V-UNIFORM ERGODICITY OF THRESHOLD AUTOREGRESSIVE NONLINEAR TIME SERIES

A Dissertation

by

THOMAS R. BOUCHER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2003

Major Subject: Statistics

V-UNIFORM ERGODICITY OF THRESHOLD AUTOREGRESSIVE NONLINEAR TIME SERIES

A Dissertation

by

THOMAS R. BOUCHER

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Daren B.H. Cline (Chair of Committee)	
Michael T. Longnecker (Member)	Joel Zinn (Member)
Dennis W. Jansen (Member)	James A. Calvin (Head of Department)

December 2003

Major Subject: Statistics

ABSTRACT

V-Uniform Ergodicity of Threshold Autoregressive Nonlinear Time Series. (December 2003)

Thomas R. Boucher, B.S., University of Massachusetts-Lowell;
M.S., University of Massachusetts-Lowell
Chair of Advisory Committee: Dr. Daren B.H. Cline

We investigate conditions for the ergodicity of threshold autoregressive time series by embedding the time series in a general state Markov chain and apply a Foster-Lyapunov drift condition to demonstrate ergodicity of the Markov chain. We are particularly interested in demonstrating V-uniform ergodicity where the test function $V(\cdot)$ is a function of a norm on the state-space.

In this dissertation we provide conditions under which the general state space chain may be approximated by a simpler system, whether deterministic or stochastic, and provide conditions on the simpler system which imply V-uniform ergodicity of the general state space Markov chain and thus the threshold autoregressive time series embedded in it. We also examine conditions under which the general state space chain may be classified as transient. Finally, in some cases we provide conditions under which central limit theorems will exist for the V-uniformly ergodic general state space chain.

To my wife, my family and all who helped along the way.

ACKNOWLEDGEMENTS

The road to a Ph.D is a long and winding one. To acknowledge all whom I should acknowledge would exhaust this space long before the task could be completed. Those I list here stand out for one reason or another but exclusion from this list should not indicate any lack of appreciation on my part. You know who you are and I know who you are.

I must express my gratitude to Dr. Daren B.H. Cline for serving as my advisor while I completed my Ph.D. His patience and graciousness were surpassed only by his dilligence. I could not have done this without him.

I would like to thank my committee members Dr. Michael T. Longnecker, Dr. Joel Zinn and Dr. Dennis W. Jansen for the benefit of their time and knowledge.

Last but not least, I extend my fullest appreciation to my wife and my family for their unconditional and much needed support. In the most trying times they were my refuge from the storm.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The increasing importance of nonlinear time series models is due to the fact these models are capable of describing many of the phenomena found in time series data that cannot be adequately described by classical linear ARMA models. There are numerous examples of these phenomena. Limit cycles describe instances where the series eventually cycles through a set of values. Jump phenomena occur when the time series suddenly 'jumps' from one fairly stable regime to another. Time-irreversibility refers to cases where the rate of increase of the time series differs from the rate of decrease. Time-varying volatility refers to instances where the volatility of the series, as evidenced by changes in the values of the series, changes over time. Amplitude-dependent volatility describes cases where the changes in volatility are related to the current amplitude of the time series. TAR(p) models were introduced by Tong and Lim (1980) to handle limit cycles in particular, but they were also shown to model jump phenomena and time-irreversiblity.

1.1 The TAR(p) Model with Delay d

Threshold autoregressive models are piecewise linear over the domain of the process; which linear piece applies depends upon the prior values of the time series. Let $\{y_t\}_{t\geq 0}$ denote the time series. As introduced by Tong and Lim (1980), the TAR(p) model of order p, delay d and thresholds $r_0,\ldots,r_l$ with $-\infty = r_0 < r_1 < \ldots < r_l = +\infty$ can be written as

---

The format and style follow that of *Journal of the American Statistical Association*.

$$y_t = \phi_1^{(i)} y_{t-1} + \cdots + \phi_p^{(i)} y_{t-p} + \xi_t, \quad r_{i-1} < y_{t-d} \leq r_i, \tag{1.1}$$

where $\phi_1^{(i)}, \ldots, \phi_p^{(i)}$, $i = 1, \ldots, l$, are constants and $\{\xi_t\}_{t \geq 0}$ are mean zero iid random variables.

The threshold autoregressive model can also be written in the more general form of the larger class of autoregressive nonlinear models:

$$y_t = f(y_{t-1}, \ldots, y_{t-p}; \xi_t), \quad p \geq d$$

with $f(\cdot)$ being an arbitrary nonlinear function. These models encompass both parametric and nonparametric models and provide us with an extraordinarily flexible family of models. Threshold autoregressive models are particularly important in light of the fact many nonlinear functions $f(\cdot)$ can be well approximated by linear functions over finite intervals.

Each threshold autoregressive model can be embedded in a Markov chain on $\mathbb{R}^p$. For specifics, if we write $X_t = (y_t, y_{t-1}, \ldots, y_{t-p+1})'$ the TAR(p) model introduced in (1.1) can be expressed as the following:

$$X_t = A_i X_{t-1} + \nu_t, \quad X_{t-1} \in R_i \tag{1.2}$$

where the space $\mathbb{R}^p$ is divided into l regions $R_i, i = 1, \ldots, l$, the R_i depending upon the thresholds r_i and the delay parameter d. The A_i are called the companion matrices and are given by:

$$A_i = \begin{pmatrix} \phi_1^{(i)} & \phi_2^{(i)} & \ldots & \phi_{p-1}^{(i)} & \phi_p^{(i)} \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & \vdots & \ldots & 1 & 0 \end{pmatrix}$$

and $\nu_t = \xi_t(1,0,\ldots,0)'$. Since the distribution of X_t given $X_0,\ldots,X_{t-1}$ depends upon X_{t-1} only, X_t is a Markov chain. The transition measure of X_t is singular w.r.t. Lebesgue measure if $p > 1$ or $d > 1$.

Stability of the nonlinear time series model is then defined as the ergodicity of the associated Markov chain. This question of stability has important ramifications for statistical inference. The existence of a stationary distribution with finite moments is crucial for proving consistency and asymptotic distributions of the parameter estimates. Standard proofs of the consistency, asymptotic normality and optimality of parameter estimates in the linear ARMA case (such as in Brockwell and Davis (1987)) require causality of the model. Generalizations of these ergodic parameter spaces for linear ARMA models to nonlinear models are often inadequate even for the simplest forms of nonlinearity in the model. In some cases these generalizations are too broad, in others they are far too restrictive; some TAR(p) models admit an unbounded ergodic parameter space (Petruccelli and Woolford (1984), Chen and Tsay (1991), Kunitomo (2001)). In order for statistical inference involving these models to be valid, it is necessary to first know the model under consideration is stationary, i.e., ergodic, making the investigation into the ergodic parameter spaces of these models of paramount importance.

1.2 Literature Review

Stability of the TAR(p) model is established for some very simple cases. Chan et al. (1985) derived necessary and sufficient ergodicity conditions for a class of multiple threshold models with delay 1. Petruccelli and Woolford (1984) did the same for a special case of a single threshold model. Guo and Petruccelli (1991) refined the results of Chan et al (1985), adding classification of the model as null recurrent or transient. Lim (1992) and separately Chen and Tsay (1991) have established necessary and sufficient conditions for geometric ergodicity of a simple case of TAR(1) models with arbitrary delay. Kunitomo (2001) has established results for some special cases of TAR(2) models, but even these very simple models are not completely characterized. This existing work reveals the valid parameter space is quite different from the product parameter space one may expect and which is often given as a sufficient condition for ergodicity, a conclusion further confirmed by the present research.

Different approaches to the stability of nonlinear time series may be taken. Authors such as Lim (1992) and Tong (1990) have taken a dynamical systems approach toward stability of nonlinear time series, linking ergodicity of the process $y_t = f(y_{t-1},\ldots,y_{t-p}) + \varepsilon_t$ to the dynamic stability of the deterministic skeleton $y_t = f(y_{t-1},\ldots,y_{t-p})$. This approach has yielded useful results where the deterministic skeleton satisfies certain regularity conditions (see Chan (1990)), such as Lipschitz continuity and exponential stability of the deterministic skeleton. Many useful models, however, do not satisfy the regularity conditions placed on the skeleton and researchers such as Cline and Pu ((1999a), (1999b), (2001) and (2002)) have noted the conditions for stability of the chain are not always the same as those for the stability of the skeleton.

An alternative method, followed by authors such as Tjøstheim (1990), Meyn and Tweedie (1993) and Cline and Pu (2001), is the previously detailed approach of embedding the time series in a Markov chain and examining stability of the time series through the ergodicity of the Markov chain. Tjøstheim (1990) also introduces the *k*-step method whereby ergodicity of the one-step chain $\{X_t\}$ is equated to the ergodicity of the *k*-step chain $\{X_{tk}\}$, where k is a finite positive integer. This is one of the approaches we use and will be explained further in the next section.

Regarding statistical inference, limit theorems for the parameter estimates depend on the existence of moments of the stationary distribution. Early work, such as that of Petruccelli and Woolford (1984) and Chan et al (1985), provided conditions on the parameter values and moment conditions on the error distribution that for particular models resulted in the existence of moments for the stationary distribution. From this they established limit theorems for strong consistency and asymptotic normality of the parameter estimates. More recent results in the case of V-uniform ergodic Markov chains (Meyn and Tweedie (1993), Cline and Pu (2001)) link the moments of the stationary distribution to the order of the test function used to satisfy the drift condition for V-uniform ergodicity. Limit theorems for the parameter estimates can then be established if the test function implies the appropriate moments of the stationary distribution exist.

1.3 Definitions/Theory

Embedding a nonlinear autoregressive process of order p, $y_t = f(y_{t-1},\ldots,y_{t-p},\varepsilon_t)$ in the Markov chain $X_t = (y_t, y_{t-1},\ldots,y_{t-p})'$ allows us to recast the problem of stability of the nonlinear time series in terms of the stability of the sequence of distributions $\{P^n(x,\cdot)\}_{n\geq 0}$, the n-step transition probabilities of the chain generated by the transition kernel $P(x,\cdot)$ from the initial distribution μ of the Markov chain.

This stability of the sequence of distributions $\{P^n(x,\cdot)\}_{n\geq 0}$ can be characterized as follows: let E denote the space, $\sigma(E)$ denote a countably generated σ-field containing E and consider an ergodic Markov chain on state-space $(E,\sigma(E))$ with transition probability $P(x,\cdot)$ and invariant probability distribution π, that is π satisfies

$$\pi(A) = \int_E \pi(dx)P(x,A), \quad \forall A \in \sigma(E).$$

If X_0 is distributed according to π, then $\{X_t\}_{t\geq 0}$ is a stationary Markov chain, following from the invariance of π. Stability of the Markov chain is equivalent to the existence of a stationary (invariant) distribution π for the chain such that the sequence of distributions $\{P^n(x,\cdot)\}$ converges to this stationary distribution π.

There are certain conditions a Markov chain must satisfy in order for the stationary distribution π to exist and be unique. The reader can consult Nummelin (1984) or Meyn and Tweedie (1993) for the following definitions.

Let $P_x(\cdot)$ denote the transition distribution of X_t given that $X_0 = x$. A Markov chain $\{X_t\}$ is said to be $\psi - irreducible$ if there exists a σ-finite measure ψ on the space such that $\forall x \in E$ and whenever $\psi(A) > 0$, we have $P_x(\tau_A < \infty) > 0$, where τ_A is the time of the first visit to the set A. A Markov chain that is ψ-irreducible for some probability measure ψ is called *irreducible* and we denote by $\sigma^+(A)$ the collection of all sets A with $\psi(A) > 0$.

A $d - cycle$ for a general state chain is a cycle of regions$(E_o, \ldots, E_{d-1})$ such that $\forall x \in E_i, P(x, E_{i+1}) = 1$ for $i = 0 \ldots d-1 (\text{mod } d)$ and the set $N = (\cup_{i=1}^d E_i)^C$ is ψ-null. A ψ-irreducible Markov chain is called *aperiodic* when $d = 1$.

A ψ-irreducible chain is said to be *Harris recurrent* if $\forall A \in \sigma^+(E)$, $\forall x \in A$ we have $P_x(X_t \in A \text{ i.o.}) = 1$. ψ-irreducible, aperiodic recurrent chains admit an invariant measure, but this measure could be infinite. If there exists a *petite* (Meyn and Tweedie, pg. 121) set C such that

$$\sup_{x \in C} E_x(\tau_C) < \infty$$

then the chain is called *positive* Harris recurrent. When combined with aperiodicity and irreducibility, our assumptions on the error terms ensure that compact sets are petite and thus we can apply drift conditions to demonstrate positive Harris recurrence. The significance of positive Harris recurrence is that it implies the invariant measure is finite and can then be suitably normalized to become a probability measure.

An *ergodic* chain is one that is positive Harris recurrent, ψ-irreducible and aperiodic. Consequently, the chain has a unique invariant probability distribution and the time series is thus stationary when the initial distribution is the invariant distribution. Verifying stationarity of the time series then becomes a question of verifying each of the conditions listed above for ergodicity of the Markov chain in which the time series is embedded. In the case of additive errors, ψ-irreducibility and aperiodicity are an easy consequence of the error distribution having a continuous density that is everywhere positive, the irreducibility measure ψ thus being Lebesgue measure (Meyn and Tweedie (1993), Cline and Pu (1998)). The question of establishing ergodicity is then reduced to establishing positive Harris recurrence.

Ergodicity implies the n-step transition probabilities $P^n(x,\cdot)$ converge to the invariant probability measure

$$\lim_{n\to\infty} \| P^n(x,\cdot) - \pi \| = 0,$$

where $\|\cdot\|$ is the total variation norm, i.e.,

$$\| P^n(x,\cdot) - \pi \| = \sup_{B\in\sigma(E)} |P^n(x,B) - \pi(B)|.$$

Nummelin (1984) informs us that this convergence is $o(1/n)$. It is advantageous to know if the convergence of $P^n(x,\cdot)$ to π occurs more quickly since we can then either assume the process has already reached stability, meaning the governing distribution is the stationary one, or at the very least from the stationary distribution we know the long-term or ergodic behavior of the chain. There are different types of ergodicity named according to the rate and manner in which the convergence of $P^n(x,\cdot)$ to π occurs.

If the convergence occurs at a geometric rate, we have the concept of geometric ergodicity:

$$\| P^n(x,\cdot) - \pi \| \leq Rr^{-n}, \quad r > 1, \quad R < \infty, \quad n \geq 1.$$

However, this convergence is not uniform; the constants depend on the initial value x. To get uniform convergence, we need to consider a stronger form of ergodicity named V-uniform ergodicity.

Following Meyn and Tweedie (1993) define the V-norm for a positive function $V \geq 1$ and any measure P as

$$\| P \|_V = \sup_{f:|f|\leq V} |P(f)|.$$

For two Markov transition functions P_1 and P_2, define the V-norm distance between P_1 and P_2 as

$$| \, \|P_1 - P_2| \|_V := \sup_{\{x\}} \frac{\| P_1(x,\cdot) - P_2(x,\cdot) \|_V}{V(x)}.$$

Consider π as the transition function $\pi(A) = \pi(x,A)$; then V-uniform convergence is defined as geometric convergence of $P^n(x,\cdot)$ to π when the distance between $P^n(x,\cdot)$ and π is measured by the V-norm distance: there exist $r > 1$, $R < \infty$ such that for all positive intergers n:

$$|\|P^n - \pi|\|_V \leq Rr^{-n}.$$

Note that since V-uniform convergence is defined in terms of the V-norm which involves a supremum over all x, the convergence is in fact uniform in x and thus the name.

Establishing the various types of ergodicity for a Markov chain is often done with the use of Foster-Lyapounov drift conditions, one of which is provided by Meyn and Tweedie (1993): for our purposes an irreducible, aperiodic Markov chain $\{X_t\}$ is geometrically ergodic if for some extended real-valued, locally bounded test function $V : E \to [1, +\infty)$, there exist $K < \infty$, $\rho < 1$ and a compact set C such that

$$E(V(X_t)|X_{t-1} = x) \leq \rho V(x) + K 1_C(x).$$

A useful equivalent condition requires (Cline and Pu (1999b), Cline and Pu (2001)):

$$\limsup_{\|x\| \to \infty} \frac{E(V(X_{t+1})|X_t = x)}{V(x)} < 1 \tag{1.3}$$

and there exists $M < \infty$ such that

$$\sup_{\|x\| \leq M} E(V(X_{t+1})|X_t = x) < \infty. \tag{1.4}$$

In fact, both these conditions establish the stronger form of ergodicity, V-uniform ergodicity, where $V(\cdot)$ is the test function used to satisfy the drift condition.

Cline and Pu (2001) list several approaches to using test functions V to satisfy drift conditions. We make use of the *directional method*, involving the use of a test function of the form $V(x) = 1 + \lambda(x) \parallel x \parallel^r$ where $r > 0$ and λ is bounded and bounded away from 0. The function $\lambda(\cdot)$ is typically chosen to depend on the direction of x. This approach has worked well with threshold models of order 1 and we apply it with success in Chapter III to threshold models of higher order. The advantage of this method can be seen when the TAR(p) model has a cycle, under appropriate conditions on the skeleton. By choosing the function $\lambda(x)$ to be constant in each region and by choosing the appropriate constants, one can ensure the expected ratio of test functions (1.3) to be less than 1 as the chain travels from region to region (Cline and Pu (1999b)).

The choice of test function $V(\cdot)$ has important implications for statistical inference. Meyn and Tweedie (1993) relate the existence of moments of the stationary distribution to the order of the test function used in demonstrating V-uniform ergodicity: for a V-uniformly ergodic chain $\{X_t\}$ and any function $g(x)$ such that $g^2 \leq V$, Meyn and Tweedie prove consistency and central limit theorems for the partial sums

$$S_n(g) = \sum_{i=1}^{n} g(X_i).$$

Spieksma and Tweedie (1994) provide a set of conditions for a countable state-space Markov chain that allow the test function $V(x)$ to be 'boosted' to an exponential test function $V_1(x) = e^{sV(x)}$, where $s > 0$. Cline and Pu (2001) generalize this result by deriving conditions for a general state-space chain to have an exponential test function of the form either $V_1(x) = e^{sV(x)}$ or $V_1(x) = e^{V(x)^s}$. For test function $V(x)$ such that $\parallel x \parallel \leq V(x) \leq M + K \parallel x \parallel$ for finite constants K and M, the existence of an exponential test function thus implies the existence of all moments of the stationary distribution (Cline and Pu (2001)). With this in hand, consistency and limit theorems for the parameter estimates can then be established.

Our project is to give conditions for ergodicity of threshold autoregressive models by embedding the time series in a general state space Markov chain and applying the Markov theory detailed in this introduction. We identify two special cases we define as cyclic and finite state chain approximated and use Foster-Lyapunov drift criteria to demonstrate V-uniform ergodicity. We also provide sufficient conditions for transience of cyclic models in some special cases and conditions under which central limit theorems will hold in the cyclic case.

CHAPTER II

STABILITY AND INFERENCE FOR CYCLIC THRESHOLD AUTOREGRESSIVE MODELS

2.1 Introduction

We approach the question of stability of the nonlinear time series by deriving conditions under which the Markov chain in which the series is embedded could be classified as either V-uniform ergodic or transient.

Suppose the threshold autoregressive model of order p described in (1.1) is embedded in a Markov chain $\{X_t\}$ as in (1.2). The space $\mathbb{R}^p$ is divided into regions $R_1,\ldots,R_l$ each region R_i with companion matrix A_i, $i=1,\ldots,l$. We define the deterministic skeleton x_t of the Markov chain X_t to be the deterministic process

$$x_t = A_i x_{t-1}, \quad x_{t-1} \in R_i, \quad i=1,\ldots,l \tag{2.1}$$

i.e., the deterministic skeleton is the process with the additive errors removed.

Define a k-cycle for the deterministic skeleton for a collection $\{i_1,\ldots,i_k\}$ of length k from $\{1,\ldots,l\}$ to be a collection of k regions $R_{i_1},\ldots,R_{i_k}$ with corresponding companion matrices $A_{i_1},\ldots,A_{i_k}$ such that $x \in R_{i_j}$ implies $A_{i_j}x \in R_{i_{j+1(\mathrm{modk})}}$. Similarly, the multi-cyclic case has a finite number of cycles $C_1,\ldots,C_m$ each of length k_i, $i=1,\ldots,m$. Since there are a finite number of cycles of finite length we can with some modifications reduce this to a k-step process with $k=\Pi_{i=1}^m k_i$.

Consider the case of a single k-cycle. Setting $x_t = x \in R_i$ and looking k transitions ahead, we have $x_t = x \in R_i$ implies $x_{t+k} = \Pi_{i=1}^k A_{i_j} x \in R_i$. This observation tells us the skeleton will shrink if it shrinks each trip through the cycle; thus, rather than look at the one-step transitions of the process $\{X_t\}$ we can consider the k-step transitions of the chain.

This is the heuristic behind the k-step strategy for demonstrating geometric ergodicity put forth by Tjøstheim (1990).

There is a clear benefit from considering the cyclic behavior of the chain. Consider that for any two matrices A and B, for an arbitrary norm $\| \cdot \|$ and where $\rho(A)$ denotes the largest eigenvalue of A in modulus

$$\rho(AB) = \lim_{n\to\infty} \| (AB)^n \|^{1/n} \leq \lim_{n\to\infty} \| A^n \|^{1/n} \times \lim_{n\to\infty} \| B^n \|^{1/n} = \rho(A) \times \rho(B). \qquad (2.2)$$

Generalizing conditions for ergodicity from the linear AR(p) case to a threshold autoregressive model with companion matrices A and B would lead to the condition $\rho(A) < 1$ and $\rho(B) < 1$, implying $\rho(A) \times \rho(B) < 1$. The condition $\rho(A) < 1, \rho(B) < 1$ implies V-uniform ergodicity for all cyclic models with companion matrices A and B, but as (2.2) shows this condition is stronger than what is necessary and leads us to miss valid models. Models whose deterministic skeleton has a k-cycle $R_{i_1}, \ldots, R_{i_k}$ with companion matrices $A_{i_1}, \ldots, A_{i_k}$ require under certain conditions only that $\rho(\| \Pi_{i=1}^{k} A_{i_j} \|) < 1$, rather than the stronger condition $\rho(\| A_{i_j} \| < 1)$ for $j = 1, \ldots, k$. We can argue analagously in the case of multiple cycles of finite length. The gain here can be tremendous; as mentioned in Chapter I certain threshold autoregressive models have been shown to have unbounded parameter spaces (Petruccelli and Woolford (1984), Kunitomo (2001)). Considering only the relevant cycles allows us to recover the full parameter space.

Define for a Markov chain $\{X_t\}$ and a constant $k < \infty$ the k-step chain to be the Markov chain $\{X_{tk}\}$. Using a drift criterion, Tjøstheim demonstrated geometric ergodicity of the k-step chain $\{X_{tk}\}$ and drew upon Nummelin (Thm. 6.14, (1984)) who equated geometric ergodicity of $\{X_{tk}\}$ with that of $\{X_t\}$. Meyn and Tweedie subsequently strengthened Tjøstheim's result by showing Tjøstheim's drift criterion implied V-uniform ergodicity of $\{X_{tk}\}$ and thus V-uniform ergodicity of $\{X_t\}$. We summarize all of this in an original lemma which restates the equivalency of V-uniform ergodicity of $\{X_t\}$ and $\{X_{tk}\}$ in terms only of the V-norm.

We use the following drift criteria to establish V-uniform ergodicity for the k-step chain:

$$\limsup_{\|x\|\to\infty} \frac{E(V(X_{n+k})|X_0 = x)}{V(x)} < 1 \tag{2.3}$$

and for all $M < \infty$

$$\sup_{\|x\|\leq M} E(V(X_{n+k})|X_0 = x) < \infty, \tag{2.4}$$

where the test function $V(\cdot) \geq 1$ is such that $V(\cdot)$ is measurable, locally bounded and $V(x) \to \infty$ as $\| x \| \to \infty$.

Analogously, transience of $\{X_t\}$ is demonstrated through its equivalence to the transience of the k-step chain $\{X_{tk}\}$. Tweedie (1976, Theorem 11.3) provides the following criteria for the transience of $\{X_t\}$: for sets B and B^c of positive measure, if there exists a non-negative function $g(x)$ with

$$E(g(X_1)|X_0 = x) \leq g(x), \quad x \in B^c \tag{2.5a}$$

$$g(x) < \inf_{y\in B} g(y), \quad x \in B^c \tag{2.5b}$$

or a bounded non-negative function $g(x)$ with

$$E(g(X_1)|X_0 = x) \geq g(x), \quad x \in B^c \tag{2.6a}$$

$$g(x) > \sup_{y\in B} g(y), \quad x \in B^c \tag{2.6b}$$

and if $\{X_t\}$ is ψ-irreducible, then $\{X_t\}$ is transient. We apply these criteria to derive conditions under which $\{X_{tk}\}$ is transient and then apply Tjøstheim (1990, Lemma 3.1) to conclude transience of $\{X_t\}$.

Inference for the threshold autoregressive model depends upon the existence of moments of the stationary distribution and upon the existence of central limit theorems for partial sums $\sum_{i=1}^{n} g(X_i)$. Results from Meyn and Tweedie (1993) link the existence of moments and central limit theorems to the order of the test function V used in establishing

V-uniform ergodicity. Cline and Pu (2001) provide conditions under which the test function V can be 'boosted' to an exponential test function $V^{'}(x) = e^{(V(x))^s}$ such that the process is $V^{'}$-uniformly ergodic. This is discussed in more detail in Section 4.

Our results are applicable to threshold autoregressive models where the stochastic process behaves asymptotically like the deterministic skeleton and whose determinstic skeleton for $\| x \|$ large exhibits finite cyclic behavior, i.e., has a finite number of cycles of finite length. We assume throughout that the cycles do not fall on any of the thresholds so that for large x there is negligible probability of the process leaving one cycle for another.

We first establish conditions under which these processes are V-uniformly ergodic, then turn our attention to transience and finally to the question of the existence of moments of the stationary distribution. Theorems establishing ergodicity are in Section 3, those for transience are in Section 4 and conditions for the test function $V(\cdot)$ to be 'boosted' to an exponential are in Section 5. First we provide some results to be used throughout the chapter.

2.2 Preliminary Results

This first result provides us with the norm we will use. It is due to Ciarlet (1982) and can be found in An and Huang (1996). The statement of the lemma is Ciarlet's; the sketch of the proof is ours.

(Ciarlet) Lemma 1. If a matrix G has $\rho(G) < 1$, then there exists a matrix norm $\| \cdot \|_m$ induced by a vector norm $\| \cdot \|_v$ and a constant $\lambda < 1$ such that

$$\| Gx \|_x \leq \| G \|_m \| x \|_x \leq \lambda \| x \|_v, \ \forall x. \tag{2.7}$$

Proof. Let $\rho(G)$ denote the eigenvalue of largest modulus for an arbitrary matrix G. It is a well-known fact (Martelli (1992), Lemma 4.2.1, for example) that $\rho(G) < 1$ implies the existence of a vector norm $\| \cdot \|_v$ such that the matrix operator norm $\| \cdot \|_m$ induced by $\| \cdot \|_v$

has $\| G \|_m < 1$. Also, $\| G \|_m < 1$ implies the existence of a constant $\lambda < 1$ with $\lambda \geq \| G \|_m$. Combining these facts with a norm inequality we have

$$\| Gx \|_x \leq \| G \|_m \| x \|_x \leq \lambda \| x \|_v, \; \forall x$$

which is the result. □

We will use this vector norm $\| \cdot \|_v$ and the matrix norm $\| \cdot \|_m$ induced by $\| \cdot \|_v$ throughout the rest of this chapter.

The following lemma establishes the V-uniform ergodicity of the one-step chain $\{X_t\}$ from that of the k-step chain $\{X_{tk}\}$. As mentioned previously, Meyn and Tweedie (1993) define for a function $V \geq 1$ the V-norm distance between two transition kernels P_1 and P_2 as

$$|||P_1 - P_2|||_{\mathbf{V}} := \sup_x \sup_{|g| \leq V} \frac{|P_1 g - P_2 g|}{V(x)}. \tag{2.8}$$

where for a kernel P we define

$$Pg := \int g(y) P(x, dy). \tag{2.9}$$

Let $P = P(X_1 \in A | X_0 = x)$ denote the transition kernel of $\{X_t\}$. Then from (2.8) and (2.9)

$$\begin{aligned} |||P|||_{\mathbf{V}} &= \sup_x \sup_{|g| \leq V} \frac{|Pg|}{V(x)} \\ &= \sup_x \sup_{|g| \leq V} \frac{|E(g(X_1)|X_0 = x)|}{V(x)} \\ &\leq \sup_x \sup_{|g| \leq V} \frac{E(|g(X_1)||X_0 = x)}{V(x)} \\ &\leq \sup_x \frac{E(V(X_1)|X_0 = x)}{V(x)}. \end{aligned} \tag{2.10}$$

So if we can show

$$\sup_x \frac{E(V(X_1)|X_0 = x)}{V(x)} < \infty, \tag{2.11}$$

where V is the test function used to show V-uniform ergodicity for the k-step chain then we have $|||P|||_{\mathbf{V}} < \infty$, where P is the transition kernel of $\{X_t\}$.

Lemma 2. Suppose for a positive integer $k < \infty$ the k-step $\{X_{tk}\}$ chain is V-uniformly ergodic and that for the one-step chain $\{X_t\}$ with transition kernel P we have $|||P|||_{\mathbf{V}} < \infty$. Then the one-step chain $\{X_t\}$ is V-uniformly ergodic as well.

