\nu,\mu;\phi)]^2 < (\alpha+\beta)^2[g(\mu,\nu;\phi)g(\nu,\mu;\phi)].$$

The conditions in Proposition 4 reflect the general condition that θ must be the square root of a positive real number. Since both the numerator and denominator are real if μ and ν are real, the test is whether the product of numerator and denominator is positive. If, however, μ, ν are a conjugate pair then so are α, β and $g(\mu,\nu;\phi), g(\nu,\mu;\phi)$ and both numerator and denominator are imaginary. The test is then that their product is real and negative.

The restrictions on the parameter ϕ given μ and ν to ensure embeddability can be determined by considering the shape and the roots of the quadratic functions: $q_1^N(1+\phi^2)+q_2^N\phi$ for the numerator and $\alpha\beta[q_1^D(1+\phi^2)+q_2^D\phi]$ for the denominator, with

$$\begin{aligned}
q_1^N &= \alpha(1-\nu^2)\mu-\beta(1-\mu^2)\nu,\\
q_2^N &= -[\alpha(1-\mu^2)(1+\nu^2)-\beta(1-\nu^2)(1+\mu^2)] = -(\alpha-\beta)(1-\mu^2\nu^2)-(\alpha+\beta))(\mu-\nu),\\
q_1^D &= [\beta(1-\mu^2)\nu-\alpha(1-\nu^2)\mu],\\
q_2^D &= -[\beta(1-\mu^2)(1+\nu^2)-\alpha(1-\nu^2)(1+\mu^2)] = (\alpha-\beta)(1-\mu^2\nu^2)-(\alpha+\beta))(\mu-\nu).
\end{aligned}$$

The sign of the numerator and denominator depends on whether they have real roots, that is on the sign of

$$(q_2^N)^2-4(q_1^N)^2=(q_2^D)^2-4(q_1^D)^2=(1-\mu^2)(1-\nu^2)[(\alpha-\beta)^2(1-\mu\nu)^2-(\alpha+\beta)^2(\mu-\nu)^2],$$

and on the signs of q_1^N q_1^D.

In the case of complex conjugate roots, $\alpha=a+ib$ and $\beta=a-ib$, with $b>0$, without loss of generality, then $\mu=e^\alpha[\cos(b)+i\sin(b)]$ while $\nu=e^\alpha[\cos(b)-i\sin(b)]$ and the above expressions become

$$\begin{aligned}
q_1^N &= -2e^a[a\sin(b)(1+e^{2a})+b\cos(b)(1-e^{2a})]i,\\
q_2^N &= -2[4ae^{2a}\cos(b)\sin(b)+b(1-e^{4a})]i,\\
q_1^D &= -2e^a[a\sin(b)(1+e^{2a})-b\cos(b)(1-e^{2a})]i,\\
q_2^D &= -2[4ae^{2a}\cos(b)\sin(b)-b(1-e^{4a})]i,\\
(q_2^N)^2 &- 4(q_1^N)^2=(q_2^D)^2-4(q_1^D)^2=4\left\{1+e^{4a}-e^{2a}[\cos^2(b)-\sin^2(b)]\right\}\times\\
&\quad [b^2(1-e^{2a})^2-4a^2e^{2a}\sin^2(b)]i^2.
\end{aligned}$$

We consider the case of the stationary CARMA(2, 1) in more detail. In the case of real roots, with $\mu\geq\nu$ without loss of generality, it can be shown (see appendix) that $q_1^N>0$ and $q_1^D>0$, meaning that both the numerator and denominator are convex. This means that both numerator and denominator have real roots and so change sign over their domain. For $\alpha\neq\beta$ these sign changes will not occur simultaneously and so there will be blind spots - two regions of ϕ for given $\mu\neq\nu$ that cannot be hit with a CARMA(2, 1).

The case of complex roots is complicated by the possibility of aliasing, which has noticeably different effects from aliasing on the autoregressive and moving average parts. Adding (or subtracting) multiples of 2π to the coefficient b has no effect on μ and ν while changing the shape of numerator and denominator, potentially violating the conditions in Proposition 4 for some values but not for others.

The embeddability literature has so far concentrated on whether it is possible to embed a discrete stock variable in a continuous time models. Not all discrete time processes are stocks, however. Since their exact discrete representation is known, see Chambers and Thornton (2011), this method opens up the possibility of considering flow variables and even vectors containing both stock and flow variables. This is worthy of further research.

