Statistica Sinica **6**(1996), 715-732

EXISTENCE AND STABILITY OF CONTINUOUS TIME THRESHOLD ARMA PROCESSES

O. Stramer, R. L. Tweedie and P. J. Brockwell

University of Iowa, Colorado State University and Royal Melbourne Institute of Technology

Abstract: In this paper we define a class of continuous-time threshold ARMA (CTARMA) processes uniquely in terms of the weak solution of a certain stochastic differential equation, and investigate stability properties of these processes. We apply criteria for stability of weak solutions (see Meyn and Tweedie (1993b), Stramer and Tweedie (1994) and Stramer and Tweedie (1996)) to CTARMA processes and thus obtain criteria for transience, Harris recurrence, positive Harris recurrence and geometric ergodicity for these processes. In order to do this it is shown that CTARMA processes satisfy suitable continuity conditions, and so can be analyzed as φ-irreducible T-processes (Meyn and Tweedie (1993b)).

Key words and phrases: Continuous-time SETARMA models, exponential ergodicity, irreducible Markov processes, non-linear time series, recurrence, stochastic differential equations, stationary distributions, transience.

1. Introduction

In recent years there has been great interest in the development of non-linear models for discrete time processes observed at equally spaced times. One class of nonlinear models which has been particularly fruitful includes the self-exciting threshold ARMA processes (SETARMA models). Basically SETARMA models are piecewise linear ARMA models in which the linear relationship varies over regimes delineated by the threshold values. A process $\{X_t\}$ is a SETARMA model of order (l, p, q) (with $0 \le q < p$) and delay parameter d (see Tong (1990)) with thresholds $-\infty = r_0 < r_1 < \cdots < r_l = \infty$ if it is a solution of the equation

$$X_t = a_0^{(i)} + \sum_{j=1}^{p} a_j^{(i)} X_{t-j} + \sum_{j=0}^{q} b_j^{(i)} e_{t-j}, \qquad r_{i-1} \le X_{t-d} < r_i, \tag{1}$$

where $a_j^{(i)}$ and $b_j^{(i)}$ $i = 1, \ldots, l$ are constants, and $\{e_t\}$ is a white noise sequence with unit variance.

Many data sets are, in fact, observations of continuous-time processes at discrete times. As pointed out in the linear case, the use of continuous-time models also facilitates the analysis of irregularly spaced data, which are common in practice (Jones (1981)). They may be partially observed with some missing observations or irregularly observed with an arbitrary sampling interval.

Continuous-time analogues of the non-linear model (1) have been developed recently (see Brockwell, Hyndman and Grunwald (1991), Tong and Yeung (1991), Brockwell and Hyndman (1992), Brockwell (1994) and Stramer, Brockwell and Tweedie (1996)). Inference for such models has been based on the conditional Gaussian likelihood of the data (see Brockwell and Hyndman (1992) and Brockwell (1994)). If the process is stationary then it is possible to compute (and numerically maximize) the unconditional Gaussian likelihood provided the first two moments of the stationary and transition distributions of the underlying state process can be computed. Necessary and sufficient conditions for the existence of a stationary distribution in the special case of a CTAR(1) process, as well as an explicit expression for the stationary mean and variance have been developed in Stramer, Brockwell and Tweedie (1996).

In this paper we find sufficient conditions for transience and positive Harris recurrence for a more general class of CTARMA(p,q) processes with constant moving average parameters and scale parameter and with $p \geq 2$. We estimate the stationary mean and variance using the approximating sequence of Brockwell and Hyndman (see Brockwell and Hyndman (1992) and Brockwell (1994)) with t large, in the case when there exists a stationary solution.

We then find conditions for geometric rates of convergence to stationarity; and this strengthens the justification for the use of the approximating sequence for estimating the stationary moments.

2. CTARMA Models: Definition and Existence

A continuous-time (linear) ARMA, or CARMA(p,q), process, with $0 \leq q < p$ is defined (see Brockwell and Hyndman (1992) and Brockwell (1994)) to be a strong solution of the p-th order linear differential equation

$$Y^{(p)}(t)+a_1Y^{(p-1)}(t)+\cdots+a_pY(t) = \sigma[W^{(1)}(t)+b_1W^{(2)}(t)+\cdots+b_qW^{(q+1)}(t)+c], \tag{2}$$

where the superscript $^{(j)}$ denotes j-fold differentiation with respect to t, $\{W(t)\}$ is standard Brownian motion and $a_1,\ldots,a_p$, $b_1,\ldots,b_q$, c and σ (the *scale parameter*) are constants. We assume that $\sigma > 0$ and $b_q \neq 0$ and define $b_j := 0$ for $j > q$. Since the derivatives $W^{(j)}(t), j > 0$ do not exist in the usual sense, we interpret (2) as being equivalent to the *observation* and *state* equations,

$$Y(t) = \sigma \mathbf{b}'\mathbf{X}(t), \quad t \geq 0, \tag{3}$$

and

$$d\mathbf{X}(t) = A\mathbf{X}(t)dt + \mathbf{e}(c\ dt + dW(t)), \tag{4}$$

where

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

and (4) is an Itô differential equation for the state vector $\mathbf{X}(t)$. We assume also that $\mathbf{X}(0)$ is independent of $\{W(t)\}$. The state-vector $\mathbf{X}(\mathbf{t})$ is in fact the vector of derivatives,

$$\mathbf{X}(t) = \begin{bmatrix} X(t) \\ X^{(1)}(t) \\ \vdots \\ X^{(p-1)}(t) \end{bmatrix},$$

of the continuous-time AR(p) process to which $\{Y(t)\}$ reduces when $b_j = 0$, $j \geq 1$.

The process $\{Y(t), t \geq 0\}$ is thus equivalently said to be a CARMA(p, q) process with parameters $(a_1, \ldots, a_p, b_1, \ldots, b_q, \sigma, c)$ if $Y(t) = \sigma[1\ b_1\ \cdots\ b_{p-1}]$ $\mathbf{X}(t)$ where $\{\mathbf{X}(t)\}$ is a strong solution of (4). The solution of (4) can be written as

$$\mathbf{X}(t) = e^{At}\mathbf{X}(0) + \int_0^t e^{A(t-u)}\mathbf{e}\, dW(u) + c\int_0^t e^{A(t-u)}\mathbf{e}\, du.$$

Hence $\mathbf{X}(t)$ is a Gaussian diffusion process which is stationary if and only if

$$\mathbf{X}(0) \sim N\Big(c\int_0^\infty e^{Ay}\mathbf{e}dy, \int_0^\infty e^{Ay}\mathbf{e}\ \mathbf{e}' e^{A'y}dy\Big)$$

and all the eigenvalues of A (i.e. the roots of $z^p + a_1 z^{p-1} + \cdots + a_p = 0$) have negative real parts (see Proposition 6.2 of Ichihara and Kunita (1974)). Note that it is easily shown, using the spectral representation of the matrix A, that $c\int_0^\infty e^{Ay}\mathbf{e}dy = a_p^{-1}c[1\ 0\ \cdots\ 0]$.

