

IRREDUCIBILITY AND CONTINUITY ASSUMPTIONS FOR POSITIVE OPERATORS WITH APPLICATION TO THRESHOLD GARCH TIME SERIES MODELS

DAREN B. H. CLINE,* *Texas A&M University*

Abstract

Suppose that $\{X_t\}$ is a Markov chain such as the state space model for a threshold GARCH time series. The regularity assumptions for a drift condition approach to establishing the ergodicity of $\{X_t\}$ typically are ϕ -irreducibility, aperiodicity, and a minorization condition for compact sets. These can be very tedious to verify due to the discontinuous and singular nature of the Markov transition probabilities. We first demonstrate that, for Feller chains, the problem can at least be simplified to focusing on whether the process can reach some neighborhood that satisfies the minorization condition. The results are valid not just for the transition kernels of Markov chains but also for bounded positive kernels, opening the possibility for new ergodic results. More significantly, we show that threshold GARCH time series and related models of interest can often be embedded into Feller chains, allowing us to apply the conclusions above.

Keywords: Feller operator; T -chain; GARCH; threshold time series

2010 Mathematics Subject Classification: Primary 60J05

Secondary 37A50; 62M10; 91B84

1. Introduction

1.1. Overview

One of the more troublesome features of studying the Markov state process for a nonlinear time series is that it can be difficult to establish irreducibility and other regularity assumptions typically needed for proving ergodicity. Usually, each type of model must be checked individually, and the process is often tedious. Since the transition probabilities for the state process are singular (often one-dimensional), one must identify the properties of some n -step transition density, where n is at least as big as the dimension of the state vector. Models with discontinuous behavior present further difficulties. For example, it is not at all clear how best to verify the regularity assumptions for a threshold AR-GARCH model consisting of different regimes with possibly widely different behavior. (See, for example, Cline (2007).)

By regularity assumptions, we mean ϕ -irreducibility, aperiodicity, and some type of minorization condition for compact sets. (See Nummelin (1984) or Meyn and Tweedie (1993).) Compact sets are of particular interest because (irreducible) time series models are usually transient only by escaping compact sets. The minorization condition is a continuity condition: it establishes that the behavior of the process is locally smooth in some sense. The (weak) Feller property, if it holds, is not enough by itself.

As it happens, these properties are closely aligned and frequently end up being verified more or less simultaneously. This can be a bit of a problem since knowledge of one property may be

Received 7 January 2009; revision received 15 September 2010.

* Postal address: Department of Statistics, Texas A&M University, College Station TX 77843-3143, USA.

Email address: dcline@stat.tamu.edu

needed to clinch the validity of another. Being global in nature, furthermore, these assumptions are not easily deduced for nonlinear models, and especially not for threshold models. One possibility is to establish a weak continuity condition first, such as the Feller property, in the hope that the stronger properties will then follow from a more local condition. Showing that this is possible is the first of our objectives for this paper.

On the other hand, the transition kernel for the state process of a threshold model is clearly not Feller. Or is it? We will show that a typical threshold model can usually be reformulated as a Feller process. This does not mean that the behavior is no longer discontinuous at the thresholds, but rather it means that we can deal with the thresholds in an elegant manner, thereby enabling the use of results for Feller chains. As a result, we will define a class of time series with the seemingly contradictory descriptor *continuous threshold model*.

In fact, we wish to verify these regularity assumptions for both stochastic and nonstochastic kernels. Supposing that $Q_0(x, A) = P(X_1 \in A \mid X_0 = x)$ is the transition kernel for a Markov chain $\{X_t\}$ and that $\tilde{w}(x, y) \geq 0$, the kernel operator

$$Q(x, A) = E(\tilde{w}(x, X_1) \mathbf{1}_{\{X_1 \in A\}} \mid X_0 = x) = \int_A \tilde{w}(x, y) Q_0(x, dy) \tag{1.1}$$

may also be of interest. While such operators may inherit irreducibility from the transition kernel, they generally do not inherit continuity without also assuming a corresponding continuity of $\tilde{w}(x, y)$.

1.2. Objectives for threshold models

To be more specific, suppose that the Markov chain $\{X_t\}$ in \mathbb{R}^m is defined by

$$X_t = z(X_{t-1}, e_t) := \sum_{j=1}^N h_j(X_{t-1}, e_t) \mathbf{1}_{C_j}(X_{t-1}), \tag{1.2}$$

where C_1, \dots, C_N are connected sets with nonempty interiors that partition \mathbb{R}^m and $h_j(x, u)$, $j = 1, \dots, N$, are vector-valued functions, continuous in x . Since $z(x, u)$ is piecewise continuous but not necessarily continuous, it is easy to see that the Markov transition operator is usually not Feller. The boundaries of the C_j s are known as *thresholds* and the C_j s themselves are called *regimes*. What is at stake is the behavior of X_t near the thresholds.

Typically, however, the model behavior (within a regime) is locally Lipschitz near the thresholds. Suppose, for example, that $\{X_t\}$ satisfies (1.2) with $e_t \stackrel{\text{i.i.d.}}{\sim} \mu$ and h_1, \dots, h_N as follows: for some $L_1 < \infty$,

$$\|h_j(x_1, u) - h_j(x_2, u)\| \leq L_1(1 + \|u\|)\|x_1 - x_2\|, \quad j = 1, \dots, N, \tag{1.3}$$

where $\|\cdot\|$ here is a norm equivalent to the Euclidean norm. Suppose that we also have, for some $\zeta \geq 0$ and locally Lipschitz \tilde{w} ,

$$w(x, u) := \tilde{w}(x, z(x, u)) \leq L_2(1 + \|u\|)^\zeta \quad \text{and} \quad E(\|e_1\|^\zeta) < \infty. \tag{1.4}$$

In the context of a drift condition for determining whether the process is ergodic, $\tilde{w}(x, y)$ might take the form $(1 + \|y\|)^\zeta / (1 + \|x\|)^\zeta$. More generally, we might have a *test function* $V(x)$ such that $m_1(1 + \|x\|)^\zeta \leq V(x) \leq m_2(1 + \|x\|)^\zeta$, and let $\tilde{w}(x, y) = V(y)/V(x)$. (For the probability transition operator itself, $\tilde{w}(x, y) \equiv 1$.) With μ as the distribution of e_t on \mathbb{E} , the transition kernel Q is thus given by

$$Q(x, A) = E(\tilde{w}(x, X_1) \mathbf{1}_A(X_1) \mid X_0 = x) = \int_{\mathbb{E}} w(x, u) \mathbf{1}_A(z(x, u)) \mu(du).$$

In order for this kernel to have sufficient ‘continuity’, there needs to be a certain amount of regularity near the thresholds. Specifically, the following are needed. Let C_j° denote the interior of C_j , and let $\mathbb{X}_\# = \bigcup_{j=1}^N C_j^\circ$. The complement of $\mathbb{X}_\#$ comprises the thresholds.

- (i) Conditions (1.3) and (1.4) hold, where $\tilde{w}(x, y)$ is positive and locally Lipschitz continuous on \mathbb{R}^{2m} , and is bounded away from 0 on compact sets.
- (ii) $P(h_j(x, e_1) \notin \mathbb{X}_\#) = 0$ for all $x \in C_j^\circ, j = 1, \dots, N$. That is, the process cannot move directly onto a threshold. This usually requires only that e_t has a continuous distribution.
- (iii) For each $\varepsilon > 0$, finite M , and $j_1 \neq j_2$, there exists a $\delta > 0$ such that $x_1, x_2 \in C_{j_1}^\circ, \|x_i\| \leq M$, and $\|x_1 - x_2\| < \delta$ imply that

$$P(h_{j_1}(x_1, e_1) \in C_{j_2}^\circ, h_{j_2}(x_2, e_1) \in C_{j_1}^\circ, \|e_1\| \leq M) < \varepsilon, \quad j = 1, \dots, N.$$

This condition says that two processes (driven by the same noise sequence) starting at adjacent points in the same regime are unlikely to move immediately to different regimes, such as opposite sides of a threshold.

- (iv) Now let \bar{C}_j be the closure of C_j , define $X_t(x)$ to be the value of the process assuming $X_0 = x$, and let $d(x)$ be the distance from x to the nearest threshold. The final condition is simply that a process starting near a threshold has sufficient probability that it eventually will not be near a threshold. To be precise, for each $x \in \bar{C}_j, j = 1, \dots, N$, and each sequence $\{x_n\}$ in C_j° such that $\|x_n - x\| \rightarrow 0$, there exist $\delta > 0$ and $k \geq 1$ so that

$$\liminf_{n \rightarrow \infty} P(d(X_k(x_n)) > \delta) > 0. \tag{1.5}$$

Note that both k and δ can depend on the sequence $\{x_n\}$.

How these properties enable us to reformulate $\{X_t\}$ in terms of a Feller process, and how to show that the resulting Feller process is sufficiently regular for discussing ergodicity, is the goal of this paper.

In summary, the objectives of this paper are to determine how the requirements mentioned in Subsection 1.1 may be simplified in the context of a Feller chain and to show that (most) threshold models can be redefined as Feller chains, both of which are to be generalized to bounded positive kernel operators. In addition, we will put the results to use for threshold GARCH models and for related models of interest. In the remainder of this section, we provide our general setting and definitions. In Section 2 we summarize certain results about the Feller and Tweedie properties, as extended to bounded positive operators, in order to establish the groundwork for additional useful results about Feller operators, also presented in Section 2. Section 3 is the heart of the paper, where we reformulate well-behaved threshold operators as Feller operators and identify the consequences. Section 4 contains the applications and Section 5 has the proofs.

1.3. General setting

Let \mathbb{X} be a locally compact separable metric space with metric ρ . In the time series setting, we would usually have $\mathbb{X} \subset \mathbb{R}^m$ with the metric obtained from the Euclidean norm. (In fact, if \mathbb{X} is a locally compact topological *vector* space, as is usual, then it is necessarily finite dimensional.) Another metric may be more useful occasionally, such as one that requires compact sets to be bounded away from the origin. We suppose throughout the paper that Q is a bounded positive

kernel on \mathbb{X} , namely $Q(\cdot, A)$ is a bounded measurable function for each Borel set $A \subset \mathbb{X}$ and $Q(x, \cdot)$ is a nontrivial (finite) measure for each $x \in \mathbb{X}$.

Letting $\mathcal{L}^\infty(\mathbb{X})$ denote the class of bounded measurable real-valued functions, Q may be interpreted as a continuous linear operator from $\mathcal{L}^\infty(\mathbb{X})$ into itself:

$$Qf(x) := \int_{\mathbb{X}} f(y) Q(x, dy). \tag{1.6}$$

In particular, $Q(x, A) = Q\mathbf{1}_A(x)$. Even when not bounded, we will call any operator determined by (1.6) a kernel. Of course, we have the usual iteration definitions: $(Q_1 Q_2)f = Q_1(Q_2 f)$ and $Q^n f = Q(Q^{n-1} f)$, with $Q^0 f := f$. A sequence $\{Q^n\}$ is called a *chain*.

In general, Q is not an integral operator, which is to say that there need be no measure that dominates $Q(x, \cdot)$ for all x . For some n , however, Q^n may be integral and that can be helpful.

As is customary, the class of Borel sets will be denoted by $\mathcal{B}(\mathbb{X})$ and the class of bounded continuous real-valued functions (with respect to ρ) will be denoted by $\mathcal{C}(\mathbb{X}, \rho)$. Where needed, we have the \mathcal{L}^∞ norm

$$\|f\| := \sup_{x \in \mathbb{X}} |f(x)| \quad \text{for } f \in \mathcal{L}^\infty(\mathbb{X})$$

and the operator norm (since Q is positive)

$$\|Q\| := \|Q\mathbf{1}_{\mathbb{X}}\| = \sup_{x \in \mathbb{X}} Q(x, \mathbb{X}).$$

When $Q(x, \mathbb{X}) = 1$ for all x , Q is of course stochastic and is the transition kernel for some Markov chain $\{X_t\}$. For example, X_t may be the state vector for a time series $\{\xi_t\}$, an embedding which enables one to investigate the stability/ergodicity properties of the time series. We are particularly interested in time series such as the threshold AR model with threshold GARCH errors. This type of model has state vector

$$X_t = (\xi_t, \dots, \xi_{t-p+1}, v_t, \dots, v_{t-q+1}, \sigma_t, \dots, \sigma_{t-r+1})$$

defined by

$$\begin{aligned} \xi_t &= a_{0,J_t} + \sum_{i=1}^p a_{i,J_t} \xi_{t-i} + v_t, & v_t &= \sigma_t e_t, \\ \sigma_t &= \left(b_{0,J_t} + \sum_{i=1}^q b_{i,J_t} v_{t-i}^2 + \sum_{i=1}^r c_{i,J_t} \sigma_{t-i}^2 \right)^{1/2}, & \text{and } J_t &= \sum_{j=1}^N j \mathbf{1}_{X_{t-1} \in C_j}, \end{aligned} \tag{1.7}$$

where the regimes C_1, \dots, C_N partition $\mathbb{X} \subset \mathbb{R}^{2p+q}$ and $\{e_t\}$ is an independent and identically distributed (i.i.d.) real-valued sequence of random variables.

Conditions for the process (1.7) to be ergodic are given in Cline (2007). Models of this type, or similar, have also been investigated by many authors, including Li and Li (1996), Lu (1996), Liu *et al.* (1997), Ling (1999), Ling and McAleer (2002), Ling (2007), and Meitz and Saikkonen (2008). A general approach for showing ergodicity of such a process involves verifying a *drift condition* (cf. Meyn and Tweedie (1993)). Our interest in this paper, however, is not in the drift condition but rather in the side (regularity) conditions mentioned in the overview. Our intention is to present a systematic, and hopefully simpler, approach to demonstrating the regularity conditions for particular threshold models.

1.4. Definitions

We use the fairly standard terminology from Meyn and Tweedie (1993), extended to positive kernel operators. Also, see Nummelin (1984).

If $a = \{a_n\}$ is a nonnegative sequence and Q is a bounded positive kernel, then let

$$K_a(x, A) := \sum_{n=0}^{\infty} a_n Q^n(x, A).$$

When Q is the Markov operator for $\{X_t\}$ and a is a probability distribution on \mathbb{Z}_+ , Meyn and Tweedie (1993) point out that K_a is the transition kernel of the *sampled chain* $\{X_{S_t}\}$ ($S_t = \sum_{i=1}^t Y_i, Y_i \stackrel{i.i.d.}{\sim} a$). Accordingly, we will call any K_a a *sampled kernel*, though it may not even be σ -finite. Bounded sampled kernels commute and satisfy (due to the Chapman–Kolmogorov equations)

$$K_a K_b = K_b K_a = K_{a*b},$$

where $a*b$ denotes convolution. When K_a is bounded and ν is any measure, νK_a is the measure defined by $\nu K_a(A) = \int K_a(x, A) \nu(dx)$.