5. CONCLUDING COMMENTS

This paper has explored different methods for deriving the exact discrete time representations of continuous time autoregressive moving average processes. We have demonstrated the equivalence between the exact discrete time representations for data generated by a continuous time ARMA$(2, 1)$ system by the methods of Bergstrom (1983), with a supporting lemma, and Chambers and Thornton (2011). The exploration of the two forms gives further insight into the effect of introducing a moving average error on the exact discrete representation. The moments of the error term remain influenced, but not determined, by the underlying autoregressive model, offering the analyst greater flexibility to model the data parsimoniously at a relatively small additional cost in computation. When applied to univariate models, the exact discrete time representation enables a mapping to be drawn from the parameters of an observed discrete ARMA to those of a continuous time process from which it may have been sampled. By considering when that mapping is to the real line rather than the complex plane, we have considered the question of embeddability.

Acknowledgements: The first author's research was supported by an Economic and Social Research Council studentship no PTA-030-2006-00395.

REFERENCES

Bergstrom, A. R. (1983) Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica* 51, 117–152.

Bergstrom, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In: Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics vol. 2*, North-Holland.

Bergstrom, A. R. (1986) The estimation of open higher-order continuous time dynamic models with mixed stock and flow data. *Econometric Theory* 2, 350–373.

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

Bergstrom, A. R. (1997) Gaussian estimation of mixed-order continuous-time dynamic models with unobservable stochastic trends from mixed stock and flow data. *Econometric Theory* 13, 467–505.

Brockwell, A. E. and Brockwell, P. J. (1998) A class of non-embeddable ARMA Processes. *Journal of Time Series Analysis* 20, 483–86.

Brockwell, P. J. (1995) A note on the embedding of discrete time ARMA Processes. *Journal of Time Series Analysis* 16, 451–60.

Brockwell, P. J. (2004) Representations of Continuous-Time ARMA Processes. *Journal of Applied Probability* 41, 375–382.

Chambers, M. J. (1999) Discrete time representation of stationary and non-stationary continuous time systems. *Journal of Economic Dynamics and Control* 23, 619–639.

Chambers, M. J. (2009) Discrete time representations of cointegrated continuous time models with mixed sample data. *Econometric Theory* 25, 1030–1049.

Chambers, M. J. and Thornton, M. A. (2011) Discrete time representations of continuous time ARMA processes. *Econometric Theory* (forthcoming).

Ercolani (nèe McGarry), J. S. (2005) The Exact Discrete Time Representation of a System of Fourth-Order Differential Equations. *Computers and Mathematics with Applications, 46, 213–230.*

He, S. W. and Wang, J. G. (1989) On embedding a discrete parameter ARMA model in a continuous parameter ARMA model.*Journal of Time Series Analysis* 10, 315–23.

Harvey, A. C. and Stock, J. H. (1985) The estimation of higher-order continuous time autoregressive models. *Econometric Theory* 1, 97–117.

Huzii, M. (2007) 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.

McCrorie, J. R. (2000) Deriving the exact discrete analog of a continuous time system. *Econometric Theory*, 16, 998-1015.

McCrorie, J. R. (2001) Interpolating exogenous variables in continuous time dynamic models. *Journal of Economic Dynamics and Control, 25, 1399–1427.*

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

Priestley, M. B. (1981) *Spectral Analysis and Time Series.* Academic Press.

Robinson, P. M. (1976) Fourier estimation of continuous time models. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 215–266. North-Holland.

Simos, T. (1996) Gaussian Estimation of a Continuous Time Dynamic Model with Common Stochastic Trends. *Econometric Theory* 12, 361–373.

Zadrozny, P. (1988) Gaussian likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies. *Econometric Theory* 4, 108–124.

