Adv. Appl. Prob. **24**, 542–574 (1992)
Printed in N. Ireland

STABILITY OF MARKOVIAN PROCESSES I: CRITERIA FOR DISCRETE-TIME CHAINS

SEAN P. MEYN, * *University of Illinois*
R. L. TWEEDIE, ** *Bond University*

Abstract

In this paper we connect various topological and probabilistic forms of stability for discrete-time Markov chains. These include tightness on the one hand and Harris recurrence and ergodicity on the other. We show that these concepts of stability are largely equivalent for a major class of chains (chains with continuous components), or if the state space has a sufficiently rich class of appropriate sets ('petite sets').

We use a discrete formulation of Dynkin's formula to establish unified criteria for these stability concepts, through bounding of moments of first entrance times to petite sets. This gives a generalization of Lyapunov–Foster criteria for the various stability conditions to hold. Under these criteria, ergodic theorems are shown to be valid even in the non-irreducible case. These results allow a more general test function approach for determining rates of convergence of the underlying distributions of a Markov chain, and provide strong mixing results and new versions of the central limit theorem and the law of the iterated logarithm.

IRREDUCIBLE MARKOV CHAINS; STOCHASTIC LYAPUNOV FUNCTIONS; ERGODICITY; RECURRENCE; CENTRAL LIMIT THEOREM; LAW OF THE ITERATED LOGARITHM; STRONG MIXING

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60J10
SECONDARY 34D25

1. Introduction and background

1.1. *Structure of the paper.* The paper is structured into nine main sections. In this introduction, we first introduce a number of related topological and probabilistic concepts of stability for a Markov chain $\mathbf{\Phi}$. The term 'stability' is not commonly used in the Markov chain literature. In dynamical systems literature, it is commonly used to mean 'asymptotic stability', i.e. convergence of sample paths to a fixed, stable, point. With Markovian systems, convergence is most likely in a distributional sense, and we use the term here to denote a range of behaviours, all of them implying the system is to some degree stable. In general terms, we say the chain $\mathbf{\Phi}$ is

Received 27 March 1990; revision received 14 May 1991.
* Postal address: Coordinated Science Laboratory, University of Illinois, 1101 West Springfield Ave, Urbana, IL 61801, USA.
** Present address: Department of Statistics, Colorado State University, Fort Collins, Colorado, CO 80523, USA.
This work was begun while the first author was visiting Bond University and developed there, at the Australian National University, and the University of Illinois. Research supported in part by the NSF initiation grant No ECS 8910088.

topologically stable if there is a positive probability that it does not leave the compact centre of the space (which we call 'non-evanescence'), or, using a stronger condition, if the distributions of $\mathbf{\Phi}$ as time evolves are tight (which we call 'bounded in probability'); we say the chain is probabilistically stable if $\mathbf{\Phi}$ returns to sets of positive measure ('Harris recurrence') or if there is a unique invariant probability measure for $\mathbf{\Phi}$ ('positive Harris recurrence'). Conceptually, these concepts are very similar, and most have been widely studied. We study conditions under which the various stability concepts are equivalent or related.

In Section 2 we identify a class of chains for which the various concepts are closely related: these '*T-chains*' are those which admit an everywhere non-trivial continuous component T, as defined in [33]. For T-chains, we show that a topological stability condition (the *tightness* hypothesis of [5]) is equivalent to a probabilistic stability condition which is a generalization of positive Harris recurrence. Specifically, if the underlying distributions are not tight, then we show that the trajectories of the Markov chain $\mathbf{\Phi}$ are either recurrent in a strong sense, or leave every compact subset of the state space.

In Section 3, the class of *petite sets* is introduced, and their properties explored. We show that when all compact sets are petite, the two sets of stability conditions are similarly related. Further, we show that all compact sets are petite if and only if the chain is an irreducible T-chain, under a mild stability condition; whilst, for irreducible chains, all compact sets are petite if the chain is Feller and the irreducibility measure satisfies mild conditions.

In Section 4, we demonstrate that stochastic stability is intimately connected with the return times to petite sets. Section 4 is perhaps the most important in the paper. Here it is shown that various versions of probabilistic stability criteria [13], [36] are in fact special cases of a discrete form of Dynkin's formula. Exploiting this fact, we derive much more general versions of the Lyapunov–Foster criteria, which allow for time-varying test functions. These extended versions can be necessary in evaluating the behaviour of complex systems (see [21], [19], [20]).

In Sections 5 and 6 we review and strengthen the known forms of ergodic behaviour for various types of stability for irreducible chains. Lyapunov–Foster criteria are developed for the various stability concepts, and to obtain rates of convergence for the underlying distributions of $\mathbf{\Phi}$.

In Section 7, we prove a general ergodic theorem for T-chains even if the chain is not irreducible, using an improved Doeblin decomposition theorem.

Finally, the existence of a test function is used in Sections 8 and 9 to develop strong mixing results, a new version of the central limit theorem and a law of the iterated logarithm for Markov chains.

1.2. *Markov chain notation.* We concentrate on discrete time here, and deal with Markov chains in, typically, the setting discussed in [31] or [26]. The results of this work are extended to the more technically difficult case of continuous-time processes in Parts II and III of the paper [23], [24]. In contrast to the purely probabilistic contexts of [26], our results also exploit the *topology* of the space on which the chains evolve.

The following is a minimal description of the systems we study. We consider a Markov chain $\mathbf{\Phi} = \{\Phi_0, \Phi_1, \cdots,\}$ evolving on a locally compact separable metric space $\boldsymbol{X}$, whose Borel σ-algebra shall be denoted $\mathscr{B}$. We use $\boldsymbol{P}_\mu$ and $\boldsymbol{E}_\mu$ to denote probabilities and expectations conditional on Φ_0 having distribution μ, and $\boldsymbol{P}_x$ and $\boldsymbol{E}_x$ when μ is concentrated at x.

A Markov transition function is an example of a *positive kernel* $K = \{K(x, A) : x \in \boldsymbol{X}, A \in \mathscr{B}\}$, which is a measurable function of its first argument for any measurable set A, and a σ-finite measure of its second argument for any point x. For two kernels K and S, the product KS is a kernel which is defined for $x \in \boldsymbol{X}$ and $A \in \mathscr{B}$ by

$$KS(x, A) = \int K(x, dy)S(y, A).$$

We assume that the transition probabilities of $\mathbf{\Phi}$ are defined by a Markov transition function denoted by $P = \{P(x, A), x \in \boldsymbol{X}, A \in \mathscr{B}\}$, where the iterates P^k, $k \in \mathbb{Z}_+$, are defined inductively by

$$P^0 \triangleq I, \qquad P^k \triangleq PP^{k-1}, \qquad k \geqq 1;$$

here I denotes the identity transition function defined for $x \in \boldsymbol{X}$ and $A \in \mathscr{B}$ by $I(x, A) = \mathbf{1}_A(x)$. For each $k, j \in \mathbb{Z}_+$, $x \in \boldsymbol{X}$ and $A \in \mathscr{B}$ we have

$$\boldsymbol{P}_x\{\Phi_{k+j} \in A \mid \sigma\{\Phi_0, \cdots, \Phi_k\}\} = P^j(\Phi_k, A) \quad \text{a.s. } [\boldsymbol{P}_x].$$

The transition law P acts on measurable functions f by the operation

$$Pf(x) \triangleq \int P(x, dy)f(y), \qquad x \in \boldsymbol{X}.$$

1.3. *Probabilistic stability*. In this paper we consider the probabilistic stability conditions of Harris recurrence and positive Harris recurrence, which is often also called (under slight extra conditions) Harris ergodicity. To define these, we need a little further notation. Letting I_B denote the kernel defined as $I_B(x, A) = \mathbf{1}_{A \cap B}(x)$, we write $\boldsymbol{P}_x\{\tau_A = k\} = [(PI_{A^c})^{k-1}P](x, A)$, where the stopping time τ_A is defined for a set $A \in \mathscr{B}$ by $\tau_A = \inf\{k \geqq 1 : \Phi_k \in A\}$. For $x \in \boldsymbol{X}$, $A \in \mathscr{B}$ we let

$$\begin{aligned} L(x, A) &= \boldsymbol{P}_x\{\tau_A < \infty\} \\ &= \boldsymbol{P}_x\{\mathbf{\Phi} \text{ enters } A\}; \\ Q(x, A) &= \boldsymbol{P}_x\left\{\bigcap_{N=1}^{\infty} \bigcup_{k=N}^{\infty} \{\Phi_k \in A\}\right\} \\ &= \boldsymbol{P}_x\{\mathbf{\Phi} \in A \text{ i.o.}\}; \end{aligned}$$

$$G(x, A) = \sum_{k=1}^{\infty} P^k(x, A)$$
$$= \boldsymbol{E}_x\Big[\sum_{k=1}^{\infty} \mathbf{1}\{\Phi_k \in A\}\Big].$$

Standard identities involving these quantities may be found in [26], [29]. In particular, we will frequently use the following well-known result (cf. [29]).

Lemma 1.1. Suppose for some sets A and B, $\inf_{x\in A} L(x, B) > 0$. Then

$$\{\boldsymbol{\Phi} \in A \text{ i.o.}\} \subset \{\boldsymbol{\Phi} \in B \text{ i.o.}\} \quad \text{a.s. } [\boldsymbol{P}_x] \tag{1}$$

and hence $Q(z, A) \leqq Q(z, B)$, for all $z \in \boldsymbol{X}$.

Following [29], a set $A \in \mathscr{B}$ is called *inessential* if $Q(x, A) = 0$ for all $x \in \boldsymbol{X}$; otherwise it is called *essential.* If A is essential, and cannot be written as a countable union of inessential sets, then it is called *properly essential.*

We say $\boldsymbol{\Phi}$ is *φ-irreducible* if there exists a finite measure φ such that $G(x, A) > 0$ (or equivalently $L(x, A) > 0$) for all $x \in \boldsymbol{X}$ whenever $\varphi\{A\} > 0$, and in this case φ is called an *irreducibility measure.* The chain $\boldsymbol{\Phi}$ will simply be called *irreducible* when the specific irreducibility measure is irrelevant.

We investigate reducible chains in Section 2. Here we first define stability concepts for irreducible chains.

Stochastic stability condition 1. Suppose that $\boldsymbol{\Phi}$ is φ-irreducible, and $Q(x, A) \equiv 1$ whenever $\varphi\{A\} > 0$. Then $\boldsymbol{\Phi}$ is called Harris recurrent [26], [35].

A subset $B \subset \boldsymbol{X}$ is called *absorbing* if $P(x, B) = 1$ for every $x \in B$. If B is absorbing, then the Markov chain $\boldsymbol{\Phi}$ may be restricted to the set B, and B is called a *Harris set* if the restricted chain is Harris recurrent. A σ-finite measure π on $\mathscr{B}$ with the property

$$\pi\{A\} = \pi P\{A\} \triangleq \int \pi(dx) P(x, A) \quad \text{for all} \quad A \in \mathscr{B}$$

will be called *invariant.* It is called *subinvariant* if we have $\pi \geqq \pi P$. It is shown in [26], [29] that if $\boldsymbol{\Phi}$ is irreducible, then a subinvariant measure exists, and if $\boldsymbol{\Phi}$ is Harris recurrent then an essentially unique invariant measure π exists. We shall use the following characterization of π (cf. [26], Proposition 5.9).