We also recall the definition of the resolvent kernel,

$$R_\lambda(x, A) := \sum_{n=0}^{\infty} \lambda^{-n-1} Q^n(x, A) \quad \text{for } \lambda \in \mathbb{C}, \lambda \neq 0,$$

which satisfies $QR_\lambda f = \lambda R_\lambda f - f$. The resolvent is finite (in fact bounded) if $|\lambda|$ is greater than the spectral radius of $Q, r(Q) := \lim_{n \rightarrow \infty} \|Q^n\|^{1/n}$. Note that $r(Q) \leq \|Q\| < \infty$.

We start with two notions of irreducibility, one of which (cf. Nummelin (1984)) has been found to be the most useful for Markov chain theory. There are others: the functional analysis literature typically uses a definition much stronger than the ones we consider here.

Definition 1.1. Fix any $\lambda > 0$.

- (i) Q is called *ϕ -irreducible* if there exists a nontrivial measure ϕ such that $\phi(A) > 0$ implies that $R_\lambda(x, A) > 0$ for all $x \in \mathbb{X}$. In this case, ϕ is called an *irreducibility measure*. If ϕ is an irreducibility measure such that Q is ϕ_1 -irreducible if and only if $\phi_1 \ll \phi$ then ϕ is called *maximal*.
- (ii) Q is called *open set irreducible* if $R_\lambda(x, A) > 0$ for every nonempty open set A and all $x \in \mathbb{X}$.

Note that the choice of λ is irrelevant to either definition. Henceforth, we will always assume that $\lambda > r(Q)$ so that R_λ is bounded. The following states that the existence of a maximal irreducibility measure is assured when Q is ϕ -irreducible (see Nummelin (1984, Proposition 2.4) and Meyn and Tweedie (1993, Theorem 4.2.2)). Throughout this paper, a proposition will denote a commonly known result from the literature (at least for stochastic kernels).

Proposition 1.1. *If Q is ϕ -irreducible then a maximal irreducibility measure exists and any maximal measure is equivalent to the measure ψ defined by*

$$\psi(A) = \phi_1 R_\lambda(A) = \int_{\mathbb{X}} R_\lambda(y, A) \phi_1(dy), \tag{1.8}$$

if $\lambda > r(Q)$ and ϕ_1 is (without any loss) chosen to be any finite irreducibility measure. Furthermore, $\psi(A) > 0$ if and only if $\psi(\{y : R_\lambda(y, A) > 0\}) > 0$.

Remark. Suppose that Q and Q_0 are related by (1.1) with $\tilde{w}(x, y) > 0$ for all $x, y \in \mathbb{X}$. It is quite evident from Definition 1.1 that Q is ϕ -irreducible if and only if Q_0 is. Although we will state our results more generally, this is a very common scenario.

Open set irreducibility is an intuitive, if slightly naive, concept. In the case of a Markov kernel, it indicates simply that the process has a positive chance of eventually being near any point, no matter where it starts. Focusing on just one such set or point, we have the following definition.

Definition 1.2. If $R_\lambda(x, A) > 0$ for all $x \in \mathbb{X}$ then A is called *reachable*. If every neighborhood A of $x_* \in \mathbb{X}$ is reachable then x_* is called *reachable*.

Clearly, if Q is ϕ -irreducible then every ϕ -positive set is reachable. One connection between reachability and ϕ -irreducibility is the following, proved in Section 5. Compare with Nummelin (1984, Proposition 2.5). Other connections will be presented in the second section.

Lemma 1.1. *If Q is ϕ -irreducible with maximal measure ϕ then every reachable set is ϕ -positive. In particular, every neighborhood of a reachable point x_* is ϕ -positive.*

Furthermore, the set of reachable points is the support of ϕ (that is, the complement of the union of all ϕ -null open sets).

This result implies that a ϕ -irreducible kernel is open set irreducible if and only if \mathbb{X} is the support of the maximal irreducibility measure. Unfortunately, this seems to be about the only benefit to asking whether *every* point is reachable.

We now discuss continuity notions. Recall that $\mathcal{C}(\mathbb{X}, \rho)$ is the collection of bounded continuous functions.

Definition 1.3. (i) Q is called *Feller* if it maps $\mathcal{C}(\mathbb{X}, \rho)$ into $\mathcal{C}(\mathbb{X}, \rho)$. In this case, the sequence $\{Q^n\}_{n \geq 0}$ is called a *Feller chain*.

(ii) Q is called *Tweedie* (T -continuous) on (\mathbb{X}, ρ) if there exist a bounded sampled kernel K_a and a positive operator $T \leq K_a$ such that $T(\cdot, A)$ is lower semicontinuous on (\mathbb{X}, ρ) for all $A \in \mathcal{B}(\mathbb{X})$ and $T(x, \mathbb{X}) > 0$ for all $x \in \mathbb{X}$. In this case, the sequence $\{Q^n\}_{n \geq 0}$ is called a *Tweedie chain* (T -chain).

Remark. A positive operator $T \leq K_a$ such that $T(\cdot, B)$ is lower semicontinuous for all $B \in \mathcal{B}(\mathbb{X})$ and $T(x, \cdot)$ is nontrivial for some x is called a *continuous component* of K_a . Obviously, the Tweedie property is that some bounded K_a has a continuous component that is nontrivial everywhere. A continuous component T such that $T \mathbf{1}_{\mathbb{X}}$ is continuous and everywhere positive is also known as a *strong Feller operator*, as it actually maps $\mathcal{L}^\infty(\mathbb{X})$ into $\mathcal{C}(\mathbb{X}, \rho)$.

The next result is generally known. Its proof relies on the following facts: a bounded function f is lower semicontinuous if and only if there exists (uniformly) bounded continuous functions $f_n \uparrow f$ (cf. Ash (1972, Theorems A.6.4, A.6.6)), if A is open then $\mathbf{1}_A$ is lower semicontinuous, and every bounded continuous function is a limit of simple functions constant on open sets.

Proposition 1.2. *The following three statements are equivalent.*

- (i) Q is Feller.
- (ii) Q maps bounded lower semicontinuous functions to bounded lower semicontinuous functions (and, likewise, bounded upper semicontinuous functions to bounded upper semicontinuous functions).

(iii) If A is open then $Q \mathbf{1}_A$ is lower semicontinuous (and if A is closed then $Q \mathbf{1}_A$ is upper semicontinuous).

Corollary 1.1. Assume that K_a is bounded. If Q is Feller and f is bounded lower semicontinuous, then $K_a f$ is lower semicontinuous. Likewise, if Q is Feller and f is bounded upper semicontinuous, then $K_a f$ is upper semicontinuous.

In the time series scenario, we can often determine that an operator is Feller by the following method. This result is easy to show, but later we will present an analogous result (Theorem 3.2) that is useful even for threshold models.

Theorem 1.1. Suppose that μ is a measure on some space \mathbb{E} , and that w and z are measurable functions from $\mathbb{X} \times \mathbb{E}$ into \mathbb{R}_+ and \mathbb{X} , respectively. Suppose also that

$$Qf(x) = \int_{\mathbb{E}} w(x, u) f(z(x, u)) \mu(du), \quad f \in \mathcal{L}^\infty(\mathbb{X}),$$

with $\sup_{x \in \mathbb{X}} \int_{\mathbb{E}} w(x, u) \mu(du) < \infty$ and $\mu(\{u : w(x, u) > 0\}) > 0$ for each $x \in \mathbb{X}$. If

- (i) $w(\cdot, u)$ and $z(\cdot, u)$ are continuous at x , μ -almost everywhere (μ -a.e.) for each $x \in \mathbb{X}$, and
- (ii) $\{w(x, \cdot)\}_{x \in K}$ is μ -uniformly integrable for every compact $K \subset \mathbb{X}$,

then Q is Feller.

Remark. For a concrete example, think of the (possibly nonlinear) time series

$$X_t = z(X_{t-1}, e_t) := a(X_{t-1}) + b(X_{t-1})e_t,$$

$\{e_t\} \stackrel{\text{i.i.d.}}{\sim} \mu$, and let $w(x, u) = \tilde{w}(x, z(x, u))$. Then let Q be the operator defined by

$$Q(x, A) = E(\tilde{w}(x, X_1) \mathbf{1}_A(X_1) \mid X_0 = x).$$

The conditions of the theorem would be satisfied if, for example, a, b , and \tilde{w} are continuous, $\max(\|a(x)\|, \|b(x)\|) \leq c_1(1 + \|x\|)$, $0 < w(x, u) \leq c_2(1 + \|u\|)^\zeta$, $\zeta > 0$, and $E(\|e_1\|^\zeta) < \infty$.

The Feller property is a longstanding and common assumption in Markov chain literature. Unfortunately, it fails to hold for threshold models such as (1.7) due to the inherent discontinuities across thresholds. Often, however, this problem can be circumvented by a redefinition of the process. (See Theorem 3.2 below.) In that event, it will be useful to ask what the Feller property can tell us about ϕ -irreducibility and T -continuity. (In general, the Feller property neither implies nor is implied by the Tweedie property.)

2. Continuity results

2.1. The role of compact sets

Section 2 is both a review of basic results found mostly in Meyn and Tweedie (1993) and an extrapolation of those results to more usable form. With an impressive presentation, Meyn and Tweedie (1993) show that the use of petite sets (see Definition 2.1(ii)) is optimal for establishing a drift condition for ergodicity of a Markov chain. The problem is that in order to demonstrate that compact sets in particular are petite, one is usually required first to show that the process

is Tweedie. This can be a complicated undertaking, especially as the same task may also be needed to establish ϕ -irreducibility. See, for example, Cline (2007, Theorem 5.1).

As indicated in the introduction, compact sets play a crucial role. A time series that ‘explodes’ (i.e. escapes all compact sets) obviously is not going to be ergodic, and it is just this sort of behavior that a drift condition prohibits. (A time series whose state process, in contrast, shrinks toward a lower-dimensional space frequently has other issues, such as nonirreducibility, which may be remedied by redefining the process.) Showing that compact sets are petite, therefore, is an important part of the procedure for verifying the regularity and stability of a time series model. It is important to keep in mind that we are assuming \mathbb{X} is locally compact.

Although the Tweedie property is exactly the right criterion for this task, we are looking for another approach, one we hope minimizes or at least systematizes the effort required. This approach will take advantage of the Feller property and the verification of a more local problem, namely the existence of an open reachable petite set. We do this in Subsection 2.4. We also discuss some of the strategies for accomplishing this and we conclude the section with a look at aperiodicity.

2.2. Small and petite sets

In this subsection, we recall the definitions for small and petite sets and summarize some of their properties in the context of bounded positive kernels. See Nummelin (1984) and Meyn and Tweedie (1993).

Definition 2.1. Suppose that $A \in \mathcal{B}(\mathbb{X})$.

- (i) A is called v_n -small if, for all $B \in \mathcal{B}(\mathbb{X})$, it satisfies the minorization condition

$$Q^n(\cdot, B) \geq \mathbf{1}_A(\cdot)v_n(B) \quad \text{for some } n \geq 1 \text{ and some measure } v_n. \quad (2.1)$$

- (ii) A is called v_a -petite if, for all $B \in \mathcal{B}(\mathbb{X})$, it satisfies the minorization condition

$$K_a(\cdot, B) \geq \mathbf{1}_A(\cdot)v_a(B) \quad \text{for some bounded } K_a \text{ and some measure } v_a. \quad (2.2)$$

Obviously, small sets are also petite. Likewise, it is clear that every subset of a small (petite) set is small (petite). It is useful to note that (2.1) and (2.2) need only to be checked for open B . In fact, if \mathbb{X} is an open or closed subset of \mathbb{R}^m , under the Euclidean metric, then they need only be checked for open *rectangles*. See Lemma 5.1 and Lemma 5.3.

A key result, noted in Proposition 5.2.4(i) and Proposition 5.5.4(i) of Meyn and Tweedie (1993) is the following.

Proposition 2.1. (i) If A is v_n -small and $Q^m(x, A) \geq \delta > 0$ for all $x \in B$, then B is v_{n+m} -small with $v_{n+m} = \delta v_n$.

(ii) If A is v_a -petite and $K_b(x, A) \geq \delta > 0$ for all $x \in B$, then B is v_{a*b} -petite with $v_{a*b} = \delta v_a$.

The existence of nontrivial small or petite sets (apparently) is not generally assured. We do have the following, however, from Proposition 2.6 of Nummelin (1984) or Theorems 5.2.2 and 5.2.4(iii) of Meyn and Tweedie (1993).

Proposition 2.2. If Q is ϕ -irreducible and ϕ is maximal, then every ϕ -positive set contains a ϕ -positive small subset. Also, every ϕ -positive small set is both v_n -small and v_n -positive for some n and v_n .

We are interested here in the consequences of the existence of a *reachable* petite set. The following characterization is surprising in its simplicity and yet apparently is a new result.

Theorem 2.1. *There exists a reachable petite set if and only if Q is ϕ -irreducible. In this case, if ϕ is maximal then every reachable set is ϕ -positive and contains a ϕ -positive small set.*

Theorem 2.1 is verified in Section 5. It is part of the key to our stated purpose of finding a localized criterion.

2.3. Tweedie chains

Meyn and Tweedie (1993) looked at the relationship between the Tweedie property (T -continuity) and other properties for a stochastic kernel Q . Though some of their proofs require a close look for bounded positive kernels, the results all extend.

This first result is proven for bounded positive kernels exactly as it is for stochastic kernels, using Meyn and Tweedie (1993, Propositions 6.2.1, 6.2.3, and 6.2.4).

Proposition 2.3. *Suppose that Q is Tweedie and that a reachable point x_* exists. Then Q is ϕ -irreducible.*

The next, very important, result is a consequence of relating the kernel Q to the topology of \mathbb{X} (cf. Meyn and Tweedie (1993, Theorem 6.2.5(i), using Propositions 6.2.3 and 6.2.4)).