APPENDIX

Proof of Proposition 1. Note that $I^{\triangle} = I$ and $X^{\triangle} + Y^{\triangle} = (X + Y)^{\triangle}$ for conformable matrices, X and Y. Establishing the proposition for the matrix exponential then reduces to establishing it for arbitrary integer powers of A and $\widetilde{A}$, which we do by induction. First suppose that for some positive integer r, $(A^{\triangle})^r = (A^r)^{\triangle}$. Evidently this is true for $r = 1$. Then note that every square matrix must commute in multiplication with powers of itself. Given this, the necessary and sufficient condition that the i, j'th block of $(A^{\triangle})^{r+1}$ should equal the i, j'th block of $(A^{r+1})^{\triangle}$ is that

$$(A^{\triangle})^{r+1}_{[i,j]} = \sum_{k=1}^{p} A^{\triangle}_{[i,k]} (A^{\triangle})^{r}_{[k,j]} = \sum_{k=1}^{p} A_{[k,p+1-i]} A^{r}_{[p+1-j,k]},$$

must equal

$$(A^{r+1})^{\triangle}_{[i,j]} = A^{r+1}_{[p+1-j,p+1-i]} = \sum_{k=1}^{p} A^{r}_{[p+1-j,k]} A_{[k,p+1-i]}.$$

In other words, the sub-matrices that give rise to the i,j'th block should 'commute in sum'. This establishes the sufficiency of (i).

Now consider the following equivalent expressions for the matrix A^{r+1}, firstly AA^r and then A^rA:

$$\begin{pmatrix} A_{p-1}A^r_{[1,1]}+A^r_{[2,1]} & A_{p-1}A^r_{[1,2]}+A^r_{[2,2]} & \cdots & A_{p-1}A^r_{[1,j]}+A^r_{[2,j]} & \cdots \\ A_{p-2}A^r_{[1,1]}+A^r_{[3,1]} & A_{p-2}A^r_{[1,2]}+A^r_{[3,2]} & \cdots & A_{p-2}A^r_{[1,j]}+A^r_{[3,j]} & \cdots \\ A_{p-3}A^r_{[1,1]}+A^r_{[4,1]} & A_{p-3}A^r_{[1,2]}+A^r_{[4,2]} & \cdots & A_{p-3}A^r_{[1,j]}+A^r_{[4,j]} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{p-i}A^r_{[1,1]}+A^r_{[i+1,1]} & A_{p-i}A^r_{[1,2]}+A^r_{[i+1,2]} & \cdots & A_{p-i}A^r_{[1,j]}+A^r_{[i+1,j]} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_0A^r_{[1,1]} & A_0A^r_{[1,2]} & \cdots & A_0A^r_{[1,j]} & \cdots \end{pmatrix};$$

$$\begin{pmatrix} \sum_{k=1}^p A^r_{[1,k]}A_{p-k} & A^r_{[1,1]} & \cdots & A^r_{[1,j-1]} & \cdots \\ \sum_{k=1}^p A^r_{[2,k]}A_{p-k} & A^r_{[2,1]} & \cdots & A^r_{[2,j-1]} & \cdots \\ \sum_{k=1}^p A^r_{[3,k]}A_{p-k} & A^r_{[3,1]} & \cdots & A^r_{[3,j-1]} & \cdots \\ \vdots & \vdots & & \vdots & \\ \sum_{k=1}^p A^r_{[i,k]}A_{p-k} & A^r_{[i,1]} & \cdots & A^r_{[i,j-1]} & \cdots \\ \vdots & \vdots & & \vdots & \\ \sum_{k=1}^p A^r_{[p,k]}A_{p-k} & A^r_{[p,1]} & \cdots & A^r_{[p,j-1]} & \cdots \end{pmatrix}.$$

We need only consider whether the 'commute in sum' property holds in the first block column. The expansion of A^rA establishes that $A^{r+1}_{[i,j]} = A^r_{[i,j-1]}, j = 2,\ldots,p, \forall i$, from which it follows from that if the condition holds for $A^r_{[i,1]},\ \ \forall r \leq \underline{r}$ it holds for all $A^r_{[i,j]},\ \ \forall r \leq \underline{r}+j-1$, but if it fails for $A^{\bar{r}}_{[i,1]}$ then it fails for $A^{\bar{r}+j-1}_{[i,j]}$.

The sufficiency of (ii) then depends on showing that $A^{r+1}_{[1,1]} = \sum_{k=1}^p A^r_{[1,k]}A_{p-k} = \sum_{k=1}^p A_{p-k}A^r_{[1,k]}$. Comparing terms along the principal diagonals above establishes $A^r_{[i,i-1]} = A_{p-i}A^r_{[1,i]} + A^r_{[i+1,i]}$ for $i = 2,3,\ldots,p-1$. Recursive substitution of this equality, ending with $A^r_{[p,p-1]} = A_0A^r_{[1,p]}$, into AA^r establishes 'commute in sum' for this block.