Proof. Since $|||P|||_{\mathbf{V}} < \infty$, there exists $M < \infty$ such that $|||P|||_{\mathbf{V}} \leq M$. Suppose w.l.o.g. that $M \geq 1$. Since $\{X_{tk}\}$ is V-uniformly ergodic it is geometrically ergodic and by Tjøstheim (1990, Lemma 3.1) so is X_t, meaning $\{X_{tk}\}$ and $\{X_t\}$ each have invariant distributions π_k and π, respectively. By Meyn and Tweedie (1993, Theorem 10.4.5), π_k is also invariant for $\{X_t\}$; π is clearly invariant for $\{X_{tk}\}$. Since the invariant distributions are unique up to constant multiples we have that $\pi(A) = \pi_k(A)$ for all sets A with $\pi(A) > 0$ and $\pi_k(A) > 0$. Denote this common invariant distribution by π.

Note that for the one-step chain $\{X_t\}$ with transition kernel P the k-step chain $\{X_{tk}\}$ has transition kernel P^k. By Meyn and Tweedie (1993, Theorem 16.0.1) the k-step chain is V-uniformly ergodic if and only if we have for some $R < \infty$, $r > 1$, and for all n that $|||(P^k)^n - \pi|||_{\mathbf{V}} \leq Rr^{-n}$. Now write

$$|||P^{kn} - \pi|||_{\mathbf{V}} = |||(P^k)^n - \pi|||_{\mathbf{V}} \leq Rr^{-n} = Rr^{-\frac{nk}{k}} = R(r^{1/k})^{-nk} = Rr_*^{-nk}, \tag{2.12}$$

where $r_* = r^{1/k} > 1$ since $r > 1$.

The invariance of π for P implies $P^j\pi = \pi$ for all integers j. Now consider that since $|||\cdot|||_{\mathbf{V}}$ is an operator norm (Meyn and Tweedie (1993), Lemma 16.1.1) we have by norm inequalities for any kernels P, P_1, P_2 and any integers j, k

$$|||P^j|||_{\mathbf{V}} \leq (|||P|||_{\mathbf{V}})^j \quad \text{and} \quad |||\mathrm{P}_1^j \mathrm{P}_2^k|||_{\mathbf{V}} \leq (|||\mathrm{P}_1|||_{\mathbf{V}})^{\mathrm{j}}(|||\mathrm{P}_2|||_{\mathbf{V}})^{\mathrm{k}}. \tag{2.13}$$

Consider π to be the kernel $\pi(x,A) := \pi(A)$ for all sets A. Then for any integers n and

$1 \leq j < k$, k fixed, using (2.12), (2.13) and the invariance of π

$$
\begin{aligned}
|||P^{kn+j} - \pi|||_{\mathbf{V}} &= |||P^j P^{kn} - P^j \pi|||_{\mathbf{V}} \\
&\leq |||P^j|||_{\mathbf{V}} |||P^{kn} - \pi|||_{\mathbf{V}} \\
&\leq (|||P|||_{\mathbf{V}})^j |||P^{kn} - \pi|||_{\mathbf{V}} \leq M^j R r_*^{-kn} \leq R' r_*^{-(kn+j)}
\end{aligned}
\tag{2.14}
$$

where $R' = RM^k r_*^k$. Then for all $n' = nk + j$ for some n, $1 \leq j < k$, we have from (2.14)

$$
|||P^{n'} - \pi|||_{\mathbf{V}} \leq R' r_*^{-n'}, \quad r^* > 1, R' < \infty
$$

which by Meyn and Tweedie (1993, Theorem 16.0.1) is true if and only if $\{X_t\}$ is V-uniformly ergodic. □

Denote the test functions used to satisfy the drift condition for V-uniform ergodicity of $\{X_t\}$ and of $\{X_{tk}\}$ by $V_1(\cdot)$ and $V_k(\cdot)$, respectively. Meyn and Tweedie (1993) pointed out the equivalence of the drift condition and V-uniform ergodicity; since $\{X_t\}$ and $\{X_{tk}\}$ are both V-uniformly ergodic for the same function V as a result of Lemma 2, this implies the test functions $V_1(\cdot)$ and $V_k(\cdot)$ are of the same order.

Two definitions are needed before proceeding on to the next lemma. Define a d-path to be a sequence of $d+1$ regions $R_0, \ldots, R_d$ with companion matrices $A_0, \ldots, A_d$ that the skeleton of the process moves through, i.e.,

$$
x \in R_i \Rightarrow A_i x \in R_{i+1}, \quad i = 0, \ldots, d.
$$

For the vector norm $\|\cdot\|_v$ and a positive function $f(\cdot)$, let the ball of radius $f(\| x \|_v)$ around $A_i x$ be denoted by

$$
B_{f(\|x\|_v)}(A_i x) = \{y : \| y - A_i x \|_v < f(\| x \|_v)\}.
$$

The next lemma assures us that by picking the initial x large enough in the d-path and with appropriate conditions on the skeleton the process will remain arbitrarily large and in the d-path in certain important regions for a finite time. The condition that the process stay in and large through the path is an important one; requiring the process to remain large ensures the error perturbations are negligible and thus the process behaves like the deterministic skeleton. We will make use of this lemma several times in proving the theorems.

Lemma 3. Suppose a d-path $R_0,\dots,R_d$ exists with companion matrices $A_0,\dots,A_d$. Suppose there exists a collection of positive, strictly increasing functions $f_i(\cdot)$ so that for $M<\infty$ with $\| x \|_v > M$, $x \in R_i$ implies $B_{f_i(\|x\|_v)}(A_i x) \subset R_{i+1}$ for $i=0,\dots,d-1$. Suppose $E|\xi_t|^r < \infty$ for some $r>0$. Assume all $\| A_i \|_m$ are finite and bounded away from zero. Then for any $\delta' > 0$, $M<\infty$ there exists $M_1<\infty$ so that for all $X_0 = x \in R_0$ with $\| x \|_v > M$ and $\| \prod_{i=0}^d A_i x \|_v > M_1$, the process stays larger than M in magnitude and stays in the d-path from time 1 to time d with probability greater than $1-d\delta'$.

Proof. Get the functions $f_i(\cdot)$. Choose $\delta'>0$ and $M<\infty$. By the assumption $E|\xi_t|^r<\infty$ there exists $M_2<\infty$ so that $P(|\xi_t| > \min_i\{f_i(M_2)\}) < \delta'/2$. Let $M^* = \max(M, M_2)$.

Let $C_k = \sum_{s=1}^{k+1}(\max_i \| A_i \|_m)^{k-s}$. Let $D = \max_{k\in\{0,\dots,d-1\}}(\| \prod_{i=k+1}^d A_i \|_m)$. Given $\delta'>0$ and since C_k, D are finite, the assumption on the errors implies there exists $M_1<\infty$ so that

$$P\left(|\xi_1| > \frac{M_1/D - M^*}{(k+1)C_k}\right) < \frac{\delta'}{2(k+1)}, \quad k=0,\dots,d-1. \tag{2.15}$$

Suppose that $X_0 = x \in R_0$. Since each $\| A_i \|_m < \infty$ we have that $\| \prod_{i=k+1}^d A_i \|_m < \infty$ for $k=0,\dots,d-1$. Then note that by a norm inequality

$$\| (\prod_{i=0}^d A_i) x \|_v \le \| \prod_{i=k+1}^d A_i \|_m \times \| (\prod_{i=0}^k A_i) x \|_v, \quad k=0,\dots,d-1. \tag{2.16}$$

Then if $\| \prod_{i=0}^d A_i x \|_v > M_1$, from (2.16)

$$\| (\prod_{i=0}^k A_i) x \|_v > \frac{M_1}{D}, \quad k=0,\dots,d-1. \tag{2.17}$$

Let $X_0 = x$ and write where the process stays in the d-path up until time $k+1$

$$X_{k+1} = (\prod_{i=0}^{k} A_i)x + \sum_{s=1}^{k+1}(\prod_{i=s+1}^{k+1} A_i)\nu_s, \quad k = 0,\ldots,d-1. \tag{2.18}$$

Let $I_k = (X_{k-1} \in R_{k-1}) \cap (\| X_{k-1} \|_v > M^*)$. Using (2.15)-(2.18), norm inequalities, Boole's inequality, subadditivity and the fact the errors are independent and identically distributed we have for M^* and for $k = 0,\ldots,d-1$

$$\begin{aligned}
&\sup_{\substack{x\in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^{d} A_i)x\|_v > M_1}} P\left((\| X_{k+1} \|_v \leq M^*) I_{k+1} \middle| X_0 = x \right) \\
\leq\ & \sup_{\substack{x\in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^{d} A_i)x\|_v > M_1}} P\left(\| (\prod_{i=0}^{k} A_i)x \|_v - \sum_{s=1}^{k+1}\prod_{i=s+1}^{k+1} \| A_i \|_m |\xi_s| \leq M^* \right) \\
\leq\ & \sup_{\substack{x\in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^{d} A_i)x\|_v > M_1}} P\left(\sum_{s=1}^{k+1} (\max_i \| A_i \|_m)^{k-s} |\xi_s| \geq \| (\prod_{i=0}^{k} A_i)x \|_v - M^* \right) \\
\leq\ & P\left(\sum_{s=1}^{k+1} (\max_i \| A_i \|_m)^{k-s} |\xi_s| \geq \frac{M_1}{D} - M^* \right) \\
\leq\ & P\left(C_k \sum_{s=1}^{k+1} |\xi_s| > \frac{M_1}{D} - M^* \right) \\
\leq\ & P\left(\max_{s\in\{1,\ldots,k+1\}} |\xi_s| > \frac{M_1/D - M^*}{(k+1)C_k} \right) \\
\leq\ & P\left(\cup_{s=1}^{k+1} |\xi_s| > \frac{M_1/D - M^*}{(k+1)C_k} \right) \\
\leq\ & (k+1)P\left(|\xi_t| > \frac{M_1/D - M^*}{(k+1)C_k} \right) \quad < \quad \frac{\delta'}{2}.
\end{aligned} \tag{2.19}$$

By subadditivity from (2.19)

$$\sup_{\substack{x\in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^{d} A_i)x\|_v > M_1}} P\left(\cup_{k=0}^{d-1} (\| X_{k+1} \|_v \leq M^*) I_k \middle| X_0 = x \right) < \frac{d\delta'}{2}. \tag{2.20}$$

As for the probability the process leaves the d-path, from the assumptions if we have $X_{k-1} = x_{k-1} \in R_{k-1}$ with $\| x_{k-1} \|_v > M^* \geq M$, then $B_{f_{k-1}(\|x_{k-1}\|_v)}(A_{k-1}x_{k-1}) \subset R_k$. Since

$M^* \geq M_2$ and $f_{k-1}(\cdot)$ is strictly increasing this implies for $k = 1, \ldots, d$

$$\begin{aligned}
\sup_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^d A_i)x\|_v > M_1}} P\left((X_k \notin R_k) I_k \middle| X_0 = x\right) &\leq \inf_{\substack{x_{k-1} \in R_{k-1} \\ \|x_{k-1}\|_v > M^*}} P(|\xi_t| > f_{k-1}(\| x_{k-1} \|_v)) \\
&\leq P(|\xi_t| > f_{k-1}(M^*)) \\
&< \frac{\delta'}{2}.
\end{aligned} \tag{2.21}$$

By subadditivity from (2.21)

$$\sup_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^d A_i)x\|_v > M_1}} P\left(\cup_{k=1}^d [(\| X_k \|_v \leq M^*) I_k] \middle| X_0 = x\right) < \frac{d\delta'}{2}. \tag{2.22}$$

Then from (2.20), (2.22) and using subadditivity the probability the process stays in the d-path and remains larger than $M^* \geq M$ is given by

$$\begin{aligned}
&\inf_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^d A_i)x\|_v > M_1}} P\left(\cap_{k=1}^d [(X_k \in R_k) \cap (\| X_k \|_v > M^*)] I_k \middle| X_0 = x\right) \\
\geq \quad & 1 - \sup_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^d A_i)x\|_v > M_1}} P\left(\cup_{k=1}^d [(X_k \notin R_k)] I_k \middle| X_0 = x\right) \\
& \quad - \sup_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\prod_{i=0}^d A_i)x\|_v > M_1}} P\left(\cup_{k=1}^d [(\| X_k \|_v \leq M^*)] I_k \middle| X_0 = x\right) \\
> \quad & 1 - d\delta'.
\end{aligned} \tag{2.23}$$

□

2.3 V-Uniform Ergodicity

Our first result on the V-uniform ergodicity of cyclic threshold autoregressive models is a revision of Tjøstheim (1990) Theorem 4.5. The original statement of the theorem was roughly this:

Tjosthiem: Theorem 1. Assume there is a k-cycle of indices $i_1 \to i_2 \to \ldots \to i_k \to i_1$ such that $x \in R_{i_j} \Rightarrow A_{i_j}x \in R_{i_{j+1}}$ for $\| x \|$ large. Moreover, assume that $\rho(\prod_{s=1}^{k} A_{i_s}) < 1$ so that there exists an integer h such that $\| (\prod_{s=1}^{k} A_{i_s})^h \| < 1$ for a matrix norm $\| \cdot \|$. If $\forall j, v \leq j \leq k+v, 1 \leq v \leq k$ we have

$$P\{(\prod_{s=v}^{j} A_{i_s}x + \sum_{u=1}^{j-v+1}(\prod_{s=u+v}^{j} A_{i_s})\varepsilon_{t+u} \notin R_{j+1})\} \bigcap(\prod_{s=v}^{j-1} A_{i_s}x + \sum_{u=1}^{j-v}(\prod_{s=u+v}^{j-1} A_{i_s})\varepsilon_{t+u} \in R_{i_j})\} = O(\| x \|^{-\varepsilon}) \tag{2.24}$$

for some $\varepsilon > 0$ as $\| x \| \to \infty$ and if there exists n such that for some u, $1 \leq u \leq n$

$$P(X_{t+u} \in \bigcup R_{i_s} | X_t = x \notin \bigcup R_{i_s}) = 1 - O(\| x \|^{-\delta}) \tag{2.25}$$

for some $\delta > 0$ as $\| x \| \to \infty$, then $\{X_t\}$ is geometrically ergodic.

Proof. see Tjøstheim (1990, Theorem 4.5) □

The condition that the process remains in a cycle once it reaches one for $\| x \|$ large is (2.24). We are guaranteed by (2.25) that we reach a cycle in a finite time for $\| x \|_v$ large. Taken together these two conditions specify the process behaves arbitrarily close to the skeleton process x_t for $\| x \|_v$ large. This implies certain conditions on the error distribution and the skeleton itself. We attempt to express these conditions more explicitly in terms of the error distribution and the behavior of the skeleton in order that they may be more easily verified.

As stated, our result will handle cases where the dynamical skeleton has a single limiting cycle of finite length. The skeleton must be such that points in the cycle are mapped onto rays in the interior of the next region in the cycle. They cannot fall on the thresholds. This allows us to bound the transition probabilities between regions within ε of either zero or one by picking $\| x \|$ to be arbitrarily large, since the larger $\| x \|_v$ is the further the points are mapped from the thresholds and the smaller the probability the errors can cause the process to change regions. Thus we can focus on the regions in the cycle when determining the condition for ergodicity.

For a vector norm $\| \cdot \|_v$ and a positive function $f(\cdot)$ recall that we denoted the ball of radius $f(\| x \|_v)$ around $A_i x$ by

$$B_{f(\|x\|_v)}(A_i x) = \{y : \| y - A_i x \|_v < f(\| x \|_v)\}.$$

To guarantee points in the cycle are mapped away from the thresholds, we suppose in (2.26) below that for some $M < \infty$ there exists a collection of positive, strictly increasing functions $f_{i_j}(\cdot)$ so that

$$\forall x \in R_{i_j}, j = 1, \ldots, k \text{ with } \| x \| > M, \quad B_{f_{i_j}(\|x\|_v)}(A_{i_j} x) \subset R_{i_{j+1(\text{mod } k)}},$$

i.e., points in the cycle must be mapped bounded away from the thresholds, with this bound increasing as $\| x \|_v \to \infty$.

Points not already in the cycle must be assured of reaching one in a finite time. We assume for $x \notin \cup_{j=1}^{k} R_{i_j}$ that under the action of the deterministic skeleton x follows a d-path $R_0, \ldots, R_{d-1}$ before x enters the cycle. We allow the d-path to vary from one x to another, but we require the length of the path $d = d(x) \leq n$ for some finite, uniform n.

We need points not in the cycle to either be mapped away from the thresholds so that the probability the errors disrupt the progress toward the cycle is negligible for large $\| x \|_v$ or we need these points to lie in regions the process hits with arbitrarily small probability and then be mapped with near certainty to one of the former regions. To accomplish this we assume in (2.27) below we can partition off the problematic subregions of each R_i and make the probability the process hits these subregions arbitrarily small by requiring $\| x \|_v$ large enough. We call these subregions $R_i^{'}$.

Formally, we suppose for each $i \notin \{i_1, \ldots, i_k\}$ there exists a positive, strictly increasing function $g_i(\cdot)$ such that $x \in R_i$ with $\parallel x \parallel_v > M$ implies either

$$B_{g_i(\|x\|)}(A_i x) \subset R_{i+1},\ i = 0, \ldots, d,$$

or that $x \in R_i^{'} \subset R_i$, where $\cup R_i^{'}$ is such that for arbitrary $\delta^{'} > 0$ there exists $M^{'} \geq M$ so that

$$\sup_{\substack{x \\ \|x\| > M^{'}}} P(X_1 \in \cup R_i^{'} | X_0 = x) < \delta^{'}.$$

In the case where one or more of the companion matrices is not full rank, the process may not remain large on certain subregions no matter the value of $\parallel x \parallel_v$ with which we begin. Where these are subregions of regions in the cycle this does not cause a problem since these subregions are taken care of when we show the test function satisfies the drift condition for V-uniform ergodicity in (1.3). The subregions of regions not in the cycle do cause a problem for us and we need to be able to write them off; that is, we handle cases where the skeleton maps points away from these regions and these regions can be made arbitrarily small so that for $\parallel x \parallel_v$ large the probability the process enters them is arbitrarily small as well.

Since we need the process to remain large to continue in its progress towards the cycle, we suppose that for arbitrary $M_4 < \infty$ there exists $M^{''} < \infty$ so that the set of points $\{x : \parallel x \parallel_v > M^{''}, \parallel (\prod_{i=1}^{d(x)} A_i) x \parallel_v \leq M_4\}$ is contained in the union of the subregions $R_i^{'}$. This will hold, for example, in cases where we can cut off tiny slices near the thresholds and have the process remain large on the remainder of the space. These tiny offending regions are 'transient' in a sense.

Theorem 1. Suppose there exists a k-cycle of regions $R_{i_1} \to R_{i_2} \to \ldots \to R_{i_k} \to R_{i_1}$ with companion matrices $A_{i_1}, \ldots, A_{i_k}$ so that for an arbitrary norm $\parallel \cdot \parallel$ there exists $M < \infty$ with $\parallel x \parallel > M$ implying $x \in R_{i_j} \Rightarrow A_{i_j} x \in R_{i_{j+1(\text{mod } k)}}$. Suppose for some $M < \infty$ there exists a collection of positive, strictly increasing functions $f_{i_j}(\cdot)$ so that

$$\forall x \in R_{i_j}, j = 1, \ldots, k \text{ with } \parallel x \parallel > M, \quad B_{f_{i_j}(\|x\|_v)}(A_{i_j} x) \subset R_{i_{j+1(\text{mod } k)}}, \tag{2.26}$$

Suppose there exists a uniform $n < \infty$ so that for each $x \notin \cup_{j=1}^{k} R_{i_j}$ with $\| x \| > M$ there exists an integer $d = d(x) \le n$ where under the action of the skeleton x follows the deterministic d-path $R_0 \to \ldots \to R_d$ with $d \in \{i_1, \ldots, i_k\}$, i.e., $x \in R_i \Rightarrow A_i x \in R_{i+1}$, $i = 0, \ldots, d-1$, before entering the cycle.

Assume for each $i \notin \{i_1, \ldots, i_k\}$ there exists a positive, strictly increasing function $g_i(\cdot)$ so $x \in R_i$ with $\| x \| > M$ implies either that

$$B_{g_i(\|x\|)}(A_i x) \subset R_{i+1},\ i = 0, \ldots, d \tag{2.27}$$

or that $x \in R_i^{'} \subset R_i$ where for arbitrary $\delta^{'} > 0$ there exists $M^{'} \ge M$ such that

$$\sup_{\substack{x \\ \|x\| > M^{'}}} P(X_1 \in \cup R_i^{'} | X_0 = x) < \delta^{'}. \tag{2.28}$$

Suppose for $M_4 < \infty$ there exists $M^{''} < \infty$ so that $\{x : \| x \|_v > M^{''}, \| (\prod_{i=1}^{d(x)} A_i) x \|_v \le M_4\}$ is contained in $\cup_{i=1}^{l} R_i^{'}$. If ξ_t has a continuous density everywhere positive, $E|\xi_t|^2 < \infty$ and $\rho(\prod_{j=1}^{k} A_{i_j}) < 1$, then $\{X_t\}$ is V-uniformly ergodic.

Proof. If ξ_t has a continuous density everywhere positive then we are assured $\{X_t\}$ is aperiodic and ψ-irreducible with the irreducibility measure being Lebesgue measure. It remains to construct a test function $V^{'}(\cdot)$ and show $\{X_t\}$ satisfies the conditions for $V^{'}$-uniform ergodicity in (2.3) and (2.4).

From the assumption $\rho(\prod_{j=1}^{k} A_{i_j}) < 1$, by Lemma 1 there exists a matrix norm $\| \cdot \|_m$ induced by a vector norm $\| \cdot \|_v$ and a constant $\lambda < 1$ so that

$$\| (\prod_{j=1}^{k} A_{i_j}) x \|_v \le \| \prod_{j=1}^{k} A_{i_j} \|_m \| x \|_v \le \lambda \| x \|_v\ . \tag{2.29}$$

Let $\{\sigma_k\}$ denote the collection of all out of cycle sequences from $\{1, \ldots, l\}$ of length k. Note that $k < \infty$ implies $\text{card}(\{\sigma_k\}) < \infty$ and $E|\xi_t| < \infty$, $\| A_i \|_m < \infty$, $i = 1, \ldots, l$, implies we can define constants $C_1, C_2 < \infty$ such that

$$C_1 > \| (\prod_{i=1}^{k} A_{i_j}) \|_m + \sum_{\{\sigma_k\}} \prod_{i=1}^{k} \| A_{\sigma_k(i)} \|_m, \tag{2.30a}$$

$$C_2 > \Big(\sum_{u=1}^{k} \prod_{i=u+1}^{k} \| A_{i_j} \|_m + \sum_{\{\sigma_k\}} \sum_{u=1}^{k} \prod_{i=u+1}^{k} \| A_{\sigma_k(i)} \|_m \Big) E|\xi_t|. \tag{2.30b}$$

Get $n < \infty$ from the assumptions. Get $M < \infty$ according to the assumptions.

Define $V(x) = \| x \|_v$ and $V_1(x) = E(V(X_n)|X_0 = x)$. Note that since the ξ_t are uncorrelated and using a norm inequality

$$\begin{aligned}
& \Big(E\big((V(X_n))^2\big|X_0 = x\big)\Big)^{1/2} \\
\leq \quad & \Big(E\big((\max_i \| A_i \|_m)^n \| x \|_v + \sum_{i=1}^{n} (\max_i \| A_i \|_m)^{n-i} |\xi_t|\big)^2\Big)^{1/2} \\
\leq \quad & (\max_i \| A_i \|_m)^n \| x \|_v + \Big(\sum_{i=1}^{n} (\max_i \| A_i \|_m)^{2(n-i)} \Big)^{1/2} (E|\xi_t|^2)^{1/2} \\
+ \quad & 2 \| x \|_v^{1/2} (E|\xi_t|)^{1/2} \Big(\sum_{i=1}^{n} (\max_i \| A_i \|_m)^{2n-i} \Big)^{1/2}.
\end{aligned} \tag{2.31}$$

Since $\| A_i \|_m < \infty$, $E|\xi_t|^2 < \infty$, then by (2.31) for large $\| x \|_v$ there exist $K_1, K_2 < \infty$ so that

$$\Big(E\big((V(X_n))^2\big|X_0 = x\big)\Big)^{1/2} \leq K_1 E(V(X_n)|X_0 = x) + K_2. \tag{2.32}$$

Likewise, since $C_1, C_2 < \infty$ there exist $K_3, K_4 < \infty$ so that

$$\Big(E((C_1 V(X_n) + C_2)^2 | X_0 = x)\Big)^{1/2} \leq K_3 E(V(X_n)|X_0 = x) + K_4. \tag{2.33}$$

Pick $\delta' > 0$ so that $\lambda + [C_1 k + (2n+1)K_3]\delta' < 1$.

To satisfy the drift condition (2.3) we are going to look at

$$\limsup_{\|x\|_v \to \infty} \frac{E(V_1(X_{k+1})|X_0 = x)}{V_1(x)} = \limsup_{\|x\|_v \to \infty} \frac{E(E[E(V(X_{n+k+1})|X_{n+1})|X_1]|X_0 = x)}{E(V(X_n)|X_0 = x)}. \tag{2.34}$$

We will proceed by bounding $E(V(X_{n+k})|X_n) = E(V(X_{n+k+1})|X_{n+1})$, splitting it into cases where X_n is in a cycle or not. After this, conditioning on X_1 allows us to deal with the cases $X_1 \in \cup_{i=1}^{l} R_i'$ and $X_1 \notin \cup_{i=1}^{l} R_i'$.

We want to bound the probability the process remains in the cycle and remains large if it begins at $X_n = x_n$ in the cycle with $\| x_n \|_v$ sufficiently large. By (2.26) for all $x_n \in \cup_{j=1}^{k} R_{i_j}$ with $\| x_n \|_v > M$ there exists a positive, strictly increasing function $f_{i_j}(\cdot)$ such that we have $B_{f_{i_j}(\|x_n\|_v)}(A_{i_j} x_n) \subset R_{i_{j+1}}$. Suppose w.l.o.g. that $x_n \in R_{i_1}$ with $\| x_n \|_v > M$. Then by this and the assumption on the errors, the conditions for Lemma 3 are satisfied implying there exist $M_1 < \infty$ and $M^* \geq M$ so that

$$\inf_{\substack{x_n \in R_{i_1} \\ \|x_n\|_v > M \\ \|(\prod_{j=1}^{k} A_{i_j}) x_n\|_v > M_1}} P\left(\cap_{j=n+1}^{n+k} [(X_j \in R_{i_j}) \cap (\| X_j \|_v > M^*)] \middle| X_n = x_n \right) > 1 - k\delta^{'}, \tag{2.35}$$

which provides a bound on the probability the process stays in the cycle and remains large if it begins in the cycle at $X_n = x_n \in \{x : \| (\prod_{j=1}^{k} A_{i_j}) x \|_v > M_1\}$. By (2.29) and (2.30b)

$$\begin{aligned} &\sup_{\substack{x_n \in R_{i_1} \\ \|x_n\|_v > M \\ \|(\prod_{j=1}^{k} A_{i_j}) x_n\|_v > M_1}} E(V(X_{n+k}) I\{\cap_{j=n+1}^{n+k} [(X_j \in R_{i_j}) \cap (\| X_j \|_v > M^*)]\} | X_n = x_n) \\ < \quad & \lambda V(x_n) + C_2. \end{aligned} \tag{2.36}$$

Note that by assumption for $M_1 < \infty$ there exists $M^{''} < \infty$ so that the set of points $\{x_n : \| x_n \|_v > M^{''}, \| (\prod_{j=1}^{k} A_{i_j}) x_n \|_v \leq M_1\}$ is contained in $\cup_{i=1}^{l} R_i^{'}$. By (2.30a), (2.30b) and (2.35)

$$\begin{aligned} &\sup_{\substack{x_n \in R_{i_1} \\ \|x_n\|_v > \max(M, M^{''}) \\ x_n \notin \cup_{i=1}^{l} R_i^{'}}} E(V(X_{n+k}) I\{\cup_{j=n+1}^{n+k} [(X_j \notin R_{i_j}) \cup (\| X_j \|_v \leq M^*)]\} | X_n = x_n) \\ < \quad & (C_1 V(x_n) + C_2) k\delta^{'}. \end{aligned} \tag{2.37}$$

Then by (2.36) and (2.37) we can say for $X_n = x_n$ in a cycle and sufficiently large:

$$\sup_{\substack{x_n \in R_{i_1} \\ \|x_n\|_v > \max(M, M^{''}) \\ x_n \notin \cup_{i=1}^{l} R_i^{'}}} E(V(X_{n+k}) | X_n = x_n) \leq \lambda V(x_n) + C_2 + (C_1 V(x_n) + C_2) k\delta^{'}. \tag{2.38}$$

To bound $E(V(X_{n+k})|X_n)$ when X_n is either not in a cycle or $\| X_n \|_v$ is not large enough, we need to consider what happens beginning at $X_0 = x$. We want to bound the probability points $X_0 = x$ with $\| x \|_v$ large and not in the cycle get to the cycle by a finite time n and remain large while doing so.

Now $\forall x$ such that $x \notin \cup_{j=1}^{k} R_{i_j}$, $x \notin \cup_{i=1}^{l} R_i^{'}$ with $\| x \|_v > M$, by assumption there exists a uniform $n < \infty$ and $d = d(x) \leq n$ such that x follows the d-path $R_0 \to \ldots \to R_d$ with $d \in \{i_1, \ldots, i_k\}$. Suppose w.l.o.g. that $X_0 = x \in R_0$ with $\| x \|_v > M$.