Proposition 2.4. *If every compact set is petite then Q is Tweedie.*

A partial converse to Proposition 2.4 also holds. The proof is based on Meyn and Tweedie (1993, Theorem 6.2.5(ii)), which relies on Proposition 2.2, and on Meyn and Tweedie (1993, Theorem 6.2.6), and may be used here without modification. Note that Proposition 2.3 is used for the case where Q is assumed open set irreducible.

Proposition 2.5. *If Q is either ϕ -irreducible or open set irreducible and Q is Tweedie, then every compact set is petite.*

Indeed, our primary interest is in showing that compact sets are petite as this is precisely what we need (in a time series setting) to establish a drift condition for ergodicity. However, Proposition 2.5 requires both ϕ -irreducibility and T -continuity. (In fact, one method (cf. Proposition 2.3) for showing ϕ -irreducibility requires first showing T -continuity.) We would like to simplify how this may be done.

We summarize with the following result.

Theorem 2.2. *The following statements are equivalent.*

- (i) Q is ϕ -irreducible and every compact set is petite.
- (ii) Q is ϕ -irreducible and Tweedie.
- (iii) Q is Tweedie and a reachable point exists.

Before moving on, we consider the existence of *open* petite sets. The proof of this is given in Section 5.

Theorem 2.3. *Suppose that Q is ϕ -irreducible and Tweedie. Then a ϕ -positive open (and reachable) petite set exists.*

2.4. Feller chains

Meyn and Tweedie (1993) also looked at Feller Markov chains that are ϕ -irreducible. The next two results again extend to bounded positive operators. See Lemma 6.2.7 and

Proposition 6.2.8, respectively, of Meyn and Tweedie (1993). Note that the support of ϕ is defined to be the complement of the union of all ϕ -null open sets.

Proposition 2.6. *If Q is ϕ -irreducible and Feller, then the closure of every petite set is petite.*

Proposition 2.7. *If Q is ϕ -irreducible and Feller, and if either an open ϕ -positive petite set exists or the support of ϕ has a nonempty interior, then every compact set is petite and Q is Tweedie.*

Note that Proposition 2.7 is a partial converse to Theorem 2.3. Part of the assertion in Proposition 2.7 relies on the following corollary to Proposition 2.2 and Proposition 2.6. See the proof of Proposition 6.2.8(ii) in Meyn and Tweedie (1993).

Proposition 2.8. *If Q is ϕ -irreducible and Feller and the support of ϕ has a nonempty interior, then an open ϕ -positive petite set exists.*

Remark. From Lemma 1.1 we see that (when Q is ϕ -irreducible) the support of ϕ has a nonempty interior if and only if there exists an open set of reachable points. However, the existence of such a set is not necessary for either Feller or Tweedie operators that are not ϕ -irreducible, as illustrated by the simple example $Q\mathbf{1}_B(x) = \mathbf{1}_B(0)$, $x \in \mathbb{X} = [-1, 1]$, $B \in \mathcal{B}([-1, 1])$.

The next lemma is analogous to Proposition 2.3 above.

Lemma 2.1. *Suppose that Q is Feller and that a reachable point x_* exists. If A is open and $R_\lambda(x_*, A) > 0$, then A is reachable.*

In particular, Q is open set irreducible if and only if $R_\lambda(x_, A) > 0$ for all nonempty open A .*

Again, our main interest here is to determine when every compact set is petite for a Feller chain without initially verifying that it is either a Tweedie chain or ϕ -irreducible, as is required for both Proposition 2.5 and Proposition 2.7. Thus, we have the following characterization of an important subclass of Feller operators. As before, the proof is given in Section 5.

Theorem 2.4. *Suppose that Q is Feller. Then the following statements are equivalent.*

- (i) *Q is ϕ -irreducible and Tweedie (and every compact set is petite).*
- (ii) *There exist a reachable point x_* and an open petite set A such that $R_\lambda(x_*, A) > 0$.*
- (iii) *There exists an open, reachable petite set.*

The key to applying Theorem 2.4 is in verifying either (ii) or (iii). Showing that a point or set is reachable is best done by a method we call *tunneling*.

Lemma 2.2. *Assume that Q is Feller. Suppose that K_{a_1}, \dots, K_{a_n} and x_0, x_1, \dots, x_n are such that $K_{a_i}(x_{i-1}, A) > 0$ for all open A containing x_i , $i = 1, \dots, n$. Let $b = a_1 * \dots * a_n$. Then, given any neighborhood A_n of x_n , there exists a neighborhood A_0 of x_0 such that $x \in A_0$ implies that $K_b(x, A_n) > 0$.*

In particular, suppose that x_ is fixed. If, for each $x \in \mathbb{X}$, there exists x_0, x_1, \dots, x_n as above with $x_0 = x$ and $x_n = x_*$, then x_* is reachable.*

To show that a set A is small (and, hence, petite), it obviously suffices to show that there exists a measure ν , $n \geq 1$, and $D \in \mathcal{B}(\mathbb{X})$ such that $Q^n(x, D) \geq \int_D g(x, y)\nu(dy)$ with g bounded away from 0 on $A \times D$. This entails finding the density g , however, so it may be preferable to accomplish it with the following result.

Lemma 2.3. *Suppose that \mathbb{X} is a locally compact subset of \mathbb{R}^m . Let $\tilde{\nu}$ be a measure on $\mathcal{B}(\mathbb{R})$, and let $c_i, d_i \in \mathbb{R}$ be such that $(c_1, d_1) \times \cdots \times (c_m, d_m) \subset \mathbb{X}$ and $\tilde{\nu}((c_i, d_i)) > 0$ for each $i = 1, \dots, m$. If there exists $n \geq 1$ such that*

$$Q^n(x, (a_1, b_1) \times \cdots \times (a_m, b_m)) \geq \prod_{i=1}^m \tilde{\nu}((a_i, b_i)) \quad \text{for all } a_i, b_i \in [c_i, d_i] \text{ and } x \in A, \quad (2.3)$$

then A is ν_n -small (and, hence, petite).

The intervals (c_i, d_i) in Lemma 2.3 may be chosen as small as one likes, subject to the constraint $\tilde{\nu}((c_i, d_i)) > 0$.

2.5. Aperiodicity

Although aperiodicity is not required for a stationary distribution to exist (in the Markov setting), it is required for ergodicity. As in the previous subsections, we wish to approach it in the context of a Feller chain.

Definition 2.2. *Suppose that Q is ϕ -irreducible. If there exist a ϕ -positive set A , n_1, n_2 with $\gcd(n_1, n_2) = 1$, and positive c such that A is ν_{n_i} -small, $i = 1, 2$, with $\nu_{n_2} = c\nu_{n_1}$, then Q is called *aperiodic*.*

The following theorem shows how the Feller property can be used to help verify aperiodicity.

Theorem 2.5. *Assume that Q is Feller, and that B is open and ν_n -small.*

- (i) *Suppose that there exist $x_* \in \mathbb{X}$ and n_1, n_2 with $\gcd(n_1, n_2) = 1$ such that $Q^{n_i}(x_*, B) > 0$, $i = 1, 2$. Then there exist an open ν_{n_1} -small and ν_{n_2} -small set with $\nu_{n_2} = c\nu_{n_1}$ for some $c > 0$.*
- (ii) *If $x_* \in B$, it suffices to have $n_1 = n$ and $Q^{n_2}(x_*, B) > 0$ with $\gcd(n, n_2) = 1$.*
- (iii) *If, in addition to (i) or (ii), x_* is reachable then Q is aperiodic, ϕ -irreducible, and Tweedie.*

For an aperiodic chain, all petite sets are small (see Meyn and Tweedie (1993, Theorem 5.5.7)).

We conclude this section with a summarization that characterizes a very workable strategy for the time series models we are interested in. Here we assume that \mathbb{X} is a closed (or at least locally compact) subset of \mathbb{R}^m . Let \mathcal{A}_δ be the open rectangles with sides of length δ and intersected with \mathbb{X} .

Theorem 2.6. *Suppose that \mathbb{X} is a locally compact subset of \mathbb{R}^m and that Q is Feller. Assume that there exist a reachable point x_* , a neighborhood B of x_* , n_1, n_2 with $\gcd(n_1, n_2) = 1$, a measure ν that is nontrivial on B , and $\delta > 0$ such that*

$$Q^{n_1}(x_*, B) > 0 \quad \text{and} \quad Q^{n_2}(y, A) \geq \nu(A) \quad \text{for all } A \in \mathcal{A}_{\delta_1}, \delta_1 \leq \delta, y \in B. \quad (2.4)$$

Then Q is aperiodic, ϕ -irreducible, and Tweedie.
 The converse is also true (even if Q is not Feller).

3. Threshold models

3.1. Extended state space

In Section 2 we discussed many results, not all new, in a fairly general context of bounded linear operators, with an emphasis on those having the Feller property. In this section we will focus more narrowly on operators that are inherently *not* Feller but have sufficient structure that their continuity properties may be assessed. Specifically, we are interested in threshold operators such as those derived from (1.1) and (1.2). The objective of the section is to identify Feller-like characteristics that may be exploited for use in other results. This will be a staged undertaking as we first extend the state space to accommodate the thresholds and regimes (Subsection 3.1), then we provide a condition for a threshold chain to be embedded in a Feller chain on the extended state space (Subsection 3.2) and, finally, we demonstrate that many threshold models are in fact Tweedie chains (Subsections 3.3 and 3.4). All the results for this section are proven in Subsection 5.2. As before, we will provide results valid for bounded positive operator chains and not just for Markov chains.

Recall that the threshold nature of the model in (1.2) is determined by the regimes C_1, \dots, C_N , which are connected sets with nonempty interiors, and by continuous functions $h_j(x, u)$, $j = 1, \dots, N$. Since $z(x, u) = \sum_{j=1}^N h_j(x, u) \mathbf{1}_{C_j}(x)$ is merely piecewise continuous, it is easy to see that the Markov transition operator is usually not Feller. This complicates the use of results such as Theorem 2.4. One might hope somehow to apply the results to the interior of the regimes, but this is not completely helpful because compact subsets of the interior are bounded away from the thresholds and our ultimate goal is to identify conditions for which subsets with compact closure in the full state space are petite.

Still, the piecewise continuity and the interior behavior should count for something. Accordingly, we will explore how to take advantage of the piecewise continuity in order to reformulate the process as a Feller chain, at least when it is appropriately well behaved. To this end, let $C_1^\circ, \dots, C_N^\circ$ be the interiors, called *components*, of C_1, \dots, C_N . We are interested in behavior that is continuous on $\mathbb{X}_\# = \bigcup_{j=1}^N C_j^\circ$ but not on \mathbb{X} itself. On the other hand, this behavior will often be somewhat regular (e.g. uniformly continuous) near the boundaries. Before we can properly describe what we mean by this, we must first identify a useful subclass of $\mathcal{C}(\mathbb{X}_\#, \rho)$ and we change the metric.

It is important to note that we require \mathbb{X} to be complete under the metric ρ in this section.

Definition 3.1. Suppose that (\mathbb{X}, ρ) is a complete separable, locally compact metric space and that $\mathbb{X}_\#$ is an open subset of \mathbb{X} .

- (i) Define $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ to be the collection of bounded real-valued functions on $\mathbb{X}_\#$ that are uniformly continuous on $K \cap \mathbb{X}_\#$ for each ρ -compact $K \subset \mathbb{X}$.
- (ii) For $x_1, x_2 \in \mathbb{X}_\#$, define the metric

$$\rho_\#(x_1, x_2) = \begin{cases} \rho(x_1, x_2) & \text{if there is a connected subset of } \mathbb{X}_\# \text{ containing both } x_1 \\ & \text{and } x_2, \\ \rho(x_1, x_2) + 1 & \text{otherwise.} \end{cases}$$

Remark. It is easily seen that $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ is complete under the norm $\| \cdot \|$, and, therefore, is a Banach space.

It is a mouthful, but one may say that a function in $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ is ρ -locally uniformly $\rho_\#$ -continuous. Although the metric $\rho_\#$ separates points that are on opposite sides of a boundary,

convergence (within a component) under $\rho_{\#}$ is the same as it was with ρ . To complete $(\mathbb{X}_{\#}, \rho_{\#})$, however, we must embed it in a larger space. Note that $\mathbb{X}_{\#}$ is the union of at most countably many disjoint, open, connected sets, say $\{C_1^{\circ}, C_2^{\circ}, \dots\}$. By identifying $\mathbb{X}_{\#}$ with $\bigcup_j (C_j^{\circ} \times \{j\})$, we may obtain its closure under $\rho_{\#}$. Formally, we have the following theorem, which not only shows how to complete $\mathbb{X}_{\#}$, but also helps to characterize $\mathcal{C}_u(\mathbb{X}_{\#}, \rho)$.

Theorem 3.1. *Suppose that $\{C_1^{\circ}, C_2^{\circ}, \dots\}$ is a collection of disjoint, open, connected subsets of a complete separable, locally compact metric space (\mathbb{X}, ρ) , $\mathbb{X}_{\#} = \bigcup_j C_j^{\circ}$, and $\rho_{\#}$ is the metric defined in Definition 3.1. We make the obvious identification of $\mathbb{X}_{\#}$ with $\hat{\mathbb{X}}_{\#} = \bigcup_j (C_j^{\circ} \times \{j\})$.*

- (i) *The completion of $\mathbb{X}_{\#}$ under $\rho_{\#}$ may be identified with $\hat{\mathbb{X}} = \bigcup_j (\bar{C}_j \times \{j\})$, where \bar{C}_j is the ρ -closure of C_j° within \mathbb{X} . We can interpret $\hat{\mathbb{X}}$ as a closed subset of $\mathbb{X} \times \mathbb{Z}$, with metric $\hat{\rho}_{\#}((x_1, j_1), (x_2, j_2)) = \rho(x_1, x_2) + |j_1 - j_2|$.*
- (ii) *Treating $(\hat{\mathbb{X}}_{\#}, \rho_{\#})$ as a subspace of $(\hat{\mathbb{X}}, \hat{\rho}_{\#})$, every function in $\mathcal{C}_u(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$ has a unique extension to a function in $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_{\#})$, and we may equate the Banach spaces $\mathcal{C}_u(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$ and $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_{\#})$.*
- (iii) *$\mathcal{C}_u(\mathbb{X}_{\#}, \rho) \subset \mathcal{C}_u(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$ (with the obvious identification of functions on $\mathbb{X}_{\#}$ and $\hat{\mathbb{X}}_{\#}$), and if each ρ -compact $K \subset \mathbb{X}$ intersects only finitely many of $\{C_1^{\circ}, C_2^{\circ}, \dots\}$ then $\mathcal{C}_u(\mathbb{X}_{\#}, \rho) = \mathcal{C}_u(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$.*

In summary, if $\mathbb{X}_{\#}$ has finitely many components then $\mathcal{C}_u(\mathbb{X}_{\#}, \rho)$ is effectively the same as $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_{\#})$.