Then note that this condition refers only to sub-matrices in the first block row of A^r. Given that the mapping holds for $r = 1$ across this block row, it holds $\forall r$ across the block row regardless of whether it holds across other block rows, establishing the sufficiency of (ii).

To see the necessity of (i) and (ii), consider the generic term $A^{r+1}_{[i,1]}$. Note that the matrices

A_{p-i} and $A^r_{[1,1]}$ are not paired together in multiplication in the bottom expression for $i > 1$. Nor are they multiplied together to form any other block element. Thus $\sum_{k=1}^p A^r_{[i,k]}A_{p-k} = \sum_{k=1}^p A_{p-k}A^r_{[i,k]}$ for given $i > 1$ if and only if $A^r_{[i,k]}A_{p-k} = A_{p-k}A^r_{[i,k]}, \ \ \forall k$. $\square$

Proof of Proposition 2. We first establish some useful mappings. For the CARMA (2, 1), we simplify the above to

$$A^{r+1} = \begin{bmatrix} A^r_{11}A_1 + A^r_{12}A_0 & A^r_{11} \\ A^r_{21}A_1 + A^r_{22}A_0 & A^r_{21} \end{bmatrix} = \begin{bmatrix} A_1A^r_{11} + A^r_{21} & A_1A^r_{12} + A^r_{22} \\ A_0A^r_{11} & A_0A^r_{12} \end{bmatrix},$$

and

$$\widetilde{A}^{r+1} = \begin{bmatrix} \widetilde{A}^r_{12}A_0 & \widetilde{A}^r_{11} + \widetilde{A}^r_{12}A_1 \\ \widetilde{A}^r_{22}A_0 & \widetilde{A}^r_{21} + \widetilde{A}^r_{22}A_1 \end{bmatrix} = \begin{bmatrix} \widetilde{A}^r_{21} & \widetilde{A}^r_{22} \\ A_0\widetilde{A}^r_{11} + A_1\widetilde{A}^r_{21} & A_0\widetilde{A}^r_{12} + A_1\widetilde{A}^r_{22} \end{bmatrix}$$

Comparing blocks, it follows that from the definition of the matrix exponential that:

$$\begin{aligned} C_{11}(s) &= A_1C_{12}(s) + C_{22}(s); \\ C_{21}(s) &= A_0C_{12}(s); \\ C_{11}(s)A_1 + C_{12}(s)A_0 &= A_1C_{11}(s) + C_{21}(s); \\ H_{22}(s) &= H_{11}(s) + H_{12}(s)A_1; \text{ and,} \\ H_{21}(s) &= H_{12}(s)A_0. \end{aligned}$$

While, from proposition 1, it follows that:

$$\begin{aligned} H_{22}(s) &= C^{\triangle}_{22}(s) &= C_{11}(s); \text{ and,} \\ H_{12}(s) &= C^{\triangle}_{12}(s) &= C_{12}(s). \end{aligned}$$

We use these mappings to establish:

$$\begin{aligned} \text{(i)} \ \ G_1 &= H_{11} + H_{12}H_{22}H_{12}^{-1} = C_{11} - C_{12}A_1 + C_{12}\left(A_1C_{12} + C_{22}\right)C_{12}^{-1} \\ &= C_{11} + C_{12}C_{22}C_{12}^{-1} = F_1; \end{aligned}$$

and,

$$\begin{aligned} \text{(ii)} \ \ G_2 &= H_{12}\left(H_{21} - H_{22}H_{12}^{-1}H_{11}\right) \\ &= C_{12}\left(C_{12}A_0 - (C_{22} + A_1C_{12})C_{12}^{-1}(C_{11} - C_{12}A_1)\right) \\ &= C_{12}\left(C_{12}A_0 - A_1C_{11} + A_1C_{12}A_1 + C_{22}A_1 - C_{22}C_{12}^{-1}C_{11}\right) \\ &= C_{12}\left(C_{21} - C_{22}C_{12}^{-1}C_{11}\right) = F_2. \end{aligned}$$

(iii) Since $\Phi(r)$ is also the sum of the identity matrix and sequential ascending powers of

A, the mappings within and between C and H apply within and between Φ and Ψ. Equality between f and g comes from applying these mappings and noting that

$$\Psi_{22} - H_{22}H_{12}^{-1}\Psi_{12} = (A_1\Phi_{12} + \Phi_{22}) - (A_1C_{12} + C_{22})C_{12}^{-1}\Phi_{12} = \Phi_{22} - C_{22}C_{12}^{-1}\Phi_{12}$$