Based on the CARMA(p, q) process, we now develop a continuous time analogue of the discrete-time SETARMA process of Tong (1990). To do this assume we are given a fixed linear function $f(x) = \sum_{i=1}^p d_i x_i$, where $d_i \geq 0$, $i = 1, \ldots, p$ are constants, and allow the parameters $a_1, \ldots, a_p, b_1, \ldots, b_q$, c and σ of the process $\{Y(t)\}$ in the defining equations (3) and (4) to depend on the linear function $f(\mathbf{X}(t))$ of the state vector $\mathbf{X}(t)$, taking fixed values in each of the l regions $r_{i-1} \leq f(\mathbf{X}(t)) < r_i$, $i = 1, \ldots, l$ where $-\infty = r_0 < r_1 < \cdots < r_l = \infty$.

In particular, the change of dynamics might depend on $f(\mathbf{X}(t)) = Y(t) = \sigma\mathbf{b}'\mathbf{X}(t)$ or, in the case when there is no MA part (i.e. $b_j = 0$ for $j \leq q$), on the velocity $X'(t)$ or the acceleration $X''(t)$ of the process $X(t)$.

In this paper we assume that only the autoregressive coefficients $a_1, \ldots, a_p$ and c change with the value of $f(\mathbf{X}(t))$. This restriction is satisfied by all threshold *autoregressive* models with constant σ. The case of non-constant σ seems much harder (although if $p = 1$, $q = 0$ it can be solved (see Stramer, Brockwell and Tweedie (1996)). We shall also restrict attention to the case of a single threshold, since the extension to more than one is straightforward.

Thus, we define $\{Y(t)\}$ to be a CTARMA(p, q) process with threshold at r, and constants $b_1, \ldots, b_q, \sigma$, analogously to (3) and (4), except that we allow the parameters $a_1, \ldots, a_p$ and c to depend on the linear function $f(\mathbf{X}(t))$ of the state vector $\mathbf{X}(t)$ in such a way that

$$a_i(f(\mathbf{X}(t))) = a_i^{(j)}, \quad i = 1, \ldots, p \ ; \quad c(f(\mathbf{X}(t))) = c^{(j)}, \tag{5}$$

where $j = 1$ or 2 according as $f(\mathbf{X}(t)) \leq r$ or $f(\mathbf{X}(t)) > r$. All of our results will be true for any segmentation of the state space into regions over which the instantaneous mean function of $\mathbf{X}(t)$ in $\{\mathbf{x} : f(\mathbf{x}) \leq r\}$ is $A^{(1)}\mathbf{X}(t) + c^{(1)}\mathbf{e}$ and in $\{\mathbf{x} : f(\mathbf{x}) > r\}$ is $A^{(2)}\mathbf{X}(t) + c^{(2)}\mathbf{e}$, where $A^{(j)}$, $j = 1, 2$ are defined in (4) with elements $a_i^{(j)}$, $i = 1, \ldots, p$.

We first show, following the argument of Brockwell (1994) and Brockwell and Stramer (1995), that (4) with $\mathbf{X}(0) = \mathbf{x} = [x_1 \ \cdots \ x_p]'$, and coefficients as defined in (5) has a unique (in law) weak solution $\{\mathbf{X}(t)\}$, and determine the distribution of $\mathbf{X}(t)$ for any given $\mathbf{X}(0) = \mathbf{x}$. These distributions determine, in particular, the joint distribution of the values of the process $\{Y(t)\}$ at times $t_1, \ldots, t_N$, given $\mathbf{X}(0)$.

Before proving this we need the following notation. Let B^{x_p} be Brownian motion with $B^{x_p}(0) = x_p$ defined on the probability space $(C[0, \infty), \mathcal{B}[0, \infty), P_{x_p})$ and let $\mathcal{F}_t{}^x = \sigma\{B^{x_p}(s), s \leq t\} \vee \mathcal{N}^{x_p}$, where $\mathcal{N}^{x_p}$ is the sigma algebra of P_{x_p}-null sets of $\mathcal{B}[0, \infty)$ and P_{x_p} denotes the law of B^{x_p}; we use E_{x_p} to denote expectation relative to P_{x_p}.

We now have

Theorem 2.1. *For any* $\mathbf{x}$*, the stochastic differential equation* (4) *with coefficients* $a_1(f(\mathbf{X}(t))), \ldots, a_p(f(\mathbf{X}(t)))$ *and* $c(f(\mathbf{X}(t)))$ *as defined in* (5) *has a weak solution with initial condition* $\mathbf{X}(0) = \mathbf{x}$ *and this solution is unique in the sense of probability law.*

Proof. Assuming that $\mathbf{X}(0) = \mathbf{x}$, we can write the components $X_1(t), \ldots, X_p(t)$ of the state vector $\mathbf{X}(t)$ in terms of $\{X_p(s), 0 \leq s \leq t\}$ using the relations $X_{p-1}(t) = x_{p-1} + \int_0^t X_p(s)ds$, $\ldots$, $X_1(t) = x_1 + \int_0^t X_2(s)ds$. The resulting functional relationship will be denoted by

$$\mathbf{X}(t) = \mathbf{F}^{\mathbf{x}}(X_p, t). \tag{6}$$

Substituting from (6) into (4), we see that it can be written in the form

$$dX_p = D^{\mathbf{x}}(X_p, t)dt + dW(t), \tag{7}$$

where $D^{\mathbf{x}}(X_p, t)$, like $\mathbf{F}^{\mathbf{x}}(X_p, t)$, depends on $\{X_p(s), 0 \le s \le t\}$.

The equations $dZ_1 = Z_2dt$, $dZ_2 = Z_3dt$, $\ldots$, $dZ_{p-1} = Z_pdt$, $dZ_p = dB^{x_p}(t)$, with $\mathbf{Z}^{\mathbf{x}}(0) = \mathbf{x} = [x_1 \ \cdots \ x_p]'$, clearly have the unique strong solution $\mathbf{Z}^{\mathbf{x}}(t) = \mathbf{F}^{\mathbf{x}}(B^{x_p}, t)$, where $\mathbf{F}^{\mathbf{x}}$ is defined as in (6). Let $D^{\mathbf{x}}$ be the functional appearing in (7) and suppose that $\hat{B}^{x_p}$ is the Itô integral defined by $\hat{B}^{x_p}(0) = x_p$ and

$$d\hat{B}^{x_p}(t) = -D^{\mathbf{x}}(B^{x_p}, t)dt + dB^{x_p}(t) = -D^{\mathbf{x}}(Z_p^{\mathbf{x}}, t)dt + dZ_p^{\mathbf{x}}(t). \tag{8}$$

If we define the new measure $\hat{P}_{x_p}$ on $(C[0,\infty), \mathcal{B}[0,\infty))$ satisfying

$$d\hat{P}_{x_p} = M^{\mathbf{x}}(B^{x_p}, t)dP_{x_p},$$

where

$$M^{\mathbf{x}}(B^{x_p}, t) = \exp\left[-\frac{1}{2}\int_0^t [D^{\mathbf{x}}(B^{x_p}, s)]^2 ds + \int_0^t D^{\mathbf{x}}(B^{x_p}, s)dB^{x_p}(s)\right],$$

then, by the Cameron-Martin-Girsanov formula (see e.g. Oksendal (1992), p.126), $\hat{B}^{x_p}$ is Brownian motion under $\hat{P}_{x_p}$. Hence, from (8) we see that $(B^{x_p}, \hat{B}^{x_p})$ is a weak solution of (7) (on $(C[0,\infty), \mathcal{B}[0,\infty), \hat{P}_{x_p}, \{\mathcal{F}_t\})$) with initial condition $X_p(0) = x_p$. Hence, $(\mathbf{Z}^{\mathbf{x}}(t), \hat{B}^{x_p}(t))$ is a weak solution (on the same probability space) of the Equation (4) with initial condition $\mathbf{X}(0) = \mathbf{x}$ and coefficients as defined in (5). By Proposition 5.3.10 of Karatzas and Shreve (1991) and by Theorem 10.2.2 of Stroock and Varadhan (1979) the weak solution is unique in law.