By (2.27) suppose there exists a positive, strictly increasing function $g_0(\cdot)$ such that $B_{g_0(\|x\|)}(A_0 x) \subset R_1$. By this and the assumption on the errors, the conditions for Lemma 3 are satisfied, implying there exists $M_2 < \infty$ so that for $M^* = \max(M, M^{''})$

$$\inf_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\Pi_{j=0}^{n} A_j)x\|_v > M_2}} P\left(\cap_{j=1}^{n} [(X_j \in R_j) \cap (\| X_j \|_v > M^*)] \middle| X_0 = x \right) > 1 - n\delta^{'}, \tag{2.39}$$

which provides a bound on the probability $x \notin \cup_{j=1}^{k} R_{i_j}$, $x \notin \cup_{i=1}^{l} R_i^{'}$ reaches a cycle by a finite time n and the process remains large while doing so. Then by (2.39) and Cauchy-Schwarz

$$\begin{aligned} &\sup_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\Pi_{j=0}^{n} A_j)x\|_v > M_2}} E\left(V(X_{n+k}) I\{(X_n \notin \cup_{j=1}^{k} R_{i_j}) \cup (\| X_n \|_v \leq \max(M, M^{''}))\} \middle| X_0 = x \right) \\ \leq &\sup_{\substack{x \in R_0 \\ \|x\|_v > M \\ \|(\Pi_{j=0}^{n} A_j)x\|_v > M_2}} \left(E\left((C_1 V(X_n) + C_2)^2 \middle| X_0 = x \right) \right)^{1/2} n\delta^{'}. \end{aligned} \tag{2.40}$$

By (2.28) and subadditivity there exists $M^{'} < \infty$ so that

$$\sup_{\substack{x \\ \|x\|_v > M^{'}}} P(X_n \in \cup_{i=1}^{l} R_i^{'} | X_0 = x) < n\delta^{'}, \tag{2.41}$$

implying by (2.30a), (2.30b), (2.41) and Cauchy-Schwarz

$$\begin{aligned} &\sup_{\substack{x \\ \|x\|_v > M'}} E\left(V(X_{n+k}) I\{X_n \in \cup_{i=1}^{l} R_i'\} \middle| X_0 = x\right) \\ &< \sup_{\substack{x \\ \|x\|_v > M'}} \left(E\left((C_1 V(X_n) + C_2)^2 \middle| X_0 = x\right)\right)^{1/2} n\delta'. \end{aligned} \tag{2.42}$$

By (2.38), (2.40) and (2.42)

$$\begin{aligned} &\sup_{\substack{x \in R_0 \\ \|x\|_v > \max(M, M') \\ \|(\prod_{j=0}^{n} A_j) x\|_v > M_2}} E\left(E(V(X_{n+k})|X_n) \middle| X_0 = x\right) \\ \leq &\sup_{\substack{x \in R_0 \\ \|x\|_v > \max(M, M') \\ \|(\prod_{j=0}^{n} A_j) x\|_v > M_2}} E\left(\lambda V(X_n) + C_2 + (C_1 V(X_n) + C_2) k\delta' \middle| X_0 = x\right) \\ + &\sup_{\substack{x \in R_0 \\ \|x\|_v > \max(M, M') \\ \|(\prod_{j=0}^{n} A_j) x\|_v > M_2}} \left(E\left((C_1 V(x_n) + C_2)^2 \middle| X_0 = x\right)\right)^{1/2} 2n\delta'. \end{aligned} \tag{2.43}$$

Note that by assumption for $M_2 < \infty$ there exists $M'''' < \infty$ so that the set of points $\{x : \| x \|_v > M'''', \| (\prod_{j=0}^{n} A_j) x \|_v \leq M_2\}$ is contained in $\cup_{i=1}^{l} R_i'$. Thus points $x \notin \cup_{i=1}^{l} R_i'$ with $\| x \|_v$ large enough remain large. One final complication remains: what if $x \in \cup_{i=1}^{l} R_i'$? By (2.28) there exists $M'' < \infty$ so that

$$\sup_{\substack{x \\ \|x\|_v > M''}} P(X_1 \in \cup_{i=1}^{l} R_i' | X_0 = x) < \delta', \tag{2.44}$$

implying by (2.30a), (2.30b), (2.44) and Cauchy-Schwarz

$$\begin{aligned} &\sup_{\substack{x \in R_0 \\ \|x\|_v > \max(M, M'', M'''')}} E\left(E(V(X_{n+k+1})|X_{n+1}) I\{X_1 \in \cup_{i=1}^{l} R_i'\} \middle| X_0 = x\right) \\ \leq &\sup_{\substack{x \in R_0 \\ \|x\|_v > \max(M, M'', M'''')}} \left(E((C_1 V(X_n) + C_2)^2 | X_0 = x)\right)^{1/2} \delta' \end{aligned} \tag{2.45}$$

and note that by (2.43) and the Markov property

$$\begin{aligned}
&\sup_{\substack{x\in R_0\\ \|x\|_v>\max(M,M'',M'''')}} E\left(E(V(X_{n+k+1})|X_{n+1})I\{X_1\not\in\cup_{i=1}^{l}R_i'\}\Big|X_0=x\right)\\
\le\;&\sup_{\substack{x_1\in R_0\\ \|x_1\|_v>\max(M,M')\\ \|(\prod_{j=0}^{n}A_j)x_1\|_v>M_2}} E\left(E(V(X_{n+k+1})|X_{n+1})\Big|X_1=x_1\right)\\
\le\;&\sup_{\substack{x\in R_0\\ \|x\|_v>\max(M,M')\\ \|(\prod_{j=0}^{n}A_j)x\|_v>M_2}} E\left(\lambda V(X_n)+C_2+(C_1V(X_n)+C_2)k\delta'\Big|X_0=x\right)\\
+\;&\sup_{\substack{x\in R_0\\ \|x\|_v>\max(M,M')\\ \|(\prod_{j=0}^{n}A_j)x\|_v>M_2}} \left(E\left((C_1V(x_n)+C_2)^2\Big|X_0=x\right)\right)^{1/2}2n\delta'.
\end{aligned}\tag{2.46}$$

Note that $E(V(X_n)|X_0=x)\to\infty$ as $\|x\|_v$ does. Then by (2.33), (2.45), (2.46), the choice of δ' and since R_0 is arbitrary we have

$$\begin{aligned}
&\limsup_{\|x\|\to\infty}\frac{E(V_1(X_{k+1})|X_0=x)}{V_1(x)}\\
=\;&\lim_{M\to\infty}\sup_i\sup_{\substack{x\in R_i\\ \|x\|_v>M}}\frac{E\left(E[(E(V(X_{n+k+1})|X_{n+1})|X_1]\Big|X_0=x\right)}{E(V(X_n)|X_0=x)}\\
\le\;&\lim_{M\to\infty}\sup_i\sup_{\substack{x\in R_i\\ \|x\|_v>M}}\frac{E(\lambda V(X_n)+C_2+(C_1V(X_n)+C_2)k\delta'|X_0=x)}{E(V(X_n)|X_0=x)}\\
+\;&\lim_{M\to\infty}\sup_i\sup_{\substack{x\in R_i\\ \|x\|_v>M}}\frac{(2n+1)\delta'[K_3E(V(X_n)|X_0=x)+K_4]}{E(V(X_n)|X_0=x)}\\
=\;&\lambda+[C_1k+(2n+1)K_3]\delta'<1.
\end{aligned}\tag{2.47}$$

Also, since $E|\xi_1|<\infty$ and $\|A_i\|_m<\infty$ for each i we have for all $N<\infty$

$$\begin{aligned}
\sup_{\|x\|_v\le N}E(V_1(X_{k+1})|X_0=x)\le&\sup_{\|x\|_v\le N}(\max_{j=1,\dots,l}\|A_j\|_m)^{n+k+1}\|x\|_v\\
&+\sum_{s=1}^{k}(\max_{j=1,\dots,l}\|A_j\|_m)^{n+k+1-s}E|\xi_t|<\infty,
\end{aligned}\tag{2.48}$$

so that by (2.47) and (2.48), (2.3) and (2.4) are satisfied by $V_1(x) = E(V(X_n)|X_0 = x)$.

Let $V^{'}(x) = 1 + V_1(x) = 1 + E(V(X_n)|X_0 = x)$ be our test function and we have $V^{'} \geq 1$, locally bounded and measurable with $V^{'}(x) \to \infty$ as $\| x \|_v \to \infty$. Since $V_1(\cdot)$ satisfies (2.3) and (2.4), so does $V^{'}(\cdot)$ and we have $\{X_{tk}\}$ is $V^{'}$-uniformly ergodic. Since

$$\begin{aligned}\sup_x \frac{E(V^{'}(X_1)|X_0 = x)}{V^{'}(x)} &= \sup_x \frac{E\left(1 + E(V(X_{n+1})|X_1)\middle|X_0 = x\right)}{1 + E(V(X_n)|X_0 = x)} \\ &= \sup_x \frac{1 + E(V(X_{n+1})|X_0 = x)}{1 + E(V(X_n)|X_0 = x)} \\ &\leq \sup_i \sup_x \frac{1 + \| A_i \|_m E(V(X_n)|X_0 = x) + E|\xi_t|}{1 + E(V(X_n)|X_0 = x)} < \infty\end{aligned} \tag{2.49}$$

we have $|||P|||_{V'} < \infty$ and so by Lemma 2 the process $\{X_t\}$ is $V^{'}$-uniformly ergodic as well. □

The next result handles the case where the dynamic skeleton has a finite number of cycles. Since the length and number of cycles are both finite there exists an integer $k < \infty$ equal to the product of the lengths of the cycles so that we can restrict our attention to the k-step chain $\{X_{tk}\}$. Once again we assume the space $\mathbb{R}^p$ can be partitioned into a finite number of regions $R_1, \ldots, R_l$. Since we are dealing with multiple cycles, we denote the cycles by $C_1, \ldots, C_m$, the length of C_i by k_i, the regions in cycle C_i by $R_1^{(i)}, \ldots, R_{k_i}^{(i)}$ and the companion matrices in cycle C_i by $A_1^{(i)}, \ldots, A_{k_i}^{(i)}$, $i = 1, \ldots, m$. The assumptions on the skeleton are the logical extensions of the assumptions on the skeleton contained in Theorem 1.

Theorem 2. Suppose there exist m cycles $C_1, \ldots, C_m$ with $m < \infty$, each cycle C_i of finite length k_i. Assume $\max_i\{k_i\} < \infty$ and each cycle C_i consists of regions $R_1^{(i)}, \ldots, R_{k_i}^{(i)}$ with companion matrices $A_1^{(i)}, \ldots, A_{k_i}^{(i)}$ such that $x \in R_j^{(i)} \Rightarrow A_j^{(i)} x \in R_{j+1 (\mathrm{mod}\ k)}^{(i)}$. In addition suppose for an arbitrary norm $\| \cdot \|$ there exists some $M < \infty$ and a collection of positive, strictly increasing functions $f_j^{(i)}(\cdot)$ so that

$$\forall x \in R_j^{(i)},\ \ j = 1,\ldots,k_i,\ \ i = 1,\ldots,m \text{ with } \| x \| > M, \quad B_{f_j^{(i)}(\|x\|)}(A_j^{(i)}x) \subset R_{j+1(\text{mod } k)}^{(i)}. \tag{2.50}$$

Suppose there exists a uniform $n < \infty$ such that for each $x \notin \cup_{i,j} R_j^{(i)}$ with $\| x \| > M$ there exists an integer $d = d(x) \leq n$ implying $x_d \in \cup_{i,j} R_j^{(i)}$. Suppose x follows the deterministic d-path $R_0 \to \ldots \to R_{d-1}$ before entering $\cup_{i,j} R_j^{(i)}$, with $R_0,\ldots,R_{d-1} \notin \cup_{i,j} R_j^{(i)}$, i.e., $x \in R_j \Rightarrow A_j x \in R_{j+1},\ j = 0,\ldots,d-1$.

Assume for each $R_i \notin \cup_{i=1}^m C_i$ either that $x \in R_i$ with $\| x \| > M$ implies the existence of a positive, strictly increasing function $g_i(\cdot)$ such that

$$B_{g_i(\|x\|)}(A_i x) \subset R_{i+1},\ i = 0,\ldots,d \tag{2.51}$$

or that for arbitrary $\delta' > 0$ there exists $M' \geq M$ such that

$$\sup_{\substack{x \\ \|x\| > M'}} P(X_1 \in R_i | X_0 = x) < \delta'. \tag{2.52}$$

Denote these latter regions by R_i' and suppose that for arbitrary $M_4 < \infty$ there exists $M'' < \infty$ so that $\{x : \| x \| > M'', \| (\prod_{i=1}^{d(x)} A_i) x \| \leq M_4\}$ is contained in $\cup R_i'$.

If ξ_t has a continuous density everywhere positive, $E|\xi_t|^2 < \infty$ and $\rho(\prod_{i=1}^{k_i} A_j^{(i)}) < 1$, $i = 1,\ldots,m$, then $\{X_t\}$ is V-uniformly ergodic.

Proof. The proof is much the same as that for the single cycle case, with some extensions. From the assumption $\rho(\prod_{j=1}^{k_i} A_j^{(i)}) < 1$ for $i = 1,\ldots,m$ we can get positive constants $\lambda_1,\ldots,\lambda_m$ from Ciarlet's Lemma (2.7) such that $\rho(\prod_{j=1}^{k_i} A_j^{(i)}) < \lambda_i < 1,\ i = 1,\ldots,m$. Note $k := \prod_1^m k_i < \infty$ and that since there are a finite number of the λ_i there exists λ such that $1 > \lambda > \max_i\{\lambda_i\}$. By Ciarlet's Lemma (2.7) for each cycle $C_i,\ i = 1,\ldots,m$ there exist vector norms $\| \cdot \|_i$ and matrix norms $\| \cdot \|_{m_i}$ so that $x \in \cup_{i=1}^m \cup_{j=1}^{k_i} R_j^{(i)}$ implies

$$\| (\prod_{j=1}^{k_i} A_j^{(i)}) x \|_i \leq \| \prod_{j=1}^{k_i} A_j^{(i)} \|_{m_i} \| x \|_i \leq \lambda \| x \|_i . \tag{2.53}$$

For $x \notin \cup_{i=1}^{m} \cup_{j=1}^{k_i} R_j^{(i)}$, the Euclidean norm will serve as the vector norm and the operator norm induced by this will serve as the matrix norm. We will denote both of these by $\| \cdot \|$.

Define $V(x) = \sum_{i=1}^{m} \| x \|_i I\{x \in C_i\} + \| x \| I\{x \notin \cup_{i=1}^{m} C_i\}$ and define the function $V_1(x) = E(V(X_n)|X_0 = x)$. To satisfy the drift condition (2.3) we are going to look at

$$\limsup_{\|x\| \to \infty} \frac{E(V_1(X_{k+1})|X_0 = x)}{V_1(x)} = \limsup_{\|x\| \to \infty} \frac{E(E[E(V(X_{n+k+1})|X_{n+1})|X_1]|X_0 = x)}{E(V(X_n)|X_0 = x)}. \tag{2.54}$$

For an integer $j < \infty$ let $\{\sigma_j\}$ be the collection of all out of cycle indices from $\{1, \dots, l\}$ of length j. Note card($\{\sigma_j\}$) $< \infty$ for each j.

Since $\| A_j^{(i)} \|_{m_i}, \| A_i \| < \infty$ and $\{\sigma_j\}$ is a finite set for each j we can choose constants $D_1, D_2 < \infty$ so that

$$D_1 > \max_i \left(\| \prod_{j=1}^{k} A_j^{(i)} \|_{m_i} + \sum_{\{\sigma_k\}} \| \prod_{i=1}^{k} A_{\sigma_k(i)} \|_{m_i} \right) \tag{2.55a}$$

$$D_2 > \max_i \left(\| \sum_{u=1}^{k} (\prod_{s=u+1}^{k} A_s^{(i)}) \|_{m_i} + \| \sum_{u=1}^{k} (\prod_{s=u+1}^{k} A_{\sigma_k(s)}) \|_{m_i} \right). \tag{2.55b}$$

By arguments similar to (2.31) - (2.33) and since $\| A_j^{(i)} \|_{m_i}, \| A_i \| < \infty$, $E|\xi_t|^2 < \infty$ and $D_1, D_2 < \infty$ we have there exist $K_5, K_6 < \infty$ so that

$$\left(E((D_1 V(X_n) + D_2)^2 | X_0 = x) \right)^{1/2} \leq K_5 E(V(X_n)|X_0 = x) + K_6. \tag{2.56}$$

Pick $\delta' > 0$ so that $\lambda + [D_1 k + (2n+1) K_5] \delta' < 1$. By arguments similar to those leading to (2.35) we have there exist $M^*, M_1 < \infty$ so that

$$\inf_i \inf_{\substack{x \in R_1^{(i)} \\ \|x\|_i > M \\ \|(\prod_{j=1}^{k} A_j^{(i)}) x\|_{m_i} > M_1}} P\left(\cap_{j=n+1}^{n+k} [(X_j \in R_j^{(i)}) \cap (\| X_j \|_i > M^*)] I_j \middle| X_n = x \right) > 1 - k\delta'. \tag{2.57}$$

The argument for (2.57) must be repeated for each $i = 1, \dots, m$, yielding a $M_1^{(i)}$ which works for that cycle. Setting $M_1 = \max(M_1^{(i)})$ gives the result. By the assumptions and arguments similar to those leading to (2.38) and by (2.57) we can say for $X_n = x_n$ in a cycle and sufficiently large:

$$\sup_{\substack{x_n \in C_i \\ \|x_n\|_i > \max(M, M'') \\ x_n \notin \cup_{i=1}^{l} R_i'}} E(V(X_{n+k})|X_n = x_n) \leq \lambda V(x_n) + D_2 + (D_1 V(x_n) + D_2)k\delta'. \tag{2.58}$$

By arguments similar to those leading to (2.45) and (2.46) we have

$$\begin{aligned} &\sup_{\substack{x \in R_0 \\ \|x\| > \max(M, M'', M'''')}} E\left(E(V(X_{n+k+1})|X_{n+1}) I\{X_1 \in \cup_{i=1}^{l} R_i'\} \middle| X_0 = x\right) \\ \leq &\sup_{\substack{x \in R_0 \\ \|x\| > \max(M, M'', M'''')}} \left(E((C_1 V(X_n) + C_2)^2 | X_0 = x)\right)^{1/2} \delta' \end{aligned} \tag{2.59}$$

and

$$\begin{aligned} &\sup_{\substack{x \in R_0 \\ \|x\| > \max(M, M'', M'''')}} E\left(E(V(X_{n+k+1})|X_{n+1}) I\{X_1 \notin \cup_{i=1}^{l} R_i'\} \middle| X_0 = x\right) \\ \leq &\sup_{\substack{x_1 \in R_0 \\ \|x_1\| > \max(M, M') \\ \|(\prod_{j=0}^{n} A_j) x_1\|_v > M_2}} E\left(E(V(X_{n+k+1})|X_{n+1}) \middle| X_1 = x_1\right) \\ \leq &\sup_{\substack{x \in R_0 \\ \|x\| > \max(M, M') \\ \|(\prod_{j=0}^{n} A_j) x\|_v > M_2}} E\left(\lambda V(X_n) + C_2 + (C_1 V(X_n) + C_2)k\delta' \middle| X_0 = x\right) \\ + &\sup_{\substack{x \in R_0 \\ \|x\| > \max(M, M') \\ \|(\prod_{j=0}^{n} A_j) x\|_v > M_2}} \left(E\left((C_1 V(x_n) + C_2)^2 \middle| X_0 = x\right)\right)^{1/2} 2n\delta'. \end{aligned} \tag{2.60}$$

Then by (2.58), (2.59), (2.60) and the choice of δ' we have

$$\begin{aligned}
&\limsup_{\|x\|\to\infty} \frac{E(V_1(X_{k+1})|X_0=x)}{V_1(x)} \\
&= \lim_{M\to\infty} \sup_{\substack{x\\ V(x)>M}} \frac{E\left(E[(E(V(X_{n+k+1})|X_{n+1})|X_1]\big|X_0=x\right)}{E(V(X_n)|X_0=x)} \\
&\leq \lim_{M\to\infty} \sup_{\substack{x\\ V(x)>M}} \frac{E(\lambda V(X_n)+D_2+(D_1V(X_n)+D_2)k\delta'|X_0=x)}{E(V(X_n)|X_0=x)} \qquad (2.61)\\
&+ \lim_{M\to\infty} \sup_{\substack{x\\ V(x)>M}} \frac{(2n+1)\delta'[K_5E(V(X_n)|X_0=x)+K_6]}{E(V(X_n)|X_0=x)} \\
&= \lambda+[D_1k+(2n+1)K_5]\delta' < 1.
\end{aligned}$$

Let $D=(\max(\| A_i \|, \max_i(\| A_i \|_{m_i})))^k$ with $D<\infty$ since each of the $\| A_i \|_{m_i}$, $\| A_i \|$ are finite and there are a finite number of them. Then since $E|\xi_t|<\infty$

$$\sup_{\substack{x\\ V(x)\leq M}} E(V_1(X_{k+1})|X_0=x) \leq \sup_{\substack{x\\ V(x)\leq M}} D\max(\| x \|_i, \| x \|)+\sum_{s=1}^{k} D^{k-s}E|\xi_s|<\infty. \qquad (2.62)$$

Let the test function be $V'(x)=1+V_1(x)=1+E(V(X_n)|X_0=x)$. Following from (2.61) and (2.62) we have that $V'(\cdot)$ satisfies (2.3), (2.4); also $V'\geq 1$ is locally bounded and measurable with $V'(x)\to\infty$ as $\| x \|\to\infty$. Thus $\{X_{tk}\}$ is V'-uniformly ergodic.

Note also that since $E|\xi_t|<\infty$

$$\begin{aligned}
\sup_x \frac{E(V'(X_1)|X_0=x)}{V'(x)} &= \sup_x \frac{E\left(1+E(V(X_{n+1})|X_1)\big|X_0=x\right)}{1+E(V(X_n)|X_0=x)} \\
&= \sup_x \frac{1+E(V(X_{n+1})|X_0=x)}{1+E(V(X_n)|X_0=x)} \\
&\leq \sup_i \sup_x \frac{1+\max(\| A_i \|_{m_i}, \| A_i \|)E(V(X_n)|X_0=x)+E|\xi_t|}{1+E(V(X_n)|X_0=x)} < \infty
\end{aligned} \qquad (2.63)$$

we have $|||P|||_{V'}<\infty$ and so by Lemma 2 the process $\{X_t\}$ is V'-uniformly ergodic as well. □

2.4 Transience

We have been able to identify some conditions under which cyclic and multi-cyclic models are transient. No doubt these conditions are stronger than what is necessary, but they are sufficient and do provide a beginning to the task of completely characterizing the parameter spaces of cyclic and multi-cyclic models.

We begin with a theorem describing conditions under which an AR(p) process is transient. Transience of cyclic and multi-cyclic processes follow as corollaries of this theorem under appropriate conditions on the skeleton that distill the asymptotic behavior of the process down to that of the k-cycle.

The theorem requires $\min_i |\lambda_i(A)| > 1$, where A is the companion matrix of the Markov chain $\{X_t\}$ in which the AR(p) process is embedded and $\lambda_i(A)$ are the eigenvalues of A. We are aware of the well-known condition for non-stationarity of an AR(p) process which is equivalent to the weaker $\rho(A) > 1$ condition for transience of $\{X_t\}$ (see Tjøstheim (1990), Theorem 4.4(ii)). However, the cyclic models demand a strict inequality in the drift conditions (2.5) and (2.6) for transience of the cycle; that is, we require either

$$E(g(X_1)|X_0 = x) < g(x), \quad x \in B^c \tag{2.64}$$

or

$$E(g(X_1)|X_0 = x) > g(x), \quad x \in B^c. \tag{2.65}$$

This requires that we assume the stronger condition $\min_i |\lambda_i(A)| > 1$.

The strict inequality is necessary to account for the extra terms corresponding to the process either leaving a cycle, not reaching a cycle by a certain time or not staying large enough for the assumptions on the skeleton to hold. The details of this are worked out in Corollary 1.

The theorem has a stronger error condition, $Ee^{|\xi_t|} < \infty$ than that used in establishing V-uniform ergodicity. This stronger error condition is necessitated by the exponential, strictly decreasing, $g(x)$ we use which gives a strict inequality in the drift condition (2.65). It is thus appropriate for application to cyclic and multi-cyclic models. The corollaries establishing transience of these models follow from this theorem.

Theorem 3. For a linear Markov chain $X_t = AX_{t-1} + \nu_t$ with $\nu_t = \xi_t(1,0,\ldots,0)'$ suppose the companion matrix A is full rank. Let λ_i, $i = 1,\ldots,\text{rank(A)}$, denote the distinct eigenvalues of A. Assume $\min_i |\lambda_i| > 1$. Suppose $Ee^{|\xi_t|} < \infty$ and that the error distribution has a density which is continuous and everywhere positive. Then the chain X_t is transient.

Proof. Since A is full rank, A^{-1} exists. The eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A, so $\min_i |\lambda_i| > 1$ implies $\rho(A^{-1}) = \max_i \frac{1}{|\lambda_i|} = \frac{1}{\min_i |\lambda_i|} < 1$. Let $y = Ax$; then $x = A^{-1}y$. Since $\rho(A^{-1}) < 1$, by Lemma 1 there exist $\lambda < 1$ and norms $\| \cdot \|_v$, $\| \cdot \|_m$ so that

$$\| A^{-1}y \|_v \leq \| A^{-1} \|_m \| y \|_v \leq \lambda \| y \|_v \,. \tag{2.66}$$

Then since $x = A^{-1}y = A^{-1}Ax$

$$\| x \|_v = \| A^{-1}Ax \|_v \leq \| A^{-1} \|_m \| Ax \|_v \leq \lambda \| Ax \|_v, \tag{2.67}$$

implying $\| Ax \|_v \geq \frac{1}{\lambda} \| x \|_v$. Write $\lambda' = \frac{1}{\lambda} > 1$ and then from (2.67) we have that
$\| Ax \|_v \geq \lambda' \| x \|_v > \| x \|$.

Since $Ee^{|\xi_t|} < \infty$ we can choose $1 < C < \infty$ with $Ee^{|\xi_t|} \leq C$. Choose $r > \frac{1}{1-\lambda'} \log(1/C)$ and note that

$$\frac{Ee^{-\|Ax+\nu_t\|_v}}{e^{-\|x\|_v}} \leq \frac{e^{-\|Ax\|_m} Ee^{|\xi_t|}}{e^{-\|x\|_v}} \leq Ce^{(1-\lambda')\|x\|_v} < Ce^{(1-\lambda')r} < 1, \quad \| x \|_v > r. \tag{2.68}$$

Let $B = \{x : \| x \|_v \leq r\}$, $B^c = \{x : \| x \|_v > r\}$. Let $g(x) = e^{-\|x\|_v}$. Then by (2.68)

$$\frac{E(g(X_1)|X_0 = x)}{g(x)} = \frac{Ee^{-\|Ax+\nu_t\|_v}}{e^{-\|x\|_v}} < 1, \quad x \in B^c \tag{2.69}$$

and so by this and the fact $g(x)$ is a strictly decreasing function we have

$$\frac{E(g(X_1)|X_0 = x)}{g(x)} < 1, \quad x \in B^c \tag{2.70a}$$

$$g(x) < \inf_{y \in B} g(y), \quad x \in B^c. \tag{2.70b}$$

Thus $g(x)$ satisfies Tweedie (1976, Theorem 11.3(i)). Since ξ_t has an error distribution which is continuous and everywhere positive, $\{X_t\}$ is ψ-irreducible, with ψ being Lebesgue measure. By Tweedie (1976, Theorem 11.3(iii)) then, $\{X_t\}$ is transient. □

With this theorem established and under similiar assumptions on the product(s) of matrices involved in the cycle(s), the transience of cyclic and multi-cyclic models are easy corollaries. For ease of exposition we consider first the case of a single cycle:

Corollary 1. Under the same assumptions on the skeleton as in Theorem 1, but with the additional assumptions $\min_i |\lambda_i(\prod_{j=1}^{k} A_{i_j})| > 1$, $(\prod_{j=1}^{k} A_{i_j})$ is full rank and supposing $Ee^{|\xi_t|} < \infty$, $\{X_t\}$ is transient.

Proof. The strategy is to show the k-step chain $\{X_{tk}\}$ is transient by demonstrating an appropriate function $g(x)$ and sets B, B^c exist that satisify Tweedie's criteria for transience:

$$\frac{E(g(X_{n+k})|X_0 = x)}{g(x)} \leq 1, \quad x \in B^c \tag{2.71a}$$

$$g(x) < \inf_{y \in B} g(y), \quad x \in B^c. \tag{2.71b}$$

Once this is established, by Tjøstheim (1990, Lemma 3.1) we have $\{X_t\}$ is transient if $\{X_{tk}\}$ is.