The point of Theorem 3.1 is that the boundary of C_j° remains separate from the boundary of C_k° in $\hat{\mathbb{X}}$, if $j \neq k$, even though they may overlap in \mathbb{X} . This separation enables us to extend the well-behaved functions in $\mathcal{C}_u(\mathbb{X}_{\#}, \rho)$ to continuous functions on $\hat{\mathbb{X}}$. We may use the notation a bit loosely, sometimes equating $(\mathbb{X}_{\#}, \rho_{\#})$ with $(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$. Likewise, we will refer interchangeably to functions on the two spaces, and we will use x to denote an element of $\hat{\mathbb{X}}$, with the understanding that it has different identities on the boundaries depending on which C_j° it is associated with. If need be, one may denote such identities by $x^{(j)}$. For the sake of clarity, we summarize the metric spaces:

- (\mathbb{X}, ρ) = complete, separable metric space,
- $(\mathbb{X}_{\#}, \rho_{\#})$ = interior of \mathbb{X} with metric that separates components,
- $(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$ = same as $(\mathbb{X}_{\#}, \rho_{\#})$, but with points labeled by the component they are in,
- $(\hat{\mathbb{X}}, \hat{\rho}_{\#})$ = completion of $(\hat{\mathbb{X}}_{\#}, \hat{\rho}_{\#})$.

Remark. As completeness is a property of the metric, not just of the topology, the results of the next two subsections will depend on the choice of the metric ρ .

3.2. Feller chains on $\hat{\mathbb{X}}$

The state process for most threshold models is not Feller, due to the discontinuous behavior of the process near the thresholds. If, however, the behavior *within* each regime is uniformly continuous near the thresholds then there is some hope of sufficient continuity as long as the state process cannot go directly to the thresholds. This is the case for most (if not nearly all) common models.

In what follows, \mathbb{X} , $\mathbb{X}_\#$, $\hat{\mathbb{X}}$, and $\hat{\mathbb{X}}_\#$ are as in the previous subsection. Also, recall that a collection $\{f_\gamma\}_{\gamma \in \Gamma}$ is *uniformly equicontinuous* with respect to $\rho_\#$ if, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that $\rho_\#(x_1, x_2) < \delta$ implies that $\sup_{\gamma \in \Gamma} |f_\gamma(x_1) - f_\gamma(x_2)| < \varepsilon$.

Theorem 3.2. *Suppose that μ is a finite measure on some space \mathbb{E} , w and z are measurable functions from $\mathbb{X} \times \mathbb{E}$ into \mathbb{R}_+ and \mathbb{X} , respectively, and that*

$$Qf(x) = \int_{\mathbb{E}} w(x, u) f(z(x, u)) \mu(du), \quad f \in \mathcal{L}^\infty(\mathbb{X}). \tag{3.1}$$

Assume also that $\|Q \mathbf{1}_{\mathbb{X}}\| = \sup_{x \in \mathbb{X}} \int_{\mathbb{E}} w(x, u) \mu(du) < \infty$, $\mu(\{u : w(x, u) > 0\}) > 0$ for each $x \in \mathbb{X}$, and $\mu(\{u : z(x, u) \notin \mathbb{X}_\#\}) = 0$ for all $x \in \mathbb{X}_\#$.

If, for every ρ -compact $K \subset \mathbb{X}$, there exist $B_n \uparrow B$, $\mu(B^c) = 0$, such that

- (i) $\{w(\cdot, u)\}_{u \in B_n}$ and $\{z(\cdot, u)\}_{u \in B_n}$ are uniformly equicontinuous on $(K \cap \mathbb{X}_\#, \rho_\#)$,
- (ii) $\{w(x, \cdot)\}_{x \in K \cap \mathbb{X}_\#}$ is μ -uniformly integrable, and
- (iii) for all n ,

$$\lim_{\delta \downarrow 0} \sup_{\substack{\rho_\#(x_1, x_2) < \delta \\ x_1, x_2 \in K \cap \mathbb{X}_\#}} \mu(\{u : \rho_\#(z(x_1, u), z(x_2, u)) \geq 1\} \cap B_n) = 0,$$

then Q maps $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ into $\mathcal{C}_u(\mathbb{X}_\#, \rho)$. In this case, we will say that Q is \mathcal{C}_u -Feller on $(\mathbb{X}_\#, \rho_\#)$.

Moreover, if each ρ -compact $K \subset \mathbb{X}$ intersects only finitely many of $\{C_1^\circ, C_2^\circ, \dots\}$ then the restriction of Q to $\mathbb{X}_\#$ has a unique extension to a Feller operator \hat{Q} on $\mathcal{L}^\infty(\hat{\mathbb{X}})$. This operator may be computed as follows. If $\hat{\rho}_\#(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$ with $x_n \in \hat{\mathbb{X}}_\#$ then (again identifying $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ with $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$)

$$\hat{Q}^k f(x) = \lim_{n \rightarrow \infty} Q^k f(x_n) \quad \text{for all } f \in \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#), k \geq 0. \tag{3.2}$$

The representation in (3.2) may not hold for $f \notin \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$.

Remark. Condition (i) of Theorem 3.2 says that $z(\cdot, u)$ is $\rho_\#$ -continuous as a function into (\mathbb{X}, ρ) , but it is not generally true (even μ -a.e.) that it may be interpreted as a $\rho_\#$ -continuous function from $(\mathbb{X}_\#, \rho_\#)$ into $(\hat{\mathbb{X}}, \hat{\rho}_\#)$. It is even more futile to hope that $z(\cdot, u)$ may be extended to a function that is μ -a.e. continuous from $(\hat{\mathbb{X}}, \hat{\rho}_\#)$ into $(\hat{\mathbb{X}}, \hat{\rho}_\#)$. That would require at least

$$\lim_{\delta \downarrow 0} \mu\left(\left\{u : \sup_{\substack{\rho_\#(x_1, x_2) < \delta \\ x_1, x_2 \in K \cap \mathbb{X}_\#}} \rho_\#(z(x_1, u), z(x_2, u)) \geq 1\right\} \cap B_n\right) = 0, \tag{3.3}$$

which usually cannot hold unless z is already continuous. (An exception would be if z maps each C_j° into a single C_k° , μ -a.e.)

Also, it is tempting to try to give an integral representation to \hat{Q} , for example, by applying Theorem 3.1(ii) to $w(\cdot, u)$ and $z(\cdot, u)$, and substituting the extensions into (3.1). But this is not generally possible, now because f need not be ρ -continuous. Consequently, Theorem 3.2 cannot be proven simply by combining Theorems 1.1 and 3.1.

On the other hand, the conditions of Theorem 3.2 essentially say that $z(\cdot, u)$ and $w(\cdot, u)$ may be considered to be continuous functions into $(\hat{\mathbb{X}}, \hat{\rho}_\#)$ *in measure*. This is enough for the result to hold.

The following is an immediate consequence of Theorem 3.2.

Theorem 3.3. *For the scenario described in Subsection 1.2, including conditions (i)–(iii), Q is \mathcal{C}_u -Feller under the Euclidean metric and its restriction to $\mathbb{X}_\#$ has a unique extension to a Feller operator \hat{Q} on $\hat{\mathbb{X}}$.*

Three additional corollaries complete this subsection.

Corollary 3.1. *Make the same assumptions as in Theorem 3.2, and let $K \subset \mathbb{X}$ be ρ -compact. Suppose that $\{f_\gamma\}_{\gamma \in \Gamma}$ is uniformly equicontinuous on $(K \cap \mathbb{X}_\#, \rho_\#)$ and that $\sup_{\gamma \in \Gamma} \|f_\gamma\| < \infty$. Then $\{Qf_\gamma\}_{\gamma \in \Gamma}$ is uniformly equicontinuous on $(K \cap \mathbb{X}_\#, \rho_\#)$.*

Corollary 3.2. *Assume that K_a is bounded. If Q is \mathcal{C}_u -Feller then $K_a(\cdot, A)$ is lower semicontinuous on $\mathbb{X}_\#$ for all open $A \subset \mathbb{X}_\#$.*

As a consequence of Corollary 3.2, some of the results about Feller operators in Sections 1 and 2 may be restated for \mathcal{C}_u -Feller operators. In particular, the following result is helpful.

Theorem 3.4. *Suppose that $\{Q^n\}$ is a \mathcal{C}_u -Feller chain. Then Lemma 2.2 and Theorem 2.5 apply to the restriction of Q to $\mathbb{X}_\#$.*

However, as compact subsets of $\mathbb{X}_\#$ are bounded away from the thresholds, the restated results still fall short of our goal to identify conditions that imply the subsets with compact closure in \mathbb{X} are petite.

3.3. Continuous threshold chains

The conditions in Theorem 3.2 are purposely stated in terms of $(\mathbb{X}_\#, \rho_\#)$ because we really want to restrict the chain to $\mathbb{X}_\#$. (Recall that we assume that $z(x, u) \in \mathbb{X}_\#, \mu$ -a.e.) For this to be useful, however, we also need to know that $\mathbb{X}_\#^c$ is not absorbing. Hence, we have the following lemma. First, some notation. We set

$$\begin{aligned} z_1(x, u_1) &= z(x, u_1), & w_1(x, u_1) &= w(x, u_1), \\ z_k(x, u_1, \dots, u_k) &= z(z_{k-1}(x, u_1, \dots, u_{k-1}), u_k), & k &\geq 2, \\ w_k(x, u_1, \dots, u_k) &= w(z_{k-1}(x, u_1, \dots, u_{k-1}), u_k) w_{k-1}(x, u_1, \dots, u_{k-1}), & k &\geq 2, \end{aligned}$$

and $\mu_k = \mu \times \dots \times \mu$ to be the product measure on \mathbb{E}^k . Thus,

$$Q^k f(x) := \int_{\mathbb{E}^k} w_k(x, u_1, \dots, u_k) f(z_k(x, u_1, \dots, u_k)) \mu_k(du_1 \dots du_k).$$

We also make the obvious definition $\hat{R}_\lambda = \sum_{n=0}^\infty \lambda^{-n-1} \hat{Q}^n$.

Lemma 3.1. *Make the same assumptions as in Theorem 3.2, and suppose that \mathbb{X} is the ρ -closure of $\mathbb{X}_\#$. Define $d(x) = \inf_{x' \in \mathbb{X}_\#^c} \rho(x, x')$, the ‘distance’ from x to $\mathbb{X}_\#^c$. If, for each $x \notin \mathbb{X}_\#$ and each $\rho_\#$ -Cauchy sequence in $\mathbb{X}_\#$ such that $\rho(x_n, x) \rightarrow 0$, there exist $\delta > 0$ and $k \geq 1$ so that*

$$\liminf_{n \rightarrow \infty} \mu_k(\{u: w_k(x_n, u_1, \dots, u_k) > \delta, d(z_k(x_n, u_1, \dots, u_k)) > \delta\}) > 0, \tag{3.4}$$

then $\hat{R}_\lambda(x, \hat{\mathbb{X}}_\#) > 0$ for all $x \in \hat{\mathbb{X}}$.

A condition such as the one in Lemma 3.1 is needed because, in general, it is not easy to compare $\hat{R}_\lambda(x, \hat{\mathbb{X}}_\#)$ to $\liminf_{\rho(x_n, x) \rightarrow 0} R_\lambda(x_n, \mathbb{X}_\#)$ and, moreover, the latter may not be helpful because $R_\lambda(x_n, \mathbb{X}_\#)$ need only be lower semicontinuous. Indeed, the conclusion of Lemma 3.1

need not hold. Consider the situation with the usual metric on $\mathbb{X} = \mathbb{R}$, $\mathbb{X}_\# = (\infty, 0) \cup (0, \infty)$, and $z(x, u) = xu$. Set $f_\lambda(y) = \min(|y|/\lambda, 1)$, with $\lambda > 0$. Clearly,

$$\hat{Q} f_\lambda(0) = \lim_{x \rightarrow 0} \int_{\mathbb{R}} w(x, u) \min(|xu|/\lambda, 1) \mu(du) = 0 \quad \text{for all } \lambda > 0,$$

implying that $\hat{Q}(0, \hat{\mathbb{X}}_\#) = \lim_{\lambda \downarrow 0} \hat{Q} f_\lambda(0) = 0$, and, thus, $\hat{R}_\lambda(0, \hat{\mathbb{X}}_\#) = 0$.

Definition 3.2. We say that $\{Q^n\}$ is a *continuous threshold chain* on (\mathbb{X}, ρ) if there exist disjoint, open connected sets $C_1^\circ, \dots, C_N^\circ$ such that Q satisfies the conditions of Theorem 3.2 and Lemma 3.1 with $\mathbb{X}_\# = \bigcup_{j=1}^N C_j^\circ$.

Remark. The choice for $C_1^\circ, \dots, C_N^\circ$ is not necessarily the obvious one from the piecewise continuity of $z(\cdot, u)$. For example, in models with a delay, the regime containing the new observation $X_t = z(X_{t-1}, e_t)$ is known in advance from the value of X_{t-1} , meaning that we must account for the regimes of the process at earlier times. Satisfying all the conditions of Theorem 3.2, as well as the condition of Lemma 3.1, thus requires carefully splitting regimes. See the examples below.

Theorem 3.5. *For the scenario described in Subsection 1.2, including conditions (i)–(iv), $\{Q^n\}$ is a continuous threshold chain under the Euclidean metric on \mathbb{X} .*

Theorem 3.2, Lemma 3.1, and Definition 3.2 lead immediately to the following theorem.

Theorem 3.6. *If $\{Q^n\}$ is a continuous threshold chain on (\mathbb{X}, ρ) then Q is \mathcal{C}_u -Feller on $(\mathbb{X}_\#, \rho_\#)$ and Q may be extended to an operator \hat{Q} that is Feller on $(\hat{\mathbb{X}}, \hat{\rho}_\#)$ with the property that $\hat{R}_\lambda(x, \hat{\mathbb{X}}_\#) > 0$ for all $x \in \hat{\mathbb{X}}$.*

By identifying the conditions needed to embed the chain into a Feller chain on $\hat{\mathbb{X}}$, we have actually better organized the method for verifying when compact sets are petite or for verifying a drift condition (since even that necessitates some continuity). See Cline (2007). We turn now to the question of compact sets being petite.