Corollary 2.2. *With the notation developed in the above proof and with* $0 \le t_1 < t_2 < \cdots < t_n < T$, *we have*

$$\hat{P}_{x_p}[(\mathbf{X}_{t_1}, \ldots, \mathbf{X}_{t_n}) \in \Gamma] \ = \ E_{x_p}[\mathbb{1}_{\{(\mathbf{F}^{\mathbf{x}}(B^{x_p}, t_1), \ldots, \mathbf{F}^{\mathbf{x}}(B^{x_p}, t_n)) \in \Gamma\}} M^{\mathbf{x}}(B^{x_p}, T)]; \quad \Gamma \in \mathcal{B}(\mathbb{R}^{pn}). \tag{9}$$

Remark 2.3. The importance of Equation (9) is that it gives the conditional distribution of the state-vector $\mathbf{X}(t)$ (and in particular of the CTARMA process, $Y(t) = \sigma \ \mathbf{b}'\mathbf{X}(t)$) given $\mathbf{X}(0) = \mathbf{x}$, as an expectation of a functional of the Brownian motion $\{B^{x_p}(t)\}$ starting at x_p.

3. Irreducibility and Continuous Components of CTARMA Processes

Our aim in the following sections is to investigate stability properties of CTARMA processes. We consider the state vector $\mathbf{X}(t)$ of a CTARMA process as a continuous time-homogeneous Markov process evolving on $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$, with transition probability $P^t(\mathbf{x}, A) = \mathsf{P}_{\mathbf{x}}(\mathbf{X}(t) \in A),\ \mathbf{x} \in \mathbb{R}^p, A \in \mathcal{B}(\mathbb{R}^p)$; here $\mathcal{B}(\mathbb{R}^p)$ denotes the Borel σ-field on $\mathbb{R}^p$.

Following Meyn and Tweedie (1993a), in order to investigate stability of $\mathbf{X}$, for a measurable set $A \subseteq \mathbb{R}^p$, we denote the hitting times and occupation times of $\mathbf{X}$ by

$$\tau_A^{\mathbf{X}} = \inf\{t \geq 0 : \mathbf{X}(t) \in A\}, \qquad \eta_A^{\mathbf{X}} = \int_0^\infty 1\!\!1\{\mathbf{X}(t) \in A\}\, dt.$$

We write

$$L^{\mathbf{X}}(\mathbf{x}, A) = \mathsf{P}_{\mathbf{x}}(\tau_A^{\mathbf{X}} < \infty); \qquad U^{\mathbf{X}}(\mathbf{x}, A) = \mathsf{E}_{\mathbf{x}}[\eta_A^{\mathbf{X}}].$$

A Markov process is called *φ-irreducible* if for the σ-finite measure φ

$$\varphi\{B\} > 0 \Longrightarrow U^{\mathbf{X}}(\mathbf{x}, B) > 0, \qquad \forall \mathbf{x} \in \mathbb{R}^p.$$

Stability properties of Markov processes are much easier to establish for T-processes, introduced in Meyn and Tweedie (1993a).

Definition 3.1. A process will be called a T-process if there exists a kernel $T(x, A)$ such that
(i) For $A \in \mathcal{B}(\mathbb{R}^p)$ the function $T(\cdot, A)$ is lower semi-continuous;
(ii) For all $x \in \mathbb{R}^p$ and $A \in \mathcal{B}(\mathbb{R}^p)$, the measure $T(x, \cdot)$ satisfies, for some probability measure a on $[0, \infty)$, $\int_0^\infty P^t(x, A)\, a(dt) \geq T(x, A)$ with $T(x, \mathbb{R}^p) > 0$ for all $x \in \mathbb{R}^p$.

We now show that the state vectors of CTARMA processes with constant σ are μ^{Leb}-irreducible T-processes where μ^{Leb} denotes Lebesgue measure. We first need a preliminary lemma concerning the process $\{\mathbf{F}^{\mathbf{x}}(B^{x_p}, t)\}$ which appears in the key representation (9).

Lemma 3.2. *If $\{B^{x_p}(t)\}$ is Brownian motion started at x_p, then the process $\{\mathbf{F}^{\mathbf{x}}(B^{x_p}, t)\}$ defined as in* (6) (*with deterministic initial state* $\mathbf{x}$) *is a Gaussian diffusion process. The covariance matrix $V(t)$ of $\mathbf{F}^{\mathbf{x}}(B^{x_p}, t)$* (*strictly positive definite for $t > 0$*) *can be written as*

$$V(t) = \int_0^t e^{H(t-y)}\mathbf{e}\ \mathbf{e}' e^{H'(t-y)} dy,$$

where

$$H = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \qquad \textit{and} \qquad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Proof. Observe that $\mathbf{F}^{\mathbf{x}}(B^{x_p}, t)$ satisfies the stochastic differential equation (4) with $c = a_1 = \cdots = a_p = 0$. Since H is the value of the coefficient matrix A when $a_1 = \cdots = a_p = 0$, it suffices, by Proposition 6.2 of Ichihara and Kunita (1974), to check that $\operatorname{rank}(H^0\mathbf{ee}', H\mathbf{ee}', \ldots, H^{p-1}\mathbf{ee}') = p$. This elementary calculation completes the proof.

We now turn to the main result of this section.

Theorem 3.3. *Any weak solution* $\mathbf{X}(t)$ *of* (4) *with coefficients as defined in* (5) *is a* μ^{Leb}*-irreducible* T*-process. Moreover, the transition probability* $P^t(\mathbf{x}, A)$ *is positive, whenever* $\mu^{\text{Leb}}(A) > 0$*, for all* $\mathbf{x} \in \mathbb{R}^p$ *and all* $t > 0$*.*

Proof. We first note that, by Corollary 2.2,

$$\hat{P}_{x_p}[(\mathbf{X}_t) \in \Gamma] = E_{x_p}[\mathbb{1}_{\{\mathbf{F}^{\mathbf{x}}(B^{x_p},t)\in\Gamma\}} M^{\mathbf{x}}(B^{x_p}, t)]; \qquad \Gamma \in \mathcal{B}(\mathbb{R}^p). \tag{10}$$

We have immediately from (10) and Lemma 3.2 that the transition probability $P^t(x, A)$ is positive whenever $\mu^{\text{Leb}}(A) > 0$, for all $\mathbf{x} \in \mathbb{R}^p$ and $t > 0$. The proof that $\mathbf{X}(t)$ is a T-process will now follow directly from Theorem 6.4 of Tweedie (1994), provided we can show that $\mathbf{X}$ has the weak Feller property, i.e. $E_{\mathbf{x}}[g(\mathbf{X}_t)]$ is a continuous function of $\mathbf{x}$, for all bounded continuous functions $g\colon \mathbb{R}^p \to \mathbb{R}$ and all $t \geq 0$.