Let

$$I_C = I(X_n \in \cup_{j=1}^{k} R_{i_j}) \tag{2.72a}$$

$$I_L = I(\| X_p \|_v > M, p = n, \ldots, n+k) \tag{2.72b}$$

$$I_D = I(X_{n+j} \in R_{i_j}, j = 1, \ldots, k). \tag{2.72c}$$

Let $g(x) = e^{-\| x \|_v}$ as in Theorem 3. By Theorem 3 and the assumptions there exists $\gamma < 1$ so that we have $\frac{E(g(X_{n+k})I_C I_L I_D | X_0 = x)}{g(x)} \leq \gamma < 1$. By Lemma 3, Theorem 1 and $(\prod_{i=1}^{k} A_{i_j})$ full rank for an arbitary δ' there exist M_1, M_2 so that $\| x \|_v > M_1$ implies $\| (\prod_{i=1}^{k} A_{i_j}) \|_v > M_2$, implying in turn $E(I_c^c + I_L^c + I_D^c | X_0 = x) < (2n+1+k)\delta'$. Now we pick $\delta' > 0$ so that $\gamma + K(2n+1+k)\delta' < 1$. Let

$$B = \{x : \| x \|_v \leq M_1\}, \quad B^c = \{x : \| x \|_v > M_1\}. \tag{2.73}$$

Since $Ee^{|\xi_t|} < \infty$ and the $\| A_i \|_m$ are finite and bounded away from zero there exists $K < \infty$ so that for $\| x \|_v > M_1$

$$\frac{\left(E[(g(X_{n+k}))^2 | X_0 = x]\right)^{1/2}}{g(x)} \leq K. \tag{2.74}$$

Then note from (2.72), (2.73), (2.74), the choice of δ' and Cauchy-Schwarz we have

$$\begin{aligned} \frac{E(g(X_{n+k})|X_0 = x)}{g(x)} &\leq \gamma + \frac{E(g(X_{n+k})(I_C^c + I_L^c + I_D^c)|X_0 = x)}{g(x)} \\ &\leq \gamma + K(2n+1+k)\delta' < 1, \quad x \in B^c \end{aligned} \tag{2.75}$$

and

$$g(x) < \inf_{y \in B} g(y), \quad x \in B^c \tag{2.76}$$

so that Tweedie (1976, Theorem 11.3(i),(iii)) is satisfied; thus $\{X_{tk}\}$ is transient and by Tjøstheim (1990, Lemma 3.1) so is $\{X_t\}$. □

The case of a finite number of cycles of finite length is similar, the modifications being obvious.

Corollary 2. Under the same assumptions on the skeleton as in Theorem 2, other than supposing $\min_i |\lambda_i(\prod_{j=1}^{k_i} A_j^{(i)})| > 1$, $(\prod_{j=1}^{k_i} A_j^{(i)})$ is full rank for each i and supposing that $Ee^{|\xi_t|} < \infty$, $\{X_t\}$ is transient.

Proof. Similiar to Corollary 1. □

At this point, contrasting the conditions for V-uniform ergodicity with those for transience, there is clearly much of the parameter space between the two. There are several reasons for this. V-uniform ergodicity is the strongest form of ergodicity; weaker conditions corresponding to weaker forms of ergodicity may occupy some of this space. Left out of our treatment is discussion of null recurrence; conditions for null recurrence may fill even more of this void. Lastly, as stated we do not doubt that our conditions for transience are stronger than is necessary. For example, Cline and Pu (2000) showed that transience occurs if a companion matrix A has $\rho(A) > 1$ and the AR coefficients are all positive or are alternating in sign with the first one negative.

We conjecture that in the cyclic case

$$\rho(\prod_{j=1}^{k} A_{i_j}) > 1$$

and in the multi-cyclic case that

$$\max_i \rho(\prod_{j=1}^{k_i} A_j^(i)) > 1$$

are sufficient conditions for transience. If true, we expect these weaker conditions for transience will fill in the remainder of the parameter space. This will be a problem for future research.

2.5 Existence of Moments

It is known (see for example Tjøstheim (1990) Lemma 6.1) that under certain conditions existence of moments for the error distribution is equivalent to the existence of moments of the stationary distribution of $\{X_t\}$. Thus under certain conditions $E|\xi_t|^n < \infty$ implies $E|X_t|^n < \infty$ for n fixed. Here we pursue conditions under which all moments of the stationary distribution and central limit theorems can be shown at once to exist.

In the first section we derived conditions under which the process $\{X_t\}$ is V-uniformly ergodic, with $V(\cdot)$ being a function of a norm on the state space. Cline and Pu (2001) provide tools for deriving conditions under which the test function $V(\cdot)$ can be boosted to an exponential test function $V_1(s,x) = e^{(V(x))^s}$ for $s > 0$. Meyn and Tweedie (1993) link the order of the test function $V(\cdot)$ to the existence of moments of the stationary distribution. The implications for statistical inference are obvious and enormous: if the test function can be boosted to an exponential function of the norm all moments of the stationary distribution exist at once.

Cline and Pu (2001, Theorem 4) assume $\{X_t\}$ is an aperiodic, ψ-irreducible T-chain in $\mathbb{R}^p$ and $V : \mathbb{R}^p \to [1,\infty)$ is locally bounded with $V(x) \to \infty$ as $\| x \| \to \infty$. If we can find a random variable $W(x)$ satisfying the following:

$$V(X_1) \leq W(x), \text{ whenever } \mathrm{X}_0 = \mathrm{x},\ \forall \mathrm{x}, \tag{2.77a}$$

$$\limsup_{\|x\|\to\infty} E(\log(W(x)/V(x))) < 0, \tag{2.77b}$$

and if

$$|\log(W(x)/V(x))| + e^{(W(x))^r - (V(x))^r} \tag{2.78}$$

is uniformly integrable for some $r > 0$ then there exists $s > 0$ with $V_1(x) = e^{(V(x))^s}$ such that $\{X_t\}$ is V_1-uniformly ergodic.

Meyn and Tweedie (1993, Theorem 17.0.1) demonstrate if $\{X_t\}$ is a V-uniform ergodic Markov chain then for any function $g(\cdot)$ with $|g| \leq V$ and $g^2 \leq V$ then we have laws of large numbers and central limit theorems for $\frac{1}{n}\sum_{i=1}^{n} g(X_i)$. Thus, the choice of the test function implies both the ergodicity and the limit laws. If the conditions of Cline and Pu (2001, Theorem 4) are satisfied and the exponential boosting of our norm-like test function $V(\cdot)$ is valid, the existence of all moments of the stationary distribution follows from the fact that any polynomial function is eventually bounded by any exponential (i.e., given any exponential $V(\cdot)$ that satisfies the drift condition, we can find finite constants K,C so that $g(x)^2 \leq KV(x) + C$ and $KV(x) + C$ will also satisfy the drift condition) .

The assumption that $\{X_t\}$ is a T-chain is easily verified for threshold autoregressive models. Since the deterministic skeleton is bounded on compact sets, if we assume the errors ξ_t have a continuous density that is everywhere positive we have that compact sets are petite. By Meyn and Tweedie (1993, Theorem 6.2.5) the fact that compact sets are petitie implies $\{X_t\}$ is a T-chain.

In view of these results, establishing the existence of moments for the stationary distribution is simply a matter of finding a random variable $W(x)$ that satisifies (2.77) and (2.78). For cyclic and multi-cyclic threshold autoregressive models it suffices to choose $V(x)$ to be the test function used in satisfying the drift condition for V-uniform ergodicity. $W(x)$ can then be gotten by piecing together the appropriate function from the steps involved in demonstrating the drift condition for V-uniform ergodicity.

We demonstrate the same conditions on the skeleton and thus the same parameter space as for V-uniform ergodicity guarantee exponential boosting is valid. Of course, to enable exponential boosting we need to exponentially boost our condition on the error distribution.

This first theorem will handle the single cycle case covered in Theorem 1.

Theorem 4. Suppose the assumptions of Theorem 1 on the deterministic skeleton of $\{X_t\}$ hold, that for some $r > 0$ that $Ee^{(|\xi_t|)^r} < \infty$ and ξ_t has a density which is continuous and everywhere positive. Then for a vector norm $\|\cdot\|_v$ there exist $0 < s < 1$ and $V''(x) = e^{(V'(x))^s} = e^{(E(V(X_n)|X_0=x)+1)^s}$ such that $\{X_t\}$ is V''-uniformly ergodic.

Proof. The strategy is to show that exponential boosting is valid for the k-step chain $\{X_{tk}\}$ and then to apply Lemma 2 to extend this boosting to the one-step chain $\{X_t\}$.

Get $V_1(x) = E(V(X_n)|X_0 = x) + 1$ from Theorem 1. Let

$$I_{C^c} = I\{(X_n \notin \cup_{i=1}^{k} R_{i_j}) \cup (\| X_n \|_v \leq M)\}, \tag{2.79a}$$

$$I_\sigma = I\{\cup_{i=1}^{k}(X_{n+i} \notin \cup_{j=1}^{k} R_{i_j})\} \tag{2.79b}$$

$$I_1 = I\{(\cup_{j=1}^{k}(\| X_{n+j} \|_v \leq M)) \cup (\cup_{j=1}^{k}(X_{n+j} \in \cup_{i=1}^{l} R_i^{'}))\} \tag{2.79c}$$

$$I_2 = I\{(\cup_{j=1}^{n}(\| X_j \|_v \leq M)) \cup (\cup_{j=1}^{n}(X_j \in \cup_{i=1}^{l} R_i^{'}))\}. \tag{2.79d}$$

Define similar to Theorem 1

$$C_1 > \| (\prod_{i=1}^{k} A_{i_j}) \|_m + \sum_{\{\sigma_k\}} \prod_{i=1}^{k} \| A_{\sigma_k(i)} \|_m, \tag{2.80a}$$

$$C_2 > \Big(\sum_{u=1}^{k} \prod_{i=u+1}^{k} \| A_{i_j} \|_m + \sum_{\{\sigma_k\}} \sum_{u=1}^{k} \prod_{i=u+1}^{k} \| A_{\sigma_k(i)} \|_m \Big). \tag{2.80b}$$

Then define for a suitable $D < \infty$

$$\begin{aligned} W(x) = \lambda E(V(X_n)|X_0 = x) + C_2|\xi_t| + 1 + C_1 D I(\| x \|_v \leq M) \\ + (C_1 + C_2)(I_\sigma + I_1) + (C_3 + C_4)(I_2 + I_{C^c}) \end{aligned} \tag{2.81}$$

and clearly $V_1(X_{n+k}) \leq W(x)$ whenever $X_n = x$.

Also, since $\log(\cdot)$ is a monotone continuous function and by Jensen's inequality we have by (2.47)

$$\limsup_{\|x\|_v \to \infty} E\left(\log\left(\frac{W(x)}{V_1(x)}\right)\right) \leq \log\left(\limsup_{\|x\|_v \to \infty} E\left(\frac{W(x)}{V_1(x)}\right)\right) < 0. \tag{2.82}$$

Since $Ee^{(|\xi_t|)^r} < \infty$ implies $E|\xi_t| < \infty$ we have

$$\sup_x E(W(x)/V_1(x)) \leq \lambda + C_2 E|\xi_t| + 1 + C_1 D + 2C_1 + 2C_2 < \infty. \tag{2.83}$$

Note also that since $\lambda < 1$

$$\frac{W(x)}{V_1(x)} > \frac{\lambda E(V(X_n)|X_0 = x) + 1}{E(V(X_n)|X_0 = x) + 1} = \lambda + \frac{1 - \lambda}{E(V(X_n)|X_0 = x) + 1} > \lambda > 0, \tag{2.84}$$

implying

$$\left|\log\left(\frac{W(x)}{V_1(x)}\right)\right| < \left|\log(\lambda)\right| < \infty \tag{2.85}$$

when $\frac{W(x)}{V_1(x)} < 1$. Then by (2.83), (2.84) and (2.85)

$$\begin{aligned} & E\left|\log\left(\frac{W(x)}{V_1(x)}\right)\right| \\ = \quad & E\Big(\log\left(\frac{W(x)}{V_1(x)}\right)\times I\Big(\frac{W(x)}{V_1(x)}\geq 1\Big) - \log\left(\frac{W(x)}{V_1(x)}\right)\times I\Big(\frac{W(x)}{V_1(x)}<1\Big)\Big) \\ < \quad & E\Big(\log\left(\frac{W(x)}{V_1(x)}\right)\times I\Big(\frac{W(x)}{V_1(x)}\geq 1\Big) - \log(\lambda)\Big) \\ < \quad & \log\Big(E\Big(\frac{W(x)}{V_1(x)}\Big)\Big) - \log(\lambda) < \infty. \end{aligned} \tag{2.86}$$

Suppose w.l.o.g that $r<1$. Then since $\lambda<1$ and for a constant $C_5 = C_5(M, C_1, C_2) < \infty$ this implies by the assumption on ξ_t

$$\sup_x E e^{(W(x))^r - (V_1(x))^r} \leq \sup_x E e^{(\lambda E(V(X_n)|X_0=x)+1)^r - (E(V(X_n)|X_0=x)+1)^r} e^{C_5|\xi_t|^r} \leq E e^{C_5|\xi_t|^r} < \infty, \tag{2.87}$$

implying that $e^{(W(x))^r - (V_1(x))^r}$ is uniformly integrable. Then by properties of the supremum so is the sum $|\log(W(x)/V_1(x))| + e^{(W(x))^r-(V_1(x))^r}$.

The conditions of Cline and Pu (2.77) are satisfied and the k-step chain $\{X_{tk}\}$ is thus V''-uniformly ergodic with $V''(x) = e^{(E(V(X_n)|X_0=x)+1)^s}$ for some $s>0$. Suppose that $s<r$ then because $Ee^{|\xi_t|} < \infty$ and the $\| A_i \|_m$ are bounded we have that

$$E\big(\parallel X_{n+1}\parallel_v \big| X_0 = x\big) \leq \sup_i \parallel A_i \parallel_m E\big(\parallel X_n \parallel_v \big| X_0 = x\big) + E|\xi_t|$$

and so we can find appropriate $N_1, N_2 < \infty$ so that

$$\sup_x \frac{E(V''(X_1)|X_0=x)}{V''(x)} < N_1 e^{N_2}$$

implying $|||P|||_{V''} < \infty$ by (2.11) and by Lemma 2 $\{X_t\}$ is V''-uniformly ergodic with test function $V''(x) = e^{(E(V(X_n)|X_0=x)+1)^s}$ for some s with $0<s<1$. $\square$

Next we move on to the multi-cyclic case discussed in Theorem 2.

Theorem 5. Suppose the assumptions of Theorem 2 on the deterministic skeleton of $\{X_t\}$ hold, that for some $r > 0$ that $Ee^{(|\xi_t|)^r} < \infty$ and ξ_t has a density which is continuous and everywhere positive. Then there exist $s > 0$, $s < 1$ and $V^{''}(x) = e^{(V^{'}(x))^s}$ such that $\{X_t\}$ is $V^{''}$-uniformly ergodic.

Proof. Similar to Theorem 4, with obvious modifications made for the assumptions in Theorem 2. □

Taken together, Theorems 4 and 5 imply that the assumptions made on the deterministic skeletons in Theorems 1 and 2 are adequate for exponential boosting when combined with an exponential condition on the error distribution. Under these conditions all moments of the stationary distribution exist and we have laws of large numbers for partial sums $\sum_{i=1}^{n} g(X_i)$, where g is any polynomial function.

CHAPTER III

ERGODICITY OF THRESHOLD AUTOREGRESSIVE MODELS THROUGH APPROXIMATION WITH A FINITE STATE MARKOV CHAIN

3.1 Introduction

3.1.1 Background

Consider the TAR(p) model $\{y_t\}_{t\geq 0}$ described in (1.1) embedded in a general state Markov chain $\{X_t\}$ according to (1.2) with the domain divided into l regions $R_1,\ldots,R_l$, each region having companion matrix A_j, $j\in\{1,\ldots,l\}$. Under certain assumptions on the general state Markov chain $\{X_t\}$ we will approximate the transitions of $\{X_t\}$ from region R_i to region R_j by the transitions of a finite state Markov chain on the states $\{1,\ldots,l\}$. We denote the finite state Markov chain by $\{J_t\}$. We will then derive ergodic conditions for $\{X_t\}$ through analysis of the simpler chain $\{J_t\}$ and incorporate the finite state chain into a test function for the general state space chain.

We are going to consider the space $\mathbb{R}^p$ to be equipped with the Euclidean norm. Let $\|\cdot\|$ denote the Euclidean norm, $\| x \|$ be the Euclidean norm of $x\in\mathbb{R}^p$ and for a matrix A let $\| A \|$ be the operator norm of A induced by the Euclidean norm. We assume all matrices have a finite operator norm.

We use the drift criteria for V-uniform ergodicity described in Chapter 1: for a locally bounded, measurable function $V\geq 1$ with $V\to\infty$ as $\| x \|\to\infty$ we require

$$\limsup_{\|x\|\to\infty}\frac{E(V(X_1)|X_0=x)}{V(x)}<1 \tag{3.1}$$

and for all $M<\infty$

$$\sup_{\|x\|\leq M} E(V(X_1)|X_0=x)<\infty. \tag{3.2}$$

In constructing the test function we use the directional method mentioned in Chapter I and detailed by Cline and Pu (2001). Our test function is of the form $V(s,x) = 1+\lambda(x) \parallel x \parallel^s$ where $s > 0$ and $\lambda(x)$ is piecewise constant, bounded and bounded away from zero, the values of $\lambda(x)$ depending upon the direction of x. The challenge in defining a test function is then to define the piecewise constants that comprise the function $\lambda(x)$.

3.1.2 Modelling $\{X_t\}$ with a finite state Markov chain

In the asymptotically deterministic case discussed in Chapter II it is possible to set up a trivial chain $\{J_t\}$ which tracks $\{X_t\}$ step by step beginning at $X_0 = x$ with a probability arbitrarily close to 1 when $\parallel x \parallel$ is large. We call this case asymptotically deterministic because for all i, j the probability of the transition from $X_0 = x \in R_i$ to $X_1 \in R_j$ can be bounded arbitrarily close to 0 or 1 by picking $\parallel x \parallel$ large enough. We can then determine conditions for stability of $\{X_t\}$ from conditions for stability of the implicit deterministic system $\{J_t\}$.

In the present chapter we explore more general cases where $\{X_t\}$ is not asymptotically deterministic and the step by step 'shadowing' of $\{X_t\}$ by $\{J_t\}$ fails. Specifically, regardless of the magnitude of $\parallel x \parallel$ there can be more than one region to which X_1 can travel, each with a probability not going to zero. We are forced to approximate the transitions of $\{X_t\}$ from region to region not with a determinstic system as we did in Chapter II, but rather with a simpler stochastic system, a finite state Markov chain.

The Markov chain $\{J_t\}$ is chosen so that for an arbitrary $\varepsilon > 0$ the transition probabilities of $\{J_t\}$ from state i to state j are within ε of the 'transition' probabilities of $\{X_t\}$ from region R_i to region R_j when $\{X_t\}$ is large. By this we mean $\{J_t\}$ is such that for an arbitrary $\varepsilon > 0$ there exists an $M < \infty$ so that

$$\sup_{i,j} \sup_{\substack{x \in R_i \\ \|x\| > M}} \left| P(X_1 \in R_j | X_0 = x) - P(J_1 = j | J_0 = i) \right| < \varepsilon. \tag{3.3}$$

The class of chains $\{X_t\}$ handled by this method is thus the class of chains amenable to this approximation.

We consider two cases. The first is where the ε-approximation holds for all states $1,\ldots,l$ and regions $R_1,\ldots,R_l$. We label this case (A1). In the second the ε-approximation holds for all recurrent states $1,\ldots,l$ and 'recurrent' regions $R_1,\ldots,R_l$. This case is labelled (A2). The case (A2) is more general and contains (A1), but exposition is helped by considering the simpler situation first and examining the issues here before moving on to the more complicated second case. Alternatively, in proving (A2) it is necessary to go through (A1) first, so (A1) can be thought of as a set of preliminary results to be used in proving (A2).

We demonstrate that under our assumptions the expectations of certain bounded, measurable functions of $\{J_t\}$ and of $\{X_t\}$ will be very close. In constructing our test function $V(s,x) = 1+\lambda(x) \parallel x \parallel^s$ we choose two particular bounded, measurable functions and we use the fact their expectations will be close in determining the values of the piecewise constant function $\lambda(x)$ used in the test function $V(s,x)$.

It is tempting to define a process $Y_t = \sum_{i=1}^{l} i\, I(X_t \in R_i)$ that keeps track of the 'states' of $\{X_t\}$ and, noting that $\log(\parallel A_i \parallel)$ describes the log-change of $\{X_t\}$ when it moves from region i, to attempt to ascertain conditions for the ergodicity of $\{X_t\}$ through appropriate conditions on the function $h(y) = \sum_{i=1}^{l} \log(\parallel A_i \parallel) I(y=i)$. However, this approach fails since $\{Y_t\}$ is not quite a finite state chain, the obvious problem being that the transition probabilities $P(Y_1 = j|Y_0 = i)$ of $\{Y_t\}$ are not constant because they depend upon where X_0 is in the region R_i. We are forced to use something slightly different. For a piecewise constant function $h(j) = \log(\parallel A_j \parallel + \delta)$ with $j = 1,\ldots,l$ we consider the function $h^{'}(X_t) = \sum_{j=1}^{l} h(j) I(X_t \in R_j)$. Note the difference in emphasis here: we have defined the function $h^{'}(\cdot)$ in terms of the process $\{X_t\}$ rather than $\{Y_t\}$. Since $\{X_t\}$ is a Markov chain and $h^{'}(\cdot)$ a bounded measurable function the Markov property holds. We use this fact to show the expectations of $h^{'}(X_t)$ and $h(J_t)$ will be close due to the ε-approximation when considered over a sufficiently long but finite time. We then demonstrate that if the appropriate condition on $h(J_t)$

$$E_\pi\big(h(J_t)\big) = \sum_{j=1}^{l} \pi_j h(j) < 0,$$

holds for every stationary distribution π of $\{J_t\}$ then $\{X_t\}$ will be V-uniformly ergodic.

3.1.3 Previous results

We make use of two results from Cline and Pu (2002) pertaining to the behavior of the long term average of the function $h(j) = \log(\parallel \mathrm{A_j} \parallel + \delta)$, where $j = 1,\ldots,l$ are the states of the Markov chain $\{J_t\}$, and of certain functions of $h(\cdot)$ which will be defined below. These are included in the proof of their Theorem 4.1; we have taken the liberty of separating them out and writing them as lemmas. These two lemmas give the necessary condition on $\{J_t\}$ and help us to define the piecewise constant terms we will use in constructing the test function $V(s,x) = 1 + \lambda(x) \parallel x \parallel^s$ that demonstrates the V-uniform ergodicity of $\{X_t\}$.

Lemma 1. Let $\{J_t\}$ be a finite-state chain on $\{1,\ldots,l\}$. Decompose the state space $S = \{1,\ldots,l\} = (\bigcup_{i=1}^{k} S_i) \bigcup T$, where each S_i is irreducible and recurrent and T is the set of all transient states. Let $\pi^{(i)}$ be the stationary distribution for S_i, where $\pi_j^{(i)} > 0$ for $j \in S_i$ and $\pi_j^{(i)} = 0$ for $j \notin S_i$. Then under the assumption

$$\Pi_{j=1}^{l} \parallel A_j \parallel^{\pi_j^{(i)}} < 1, \quad \forall i \in \{1,\ldots,k\}$$

we can define $h(j) = \log(\parallel \mathrm{A_j} \parallel + \delta)$, $j \in \{1,\ldots,l\}$, where δ is chosen such that $\pi^{(i)} h = \sum_j \pi_j^{(i)} h(j) < 0$, $i = 1,\ldots,k$ and there exists a finite n such that

$$\frac{1}{n}\sum_{t=1}^{n} E(h(J_t)|J_0 = i) < 0, \quad \forall i \in \{1,\ldots,l\}. \tag{3.4}$$

Proof. See Cline and Pu (2002) Theorem 4.1. □

Lemma 2. Following Lemma 2, define for each i

$$\tilde{h}(i) = \sum_{t=0}^{n-1} \frac{n-t}{n} E(h(J_t)|J_0 = i) \tag{3.5}$$

and for $s > 0$ let $H_1(s,i) = e^{s\tilde{h}(i)}$. Then there exists s_1 such that for $s < s_1$

$$\sup_i E\left(\frac{H_1(s,J_1)(\| A_i \| + \delta)^s}{H_1(s,i)} \middle| J_0 = i \right) < 1. \tag{3.6}$$

Proof. See Cline and Pu (2002) Theorem 4.1. □

3.2 Results

This first original lemma assures us that by picking $X_0 = x$ large enough and restricting our attention to certain subregions of the space, the process will remain large for a finite time with high probability if it remains in these subregions. It will be necessary for the process to remain large in order that our conditions may hold.

This lemma serves a purpose similar to that of Lemma 3 in Chapter II but the lemma is different and contrasting the two lemmas points up the difference in the classes of models considered in Chapter II versus those considered here. In Chapter II the simpler stochastic system used to approximate the transitions of $\{X_t\}$ from region to region was in fact deterministic; thus the requirement in Lemma 3 Chapter II that $\{X_t\}$ be mapped to a particular region. Here there is no such requirement because the simpler stochastic system used to approximate the transitions of $\{X_t\}$ from region to region is stochastic; there can be more than one region $\{X_t\}$ can be mapped to, each with a positive probability regardless of the magnitude of $\| x \|$. This explains the requirement below that $\{X_t\}$ be mapped to a particular *collection* of regions, not a particular region.

The lemma is trivially satisfied if the companion matrices are all of full rank but unfortunately this is not always the case.

Lemma 3. Assume there exists $r > 0$ for which $E|\xi_t|^r < \infty$. For all R_i, $i = 1,\dots,l$ define for an arbitrary $\delta_3 > 0$, $M < \infty$

$$R_i(\delta_3) = \{x \in R_i : \| x \| > M, \| A_i x \| > \delta_3 \| x \|\}, \quad i = 1,\dots,l. \tag{3.7}$$

Then given a finite $n < \infty$ if $t \le n$, there exists $D = D(t,\delta_3) < \infty$ such that

$$\inf_i \inf_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\left(\cap_{i=1}^{t} \{ (\| X_i \| > M) \cap (\cap_{j=0}^{i-1} (X_j \in \cup_{i=1}^{l} R_i(\delta_3))) \} \middle| X_0 = x \right) > 1 - n\delta_3. \tag{3.8}$$

Proof. For an arbitrary $\delta_3 > 0$ suppose for some $i \in \{1,\dots,l\}$ that $X_0 = x \in R_i(\delta_3)$. By the assumption $E|\xi_t|^r < \infty$ we can pick $1 < C < \infty$ so that

$$P\big(|\xi_t| > (\delta_3 C - 1)M\big) < \delta_3. \tag{3.9}$$

Consider that for $\| x \| > CM$ by (3.7) and (3.9)

$$\begin{aligned}
\sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > CM}} P(\| X_1 \| \le M | X_0 = x) &\le \sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > CM}} P(\| A_i x \| - |\xi_t| \le M) \\
&= \sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > CM}} P(|\xi_t| \ge \| A_i x \| - M) \\
&\le \sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > CM}} P(|\xi_t| \ge \delta_3 \| x \| - M) \\
&\le P(|\xi_t| \ge (\delta_3 C - 1)M) \\
&< \delta_3.
\end{aligned} \tag{3.10}$$

Likewise, $X_0 = x \in R_i(\delta_3)$ with $\| x \| > C^2 M$ implies using (3.9)

$$\begin{aligned}
\sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > C^2 M}} P(\| X_1 \| \le CM | X_0 = x) &\le P(|\xi_t| \ge (\delta_3 C - 1)CM) \\
&\le P(|\xi_t| \ge (\delta_3 C - 1)M) \\
&< \delta_3
\end{aligned} \tag{3.11}$$

and by (3.10) $X_1 = x_1 \in \cup_{i=1}^{l} R_i(\delta_3)$ with $\| x_1 \| > CM$ implies

$$\sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > C^2 M}} P\big((\| X_2 \| \leq M) \cap (X_1 \in \cup_{i=1}^{l} R_i(\delta_3)) \cap (\| X_1 \| > CM) | X_0 = x\big) < \delta_3. \quad (3.12)$$

Let $I_1(t) = \cap_{k=0}^{t-1}(X_k \in \cup_{i=1}^{l} R_i(\delta_3))$ and $I_2(t) = \cap_{k=0}^{t-1}(\| X_k \| > C^{t-j}M)$. By induction then for $j \in \{1, \ldots, t\}$

$$\begin{aligned} & \sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > C^t M}} P\left((\| X_j \| \leq C^{t-j}M) \cap I_1(j) \cap I_2(j) \middle| X_0 = x\right) \\ \leq \quad & P(|\xi_t| \geq (\delta_3 C - 1)C^{t-j}M) \quad (3.13) \\ \leq \quad & P(|\xi_t| \geq (\delta_3 C - 1)M) \\ < \quad & \delta_3. \end{aligned}$$

Let $D = C^t$. Then we have from (3.13), using DeMorgan's laws and Boole's inequality

$$\begin{aligned} & \inf_i \inf_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\left(\cap_{i=1}^{t} \{(\| X_i \| > M) \cap I_1(i) \cap I_2(i)\} \middle| X_0 = x\right) \\ \geq \quad & 1 - \sum_{j=1}^{t} \sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > C^t M}} P\left((\| X_j \| \leq C^{t-j}M) \cap I_1(j) \cap I_2(j) \middle| X_0 = x\right) \quad (3.14) \\ > \quad & 1 - t\delta_3. \end{aligned}$$

Since $C > 1$ and $t \leq n$ the result (3.8) follows. □

3.2.1 Case 1: uniform ε-bounds.

We assume the probabilities governing the transitions of $\{X_t\}$ from region to region can be approximated for all regions to within an arbitrary $\varepsilon > 0$ by taking $\| x \|$ large enough. Since the number of regions is finite the ε used is uniform over the entire space.