Lemma 1.2. If $\boldsymbol{\Phi}$ is Harris recurrent with invariant probability measure π, then for any measurable A and $f \geqq 0$, with $\pi(A) > 0$,

$$\pi(f) = \int_A \pi(dy) \boldsymbol{E}_y\Big[\sum_{k=1}^{\tau_A} f(\Phi_{k-1})\Big] = \int_A \pi(dy) \boldsymbol{E}_y\Big[\sum_{k=1}^{\tau_A} f(\Phi_k)\Big]. \tag{2}$$

In particular, choosing $f = 1$, we see that π is finite if, for some A with $\infty > \pi(A) > 0$,

$$\sup_{y \in A} \boldsymbol{E}_y[\tau_A] < \infty. \tag{3}$$

If the invariant measure is finite, then it may be normalized to a probability measure, and in practice this is the main stable situation of interest.

Stochastic stability condition 2. Suppose that $\boldsymbol{\Phi}$ is Harris recurrent, and π is finite. Then Φ is called positive Harris recurrent.

The rationale for these classical recurrence conditions being thought of as stability conditions comes from Section 2 and Section 3, where we see that they are related to the following results of a topological stability nature.

1.4. *Topological stability.* With probabilistic stability one has 'finiteness' in terms of return visits to sets of positive measure of some sort, where the measure is often dependent on the chain; with topological stability the sets of interest are compact sets which are defined by the structure of the space rather than of the chain. In this section we introduce stability concepts which relate to such links between the chain and the topology on the space.

We say that a sample path of $\boldsymbol{\Phi}$ converges to ∞ (denoted $\boldsymbol{\Phi} \to \infty$) if the trajectory visits each compact set only finitely often.

Since $\boldsymbol{X}$ is locally compact and separable, it follows from Lindelöf's theorem [15] that there exists a countable collection of open precompact sets $\{O_n : n \in \mathbb{Z}_+\}$ such that

$$\{\boldsymbol{\Phi} \to \infty\} = \bigcap_{n=0}^{\infty} \{\boldsymbol{\Phi} \in O_n \text{ i.o.}\}^c.$$

In particular, then, the event $\{\boldsymbol{\Phi} \to \infty\}$ lies in $\mathscr{B}(\boldsymbol{X}^{\mathbb{Z}_+})$.

Topological stability condition 1. A Markov chain $\boldsymbol{\Phi}$ will be called non-evanescent if $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} = 0$ for each $x \in \boldsymbol{X}$.

This is a recurrence-like condition: the terminology is loosely based on Vere-Jones [38]. Note that such chains are called non-explosive by Meyn [17], and non-dissipative by Tweedie [34], but both of these terminologies have other implications in related literature.

Our second topological condition is more closely related to the distributions of $\{\Phi_k\}$ as $k \to \infty$.

Topological stability condition 2. The chain $\boldsymbol{\Phi}$ will be called bounded in probability if for each initial condition $x \in \boldsymbol{X}$ and each $\varepsilon > 0$, there exists a compact subset $K \subset \boldsymbol{X}$ such that

$$\liminf_{k \to \infty} \boldsymbol{P}_x\{\Phi_k \in K\} \geqq 1 - \varepsilon.$$

The chain $\mathbf{\Phi}$ will be called μ-bounded in probability if, for some measure μ, this condition holds for μ a.e. initial condition $x \in \mathbf{X}$.

The chain $\mathbf{\Phi}$ is bounded in probability by definition if its distributions $\{P^k\}$ are tight for each fixed initial condition: the consequences of this are well known (see [5]). The nomenclature of boundedness in probability is derived from dynamical systems literature (see [22] for a detailed discussion). The concepts appear to be relatively unexplored in discrete-time Markov chain theory, although [16] uses the idea of non-evanescence and actually calls it recurrence. These conditions, in the continuous-time context, have also been called 'recurrence' conditions: see, for example, [14]. It is this sort of plurality of nomenclature that had led us to adopt the term 'stability' to cover the various related concepts: as a justification we note that non-evanescence is related to the concept of Lagrange-stability in the dynamical systems literature [4], whilst a Markov chain induces a dynamical system whose state space consists of probability measures, and in this context tightness is equivalent to Lagrange 'stability'.

We show in the next section that non-evanescence can be closely connected with Harris recurrence under suitable conditions. Clearly, though, since the stochastic stability conditions we use require irreducibility, they are not in general equivalent to the topological conditions. In Section 2 we define a class of chains ('T-chains') for which the stability conditions are equivalent in the irreducible case, and we show that for T-chains which are reducible, the topological stability conditions imply, essentially, a decomposition of the space into a countable number of appropriately stochastically stable subsets, namely Harris sets.

In Section 3, we give a different approach. We define a class of sets ('petite sets', extending the idea of small sets of [26]) which we show to have the property that, if all compact sets are petite, then these and other stability conditions are closely related. We also indicate conditions under which all compact sets are in fact petite.

2. Continuous components, T-chains and stability

2.1. *The definition and existence of T-chains.* Here we show that, for a class of chains introduced in [33], the forms of stability above are closely related. For a probability a on $\mathbb{Z}_+$ we define the Markov transition function K_a as

$$K_a \triangleq \sum_{i=1}^{\infty} a(i) P^i. \tag{4}$$

For two probabilities a and b on $\mathbb{Z}_+$ we have $K_a K_b = K_{a \star b}$, where $a \star b$ denotes the convolution of a and b. We let $a_\varepsilon \triangleq (1-\varepsilon)^{-1}(1, \varepsilon, \varepsilon^2, \cdots)$ for $0 < \varepsilon < 1$, and $\mathbf{e}_j$ denote the probability on $\mathbb{Z}_+$ which is supported on $j \in \mathbb{Z}_+$. We let K_ε denote K_a in the specific case where $a = a_\varepsilon$, and $a \star \varepsilon$ denote $a \star a_\varepsilon$. In this situation we always assume ε is a fixed positive number. We begin with a definition.

A kernel T is called a continuous component of a function $K:(\boldsymbol{X}, \mathcal{B}) \to \mathbb{R}_+$ if

(i) For $A \in \mathcal{B}$ the function $T(\cdot, A)$ is lower semi-continuous;

(ii) For all $x \in \boldsymbol{X}$ and $A \in \mathcal{B}$, the measure $T(x, \cdot)$ satisfies $K(x, A) \geqq T(x, A)$.

The continuous component T is called non-trivial at x if $T(x, \boldsymbol{X}) > 0$.

In this section we will be particularly interested in continuous components of the Markov transition function K_a, as defined in (4).

A chain will be called a T-chain if, for some a, the Markov transition function K_a admits a continuous component T which is non-trivial for all $x \in \boldsymbol{X}$.

Three important classes of Markov chains are those evolving on a countable state space, those which are essentially random walks, and those obtained from stochastic systems: from Proposition 3.2 below we see that a Markov chain on a countable state space is always a T-chain. It is shown in [33] that random walks are T-chains if and only if the increment distribution is 'spread out' with respect to the Haar measure on the space.

If $\boldsymbol{X}$ is a smooth (C^∞) manifold and the Markov chain is generated by the *Markovian system*

$$\Phi_{k+1} = F(\Phi_k, w_{k+1}), \qquad k \in \mathbb{Z}_+ \tag{5}$$

with $\mathbf{w}$ an i.i.d. sequence, and $F: \boldsymbol{X} \times \mathbb{R}^p \to \boldsymbol{X}$ smooth, then $\boldsymbol{\Phi}$ is a T-chain if the distribution of w_0 possesses a lower semi-continuous density, and an associated control system is forward accessible (combine Theorem 2.1(iii) of [18] with Proposition 3.2). Forward accessibility is a generalization of the controllability condition from linear systems theory. Hence if $\boldsymbol{X} = \mathbb{R}^d$ and the function F appearing in (5) is linear, so that $\Phi_{k+1} = A\Phi_k + Bw_{k+1}$ for matrices A and B of appropriate dimension, then $\boldsymbol{\Phi}$ is a T-chain if the *controllability matrix* $[A^{d-1}B|\cdots|AB|\,B]$ has rank d. In practice, then, many chains of interest are indeed T-chains, and the results below can be widely applied.

We first show that for a T-chain, either sample paths converge to ∞ or they enter a 'stochastically recurrent' part of the space. For a set $A \in \mathcal{B}$, we let $A^0 = \{x \in \boldsymbol{X} : L(x, A) = 0\}$. It is easy to show that the set A^0 is either empty or absorbing.

Proposition 2.1. Suppose that $\boldsymbol{\Phi}$ is a T-chain. If $A \in \mathcal{B}$ is not properly essential, then for each $x \in \boldsymbol{X}$, $\boldsymbol{P}_x\{\{\boldsymbol{\Phi} \to \infty\} \cup \{\boldsymbol{\Phi} \text{ enters } A^0\}\} = 1$.

Proof. Let $A = \bigcup B_j$, with each B_j inessential; then from Lemma 1.1, the sets $B_{ij} = \{x \in \boldsymbol{X} : L(x, B_i) > j^{-1}\}$ are also inessential, for any i, j. Thus $A^{0^c} = \bigcup A_i$ where each A_i is inessential. Since T is lower semi-continuous, the sets $O_{ij} \triangleq \{x \in \boldsymbol{X} : T(x, A_i) > j^{-1}\}$ are open, as is $U_j \triangleq \{x \in \boldsymbol{X} : T(x, A^0) > j^{-1}\}$, $i, j \in \mathbb{Z}_+$. Since T is everywhere non-trivial we have for all $x \in \boldsymbol{X}$, $T\left(x, \left(\bigcup A_j\right) \cup A^0\right) = T(x, \boldsymbol{X}) > 0$ and hence the sets $\{O_{ij}, U_j\}$ form an open cover of $\boldsymbol{X}$.

Let C be a compact subset of $\boldsymbol{X}$, and $\{U_N, O_{iN}: 1 \leqq i \leqq N\}$ a finite subcover. Since A_i is inessential, and

$$L(x, A_i) \geqq K_a(x, A_i) \geqq T(x, A_i) \geqq j^{-1} \qquad x \in O_{ij} \tag{6}$$

we may conclude from (1) that each of the sets O_{ij} is inessential. It follows that, with probability 1, every trajectory that enters C infinitely often must enter U_N infinitely often. That is, $\{\boldsymbol{\Phi} \in C \text{ i.o.}\} \subset \{\boldsymbol{\Phi} \in U_N \text{ i.o.}\}$ a.s. But since $L(x, A^0) > 1/N$ for $x \in U_N$ we have by (1) again, $\{\boldsymbol{\Phi} \in U_N \text{ i.o.}\} \subset \{\boldsymbol{\Phi} \in A^0 \text{ i.o.}\}$ a.s., and this completes the proof.

Corollary. If $\boldsymbol{\Phi}$ is a non-evanescent T-chain, then $\boldsymbol{X}$ is properly essential.

In fact, as we now show, a much more explicit statement linking recurrence to non-evanescence is possible for T-chains.