In the notation of (6) and (7), let

$$\tilde{\mathbf{F}}^{\mathbf{x}}(\tilde{B}, t) = \mathbf{F}^{\mathbf{x}}(B^{x_p}, t); \qquad \tilde{M}^{\mathbf{x}}(\tilde{B}, t) = M^{\mathbf{x}}(B^{x_p}, t),$$

where $B^{x_p} = \tilde{B} + x_p$ and $\tilde{B}$ is a standard Brownian motion (with $\tilde{B}(0) = 0$). Then

$$E_{\mathbf{x}}[g(\mathbf{X}_t)] = E_0[g(\tilde{\mathbf{F}}^{\mathbf{x}}(\tilde{B}, t))\tilde{M}^{\mathbf{x}}(\tilde{B}, t)].$$

It is easy to show that, with probability one, as $\mathbf{x_n} \to \mathbf{x}$

$$g(\tilde{\mathbf{F}}^{\mathbf{x}_n}(\tilde{B}, t))\tilde{M}^{\mathbf{x}_n}(\tilde{B}, t) \to g(\tilde{\mathbf{F}}^{\mathbf{x}}(\tilde{B}, t))\tilde{M}^{\mathbf{x}}(\tilde{B}, t).$$

By Example 16.21 of Billingsley ((1986), p.223), it then suffices for the Feller property to show that $\{g(\tilde{\mathbf{F}}^{\mathbf{x}_n}(\tilde{B}, t))\tilde{M}^{\mathbf{x}_n}(B, t)\}$ is uniformly integrable. According to Corollary 3.5.16 of Karatzas and Shreve (1991) $M^{\mathbf{x}}(B^{x_p}, t)$ is a martingale under the measure P_{x_p} and hence for each $\mathbf{x} \in \mathbb{R}^p$, $E_0[\tilde{M}^{\mathbf{x}}(\tilde{B}, t)] = 1$. We conclude from the positivity of $\tilde{M}^{\mathbf{x}}(\tilde{B}, t)$ and from Example 16.21 of Billingsley ((1986), p.223), that $\{\tilde{M}^{\mathbf{x}_n}(\tilde{B}, t)\}$ is also uniformly integrable. Finally we observe that $g(\tilde{M}^{\mathbf{x}_n}(\tilde{B}, t))$ is uniformly bounded in n and hence $\{g(\tilde{\mathbf{F}}^{\mathbf{x}_n}(\tilde{B}, t))\tilde{M}^{\mathbf{x}_n}(\tilde{B}, t)\}$ is uniformly integrable. Therefore $\mathbf{X}$ has the weak Feller property and this completes the proof.

4. Recurrence and Ergodicity Properties

It is known that if $\mathbf{X}$ is a time-homogeneous irreducible Markov process on $\mathbb{R}^p$ then we have a dichotomy between recurrence and transience (see Tweedie (1994)). The process $\mathbf{X}$ is called Harris recurrent if for some σ-finite measure μ, $L^{\mathbf{X}}(x, A) \equiv 1$ whenever $\mu\{A\} > 0$, and transient if there exists a countable cover of $\mathbb{R}^p$ by sets A_j such that $U^{\mathbf{X}}(x, A_j) \le M_j < \infty$ for all j. For a discussion of these concepts see Meyn and Tweedie (1993a) and Tweedie (1994). It is also known (cf. Azéma, Duflo and Revuz (1967) and Getoor (1979)) that if $\mathbf{X}$ is Harris recurrent then an essentially unique invariant measure π exists. If the invariant measure is *finite*, then it may be normalized to a probability measure; in this case X is called *positive Harris recurrent*.

In general (under some aperiodicity assumptions) we then have convergence of P^t to π. However, for the models here we can say much more. For a function $V \ge 1$ we say that X is *V-exponentially ergodic* if there exists a constant $\beta < 1$ and $R < \infty$ such that

$$\|P^t(x,\cdot) - \pi\|_V \le V(x)\, R\, \beta^t \quad \text{for all} \quad t \in \mathbb{R}_+,\ x \in \mathbb{R}^p,$$

where the V-norm $\|\cdot\|_V$ is defined by

$$\|P^t(x,\cdot) - \pi\|_V = \sup_{|g|\le V}\Big|\int_{\mathbb{R}^p} P^t(x,dy)g(y) - \int_{\mathbb{R}^p} \pi(dy)g(y)\Big|.$$

In this section we find criteria for ergodicity and transience of the state vectors of CTARMA processes in terms of $A^{(j)}$, $j = 1, 2$, where $A^{(j)}$ are defined in (4) with elements $a_i^{(j)}$. All of our criteria will be based on use of the Itô formula leading to "drift conditions" on the generator of the process, and we describe this briefly here for completeness.

Let $V \in C^2(\mathbb{R}^p)$, where $C^2(\mathbb{R}^p)$ is the class of functions on $\mathbb{R}^p$ with continuous first and second partial derivatives. Then

$$dV(X(t)) = L_V(X(t))dt + \frac{\partial V(X(t))}{\partial X_p(t)} dW(t),$$

where L_V is the second-order differential operator defined as

$$\begin{aligned}(L_V)(x) \;=\; & \frac{1}{2}\,\frac{\partial^2 V(x)}{\partial x_p^2} - a_1^j\, x_p \frac{\partial V(x)}{\partial x_p} - \cdots - a_p^j\, x_1 \frac{\partial V(x)}{\partial x_p} \\ & + c^j \frac{\partial V(x)}{\partial x_p} + x_p \frac{\partial V(x)}{\partial x_{p-1}} + \cdots + x_2 \frac{\partial V(x)}{\partial x_1} \qquad (11)\end{aligned}$$

and $j = 1$ or 2 according as $f(x) \le r$ or $f(x) > r$. It is easy to verify that the "local mean drift" $L_V(X(t))$ of $V(X(t))$ is the extended generator as defined in Meyn and Tweedie (1993b) and hence we can apply the drift criteria for general continuous time Markov processes (Meyn and Tweedie (1993b)) and in particular for weak solutions to stochastic differential equations (Stramer and Tweedie (1994b)) to our models. From this approach we will be able to conclude that a CTARMA process is exponentially ergodic if there is a compact set C towards which the process drifts, in the sense that for some constant $c \geq 0$ and constant $d < \infty$, $L_V(x) \leq -cV(x) + d\mathbb{1}_C(x)$, and conversely, if there is outward drift in the sense that for some bounded function $h \geq 0$, $L_V(x) \geq h(x)\mathbb{1}_{C^c}(x)$ then the process is transient. In this section and the next we use specific forms of the test function V in these results to classify the CTARMA models.

If $x = [x_1 \ \cdots \ x_p]$ and $y = [y_1 \ \cdots \ y_p]$ then we use the inner product notation $(x, y) = \sum_{i=1}^{p} x_i\, y_i$, and $\|x\|^2 = (x, x)$. We will need a preliminary well known lemma.