Assumption 1. **(A1)** *Suppose there exists a finite state Markov chain* $\{J_t\}$ *on the states* $\{1, \ldots, l\}$ *so that for arbitrary* $\varepsilon > 0$ *there exists* $M < \infty$ *with*

$$\sup_{i,j} \sup_{\substack{x \in R_i \\ \|x\| > M}} \left| P(X_1 \in R_j | X_0 = x) - P(J_1 = j | J_0 = i) \right| < \varepsilon. \quad (3.15)$$

Also, suppose $\{J_t\}$ *is such that*

$$\Pi_{i=1}^{l}(\| A_i \|)^{\pi_i} < 1, \tag{3.16}$$

for every stationary distribution π *of* $\{J_t\}$.

Based on (A1) and Lemmas 1-3 we can now introduce the functions that will define our piecewise constant function $\lambda(x)$ and we can demonstrate that for large x the expectations of these functions will be arbitrarily close.

Constructing the test function $V(s,x)$ is complicated by the fact that rather than a deterministic system, the approximating 'skeleton' is a stochastic system with transition probabilities arbitrarily close to the 'transition' probabilities of $\{X_t\}$. This requires that we must rely on expectations of the processes over the entire collection of states. We cannot rely on the pathwise behavior of the processes as we did in Chapter II.

Lemma 4. Suppose the conditions given in (A1) and Lemmas 1-3 hold for some $r > 0$. Define $h^{'}(x) = \sum_{j=1}^{l} h(j)I(x \in R_j)$. For a fixed $n < \infty$ let

$$\tilde{h}^{'}(x) = \sum_{t=0}^{n} \frac{n-t}{n} E(h^{'}(X_t)|X_0 = x,), \tag{3.17a}$$

$$\tilde{h}(j) = \sum_{t=0}^{n} \frac{n-t}{n} E(h(J_t)|J_0 = i). \tag{3.17b}$$

Then for arbitrary $\delta^{'} > 0$ there exists $D, M, n < \infty$ and $\delta_3 > 0$ so that

$$\sup_{i} \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \left| E(\tilde{h}(J_1) - \tilde{h}(i)|J_0 = i) - E(\tilde{h}^{'}(X_1) - \tilde{h}^{'}(x)|X_0 = x) \right| < \delta^{'}. \tag{3.18}$$

Proof. Let $N = \max_j |h(j)| = \max_j |\log(\| A_j \| + \delta)|$. Get $\{J_t\}$ according to (A1) and $n < \infty$ from Lemma 1 (3.4). Given $\delta^{'} > 0$ pick $\varepsilon > 0$ so that $(n+1)Nl^2\varepsilon < \delta^{'}$. Get $M < \infty$ from (A1). Pick $\delta_3 > 0$ so that $lN((n+1)l\varepsilon + n\delta_3) < \delta^{'}$. With $\delta_3 > 0$ and r from the assumptions get $D = D(n, \delta_3) > 1$ from Lemma 3 (3.8). Since we have defined $D = D(n, \cdot)$, then by Lemma 3 (3.8) holds $\forall t \leq n$.

Define $I_t = \cap_{i=1}^{t}\{(\| X_i \| > M) \cap (\cap_{j=0}^{i-1}(X_j \in \cup_{i=1}^{l} R_i(\delta_3)))\}$; then following from (3.15) we have for each i, j and $t \leq n$

$$\sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\left((X_t \in R_j) \cap I_{t-1} | X_0 = x\right) < P(J_t = j | J_0 = i) + (n+1)l\varepsilon \tag{3.19}$$

and

$$\inf_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\left((X_t \in R_j) \cap I_{t-1} | X_0 = x\right) > P(J_t = j | J_0 = k) - (n+1)l\varepsilon. \tag{3.20}$$

Lemma 3 (3.8) combined with (3.19) and (3.20) implies for $t \leq n$ and for all j

$$\inf_i \inf_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\left(X_t \in R_j | X_0 = x\right) > P(J_t = j | J_0 = i) - (n+1)l\varepsilon - n\delta_3, \tag{3.21a}$$

$$\sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\left(X_t \in R_j | X_0 = x\right) < P(J_t = j | J_0 = i) + (n+1)l\varepsilon + n\delta_3. \tag{3.21b}$$

Recall from (3.17) we have defined for all i, for all $x \in R_i$

$$\tilde{h}'(x) = \sum_{t=0}^{n} \frac{n-t}{n} E(h'(X_t) | X_0 = x), \quad \tilde{h}(j) = \sum_{t=0}^{n} \frac{n-t}{n} E(h(J_t) | J_0 = i).$$

Notice that since both $h(\cdot)$ and $h'(\cdot)$ are bounded, measurable functions the Markov property applies and we have for all i, for all $x \in R_i$

$$E(\tilde{h}'(X_1) - \tilde{h}'(x) | X_0 = x) = \frac{1}{n} \sum_{t=1}^{n} E(h'(X_t) | X_0 = x) - h'(x) \tag{3.22a}$$

$$E(\tilde{h}(J_1) - \tilde{h}(i) | J_0 = i) = \frac{1}{n} \sum_{t=1}^{n} E(h(J_t) | J_0 = i) - h(i) \tag{3.22b}$$

Consider for all i, for all $x \in R_i$

$$\frac{1}{n} \sum_{t=1}^{n} E(h'(X_t) | X_0 = x) = \frac{1}{n} \sum_{t=1}^{n} \sum_{j=1}^{l} h(j) P(X_t \in R_j | X_0 = x) \tag{3.23a}$$

$$\frac{1}{n} \sum_{t=1}^{n} E(h(J_t) | J_0 = i) = \frac{1}{n} \sum_{t=1}^{n} \sum_{j=1}^{l} h(j) P(J_t = j | J_0 = i) \tag{3.23b}$$

Recall that $N = \max_j |h(j)|$. We have for all i from (3.21), (3.22) and (3.23) and since $x \in R_i$ implies $h^{'}(x) = h(i)$

$$
\begin{aligned}
& \frac{1}{n}\sum_{t=1}^{n} E(h(J_t)|J_0 = i) - h(i) - (n+1)l^2 N\varepsilon - lNn\delta_3 \\
\leq \quad & \frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{l} h(j)(P(J_t = j|J_0 = i) - (n+1)l\varepsilon - n\delta_3) - h(i) \\
< \quad & \inf_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} \frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{l} h^{'}(j)P(X_t \in R_j|X_0 = x) - h^{'}(x) \\
\leq \quad & \sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} \frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{l} h^{'}(j)P(X_t \in R_j|X_0 = x) - h^{'}(x) \\
< \quad & \frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{l} h(j)(P(J_t = j|J_0 = i) + (n+1)l\varepsilon + n\delta_3) - h(i) \\
\leq \quad & \frac{1}{n}\sum_{t=1}^{n} E(h(J_t)|J_0 = i) - h(i) + (n+1)l^2 N\varepsilon + lNn\delta_3.
\end{aligned}
\tag{3.24}
$$

So from (3.22), (3.23), (3.24) and recalling the choices of ε and δ_3 we have the conclusion (3.18)

$$
\sup_i \sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} \left| E(\tilde{h}(J_1) - \tilde{h}(i)|J_0 = i) - E(\tilde{h}^{'}(X_1) - \tilde{h}^{'}(x)|X_0 = x) \right| < \delta^{'}.
$$

□

Now that we have shown the functions $\tilde{h}^{'}(X_t)$ and $\tilde{h}(J_t)$ are close in expectation, all that remains is to use this result to build a test function for $\{X_t\}$ demonstrating V-uniform ergodicity.

Theorem 1. Suppose the assumptions in (A1) and Lemmas 1-4 hold. Suppose as well that for arbitrary $\delta_4 > 0$ there exists $M^{'} < \infty$ so that the $R_i(\delta_3)$ for $i = 1, \ldots, l$ as defined in Lemma 3 exist with

$$
\sup_i \sup_{\substack{x\in R_i\\ \|x\|>M^{'}}} P(X_1 \notin \cup_{i=1}^{l} R_i(\delta_3)|X_0 = x) < \delta_4. \tag{3.25}
$$

Then there exists an $s > 0$ such that $\{X_t\}$ is V-uniformly ergodic with the test function $V(s,x) = 1 + \lambda(x) \parallel x \parallel^s$, where $\lambda(x)$ is piecewise constant, bounded and bounded away from zero.

Proof. Get $\{J_t\}$ from (A1). Get n from Lemma 2 such that (3.4) is true. Recall from (3.6) that there exist $s_1 < 1$ and $\beta < 1$ such that $\forall s < s_1$

$$\sup_i E\left(\frac{H_1(s,J_1)(\parallel A_i \parallel + \delta)^s}{H_1(s,i)}\middle| J_0 = i\right) \leq \beta < 1. \tag{3.26}$$

Since $\max_i(\parallel A_i \parallel + \delta)^s \to 1$ as $s \to 0$, then for an arbitrary $\delta_1 > 0$ there exists $s_2 > 0$ such that $\forall s < s_2$ we have

$$\max_i(\parallel A_i \parallel + \delta)^s < 1 + \delta_1. \tag{3.27}$$

Choose $\delta_1 > 0$ and then δ_2 so that

$$\delta_2 < \frac{1-\beta}{2(1+\delta_1)}, \tag{3.28}$$

then pick $\delta' > 0$ so that

$$\delta' < \frac{1-\beta}{(1+\delta_1)} - 2\delta_2. \tag{3.29}$$

Finally, pick $\varepsilon > 0$ so that $(n+1)l^2N\varepsilon < \delta'$ and $\delta_3 > 0$ so that $lN((n+1)l\varepsilon + n\delta_3) < \delta'$. For $\varepsilon > 0$ get M from (A1).

Let $H_1'(s,x) = e^{s\tilde{h}'(x)}$. Note that if we let $N = \max_j |h(j)|$, then for all x

$$\tilde{h}'(x) = \sum_{t=0}^{n} \frac{n-t}{n} E(h'(X_t)|X_0 = x) \leq nN \tag{3.30}$$

so that $1 \leq e^{s\tilde{h}'(x)} \leq e^{snN}$ and so we have from (3.30) that for some $K < \infty$

$$\sup_i \sup_{\substack{x \in R_i \\ \|x\| > M}} E\left(\frac{H_1'(s,X_1)(\parallel A_i \parallel + \delta)^s}{H_1'(s,x)}\middle| X_0 = x\right) \leq K \tag{3.31}$$

Given K from (3.26)-(3.31) we can find $\delta_4 > 0$ so that $\beta + (1+\delta_1)(\delta' + 2\delta_2) + K\delta_4 < 1$ and $M' < \infty$ so that (3.25) is satisfied.

Given δ' from (3.29) there exists $D = D(n, \delta_3) < \infty$ from (3.8) such that we have from (A1) and Lemma 4 (3.18)

$$\sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \left| E(\tilde{h}(J_1) - \tilde{h}(i) | J_0 = i) - E(\tilde{h}'(X_1) - \tilde{h}'(x) | X_0 = x) \right| < \delta'. \tag{3.32}$$

Note that for $y > 0$, fixed

$$\frac{y^s - 1}{s} \to \log(y), \quad \text{as} \quad \text{s} \to 0. \tag{3.33}$$

This limit is not uniform in y, but since the number of states l is finite, we can make the limit uniform when working with $\{J_t\}$, i.e., for arbitrary $\delta_2 > 0$ there exists s_3 such that $\forall s < s_3$, for all i

$$E(e^{s(\tilde{h}(J_1) - \tilde{h}(i))} | J_0 = i) > 1 + sE(\tilde{h}(J_1) - \tilde{h}(i) | J_0 = i) - \delta_2. \tag{3.34}$$

By (3.30) we have that $E(e^{s(\tilde{h}'(X_1) - \tilde{h}'(x))} | X_0 = x)$ is bounded and thus

$$\sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} E(e^{s(\tilde{h}'(X_1) - \tilde{h}'(x))} | X_0 = x) < \infty. \tag{3.35}$$

Also, (3.30) implies that

$$E(\tilde{h}'(X_1) - \tilde{h}'(x) | X_0 = x) \leq 2nN. \tag{3.36}$$

Taken together, (3.35) and (3.36) tell us that

$$E\left((\tilde{h}'(X_1) - \tilde{h}'(x))^2 e^{s(\tilde{h}'(X_1) - \tilde{h}'(x))} \middle| X_0 = x\right) \tag{3.37}$$

is bounded as well. Using a Taylor series expansion around zero we have

$$\begin{aligned} & E(e^{s(\tilde{h}'(X_1) - \tilde{h}'(x))} | X_0 = x) \\ < \quad & 1 + sE(\tilde{h}'(X_1) - \tilde{h}'(x) | X_0 = x) + \frac{s^2}{2!} E\left((\tilde{h}'(X_1) - \tilde{h}'(x))^2 e^{s^*(\tilde{h}'(X_1) - \tilde{h}'(x))} \middle| X_0 = x\right) \end{aligned} \tag{3.38}$$

for some $s^* \in [0,s]$. Thus, given $\delta_2 > 0$ as in (3.28) and (3.34) by (3.38) we can pick s_4 small enough so that $\forall s < s_4$ we have

$$\sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} E(e^{s(\tilde{h}'(X_1)-\tilde{h}'(x))}|X_0=x) \leq 1 + \sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} sE(\tilde{h}'(X_1)-\tilde{h}'(x)|X_0=x)+\delta_2. \tag{3.39}$$

Pick $s_5 = \min\{s_1,s_2,s_3,s_4\}$, then from (3.32), (3.34) and (3.39) for all $s < s_5$, for all i

$$\begin{aligned}
&\sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} E(e^{s(\tilde{h}'(X_1)-\tilde{h}'(x))}|X_0=x) - E(e^{s(\tilde{h}(J_1)-\tilde{h}(i))}|J_0=i) \\
< \quad &\sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} s(E(\tilde{h}'(X_1)-\tilde{h}'(x)|X_0=x) - E(\tilde{h}(J_1)-\tilde{h}(i)|J_0=i)) + 2\delta_2 \qquad (3.40)\\
< \quad & s\delta' + 2\delta_2.
\end{aligned}$$

Recalling that $H_1'(s,x) = e^{s\tilde{h}'(x)}$, from (3.40) for all $s < s_5$, for all i

$$\sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} E\left(\frac{H_1'(s,X_1)}{H_1'(s,x)}\middle|X_0=x\right) - E\left(\frac{H_1(s,J_1)}{H_1(s,i)}\middle|J_0=i\right) < s\delta' + 2\delta_2. \tag{3.41}$$

Equivalently, using (3.26), (3.27), supposing $s < s_5 \leq 1$ and recalling the choices of δ_1,δ_2 and δ' we have for all i

$$\begin{aligned}
&\sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} E\left(\frac{H_1'(s,X_1)(\| A_i \| + \delta)^s}{H_1'(s,x)}\middle|X_0=x\right) \\
< \quad & E\left(\frac{H_1(s,J_1)(\| A_i \| + \delta)^s}{H_1(s,i)}\middle|J_0=i\right) + (1+\delta_1)(s\delta' + 2\delta_2) \qquad (3.42)\\
< \quad & \beta + (1+\delta_1)(s\delta' + 2\delta_2) \\
< \quad & \beta + (1+\delta_1)(\delta' + 2\delta_2) < 1
\end{aligned}$$

and we have from (3.31), (3.42) and the definition of δ_4, M' in (3.25) that

$$\begin{aligned}
&\limsup_{\|x\|\to\infty} E\left(\frac{H_1'(s,X_1)(\| A_i \| + \delta)^s}{H_1'(s,x)}\middle|X_0=x\right) \\
\leq \quad & \sup_i \lim_{M\to\infty} \sup_{\substack{x\in R_i\\ \|x\|>DM}} E\left(\frac{H_1'(s,X_1)(\| A_i \| + \delta)^s}{H_1'(s,x)}\middle|X_0=x\right) \qquad (3.43)\\
< \quad & \beta + (1+\delta_1)(\delta' + 2\delta_2) + K\delta_4 < 1.
\end{aligned}$$

Define $H_2(s,x) = \| x \|^s$, with $s < \min(r, s_5, 1)$. Note that under the assumption $E|\xi_t|^r < \infty$ we have

$$\begin{aligned}
& \limsup_{\|x\|\to\infty} E\left(\frac{H_2(s,X_1)}{(\| A_i \| + \delta)^s H_2(s,x)} \middle| X_0 = x \right) \\
= \quad & \sup_i \lim_{M\to\infty} \sup_{\substack{x\in R_i \\ \|x\|>M}} E\left(\frac{H_2(s,X_1)}{(\| A_i \| + \delta)^s H_2(s,x)} \middle| X_0 = x \right) \\
\le \quad & \sup_i \lim_{M\to\infty} \sup_{\substack{x\in R_i \\ \|x\|>M}} \frac{\| A_i \|^s \| x \|^s}{(\| A_i \| + \delta)^s \| x \|^s} + \sup_i \lim_{M\to\infty} \sup_{\substack{x\in R_i \\ \|x\|>M}} \frac{E|\xi|^s}{(\| A_i \| + \delta)^s \| x \|^s} \\
< \quad & 1.
\end{aligned} \tag{3.44}$$

Define $V(s,x) = (H_1^{'}(s,x) H_2(s,x))^{1/2}$; then we have $\forall s < \min(r, s_5, 1)$, using (3.43), (3.44) and Cauchy-Schwarz

$$\limsup_{\|x\|\to\infty} E\left(\frac{V(s,X_1)}{V(s,x)} \middle| X_0 = x \right) < 1. \tag{3.45}$$

Also, since $e^{s\tilde{h}'(x)}$ is bounded for $M < \infty$ we have

$$\begin{aligned}
\sup_{\|x\|\le M} E(V(s,X_1)|X_0 = x) &= \sup_{\|x\|\le M} E((H_1^{'}(s,X_1) H_2(s,X_1))^{1/2} | X_0 = x) \\
&= \sup_{\|x\|\le M} E((e^{s\tilde{h}'(X_1)})^{1/2} (\| X_1 \|^s)^{1/2} | X_0 = x) \\
&\le \sup_i \sup_{\|x\|\le M} E((e^{s\tilde{h}'(X_1)})^{1/2} (\| A_i \|^s \| x \|^s + |\xi_t|^s)^{1/2} | X_0 = x) \\
&< \infty.
\end{aligned} \tag{3.46}$$

Let $V_1(s,x) = 1 + V(s,x)$; then $V_1 \ge 1$, V_1 is locally bounded and measurable and from (3.45), (3.46) we have

$$\limsup_{\|x\|\to\infty} E\left(\frac{V_1(s,X_1)}{V_1(s,x)} \middle| X_0 = x \right) < 1, \tag{3.47a}$$

$$\sup_{\|x\|\le M} E(V_1(s,X_1)|X_0 = x) < \infty, \tag{3.47b}$$

so we have by (3.1) and (3.2) that $\{X_t\}$ is V_1-uniformly ergodic. $\square$

3.2.2 Case 2: regions where ε-bounds do not hold correspond to transient states of $\{J_t\}$.

Placing ε-bounds on the probabilities of $\{X_t\}$ transitioning from region to region will often not be possible for all regions. Our test function relies on specific bounded, measurable functions of $\{X_t\}$ and $\{J_t\}$ that we require to be close in expectation. This will not be true if the probabilities are not close for all recurrent regions. Assuming the regions/states where the approximation does not hold are transient removes this problem.

Since $\{J_t\}$ is a finite state chain the collection of transient states of $\{J_t\}$ is finite and therefore uniformly transient, implying $\{J_t\}$ leaves these states in a finite time with a probability arbitrarily close to 1. This observation tells us that we need only wait a finite time and then we are back in Case 1. The following results are here essentially to deal with the complications created by having to wait a finite time for the processes to reach states/regions where ε-approximation is possible. First we modify (A1) to include the assumption the regions where the ε-approximation does not hold are 'transient'.

Assumption 2. **(A2)** *Suppose there exists a finite state Markov chain $\{J_t\}$ on $\{1,\ldots,l\}$ with G consisting of the recurrent and T consisting of the transient states for $\{J_t\}$. Suppose further there exists $t^* < \infty$ so that for arbitrary $\varepsilon > 0$ there exists $M < \infty$ with*

$$\sup_{j} \sup_{i\in G} \sup_{\substack{x\in R_i \\ \|x\|>M}} \left| P(X_1 \in R_j | X_0 = x) - P(J_1 = j | J_0 = i) \right| < \varepsilon \tag{3.48}$$

and

$$\sup_{i\in T} \sup_{\substack{x\in R_i \\ \|x\|>M}} P(X_{t^*} \in \cup_{k\in T} R_k | X_0 = x) < \varepsilon. \tag{3.49}$$

Suppose also that

$$\Pi_{i=1}^{l} \parallel A_i \parallel^{\pi_i} < 1, \tag{3.50}$$

for every stationary distribution π of $\{J_t\}$.

Assumption (A2) and Lemma 3 lead to the next result, which demonstrates that under the modified set of assumptions both $\{X_t\}$ and $\{J_t\}$ leave the regions/states where the

ε-approximation does not hold in a finite time with a probability arbitrarily close to 1, provided $\| x \|$ is large enough. As a consequence the Markov property guarantees the expectations of bounded, measurable functions of each chain will be similiar if we look at them for a suitably long but finite time. This is necessary because we define our piecewise constant function $\lambda(x)$ from the expectations of specific bounded, measurable functions of each process.

Lemma 5. For $\delta_3 > 0$, $M < \infty$ let $I_t = \cap_{i=1}^{t}\{(\| X_i \| > M) \cap (\cap_{j=0}^{i-1}(X_j \in \cup_{i=1}^{l} R_i(\delta_3)))\}$. If $|\xi_t|^r < \infty$ for some $r > 0$ then under (A2) for arbitrary $\varepsilon > 0$, with $M < \infty$, $t^* < \infty$ from (A2) and $1 < D < \infty$ from Lemma 3 there exists $t^{'} < \infty$ so that both of the following hold:

$$\sup_{i} \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} P\Big((X_{t'} \in \cup_{k \in T} R_k) \cap I_{t'-1} | X_0 = x\Big) < t^* \varepsilon \tag{3.51a}$$

$$\sup_{i} \sup_{j \in T} P(J_{t'} = j | J_0 = i) < \varepsilon \tag{3.51b}$$

Proof. Since T consists of the transient states of $\{J_t\}$ and $\{J_t\}$ is a finite state chain, we have that T is uniformly transient, meaning for arbitrary $\varepsilon > 0$ there exists a $t^{**} < \infty$ so that

$$\sup_{j \in T} \sup_{i \in T} P(J_{t^{**}} = j | J_0 = i) < \varepsilon. \tag{3.52}$$

Since $\{J_t\}$ is a finite state chain we can decompose the states $\{1, \ldots, l\}$ into $(\bigcup_{i=1}^{k} S_i) \bigcup T$, where $G = \bigcup_{i=1}^{k} S_i$ with each S_i irreducible and recurrent. This implies that

$$\sup_{j \in T} \sup_{i \in G} P(J_1 = j | J_0 = i) = 0 \tag{3.53}$$

since if $\sup_{j \in T} \sup_{i \in G} P(J_1 = j | J_0 = i) > 0$, then one of the S_i communicates with T, making either S_i transient or T recurrent which is a contradiction. From this and (3.48) in (A2)

$$\sup_{i \in G} \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > M}} P(X_1 \in \cup_{k \in T} R_k | X_0 = x) < \varepsilon. \tag{3.54}$$

From (3.49) in (A2) there exists $t^* < \infty$ so that for $\varepsilon > 0$ there exists $M < \infty$ with

$$\sup_{i\in T} \sup_{\substack{x\in R_i \\ \|x\|>M}} P(X_{t^*} \in \cup_{k\in T} R_k | X_0 = x) < \varepsilon. \tag{3.55}$$

If $t^* = t^{**}$ then set $t^{'} = t^* = t^{**}$. If $t^{**} < t^*$ we can set $t^{'} = t^*$ since immediately by (3.52), (3.53), (3.54) and (3.55) we will have

$$\sup_{i} \sup_{\substack{x\in R_i(\delta_3) \\ \|x\|>M}} P((X_{t'} \in \cup_{k\in T} R_k) \cap I_{t'-1} | X_0 = x) < t^*\varepsilon \tag{3.56a}$$

$$\sup_{i} \sup_{j\in T} P(J_{t'} = j | J_0 = i) < \varepsilon. \tag{3.56b}$$

If $t^{**} > t^*$ we can set $t^{'} = t^{**}$ as shown by the following. Consider that by the time homogeneous Markov property

$$P(X_{t^{**}} \in \cup_{k\in T} R_k | X_{t^{**}-t^*} = x) = P(X_{t^*} \in \cup_{k\in T} R_k | X_0 = x). \tag{3.57}$$

Then we have from (3.56a), (3.57) by iterating the expectation

$$\begin{aligned}
& \sup_{i} \sup_{\substack{x\in R_i(\delta_3) \\ \|x\|>DM}} P((X_{t^{**}} \in \cup_{k\in T} R_k) \cap I_{t^{**}-1} | X_0 = x) \\
= \; & \sup_{i} \sup_{\substack{x\in R_i(\delta_3) \\ \|x\|>DM}} E\left(P((X_{t^{**}} \in \cup_{k\in T} R_k) \cap I_{t^{**}-1} | X_{t^{**}-t^*}) \Big| X_0 = x \right) \\
\leq \; & \sup_{i} \sup_{\substack{x\in R_i(\delta_3) \\ \|x\|>DM}} E\Big(\sup_{i} \sup_{\substack{x_{t^{**}-t^*}\in R_i(\delta_3) \\ \|x_{t^{**}-t^*}\|>M}} P((X_{t^{**}} \in \cup_{k\in T} R_k) \cap I_{t^{**}-1} | X_{t^{**}-t^*} = x_{t^{**}-t^*}) \Big| X_0 = x \Big) \\
< \; & t^*\varepsilon.
\end{aligned} \tag{3.58}$$

Of course, if $t^{'} = t^{**}$ immediately from (3.52) and (3.53) we have

$$\sup_{i} \sup_{j\in T} P(J_{t'} = j | J_0 = i) < \varepsilon. \tag{3.59}$$

Putting the cases together we can set $t^{'} = \max(t^*, t^{**})$ and the result follows from (3.56), (3.58) and (3.59). □

Lemma 6 is Lemma 4 rewritten to account for the complications induced by the 'transient' regions whose transition probabilities cannot be approximated. Lemma 6 makes use of Lemma 5 as well in its proof.

We make use of the same $\tilde{h}'(X_t)$ and $\tilde{h}(J_t)$ and demonstrate that despite the presence of transient states for $\{J_t\}$ whose transition probabilities cannot approximate the probabilities of $\{X_t\}$ transitioning from the corresponding regions, Lemma 5 implies $\tilde{h}'(X_t)$ and $\tilde{h}(J_t)$ are guaranteed to be close in expectation on the 'important' regions of the space described in Lemma 3 when averaged over a sufficiently long time and the process $\{X_t\}$ is large. Once we have this, we will use these functions $\tilde{h}'(\cdot), \tilde{h}(\cdot)$ to define the piecewise constants in our test function.

Lemma 6. Suppose the conditions given in (A2) and in Lemmas 1-3 and 5 hold for arbitrary $\varepsilon > 0$ and some $r > 0$. Define $h'(x) = \sum_{j=1}^{l} h(j) I(x \in R_j)$. For a fixed $n < \infty$ let

$$\tilde{h}'(x) = \sum_{t=0}^{n} \frac{n-t}{n} E(h'(X_t)|X_0 = x,), \tag{3.60a}$$

$$\tilde{h}(j) = \sum_{t=0}^{n} \frac{n-t}{n} E(h(J_t)|J_0 = i). \tag{3.60b}$$

Then for arbitrary $\delta' > 0$ there exist $D, n < \infty$ so that

$$\sup_i \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \left| E(\tilde{h}(J_1) - \tilde{h}(i)|J_0 = i) - E(\tilde{h}'(X_1) - \tilde{h}'(x)|X_0 = x) \right| < \delta'. \tag{3.61}$$

Proof. Let $N = \max_j |h(j)|$. Get $\{J_t\}$ and t^* from (A2). Given $\delta' > 0$ pick $\varepsilon_1 > 0$ so that $Nt^*\varepsilon_1 < \delta'$. Given $\varepsilon_1 > 0$ get $M_1 < \infty$ from (A2) and t' from Lemma 5. Get n_1 from Lemma 1 and pick $n \geq n_1$ so that $Nt^*\varepsilon_1 + \frac{2N(t'-1)}{n} < \delta'$. Pick $\varepsilon > 0$ so that $(n+1)l^2N\varepsilon + Nt^*\varepsilon_1 + \frac{2N(t'-1)}{n} < \delta'$ and get $M \geq M_1$ from (A2). Pick $\delta_3 > 0$ so that

$$lN((n+1)l\varepsilon + n\delta_3) + Nt^*\varepsilon_1 + Nt'\delta_3 + \frac{2N(t'-1)}{n} < \delta'. \tag{3.62}$$

Get $D = D(n, \delta_3)$ from Lemma 3.