(iv) To show equivalence between h_t and η_t, we utilise the mappings from proposition 1, to write,

$$\begin{aligned} h_t &= \int_{t-1}^{t} \{C_{12}(t-r) + C_{11}(t-r)\Theta_1\}\, u(r)dr \\ &+ C_{12}\int_{t-2}^{t-1} \left\{C_{11}(t-1-r) - C_{11}C_{12}^{-1}C_{12}(t-1-r)\right\} u(r)dr \\ &+ C_{12}\int_{t-2}^{t-1} \left\{A_0C_{12}(t-1-r) + (A_1 - C_{11}C_{12}^{-1})C_{11}(t-1-r)\right\} \Theta_1 u(r)dr. \end{aligned}$$

Since $C_{11}(t-1-r) - C_{11}C_{12}^{-1}C_{12}(t-1-r) = C_{22}(t-1-r) - C_{22}C_{12}^{-1}C_{12}(t-1-r)$, results from substituting $C_{11}(t-1-r) = A_1C_{12}(t-1-r) + C_{22}(t-1-r)$ into the second line and by substituting $A_0C_{12}(t-1-r) = C_{21}(t-1-r)$ and $(A_1 - C_{11}C_{12}^{-1}) = C_{22}C_{12}^{-1}$ into the final line establishes equivalence. □

Proof of Proposition 3. In this scalar case we base our analysis around Lemma 1. Firstly we factorise (11)

$$(D-\alpha)(D-\beta)x(t) = u(t) + \theta Du(t),$$

then repeated solution of the differential equations gives the exact discrete representation,

$$(1 - e^{\alpha}L)(1 - e^{\beta}L)x_t = \eta_t,$$

where

$$\eta_t = \int_{s=t-1}^{t}\int_{r=s-1}^{s} e^{\beta(t-s)}e^{\alpha(s-r)}[u(r) + \theta Du(r)]drds.$$

The mapping in (i) follows directly. The mappings in (ii) and (iii) depend on matching the covariance structure of η_t to that of $(1-\phi L)e_t$. To this end, note that

$$\begin{aligned} &\int_{s=t-1}^{t}\int_{r=s-1}^{s} e^{\beta(t-s)}e^{\alpha(s-r)}\zeta(r)drds \\ =& \int_{r=t-1}^{t}\int_{s=r}^{t} e^{\beta t}e^{-\alpha r}e^{(\alpha-\beta)s}\zeta(r)dsdr + \int_{r=t-2}^{t-1}\int_{s=t-1}^{r+1} e^{\beta t}e^{-\alpha r}e^{(\alpha-\beta)s}\zeta(r)dsdr \\ =& \int_{r=t-1}^{t} \frac{1}{\alpha-\beta}[e^{\alpha(t-r)} - e^{\beta(t-r)}]\zeta(r)dr - \int_{r=t-2}^{t-1} \frac{1}{\alpha-\beta}[e^{\beta}e^{\alpha(t-1-r)} - e^{\alpha}e^{\beta(t-1-r)}]\zeta(r)dr, \end{aligned}$$

and that by Lemma 1,

$$\int_{r=s-1}^{s} e^{\alpha(s-r)} D\zeta(r)dr = \zeta(s) - e^{\alpha}\zeta(s-1) + \alpha \int_{r=s-1}^{s} e^{\alpha(s-r)}\zeta(r)dr.$$

Using these expressions we can write η_t as

$$\begin{aligned}
& \int_{r=t-1}^{t} \left\{ \frac{1+\alpha\theta}{\alpha-\beta} \left[e^{\alpha(t-r)} - e^{\beta(t-r)} \right] + \theta e^{\beta(t-r)} \right\} u(r)dr \\
& - \int_{r=t-2}^{t-1} \left\{ \frac{1+\alpha\theta}{\alpha-\beta} \left[e^{\alpha(t-1-r)} e^{\beta} - e^{\alpha} e^{\beta(t-1-r)} \right] + \theta e^{\alpha} e^{\beta(t-1-r)} \right\} u(r)dr \\
= \; & \frac{1}{\alpha-\beta} \int_{s=0}^{1} \left\{ \left[(1+\alpha\theta)e^{\alpha s} - (1+\beta\theta)e^{\beta s} \right] u(t-s) \right. \\
& \left. - \left[(1+\alpha\theta)e^{\alpha s}e^{\beta} - (1+\beta\theta)e^{\alpha}e^{\beta s} \right] u(t-s-1) \right\} ds,
\end{aligned}$$