Lemma 4.1. (i) *Let M be a symmetric matrix and let $\lambda_{min}^{(M)}$ and $\lambda_{max}^{(M)}$ be the smallest and largest eigenvalues of M, respectively. Then*

$$\lambda_{min}^{(M)}\|x\|^2 \leq x^T Mx \leq \lambda_{max}^{(M)}\|x\|^2$$

(ii) *Let A be an $n \times n$ matrix and let B be a symmetric $n \times n$ matrix. Then $(Bx, Ax) = \frac{1}{2}x^T Nx$ where $N = A^T B + BA$.*

We first give a sufficient condition for V-exponential ergodicity.

Theorem 4.2. *If there exists a positive definite matrix P such that $-N^{(j)}$, $j = 1, 2$ are positive definite matrices, where $N^{(j)} = (A^{(j)})^T P + P(A^{(j)})$, then the state vector* **X** *of the CTARMA(p,q) process is V-exponentially ergodic with $V = \|x\|^2 + 1$.*

Proof. We first show that inequality (15) of Stramer and Tweedie (1996) is satisfied for large $\|x\|$ and this choice of V. Setting $B = P$, we have from Lemma 4.1 that

$$(Bx, A^{(j)}x) \leq \frac{1}{2}\lambda_{max}^{(N^{(j)})}\|x\|^2; \qquad (Bx, x) \geq \lambda_{min}^{(B)}\|x\|^2.$$

Using the assumption that $-N^{(j)}$, $j = 1, 2$ and B are positive definite, we have that $\lambda_{max}^{(N^{(j)})} < 0$ and $\lambda_{min}^{(B)} > 0$ and hence (15) of Stramer and Tweedie (1996) holds for $\|x\|$ large. We next note that by Theorem 3.3 any skeleton chain is a μ^{Leb}-irreducible T-process. The proof follows now from Proposition 4.2 of Stramer and Tweedie (1996).

We first note that the condition of Theorem 4.2 does not depend on the segmentation of the state space and in particular does not depend on **b**, where **b** (the MA part of the CTARMA process) is defined in (4).

We also note that this theorem gives sufficient conditions for existence of a second order stationary solution as well as V-norm convergence of the distributions for each initial condition. In particular this shows that

$$\lim_{t\to\infty} E_x(Y_t) = \int_{\mathbb{R}^p} \mathbf{b}'\mathbf{x}\,\pi(d\mathbf{x}) \tag{12}$$

and

$$\lim_{t\to\infty} E_x(Y_t{}^2) = \int_{\mathbb{R}^p} (\mathbf{b}'\mathbf{x})^2\,\pi(d\mathbf{x}), \tag{13}$$

where π denotes the unique invariant probability measure for $\mathbf{X}$ and the CTARMA process Y_t is defined in (3).

For the linear case (i.e. $A^{(1)} = A^{(2)} = A$), the condition of the theorem holds when all the eigenvalues of A have negative real parts (see Theorem 8-20 of Chen (1984)). We conjecture that this is true for CTARMA processes also, but have not so far proved this. However, in the case $p = 2$ we will use a different method in Section 5 to prove that if all the eigenvalues of $A^{(1)}$ and $A^{(2)}$ have negative real parts and the change of dynamics of the process $X(t)$ depends on either the state of the process at time t or the state of its velocity at time t, then the state vector $\mathbf{X}(t) = [X(t)\ X'(t)]'$ of the CTAR(2) process is V-exponentially ergodic.

We next give a sufficient condition for transience.

Theorem 4.3. *The state vector* $\mathbf{X}$ *of CTARMA*(p,q) *process is transient if there exists a positive definite matrix* P *such that for* $j = 1, 2$ $N^{(j)} = (A^{(j)})^T P + P(A^{(j)})$ *are positive definite matrices.*

Proof. We infer from Lemma 4.1 that if we set $B = P$ then

$$(Bx, A^{(j)}x) \geq \frac{1}{2}\lambda_{min}^{(N^j)}\|x\|^2; \qquad \frac{(B^2x, x)}{(Bx, x)} \leq \frac{\lambda_{max}^{(B^2)}}{\lambda_{min}^{(B)}}.$$

Since $\lambda_{min}^{(N^j)} > 0$, $\lambda_{max}^{(B^2)} > 0$ and $\lambda_{min}^{(B^2)} > 0$, it follows from Proposition 3.5 of Stramer and Tweedie (1996), $\mathbf{X}$ is transient.

Remark 4.4. For the linear case (i.e. $A^{(1)} = A^{(2)} = A$), the condition of the theorem holds when all the eigenvalues of A have positive real parts (see Theorem 8-20 of Chen (1984)).

5. The One and Two Dimensional Cases

We can give more explicit results when $p \leq 2$. In the one dimensional case ($p = 1$), we showed in Stramer, Brockwell and Tweedie (1996) that we could identify explicitly conditions for recurrence, transience and exponential ergodicity. There we proved, in particular

Theorem 5.1. *Let X be a CTAR(1) process. Then X is Harris recurrent if and only if*

$$\lim_{|x|\to\infty}[a(x)x^2 - 2c(x)x] < 0.$$

Moreover, if this condition is satisfied then
(i) *X is positive Harris recurrent and the stationary distribution has probability density*

$$\pi(x) = k\sigma^{-2}(x)\exp\{-\sigma^{-2}(x)[a(x)x^2 - 2c(x)x]\} \tag{14}$$

and k is the uniquely determined constant such that $\int_{-\infty}^{\infty}\pi(x)dx = 1$.
(ii) *X is V-exponentially ergodic, where $V = x^{2m}+1$ or $V = \exp\{x^{2m}\}+1$ for $m \in \mathbb{Z}_+$.*

Note also that, in the case of $p = 1$, we can in fact allow σ to depend also on the threshold. (See Stramer, Brockwell and Tweedie (1996) for details.)

We now discuss the CTAR(2) process in more detail. Here it is possible to give more results than in the case when $p > 2$, although the results are not as complete as those of Theorem 5.1 for CTAR(1) processes.

We first give a sufficient condition for the state vector $\mathbf{X}$ of the CTAR(2) process to be V-exponentially ergodic.

Theorem 5.2. *If all the eigenvalues of*

$$A^{(1)} = \begin{bmatrix} 0 & 1 \\ -a_2^1 & -a_1^1 \end{bmatrix}, \qquad A^{(2)} = \begin{bmatrix} 0 & 1 \\ -a_2^2 & -a_1^2 \end{bmatrix} \tag{15}$$

(*i.e. the roots of $z^2 + a_1^{(j)}z + a_2^{(j)} = 0$ for $j = 1, 2$) have negative real parts and $f(x)$ as defined in (5) is either x_1 or x_2, then the state vector $\mathbf{X}$ of the CTAR(2) process is V-exponentially ergodic where $V(x) = x^t B(f(x))\,x + 1$, $x = [x_1\ x_2]$ and $B(f(x))$ is positive definite.*

Proof. By Proposition 6.1 of Meyn and Tweedie (1993b) it is suffices to prove that
(i) some skeleton chain is an irreducible T-process;
(ii) there exists a function $V \in C^2(\mathbb{R}^2)$ such that V is positive, $V(x) \to \infty$ as $x \to \infty$, and for some $c > 0$ and $d < \infty$,

$$(L_V)(x) \le -cV(x) + d \qquad x \in \mathbb{R}^2, \tag{16}$$

where L_V is the second-order differential operator defined as in (11) with $p = 2$: that is

$$(L_V)(x) = \frac{1}{2}\frac{\partial^2 V(x)}{\partial x_2^2} - a_1^{(j)}x_2\frac{\partial V(x)}{\partial x_2} - a_2^{(j)}x_1\frac{\partial V(x)}{\partial x_2} + c^{(j)}\frac{\partial V(x)}{\partial x_2} + x_2\frac{\partial V(x)}{\partial x_1} \tag{17}$$

and $j = 1$ or 2 according as $f(x) \le r$ or $f(x) > r$.
But (i) follows directly from Theorem 3.3, so we need only prove that (ii) holds.