From (3.22) we have for all i, for all $x \in R_i$

$$E(\tilde{h}'(X_1) - \tilde{h}'(x)|X_0 = x) = \frac{1}{n}\sum_{t=1}^{n} E(h'(X_t)|X_0 = x) - h'(x), \tag{3.63a}$$

$$E(\tilde{h}(J_1) - \tilde{h}(i)|J_0 = i) = \frac{1}{n}\sum_{t=1}^{n} E(h(J_t)|J_0 = i) - h(i). \tag{3.63b}$$

Note that for all i, for all $x \in R_i$

$$\frac{1}{n}\sum_{t=1}^{t'-1} E\left(h'(X_t)\middle|X_0 = x\right) \le \frac{N(t'-1)}{n}, \tag{3.64a}$$

$$\frac{1}{n}\sum_{t=1}^{t'-1} E\left(h(J_t)\middle|J_0 = i\right) \le \frac{N(t'-1)}{n}. \tag{3.64b}$$

Define $I_t = \cap_{i=1}^{t}\{(\| X_i \| > M) \cap [\cap_{j=0}^{i-1}(X_j \in \cup_{i=1}^{l} R_i(\delta_3))]\}$. Under (A2) and given $\varepsilon_1 > 0$, $\delta_3 > 0$ we have from Lemma 5

$$\begin{aligned}
&\sup_i \sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} \frac{1}{n}\sum_{t=t'}^{n} E\left(h'(X_t) I\{(X_{t'} \in \cup_{k\in T} R_k) \cap I_{t'-1}\}\middle|X_0 = x\right)\\
\le\ & \sup_i \sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} N \times E\left(I\{(X_{t'} \in \cup_{k\in T} R_k) \cap I_{t'-1}\}\middle|X_0 = x\right)\\
<\ & N t^{*} \varepsilon_1.
\end{aligned} \tag{3.65}$$

Applying the Markov property and from Lemma 4 (3.21) by an argument similar to that leading to (3.24) if $X_{t'} \in \cup_{k\in G} R_k$ we have for all i

$$\begin{aligned}
&\sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} \frac{1}{n}\sum_{t=t'}^{n} E\left(h'(X_t) I\{(X_{t'} \in \cup_{k\in G} R_k) \cap I_{t'-1}\}\middle|X_0 = x\right)\\
\le\ & \sup_{\substack{x\in R_i(\delta_3)\\ \|x\|>DM}} \frac{1}{n}\sum_{t=t'}^{n} E\left(E(h'(X_t)|X_{t'}) I\{(X_{t'} \in \cup_{k\in G} R_k) \cap I_{t'-1}\}\middle|X_0 = x\right)\\
<\ & \frac{1}{n}\sum_{t=t'}^{n} E\left(h(J_t) I\{J_{t'} \in G\}\middle|J_0 = i\right) + lN((n+1)l\varepsilon + n\delta_3)\\
\le\ & \frac{1}{n}\sum_{t=t'}^{n} E\left(h(J_t)\middle|J_0 = i\right) + lN((n+1)l\varepsilon + n\delta_3).
\end{aligned} \tag{3.66}$$

By Lemma 3 (3.8) with $n = t^{'}$

$$\sup_{i} \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \frac{1}{n} \sum_{t=t^{'}}^{n} E\left(h^{'}(X_t) I\{I^c_{t^{'}-1}\} \middle| X_0 = x\right) \leq N t^{'} \delta_3. \tag{3.67}$$

Then from (3.64a), (3.64b), (3.65), (3.66), and (3.67) we have for all i

$$\begin{aligned} & \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \frac{1}{n} \sum_{t=1}^{n} E(h^{'}(X_t)|X_0 = x) \\ < \quad & \frac{1}{n} \sum_{t=1}^{n} E\left(h(J_t) \middle| J_0 = i\right) + lN((n+1)l\varepsilon + n\delta_3) + Nt^{*}\varepsilon_1 + \frac{2N(t^{'}-1)}{n} + Nt^{'}\delta_3. \end{aligned} \tag{3.68}$$

By similar arguments we have for all i

$$\begin{aligned} & \inf_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \frac{1}{n} \sum_{t=1}^{n} E(h^{'}(X_t)|X_0 = x) \\ > \quad & \frac{1}{n} \sum_{t=1}^{n} E\left(h(J_t) \middle| J_0 = i\right) - lN((n+1)l\varepsilon + n\delta_3) - Nt^{*}\varepsilon_1 - \frac{2N(t^{'}-1)}{n} - Nt^{'}\delta_3 \end{aligned} \tag{3.69}$$

or from (3.62), (3.63), (3.68) and (3.69) we have the conclusion

$$\begin{aligned} & \sup_{i} \sup_{\substack{x \in R_i(\delta_3) \\ \|x\| > DM}} \left| \frac{1}{n} \sum_{t=1}^{n} E(h^{'}(X_t)|X_0 = x) - \frac{1}{n} \sum_{t=1}^{n} E(h(J_t)|J_0 = i) \right| \\ < \quad & lN((n+1)l\varepsilon + n\delta_3) + Nt^{*}\varepsilon_1 + Nt^{'}\delta_3 + \frac{2N(t^{'}-1)}{n} < \delta^{'}. \end{aligned} \tag{3.70}$$

□

Theorem 2 handles the case of (A2). The proof was complicated by the fact we must wait a finite time for the processes $\{X_t\}$ and $\{J_t\}$ to get to the recurrent regions/states, requiring that we wait longer but still a finite time for the expectations of $\tilde{h}(J_t)$ and $\tilde{h}^{'}(X_t)$ to be sufficiently close with arbitrarily high probability. These issues were handled in Lemmas 5 and 6.

Theorem 2. Suppose the assumptions in (A2), Lemmas 1-3, 5 and 6 hold. Suppose as well that for arbitrary $\delta_3, \delta_4 > 0$ we can find $M < \infty$ such that $R_i(\delta_3)$ for $i = 1, \ldots, l$ as defined in Lemma 4 exist with

$$\sup_{i} \sup_{\substack{x\in R_i \\ \|x\|>M}} P(X_1 \in \cup_{i=1}^{l} R_i(\delta_3)|X_0 = x) < \delta_4. \tag{3.71}$$

Then under (A2) there exists an $s > 0$ such that $\{X_t\}$ is V-uniformly ergodic with test function $V(s,x) = 1+\lambda(x) \parallel x \parallel^s$, where $\lambda(x)$ is piecewise constant.

Proof. The complication created by the existence of a finite number of transient regions where ε-approximation of the 'transition' probabilities was not feasible was handled in Lemmas 5 and 6, so that we have regardless

$$\sup_{i} \sup_{\substack{x\in R_i(\delta_3) \\ \|x\|>DM}} \left| E(\tilde{h}(J_1) - \tilde{h}(i)|J_0 = i) - E(\tilde{h}'(X_1) - \tilde{h}'(x)|X_0 = x) \right| < \delta' \tag{3.72}$$

The remainder of the proof is the same as that for Theorem 1. □

In this chapter, we have demonstrated that under certain conditions on the general state space chain $\{X_t\}$ we can approximate its movements by those of a finite state chain $\{J_t\}$ and derive a condition for V-uniform ergodicity of $\{X_t\}$ through analysis of the more tractable chain $\{J_t\}$.

CHAPTER IV

EXAMPLES

4.1 Multi-Cyclic

4.1.1 Implications and method

The heuristic behind the work in Chapter II on V-uniform ergodicity of $\{X_t\}$ is the following:

1. Comparatively, the errors ξ_t become smaller in magnitude and less significant as $\| x \|_V$ increases.

2. Under certain conditions on the skeleton the eventual behavior of the process when the process is large mirrors that of the deterministic skeleton due to the observation in (1) above.

3. Thus, conditions for ergodicity of $\{X_t\}$ in this situation can be derived from the conditions for stability of the skeleton.

4. In particular, if the skeleton contains cycle(s) then the condition for ergodicity of $\{X_t\}$ is that the product(s) of companion matrices corresponding to regions in the cycle(s) have eigenvalue of maximum modulus smaller than 1.

In Chapter II we summarized this in a set of assumptions and verified in Theorems 1 and 2 the conditions do in fact establish V-uniform ergodicity of $\{X_t\}$. The results of Chapter II provide us with an algorithm for addressing the question of ergodicity of a threshold autoregressive time series:

1. Verify the assumptions on ξ_t.

2. Embed the time series $\{y_t\}$ of order p in a general state space Markov chain $\{X_t\}$ on $\mathbb{R}^p$.

3. Identify the skeleton of $\{X_t\}$. Label regions in $\mathbb{R}^p$ according to the companion matrix that applies in each.

4. Analyze the dynamics of the skeleton of $\{X_t\}$. Determine which regions are mapped to which. If necessary, subdivide the regions further so that entire regions are mapped to entire regions when $\| x \|$ is large. Suppose the regions are $R_1, \ldots, R_l$.

5. Identify regions that comprise the cycle(s) for the skeleton of $\{X_t\}$ and verify that those not in the cycle(s) are mapped to the cycle(s) along a d-path $R_0 \rightarrow R_1 \rightarrow R_d$ for some finite d.

6. We want points x in the cycle(s) to be mapped bounded away from the thresholds to the interior of the next region in the cycle(s) as in (2.50). Where necessary to make this so, cut out tiny cones from the regions and label them $R_i^{'}$, the i referring to the region in question. Verify these regions $R_i^{'}$ are 'transient' as in (2.52).

7. We want points not in the cycle to either be mapped into the interior of the next region in the d-path as in (2.51) or to be transient as in (2.52). Cut out small cones from these regions not in the cycle in order to make (2.51) true. Call these small cones $R_i^{'}$ as well, the i referring to the region in question, and verify these regions $R_i^{'}$ are transient as in (2.52).

8. Do so keeping in mind that for points x, we require for an arbitrary $M_4 < \infty$ we can pick $\| x \| > M^{''}$ large enough so that $S = \{x : \| x \| > M^{''}, \| (\pi_{i=1}^{p(x)} A_i) x \| \leq M_4\}$ is contained in $\cup_{i=1}^{l} R_i^{'}$. The set S will be an issue where one or more of the matrices in the cycle(s) is not of full rank. Be certain S is included in $\cup_{i=1}^{l} R_i^{'}$.

9. Then by the appropriate theorem, the condition for V-uniform ergodicity of $\{X_t\}$ is that the eigenvalue(s) of maximum modulus of the product(s) of companion matrices that comprise the cycle(s) be less than 1.

4.1.2 Example

As an example of the multi-cyclic methods consider the TAR(2;1;1) model

$$y_t = \begin{cases} a_1 y_{t-1} + a_2 y_{t-2} + \xi_t, & y_{t-1} \geq d, y_{t-2} \geq 0 \\ b_1 y_{t-1} + \xi_t, & y_{t-1} < d \\ c_1 y_{t-1} + \xi_t, & y_{t-1} \geq d, y_{t-2} < 0 \end{cases} \tag{4.1}$$

Suppose $\xi_t \sim N(0, \sigma^2)$. Then since $\sigma^2 < \infty$ we have $E|\xi_t|^2 < \infty$. We will analyze the case $a_1 > 0$, $a_2 > 0$, $b_1 < 0$, $c_1 < 0$, $d > 0$. Embed y_t in a Markov chain by writing:

$$X_t = (y_t, y_{t-1})', \quad \nu_t = (\xi_t, 0)' \tag{4.2}$$

and define the companion matrices by

$$A = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & 0 \\ 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} c_1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{4.3}$$

Then

$$X_t = AX_{t-1}I(y_{t-1} \geq d, y_{t-2} \geq 0) + BX_{t-1}I(y_{t-1} < d) + CX_{t-1}I(y_{t-1} \geq d, y_{t-2} < 0) + \nu_t \tag{4.4}$$

is the general state Markov chain on $\mathbb{R}^2$, which we will think of as the (y_{t-1}, y_{t-2})-plane. From (4.4) the skeleton of $\{X_t\}$ is thus

$$X_t = AX_{t-1}I(y_{t-1} \geq d, y_{t-2} \geq 0) + BX_{t-1}I(y_{t-1} < d) + CX_{t-1}I(y_{t-1} \geq d, y_{t-2} < 0) \tag{4.5}$$

Define the following regions:

$$R_1 = \{(y_{t-1}, y_{t-2}) : y_{t-1} \geq d,\ y_{t-2} \geq 0\} \tag{4.6a}$$

$$R_2 = \{(y_{t-1}, y_{t-2}) : y_{t-1} < d\} \tag{4.6b}$$

$$R_3 = \{(y_{t-1}, y_{t-2}) : y_{t-1} \geq d,\ y_{t-2} < 0\} \tag{4.6c}$$

A depiction of the partition of $\mathbb{R}^2$ into these regions and the companion matrix that applies in each can be seen in Figure 1.

Let $R_i \to R_j$ denote $x_{t-1} \in R_i \Rightarrow x_t \in R_j$ is dictated by the skeleton (4.5). The dynamics for the skeleton are:

$$R_1 \to R_1,\ R_2 \to R_2,\ R_2 \to R_3,\ R_3 \to R_2.$$

The region R_2 feeds into two different regions. This is a problem since our results require each region have a unique successor region. Consider this further. Note that points $(y_{t-1}, y_{t-2})' \in R_2 \cup R_3$ with small $|y_{t-1}|$ remain small:

1. Suppose $x = (y_{t-1}, y_{t-2})'$ such that $y_{t-1} < 0$; then since $Bx = y_{t-1}(b_1, 1)'$ we have that Bx remains in R_2 if $y_{t-1} \geq d/b_1$.

2. Suppose $x = (y_{t-1}, y_{t-2})'$ such that $0 < y_{t-1} < d$; then by (1) above, Bx remains in the $R_2 \to R_2$ cycle if $y_{t-1} \leq d/b_1^2$.

3. Suppose $x = (y_{t-1}, y_{t-2})'$ such that $y_{t-1} \geq d$; then since $Cx = y_{t-1}(c_1, 1)'$ we have that Cx maps to the $R_2 \to R_2$ cycle if $y_{t-1} \leq d/(c_1 b_1)$.

There are several cases here depending upon the values of b_1, c_1. The case $b_1 c_1 > 1$, $b_1 < 1$ is depicted in Figure 2.

Now for some $M_4 < \infty$, the requirement $\| BCx \| \leq M_4$ implies $\| c_1 y_{t-1}(b_1, 1)' \| \leq M_4$ or $|y_{t-1}| \leq M_4/(c_1^2(b_1^2 + 1))$. Referring to the observations in (1)-(3) above, by picking M_4 large enough we can cover the entire region involved in the $R_2 \to R_2$ cycle by

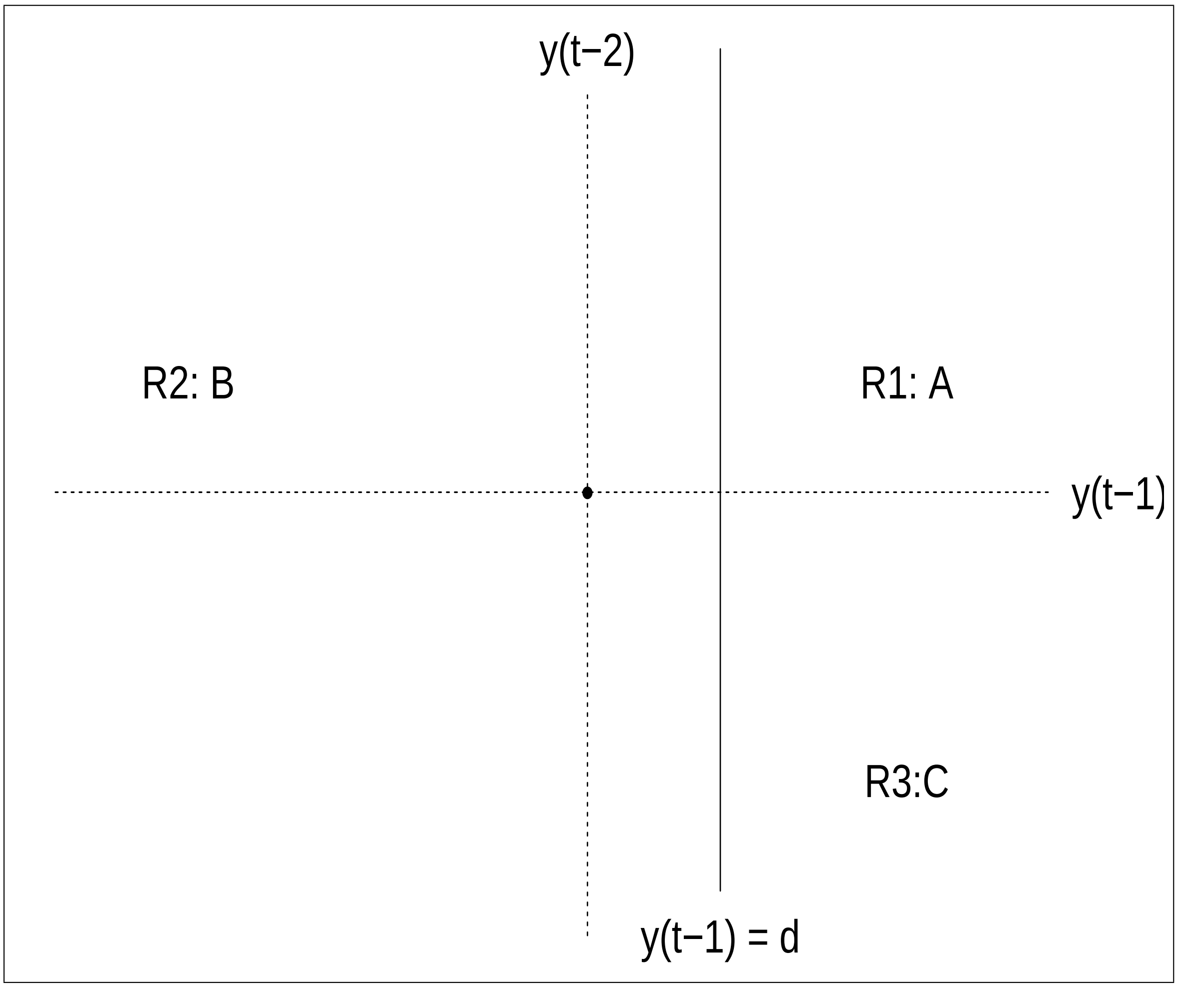

Figure 1. Regions for the TAR(2;1;1) example and their companion matrices.

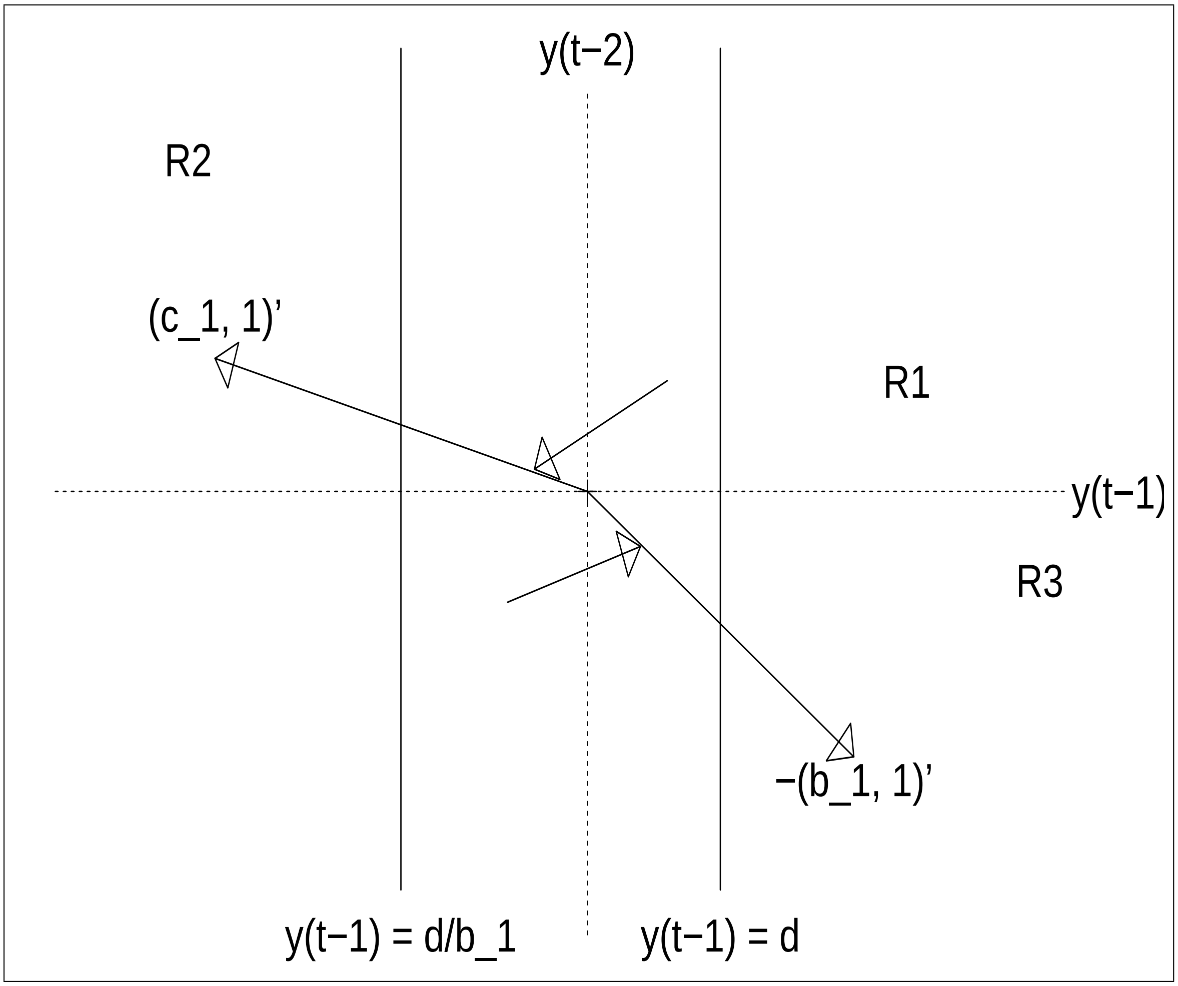

Figure 2. Middle regions for the TAR(2;1;1) example and their dynamics.

$\{x = (y_{t-1}, y_{t-2})^{'} : \parallel BCx \parallel \le M_4\}$, which defines a strip around the y_{t-2}-axis. Define

$$R_2(M_4) = \{(y_{t-1}, y_{t-2})^{'} \in R_2 : |y_{t-1}| \le M_4/(c_1^2(b_1^2+1))\}, \tag{4.7a}$$

$$R_3(M_4) = \{(y_{t-1}, y_{t-2})^{'} \in R_3 : |y_{t-1}| \le M_4/(c_1^2(b_1^2+1))\}. \tag{4.7b}$$

We are going to see that we can ignore these regions $R_2(M_4)$ and $R_3(M_4)$. The important cycles will be $R_1 \to R_1$ and $R_2 \to R_3 \to R_2$. Since we have two cycles we will apply Theorem 2 in Chapter II.

We want to define regions $R_1^{'}$, $R_2^{'}$ and $R_3^{'}$ so that conditions (2.50) and (2.52) are satisfied. Slicing out cones near the y_{t-2}-axis will ensure that points outside of these cones will be mapped bounded away from the thresholds into the interior of the next region in the cycle according to (2.50). The form of the companion matrices B and C dictate that $|y_{t-1}| \to \infty$ as $\parallel x \parallel$ does; we need this to be true for points outside of these cones. We will update our regions R_1, R_2 and R_3 to exclude the cones $R_1^{'}$, $R_2^{'}$ and $R_3^{'}$.

For $\delta > 0$ we define the regions $R_1^{'}$, $R_2^{'}$ and $R_3^{'}$ as:

$$R_1^{'} = \{(y_{t-1}, y_{t-2})^{'} : y_{t-2} \ge y_{t-1}/\delta + d,\ y_{t-1} \ge d\} \tag{4.8a}$$

$$R_2^{'} = \{(y_{t-1}, y_{t-2})^{'} : y_{t-2} \ge y_{t-1}/\delta,\ 0 < y_{t-1} < d\} \tag{4.8b}$$

$$\cup \{(y_{t-1}, y_{t-2})^{'} : y_{t-2} \le -y_{t-1}/\delta,\ 0 < y_{t-1} < d\} \tag{4.8c}$$

$$\cup \{(y_{t-1}, y_{t-2})^{'} : y_{t-2} \le y_{t-1}/\delta,\ y_{t-1} \le 0\} \tag{4.8d}$$

$$\cup \{(y_{t-1}, y_{t-2})^{'} : y_{t-2} \ge -y_{t-1}/\delta,\ y_{t-1} \le 0\} \tag{4.8e}$$

$$R_3^{'} = \{(y_{t-1}, y_{t-2})^{'} : y_{t-2} \le -y_{t-1}/\delta + d,\ y_{t-1} \ge d\} \tag{4.8f}$$

A depiction of these regions is in Figure 3.

Now for an arbitrary $\delta^{'} > 0$ we can clearly pick $\delta > 0$ small enough and $M^{'} < \infty$ so that

$$\sup_{\substack{x \\ \|x\| > M^{'}}} P(X_1 \in R_1^{'} \cup R_2^{'} \cup R_3^{'} | X_0 = x) < \delta^{'} \tag{4.9}$$

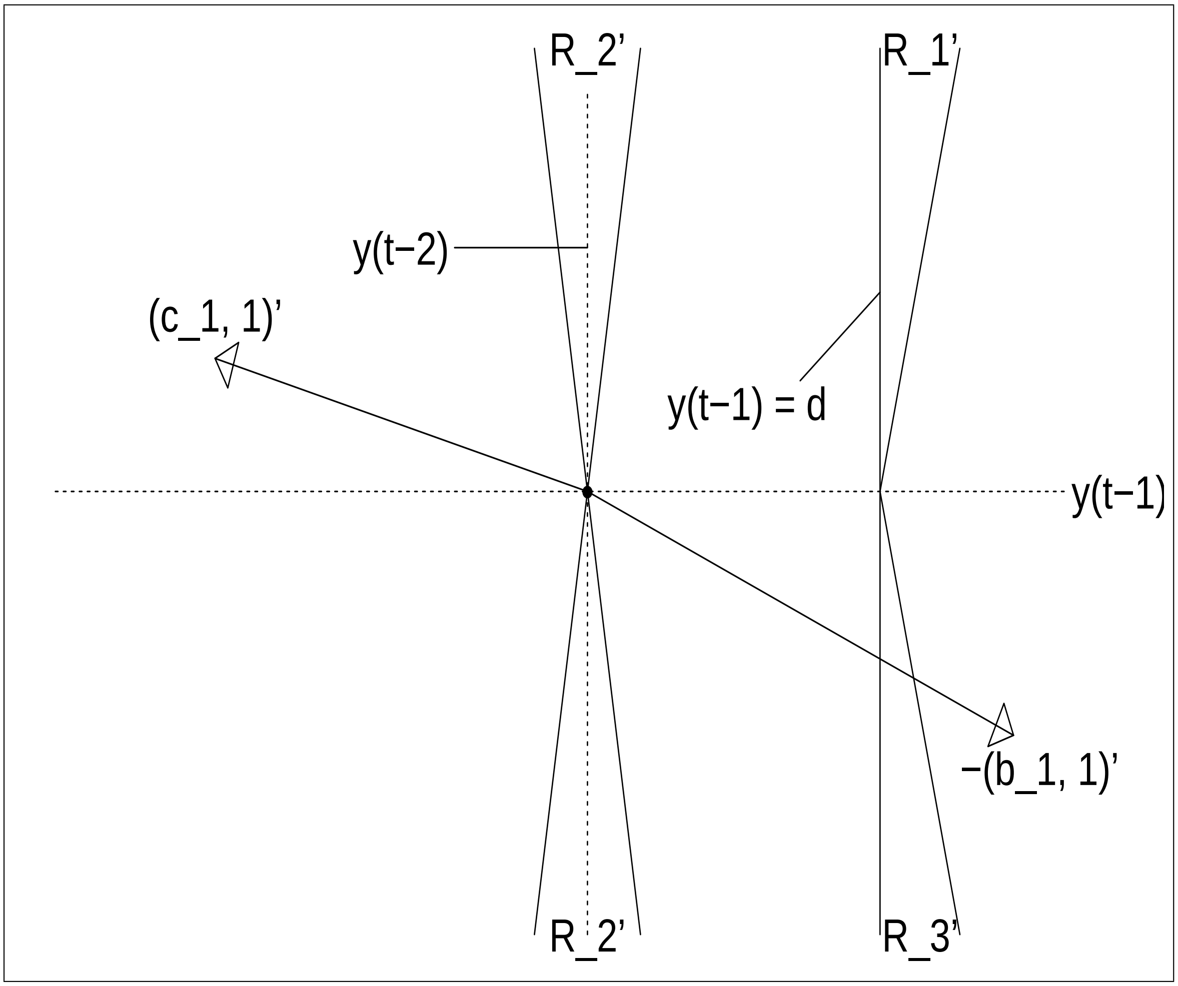

Figure 3. New partition of regions for the TAR(2;1;1) example.

and we have that (2.52) will be satisfied.

We need to verify for some $M < \infty$, that for all $x = (y_{t-1}, y_{t-2})' \in R_1$ with $\| x \| > M$ there exists a positive, strictly increasing function $f_1(\cdot)$ so that $B_{f_1(\|x\|)}(Ax) \subset R_1$. The matrix A maps R_1 into itself. The worst cases occur along the boundaries of R_1, where $(y_{t-1}, y_{t-2})'$ is such that either $y_{t-2} = y_{t-1}/\delta$ or $y_{t-2} = 0$. Since

$$A \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} = y_{t-1} \begin{pmatrix} a_1 + a_2/\delta \\ 1 \end{pmatrix}, \quad A \begin{pmatrix} y_{t-1} \\ 0 \end{pmatrix} = y_{t-1} \begin{pmatrix} a_1 \\ 1 \end{pmatrix} \tag{4.10}$$

then points $(y_{t-1}, y_{t-2})'$ such that either $y_{t-2} = y_{t-1}/\delta$ or $y_{t-2} = 0$ are mapped to the interior of R_1 along the rays $y_{t-1}(a_1 + a_2/\delta, 1)'$ and $y_{t-1}(a_1, 1)'$ respectively. The distance from these rays to the boundary of R_1 increases as $\| x \|$ does, implying there exists a positive, strictly increasing function $f_1(\cdot)$ such that $B_{f_1(\|x\|)}(Ax) \subset R_1$ and this $f_1(\cdot)$ will work for all $(y_{t-1}, y_{t-2})' \in R_1$ regardless of the value of M.