2.2. *The Doeblin decomposition.* It is known that T-chains admit a version of the celebrated Doeblin decomposition theorem, as shown in [33]: we now strengthen this result, linking it to our topological stability conditions. The Doeblin decomposition breaks $\boldsymbol{X}$ into a countable union of Harris sets as defined in Section 1.3, and a countable union of *inessential* sets.

Theorem 2.1 (decomposition theorem). Suppose that $\boldsymbol{\Phi}$ is a T-chain. Then

(i) $\boldsymbol{X}$ can be decomposed into the disjoint union $\boldsymbol{X} = \sum_{i \in I} H_i + E$ where the index set I is countable, each H_i is a Harris set, and E is not properly essential;

(ii) For each compact set $C \subset \boldsymbol{X}$, $H_i \cap C = \varnothing$ for all but a finite number of $i \in I$.

(iii) For each initial condition $x \in \boldsymbol{X}$,

$$\boldsymbol{P}_x\left\{\{\boldsymbol{\Phi} \to \infty\} \cup \left\{\boldsymbol{\Phi} \text{ enters } \sum_{i \in I} H_i\right\}\right\} = 1,$$

and hence the index set I is non-empty if and only if $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} < 1$ for some initial condition $x \in \boldsymbol{X}$;

(iv) If $\boldsymbol{\Phi}$ is non-evanescent then $L(x, \sum_{i \in I} H_i) \equiv 1$. For each $i \in I$, the Harris set H_i is positive if and only if $\limsup_{k \to \infty} P^k(x, C) > 0$ for some $x \in H_i$, and some compact set $C \subset \boldsymbol{X}$.

Proof. Result (i) is taken from [33].

To see (ii), let $D_i = \{x : T(x, H_i) > 0\}$, $i \in I$, and let $D_E = \{x : T(x, E) > 0\}$. Clearly by (i), C is covered by $\cup_i D_i \cup D_E$, and so there exists N such that $C \subset \cup_1^N D_i \cup D_E$. Since H_i is absorbing we have $H_i \cap D_E = \varnothing$ and $H_i \cap D_j = \varnothing$, $i \neq j$. Hence $C \cap H_i = \varnothing$ for $i > N$, as required.

The result (iii) follows from Proposition 2.1, with $E = A$, and (iv) is a direct consequence of Proposition 3.4 of [33] (which states that compact sets have finite π-measure) and Corollary 6.7(iii) of [26].

Remark. We note explicitly, since we use it in Part II, that this decomposition relies only on the hitting-time probabilities $L(x, A)$ possessing an everywhere non-trivial continuous component, as was the case in [33], rather than on the full power of the T-chain assumptions. In the irreducible case this gives the following connections immediately.

Corollary. Suppose that $\mathbf{\Phi}$ is an irreducible T-chain. Then
(i) $\mathbf{\Phi}$ is non-evanescent if and only if $\mathbf{\Phi}$ is Harris recurrent;
(ii) $\mathbf{\Phi}$ is bounded in probability if and only if $\mathbf{\Phi}$ is positive Harris recurrent.

Proof. (i) follows immediately from (iii) of Theorem 2.1. To see (ii), note that if $\mathbf{\Phi}$ is bounded in probability, then from (iv) of Theorem 2.1, the Harris set H is positive, whilst if $\mathbf{\Phi}$ is a positive Harris-recurrent chain, then the standard limit theorems show that for all initial values the distributions P^k are tight.

From the decomposition theorem we also see that in the reducible case, if $\mathbf{\Phi}$ is a T-chain, then $\mathbf{\Phi}$ is non-evanescent if and only if every trajectory enters some Harris set with probability 1. Again, a stronger result is in fact true.

Theorem 2.2. Suppose that $\mathbf{\Phi}$ is a T-chain. Then $\mathbf{\Phi}$ is bounded in probability if and only if it is non-evanescent, and every Harris set is positive.

Proof. The proof of Theorem 2.2 in [18] can be adapted, given the results of the decomposition theorem, and we omit the details.

3. Petite sets

3.1. *Petite sets and stability conditions.* We now define a class of sets which will be critical in linking stability properties of the chain with various criteria for those properties to hold.

A set $A \in \mathscr{B}$ and a sub-probability measure φ on $B(X)$ are called petite if for some probability a on $\mathbb{Z}_+$, $K_a(x, \cdot) \geqq \varphi(\cdot)$ for all $x \in A$.

If $\mathbf{\Phi}$ is positive Harris recurrent with invariant probability π, then the class of petite sets which have positive π-measure extends the class of small sets, defined by [26], which are petite sets with $a = \mathbf{e}_j$ for some j. We can immediately show that the existence of sufficient compact petite sets implies that $\mathbf{\Phi}$ is a T-chain.

Proposition 3.1. If an open petite set A exists, then K_a possesses a continuous component for some a, non-trivial on all of A.

Proof. Since A is petite, there exists a probability a on $\mathbb{Z}_+$ and a sub-probability φ such that $K_a(\cdot, \cdot) \geqq \mathbf{1}_A(\cdot)\varphi\{\cdot\}$. It is easy to see that $T(\cdot, \cdot) \triangleq \mathbf{1}_A(\cdot)\varphi\{\cdot\}$ is a continuous component of K_a. It certainly satisfies the first and third conditions of the definition of a T-chain. If $B \in \mathscr{B}$, then $T(\cdot, B) = \varphi\{B\}\mathbf{1}_A(\cdot)$ and since the set A is open, the second condition is also satisfied.

Proposition 3.2. Suppose that for each $x \in \boldsymbol{X}$ there exists a probability a_x on $\mathbb{Z}_+$ such that K_{a_x} possesses a continuous component T_x which is non-trivial at x. Then $\boldsymbol{\Phi}$ is a T-chain.

Proof. For each $x \in \boldsymbol{X}$, let W_x denote the open set $W_x = \{y \in \boldsymbol{X} : T_x(y, \boldsymbol{X}) > 0\}$. Observe that by the given hypotheses, $x \in W_x$ for each $x \in \boldsymbol{X}$. By Lindelöf's theorem there exists a countable subcollection of sets $\{W_i : i \in \mathbb{Z}_+\}$ and corresponding kernels T_i and K_{a_i} such that $\bigcup W_i = \boldsymbol{X}$. Letting

$$T = \sum_{k=0}^{\infty} 2^{-k-1} T_k \quad \text{and} \quad a = \sum_{k=0}^{\infty} 2^{-k-1} a_k,$$

it follows that $K_a \geqq T$, and hence satisfies the conclusions of the proposition.

Theorem 3.1. If every compact set is petite, then $\boldsymbol{\Phi}$ is a T-chain.

Proof. Since $\boldsymbol{X}$ is σ-compact, there is a countable covering of open petite sets, and the result follows from Proposition 3.1 and Proposition 3.2.

In what follows we require the following preliminary result, the proof of which we leave to the reader.

Lemma 3.1. Suppose that A is petite, and that for some probability a on $\mathbb{Z}_+$ and a set $B \in \mathcal{B}$, $\inf_{x \in B} K_a(x, A) > 0$. Then B is also petite.

We now use Theorem 2.1 to give a stronger link between the existence of petite compact sets and T-chain properties.

Theorem 3.2. (i) Suppose that $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} < 1$ for one x. Then every compact set is petite if and only if $\boldsymbol{\Phi}$ is an irreducible T-chain.

(ii) Suppose that $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} < 1$ for all $x \in \boldsymbol{X}$. Then $\boldsymbol{\Phi}$ is a T-chain if and only if every compact set admits a finite cover by open petite sets.

Proof. (i) Suppose that $\boldsymbol{\Phi}$ is φ-irreducible. Let A be a φ-positive petite set, and let K_a have an everywhere non-trivial continuous component T. By irreducibility $K_\varepsilon(x, A) > 0$, and hence $K_{a \star \varepsilon}(x, A) = K_a K_\varepsilon(x, A) \geqq TK_\varepsilon(x, A) > 0$ for all $x \in \boldsymbol{X}$. The function $TK_\varepsilon(\cdot, A)$ is lower semi-continuous and positive everywhere on $\boldsymbol{X}$. Hence $K_{a \star \varepsilon}(x, A)$ is uniformly bounded from below on compact subsets of $\boldsymbol{X}$. Lemma 3.1 completes the proof of the 'if' part.

Conversely, if every compact set is petite, then by Theorem 3.1, $\boldsymbol{\Phi}$ is a T-chain. Moreover from the Decomposition Theorem 2.1(i) $\boldsymbol{X} = \sum H_i + E$. If $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} < 1$ for some x, then from the Decomposition Theorem 2.1(iii), $\sum H_i \neq \emptyset$. Let $y \in H_1$, and $x \in \boldsymbol{X}$ arbitrary. Then the set $\{x\} \cup \{y\}$ is compact and hence petite. Moreover, if φ is an associated petite measure then since H_1 is absorbing, $\varphi\{H_1^c\} \leqq K_a(y, H_1^c) = 0$ and hence $L(x, H_1) \geqq \varphi\{H_1\} > 0$. This shows that $L(x, H_1) > 0$ for all x, so that $\boldsymbol{\Phi}$ is π_1-irreducible, where π_1 is the invariant measure for H_1.

(ii) Under the given conditions it follows from (iii) of the Decomposition Theorem 2.1 that $L(x, \sum H_i) > 0$ for all x. Let π_i denote the invariant measure on H_i, and let $\{A_i\}$ denote π_i-positive petite sets with $A_i \subset H_i$.

Without loss of generality we can assume that $T(x, \sum A_i) > 0$ for all x. Indeed, we have $K_{a\star\varepsilon} = K_a K_\varepsilon \geqq TK_\varepsilon = T'$ (say). Since $K_\varepsilon(x, \sum A_i) > 0$ holds for any $x \in \boldsymbol{X}$, we may replace a by $a \star \varepsilon$, and T by T' to give the desired continuous component. If C is compact, then there exists N such that

$$C \subset \bigcup_{i=1}^{N} O_i^N \triangleq \bigcup_{i=1}^{N} \{x : T(x, A_i) \geqq 1/N\}.$$

Since by Lemma 3.1 each O_i^N is petite, this completes the 'only if' part. The converse of (ii) is a direct consequence of Propositions 3.1 and 3.2.

Finally, we link the properties of petite sets, T-chains and stability conditions in the following theorem.

Theorem 3.3. Suppose that every compact subset of $\boldsymbol{X}$ is petite.

(i) A Harris set $H \subset \boldsymbol{X}$ exists and is unique if and only if $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} < 1$ for some $x \in \boldsymbol{X}$, and if this is the case, then for all $y \in \boldsymbol{X}$,

$$\boldsymbol{P}_y\{\{\boldsymbol{\Phi} \to \infty\} \cup \{\boldsymbol{\Phi} \text{ enters } H\}\} = 1;$$

(ii) $\boldsymbol{\Phi}$ is Harris recurrent if and only if it is non-evanescent;

(iii) $\boldsymbol{\Phi}$ is positive Harris recurrent if and only if it is bounded in probability.

Proof. Since every compact set is petite, we see from Theorem 3.1 that $\boldsymbol{\Phi}$ is a T-chain.