3.4. Tweedie chains on $\hat{\mathbb{X}}$

We may now state the definitive, key result of the paper, as it identifies the petite sets for the restriction of Q to $\mathbb{X}_\#$.

Theorem 3.7. *Suppose that $\{Q^n\}$ is a continuous threshold chain and that there exists an open petite set $A \subset \mathbb{X}_\#$ with $R_\lambda(x, A) > 0$ for all $x \in \mathbb{X}_\#$. Then \hat{Q} is ϕ -irreducible and Tweedie on $(\hat{\mathbb{X}}, \hat{\rho}_\#)$, and every $\hat{\rho}_\#$ -compact set is petite.*

In this case, we say that Q is \mathcal{C}_u -Tweedie on $(\mathbb{X}_\#, \rho_\#)$, and we observe that Q is $\phi_\#$ -irreducible on $\mathbb{X}_\#$ for some measure $\phi_\#$, and the petite sets in $\mathbb{X}_\#$ include all sets whose ρ -closure in \mathbb{X} is compact.

Remark. The \mathcal{C}_u -Tweedie property on $(\mathbb{X}_\#, \rho_\#)$ is stronger than the Tweedie property on $(\mathbb{X}_\#, \rho_\#)$, when Q is ϕ -irreducible, because compact sets in the latter are bounded away from the boundaries. The \mathcal{C}_u -Tweedie property is useful in a time series setting because we want to include sets with compact closure that touch the boundaries. One advantage to defining the \mathcal{C}_u -Feller and \mathcal{C}_u -Tweedie properties on $(\mathbb{X}_\#, \rho_\#)$ is that we are allowed to restrict the chain to $\mathbb{X}_\#$.

Returning once again to the scenario in Subsection 1.2, we see that both irreducibility and continuity (in terms of a minorization condition on compact sets) boil down to finding an open petite set that is accessible from any state in $\mathbb{X}_\#$.

4. Examples

4.1. An unusual bilinear model

This example is simple, but we use it to help clarify several points about Theorem 3.2, as well as to demonstrate how to check the continuous threshold conditions. The model is given by (1.2) with $\mathbb{X} = \mathbb{E} = \mathbb{R}$, $\mathbb{X}_\# = (-\infty, 0) \cup (0, \infty)$, and

$$z(x, u) = \begin{cases} h_1(x, u) := \min(xu, a_1 + b_1u + c_1xu) & \text{if } x \leq 0, \\ h_2(x, u) := \min(xu, a_2 + b_2u + c_2xu) & \text{if } x > 0. \end{cases}$$

We assume that $a_j \neq 0$ and $b_j \neq 0$, $j = 1, 2$, and that $e_t \stackrel{\text{i.i.d.}}{\sim} \mu$, where μ has a bounded positive density. Here we let Q be the probability transition kernel for the model, so $w(x, u) \equiv 1$ and

$$Q(x, B) = \mu(\{u : z(x, u) \in B\}) = P(z(x, e_1) \in B).$$

This fits the situation discussed at the end of Subsection 1.2 so we will check the four conditions described there. Clearly, condition (i) holds, and $P(h_j(x, e_1) = 0) = 0$ if $x \neq 0$, so condition (ii) holds.

Next, we demonstrate condition (iii). Fix $M < \infty$, and let $L = \max(|c_1|, |c_2|, 1)$. Suppose that $\rho_\#(x_1, x_2) < \delta \leq (1/4ML) \min_{j=1,2} |a_j|$. That is, $|x_1 - x_2| < \delta$ and $x_1x_2 > 0$. Assume that $(-1)^j x_1 > 0$. Then we are concerned with the event $h_j(x_1, e_1)h_j(x_2, e_1) < 0$, $|e_1| \leq M$. However, this event can only occur when both $x_1e_1 > 0, x_2e_1 > 0$ and $(a_j + b_je_1 + c_jx_1e_1)(a_j + b_je_1 + c_jx_2e_1) < 0$ (and $c_j \neq 0$). Without loss, suppose that $c_jx_1e_1 < c_jx_2e_1$. The event becomes

$$a_j + b_je_1 + c_jx_1e_1 < 0 < a_j + b_je_1 + c_jx_2e_1, \quad x_1e_1 > 0, \quad x_2e_1 > 0, \quad |e_1| \leq M. \tag{4.1}$$

Since $\delta \leq |a_j|/4ML$, the event is not possible if $\min_{i=1,2} |b_j + c_jx_i| < |a_j|/2M$. Otherwise, it implies that e_1 is contained in a specific interval of length at most $2\delta LM$. We thus have

$$P(h_j(x_1, e_1)h_j(x_2, e_1) < 0, |e_1| \leq M) \leq \sup_{y \in \mathbb{R}} P(y \leq e_1 \leq y + 2\delta LM) \quad \text{if } \rho_\#(x_1, x_2) < \delta,$$

and this can be made arbitrarily small since μ has a bounded density. So condition (iii) holds. On the other hand, if $e_1(a_j + b_je_1) \neq 0$ then there always exist x_1, x_2 (depending on e_1) with $\rho_\#(x_1, x_2) < \delta$ such that (4.1) holds and, hence, $\rho_\#(z(x_1, e_1), z(x_2, e_1)) \geq 1$. Thus, (3.3) cannot be true.

Finally, $|x| < \delta/L$ implies that

$$P(|z(x, e_1)| > \delta) \geq P(\min(a_1 + b_1e_1, a_2 + b_2e_1) < -2\delta) > 0,$$

which verifies (1.5) with $k = 1$. In conclusion, we have shown that this is a continuous threshold model.

Now we look closely at the operator \hat{Q} . Although it is not generally necessary to evaluate it, we want to see why \hat{Q} must be defined by computation (3.2). Let 0^- and 0^+ be the two identities for 0, according to which component interval it is associated with. We can represent the extended

state space by $\hat{\mathbb{X}} = (-\infty, 0^-] \cup [0^+, \infty)$. For $\lambda > 0$, let $f_\lambda(x) = (1 - x/\lambda) \mathbf{1}_{(0,\lambda)}(x)$. Then f_λ is $\rho_\#$ -continuous on $\mathbb{X}_\#$ and its extension to $\hat{\mathbb{X}}$ is $f_\lambda^*(x) = (1 - x/\lambda) \mathbf{1}_{[0^+, \lambda)}(x)$. By dominated convergence and the continuity of μ ,

$$\begin{aligned} \hat{Q} f_\lambda^*(0^+) &= \lim_{x \downarrow 0} Q f_\lambda(x) \\ &= \lim_{x \downarrow 0} \mathbb{E} \left(\left(1 - \frac{1}{\lambda} \min(xe_1, a_2 + b_2e_1 + c_2xe_1) \right) \mathbf{1}_{\{0 < \min(xe_1, a_2 + b_2e_1 + c_2xe_1) < \lambda\}} \right) \\ &= \mathbb{P}(e_1 > 0, a_2 + b_2e_1 > 0). \end{aligned}$$

As this is the same for all λ , we may deduce that $\hat{Q}(0^+, (0, \infty)) = 0$ and $\hat{Q}(0^+, \{0^+\}) = \mathbb{P}(e_1 > 0, a_2 + b_2e_1 > 0)$. Likewise, let $g_\lambda(x) = (1 + x/\lambda) \mathbf{1}_{(-\lambda, 0)}(x)$. Then we may compute

$$\begin{aligned} \hat{Q}(0^+, \{0^-\}) &= \lim_{\lambda \downarrow 0} \lim_{x \downarrow 0} Q g_\lambda(x) \\ &= \lim_{\lambda \downarrow 0} \left(\mathbb{E} \left(\left(1 + \frac{1}{\lambda} (a_2 + b_2e_1) \right) \mathbf{1}_{\{-\lambda < a_2 + b_2e_1 < 0\}} \right) + \mathbb{P}(e_1 < 0, a_2 + b_2e_1 > 0) \right) \\ &= \mathbb{P}(e_1 < 0, a_2 + b_2e_1 > 0), \end{aligned}$$

while a similar argument gives

$$\hat{Q}(0^+, (y_1, y_2)) = \mathbb{P}(y_1 < a_2 + b_2e_1 < y_2) \quad \text{for } y_1 < y_2 < 0.$$

On the other hand, suppose that we define the extension of $z(\cdot, u)$ to $\hat{\mathbb{X}}$ in the obvious way, so that $z^*(0^+, u) = \min(0, a_2 + b_2u)$. It would be a mistake to equate $\hat{Q}(0^+, \{0^+\})$ either with

$$\lim_{\lambda \downarrow 0} \mathbb{E}(f_\lambda(z^*(0^+, e_1))) = \lim_{\lambda \downarrow 0} \mathbb{E}(f_\lambda(\min(0, a_2 + b_2e_1))) = 0$$

or with

$$\lim_{\lambda \downarrow 0} \mathbb{E}(f_\lambda^*(z^*(0^+, e_1))) = \lim_{\lambda \downarrow 0} \mathbb{E}(f_\lambda^*(\min(0, a_2 + b_2e_1))) = \mathbb{P}(a_2 + b_2e_1 > 0).$$

4.2. The threshold GARCH model with delay

From the previous subsection, it is clear that checking condition (iii) in Subsection 1.2 (equivalently, condition (iii) of Theorem 3.2) is likely to be more involved than checking the other conditions. In addition, it is sometimes necessary to redefine the regimes so that the conditions will hold. This is usually necessary in the case of a model with delay, as we now demonstrate.

Let b_{ji} and c_{ji} be nonnegative constants such that $b_{j0} > 0$, $b_{j1} + \dots + b_{jp} > 0$, and $c_{j1} + \dots + c_{jq} > 0$, $j = 1, 2$. For $x = (x_1, \dots, x_{p+q}) \in \mathbb{X} = \mathbb{R}^p \times \mathbb{R}_+^q$, define

$$\sigma(x) = \sum_{j=1}^2 \left(\mathbf{1}_{\{(-1)^j x_m > 0\}} \left(b_{j0} + \sum_{i=1}^p b_{ji} x_i^2 + \sum_{i=1}^q c_{ji} x_{p+i}^2 \right)^{1/2} \right).$$

A threshold GARCH(p, q) time series with delay $m \leq p$ is given by

$$v_t = \sigma_t e_t, \quad \sigma_t = \sigma(v_{t-1}, \dots, v_{t-p}, \sigma_{t-1}, \dots, \sigma_{t-q}),$$

where $\{e_t\}$ is an i.i.d. sequence of random variables on $\mathbb{E} = \mathbb{R}$ with continuous distribution μ . We assume without any loss that $\mu(\{u : |u| > 1\}) > 0$. (We will make additional assumptions about the parameters and μ later.)

The delay refers to the fact that the choice of regime 1 or 2 depends only on v_{t-m} . Let

$$z(x, u) = (\sigma(x)u, x_1, \dots, x_{p-1}, \sigma(x), x_{p+1}, \dots, x_{p+q-1}),$$

where $x = (x_1, \dots, x_{p+q})$. For this model, the state vector is

$$X_t = z(X_{t-1}, e_t) = (v_t, \dots, v_{t-p+1}, \sigma_t, \dots, \sigma_{t-q+1}).$$

Because of the delay (in the case $m > 1$), additional thresholds are needed to express the model in the form required for our results. We define new regime components

$$C_j^\circ = \{x \in \mathbb{X} : (-1)^j x_i > 0, i = 1, \dots, m, \text{ and } x_i > 0, i = p + 1, \dots, p + q\}$$

for $j = (j_1, \dots, j_m) \in \{1, 2\}^m$, and set $\mathbb{X}_\# = \bigcup_{j \in \{1, 2\}^m} C_j^\circ$. Of course, for $x \in C_j^\circ$, the choice of parameters depends only on j_m . That is, $x \in C_j^\circ$ implies that

$$\sigma(x) = v_{j_m}(x) := \left(b_{j_m 0} + \sum_{i=1}^p b_{j_m i} x_i^2 + \sum_{i=1}^q c_{j_m i} x_{p+i}^2 \right)^{1/2}$$

and

$$z(x, u) = h_j(x, u) := (v_{j_m}(x)u, x_1, \dots, x_{p-1}, v_{j_m}(x), x_{p+1}, \dots, x_{p+q-1}).$$

Also, suppose that $w(x, u) = (1 + |z(x, u)|)^\zeta / (1 + |x|)^\zeta$ and $E(|e_1|^\zeta) < \infty, \zeta \geq 0$. Define Q by

$$Qf(x) = E\left(\frac{(1 + |X_1|)^\zeta}{(1 + |x|)^\zeta} f(X_1) \mid X_0 = x \right) = \int_{\mathbb{R}} w(x, u) f(z(x, u)) \mu(du). \tag{4.2}$$

As in the previous example, we need only check the conditions described at the end of Subsection 1.2. Condition (i) obviously holds. Let $L_1 \geq 1$ be as in (1.3).

Clearly, if $x \in \mathbb{X}_\#$ then $z(x, u) \notin \mathbb{X}_\#$ if and only if $u = 0$, which is a μ -null event. This is condition (ii). Additionally, we note that $\rho_\#(x, x') < 1$ means that the first m components of x and x' have the same sign. Hence, $\rho_\#(x, x') < 1/L_1 \leq 1$ implies that $\rho_\#(z(x, u), z(x', u)) < 1$, μ -a.e. since, except when $u = 0$, $\sigma(x)u$ and $\sigma(x')u$ must have the same sign while the next $m - 1$ components of $z(x, u)$ and $z(x', u)$ do by assumption. Therefore, condition (iii) holds and Q is \mathcal{E}_u -Feller on $(\mathbb{X}_\#, \rho_\#)$.

Next, the distance from x to the nearest threshold or boundary is

$$d(x) = \min(|x_1|, \dots, |x_m|, x_{p+1}, \dots, x_{p+q}).$$

Note that $\sigma(x) \geq \tilde{b} := \min(b_{10}^{1/2}, b_{20}^{1/2}) > 0$ for all x . Let $k = \max(p, q)$. It is now easy to see that $|u_i| > 1, i = 1, \dots, k$, implies that $d(z_k(x, u_1, \dots, u_k)) \geq \tilde{b}$, no matter how small $d(x)$ is. This verifies that condition (iv) holds. Thus, $\{Q^n\}$ is a continuous threshold chain.