We first assume that $f(x) = x_1$. Let ϵ be a constant such that

$$0 < \epsilon < \min\Big(\frac{a_1^1}{M}, \frac{a_1^2}{M}\Big),$$

where $M > 2$ is an arbitrary constant. For any $\delta > 0$ we define

$$P(x_1) = \left[\begin{array}{ll} p_{11}(x_1) & \epsilon \\ \epsilon & \frac{1}{2} \end{array} \right],$$

where $p_{11}(x_1) = a_1^{(j)}\,\epsilon + \frac{a_2^{(j)}}{2}$ and $j = 1$ or 2 according as $x_1 \le r - \delta$ or $x_1 > r + \delta$. For $r - \delta < x_1 \le r + \delta$, we define $p_{11}(x_1)$ such that $p_{11}(x_1) \in C^2(\mathbb{R}^2)$. We next define

$$V_1(x) := x^t P(x_1) x = p_{11}(x_1) x_1^2 + 2\epsilon x_1\, x_2 + \frac{1}{2} x_2^2. \tag{18}$$

By Equation (17)

$$\begin{aligned}(L_{V_1})(x) \;=\; & \frac{1}{2} - a_1(x_1)\, x_2(2\epsilon\, x_1 + x_2) - a_2(x_1)\, x_1(2\epsilon\, x_1 + x_2) + c(x_1)(2\epsilon\, x_1 + x_2) \\ & + x_2(x_1^2 {p_{11}}'(x_1) + 2x_1\, p_{11}(x_1) + 2\epsilon\, x_2).\end{aligned}$$

It is easy to see that for large values of $\|x\|$ we get:
(a1) if $x_1 \le r - \delta$ or $x_1 > r + \delta$ then

$$(L_{V_1})(x) \sim -x^t\, N^{(j)}\, x \quad \text{where} \quad N^{(j)} = \left[\begin{array}{ll} 2\epsilon\, a_2^{(j)} & 0 \\ 0 & a_1^{(j)} - 2\epsilon \end{array} \right]$$

and $j = 1$ or 2 according as $x_1 \le r - \delta$ or $x_1 > r + \delta$. We assume that M is large enough so that $N^{(j)}$ is positive definite for $j = 1$ and 2.
(a2) if $r - \delta < x_1 \le r + \delta$ then $(L_{V_1})(x) \sim -x_2^2\,(a_1^{(j)} - 2\epsilon)$, and $j = 1$ or 2 according as $r - \delta < x_1 \le r$ or $r < x_1 \le r + \delta$.

Let $a_{min} = \min\{a_1^{(j)} - 2\epsilon : j = 1, 2\}$ and $K = \min\{\lambda_{min}^{(N^1)}, \lambda_{min}^{(N^2)}, a_{min}\}/2$. By Lemma 4.1 and positivity of K we get that there exists some constant d such that

$$(L_{V_1})(x) \le -K\|x\|^2 + d \qquad x \in \mathbb{R}^2.$$

It is readily seen that $P(x_1)$ is positive definite for all $x_1 \in \mathbb{R}$ and hence we have from Lemma 4.1 that there exist some constants $\hat{K}$ and $\hat{d}$ such that

$$(L_{V_1})(x) \le -\hat{K} V_1(x) + \hat{d} \qquad x \in \mathbb{R}^2.$$

This completes the proof of (ii) for $f(x) = x_1$.

We next assume that $f(x) = x_2$. For any $\delta > 0$ we define

$$P(x_2) = \begin{bmatrix} 1 & \epsilon \\ \epsilon & p_{22}(x_2) \end{bmatrix},$$

where $p_{22}(x_2) = (1 - a_1^{(j)}\epsilon)/a_2^j$ and $j = 1$ or 2 according as $x_2 \le r-\delta$ or $x_2 > r+\delta$. For $r - \delta < x_2 \le r + \delta$, we again define $p_{22}(x_2)$ such that $p_{22}(x_2) \in C^2(\mathbb{R}^2)$. We also assume that $\epsilon > 0$ is small enough so that $P(x_2)$ is positive definite for all $x_2 \in C^2(\mathbb{R}^2)$. We now define

$$V_2(x) := x^t P(x_2)x = x_1^2 + 2\epsilon x_1\, x_2 + p_{22}(x_2)x_2^2. \tag{19}$$

Using the same argument as for $f(x) = x_1$ we can show that for large values of $\|x\|$ we get:
(b1) if $x_2 \le r - \delta$ or $x_2 > r + \delta$ then $(L_{V_2})(x) \sim -x^t\, N^{(j)}\, x$, where

$$N^{(j)} = \begin{bmatrix} 2\epsilon\, a_2^{(j)} & 0 \\ 0 & 2a_1^{(j)}(1 - a_1^{(j)}\epsilon)(a_2^{(j)})^{-1} - 2\epsilon \end{bmatrix}$$

and $j = 1$ or 2 according as $x_2 \le r - \delta$ or $x_2 > r + \delta$. We assume that ϵ is small enough so that $N^{(j)}$ is positive definite for $j = 1$ and 2.
(b2) if $r - \delta < x_2 \le r + \delta$ then $(L_{V_2})(x) \sim -x_1^2\, 2\epsilon\, a_2^{(j)}$, and $j = 1$ or 2 according as $r - \delta < x_2 \le r$ or $r < x_2 \le r + \delta$.
We now use Lemma 4.1 twice to show that (ii) holds for $f(x) = x_2$.

Finally, since (ii) holds for $f(x) = x_1$ and $f(x) = x_2$ we conclude that the process $X(t)$ is V*-exponentially ergodic* where $V(x) = V_1(x) + 1$ if $f(x) = x_1$ and $V(x) = V_2(x) + 1$ if $f(x) = x_2$.

Remark 5.3. We now sketch an extension of the proof of Theorem 5.2 to indicate how to find a broader class of functions V such that the state vector $\mathbf{X}$ of a CTAR(2) processes is V-exponentially ergodic. The basic idea is first to note that by using Theorem 8-20 of Chen (1984) we can find two test functions V_{1j}, $j = 1, 2$ for the linear cases

$$d\mathbf{X}^{(j)}(t) = A^{(j)}\mathbf{X}(t)dt + \mathbf{e}(c^{(j)}\, dt + dW(t)),$$

such that (16) holds for some $c_j > 0$ and $d_j < \infty$, $j = 1, 2$. We then define the function V_1 as V_{11} if $f(x) \le r$ and V_{12} if $f(x) > r$. Since V_1 is not in $C^2(\mathbb{R}^2)$ we smooth $V_1(x)$ in the region $r - \delta < f(x) \le r + \delta$ to give a viable test function V which is in $C^2(\mathbb{R}^2)$.