(i) If H exists and is unique, then from Theorem 2.1(iii), for some $x \in \boldsymbol{X}$, $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} = 0$, and then as required, for all $y \in \boldsymbol{X}$, $\boldsymbol{P}_y\{\{\boldsymbol{\Phi} \to \infty\} \cup \{\boldsymbol{\Phi} \text{ enters } H\}\} = 1$.

Conversely, if for some $x \in \boldsymbol{X}$, $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} < 1$, then $\boldsymbol{\Phi}$ is an irreducible T-chain, from Theorem 3.2. Hence there is exactly one Harris set, from Theorem 2.1(iii) again.

(ii) follows directly from (i), and Theorem 2.1(iii).

(iii) follows from (i) and from Theorem 2.1(iv), as in the corollary to that theorem.

We have shown that in general, chains for which all compact sets are petite are T-chains. We now show that a more conventional class of chains, Feller chains, also give a rich set of petite sets under suitable auxiliary conditions.

3.2. *Feller chains and petite sets.* A Markov chain $\boldsymbol{\Phi}$ is said to have the *Feller property* if Pf is continuous for every bounded and continuous function $f : \boldsymbol{X} \to \mathbb{R}$. We require the following lemma for petite sets for Feller chains.

Lemma 3.2. If $\boldsymbol{\Phi}$ is a φ-irreducible Feller chain, then the closure of every petite set is petite.

Proof. By Lemma 3.1, Proposition 2.11(iv) of [26] (which states that $\boldsymbol{X}$ may be written as a countable union of small sets), and regularity of probability measures on $\mathscr{B}$ (i.e. a set $A \in \mathscr{B}$ may be approximated from within by compact sets), the set A is petite if and only if there exists a probability a on $\mathbb{Z}_+$, $\delta > 0$, and a compact petite set $C \subset \boldsymbol{X}$ such that $K_a(x, C) \geqq \delta$ for all $x \in A$. Since the function $K_a(x, C)$ is upper semicontinuous (Lemma 4.1 of [9]), we have $\inf_{x \in \bar{A}} K_a(x, C) = \inf_{x \in A} K_a(x, C)$, and this shows that the closure of a petite set is petite.

We can now provide the auxiliary conditions under which all compact sets are petite for a Feller chain.

Theorem 3.4. Suppose that $\boldsymbol{\Phi}$ is φ-irreducible. Then either of the conditions below implies that all compact subsets of $\boldsymbol{X}$ are petite:
(i) $\boldsymbol{\Phi}$ has the Feller property and an open φ-positive petite set exists;
(ii) $\boldsymbol{\Phi}$ has the Feller property and supp φ has non-empty interior.

Proof. The proof is similar to Theorem 4.1 of [9] or Theorem 1 of [30].

(i) Let A be an open petite set of positive φ-measure. Then $K_\varepsilon(\cdot, A)$ is lower semicontinuous and positive everywhere, and hence bounded from below on compact sets. Lemma 3.1 completes the proof.

(ii) Let A be a φ-positive petite set, and define $A_k \triangleq \mathrm{cl}\,\{x : K_\varepsilon(x, A) \geqq 1/k\} \cap \operatorname{supp} \varphi$. By Lemmas 3.1 and 3.2, each A_k is petite. Since supp φ has non-empty interior it is of the second category, and hence there exists $k \in \mathbb{Z}_+$ and an open set $O \subset A_k \subset \operatorname{supp} \varphi$. The set O is a φ-positive petite set, and hence we may apply (i) to conclude (ii).

4. Criteria for stability

4.1. *Petite sets and stability conditions.* It is well known that forms of stochastic stability are closely related to the return time behaviour of the chain on sets in the 'centre' of the space: individual points on countable spaces, compacta on topological spaces, and so on. By deriving conditions for appropriately finite return times to such sets, criteria for stability have been found in [35], [36] and [26]. We show here that petite sets play this same role, and then proceed to develop criteria for them to have finite return times.

Theorem 4.1. (i) The chain $\boldsymbol{\Phi}$ is Harris recurrent if and only if a petite set A exists with $L(x, A) = 1$ for all $x \in \boldsymbol{X}$.

(ii) The chain $\boldsymbol{\Phi}$ is positive Harris recurrent if and only if a petite set A exists with $L(x, A) = 1$ for all $x \in \boldsymbol{X}$, and $\sup_{x \in A} \boldsymbol{E}_x[\tau_A] < \infty$.

(iii) If $\boldsymbol{\Phi}$ possesses a sub-invariant measure π, then every petite set has finite π-measure.

Proof. We first prove (iii), which does not depend on any irreducibility hypothesis. If A is petite we have $K_a(x, B) \geqq \varphi\{B\}$ for all $x \in A$, $B \in \mathscr{B}$. Suppose that π is a sub-invariant measure with $\pi\{A\} \neq 0$. It is easy to see that $\varphi < \pi$, and since π is σ-finite there exists a set B such that $\pi\{B\} < \infty$ and $\varphi\{B\} > 0$. In this case we have $\infty > \pi\{B\} \geqq (\pi I_A K_a)\{B\} \geqq \pi\{A\}\varphi\{B\}$. To see (i), note that the chain is irreducible if a petite set A exists with $L(x, A) = 1$ for all $x \in \boldsymbol{X}$. By (iii), and the results of Sections 3 and 5 of [35], the chain is Harris recurrent. The converse is well known. To see (ii), assume that the chain is positive recurrent. From Proposition 5.13 of [26] there is an increasing sequence of petite sets satisfying the conditions stated. Conversely, since the chain is irreducible as in the proof of (i), from (iii) and Section 5 of [35], it is straightforward that the chain is positive recurrent as required.

In this sense petite sets mimic the behaviour we would expect of compact sets, and they play very much this role for the chain $\boldsymbol{\Phi}$. As in the proof of Theorem 4.1(i) and (ii), Theorem 4.1(iii) implies that petite sets are status sets for the class of irreducible Markov chains [35], and for irreducible chains it also follows from Theorem 4.1(ii) that petite sets are special instances of the test sets used in [37].

Having thus indicated that finiteness of hitting times for petite sets implies various forms of stability, we give criteria for such hitting times to be appropriately finite.

4.2. *Dynkin's formula.* The key result of the paper is a discrete form of Dynkin's formula. This provides the appropriate tool to bound the hitting times on sets, as required by, for example, the results of Theorem 4.1. This approach allows us to use much more general 'test functions' (which depend on the whole history of the chain rather than on the present position alone) than have appeared previously in the literature.

Dynkin's formula will be shown to yield necessary and sufficient conditions for positive recurrence and recurrence; criteria for geometric rates of convergence of the distributions of the chain; connections between stability and mixing conditions; and sample path ergodic theorems such as the central limit theorem and law of the iterated logarithm. Define

$$\mathscr{F}_k = \sigma\{\Phi_0, \cdots, \Phi_k\}, \tag{7}$$

let A denote a measurable subset of $\boldsymbol{X}$, and let

$$\sigma_A = \min(k \geqq 0 : \Phi_k \in A). \tag{8}$$

Let $\{Z_k, \mathscr{F}_k\}$ be an adapted sequence of positive random variables. We shall assume that for each k, Z_k is a fixed Borel-measurable function of $(\Phi_0, \cdots, \Phi_k)$. We abuse notation by letting Z_k denote both the random variable and the function of $\boldsymbol{\Phi}$. For any stopping time τ define $\tau^n \triangleq \min\{n, \tau, \inf\{k \geqq 0 : Z_k \geqq n\}\}$. The random time τ^n is a stopping time since it is the minimum of stopping times, and the random variable $\sum_{i=1}^{\tau^n} Z_{i-1}$ is essentially bounded.

Discrete Dynkin's formula. For each $x \in \boldsymbol{X}$ and $n \in \mathbb{Z}_+$,

$$\boldsymbol{E}_x[Z_{\tau^n}] = \boldsymbol{E}_x[Z_0] + \boldsymbol{E}_x\left[\sum_{i=1}^{\tau^n} (\boldsymbol{E}[Z_i \mid \mathcal{F}_{i-1}] - Z_{i-1})\right].$$

Proof. For each $n \in \mathbb{Z}_+$,

$$\begin{aligned} Z_{\tau^n} &= Z_0 + \sum_{i=1}^{\tau^n} (Z_i - Z_{i-1}) \\ &= Z_0 + \sum_{i=1}^{n} \mathbf{1}\{\tau^n > i\}(Z_i - Z_{i-1}). \end{aligned}$$

Taking expectations and noting that $\{\tau^n \geqq i\} \in \mathcal{F}_{i-1}$, we obtain

$$\begin{aligned} \boldsymbol{E}_x[Z_{\tau^n}] &= \boldsymbol{E}_x[Z_0] + \boldsymbol{E}_x\left[\sum_{i=1}^{n} \boldsymbol{E}_x[Z_i - Z_{i-1} \mid \mathcal{F}_{i-1}]\mathbf{1}\{\tau^n \geqq i\}\right] \\ &= \boldsymbol{E}_x[Z_0] + \boldsymbol{E}_x\left[\sum_{i=1}^{\tau^n} (\boldsymbol{E}_x[Z_i \mid \mathcal{F}_{i-1}] - Z_{i-1})\right]. \end{aligned}$$

The following result is an easy consequence.

Corollary. Suppose that there exists a sequence of positive measurable functions $\{\varepsilon_k : k \geqq 0\}$ on $\boldsymbol{X}$ such that

(i) $\varepsilon_{k+1}(x) \leqq C\varepsilon_k(x)$ for some $C < \infty$, and all $k \in \mathbb{Z}_+$, $x \in A^c$;

(ii) $\boldsymbol{E}[Z_{k+1} \mid \mathcal{F}_k] \leqslant Z_k - \varepsilon_k(\Phi_k)$, $\sigma_A > k$.

Then

$$\boldsymbol{E}_x\left[\sum_{i=1}^{\tau_A} \varepsilon_{i-1}(\Phi_{i-1})\right] \leqq \begin{cases} Z_0(x), & x \in A^c; \\ \varepsilon_0(x) + CPZ_0(x), & x \in \boldsymbol{X}. \end{cases}$$

Proof. Let Z_k and ε_k denote the random variables $Z_k(\Phi_0, \cdots, \Phi_k)$ and $\varepsilon_k(\Phi_k)$ respectively. By hypothesis, $\boldsymbol{E}[Z_k \mid \mathcal{F}_{k-1}] - Z_{k-1} \leqq -\varepsilon_{k-1}$ whenever $1 \leqq k \leqq \sigma_A$. Hence we have by Dynkin's formula

$$0 \leqq \boldsymbol{E}_x[Z_{\tau_A^n}] \leqq Z_0(x) - \boldsymbol{E}_x\left[\sum_{i=1}^{\tau_A^n} \varepsilon_{i-1}\right], \qquad x \in A^c,\ n \in \mathbb{Z}_+.$$

By the monotone convergence theorem it follows that for all initial conditions

$$\boldsymbol{E}_x\left[\sum_{i=1}^{\tau_A} \varepsilon_{i-1}\right] \leqq Z_0(x), \qquad x \in A^c.$$

This proves the corollary for $x \in A^c$. For arbitrary x, we have

$$\begin{aligned} \boldsymbol{E}_x\left[\sum_{i=1}^{\tau_A} \varepsilon_{i-1}\right] &= \varepsilon_0 + \boldsymbol{E}_x\left[\boldsymbol{E}_{\Phi_1}\Big(\sum_{i=1}^{\tau_A} \varepsilon_i(\Phi_{i-1})\Big)\mathbf{1}\{\Phi_1 \in A^c\}\right] \\ &\leqq \varepsilon_0 + CPZ_0(x). \end{aligned}$$

4.3. *A criterion for non-evanescence.* We first consider a criterion for non-evanescence, which extends the results of Section 10 of [35]. This will provide a drift condition as in Theorem 10.1 of [35], but instead of considering a fixed test function V of the state Φ_k, we let $\{V_k : k \in \mathbb{Z}_+\}$ denote a family of positive Borel-measurable functions $V_k : \boldsymbol{X}^{k+1} \to \mathbb{R}_+$. This generalization is of more than academic value: see [21], [19], [20] for applications where a fixed function does not suffice.