Now we make further assumptions: μ has a positive bounded density, and

$$b_{j_m p} > 0, \quad c_{j_m q} > 0, \quad \text{and} \quad c_{j_m 1} + \dots + c_{j_m q} < 1, \quad \text{for either } j_m = 1 \text{ or } j_m = 2. \tag{4.3}$$

Without loss of generality, we assume that the above holds with $j_m = 2$. For any $x \in \mathbb{X}$,

$$P(|v_1 - 1| < \delta \mid X_0 = x) = P(|\sigma(x)e_1 - 1| < \delta) > 0$$

for every $\delta > 0$, implying by iteration that $P(\max_{i \leq p} |v_i - 1| < \delta \mid X_0 = x) > 0$. Let $\sigma_*^2 = (b_{20} + \dots + b_{2p}) / (1 - c_{21} - \dots - c_{2q})$ and $x_* = (1, \dots, 1, \sigma_*, \dots, \sigma_*)$ (which is a

solution of $x = z(x, 1/\sigma(x))$. Note that

$$|\sigma^2(x) - \sigma_*^2| \leq (b_{21} + \dots + b_{2p}) \max_{i \leq p} |x_i^2 - 1| + (c_{21} + \dots + c_{2p}) \max_{p < i \leq p+q} |x_i^2 - \sigma_*^2|.$$

Since $c_{21} + \dots + c_{2p} < 1$ and setting $\delta = \varepsilon/2(b_{21} + \dots + b_{2p})$, these facts mean that there is always some n (depending on x and ε) such that

$$P(\|X_k - x_*\| < \varepsilon \mid X_0 = x) > 0 \quad \text{for all } k \geq n.$$

Let $A_\varepsilon = \{x' \in \mathbb{X}_\# : \|x' - x_*\| < \varepsilon\}$, $\varepsilon > 0$. Noting that $w(x, y)$ is positive, it follows that $Q^k(x, A_\varepsilon) > 0$ for all $k \geq n$. Therefore, every neighborhood of x_* is reachable.

Now let μ_+ be the subprobability measure that is μ restricted to \mathbb{R}_+ and define Q_+ just as Q is in (4.2) but with μ_+ replacing μ . Observe that Q_+ is a kernel for $\mathcal{L}(\mathbb{R}_+^{p+q})$. Moreover, Q_+ is Feller by Theorem 1.1.

Since $Q \geq Q_+$, any subset of \mathbb{R}_+^{p+q} that is small for Q_+ will also be small for Q . Condition (4.3), with $j_m = 2$, suffices for $Q_+^{p+2q}(x, \cdot)$ to be absolutely continuous with a transition density $q_{p+2q}(x, y)$ that is locally bounded away from 0 on some $A_\varepsilon \times A_\varepsilon$. See the related argument in the proof of Cline (2007, Theorem 5.2). Define, with $k = p + 2q$,

$$\nu_k(B) = \int_{B \cap A_\varepsilon} \min_{x \in A_\varepsilon} q_k(x, y) \, dy \quad \text{for } B \in \mathcal{B}(\mathbb{R}_+^{p+q})$$

and $\nu_{k+1} = \min_{x \in A_\varepsilon} Q_+(x, A_\varepsilon) \nu_k$. It follows immediately that A_ε is both ν_k -small and ν_{k+1} -small. (See also Theorem 2.5 applied to Q_+ .) In particular, Q has an open, reachable petite set.

We conclude, finally, that Q is ϕ -irreducible, aperiodic, and \mathcal{C}_u -Tweedie under the Euclidean metric.

4.3. The collapsed threshold GARCH model

Let X_t^* be the state vector for a threshold GARCH process as in the previous subsection, but assume here that $b_{01} = b_{02} = 0$. We also define $w(x, u) = \|z(x, u)\|$. The Markov chain

$$\theta_t = \frac{z(\theta_{t-1}, e_t)}{w(\theta_{t-1}, e_t)} = \frac{X_t^*}{\|X_t^*\|}$$

is known as the *collapsed* threshold GARCH model, and it exists on the unit sphere Θ in $\mathbb{R}^p \times \mathbb{R}_+^q$. This process is of interest for studying ergodicity and other properties of the ordinary threshold GARCH model $\{X_t\}$ (with $b_{01} > 0$ and $b_{02} > 0$). Again, see Cline (2007). Specifically, we look at the kernel

$$Q_*(\theta, B) = E(w(\theta_0, e_1) \mathbf{1}_{\{\theta_1 \in B\}} \mid \theta_0 = \theta) \quad \text{for } B \in \mathcal{B}(\Theta).$$

Since $w(\cdot, u)$ is only piecewise continuous, it is necessary to check the conditions in Theorem 3.2 and Lemma 3.1 in order to verify that Q_* is ϕ -irreducible, aperiodic, and \mathcal{C}_u -Tweedie. However, the argument is similar to that above, with the restriction to the unit sphere compensating for the lack of intercept parameters b_{01} and b_{02} .

4.4. The threshold AR model with threshold GARCH errors

The time series (1.7) combines a threshold GARCH model with a threshold autoregression. Ordinarily, the thresholds (and regimes) for both depend only on the values (signs) of

$\xi_{t-1}, \dots, \xi_{t-p}$. Nevertheless, the method in Subsection 4.2 may again be used. The principle complication is in choosing an appropriate reachable point. The specifics involve considerable algebra, in general, and we leave them for the reader to work out.

There are many variations of threshold models for which the results in this section may apply. A certain amount of algebra is unavoidable and, as yet, the models still must be considered case by case. It is hoped, however, that our approach will lead to further results encompassing larger classes of models.

5. Proofs

5.1. Verifying the Tweedie property

In this subsection we gather the arguments leading to Theorem 2.4, and to the results describing a practical strategy.

Proof of Lemma 1.1. This follows from Proposition 1.1. Let ψ be defined as in (1.8) (with ϕ_1 being any finite measure equivalent to ϕ). If A is reachable then

$$\psi(A) = \int_{\mathbb{X}} R_\lambda(x, A)\phi_1(dx) > 0$$

and, hence, $\phi(A) > 0$, since ϕ is maximal.

If x is in the support then every neighborhood of x is ϕ -positive and, hence, x is reachable. On the other hand, if x is not in the support then some neighborhood A of x is ϕ -null. By the first part of the proof, x cannot be reachable.

The next result is useful to note.

Lemma 5.1. *Let \mathcal{A} be a (countable) base of open sets for \mathbb{X} that is closed under finite unions. Suppose that ν and $\tilde{\nu}$ are finite measures on $\mathcal{B}(\mathbb{X})$ such that $\nu(A) \geq \tilde{\nu}(A)$ for all $A \in \mathcal{A}$. Then $\nu(B) \geq \tilde{\nu}(B)$ for all $B \in \mathcal{B}(\mathbb{X})$.*

Proof. If A is any open set then there exists $A_n \in \mathcal{A}$, $n \geq 1$, such that $A = \bigcup_{n=1}^\infty A_n$. In fact, since \mathcal{A} is closed under finite unions, we may assume without loss that $A_n \uparrow A$. Monotone convergence immediately gives us $\nu(A) \geq \tilde{\nu}(A)$.

Now let \mathcal{O} denote the collection of open sets. Since the measures are finite and regular, we conclude that

$$\nu(B) = \inf_{\{A \in \mathcal{O} : B \subset A\}} \nu(A) \geq \inf_{\{A \in \mathcal{O} : B \subset A\}} \tilde{\nu}(A) = \tilde{\nu}(B) \quad \text{for any } B \in \mathcal{B}(\mathbb{X})$$

(cf. Ash (1972, Corollary 4.3.7)).

For the following, represent $x \in \mathbb{X}$ by $x = (x_1, \dots, x_m)$. We define the open rectangles in $\mathbb{X} \subset \mathbb{R}^m$,

$$A_{\delta,x} = \left\{ y \in \mathbb{X} : \max_{i \leq m} |y_i - x_i| < \frac{\delta}{2} \right\} \quad \text{for } x \in \mathbb{X},$$

and the half-open rectangles,

$$H_{\delta,x} = \left\{ y \in \mathbb{X} : -\frac{\delta}{2} < y_i - x_i \leq \frac{\delta}{2}, i = 1, \dots, m \right\} \quad \text{for } x \in \mathbb{X}.$$

Also, let $\mathcal{A}_\delta = \{A_{\delta,x} : x \in \mathbb{X}\}$ and $\mathcal{H}_\delta = \{H_{\delta,x} : x \in \mathbb{X}\}$.

Lemma 5.2. *Suppose that $\mathbb{X} \subset \mathbb{R}^m$, with \mathbb{X} locally compact, and suppose that $C \subset \mathbb{X}$ is compact. Let ν be a finite measure on \mathbb{X} . Then, for each $\varepsilon > 0$ and $\delta > 0$, there exist $A_1, \dots, A_n \in \mathcal{A}_{\delta_1}$ such that $\delta_1 < \delta$, $C \subset \bigcup_{i=1}^n A_i$, and $\sum_{i=1}^n \nu(A_i) < \nu(C) + \varepsilon$.*

Remark. If \mathbb{X} is an open or closed subset of \mathbb{R}^m (and \mathbb{R}^m is locally compact under the metric) then \mathbb{X} also is locally compact (see Ash (1972, Corollary 8.3)).

Proof of Lemma 5.2. Choose open A containing C such that $\nu(A) < \nu(C) + \varepsilon/2$. Note that \mathbb{X} can be partitioned by sets in \mathcal{H}_δ . Indeed, since C is compact and A is open, we can choose $\delta_2 < \delta$ so small that there exist disjoint $H_1, \dots, H_n \in \mathcal{H}_{\delta_2}$ with $C \subset \bigcup_{i=1}^n H_i \subset A$. Next, we choose $\delta_1 \in (\delta_2, \delta)$ and $A_i \in \mathcal{A}_{\delta_1}$ such that $H_i \subset A_i$ and $\nu(A_i) < \nu(H_i) + \varepsilon/2n$. It follows that $C \subset \bigcup_{i=1}^n A_i$ and

$$\sum_{i=1}^n \nu(A_i) < \sum_{i=1}^n \nu(H_i) + \frac{\varepsilon}{2} \leq \nu(A) + \frac{\varepsilon}{2} < \nu(C) + \varepsilon,$$

as desired.

Lemma 5.3. *Assume that $\mathbb{X} \subset \mathbb{R}^m$ with \mathbb{X} locally compact. Suppose that ν and $\tilde{\nu}$ are finite measures on \mathbb{X} . Fix $\delta > 0$. If*

$$\nu(B) \geq \tilde{\nu}(B) \quad \text{for all } B \in \mathcal{A}_{\delta_1}, \delta_1 \leq \delta,$$

then

$$\nu(B) \geq \tilde{\nu}(B) \quad \text{for all } B \in \mathcal{B}(\mathbb{X}).$$

Proof. Suppose that there exists open $B \subset \mathbb{X}$ such that $\nu(B) < \tilde{\nu}(B)$. Then, for some $\varepsilon > 0$, there exists compact $C \subset B$ such that $\tilde{\nu}(C) > \nu(B) + \varepsilon$. By Lemma 5.2, there exist $\delta_1 < \delta$ and $A_1, \dots, A_n \in \mathcal{A}_{\delta_1}$ such that $C \subset \bigcup_{i=1}^n A_i$ and $\sum_{i=1}^n \nu(A_i) < \nu(C) + \varepsilon$. But then

$$\tilde{\nu}(C) \leq \tilde{\nu}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \tilde{\nu}(A_i) \leq \sum_{i=1}^n \nu(A_i) < \nu(C) + \varepsilon \leq \nu(B) + \varepsilon < \tilde{\nu}(C),$$

a contradiction. Thus, $\nu(B) \geq \tilde{\nu}(B)$ for all open B .

The conclusion now follows from Lemma 5.1.

Lemma 5.4. *Suppose that A is ν_a -petite. Then there exists a measure $\nu_b \gg \nu_a$ such that A is ν_b -petite and every reachable set is ν_b -positive.*

Proof. Since A is ν_a -petite, it follows that $K_a f(x) \geq \nu_a f = \int f(y) \nu_a(dy)$, if $x \in A$, for any nonnegative bounded function f . In particular, with $\lambda > r(Q)$,

$$K_a R_\lambda(x, B) = \int R_\lambda(y, B) K_a(x, dy) \geq \int R_\lambda(y, B) \nu_a(dy) = \nu_a R_\lambda(B) \quad \text{for } x \in A,$$

which shows that A is ν_b -petite with $\nu_b = \nu_a R_\lambda$ and $K_b = K_a R_\lambda$. (Note that $b = a * c$, where $c_n = \lambda^{-n-1}$.) Observe that $\nu_b \geq \lambda^{-1} \nu_a$ so $\nu_a \ll \nu_b$.

Additionally, $\nu_b(B) = \int R_\lambda(y, B) \nu_a(dy) > 0$ whenever B is a reachable set.

Lemma 5.5. *Suppose that A is ν_a -petite and that $R_\lambda(x, A) > 0$, $\lambda > r(Q)$, for all $x \in \mathbb{X}$. Then Q is ν_b -irreducible with maximal measure $\nu_b = \nu_a R_\lambda$, and $\nu_b(A) > 0$.*

Proof. Clearly, if $\nu_a(B) > 0$ then

$$R_\lambda K_a(x, B) \geq \int_A K_a(y, B) R_\lambda(x, dy) \geq \nu_a(B) R_\lambda(x, A) > 0.$$

Thus, Q is ν_a -irreducible and, by Proposition 1.1, ν_b is a maximal irreducibility measure. By Lemma 1.1 we also have $\nu_b(A) > 0$.

We combine the last two results for the proof of Theorem 2.1.

Proof of Theorem 2.1. Suppose that A is a ν_a -petite reachable set. Then Q is ν_a -irreducible by Lemma 5.5. We note that $\nu_b = \nu_a R_\lambda$ is a maximal irreducibility measure and, thus, every reachable set is ν_b -positive by Lemma 5.4 (or by Lemma 1.1) and contains a ν_b -positive small set by Proposition 2.2. Conversely, if Q is ϕ -irreducible then Proposition 2.2 implies the existence of a ϕ -positive small (hence, petite) set.

Proof of Theorem 2.2. (i) implies (ii) by Proposition 2.4, (ii) implies (iii) by Lemma 1.1, (iii) implies (ii) by Proposition 2.3, and (ii) implies (i) by Proposition 2.5.