If say $f(x) = x_1$ (the case $f(x) = x_2$ is similar) then any test functions V_{1j}, $j = 1, 2$ for which $V(x)$ is of order $-K\, x_2^2 + d$ for constants $K > 0$ and $d < \infty$

when both $r-\delta < x_1 \le r+\delta$ and also x_2 is large will provide in this way a function $\hat{V} = V+1$ such that $\mathbf{X}$ is $\hat{V}$ exponentially ergodic.

Remark 5.4. We also remark that the proof of Theorem 5.2 holds under the assumption that in the region $r-\delta < f(x) < r+\delta$, only one variable x_1 or x_2 is unbounded (see (a2) and (b2)). Therefore this proof cannot be extended to the case where $p=2$ and $f(x) = d_1x_1 + d_2x_2$, $d_1, d_2 > 0$, or to the higher order case where $p > 2$.

We conclude with a sufficient condition for the state vector $\mathbf{X}$ of the CTAR(2) process to be transient.

Theorem 5.5. *If all the eigenvalues of $A^{(j)}$, $j = 1,2$ have positive real parts (i.e. $a_1^{(j)} < 0$, $a_2^{(j)} > 0$) and $f(x)$ as defined in* (5) *is either x_1 or x_2, then the state vector* $\mathbf{X}$ *of a CTAR*(2) *process is transient.*

Proof. By Theorem 3.3 and Theorem 3.3 of Stramer and Tweedie (1994) it suffices to prove that there exists a closed set C in $\mathbb{R}^2$ with C^c nonempty and a function W bounded on C^c such that $(L_W)(x) \ge 0$, $x \in C^c$. Let $W_j(x) = 1 - V_j(x)^{-\alpha}$, $j = 1,2$ where $\alpha > 0$ is a constant and $V_j(x)$ are defined in (18) and (19) except that $\epsilon < 0$ is a constant such that $P(f(x))$ and $-N^{(j)}$, $j = 1,2$ as defined in Theorem 5.2 for $f(x) = x_1$ and $f(x) = x_2$ are positive definite. The proof is now analogous to the proof of Theorem 5.2.

Remark 5.6. Theorem 5.2 and Theorem 5.5 cover most common cases. The type of behavior which is not covered is when all the eigenvalues of $A^{(1)}$ have positive real parts and all the eigenvalues of $A^{(2)}$ have negative real parts or vice versa and when for some $j = 1,2$ not all the eigenvalues of $-A^{(j)}$ or $A^{(j)}$ have negative real parts. The stability properties of $\mathbf{X}$ for these cases are still unclear.

The approximation sequence approach of Brockwell and Hyndman (see Brockwell and Hyndman (1992) and Brockwell (1994)) provides one method of simulating approximate sample paths of CTARMA procrsses. In Figure 1 and Figure 2 we use this approximation with $n = 10$ to give two sample paths of CTAR(2) processes which have the same coefficient $\sigma(y) \equiv 1$ but different $a_1(x)$, $a_2(x)$ and $c(x)$. The first process has

$$a_1(y) = \begin{cases} 2, & \text{if } y < 0, \\ 1, & \text{if } y \ge 0, \end{cases} \quad a_2(y) = \begin{cases} 1.5, & \text{if } y < 0, \\ 0.5, & \text{if } y \ge 0, \end{cases} \quad c(y) = \begin{cases} 2, & \text{if } y < 0, \\ 0, & \text{if } y \ge 0, \end{cases} \tag{20}$$

giving vector drift so that all the eigenvalues of A^1, A^2 as defined in (15) have negative real parts. The second has

$$a_1(y) = \begin{cases} -.2, & \text{if } y < 0, \\ -.1, & \text{if } y \ge 0, \end{cases} \quad a_2(y) = \begin{cases} .15, & \text{if } y < 0, \\ 0.5, & \text{if } y \ge 0, \end{cases} \quad c(y) = \begin{cases} .2, & \text{if } y < 0, \\ 0, & \text{if } y \ge 0, \end{cases}$$

giving vector drift so that all the eigenvalues of $A^{(1)}$, $A^{(2)}$ have positive real parts. The sample paths show that the first process, being Harris recurrent, returns to neighborhoods of $[0,0]'$. The second, being transient, leaves all such neighborhoods.

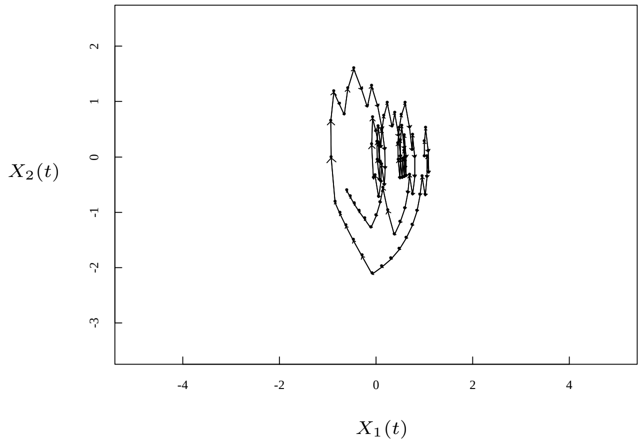

Figure 1. CTAR(2) path with $0 \le t \le 10$, $a_1^{(1)} = 2$, $a_1^{(2)} = 1$, $a_2^{(1)} = 1.5$, $a_2^{(2)} = 0.5$, $c^{(1)} = 2$, $c^{(2)} = 0$, $\sigma = 1$, $\mathbf{X}'(0) = (1,0)$.

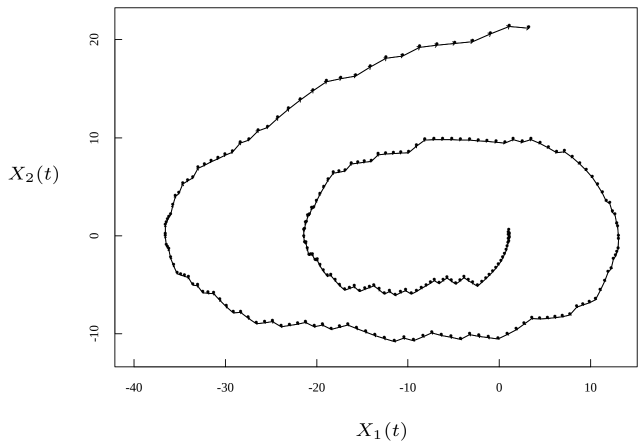

Figure 2. CTAR(2) path with $0 \le t \le 23$, $a_1^{(1)} = -.2$, $a_1^{(2)} = -.1$, $a_2^{(1)} = .15$, $a_2^{(2)} = 0.5$, $c^{(1)} = .2$, $c^{(2)} = 0$, $\sigma = 1$, $\mathbf{X}'(0) = (1,0)$.

6. Estimation of Stationary Moments

We consider finally the estimation of the stationary mean and the stationary variance. For CTAR(1) processes, the representation (14) determines the stationary moments and provides us with a method to calculate the stationary mean and variance (see Stramer, Brockwell and Tweedie (1996)). For CTARMA(p, q) processes with $p > 1$ we have not yet found an explicit expression for the stationary mean and variance as we did in the case when $p = 1$. One way around this problem is to simulate a long series and use ergodicity. To illustrate this we consider the CTAR(2) process in Example 6.1.