In order to derive criteria for non-evanescence, which is defined in terms of compact sets, we need to adapt the sequence $\{V_k\}$ more strongly to the topology. A function V is called *norm-like* if $V(x) \to \infty$ as $x \to \infty$: this means that the level sets $\{x : V(x) \leqq B\}$ are precompact for each $B > 0$. The usual choice for such functions on Euclidean space is the norm itself, or monotonic functions of the norm, hence our terminology. By imposing the appropriate 'drift condition' on the stochastic process $V_k = V_k(\Phi_0, \cdots, \Phi_k)$, we obtain bounds on moments of first entrance times to certain subsets of $\boldsymbol{X}$.

Assume that there exists a norm-like function $V : \boldsymbol{X} \to \mathbb{R}_+$ with the property

$$V_k(x_0, \cdots, x_k) \geqq V(x_k) \geqq 0 \tag{9}$$

for all $k \in \mathbb{Z}_+$ and all $x_i \in \boldsymbol{X}$. The criterion we consider for such a $\{V_k\}$ is the following.

> There exists a compact set $A \subset \boldsymbol{X}$, and an adapted sequence $\{V_k, \mathscr{F}_k\}$ satisfying (9), such that
>
> $$\boldsymbol{E}_x[V_{k+1} \mid \mathscr{F}_k] \leqq V_k \quad \text{a.s. } [\boldsymbol{P}_{\Phi_0}], \tag{DD1}$$
>
> when $\sigma_A > k$, $k \in \mathbb{Z}_+$.

Here the term DD refers to 'discrete drift'.

Theorem 4.2. If condition (DD1) holds then $\boldsymbol{\Phi}$ is non-evanescent.

Proof. Suppose that in fact $\boldsymbol{P}_x\{\boldsymbol{\Phi} \to \infty\} > 0$ for some $x \in \boldsymbol{X}$. Then, since the set A is compact, there exists $N \in \mathbb{Z}_+$ with

$$\boldsymbol{P}_x\{\{\Phi_k \in A^c, k \geqq N\} \cap \{\boldsymbol{\Phi} \to \infty\}\} > 0.$$

Hence letting $\mu = P^N(x, \cdot)$, we have by conditioning at time N,

$$\boldsymbol{P}_\mu\{\{\sigma_A = \infty\} \cap \{\boldsymbol{\Phi} \to \infty\}\} > 0. \tag{10}$$

We now show that (10) leads to a contradiction. Let $m_i = V_i \mathbf{1}\{\sigma_A \geqq i\}$. Using the fact that $\{\sigma_A \geqq k\} \in \mathscr{F}_{k-1}$, we may show that $(m_k, \mathscr{F}_k)$ is a positive supermartingale:

$$\boldsymbol{E}[m_k \mid \mathscr{F}_{k-1}] = \mathbf{1}\{\sigma_A \geqq k\}\boldsymbol{E}[V_k \mid \mathscr{F}_{k-1}] \leqq \mathbf{1}\{\sigma_A \geqq k\}V_{k-1} \leqq m_{k-1}.$$

Hence there exists an almost surely finite random variable m_∞ such that $m_k \to m_\infty$ as $k \to \infty$.

There are two possibilities for the limit m_∞. Either $\sigma_A < \infty$ in which case $m_\infty = 0$, or $\sigma_A = \infty$ in which case $\limsup_{k\to\infty} V(\Phi_k) = m_\infty < \infty$ and in particular $\boldsymbol{\Phi} \not\to \infty$. That is, we have shown that

$$P_\mu\{\{\sigma_A < \infty\} \cup \{\mathbf{\Phi} \to \infty\}^c\} = 1,$$

which clearly contradicts (10). Hence $\mathbf{\Phi}$ is non-evanescent.

4.4. *Criteria for boundedness in probability and positivity.* It is perhaps surprisingly easier to derive conditions for positive recurrence than for recurrence.

The criteria derived here extend the results of Sections 6 and 9 of [35], and again we consider $\{V_k = V_k(\Phi_0, \cdots, \Phi_k)\}$ to be a stochastic process of the form described above, and let A and f be measurable. We consider here the Lyapunov–Foster drift condition.

Positive drift condition. For some $\varepsilon > 0$,

(DD2) $$\boldsymbol{E}[V_{k+1} \mid \mathscr{F}_k] \leqq V_k - \varepsilon f(\Phi_k) \quad \text{a.s. } [\boldsymbol{P}_{\Phi_0}], \qquad k \in \mathbb{Z}_+$$

when $\sigma_A > k$.

Condition (DD2) is written in its general and hence most complicated form. Suppose for example as in [35] that $V_k = V(\Phi_k)$ where V is a simple positive measurable function on $\boldsymbol{X}$: Condition (DD2) then becomes the following.

For some $\varepsilon > 0$,

(DD2′) $$PV \leqq V - \varepsilon f \quad \text{on} \quad A^c.$$

This hypothesis, with $f = 1$, is well known to have powerful implications in stability theory: see for example [26] and [36].

Theorem 4.3. Suppose that A and $f \geqq 0$ are measurable, and that (DD2) is satisfied. Then

$$\boldsymbol{E}_x\left[\sum_{i=1}^{\tau_A} f(\Phi_{i-1})\right] \leqq \begin{cases} \varepsilon^{-1} V_0(x), & x \in A^c; \\ f(x) + \varepsilon^{-1} PV_0(x), & x \in \boldsymbol{X}. \end{cases}$$

Proof. Suppose that $\{V_k\}$ satisfies (DD2), and let $Z_k = V_k$. Then $\boldsymbol{E}[Z_k \mid \mathscr{F}_{k-1}] - Z_{k-1} \leqq -\varepsilon f(\Phi_{k-1})$, and hence the result follows by the corollary to Dynkin's formula.

These results, together with the classification in Section 2, immediately give the following result for reducible chains.

Theorem 4.4. Suppose that $\mathbf{\Phi}$ admits a sub-invariant measure μ, and that (DD2) is satisfied for some $A \in \mathscr{B}$ with $\mu(A)$ finite, and some $f \geqq 1$, with PV_0 bounded on A. Then μ is a finite invariant measure for the chain, and $\mathbf{\Phi}$ is μ-bounded in probability.

Proof. From Theorem 4.3, and imitating the proof of Theorem 1 of [37], we find μ is a finite invariant measure. From the σ-compactness of $\boldsymbol{X}$, the finiteness of μ, and using Fatou's lemma on the invariant equations, the result follows.

Corollary. Suppose that $\mathbf{\Phi}$ is a Feller chain and that (DD2) is satisfied for some compact $A \in \mathscr{B}$ and some $f \geqq 1$, with PV_0 bounded on A. Then there exists an invariant probability μ for the chain, and $\mathbf{\Phi}$ is μ-bounded in probability.

Proof. This follows as in [37], and is also related to results of [12].

We can do much better if $\mathbf{\Phi}$ is a T-chain.

Theorem 4.5. Suppose that $\mathbf{\Phi}$ is a T-chain satisfying (DD2) for some compact $A \in \mathscr{B}$, and some $f \geqq 1$, with PV_0 bounded on A. Then $\mathbf{\Phi}$ is non-evanescent, there is a finite number of Harris sets all of which are positive in the Doeblin decomposition, and hence $\mathbf{\Phi}$ is bounded in probability.

Proof. We see from Theorem 4.3 that under the conditions of Theorem 4.5, $L(x, A) \equiv 1$. From Lemma 1.1 we have $Q(x, A) = 1$ for all x, so that $\mathbf{\Phi}$ is non-evanescent. Since Harris sets are absorbing, $Q(x, A) = 1$, $x \in H_i$ implies $H_i \cap A \neq \varnothing$ for any Harris set H_i. Parts (ii) and (iii) of the decomposition theorem then show that the index set I is finite and non-empty. From Theorem 4.3, there exists a finite M such that $\sup_{x \in A} E_x(\tau_A) \leqq M$, and hence from Proposition 4.1 of [35], $\limsup_{k \to \infty} P^k(x, A) > 0$ for all x. From (iv) of the Decomposition Theorem 2.1, then, the Harris sets are all positive as required.

If the set A in (DD2) is petite, we reach our strongest positivity statement.

Theorem 4.6. Suppose that A is petite, $f \geqq 1$ is measurable, and that (DD2) is satisfied with $f + PV_0$ bounded on A. Then the chain is positive Harris recurrent and $\pi(f)$ is finite.

Proof. That $\mathbf{\Phi}$ is positive Harris recurrent follows from Theorem 4.1 and Theorem 4.3. That $\pi(f)$ is finite follows then from the representation of π in Lemma 1.2.

5. Criteria for recurrence and ergodicity: irreducible chains

The topological conditions of non-evanescence and boundedness in probability do not require any irreducibility. We have seen that, in the irreducible case, they equate to Harris recurrence and positive Harris recurrence under mild conditions.

In this section we review and extend known results which refine these irreducible chain concepts of stability, and which, in particular, relate the moments of hitting times on petite sets to rates of convergence to stable regimes. We can then use the Lyapunov–Foster approach to find criteria for these rates of convergence to hold.

We also turn to the consequences of positive recurrence in particular when the chain is irreducible. Not only does a finite invariant measure π exist in this context, but the distribution of $\{\Phi_k\}$ tends to π as $k \to \infty$, and does so in the total variation norm [26]. Thus in linking boundedness in probability with positive Harris recurrence, we have moved from weak convergence of $\{P^k\}$, which is often implied by tightness, to a much stronger form of convergence. This becomes even more apparent in Section 7 for reducible chains.

A criterion for Harris recurrence is easy to establish from our previous results.

Theorem 5.1. If there exists V satisfying (DD1) and every compact set is petite, then the chain is Harris recurrent.

Proof. Note that Φ is non-evanescent from Theorem 4.2; the result follows from Theorem 3.3.

This Lyapunov–Foster result strengthens results of [16]; see also Theorem 10.1 of [35]. The new approach has replaced the somewhat artificial requirements of [35], [16] that the level sets of the test function be 'status sets' for Φ with the more natural requirement that the compact sets are petite, which is clearly more closely allied to the structure of the chain itself.