Next we need to verify for some $M < \infty$, for all $x = (y_{t-1}, y_{t-2})' \in R_2$ with $\| x \| > M$ there exists a positive, strictly increasing function $f_2(\cdot)$ so that $B_{f_2(\|x\|)}(Bx) \subset R_3$. Since for all $x = (y_{t-1}, y_{t-2})' \in R_2$ we have $Bx = y_{t-1}(b_1, 1)'$ the worst case for $(y_{t-1}, y_{t-2})'$ with $\| (y_{t-1}, y_{t-2})' \| = M$ occurs at the infimum of $|y_{t-1}|$ in R_2 which occurs along the rays $y_{t-2} = y_{t-1}/\delta$ and $y_{t-2} = -y_{t-1}/\delta$. The distance from $y_{t-1}(b_1, 1)'$ to the rays $y_{t-2} = y_{t-1}/\delta$ and $y_{t-2} = -y_{t-1}/\delta$ increases as $|y_{t-1}|$ and thus $\| x \|$ does, implying there exists a positive, strictly increasing function $f_2(\cdot)$ that satisfies our requirement. Whichever $f_2(\cdot)$ works along these rays will work for all $(y_{t-1}, y_{t-2})' \in R_2$.

Finally, we need to verify for some $M < \infty$ that for all $x = (y_{t-1}, y_{t-2})' \in R_3$ with $\| x \| > M$ there exists a positive, strictly increasing function $f_3(\cdot)$ so that $B_{f_3(\|x\|)}(Cx) \subset R_2$. The argument is similar to that for R_2. For all $x = (y_{t-1}, y'_{t-2} \in R_3$ we have $Cx = y_{t-1}(c_1, 1)'$, so the worst case for $(y_{t-1}, y_{t-2})'$ with $\| (y_{t-1}, y_{t-2})') \| = M$ occurs at the infimum of $|y_{t-1}|$ in R_3 which occurs along the ray $y_{t-2} = -y_{t-1}/\delta$. Whichever $f_3(\cdot)$ works along this ray will work for all $(y_{t-1}, y_{t-2})' \in R_3$.

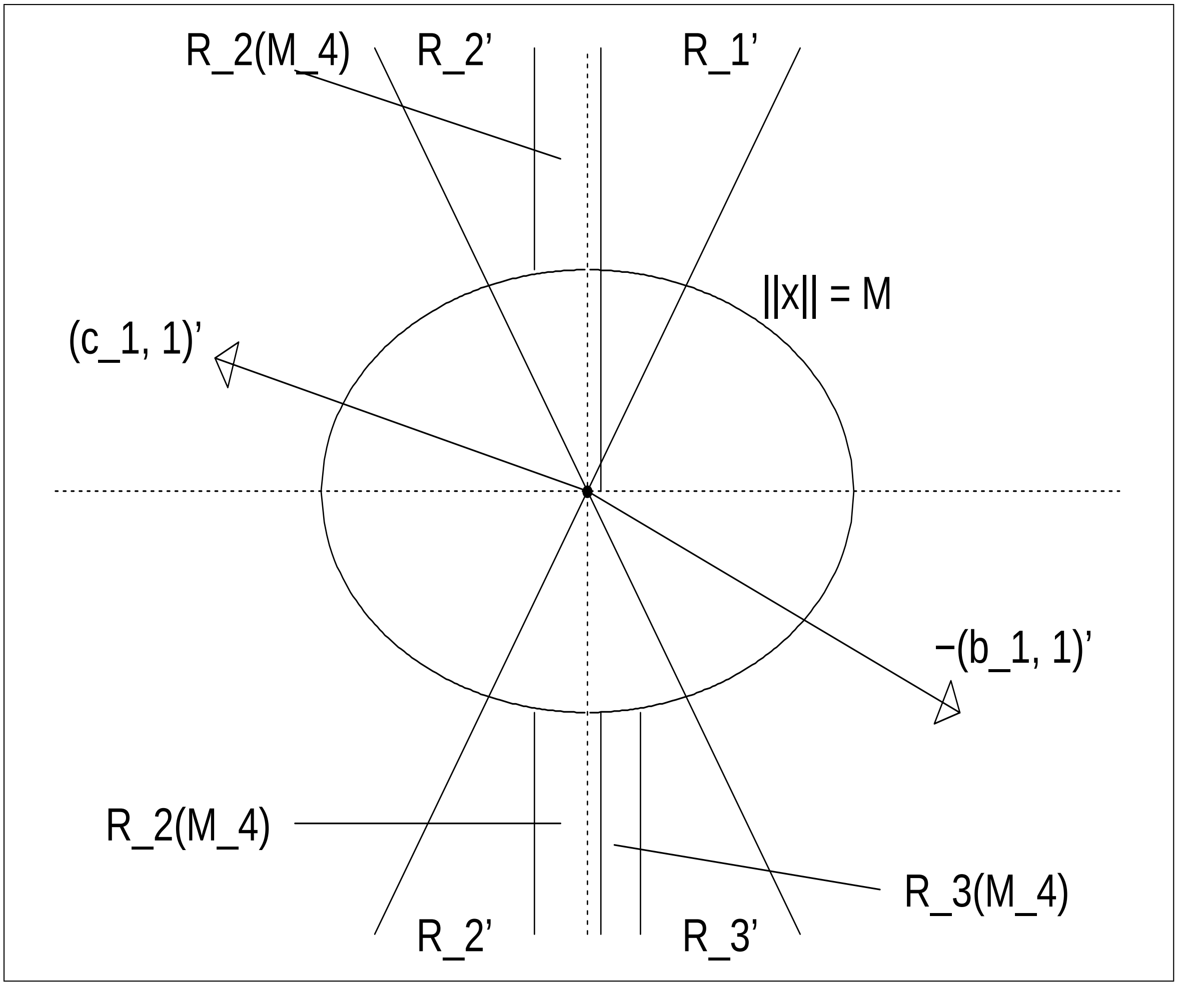

Figure 4. Regions for the TAR(2;1;1) example with large $\| x \|$.

As for R_1', R_2', R_3', these regions are not in the cycle and such uniform bounds do not exist. By the definition of these regions in (4.9) we have

$$\sup_{\substack{x \\ \|x\|>M}} P(X_1 \in R_1' \cup R_2' \cup R_3' | X_0 = x) < \delta' \tag{4.11}$$

and for large enough $\| x \|$ the region $\{x = (y_{t-1}, y_{t-2})' : \| BCx \| \leq M_4\}$ is contained in $R_1' \cup R_2' \cup R_3'$. A depiction of this is in Figure 4.

We have satisfied all the assumptions behind Theorem 2:

1. $\xi_t \sim N(0,\sigma^2)$ which has a continuous density everywhere positive and since $\sigma^2 < \infty$ we have $E|\xi_t|^2 < \infty$.

2. The regions R_1, R_2, R_3 comprising the cycles have the requisite bounds as in (2.50).

3. The regions $R_1^{'}, R_2^{'}, R_3^{'}$ not in the cycle do not have a bound but do satisfy (2.52) and do contain the sets $\{x = (y_{t-1}, y_{t-2})^{'} : \parallel x \parallel_v > M^{''}, \parallel BCx \parallel_v \leq M_4\}$ and $\{x = (y_{t-1}, y_{t-2})^{'} : \parallel x \parallel_v > M^{''}, \parallel Ax \parallel_v \leq M_4\}$ (which is the empty set for $M^{''}$ large enough: since A is full rank, the set of $x = (y_{t-1}, y_{t-2})^{'}$ such that $\parallel Ax \parallel \leq M_4$ is a bounded set).

4. All $(y_{t-1}, y_{t-2})^{'}$ are mapped by the skeleton to $R_1 \cup R_2 \cup R_3$ in a finite time.

Then by Theorem 2 in Chapter II, for the model (4.1) in the case where $a_1 > 0$, $a_2 > 0$, $b_1 < 0$ and $c_1 < 0$ if

$$\rho(A) < 1 \Leftrightarrow a_1 + a_2 < 1, \quad \rho(BC) < 1 \Leftrightarrow b_1 c_1 < 1$$

then $\{X_t\}$ is V-uniformly ergodic.

Note the ergodic parameter space is unbounded, contrary to what we would expect through analogy with the case of a linear time series:

$$\rho(A) < 1 \Leftrightarrow a_1 + a_2 < 1, \quad \rho(B) < 1 \Leftrightarrow -1 < b_1 < 0, \quad \rho(C) < 1 \Leftrightarrow -1 < c_1 < 0,$$

illustrating the point made in Chapter II.

4.2 Finite State Chain Approximation

4.2.1 Implications and method

The heuristic behind the work in Chapter III on V-uniform ergodicity of $\{X_t\}$ is the following:

1. $\sum_{i=1}^{l} \log(\| A_i \|) I(X_{t-1} \in R_i)$ describes the log-change in the process as the process moves from X_{t-1} to X_t. The expected log-change in the process as it moves from X_{t-1} to X_t is given by

$$E_{X_{t-2}}\left(\sum_{i=1}^{l} \log(\| A_i \|) I(X_{t-1} \in R_i)\right). \tag{4.12}$$

2. Under the assumptions in either (A1) or (A2), when $\| x \|$ is large the expected log-change (4.12) considered over $R_1, \ldots, R_l$ from X_{t-1} to X_t when $X_{t-2} = x \in R_i$, is close to $E_{J_{t-1}}(\log(\| A_{J_t} \|))$. By ergodicity of $\{J_t\}$, $E_{J_{t-1}}(\log(\| A_{J_t} \|))$ will converge to $E_\pi(\log(\| A_{J_t} \|))$, where π is the stationary distribution of $\{J_t\}$ and $E_\pi(\cdot)$ denotes the expectation with respect to π.

3. Thus, averaging (4.12) over a sufficiently long but finite time will make it arbitrarily close to $E_\pi(\log(\| A_{J_t} \|))$. The condition

$$E_\pi(\log(\| A_{J_t} \|)) = \sum_{i=1}^{l} \pi_i \log(\| A_i \|) < 0$$

for all stationary distributions π of J_t will guarantee V-uniform ergodicity of $\{X_t\}$.

This is the basic idea; complications were introduced according to whether the ε-approximation held for all states/regions or only for certain regions, and according to the amount of time it took for the ergodicity described in (2) to take effect. The details of all this was worked out in the lemmas and theorems of Chapter III.

The results of Chapter III provide us with an algorithm for addressing the question of ergodicity of a threshold autoregressive nonlinear time series:

1. Verify the assumption $E|\xi_t|^r < \infty$.

2. Embed the time series $\{y_t\}$ of order p in a general state Markov chain $\{X_t\}$ on $\mathbb{R}^p$.

3. Identify the skeleton. Analyze its dynamics and determine distinct regions so that entire regions are mapped to entire regions. Suppose these regions are $R_1, \ldots, R_l$.

4. Consider the errors. Cut tiny cones out near the region boundaries and find $M < \infty$ such that $\| x \| > M$ implies we can either ε-bound the transition probabilities between regions $P(X_t \in R_j | X_{t-1} = x \in R_i)$ as specified in (A1) and (A2) or that the regions R_i where the transition probabilities cannot be bounded by ε are transient.

5. Construct the finite state chain J_t and verify the appropriate assumptions on it.

6. Find the stationary distributions π of J_t and derive the condition for ergodicity.

4.2.2 Example

As an example of the application of the finite state chain approximation methods consider the TAR(2;1) model

$$y_t = \begin{cases} a_1 y_{t-1} + a_2 y_{t-2} + \xi_t, & y_{t-2} \geq \frac{1}{b_1} y_{t-1} \\ \\ b_1 y_{t-1} + \xi_t, & y_{t-1} < \frac{1}{b_1} y_{t-1} \end{cases} \tag{4.13}$$

Assume $\xi_t \sim N(0, \sigma^2)$. Since $\sigma^2 < \infty$ we have $E|\xi_t|^2 < \infty$ and the assumption on ξ_t is satisfied.

There are several cases to consider. We are going to suppose $b_1 < 0$, $b_2 = 0$, $a_1, a_2 > 0$, $a_1 b_1 + a_2 > 0$ and $b_1^2 > a_1 b_1 + a_2$.

Embed y_t in a Markov chain by writing:

$$X_t = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}, \quad \nu_t = \begin{pmatrix} \xi_t \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{4.14}$$

Then

$$X_t = AX_{t-1}I(y_{t-2} \geq (1/b_1)y_{t-1}) + BX_{t-1}I(y_{t-2} < (1/b_1)y_{t-1}) + \nu_t \tag{4.15}$$

is the general state Markov chain on $\mathbb{R}^2$. Note the threshold $y_{t-2} = \frac{1}{b_1}y_{t-1}$ is the eigenvector of the B matrix.

Let $R_A = \{(y_{t-1}, y_{t-2}) : y_{t-2} \geq (1/b_1)y_{t-1}\}$, $R_B = \{(y_{t-1}, y_{t-2}) : y_{t-2} < (1/b_1)y_{t-1}\}$. The first task is to analyze the skeleton of the process, which from (4.15) is

$$x_t = Ax_{t-1}I(x_{t-1} \in R_A) + Bx_{t-1}I(x_{t-1} \in R_B). \tag{4.16}$$

This defines a deterministic system in $\mathbb{R}^2$. Our task is to determine distinct regions dictated by the dynamics of the skeleton. A couple of things to note here:

1. $b_1^2 > a_1b_1 + a_2$ implies $\frac{b_1-a_1}{a_2} < \frac{1}{b_1}$, so the ray $y_{t-2} = \frac{b_1-a_1}{a_2}y_{t-1}$ lies above the threshold $y_{t-2} = \frac{1}{b_1}y_{t-1}$ when $y_{t-1} < 0$.

2. For (y_{t-1}, y_{t-2}) such that $y_{t-2} = \frac{1}{b_1}y_{t-1}$ with $y_{t-1} < 0$

$$A\begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} = \begin{pmatrix} a_1y_{t-1} + a_2y_{t-2} \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} a_1y_{t-1} + \frac{a_2}{b_1}y_{t-1} \\ y_{t-1} \end{pmatrix} = y_{t-1}\begin{pmatrix} a_1 + \frac{a_2}{b_1} \\ 1 \end{pmatrix}. \tag{4.17}$$

   Since $y_{t-1} < 0$ and $a_1b_1 + a_2 > 0$ imply $a_1 + a_2/b_1 < 0$, we have that (4.17) lies to the right of the y_{t-2}-axis.

With the help of these observations we can define the following 5 regions:

$$R_1 = \{(y_{t-1}, y_{t-2}) : y_{t-1} \geq 0,\ y_{t-2} \geq (1/b_1)y_{t-1}\} \tag{4.18a}$$

$$R_2 = \{(y_{t-1}, y_{t-2}) : y_{t-1} \geq 0,\ y_{t-2} < (1/b_1)y_{t-1}\} \tag{4.18b}$$

$$R_3 = \{(y_{t-1}, y_{t-2}) : y_{t-1} < 0,\ y_{t-2} < (1/b_1)y_{t-1}\} \tag{4.18c}$$

$$R_4 = \{(y_{t-1}, y_{t-2}) : y_{t-1} < 0,\ (1/b_1)y_{t-1} \leq y_{t-2} \leq ((b_1 - a_1)/(a_2))y_{t-1}\} \tag{4.18d}$$

$$R_5 = \{(y_{t-1}, y_{t-2}) : y_{t-1} < 0,\ ((b_1 - a_1)/(a_2))y_{t-1} < y_{t-2}\} \tag{4.18e}$$

A depiction of the partition of $\mathbb{R}^2$ into these regions can be seen in Figure 5. For brevity we denote the ray $y_{t-2} = ((b_1 - a_1)/a_2)y_{t-1}$ with $y_{t-1} < 0$ by L_1 and the ray $y_{t-1} = (1/b_1)y_{t-1}$ with $y_{t-1} \geq 0$ by L_2.

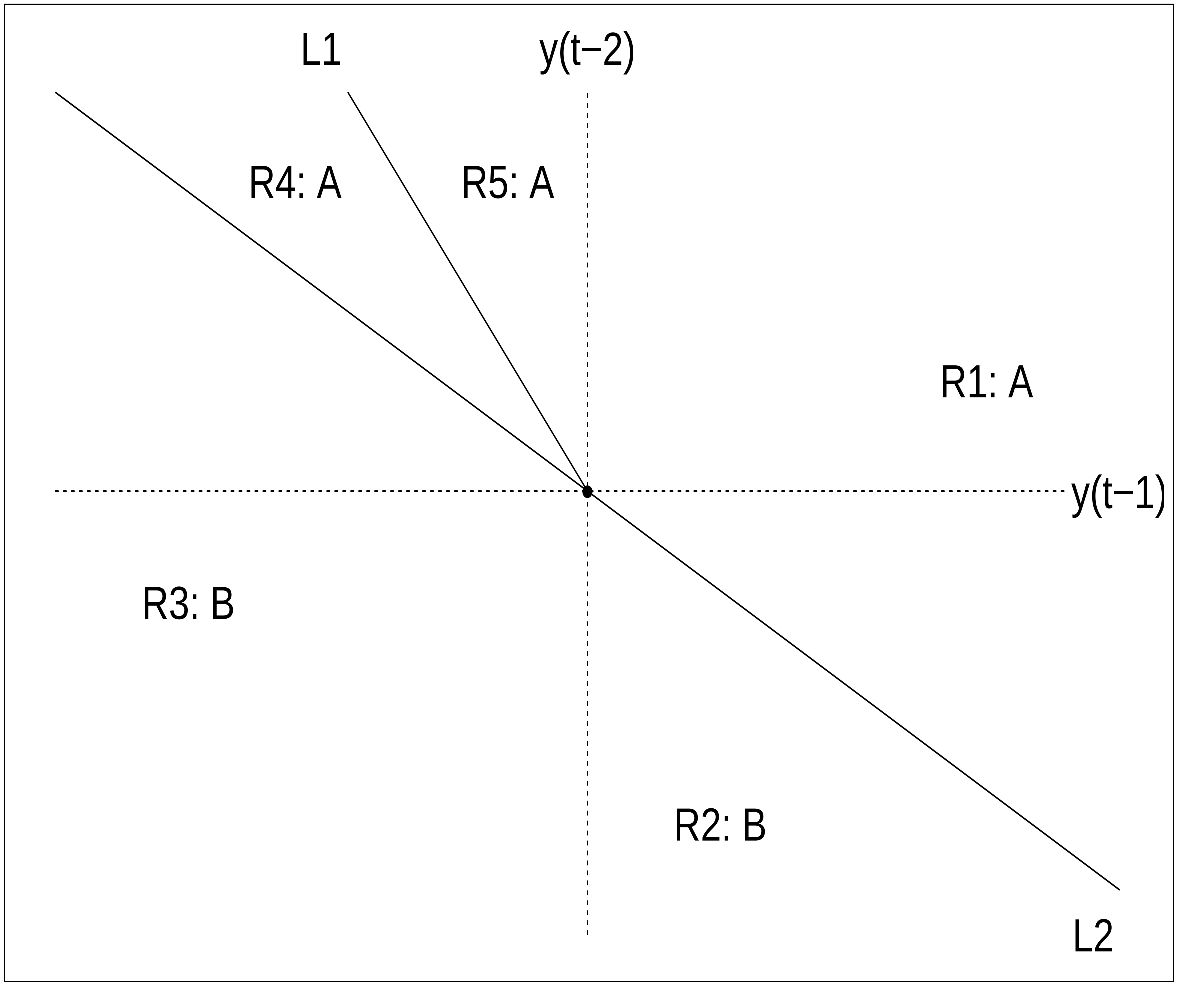

Figure 5. Regions for the TAR(2;1) example and their companion matrices.

Let $R_i \rightarrow R_j$ denote $x_{t-1} \in R_i \Rightarrow x_t \in R_j$ is dictated by (4.16) and denote by T the threshold, i.e., the set of all (y_{t-1}, y_{t-2}) such that $y_{t-2} = (1/b_1)y_{t-1}$. We have included T in $R_1 \cup R_4$ but it is useful to consider it separately for a moment. The dynamics for the skeleton are:

$$R_1 \rightarrow R_1,\ R_2 \rightarrow T,\ R_3 \rightarrow T,\ R_4 \rightarrow R_2,\ R_5 \rightarrow R_1.$$

When we add the errors ξ_t back in, several issues arise:

1. All $x \in R_2 \cup R_3$ are mapped by B to the threshold $y_{t-2} = \frac{1}{b_1} y_{t-1}$. When we add the errors back in, then regardless of the magnitude of $\| x \|$ the errors give $\{X_t\}$ an equal chance of being bumped off the threshold in either direction.

2. The ray $y_{t-2} = \frac{b_1 - a_1}{a_2} y_{t-1}$ is mapped to the threshold by A. When the errors are taken into account, these points have an equal chance of being bumped off the threshold in either direction.

3. All $x \in R_2 \cup R_3$ either on or near the y_{t-2}-axis are mapped by B near the origin and their transition probabilites vary depending upon how close they are to the y_{t-2}-axis, but do not depend on $\| x \|$ per se, only on $|y_{t-1}|$.

(1) and (2) imply it is necessary to model $\{X_t\}$ with a stochastic system rather than an asymptotically deterministic one. We will get to this in a moment but first let us address concern (3). Because no points are mapped near the y_{t-2}-axis by either A or B, we can cut out narrow cones around the y_{t-2}-axis, call them $R_2^{'}$ and $R_3^{'}$, and in effect throw them away, meaning we can construct $R_2^{'}$ and $R_3^{'}$ so that for an arbitrary $\varepsilon > 0$:

$$\sup_i \sup_{\substack{x \in R_i \\ \|x\| > M}} P(X_1 \in R_2^{'} | X_0 = x) < \varepsilon \tag{4.19a}$$

$$\sup_i \sup_{\substack{x \in R_i \\ \|x\| > M}} P(X_1 \in R_3^{'} | X_0 = x) < \varepsilon. \tag{4.19b}$$

This satisfies condition (3.71) of Theorem 2 in Chapter III. To see that we can do this, consider that the worst case, that is, the points mapped closest to $R_2 \cup R_3$, occurs for points that lie along the ray $y_{t-2} = \frac{1}{b_1} y_{t-1}$. We've already seen in (4.17) that these points are mapped to $y_{t-1}(a_1 + \frac{a_2}{b_1}, 1)^{'}$ by A. For an arbitrary $\delta > 0$ define the boundaries of $R_2^{'}$ to be the rays $(0, y_{t-2})$ and $(-\delta y_{t-2}, y_{t-2})$ for $y_{t-2} < 0$. Likewise, define the boundaries of $R_3^{'}$ to be the rays $(0, y_{t-2})$ and $(\delta y_{t-2}, y_{t-2})$ for $y_{t-2} < 0$.

For any $(y_{t-1}, y_{t-2})^{'}$ such that $y_{t-2} = \frac{1}{b_1} y_{t-1}$ and $y_{t-1} < 0$ we have from (4.17)

$$A \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} = y_{t-1} \begin{pmatrix} a_1 + \frac{a_2}{b_1} \\ 1 \end{pmatrix} := \begin{pmatrix} y_{t-1}^{'} \\ y_{t-2}^{'} \end{pmatrix} \tag{4.20}$$

where we denote the updated $(y_{t-1}, y_{t-2})^{'}$ by $(y_{t-1}^{'}, y_{t-2})^{'}$. Note that $y_{t-2}^{'} = y_{t-1}$. Consider these mapped points $(y_{t-1}^{'}, y_{t-2})^{'}$ versus the boundaries of $R_2^{'}$ and $R_3^{'}$, $(-\delta y_{t-2}^{'}, y_{t-2})^{'}$ and $(\delta y_{t-2}^{'}, y_{t-2})^{'}$, respectively. Since $\nu_t = (\xi_t, 0)^{'}$, the errors perturb the process horizontally; thus, to consider whether $X_t \in R_2 \cup R_3$ we need to consider the horizontal distance from $y_{t-1}^{'}$ to the interval $(-\delta y_{t-2}^{'}, \delta y_{t-2})^{'}$.

Since $y_{t-1} = y_{t-2}^{'}$ and $y_{t-1}^{'} = y_{t-1}(a_1 + a_2/b_1)$ we have that $X_t \in R_2^{'} \cup R_3^{'}$ if ξ_t is in the interval $(|y_{t-1}|(a_1 + a_2/b_1 - \delta),\ |y_{t-1}|(a_1 + a_2/b_1 + \delta))$. Now for $(y_{t-1}, y_{t-2})^{'}$ on the ray $y_{t-2} = \frac{1}{b_1} y_{t-1}$ and for any $M < \infty$

$$\| (y_{t-1}, y_{t-2})^{'} \| > M \Leftrightarrow \ldots \Leftrightarrow |y_{t-1}| > M \sqrt{\frac{b_1^2}{1 + b_1^2}}. \tag{4.21}$$

Thus for an arbitrary $\varepsilon > 0$ if we pick M_1 large enough so that

$$P\left(\xi_t \geq M_1 \sqrt{\frac{b_1^2}{b_1^2 + 1}} \left(a_1 + \frac{a_2}{b_1} - \delta\right)\right) < \varepsilon \tag{4.22}$$

and set $\| (y_{t-1}, y_{t-2})^{'} \| > M_1$ then by (4.21) and (4.22) we have

$$\begin{aligned} & P\left(\xi_t \in (|y_{t-1}|(a_1 + a_2/b_1 - \delta), |y_{t-1}|(a_1 + a_2/b_1 + \delta))\right) \\ \leq \quad & P\left(\xi_t \geq M_1 \sqrt{\frac{b_1^2}{b_1^2 + 1}} \left(a_1 + \frac{a_2}{b_1} - \delta\right)\right) < \varepsilon. \end{aligned} \tag{4.23}$$

For this M_1 we have

$$\sup_i \sup_{\substack{x \in R_i \\ \|x\| > M_1}} P(X_1 \in R_2^{'} | X_0 = x) < \varepsilon \tag{4.24a}$$

$$\sup_i \sup_{\substack{x \in R_i \\ \|x\| > M_1}} P(X_1 \in R_3^{'} | X_0 = x) < \varepsilon. \tag{4.24b}$$

Once $R_2^{'}$ and $R_3^{'}$ are established, we need to cut out a cone $R_5^{'}$ from R_5 so that for some $M_2 < \infty$

$$\inf_{\substack{x \in R_5 \\ \|x\| > M_2}} P(X_1 \in R_1 | X_0 = x) > 1 - \varepsilon \tag{4.25}$$

and a cone $R_4^{'}$ from R_4 so that for some $M_3 < \infty$

$$\sup_{\substack{x \in R_4 \\ \|x\| > M_3}} P(X_1 \notin R_2 | X_0 = x) < \varepsilon/2. \tag{4.26}$$

For large enough M_1, M_2 we can make these cones $R_4^{'}$ and $R_5^{'}$ very small. This serves two purposes: we can then bound all transition probabilities from R_4 and R_5 to within ε, and even though we cannot uniformly bound the transition probabilities in $R_4^{'}$ and $R_5^{'}$, these regions will be transient and satisfy the assumptions made in (A2).

Can we set up the cones $R_4^{'}$ and $R_5^{'}$ so that the desired conditions in (4.25) and (4.26) hold? Consider the case of R_5. For an arbitrary $\delta > 0$ cut the cone $R_5^{'}$ out of R_5 by defining the boundaries of $R_5^{'}$ to be the rays $y_{t-2} = \frac{b_1 - a_1}{a_2} y_{t-1}$ and $y_{t-2} = (\frac{b_1 - a_1}{a_2} + \delta) y_{t-1}$. Note that

$$A \begin{pmatrix} y_{t-1} \\ \frac{b_1 - a_1}{a_2} y_{t-1} \end{pmatrix} = \begin{pmatrix} a_1 y_{t-1} + (b_1 - a_1) y_{t-1} \\ y_{t-1} \end{pmatrix} = y_{t-1} \begin{pmatrix} b_1 \\ 1 \end{pmatrix}, \tag{4.27}$$

which lies on the threshold. So then we have from (4.27)

$$\begin{aligned} A \begin{pmatrix} y_{t-1} \\ (\frac{b_1 - a_1}{a_2} + \delta) y_{t-1} \end{pmatrix} &= A \begin{pmatrix} y_{t-1} \\ \frac{b_1 - a_1}{a_2} y_{t-1} \end{pmatrix} + A \begin{pmatrix} 0 \\ \delta y_{t-1} \end{pmatrix} \\ &= y_{t-1} \begin{pmatrix} b_1 \\ 1 \end{pmatrix} + \begin{pmatrix} a_2 \delta y_{t-1} \\ 0 \end{pmatrix}. \end{aligned} \tag{4.28}$$

Thus if $\xi_t \le -a_2 \delta |y_{t-1}|$ then $(y_{t-1}, y_{t-2})^{'}$ such that $y_{t-2} = ((b_1 - a_1)/a_2 + \delta) y_{t-1}$ is mapped to R_2. By the assumption $\xi_t \sim N(0, \sigma^2)$, $P(\xi_t \le -a_2 \delta |y_{t-1}|) = P(\xi_t \ge a_2 \delta |y_{t-1}|) < \varepsilon$ for $|y_{t-1}|$ large. By picking $|y_{t-1}|$ large along the ray $y_{t-1} = (\frac{b_1 - a_1}{a_2} + \delta) y_{t-1}$ and setting

$M_2 = \| (y_{t-1}, y_{t-2})' \|$ we have $M_2 < \infty$ large enough so that the probability of being mapped to R_2 is less than ε and we have

$$\inf_{\substack{x \in R_5 \\ \|x\| > M_2}} P(X_1 \in R_1 | X_0 = x) > 1 - \varepsilon. \tag{4.29}$$

A similar argument for the case of R_4 will reveal that we can cut R_4' out of R_4 by defining for the same $\delta > 0$ the boundaries to be $y_{t-2} = \frac{b_1 - a_1}{a_2} y_{t-1}$ and $y_{t-2} = (\frac{b_1 - a_1}{a_2} - \delta) y_{t-1}$ so that for $M_3 < \infty$ large enough

$$\sup_{\substack{x \in R_4 \\ \|x\| > M_3}} P(X_1 \notin R_2 | X_0 = x) < \varepsilon/2. \tag{4.30}$$

Note here that the assumption $a_1 b_1 + a_2 > 0$ implies the threshold $y_{t-2} = \frac{1}{b_1} y_{t-1}$ is mapped by A to the right of R_2', so by a similar argument there will exist a $M_4 < \infty$ so that

$$\sup_{\substack{x \in R_4 \\ \|x\| > M_4}} P(X_1 \in R_2' \cup R_3' \cup R_3 | X_0 = x) < \varepsilon/2. \tag{4.31}$$

The depiction of the space $\mathbb{R}^2$ with the new partition is in Figure 6.