Proof of Theorem 2.3. Since \mathbb{X} is a locally compact separable metric space, there exists a sequence of open sets $A_n \uparrow \mathbb{X}$ with compact closure \bar{A}_n , and $\bar{A}_{n-1} \subset A_n \subset \bar{A}_n$ (see Ash (1972, Theorem A5.15)). Clearly, some \bar{A}_{n-1} is ϕ -positive. By Proposition 2.5, \bar{A}_n is petite and, therefore, A_n is both petite and ϕ -positive. By definition, it is also reachable.

Proof of Lemma 2.1. (Following the proof of Proposition 6.2.1 in Meyn and Tweedie (1993).) Observe that $\mathbf{1}_A$ is lower semicontinuous. By Corollary 1.1, $R_\lambda(x, A)$ is a lower semicontinuous function of x and, therefore, bounded away from 0 in some neighborhood B of x_* . Hence,

$$R_\lambda^2(x, A) \geq \int_B R_\lambda(y, A) R_\lambda(x, dy) > 0 \quad \text{for all } x \in \mathbb{X}.$$

Since $R_\lambda^2(x, A) > 0$ if and only if $R_\lambda(x, A) > 0$, this verifies that A is reachable.

It follows immediately that $R_\lambda(x_*, A) > 0$ for all open A implies that Q is open set irreducible, and the converse holds trivially.

Lemma 5.6. *Suppose that Q is Feller and that there exists an open petite set A . Let $B = \{x: R_\lambda(x, A) > 0\}$. If C is a compact subset of B then it is petite.*

Proof. By assumption, $R_\lambda(x, A)$ is lower semicontinuous and positive everywhere on B . Being lower semicontinuous, it achieves its minimum on compact sets (see Ash (1972, Theorem A.6.3)). Given compact $C \subset B$, let $\delta = \inf_{x \in C} R_\lambda(x, A) > 0$. Since A is petite, Proposition 2.1 shows that C is also petite.

The following, finally, is the proof of our alternative to Proposition 2.5 and Proposition 2.7.

Proof of Theorem 2.4. Even without the Feller property, (i) implies the existence of an open, reachable petite set A , by Theorem 2.3. By Lemma 1.1, we need only ϕ -irreducibility to assert the existence of a reachable point x_* . Clearly, $R_\lambda(x_*, A) > 0$ follows. Thus, (i) implies (ii).

Now assume that we have x_* and A as in (ii). Since $R_\lambda(x_*, A)$ is lower semicontinuous, there exists a neighborhood B of x_* such that $\delta = \inf_{x \in B} R_\lambda(x, A) > 0$. By Proposition 2.1(ii), B is petite. Since x_* is reachable, B is also, verifying that (ii) implies (iii).

Finally, suppose that an open, reachable petite set A exists. From Lemma 5.5 we may conclude that Q is ϕ -irreducible. In addition, we may apply Lemma 5.6 to $B = \mathbb{X}$ to conclude that every compact set is petite. By Proposition 2.4, Q is Tweedie. Hence, (iii) implies (i).

Proof of Lemma 2.2. Since Q is Feller, we may recursively choose a neighborhood A_{i-1} of x_{i-1} such that $K_{a_i}(x, A_i) > 0$ for all $x \in A_{i-1}$. Define $b_i = a_i * \cdots * a_n$. Since

$$K_{b_i}(x, A_n) \geq \int_{A_i} K_{b_{i+1}}(y, A_n) K_{a_i}(x, dy),$$

it is easy by induction (in reverse order) to show that $K_{b_i}(x, A_n) > 0$ for all $x \in A_{i-1}$, $i = 1, \dots, n$. The conclusion thus holds with $b = b_1$.

If such a sequence exists for $x_n = x_*$, $x_0 = x$ (n depending on x) then $R_\lambda(x, A) > 0$ for each neighborhood A of x_* . Hence, if this is possible for all $x \in \mathbb{X}$, we deduce that x_* is reachable.

Proof of Lemma 2.3. Let $\tilde{\nu}_m := \tilde{\nu} \times \cdots \times \tilde{\nu}$ be the product measure on \mathbb{R}^m , and define ν_n to be the restriction of $\tilde{\nu}_m$ to $D = (c_1, d_1) \times \cdots \times (c_m, d_m)$. We simply observe that (2.3) and Lemma 5.3 imply that $Q^n(x, B) \geq \nu_n(B)$ for all $B \in \mathcal{B}(\mathbb{X})$. Thus, A is ν_n -small.

Proof of Theorem 2.5. (i) Since $Q^{n_i}(\cdot, B) > 0$ is lower semicontinuous, there exists a neighborhood A of x_* such that $\delta_i = \inf_{x \in A} Q^{n_i}(x, B) > 0$ for each i . By Proposition 2.1(i), A is ν_{n_i} -small with $\nu_{n_i} = \delta_i \nu$. Thus, the conclusion holds with $c = \delta_2/\delta_1$.

(ii) The argument is the same, except now we can choose open A with $x_* \in A \subset B$ and $\delta_2 = \inf_{x \in A} Q^{n_2}(x, B) > 0$. Since A is ν_n -small, the conclusion holds with $n_1 = n$ and $c = \delta_2$.

(iii) If x_* is reachable then so is A , and, thus, Q is ϕ -irreducible and Tweedie by Theorem 2.4. By Lemma 5.5 we can choose ϕ so that A is ϕ -positive. Thus, Q is aperiodic by Definition 2.2.

Proof of Theorem 2.6. By Lemma 5.3, the inequality in (2.4) holds for all measurable A and, thus, B is ν_{n_2} -small. Hence, the first implication is true by Theorem 2.5(ii) and (iii).

For the converse, Theorem 2.2 shows that there exists a reachable point x_* and that every neighborhood of x_* with compact closure is petite. By Meyn and Tweedie (1993, Theorem 5.5.7), if A is a petite neighborhood of x_* then it is also ν_n -small for some n , and we can choose ν_n so that it is ν_n positive, by Proposition 2.2 and Theorem 2.1. Since Q is aperiodic, then there also exists a ν_{n_1} -positive, ν_{n_1} -small, and ν_{n_2} -small set A_1 with $\nu_{n_1} = \nu_{n_2}$ and $\gcd(n_1, n_2) = 1$. Moreover, we know that both ν_n and ν_{n_1} may be chosen to be maximal. Thus, A_1 is also ν_n -positive and it follows from Proposition 2.1(i) that A is ν_{n+n_1} -small and ν_{n+n_2} -small with $\nu_{n+n_1} = \nu_{n+n_2} = \nu_n(A_1) \nu_{n_1}$. Since $\gcd(n + n_1, n + n_2) = 1$, we are done.

5.2. Reformulating threshold models

We now present the proofs for the material in Section 3, where we studied continuous threshold models.

Lemma 5.7. *Let (\mathbb{X}, ρ) and $(\mathbb{Y}, \tilde{\rho})$ be complete metric spaces. Suppose that A is a subset of \mathbb{X} and that \bar{A} is its closure under ρ . Assume also that $f_\gamma: A \mapsto \mathbb{Y}$ and $\{f_\gamma\}_{\gamma \in \Gamma}$ is uniformly equicontinuous on A . Then there exists a unique extension $f_\gamma^*: \bar{A} \mapsto \mathbb{Y}$ such that $\{f_\gamma^*\}_{\gamma \in \Gamma}$ is uniformly equicontinuous on \bar{A} .*

The proof of this lemma is a straightforward exercise in convergence which we leave to the reader.

Remark. Here \bar{A} need not be compact for the result above to hold, nor does the family $\{f_\gamma\}_{\gamma \in \Gamma}$ need to be bounded. However, to be able to claim any uniform convergence, by Arzela's theorem, we would need both that \bar{A} is compact and that $\{f_\gamma\}_{\gamma \in \Gamma}$ is pointwise bounded.

If $\mathbb{Y} = \mathbb{R}$ with the usual metric then, obviously, $\sup_{x \in \bar{A}} |f_Y^*(x)| = \sup_{x \in A} |f_Y(x)|$.

Corollary 5.1. *Let (\mathbb{X}, ρ) be a complete separable, locally compact metric space, and suppose that $\mathbb{X}_\#$ is an open subset whose closure under ρ is \mathbb{X} . Then we may equate the Banach spaces $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ and $\mathcal{C}(\mathbb{X}, \rho)$.*

Proof. Let $f \in \mathcal{C}_u(\mathbb{X}_\#, \rho)$, and let K be any compact subset of \mathbb{X} , so that f is uniformly continuous on $K \cap \mathbb{X}_\#$. By Lemma 5.7, f has a unique extension to a function f^* that is continuous on $\overline{K \cap \mathbb{X}_\#}$. Since this is true for arbitrary compact K , f has a unique extension f^* , continuous on all of \mathbb{X} .

Now we simply observe that $\|f_n - f\| \rightarrow 0$ in $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ requires the sequence to be uniformly equicontinuous on $K \cap \mathbb{X}_\#$, for compact K , and, therefore, $\{f_n^*\}$ is uniformly equicontinuous on $K_1 = \overline{K \cap \mathbb{X}_\#}$, by Lemma 5.7. This implies that the sequence converges uniformly on K_1 , by Arzela’s theorem (cf. Munkres (1975, p. 279)). In particular, the limit is equal to f^* on K_1 . Not only that, but

$$\|(f_n^* - f^*)\mathbf{1}_{K_1}\| = \sup_{x \in K_1} |f_n^*(x) - f^*(x)| = \sup_{x \in K \cap \mathbb{X}_\#} |f_n(x) - f(x)| = \|(f_n - f)\mathbf{1}_{K \cap \mathbb{X}_\#}\|.$$

As K is arbitrary, it must be that

$$\|f_n^* - f^*\| = \|f_n - f\| \rightarrow 0.$$

This completes the proof.

Proof of Theorem 3.1. (i) A sequence in $\mathbb{X}_\#$ is $\rho_\#$ -Cauchy if and only if it is (eventually) contained in some C_j° and is ρ -Cauchy. Such a sequence converges to a point in \bar{C}_j and all points in \bar{C}_j are such limits. This shows that the completion of $\mathbb{X}_\#$ under $\rho_\#$ may be identified with $\hat{\mathbb{X}}$ and, correspondingly, $\rho_\#$ with $\hat{\rho}_\#$. That $\hat{\mathbb{X}}$ is closed in $\mathbb{X} \times \mathbb{Z}$ under $\hat{\rho}_\#$ is obvious.

(ii) This is proven as Lemma 5.7 and Corollary 5.1 above, with $(\hat{\mathbb{X}}, \hat{\rho}_\#)$ replacing (\mathbb{X}, ρ) and $\mathbb{Y} = \mathbb{R}$. Note that $(\hat{\mathbb{X}}, \hat{\rho}_\#)$ is locally compact and separable, as well as complete.

(iii) Suppose that \hat{K} is $\hat{\rho}_\#$ -compact in $\hat{\mathbb{X}}$. Then $\hat{K} \cap \hat{\mathbb{X}}_\#$ corresponds to a set $K \cap \mathbb{X}_\#$, where K is ρ -compact within \mathbb{X} . Any $f \in \mathcal{C}_u(\mathbb{X}_\#, \rho)$ therefore has a counterpart that is uniformly $\hat{\rho}_\#$ -continuous on $\hat{K} \cap \hat{\mathbb{X}}_\#$ and, thus, is in $\mathcal{C}_u(\hat{\mathbb{X}}_\#, \hat{\rho}_\#)$.

Now suppose that each ρ -compact $K \subset \mathbb{X}$ intersects only finitely many of $\{C_1^\circ, C_2^\circ, \dots\}$. Let $\hat{K}_\#$ be the set within $\hat{\mathbb{X}}$ that is identified with $K \cap \mathbb{X}_\#$. It is of the form $\hat{K} \cap \hat{\mathbb{X}}_\#$, where \hat{K} is the $\hat{\rho}_\#$ -closure of $\hat{K}_\#$, and it is $\hat{\rho}_\#$ -compact since only finitely many components are intersected. Thus, as in the argument above, $\mathcal{C}_u(\hat{\mathbb{X}}_\#, \hat{\rho}_\#) \subset \mathcal{C}_u(\mathbb{X}_\#, \rho)$.

Proof of Theorem 3.2. Since $\|Q\mathbf{1}_\mathbb{X}\| < \infty$ and $\mu(\{u : z(x, u) \notin \mathbb{X}_\#\}) = 0$ for all $x \in \mathbb{X}_\#$ by assumption, $Qf(x) = Q(f\mathbf{1}_{\mathbb{X}_\#})(x)$ for any $x \in \mathbb{X}_\#$. Hence, we may also think of Q as a bounded linear operator from $\mathcal{L}^\infty(\mathbb{X}_\#)$ into $\mathcal{L}^\infty(\mathbb{X}_\#)$.