Our next result considers criteria for, and consequences of, positive Harris recurrence. We provide an extension of the standard convergence result, and for this we need the following extension of the total variation norm. For any positive measurable function $f \geqq 1$ and any signed measure μ on $\mathscr{B}$ we write $\|\mu\|_f = \sup_{|g| \leqq f} |\mu(g)|$. Note that the total variation norm $\|\mu\|$ is $\|\mu\|_f$ in the special case where $f \equiv 1$.

Theorem 5.2. Let $f \geqq 1$ be measurable. Then

(i) The following are equivalent:

(a) the chain $\mathbf{\Phi}$ is positive Harris recurrent with invariant probability measure π, and $\pi(f) < \infty$.

(b) there exists a petite set A such that $L(x, A) \equiv 1$ and

$$\sup_{y \in A} \boldsymbol{E}_y \left[\sum_{k=1}^{\tau_A} f(\Phi_{k-1}) \right] < \infty. \tag{11}$$

(ii) If either of the equivalent conditions (a) or (b) holds then

$$\boldsymbol{E}_x \left[\sum_{k=1}^{\tau_A} f(\Phi_{k-1}) \right] < \infty \tag{12}$$

for a.e. $x \in \boldsymbol{X}$ $[\pi]$;

(iii) If $\mathbf{\Phi}$ is positive Harris recurrent with invariant probability measure π and period $m \geqq 1$ (cf. [29]), and if $\pi(f) < \infty$, then for any initial condition x satisfying (12), the distributions converge in f-norm:

$$\left\| \frac{1}{m} \sum_{i=1}^{m} P^{i+l}(x, \cdot) - \pi \right\|_f \to 0. \tag{13}$$

Proof. (i) and (ii): If $\mathbf{\Phi}$ is positive Harris recurrent, then (11) and (12) hold for some petite set by Proposition 5.13 of [26]. Conversely, if (b) holds, then Theorem

4.1 implies that $\mathbf{\Phi}$ is Harris recurrent with invariant measure π. Theorem 4.1(iii) together with (11) gives $\pi\{\boldsymbol{X}\} \leqq \pi(f) < \infty$ from the form of π in Lemma 1.2.

(iii) This convergence result follows from the case $R = 1$ of [28]: note that the null set N_g in that result can be taken as the set on which (12) fails, when $R = 1$, since the set N_1 in that proof can be taken as empty in this case.

Most of this result is essentially known, although it is not always clearly stated in the literature that (13) holds for all f, whether bounded or not, for which $\pi(f) < \infty$.

When $\mathbf{\Phi}$ is aperiodic (that is, the period $m = 1$), and (13) holds for some $f \geqq 1$ and all x, then $\mathbf{\Phi}$ is called *Harris ergodic.*

Theorem 5.3. Suppose that A is petite, that $f \geqq 1$ is measurable, and that (DD2) is satisfied with $f + PV_0$ bounded on A. Then $\mathbf{\Phi}$ is positive Harris recurrent, and (13) holds for all x and for the same f, so that in the aperiodic case, $\mathbf{\Phi}$ is Harris ergodic.

Proof. This follows from Theorem 4.3, and Theorem 5.2. Notice that we do not have to assume irreducibility when A is petite, and that the π-null set on which (12) fails is empty under the conditions of the theorem, so Harris ergodicity does hold in the aperiodic case.

This result extends Theorem 9.1 of [35] from the situation where V is a simple function on $\boldsymbol{X}$.

6. Criteria for geometric recurrence and geometric ergodicity

We now turn to the concepts of *geometric recurrence and geometric ergodicity* which are refinements of the idea of positive recurrence and Harris ergodicity (cf. [26], [27], [36]).

Stochastic stability condition 3. Suppose that for some measurable $f \geqq 1$, a petite set A, and $r > 1$,

$$\sup_{x \in A} \boldsymbol{E}_x\left[\sum_{k=1}^{\tau_A} r^k f(\Phi_{k-1})\right] < \infty, \tag{14}$$

and $L(x, A) \equiv 1$. Then $\mathbf{\Phi}$ is called geometrically recurrent.

Theorem 6.1. Suppose that the chain $\mathbf{\Phi}$ satisfies (14) for some A and r, and is hence geometrically recurrent. Then there exists $\rho < 1$, $R < \infty$, and a probability measure π such that

$$\left\|\frac{1}{m}\sum_{i=1}^{m} P^{i+l}(x, \cdot) - \pi\right\|_f \leqq R\boldsymbol{E}_x\left[\sum_{k=1}^{\tau_A} r^k f(\Phi_{k-1})\right]\rho^l, \qquad x \in \boldsymbol{X}, \qquad l \in \mathbb{Z}_+, \tag{15}$$

where the right-hand side is finite for a.e. $x \in \boldsymbol{X}\,[\pi]$. Conversely, if $\mathbf{\Phi}$ is Harris recurrent and (15) holds for some $r > 1$, $R < \infty$, $\rho < 1$, π a probability, and A petite, with the right-hand side finite for a.e. $x \in \boldsymbol{X}\,[\pi]$, then $\mathbf{\Phi}$ is geometrically recurrent.

To prove the result, we consider the 'split chain' [25], [26], [1]. This device allows us to shift our analysis from petite sets, to a recurrent atom lying in an enlarged state space. An *atom* $\boldsymbol{\alpha}$ is by definition a subset of the state space for which $P(x, \cdot) = P(y, \cdot)$ for all $x, y \in \boldsymbol{\alpha}$. Hence an atom is essentially a single state, and by the strong Markov property, the trajectories of the chain between visits to an atom are independent and identically distributed. Such a splitting is only possible in the strongly aperiodic case, for which we take the following definition from [1].

Strong aperiodicity. There exists a set C, a probability ν with $\nu\{C\} = 1$, and $\delta > 0$ such that $L(x, C) > 0$ for all $x \in \boldsymbol{X}$ and $P(x, A) \geqq \delta\nu\{A\}$, $x \in C$, $A \in \mathscr{B}$.

Under the assumptions imposed in this paper, the k-step chain $\{\Phi_{nk} : n \in \mathbb{Z}_+\}$ will be strongly aperiodic for some $k \geqq 1$ if $\boldsymbol{\Phi}$ is irreducible and aperiodic (see [29] where the set C is called a *C-set*).

In the proof of Theorem 6.1, and in the remaining results of the paper we consider only strongly aperiodic Markov chains. The general case follows by decomposing the state space into periodic classes and considering the k-step chain, but there is some work in relating the two chains [26]. We refer the reader to [22] for a proof in the more general case.

We require here, and in the proofs of several other results in the paper, the following three lemmas which we give in general form.

Lemma 6.1. Suppose that $\boldsymbol{\Phi}$ is strongly aperiodic, that $A \subset \boldsymbol{X}$ is petite, $f : \boldsymbol{X} \to [1, \infty]$ is measurable, and that for a constant $r > 1$, the function

$$V_A(x) \triangleq \boldsymbol{E}_x\left[\sum_{k=1}^{\tau_A} f(\Phi_{k-1}) r^k\right] \qquad x \in \boldsymbol{X}, \tag{16}$$

satisfies $\sup_{x \in A} V_A(x) < \infty$. Then for a split chain with atom $\boldsymbol{\alpha}$, there exists a constant $C_0 < \infty$ such that

$$\boldsymbol{E}_x\left[\sum_{k=1}^{\tau_{\boldsymbol{\alpha}}} f(\Phi_{k-1}) r^k\right] \leq C_0 V_A(x), \qquad x \in \boldsymbol{X}.$$

Proof. The proof of this result follows from a straightforward extension of the geometric trials argument used in [27] and so will be omitted. This technique requires that the C-set satisfy $\varphi\{C\} > 0$, where φ is the petite measure associated with A. This property may be assumed without loss of generality: replace φ with φK_ε. By definition $K_\varepsilon(x, C) > 0$ for all x so that the petite measure φK_ε satisfies the desired condition.

When the function V_A defined in Lemma 6.1 is finite-valued we also obtain bounds on moments of $V_A(\Phi_k)$ as follows.

Lemma 6.2. Suppose that $A \subset \boldsymbol{X}$ and $f: \boldsymbol{X} \to [1, \infty]$ are measurable, and that for a constant $r > 1$, $B_v \triangleq \sup_{x \in A} V_A(x) < \infty$. Then

(i) $PV_A(x) \leqq r^{-1} V_A(x) + B_v$, $x \in \boldsymbol{X}$;

(ii) $\sup_{k \in \mathbb{Z}_+} \boldsymbol{E}_x[V_A(\Phi_k)] < \infty$, when $V_A(x) < \infty$;

(iii) If $L(x, A) \equiv 1$ then $\int V_A \, d\pi \leqq B_v/(r-1) < \infty$ for any invariant probability π.

Proof. Results (ii)–(iii) follow easily from the bound (i) on PV. To obtain (i), we condition at $k = 1$:

$$PV_A(x) \leqq \boldsymbol{E}_x\left[\sum_{n=2}^{\tau_A} r^{n-1} f(\Phi_{n-1}) \mathbf{1}\{\tau_A \geqq 2\}\right] + \sup_{x \in A} \boldsymbol{E}_x\left[\sum_{n=1}^{\tau_A} r^n f(\Phi_{n-1})\right]$$
$$\leqq r^{-1} V_A + B_v.$$

Lemma 6.3. Let $\{a_k : k \in \mathbb{Z}_+\}$ be a sequence of complex numbers, and suppose that the function $f(z) \triangleq \sum a_k z^k$ is analytic in a neighborhood of the closed unit disk $\bar{U} \subset \mathbb{C}$, and $|f(z)| \leqq C$ for some $C < \infty$, and all $|z| \leqq r < \infty$. Then $|a_k| \leqq C r^{-k}$ for all $k \in \mathbb{Z}_+$.

Proof. By Parseval's theorem and the boundedness assumption we have

$$C^2 \geqq \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_{k=0}^{\infty} a_k r^k e^{i\theta} \right|^2 d\theta = \sum_{k=0}^{\infty} |a_k|^2 r^{2k}.$$

Proof of Theorem 6.1. The converse follows from Theorem 6.14 of [26], where it is also shown that (14) implies a geometric rate of convergence in the ergodic theorem. The main new result here is the upper bound on the rate of convergence (see also [6]). We first note that under (14), $\boldsymbol{\Phi}$ is positive Harris recurrent with invariant probability π, and, from Lemma 6.2,

$$\boldsymbol{E}_x\left[\sum_{k=1}^{\tau_A} r^k f(\Phi_{k-1})\right] < \infty \tag{17}$$

for a.e. $x \in \boldsymbol{X}[\pi]$. Suppose now that (14) holds, and fix $x \in \boldsymbol{X}$ satisfying (17). Fix $|g| \leqq f$, and let $\bar{g}$ denote the centered function $g - \int g \, d\pi$. Lemma 6.2(i) implies that the function

$$G(z) = \sum_{n=0}^{\infty} \boldsymbol{E}_x[\bar{g}(\Phi_n)] z^n \tag{18}$$

is analytic in the open unit ball $U \subset \mathbb{C}$.