Example 6.1. Let $Y(t) = [1\ 0]\mathbf{X}(t)$ be the CTAR(2) process, where the components X_1, X_2 of the state vector $\mathbf{X}_t$ satisfy

$$dX_1 = X_2(t)dt$$

$$dX_2 = [-a_2(X_1(t))X_1(t) - a_1(X_1(t))X_2(t) + c(X_1(t))]dt + dW(t),$$

with $a_1(y)$, $a_2(y)$, and $c(y)$ defined as in (20).

We again use the approximating sequence of Brockwell and Hyndman (see Brockwell and Hyndman (1992) and Brockwell (1994)), with $n = 10$, to estimate

$$m(\mathbf{x}, t) = E(Y(t)|\mathbf{X}(0) = \mathbf{x}); \qquad v(\mathbf{x}, t) = E(Y^2(t)|\mathbf{X}(0) = \mathbf{x}) - (m(\mathbf{x}, t))^2.$$

The estimation of the functions $m((0, x), t)$ and $v((x, 0), t)$ for $t = 1, 3, 5, 7$ are shown in Figure 3 and Figure 4 respectively. As expected, the values for larger t are closer to each other, and for $t = 7$ there is little distinguishable difference, indicating the effect of the initial value has worn off and the asymptotic effect of (12) and (13) holds. Thus we can estimate the stationary mean and variance using the approximating sequence method, for values of $t \geq 7$ in this case.

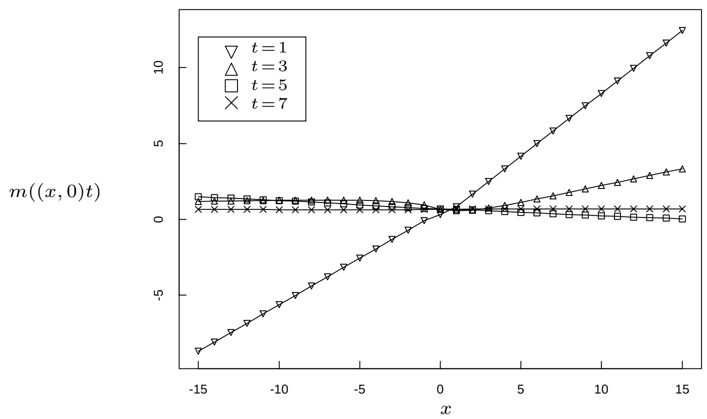

Figure 3. Conditional mean $m((x, 0), t)$, $-15 \leq x \leq 15$ for Example 6.1 with lead times $t = 1, 3, 5$ and 7.

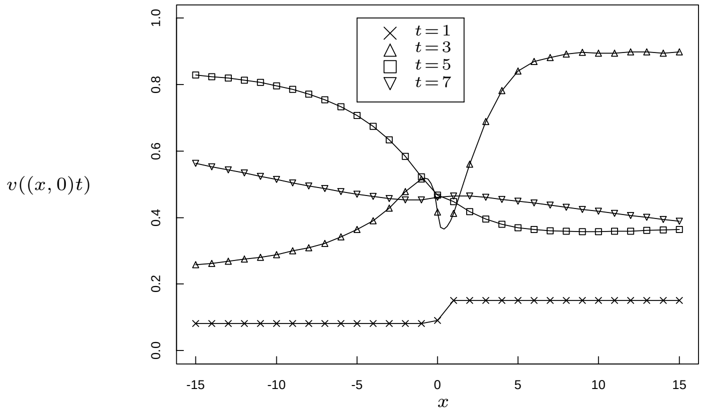

Figure 4. Conditional variance $v((x,0),t)$, $-15 \leq x \leq 15$, for Example 6.1 with lead times $t = 1, 3, 5$ and 7.

Acknowledgement

We are grateful to the referee for a careful reading of the manuscript which has substantially improved the paper, and in particular for suggesting the extended models where the threshold is allowed to depend on the velocity as well as the position of the process. This work was supported in part by NSF Grants DMS 9504798, DMS 9100392 and 9105745.

References

Azéma, J., Duflo, M. and Revuz, D. (1967). Measure invariante sur les classes récurrentes des processus de Markov. *Z. Wahrscheinlichkeitstheorie und Verw. Geb.* **8**, 157-181.

Billingsley, P. (1986). *Probability*, 2nd edition. Springer-Verlag, New York.

Brockwell, P. J. and Hyndman, R. J. (1992). On continuous time threshold autoregression. *Internat. J. Forecasting* **8**, 157-173.

Brockwell, P. J., Hyndman, R. J. and Grunwald, G. K. (1991). Continuous time threshold autoregressive models. *Statist. Sinica* **1**, 401-410.

Brockwell, P. J. (1994). On continuous time threshold ARMA processes. *J. Statist. Plann. Inference* **39**, 291-303.

Brockwell, P. J. and Stramer, O. (1995). On the approximation of continuous time threshold ARMA processes. *Ann. Inst. Statist. Math.* **47**, 1-20.

Chen, C. T. (1984). *Linear System Theory and Design.* Saunders College Publishing, Florida.

Getoor, R. K. (1979). Transience and recurrence of Markov jrocesses. *Sém. de Probabilités* **XIV**, 397-409. Springer-Verlag, Berlin, Heidelberg, New York.

Ichihara, K. and Kunita, H. (1974). A classification of the second order degenerate elliptic operators and its probabilistic characterization. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **30**, 235-254.

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

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York.

Meyn, S. P. and Tweedie, R. L. (1993a). Stability of Markovian processes II: Continuous Time Processes and Sampled Chains. *Adv. Appl. Probab.* **25**, 487-517.

Meyn, S. P. and Tweedie, R. L. (1993b). Stability of Markovian processes III: Foster-Lyapunov Criteria for Continuous Time Processes. *Adv. Appl. Probab.* **25**, 518-548.

Oksendal, B. (1992). *Stochastic Differential Equations.* Third edition, Springer-Verlag, New York.

Stramer, O., Brockwell, P. J. and Tweedie, R. L. (1996). Continuous time threshold AR(1) processes. *Adv. Appl. Probab.* (to appear).

Stramer, O. and Tweedie, R. L. (1994). Stability and instability of continuous time Markov processes. In *Probability, Statistics and Optimization : A Tribute to Peter Whittle* (Edited by F. P. Kelly), 173-184. John Wiley and Sons.

Stramer, O. and Tweedie, R. L. (1996). Existence and stability of weak solutions to stochastic differential equations with non-smooth coefficients. *Statist. Sinica* (under revision).

Stroock, D. W. and Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer-Verlag, Berlin.

Tong, H. (1990). *Non-linear Time Series : A Dynamical System Approach.* Clarendon Press, Oxford.

Tong, H. and Yeung, I. (1991). Threshold autoregressive modelling in continuous time. *Statist. Sinica* **1**, 411-430.

Tweedie, R. L. (1994). Topological conditions enabling use of Harris methods in discrete and continuous time. *Acta Applic. Math.* **34**, 175-188.

Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242, U.S.A.

Department of Statistics, Colorado State University, Fort Collins, CO 80523, U.S.A.

Department of Mathematics, Royal Melbourne Institute of Technology, Melbourne 3001, Australia.

(Received February 1994; accepted July 1995)