Let $M = \max(M_1, M_2, M_3, M_4)$ and let us stop here to summarize what we have established thus far for the regions R_1, R_2, R_3, R_4, R_5

$$\inf_{\substack{x \in R_1 \\ \|x\| > M}} P(X_1 \in R_1 | X_0 = x) > 1 - \varepsilon, \quad \sup_{\substack{x \in R_1 \\ \|x\| > M}} P(X_1 \in R_3 \cup R_4 \cup R_4' \cup R_5 \cup R_5' | X_0 = x) < \varepsilon \tag{4.32}$$

$$\inf_{\substack{x \in R_2 \\ \|x\| > M}} P(X_1 \in R_3 | X_0 = x) = \sup_{\substack{x \in R_2 \\ \|x\| > M}} P(X_1 \in R_3 | X_0 = x) = 1/2 \tag{4.33a}$$

$$\inf_{\substack{x \in R_2 \\ \|x\| > M}} P(X_1 \in R_4 | X_0 = x) > 1/2 - \varepsilon \tag{4.33b}$$

$$\sup_{\substack{x \in R_2 \\ \|x\| > M}} P(X_1 \in R_4' \cup R_5 \cup R_5' \cup R_1) < \varepsilon \tag{4.33c}$$

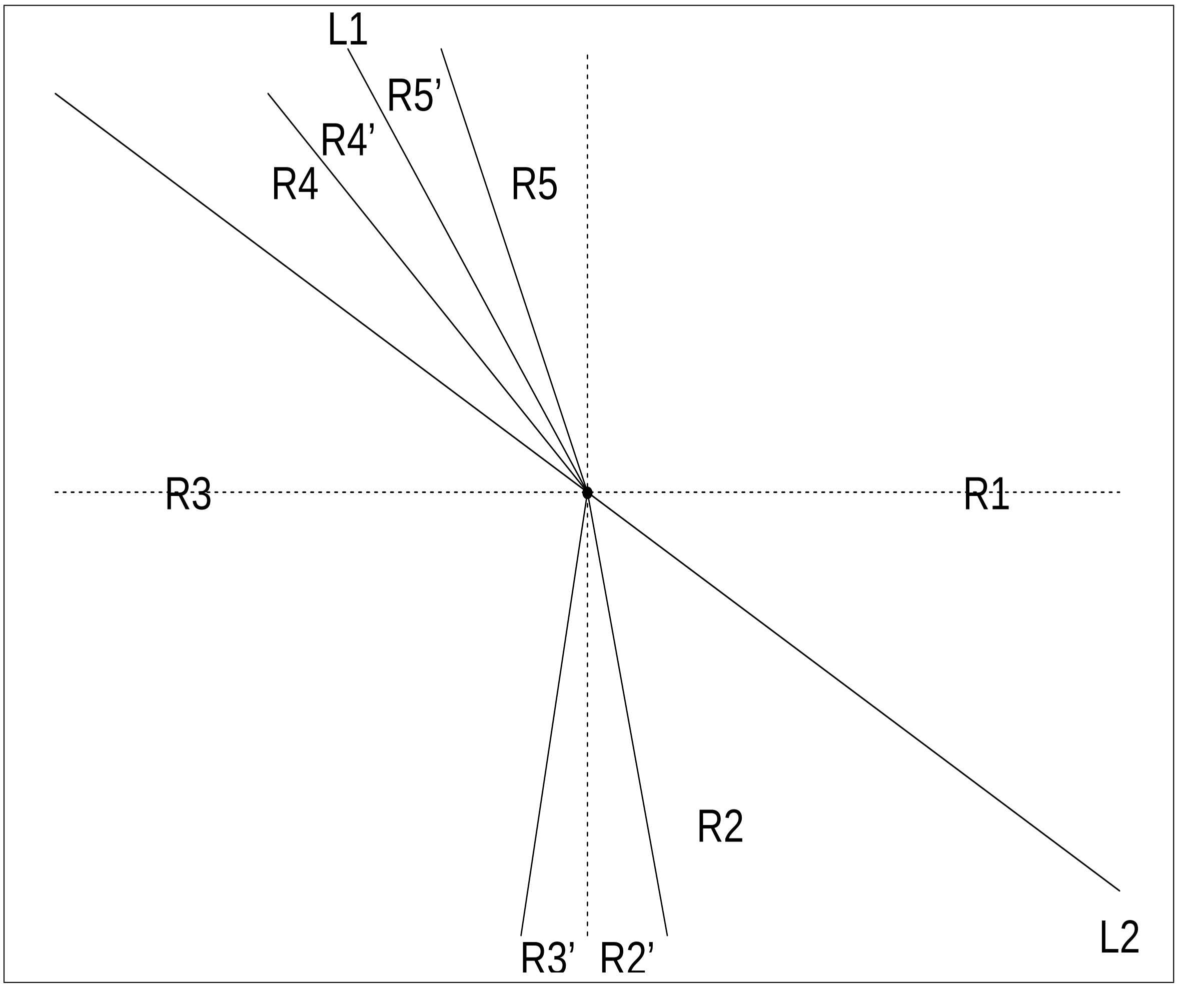

Figure 6. New partition of regions for the TAR(2;1) example.

$$\inf_{\substack{x\in R_3\\ \|x\|>M}} P(X_1 \in R_1|X_0 = x) = \sup_{\substack{x\in R_3\\ \|x\|>M}} P(X_1 \in R_1|X_0 = x) = 1/2 \tag{4.34a}$$

$$1/2-\varepsilon < \inf_{\substack{x\in R_3\\ \|x\|>M}} P(X_1 \in R_2|X_0 = x) \leq \sup_{\substack{x\in R_3\\ \|x\|>M}} P(X_1 \in R_2|X_0 = x) < 1/2 \tag{4.34b}$$

$$\sup_{\substack{x\in R_3\\ \|x\|>M}} P(X_1 \in R_2^{'} \cup R_3^{'} \cup R_3|X_0 = x) < \varepsilon \tag{4.34c}$$

$$\inf_{\substack{x\in R_4\\ \|x\|>M}} P(X_1 \in R_2|X_0 = x) > 1-\varepsilon, \quad \sup_{\substack{x\in R_4\\ \|x\|>M}} P(X_1 \in R_1 \cup R_2^{'} \cup R_3^{'} \cup R_3|X_0 = x) < \varepsilon \tag{4.35}$$

$$\inf_{\substack{x\in R_5\\ \|x\|>M}} P(X_1 \in R_1|X_0 = x) > 1-\varepsilon, \quad \sup_{\substack{x\in R_5\\ \|x\|>M}} P(X_1 \in R_2 \cup R_2^{'} \cup R_3^{'} \cup R_3|X_0 = x) < \varepsilon \tag{4.36}$$

Now to handle the small cones $R_2^{'}$, $R_3^{'}$, $R_4^{'}$ and $R_5^{'}$. Points in $R_2^{'}$ can go to R_1, R_3, R_4, $R_4^{'}$, R_5, $R_5^{'}$ with probabilites depending on where x is in $R_2^{'}$. We cannot ε-approximate these probabilities because of this, but it will not matter since $R_2^{'}$ was designed to be transient. Denote these probabilities by $\alpha_i(x)$ with $\sum_{i=1}^{6} \alpha_i(x) = 1$:

$$\alpha_i(x) = P(X_1 \in R_i|X_0 = x \in R_2^{'}), \;\; i = 1, \ldots, 5. \tag{4.37}$$

$R_3^{'}$ can go to R_1, R_2, $R_2^{'}$, $R_3^{'}$, R_3 with probabilities varying depending on where x is in $R_3^{'}$. We cannot ε-approximate these probabilities because of this; however by the definition of $R_2^{'}$ and $R_3^{'}$ we can bound these probabilities by ε for $\| x \| > M$:

$$\sup_{\substack{x\in R_3^{'}\\ \|x\|>M}} P(X_1 \in R_{2'}|X_0) < \varepsilon \tag{4.38a}$$

$$\sup_{\substack{x\in R_3^{'}\\ \|x\|>M}} P(X_1 \in R_{3'}|X_0) < \varepsilon. \tag{4.38b}$$

Denote the remaining probabilites by $\beta_i(x)$ with $\sum_{i=1}^{3}\beta_i(x) = 1-2\varepsilon$:

$$\beta_i(x) = P(X_1 \in R_i | X_0 = x \in R_3^{'}), \;\; 1 = 1,2,3. \tag{4.39}$$

$R_4^{'}$ can go to $R_1, R_2, R_2^{'}, R_3^{'}, R_3$ with probabilities depending on where x is in $R_4^{'}$, but we can find a $M_5 < \infty$ so that

$$\sup_{\substack{x\in R_4^{'} \\ \|x\|>M_5}} P(X_1 \in R_2^{'} \cup R_3^{'} \cup R_3 | X_0 = x) < \varepsilon. \tag{4.40}$$

As for the other probabilities since they depend on where x is in $R_4^{'}$ we can only say

$$1/2-\varepsilon < \inf_{\substack{x\in R_4^{'} \\ \|x\|>M}} P(X_1 \in R_2 | X_0 = x) \leq \sup_{\substack{x\in R_4^{'} \\ \|x\|>M}} P(X_1 \in R_2 | X_0 = x) < 1-\varepsilon \tag{4.41a}$$

$$\varepsilon < \inf_{\substack{x\in R_4^{'} \\ \|x\|>M}} P(X_1 \in R_1 | X_0 = x) \leq \sup_{\substack{x\in R_4^{'} \\ \|x\|>M}} P(X_1 \in R_1 | X_0 = x) < 1/2. \tag{4.41b}$$

Denote these probabilities by $\gamma(x)$ and $1-\gamma(x)$:

$$P(X_1 \in R_2 | X_0 = x \in R_4^{'}) = \gamma(x), \; P(X_1 \in R_1 | X_0 = x \in R_4^{'}) = 1-\gamma(x). \tag{4.42}$$

Finally, $R_5^{'}$ can go to $R_1, R_2, R_2^{'}, R_3^{'}, R_3$ with probabilities varying depending on where x is in $R_5^{'}$, but we can find a $M_6 < \infty$ so that

$$\sup_{\substack{x\in R_5^{'} \\ \|x\|>M_6}} P(X_1 \in R_2^{'} \cup R_3^{'} \cup R_3 | X_0 = x) < \varepsilon. \tag{4.43}$$

As for the other probabilities, we can only say

$$\varepsilon < \inf_{\substack{x\in R_5^{'} \\ \|x\|>M_6}} P(X_1 \in R_2 | X_0 = x) \leq \sup_{\substack{x\in R_5^{'} \\ \|x\|>M_6}} P(X_1 \in R_2 | X_0 = x) < 1/2-\varepsilon \tag{4.44a}$$

$$1/2 < \inf_{\substack{x\in R_5^{'} \\ \|x\|>M_6}} P(X_1 \in R_1 | X_0 = x) \leq \sup_{\substack{x\in R_5^{'} \\ \|x\|>M_6}} P(X_1 \in R_1 | X_0 = x) < 1-\varepsilon \tag{4.44b}$$

since these probabilities depend on where x is in R_5'. Denote these probabilities by $\eta(x)$ and $1-\eta(x)$:

$$P(X_1 \in R_2 | X_0 = x \in R_5') = \eta(x),\ P(X_1 \in R_1 | X_0 = x \in R_5') = 1-\eta(x). \tag{4.45}$$

It is convenient to summarize these transition probabilities in tabular form:

(4.46)

	1	2	2′	3	3′	4	4′	5	5′
1	$1-\varepsilon$	0	0	$\varepsilon/5$	0	$\varepsilon/5$	$\varepsilon/5$	$\varepsilon/5$	$\varepsilon/5$
2	$\varepsilon/4$	0	0	$1/2$	0	$1/2-\varepsilon$	$\varepsilon/4$	$\varepsilon/4$	$\varepsilon/4$
2′	α_1	0	0	α_2	0	α_3	α_4	α_5	α_6
3	$1/2$	$1/2-\varepsilon$	$\varepsilon/3$	$\varepsilon/3$	$\varepsilon/3$	0	0	0	0
3′	β_1	β_2	ε	β_3	ε	0	0	0	0
4	$\varepsilon/4$	$1-\varepsilon$	$\varepsilon/4$	$\varepsilon/4$	$\varepsilon/4$	0	0	0	0
4′	γ	$1-\gamma-\varepsilon$	$\varepsilon/3$	$\varepsilon/3$	$\varepsilon/3$	0	0	0	0
5	$1-\varepsilon$	$\varepsilon/4$	$\varepsilon/4$	$\varepsilon/4$	$\varepsilon/4$	0	0	0	0
5′	η	$1-\eta-\varepsilon$	$\varepsilon/3$	$\varepsilon/3$	$\varepsilon/3$	0	0	0	0

Then we can form the transition probability matrix for $\{J_t\}$ by letting $\varepsilon \to 0$:

(4.47)

	1	2	2′	3	3′	4	4′	5	5′
1	1	0	0	0	0	0	0	0	0
2	0	0	0	$1/2$	0	$1/2$	0	0	0
2′	α_1	0	0	α_2	0	α_3	α_4	α_5	α_6
3	$1/2$	$1/2$	0	0	0	0	0	0	0
3′	β_1	β_2	0	β_3	0	0	0	0	0
4	0	1	0	0	0	0	0	0	0
4′	γ	$1-\gamma$	0	0	0	0	0	0	0
5	1	0	0	0	0	0	0	0	0
5′	η	$1-\eta$	0	0	0	0	0	0	0

Since states $2^{'}$ and $3^{'}$ are immediately mapped to other states and never return, it follows the states $2^{'}$ and $3^{'}$ are transient for $\{J_t\}$. Since only $2^{'}$ maps to $4^{'}$ and $5^{'}$, it follows $4^{'}$ and $5^{'}$ are transient for $\{J_t\}$ as well. Define $G = \{1,2,3,4,5\}$ and $T = \{2^{'},3^{'},4^{'},5^{'}\}$. The states in T are transient for $\{J_t\}$.

Turning our attention to $\{X_t\}$, we verify the assumptions in (A2). The ε-approximation in (1.5) is satisified by the regions/states corresponding to G. We need to verify (1.6): there exists $t^* < \infty$ so that

$$\sup_{i \in T} \sup_{\substack{x \in R_1 \\ \|x\| > M}} P(X_{t^*} \in \cup_{k \in T} R_k | X_0 = x) < \varepsilon. \tag{4.48}$$

A glance at (4.46) reveals regions $R_2^{'}$ and $R_3^{'}$ satisfy this condition immediately. $R_2^{'}$ maps to $R_4^{'}$ and $R_5^{'}$. The only other regions which map to $R_4^{'}$ or $R_5^{'}$ are R_1 and R_2 and these do so with arbitrarily small probabilities. Clearly then we can find a $t^* < \infty$ so that (4.48) is satisified.

Thus, we can apply Theorem 2 in Chapter II and under the condition on $\{J_t\}$

$$\Pi_{i=1}^{l} (\| A_i \|)^{\pi_i} < 1, \tag{4.49}$$

where π is any stationary distribution of $\{J_t\}$, we will have that $\{X_t\}$ is V-uniformly ergodic, with $V(\cdot)$ specified by Theorem 2.

Examination of the transition matrix of $\{J_t\}$ reveals that $\{1\}$ is the only closed state and every other state maps into $\{1\}$. Since $\{J_t\}$ is a finite state chain, this means that $\{1\}$ is the only recurrent state and so the stationary distribution π with have $\pi_1 = 1$ and zeroes everywhere else. The condition for V-uniform ergodicity of $\{X_t\}$ in the case $b_1 < 0$, $b_2 = 0$, $a_1, a_2 > 0$, $a_1 b_1 + a_2 > 0$, $b_1^2 > a_1 b_1 + a_2$ is then $\| A \| < 1$, that is, $a_1 + a_2 < 1$.

Note that we only require b_1 to be such that $b_1 < 0$, $a_1 b_1 + a_2 > 0$ and $b_1^2 > a_1 b_1 + a_2$. In particular, we do not require $|b_1| < 1$ which would be the condition generalizing from the linear case.

CHAPTER V

FUTURE RESEARCH

We propose to use countable state chains to approximate general state space chains in the case where recurrent regions do not allow ε-approximation of their transition probabilities. The task here would be to verify the countable chain is ergodic, to find the stationary distribution of the countable state space chain and to use that stationary distribution to identify an ergodicity condition for $\{X_t\}$.

For details, consider that in cases where regions that do not allow ε-approximation by a finite state chain are recurrent the finite state chain approximation is useless since the stationary distributions of finite state Markov chains (or any Markov chain, for that matter) will be different according to the differing transition probabilities for these recurrent regions. Thus the condition for ergodicity of $\{X_t\}$

$$\sum_{j=1}^{l} \pi_j \log(\| A_j \|) < 0, \tag{5.1}$$

where π is the stationary distribution of $\{J_t\}$, will have no relevance in the case where the regions that do not allow ε-approximation are recurrent since π does not accurately describe the long term behavior of the transitions of $\{X_t\}$ from region to region. In order to glean conditions for ergodicity of $\{X_t\}$ from the stationary distribution of $\{J_t\}$ it is thus necessary to somehow ε-approximate the 'transition' probabilities for these 'recurrent' regions of the state space of $\{X_t\}$.

5.1 Countable State Markov Chains

We propose to ε-approximate the transition probabilities with a countable state Markov chain rather than a finite state Markov chain. In certain situations the recurrent regions whose transition probabilites cannot be ε-approximated by a finite partition of the region will admit such ε-approximation under a countable partition. In these situations it is conjectured the condition on $\{J_t\}$ given in (5.1) can be suitably generalized to

$$\sum_{j\geq 0} \pi_j \log(\| A_j \|) < 0 \tag{5.2}$$

and that this condition can be used to demonstrate V-uniform ergodicity of $\{X_t\}$. There are several issues that arise.

Because we were dealing with a finite state chain $\{J_t\}$ in Chapter III, the number of transient states was thereby finite and thus uniformly transient, i.e., as it was expressed in (3.52) for arbitrary $\varepsilon > 0$ there exists a $t^{**} < \infty$ so that

$$\sup_{j\in T} \sup_{i\in T} P(J_{t^{**}} = j | J_0 = i) < \varepsilon. \tag{5.3}$$

This uniform transience condition on the states where the ε-approximation does not hold was crucial in proving Lemma 1 and Lemma 5 in Chapter III and was therefore critical in proving the results on V-uniform ergodicity of $\{X_t\}$ found in Chapter III.

When $\{J_t\}$ is a countable state chain the number of transient states need not be finite and therefore the transient states are not necessarily uniformly transient. This presents a problem. In order to be able to extend the results in Chapter III to the case of a countable state $\{J_t\}$, it is necessary that we get away in a finite time with a probability arbitrarily close to 1 from the states where the ε-approximation does not hold. We thus require these states of $\{J_t\}$ be uniformly transient (which is trivially satisfied if they are finite in number).

We require more of $\{J_t\}$. The proof in Cline and Pu (2002) of Lemma 1 in Chapter III that there exists a finite n so that

$$\frac{1}{n}\sum_{t=1}^{n} E\left(h(J_t)\middle| J_0 = i\right) < 0, \quad \forall i \tag{5.4}$$

required not only that the transient states be uniformly transient, but also that the irreducible pieces of $\{J_t\}$ (or $\{J_t\}$ itself in the case of irreducibility) be uniformly ergodic. Where $\{J_t\}$

is a finite state chain this again follows automatically from the fact the number of recurrent states is finite. In the case where $\{J_t\}$ is a countable state space chain this does not follow and must be assumed. We therefore require that $\{J_t\}$ be uniformly ergodic. This is a logical assumption when one considers the convergence due to the ergodicity must take place over the entire space in a finite time with a probability arbitrarily close to 1.

Let us summarize these assumptions on $\{J_t\}$

Assumption 1. **(A3)** *Suppose there exists a uniformly ergodic countable state Markov chain $\{J_t\}$ with G consisting of the recurrent, T consisting of the transient states for $\{J_t\}$ and that the transient states T are uniformly transient. Suppose further for arbitrary $\varepsilon > 0$ there exists $M < \infty$ such that*

$$\sup_{j} \sup_{i \in G} \sup_{\substack{x \in R_i \\ \|x\| > M}} \left| P(X_1 \in R_j | X_0 = x) - P(J_1 = j | J_0 = i) \right| < \varepsilon \tag{5.5}$$

and there exists $t^* = t^*(M) < \infty$ such that

$$\sup_{i \in T} \sup_{\substack{x \in R_i \\ \|x\| > M}} P(X_{t^*} \in \cup_{k \in T} R_k | X_0 = x) < \varepsilon. \tag{5.6}$$

Suppose also that

$$\Pi_{i \geq 1} (\| A_i \|)^{\pi_i} < 1, \tag{5.7}$$

where π is any stationary distribution of $\{J_t\}$.

Under these assumptions we conjecture the condition (5.7) will imply the V-uniform ergodicity of $\{X_t\}$ through arguments similar to those in Chapter III. The challenge becomes verifying that $\{J_t\}$ satisfies the assumptions made of it and finding the stationary distribution of $\{J_t\}$.

There is still much to do here. Classifying an arbitrary countable state Markov chain as ergodic or not, and if ergodic whether it is uniformly ergodic or not, and if so finding the stationary distribution requires further work. We can suppose that $\{J_t\}$ is irreducible and aperiodic; if $\{J_t\}$ is not irreducible, we must decompose the state space into irreducible pieces and find the stationary distribution for each. If $\{J_t\}$ is periodic with period d we can look at the d-step chain $\{J_t\}$ which will be aperiodic. Perhaps $\{J_t\}$ will have a special structure we can exploit. These suggestions may help simplify the task of determining whether $\{J_t\}$ is uniformly ergodic.

The proposed approach raises new questions which need to be answered. Looking at the glass as half full rather than half empty, the general state space Markov chain $\{X_t\}$ has been reduced to a simpler, countable state Markov chain $\{J_t\}$ which will be easier to analyze and simulate. This indicates promise for future research.

CHAPTER VI

CONCLUSIONS

We derived conditions for the ergodicity of threshold autoregressive time series by embedding the time series in a general state Markov chain and applied a Foster-Lyapunov drift condition to demonstrate ergodicity of the Markov chain. In particular we were interested in demonstrating V-uniform ergodicity where the test function $V(\cdot)$ depends upon a norm.

In this dissertation we provided conditions under which the general state space chain may be approximated by a simpler system and provided conditions on the simpler system which imply V-uniform ergodicity of the general state space Markov chain and thus the threshold autoregressive time series embedded in it. We also examined conditions under which the general state space chain and thus the nonlinear time series embedded in it may be classified as transient. Finally, we provided conditions under which central limit theorems will exist for the general state space chain and by implication for the associated threshold autoregressive time series.

REFERENCES

An, H. Z., and Chen, S. G. (1997), "A Note on the Ergodicity of Non-linear Autoregressive Model", *Statistics & Probability Letters*, 34, 365-372.

An, H. Z., and Huang, F. C. (1996), "The Geometrical Ergodicity of Nonlinear Autoregressive Models", *Statistica Sinica*, 6, 943-956.

Billingsley, P. (1999), *Convergence of Probability Measures*, New York; Wiley, John and Sons.

Brockwell, P. J., and Davis, R. A. (1987), *Time Series: Theory and Methods*, New York; Springer-Verlag.

Chan, K. S., (1990), "Deterministic Stability, Stochastic Stability, and Ergodicity", Appendix 1 in *Non-linear Time Series Analysis: A Dynamical System Approach*, ed. H. Tong, London, U.K.; Oxford University Press.

Chan, K. S., Petruccelli, J. D., Tong, H., Woolford, S. W. (1985),"A Multiple-threshold AR(1) Model", *J. Appl. Prob.*, 22, 267-279.

Chan, K. S., and Tong, H. (1994), "A Note on Noisy Chaos", *Journal of the Royal Statistical Society*, B 56, No. 2, 301-311.

Chan, K. S., and Tong, H. (2001), *Chaos: A Statistical Perspective*, New York; Springer-Verlag.

Chen, R., and Tsay, S. R. (1991), "On the Ergodicity of TAR(1) Processes", *The Annals of Applied Probability*, 1, No. 4, 613-634.

Ciarlet, P. G. (1982), *Intrduction a la Analyse Numerique Matricielle et a l'Optimisation*, Paris, France; Masson.

Cline, D. B. H., and Pu, H. (1998), "Verifying Irreducibility and Continuity of a Nonlinear Time Series", *Statistics & Probability Letters*, 40, 139-148.

Cline, D. B. H., and Pu, H. (1999a), "Stability of Nonlinear AR(1) Time Series With Delay", *Stochastic Processes and their Applications*, 82, 307-333.

Cline, D. B. H., and Pu, H. (1999b), "Geometric Ergodicity of Nonlinear Time Series", *Statistica Sinica*, 9, 1103-1118.

Cline, D. B. H., and Pu, H. (2000), "Geometric Transience of Nonlinear Time Series", *Statistica Sinica*, 11, 273-287.

Cline, D. B. H., and Pu, H. (2001), "Stability of Nonlinear Time Series: What Does Noise Have to do With It?", *Selected Proceedings of the Symposium on Inference for Stochastic Processes*, 37, 151-170.

Cline, D. B. H., and Pu, H. (2002), "Stability of Threshold-like ARMA Time Series", *(submitted for publication)*, Dept. of Statistics, Texas A&M University, USA.

Diaconis, P., and Freedman, D. (1999), "Iterated Random Functions", *SIAM Review*, 41, No. 1, 45-76.

Durrett, R. (1996), *Probability: Theory and Examples*, Pacific Grove, CA; Wadsworth.

Guo, M., and Petruccelli, J. D. (1991), "On the Null Recurrence and Transience of a First Order SETAR Model", *J. Appl. Prob.*, 28, 584-592.

Kunitomo, N. (2001), "On Ergodicity of Some TAR(2) Processes", *(submitted for publication)*, Dept. of Economics, University of Tokyo, Japan.

Lim, K. S. (1992), "On the Stability of a Threshold AR(1) Without Intercepts", *Journal of Time Series Analysis*, 13, No. 2, 119-132.

Martelli, M. (1992), *Discrete Dynamical Systems and Chaos*, New York; J. Wiley and Sons.

Meyn, S. P., and Tweedie, R. L. (1993), *Markov Chains and Stochastic Stability*, London, U.K.; Springer-Verlag.

Nummelin, E. (1984), *General Irreducible Markov Chains and Non-negative Operators*, Cambridge, U.K.; Cambridge University Press.

Petruccelli, J. D., and Woolford, S. W. (1984), "A Threshold AR(1) Model", *J. Appl. Prob.*, 21, 270-286.

Priestley, M. B. (1980), "State-dependent Models: A General Approach to Non-linear Time Series Analysis", *Journal of Time Series Analysis*, 1, No. 1, 47-71.

Resnick, S. I. (1992), *Adventures in Stochastic Process*, Boston, MA; Birkhauser.

Spieksma, F. M., and Tweedie, R. L. (1994), "Strengthening Ergodicity to Geometric Ergodicity for Markov Chains", *Stoch. Models*, 10, 45-74.

Tjostheim, D. (1990), "Non-linear Time Series and Markov Chains", *Adv. Appl. Probab.*, 22, 587-611.

Tong, H. (1983), "Threshold Models in Nonlinear Time Series Analysis", *Lecture Notes in Statistics*, No. 21, Heidelburg, Germany; Springer-Verlag.

Tong, H. (1990), *Non-linear Time Series Analysis: A Dynamical System Approach*, London, U.K.; Oxford University Press.

Tong, H., and Lim, K. S. (1980), "Threshold Autoregression, Limit Cycles and Cyclical Data", *Journal of the Royal Statistical Society B: Methodological*, 42, 245-292

Tweedie, R. L. (1976), "Criteria for Classifying General Markov Chains", *Advances in Applied Probability*, 8, 737-771.

VITA

Thomas R. Boucher was born in Derry, NH. He received Bachelor of Science degrees in philosophy and in mathematics from the University of Massachusetts-Lowell in June 1994 and June 1996, respectively. He received a Master of Science in applied mathematics in June, 1998. In 1998 he entered the Ph.D program in statistics at Texas A&M University. He received his Ph.D degree in December 2003. His permanent street address is 156 Thomas Hill Rd. Pembroke, VA 24136.