By Lemma 6.1, we may turn to the split state space and assume that an atom $\boldsymbol{\alpha}$ exists, so that the function G may be written in the standard way as the sum of an analytic function and the product of two analytic functions with domain U, where the interchange of summation and expectation is justified by Fubini's theorem:

$$G(z) = \boldsymbol{E}_x\left[\sum_{n=0}^{\tau_{\boldsymbol{\alpha}}} \bar{g}(\Phi_n) z^n\right] + \boldsymbol{E}_x[z^{\tau_{\boldsymbol{\alpha}}}] \sum_{n=1}^{\infty} \boldsymbol{E}_{\boldsymbol{\alpha}}[\bar{g}(\Phi_n)] z^n.$$

By Corollary 2.4 of [27] we have for some $r_0 > 1$, and all $|z| \leqq r_0$,

$$\left|\sum_{n=1}^{\infty} \boldsymbol{E}_{\boldsymbol{\alpha}}[\bar{g}(\Phi_n)]z^n\right| \leqq \sum_{n=1}^{\infty} \|P^n(\boldsymbol{\alpha}, \cdot) - \pi\|_f r_0^n < \infty.$$

Hence by Lemma 6.1, there exists $C_0, C_1 < \infty$, and a neighborhood $N \supset \bar{U}$, all independent of g satisfying the conditions of the theorem and x satisfying (17), such that G is analytic in N and for each $z \in N$,

$$\begin{aligned} |G(z)| &\leqq C_1 \boldsymbol{E}_x\left[\sum_{k=0}^{\tau_{\boldsymbol{\alpha}}} f(\Phi_k) r^k\right] \\ &\leqq C_0 C_1 \boldsymbol{E}_x\left[\sum_{k=0}^{\tau_A} f(\Phi_k) r^k\right]. \end{aligned}$$

Lemma 6.3 together with this bound completes the proof.

We now see that a simple and often easily verified extension of (DD2) gives a general criterion for these stability regimes to hold.

For some $\varepsilon > 0$, $\lambda < 1$,

(DD3) $\quad \boldsymbol{E}[V_{k+1} \mid \mathscr{F}_k] \leqq \lambda V_k - \varepsilon f(\Phi_k) \quad \text{a.s. } [\boldsymbol{P}_{\Phi_0}], \qquad k \in \mathbb{Z}_+$

when $\sigma_A > k$.

Condition (DD3), like Condition (DD2), is given in a general form. Suppose for example as in [36] that $V_k = V(\Phi_k)$ where V is merely a positive measurable function on $\boldsymbol{X}$: Condition (DD3) then becomes the following.

(DD3′) For some $\varepsilon > 0$, $\lambda < 1$,

$$PV \leqq \lambda V - \varepsilon f \quad \text{on } A^c.$$

Theorem 6.2. Suppose that A and $f \geqq 0$ are measurable, and that (DD3) is satisfied. Then with $r \triangleq \lambda^{-1}$,

$$\boldsymbol{E}_x\left[\sum_{i=1}^{\tau_A} r^i f(\Phi_{i-1})\right] \leqq \begin{cases} \varepsilon^{-1} V_0(x), & x \in A^c; \\ rf(x) + \varepsilon^{-1} r P V_0(x), & x \in \boldsymbol{X}. \end{cases}$$

Proof. Suppose that $\{V_k\}$ satisfies (DD3), and let $Z_k = r^k V_k$ where $r = \lambda^{-1}$. Then $\boldsymbol{E}[Z_k \mid \mathscr{F}_{k-1}] - Z_{k-1} \leqq -\varepsilon r^k f(\Phi_{k-1})$, and hence the result follows by the corollary to Dynkin's formula.

When $\boldsymbol{\Phi}$ is aperiodic and (15) holds with the right-hand side finite for all x, then $\boldsymbol{\Phi}$ is called *geometrically ergodic*.

These results, together with the classification in Theorem 6.1, immediately give the following.

Corollary. Suppose that A is petite, $f \geqq 1$ is measurable, and that (DD3) is satisfied with $f + PV_0$ bounded on A. Then the chain $\boldsymbol{\Phi}$ satisfies (14) and (17) with the same f, and all $x \in \boldsymbol{X}$. Hence $\boldsymbol{\Phi}$ is geometrically recurrent.

In the aperiodic case there exists $R < \infty$ and $\rho < 1$ such that

$$\|P^n(x, \cdot) - \pi\|_f \leqq R(PV_0(x) + f(x))\rho^n, \qquad n \in \mathbb{Z}_+, \qquad x \in \boldsymbol{X}.$$

Hence in this case $\boldsymbol{\Phi}$ is geometrically ergodic.

Following the matrix definition as in [32] we define the f-operator norm, for a positive function f and a kernel G, as

$$\|G\|_{f,f} = \sup_{x \in X} \frac{\|G(x, \cdot)\|_f}{f(x)}.$$

Corollary. Suppose that A is petite, $f \geqq 1$ is measurable, and that (DD3′) is satisfied with $f + PV$ bounded on A. Then in the aperiodic case $\|P^n - \pi\|_{f+V,f+V} \leqq R\rho^n$ for some $\rho < 1$, $R < \infty$.

Proof. For ε small enough (DD3′) can be written as

$$PV \leqq (\lambda + \varepsilon)V - \varepsilon(V + f).$$

From the previous corollary and (DD3′),

$$\|P^n - \pi\|_{f+V} \leqq R(PV + V + f)\rho^n \leqq R'(V + f)\rho^n$$

which is the result.

This equivalence of geometric ergodicity and the geometric rate of convergence in the stronger operator norm is proved in [32] in the countable state space situation. On topological spaces, the conditions of (DD3) are often satisfied, in practice, by verifying the following slightly stronger condition.

There exists $\lambda < 1$, $L < \infty$ and an adapted sequence $\{V_k, \mathscr{F}_k\}$ satisfying (9), such that

(DD4) $$\boldsymbol{E}_x[V_{k+1} \mid \mathscr{F}_k] \leqq \lambda V_k + L \quad \text{a.s. } [\boldsymbol{P}_{\Phi_0}],$$

for all $k \in \mathbb{Z}_+$ and all initial conditions $x \in \boldsymbol{X}$.

Theorem 6.3. Suppose that for the Markov chain $\boldsymbol{\Phi}$, all compact subsets of $\boldsymbol{X}$ are petite. If condition (DD4) holds, and if V_0 is uniformly bounded on compact subsets of $\boldsymbol{X}$, then the conclusions of the corollary to Theorem 6.2 hold with $f = V + 1$, where V is defined in (9). In the aperiodic case we may improve the upper bound on the rate of convergence: there exists $R < \infty$ and $\rho < 1$ such that

$$\|P^n(x, \cdot) - \pi\|_f \leqq R(V_0(x) + 1)\rho^n, \qquad n \in \mathbb{Z}_+, \qquad x \in \boldsymbol{X}.$$

Proof. Let $\lambda < \rho < 1$, and define the precompact set A and the constant $\varepsilon > 0$ by

$$A = \left\{x \in \boldsymbol{X} : V(x) \leqq \frac{2L}{\rho - \lambda} + 1\right\} \qquad \varepsilon = \frac{\rho - \lambda}{2}.$$

Then for all $k \in \mathbb{Z}_+$,

$$E[V_{k+1} \mid \mathscr{F}_k] \leqq \rho V_k + \left\{[L + (\rho - \lambda)] - \frac{\rho - \lambda}{2}(V(\Phi_k) + 1)\right\} - \frac{\rho - \lambda}{2}(V(\Phi_k) + 1).$$

Hence $E[V_{k+1} \mid \mathscr{F}_k] \leqq \rho V_k - \varepsilon f(\Phi_k)$ when $\Phi_k \in A^c$. This shows that (DD3) holds. We now show that the bound on V_0 implies (17) with $f = V + 1$.

Let $r = \rho^{-1}$, $Z_k = r^k V_k$, so that

$$E[Z_{k+1} \mid \mathscr{F}_k] \leqq Z_k - \varepsilon r^{k+1} f(\Phi_k),$$

when $\Phi_k \in A^c$. Then we may use Dynkin's formula to deduce that for all $x \in X$,

$$\begin{aligned} 0 \leqq E_x[z_{\tau_A^m}] &= E_x[Z_1] + E_x\left[\left(\sum_{k=2}^{\tau_A^m} E[Z_k \mid \mathscr{F}_{k-1}] - Z_{k-1}\right)\mathbf{1}(\tau_A \geqq 2)\right] \\ &\leqq E_x[Z_1] - E_x\left[\sum_{k=2}^{\tau_A^m} \varepsilon r^k f(\Phi_{k-1})\mathbf{1}(\tau_A \geqq 2)\right]. \end{aligned}$$

This and the monotone convergence theorem shows that for all $x \in X$,

$$E_x\left[\sum_{k=1}^{\tau_A} r^k f(\Phi_{k-1})\right] \leqq \varepsilon^{-1} r E_x[V_1] + rV(x).$$

This completes the proof, since $E_x[V_1] + V(x) \leqq \lambda V_0(x) + L + V_0(x)$ by (DD4) and (9).

Theorem 6.3 is a generalization of Proposition 3.1 of [17].

7. Stability and ergodic theorems: non-irreducible chains

Here we combine the strengthened version of the decomposition theorem in Section 2.2 with the stability concepts of Section 3 to indicate the strength of the ergodic theorems which can be deduced in the non-irreducible case.

Let $\{H_i : i \in I\}$ denote the Harris sets constructed in the decomposition theorem. If each Harris set is positive and if $L(x, E^c) = 1$ for all x, then we define the Markov transition function Π and the random probability $\tilde{\pi}$ by

$$\begin{aligned} \Pi(x, A) &\triangleq \sum_{i \in I} L(x, H_i)\pi_i\{A\} \\ \tilde{\pi}\{A\} &\triangleq \sum_{i \in I} \mathbf{1}\{\Phi \text{ enters } H_i\}\pi_i\{A\} \end{aligned} \tag{19}$$

where π_i is the invariant probability supported on H_i. Observe that since H_i is absorbing, $L(x, H_i) = Q(x, H_i)$. By Theorem 3.4 of [26] the function $Q(\cdot, H_i)$ is *harmonic,* that is, $\int P(\cdot, dy)Q(y, H_i) = Q(\cdot, H_i)$ and hence we have the relations

$$\Pi P = P\Pi = \Pi. \tag{20}$$

In Theorem 4.5 we showed how Lyapunov–Foster criteria can be constructed to give boundedness in probability. We see in Theorem 7.1 that such chains have nice ergodic properties.