The first claim is proved if we show that $Qf(\cdot)$ is uniformly continuous on $K \cap \mathbb{X}_\#$ for all $f \in \mathcal{C}_u(\mathbb{X}_\#, \rho)$ and all ρ -compact $K \subset \mathbb{X}$. Accordingly, fix $f \in \mathcal{C}_u(\mathbb{X}_\#, \rho)$, ρ -compact $K \subset \mathbb{X}$, and $\varepsilon > 0$. Let B_1, B_2, \dots be as indicated in the theorem statement. By (ii), choose $B_n \subset \mathbb{E}$ such that $\mu(B_n^c)$ is sufficiently small for

$$\sup_{x \in K \cap \mathbb{X}_\#} \int_{B_n^c} w(x, u)\mu(du) < \frac{\varepsilon}{6\|f\|}.$$

Then, if $x_1, x_2 \in K \cap \mathbb{X}_\#$,

$$\int_{B_n^c} |w(x_1, u)f(z(x_1, u)) - w(x_2, u)f(z(x_2, u))|\mu(du) < \frac{\varepsilon}{3}. \tag{5.1}$$

Now define

$$D = D(x_1, x_2) = \{u : z(x_i, u) \in \mathbb{X}_\#, i = 1, 2, \} \cap \{u : \rho_\#(z(x_1, u), z(x_2, u)) \geq 1\}.$$

By (ii) and (iii), we can choose $\delta_1 > 0$ such that $x_1, x_2 \in K_1 \cap \mathbb{X}_\#$ with $\rho_\#(x_1, x_2) < \delta_1$ implies that $\mu(B_n \cap D)$ is sufficiently small so that

$$\int_{B_n \cap D} |w(x_1, u)f(z(x_1, u)) - w(x_2, u)f(z(x_2, u))|\mu(du) < \frac{\varepsilon}{3}. \tag{5.2}$$

Next, let

$$M = \sup_{x \in \mathbb{X}_\#} \int_{\mathbb{E}} w(x, u)\mu(du),$$

and choose $\gamma > 0$ such that $x_1, x_2 \in K \cap \mathbb{X}_\#$ with $\rho_\#(x_1, x_2) < \gamma$ implies that $|f(x_1) - f(x_2)| < \varepsilon/6M$. Choose $\delta \in (0, \delta_1)$ such that (by (i)) $x_1, x_2 \in K \cap \mathbb{X}_\#$ with $\rho_\#(x_1, x_2) < \delta$ implies that

$$\sup_{u \in B_n} \rho(z(x_1, u), z(x_2, u)) < \gamma \quad \text{and} \quad \sup_{u \in B_n} |w(x_1, u) - w(x_2, u)| < \frac{\varepsilon}{6\|f\|\mu(B_n)},$$

and, hence, that $\rho_\#(z(x_1, u), z(x_2, u)) < \gamma$ when $u \in B_n \cap D^c$. Thus,

$$\begin{aligned} & \int_{B_n \cap D^c} |w(x_1, u)f(z(x_1, u)) - w(x_2, u)f(z(x_2, u))|\mu(du) \\ & \leq \int_{B_n \cap D^c} (w(x_1, u)|f(z(x_1, u)) - f(z(x_2, u))| + \|f\| |w(x_1, u) - w(x_2, u)|)\mu(du) \\ & < \int_{B_n \cap D^c} \left(\frac{\varepsilon}{6M} w(x_1, u) + \frac{\varepsilon}{6\|f\|\mu(B_n)} \|f\| \right) \mu(du) \\ & \leq \frac{1}{3}\varepsilon. \end{aligned} \tag{5.3}$$

Finally, from (5.1)–(5.3),

$$\begin{aligned} |Qf(x_1) - Qf(x_2)| & \leq \int_{\mathbb{E}} |w(x_1, u)f(z(x_1, u)) - w(x_2, u)f(z(x_2, u))|\mu(du) \\ & < \varepsilon \quad \text{for all } x_1, x_2 \in K \cap \mathbb{X}_\# \text{ such that } \rho_\#(x_1, x_2) < \delta. \end{aligned}$$

Hence, $Qf(\cdot)$ is uniformly continuous on $K \cap \mathbb{X}_\#$ and the first claim is proved.

Next, assume that each ρ -compact $K \subset \mathbb{X}$ intersects only finitely many of $\{C_1^\circ, C_2^\circ, \dots\}$ so that $\mathcal{C}_u(\mathbb{X}_\#, \rho)$ is identified with $\mathcal{C}_u(\hat{\mathbb{X}}_\#, \hat{\rho}_\#)$ by Theorem 3.1. For the remaining argument, we take this identification literally. Additionally, as long as we maintain a restriction to $\hat{\mathbb{X}}_\#$, we may treat Q as a positive kernel operator from $\mathcal{L}^\infty(\hat{\mathbb{X}}_\#)$ into $\mathcal{L}^\infty(\hat{\mathbb{X}}_\#)$.

Suppose that $f^* \in \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$. Clearly, its restriction $f = f^*$ on $\hat{\mathbb{X}}_\#$ is in $\mathcal{C}_u(\hat{\mathbb{X}}_\#, \hat{\rho}_\#)$ and, by Lemma 5.7 and Corollary 5.1, f^* is the unique extension of f to $\hat{\mathbb{X}}$. Since $Qf \in \mathcal{C}_u(\hat{\mathbb{X}}_\#, \hat{\rho}_\#)$, Qf also has a unique $\hat{\rho}_\#$ -continuous extension $(Qf)^*$ to $\hat{\mathbb{X}}$, again by Lemma 5.7

and Corollary 5.1. Define $\hat{Q}f^* = (Qf)^*$. Obviously, \hat{Q} is linear, positive, and bounded on $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$ and $\|\hat{Q}f^*\| = \sup_{x \in \hat{\mathbb{X}}_\#} |Qf(x)|$.

Now suppose that $f_m^* \in \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$, with restriction f_m , and that $f_m^* \downarrow 0$ as $m \rightarrow \infty$. For any compact $K \subset \hat{\mathbb{X}}$, we know that $f_m \downarrow 0$ uniformly on $K \cap \hat{\mathbb{X}}_\#$. By Arzela's theorem and Corollary 3.1, which depends only on the first part of this proof, $Qf_m \downarrow 0$ uniformly on $K \cap \hat{\mathbb{X}}_\#$ also. In particular, suppose that $x_n \rightarrow x \in \hat{\mathbb{X}}$, $x_n \in \hat{\mathbb{X}}_\#$. Then $Qf_m(x_n) \downarrow 0$ uniformly (in n) as $m \rightarrow \infty$, implying that $\lim_{m \rightarrow \infty} \hat{Q}f_m^*(x) = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} Qf_m(x_n) = 0$. Therefore, for each $x \in \hat{\mathbb{X}}$, the linear functional $\hat{Q}f^*(x)$ extends uniquely to a finite regular measure $\hat{Q}(x, \cdot)$ on $\hat{\mathbb{X}}$. (See, for example, Ash (1972, Theorem 4.2.9 and Corollary 4.3.7).) Moreover, by construction (using $f_{nm}^* \uparrow \mathbf{1}_{A_n}$ and $\mathbf{1}_{A_n} \downarrow \mathbf{1}_B$, with continuous f_{nm}^* , open A_n), $\hat{Q}(\cdot, B)$ is measurable for each $B \in \mathcal{B}(\hat{\mathbb{X}})$. We have thus defined a unique extension to a kernel operator on $\mathcal{L}^\infty(\hat{\mathbb{X}})$. Since A open in $\hat{\mathbb{X}}_\#$ is also open in $\hat{\mathbb{X}}$, the same construction ensures that $\hat{Q}(x, B)$ agrees with $Q(x, B)$ for $x \in \hat{\mathbb{X}}_\#, B \in \mathcal{B}(\hat{\mathbb{X}}_\#)$.

By definition, \hat{Q} maps $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$ into $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$, so it is Feller. Suppose that $f \in \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$ and $k \geq 1$. Then $\hat{Q}^{k-1}f \in \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$. Recall that $\mu(\{u : z(x, u) \notin \mathbb{X}_\#\}) = 0$ for all $x \in \hat{\mathbb{X}}_\#$. Applying this iteratively, $\hat{Q}^{k-1}f(x) = Q^{k-1}f(x)$ for all $x \in \hat{\mathbb{X}}_\#$ (letting f also denote its own restriction to $\hat{\mathbb{X}}_\#$). Hence,

$$\hat{Q}^k f(x) = \lim_{n \rightarrow \infty} Q((\hat{Q}^{k-1}f)\mathbf{1}_{\hat{\mathbb{X}}_\#})(x_n) = \lim_{n \rightarrow \infty} Q((Q^{k-1}f))(x_n) = \lim_{n \rightarrow \infty} Q^k f(x_n),$$

which is (3.2).

Proof of Corollary 3.1. The proof is the same as the first half of the theorem proof, but with $\sup_{\gamma \in \Gamma} \|f_\gamma\|$ replacing $\|f\|$.

Proof of Corollary 3.2. Suppose that A_1, A_2, \dots is a sequence of precompact open sets such that $A_n \uparrow A$. Since A_n is open, $\mathbf{1}_{A_n}$ is lower semicontinuous and there exist continuous $f_{n,i} \uparrow \mathbf{1}_{A_n}$ as $i \rightarrow \infty$. As A_n is precompact, we also have $f_{n,i} \in \mathcal{C}_u(\mathbb{X}_\#, \rho)$. Therefore, $f_i := \max_{n \leq i} f_{n,i} \in \mathcal{C}_u(\mathbb{X}_\#, \rho)$ and $f_i \uparrow \mathbf{1}_A$. Hence, each $Q^k f_i$ is continuous and $Q^k \mathbf{1}_A = \lim_{i \rightarrow \infty} Q^k f_i$ is lower semicontinuous. Furthermore, $K_a \mathbf{1}_A = \sup_n \sum_{k=0}^n a_k Q^k \mathbf{1}_A$ is lower semicontinuous.

Proof of Theorem 3.4. By Corollary 3.2, each of the open sets in the proofs of Lemma 2.2 and Theorem 2.5 may be taken to be subsets of $\hat{\mathbb{X}}_\#$.

Proof of Lemma 3.1. We need only consider $x \notin \hat{\mathbb{X}}_\#$ since, by assumption, $\hat{Q}(x, \hat{\mathbb{X}}_\#) = Q(x, \mathbb{X}_\#) > 0$ for all $x \in \hat{\mathbb{X}}_\#$. Each element in $\hat{\mathbb{X}}$ is a $\rho_\#$ -Cauchy limit of a sequence of elements in $\hat{\mathbb{X}}_\#$ and, thus, corresponds to a ρ -limit of the same sequence in \mathbb{X} . Note that this means that the sequence (ultimately) lies within some fixed C_j° . As (3.4) depends only on the values of $w(\cdot, u)$ and $z(\cdot, u)$ within C_j° , it also applies if we replace $\mathbb{X}_\#$ with $\hat{\mathbb{X}}_\#$ and ρ with $\hat{\rho}_\#$.

For the sake of clarity, we let $f \in \mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_\#)$ denote its own restriction to $\hat{\mathbb{X}}_\#$.

Assume that $\{x_n\}$ is a sequence in $\hat{\mathbb{X}}_\#$ such that $\hat{\rho}_\#(x_n, x) \rightarrow 0$, $x \notin \hat{\mathbb{X}}_\#$, and satisfies (3.4). We will show that $\hat{Q}^k(x, \hat{\mathbb{X}}_\#) > 0$, which suffices for the conclusion to hold. Let $f(y) = \min(d(y)/\delta, 1) \leq \mathbf{1}_{\hat{\mathbb{X}}_\#}(y)$, which is $\hat{\rho}_\#$ -continuous and vanishes off $\hat{\mathbb{X}}_\#$. Clearly, (3.4) implies that

$$\begin{aligned} & \lim_{n \rightarrow \infty} Q^k f(x_n) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{E}^k} w_k(x_n, u_1, \dots, u_k) \min\left(\frac{d(z_k(x_n, u_1, \dots, u_k))}{\delta}, 1\right) \mu_k(du_1 \cdots du_k) \\ &> 0. \end{aligned}$$

By (3.2),

$$\hat{Q}^k(x, \hat{\mathbb{X}}_{\#}) \geq \hat{Q}^k f(x) = \lim_{n \rightarrow \infty} Q^k f(x_n) > 0,$$

as required.

Proof of Theorem 3.7. From Theorem 3.6, we know that \hat{Q} is Feller on $\mathcal{C}(\hat{\mathbb{X}}, \hat{\rho}_{\#})$ and $\hat{R}_{\lambda}(x, \hat{\mathbb{X}}_{\#}) > 0$ for all $x \in \hat{\mathbb{X}}$. Also, $\hat{R}_{\lambda}(y, A) = R_{\lambda}(y, A) > 0$ for $y \in \hat{\mathbb{X}}_{\#}$ so that

$$\hat{R}_{\lambda}^2(x, A) \geq \int_{\hat{\mathbb{X}}_{\#}} \hat{R}_{\lambda}(y, A) \hat{R}_{\lambda}(x, dy) > 0 \quad \text{for } x \in \hat{\mathbb{X}}.$$

This shows that $\hat{R}_{\lambda}(x, A) > 0$ for all $x \in \hat{\mathbb{X}}$. By Theorem 2.4, \hat{Q} is ϕ -irreducible and T -continuous, and every $\hat{\rho}_{\#}$ -compact set is petite.

As in the proof of Theorem 3.2, ρ -compact $K_1 \subset \mathbb{X}$ and $\hat{\rho}_{\#}$ -compact $K_2 \subset \hat{\mathbb{X}}$ coincide when $K_1 \cap \mathbb{X}_{\#}$ (in \mathbb{X}) is identified with $K_2 \cap \hat{\mathbb{X}}_{\#}$ (in $\hat{\mathbb{X}}$). Thus, the petite sets for the restriction of Q to $\mathbb{X}_{\#}$ include those of the form $K_1 \cap \mathbb{X}_{\#}$.

Finally, let $\phi_{\#}$ be the measure ϕ restricted to $\mathbb{X}_{\#}$ and defined on $\mathcal{B}(\mathbb{X}_{\#})$. Then, for $B \in \mathcal{B}(\mathbb{X}_{\#})$,

$$\begin{aligned} \phi_{\#}(B) > 0 &\iff \phi(B) > 0 \\ &\implies \hat{R}_{\lambda}(x, B) > 0 \\ &\iff R_{\lambda}(x, B) > 0 \quad \text{for all } x \in \mathbb{X}_{\#}. \end{aligned}$$

Thus, Q is $\phi_{\#}$ -irreducible.

References

- ASH, R. B. (1972). *Real Analysis and Probability*. Academic Press, New York.
- CLINE, D. B. H. (2007). Stability of nonlinear stochastic recursions with application to nonlinear AR-GARCH models. *Adv. Appl. Prob.* **39**, 462–491. (Correction: **39** (2007), 1115–1116.)
- LI, C. W. AND LI, W. K. (1996). On a double-threshold autoregressive heteroscedastic time series model. *J. Appl. Econometrics* **11**, 253–274.
- LING, S. (1999). On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model. *J. Appl. Prob.* **36**, 688–705.
- LING, S. (2007). A double AR(p) model: structure and estimation. *Statistica Sinica* **17**, 161–175.
- LING, S. AND McALEER, M. (2002). Stationarity and the existence of a family of GARCH processes. *J. Econometrics* **106**, 109–117.
- LIU, J., LI, W. K. AND LI, C. W. (1997). On a threshold autoregression with conditional heteroscedastic variances. *J. Statist. Planning Infer.* **62**, 279–300.
- LU, Z. (1996). A note on the geometric ergodicity of autoregressive conditional heteroscedasticity (ARCH) model. *Statist. Prob. Lett.* **30**, 305–311.
- MEITZ, M. AND SAIKKONEN, P. (2008). Stability of nonlinear AR-GARCH models. *J. Time Ser. Anal.* **29**, 453–475.
- MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- MUNKRES, J. R. (1975). *Topology: A First Course*. Prentice-Hall, Englewood Cliffs, NJ.
- NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-negative Operators*. Cambridge University Press.