the final equality following a change of variables. It then follows that $E\left\{\eta_t^2\right\}$

$$\begin{aligned}
= \; & \frac{\sigma_u^2}{(\alpha-\beta)^2} \int_0^1 \left\{ \left[(1+\alpha\theta)e^{\alpha s} - (1+\beta\theta)e^{\beta s} \right]^2 + \left[(1+\alpha\theta)e^{\alpha s}e^{\beta} - (1+\beta\theta)e^{\alpha}e^{\beta s} \right]^2 \right\} ds \\
= \; & \frac{\sigma_u^2}{(\alpha-\beta)^2} \int_{s=0}^1 \left\{ (1+\alpha\theta)^2 e^{2\alpha s}(1+e^{2\beta}) \right. \\
& \left. -2(1+\alpha\theta)(1+\beta\theta)(1+e^{(\alpha+\beta)})e^{(\alpha+\beta)s} + (1+\beta\theta)^2(1+e^{2\alpha})e^{2\beta s} \right\} ds \\
= \; & \frac{-\sigma_u^2}{(\alpha-\beta)^2} \left\{ \frac{(1+\alpha\theta)^2}{2\alpha}(1-e^{2\alpha})(1+e^{2\beta}) \right. \\
& \left. -2\frac{(1+\alpha\theta)(1+\beta\theta)}{\alpha+\beta}(1-e^{2(\alpha+\beta)}) + \frac{(1+\beta\theta)^2}{2\beta}(1+e^{2\alpha})(1-e^{2\beta}) \right\} \\
= \; & \frac{-\sigma_u^2}{2\alpha\beta(\alpha+\beta)(\alpha-\beta)^2} \left\{ \beta(\alpha+\beta)(1+\alpha^2\theta^2)(1-\mu^2)(1+\nu^2) \right. \\
& \left. -4\alpha\beta(1+\alpha\beta\theta^2)(1-\mu^2\nu^2) + \alpha(\alpha+\beta)(1+\beta^2\theta^2)(1+\mu^2)(1-\nu^2) \right\} \\
= \; & \frac{\sigma_u^2}{2\alpha\beta(\alpha^2-\beta^2)} \left\{ \beta(1-\alpha^2\theta^2)(1-\mu^2)(1+\nu^2) - \alpha(1-\beta^2\theta^2)(1+\mu^2)(1-\nu^2) \right\},
\end{aligned}$$

where a common factor of $(\alpha-\beta)$ has been cancelled. Note that terms in θ^1 have cancelled.

Using similar methods, $E\left\{\eta_t\eta_{t-1}\right\}$

$$\begin{aligned}
= \; & \frac{-\sigma_u^2}{(\alpha-\beta)^2} \int_0^1 \left\{ \left[(1+\alpha\theta)e^{\alpha s} - (1+\beta\theta)e^{\beta s} \right] \left[(1+\alpha\theta)e^{\alpha s}e^{\beta} - (1+\beta\theta)e^{\alpha}e^{\beta s} \right] \right\} ds \\
= \; & \frac{-\sigma_u^2}{(\alpha-\beta)^2} \int_0^1 \left\{ (1+\alpha\theta)^2 e^{2\alpha s} e^{\beta} \right. \\
& \left. -(1+\alpha\theta)(1+\beta\theta)(e^{\alpha}+e^{\beta})e^{(\alpha+\beta)s} + (1+\beta\theta)^2 e^{\alpha} e^{2\beta s} \right\} ds \\
= \; & \frac{\sigma_u^2}{(\alpha-\beta)^2} \left\{ \frac{(1+\alpha\theta)^2}{2\alpha}(1-e^{2\alpha})e^{\beta} \right. \\
& \left. -\frac{(1+\alpha\theta)(1+\beta\theta)}{\alpha+\beta}(e^{\alpha}+e^{\beta})(1-e^{(\alpha+\beta)}) + \frac{(1+\beta\theta)^2}{2\beta}e^{\alpha}(1-e^{2\beta}) \right\} \\
= \; & \frac{\sigma_u^2}{2\alpha\beta(\alpha+\beta)(\alpha-\beta)^2} \left\{ \beta(\alpha+\beta)(1+\alpha^2\theta^2)(1-\mu^2)\nu \right.
\end{aligned}$$

$$
\begin{aligned}
& -2\alpha\beta(1+\alpha\beta\theta^2)(\mu+\nu)(1-\mu\nu)+\alpha(\alpha+\beta)(1+\beta^2\theta^2)\mu(1-\nu^2)\} \\
= \; & \frac{\sigma_u^2}{2\alpha\beta(\alpha^2-\beta^2)}\left\{-\beta(1-\alpha^2\theta^2)(1-\mu^2)\nu+\alpha(1-\beta^2\theta^2)\mu(1-\nu^2)\right\},
\end{aligned}
$$

where terms in θ^1 have again cancelled. Equating

$$
\frac{E\left\{\eta_t\eta_{t-1}\right\}}{E\left\{\eta_t^2\right\}}=-\frac{\phi}{1+\phi^2},
$$

and solving for θ produces the mapping in (ii). The mapping in (iii) comes from equating the variances of the two processes.

□

The Covariance Stationary CARMA (2, 1)

We take, without loss of generality, $0>\alpha\geq\beta$, implying $0<\nu\leq\mu<1$. Given this it is straightforward to show that the denominator is convex since $\alpha\beta\geq 0$ and $q_1^D\geq 0$ since

$$
\begin{aligned}
(1-\mu^2)\nu-(1-\nu^2)\mu &= (\nu-\mu)(1+\mu\nu)\leq 0\Rightarrow \\
\alpha(1-\mu^2)\nu-\beta(1-\nu^2)\mu &\geq \alpha(1-\mu^2)\nu-\alpha(1-\nu^2)\mu\geq 0.
\end{aligned}
$$

At the same time $q_1^N>$ is equal to

$$
\begin{aligned}
\beta(1-\mu^2)\nu-\alpha(1-\nu^2)\mu &= \beta(1-e^{2\alpha})e^{\beta}-\alpha(1-e^{2\beta})e^{\alpha}= \\
-2\alpha\beta e^{\beta}\int_0^1 e^{2\alpha s}ds+2\alpha\beta e^{\alpha}\int_0^1 e^{2\beta s}ds &= 2\alpha\beta e^{\alpha}e^{\beta}\int_{-1/2}^{1/2}(e^{2\beta s}-e^{2\alpha s})ds,
\end{aligned}
$$

which is non-negative.

To show that both have real roots consider that $(1-\mu^2)(1-\nu^2)>0$ and that $[(\alpha-\beta)^2(1-\mu\nu)^2-(\alpha+\beta)^2(\mu-\nu)^2]$ is the difference between two squares. One of its factors

$$
(\alpha-\beta)(1-\mu\nu)-(\alpha+\beta)(\mu-\nu)\geq 0,
$$

is a negative number subtracted from a positive number. For the other, note that

$$
\begin{aligned}
& (\alpha-\beta)(1-\mu\nu)+(\alpha+\beta)(\mu-\nu)=(\alpha-\beta)(1-e^{\alpha+\beta})+(\alpha+\beta)e^{\beta}(e^{\alpha-\beta}-1) \\
= \; & (\beta^2-\alpha^2)\int_{s=0}^{1}[e^{(\alpha+\beta)s}-e^{\beta}e^{(\alpha-\beta)s}]ds \\
= \; & (\beta^2-\alpha^2)e^{(\alpha+\beta)/2}\int_{s=-1/2}^{1/2}[e^{(\alpha+\beta)s}-e^{(\alpha-\beta)s}]ds \\
= \; & (\beta^2-\alpha^2)e^{(\alpha+\beta)/2}\left\{\int_{s=0}^{1/2}e^{\alpha s}[e^{\beta s}-e^{-\beta s}]ds+\int_{s=-1/2}^{0}e^{\alpha s}[e^{\beta s}-e^{-\beta s}]ds\right\}
\end{aligned}
$$

$$= \quad (\beta^2 - \alpha^2) e^{(\alpha+\beta)/2} \int_{s=0}^{1/2} [e^{\alpha s} - e^{-\alpha s}][e^{\beta s} - e^{-\beta s}] ds \geq 0,$$

since both $[e^{\alpha s} - e^{-\alpha s}]$ and $[e^{\beta s} - e^{-\beta s}]$ are negative for $s \in (0, \frac{1}{2}]$.