Theorem 7.1. Suppose that $\mathbf{\Phi}$ is a T-chain, and is bounded in probability. Then

(i) $L(x, \sum H_i) \equiv 1$ and each Harris set H_i is positive with invariant probability measure π_i. We have the relations

$$\tilde{\pi}\{A\} = \lim_{k\to\infty} \Pi(\Phi_k, A) \quad \text{a.s. and} \quad \Pi(x, A) = \boldsymbol{E}_x[\tilde{\pi}\{A\}]; \tag{21}$$

for each $x \in \boldsymbol{X}$

$$\frac{1}{N}\sum_{k=1}^{N} P^k(x, \cdot) \to \Pi(x, \cdot) \quad \text{in total variation norm;}$$

and for each $g \in L^1(\boldsymbol{X}, \mathscr{B}, \Pi(x, \cdot))$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} g(\Phi_k) = \int g\, d\tilde{\pi} \quad \text{a.s. } [\boldsymbol{P}_x].$$

(ii) Suppose that the state space $\boldsymbol{X}$ is compact. Then there exists $\rho < 1$, $R < \infty$, and $m \geqq 1$ such that

$$\left\| \frac{1}{m}\sum_{i=1}^{m} P^{k+i}(x, \cdot) - \Pi(x, \cdot) \right\| \leqq R\rho^k, \qquad x \in \boldsymbol{X}, \qquad k \in \mathbb{Z}_+. \tag{22}$$

Proof. Under the conditions of (i) it follows from the decomposition theorem that every Harris set is positive. The relations (21) follow directly from the fact that for each $x \in \boldsymbol{X}$,

$$\lim_{k\to\infty} L(\Phi_k, A) = \mathbf{1}\{\mathbf{\Phi} \in A \text{ i.o.}\} \quad \text{a.s. } [\boldsymbol{P}_x].$$

Let τ denote the first-entrance time to $E^c = \sum H_i$. By the strong Markov property, for each $g \in L^1(\boldsymbol{X}, \mathscr{B}, \Pi(x, \cdot))$

$$\boldsymbol{P}_x\left\{\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} g(\Phi_k) = \int g\, d\tilde{\pi}\right\} = \boldsymbol{E}_x\left[\boldsymbol{P}_{\Phi_\tau}\left\{\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} g(\Phi_k) = \int g\, d\tilde{\pi}\right\}\right].$$

By a result of [2],

$$\boldsymbol{P}_{\Phi_\tau}\left\{\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} g(\Phi_k) = \int g\, d\tilde{\pi}\right\} = 1,$$

and this proves the sample path limit in (i).

The other limit in (i) is proved in a similar way, although the technical details are somewhat more difficult. Write the event that Φ_k lies in a given set A as the disjoint union

$$\boldsymbol{P}_x\{\Phi_k \in A\} = \boldsymbol{P}_x\left\{\{\Phi_k \in A\} \cap \{\tau > k\} \cup \bigcup_{i=1}^{k} \{\Phi_k \in A\} \cap \{\tau = i\}\right\}.$$

For ease of manipulation, translate this expression into operator-theoretic notation [26], to obtain

$$P^k = P^k I_E + \sum_{i=1}^{k} [(PI_E)^{i-1} PI_{E^c} P^{k-i}], \tag{23}$$

which by (20) implies that

$$\begin{aligned} P^k - \Pi &= P^k(I - \Pi) \\ &= \sum_{i=1}^{k} [(PI_E)^{i-1} PI_{E^c}(P^{k-i} - \Pi)] + P^k I_E(I - \Pi). \end{aligned} \tag{24}$$

Hence for all $M \in \mathbb{Z}_+$,

$$\begin{aligned} \frac{1}{M}\sum_{k=1}^{M} (P^k - \Pi) &= \frac{1}{M}\sum_{k=1}^{M}\sum_{i=1}^{k} [(PI_E)^{i-1} PI_{E^c}(P^{k-i} - \Pi)] + \frac{1}{M}\sum_{k=1}^{M} P^k I_E(I - \Pi) \\ &= \frac{1}{M}\sum_{i=1}^{M}\sum_{k=i}^{M} [(PI_E)^{i-1} PI_{E^c}(P^{k-i} - \Pi)] + \frac{1}{M}\sum_{k=1}^{M} P^k I_E(I - \Pi) \end{aligned}$$

from which it follows that

$$\begin{aligned} \left\|\frac{1}{M}\sum_{k=1}^{M} (P^k - \Pi)\right\| &\leqq \left\|\sum_{i=1}^{M} \frac{M-i}{M} (PI_E)^{i-1} PI_{E^c}\left[\frac{1}{M-i}\sum_{k=0}^{M-i} (P^k - \Pi)\right]\right\| \\ &\quad + \frac{1}{M}\sum_{k=1}^{M} P^k(x, E) \end{aligned} \tag{25}$$

where $\|\cdot\|$ denotes the total variation norm. The second summand converges to 0 as $M \to \infty$ since $P^k(x, E)$ converges to 0. As for the first term, write

$$\begin{aligned} &\left\|\sum_{i=1}^{M} \frac{M-i}{M} (PI_E)^{i-1} PI_{E^c}\left[\frac{1}{M-i}\sum_{k=0}^{M-i} (P^k - \Pi)\right]\right\| \\ &\quad \leqq \left\|\sum_{i=1}^{N} \frac{M-i}{M} (PI_E)^{i-1} PI_{E^c}\left[\frac{1}{M-i}\sum_{k=0}^{M-i} (P^k - \Pi)\right]\right\| + \sum_{i=N+1}^{\infty} \|(PI_E)^{i-1} PI_{E^c}\| \end{aligned}$$

where $N < M$ is fixed. The first term converges to 0 as $M \to \infty$ by a standard total variation norm limit theorem for positive Harris chains (cf. [26]) and the dominated convergence theorem. The second term is equal to $\boldsymbol{P}_{(\cdot)}\{\tau > N\}$ and hence converges to 0 as $N \to \infty$.

We now prove (ii). Let K_i denote the compact set $K_i \triangleq \{y \in \boldsymbol{X} : T(y, H_i^c) = 0\}$. The inclusion $H_i \subset K_i$ follows from the fact that H_i is absorbing. Let $A \subset H_i$ be a petite set of positive π_i-measure. Then by the definition of Harris recurrence, $K_\varepsilon(y, A) > 0$ for each $y \in H_i$. Furthermore, since T is non-trivial, $T(z, H_i) > 0$ for each $z \in K_i$. Hence by lower semi-continuity we have for some $\delta > 0$,

$$K_{a\star\varepsilon}(z, A) \geqq \int_{H_i} T(z, dy) K_\varepsilon(y, A) \geqq \delta$$

for all $z \in K_i$, and hence for $z \in H_i$. By Lemma 3.1 the Harris set H_i is petite, and hence by Theorem 6.1 there exists $r \in \mathbb{Z}_+$, and $\rho_0 < 1$ such that (22) holds for initial conditions $x \in E^c = \bigcup H_i$.

Observe that since $\mathbf{\Phi}$ is necessarily bounded in probability, it follows from part (iii) of the decomposition theorem that $K_\varepsilon(x, E^c) > 0$ for every $x \in \mathbf{X}$. Repeating the previous argument, we may deduce that $K_{a\star\varepsilon}(x, E^c)$ is bounded away from 0 on $\mathbf{X}$. Hence for some $N \geqq 1$, and $\delta > 0$,

$$\frac{1}{N} \sum_{k=1}^{N} P^k(x, E^c) \geqq \delta, \qquad x \in \mathbf{X}$$

and since $P^k(x, E^c)$ is an increasing function of k, it follows that $P^N(x, E^c) \geqq \delta$. This implies that $P^{kN}(x, E) \leqq (1-\delta)^k$, and since $P^k(x, E)$ is decreasing, there exists $M_1 < \infty$, and $\rho_1 < 1$ such that $P^k(x, E) \leqq M_1 \rho_1^k$ for all $x \in \mathbf{X}$, $k \geqq 0$. Proceeding as in the proof of (i), we may estimate as follows. For any $l \in \mathbb{Z}_+$,

$$\begin{aligned} \left\| \frac{1}{m} \sum_{k=l+1}^{l+m} (P^k - \Pi) \right\| &\leqq \left\| \sum_{i=1}^{l} (PI_E)^{i-1} PI_{E^c} \left[\frac{1}{m} \sum_{k=l+1}^{l+m} (P^{k-i} - \Pi) \right] \right\| \\ &\quad + \frac{1}{m} \sum_{k=l+1}^{l+m} \sum_{i=l+1}^{l+m} [\|P^k I_E\| + \|P^{i-1} I_E\|]. \end{aligned} \tag{26}$$

The second term is less than or equal to $2m \, \|P^l(x, E)\| \leqq 2mM_1\rho_1^l$ for all x. As for the first term, since (22) holds with $\rho = \rho_0$, $M = M_0$ for initial conditions $x \in E^c$, we have for all $l \geqq 1$,

$$\begin{aligned} \left\| \sum_{i=1}^{l} (PI_E)^{i-1} PI_{E^c} \left[\frac{1}{m} \sum_{k=l+1}^{l+m} (P^{k-i} - \Pi) \right] \right\| &\leqq \sum_{i=1}^{l} P^{i-1}(x, E)(M_0 \rho_0^{l-i}) \\ &\leqq \sum_{i=1}^{l} (M_1 \rho_1^{i-1}(M_0 \rho_0^{l-i}) \\ &\leqq M\rho^l \end{aligned}$$

where $1 > \rho > \max(\rho_0, \rho_1)$, and M is a suitably large constant.

8. Criteria for strong mixing

In this section we demonstrate that a Markov chain satisfies a mixing condition when a suitable Lyapunov–Foster criterion is satisfied. The results of this section will be applied to prove limit theorems in the final section.

Suppose that for a measurable function $f: \boldsymbol{X} \to \mathbb{R}$ the process $\{V_k\}$ satisfies the following condition.

For some $\lambda < 1$, $\varepsilon > 0$ and a petite set A,

(DD5) $$\boldsymbol{E}[V_{k+1} \mid \mathcal{F}_k] \leqq \lambda V_k - \varepsilon(1 + f^2(\Phi_k)) \quad \text{a.s. } [\boldsymbol{P}_{\Phi_0}],$$

when $\sigma_A > k$, $k \in \mathbb{Z}_+$.

If (DD4) and (9) hold, and if $\varepsilon(1+f^2) \leqq V$, then (DD5) may be obtained by the same argument used in the proof of Theorem 6.3.

The stochastic process $\mathbf{x}$ is called *strong mixing* if there exists a sequence of positive numbers $\{\delta(n): n \geqq 0\}$ tending to 0 for which

$$\sup |\boldsymbol{E}[g(x_k)h(x_{n+k})] - \boldsymbol{E}[g(x_k)]\boldsymbol{E}[h(x_{n+k})]| \leqq \delta(n), \qquad n \in \mathbb{Z}_+,$$

where the supremum is taken over all $k \in \mathbb{Z}_+$, and all measurable g and h such that $|g(x)|$, $|h(x)| \leqq 1$ for all $x \in \boldsymbol{X}$. Observe that when $\boldsymbol{\Phi}$ is a *stationary Markov chain*, strong mixing becomes

$$\sup |\boldsymbol{E}_\pi[\bar{g}(\Phi_0)\bar{h}(\Phi_n)]| \leqq \delta(n), \qquad n \in \mathbb{Z}_+,$$

where π is an invariant probability, and $\bar{g}, \bar{h}$ denote the centered functions $g - \int g\, d\pi$, $h - \int h\, d\pi$ respectively.

It is shown in [3] that a positive Harris-recurrent Markov chain is always strong mixing for some δ. The next result strengthens this result when condition (DD5) is satisfied in two directions: a geometric rate of convergence for δ is established, and the supremum over the class of bounded functions in the definition of strong mixing is extended to include any function which is dominated by f.

We consider here the strongly aperiodic case. We have recently learned that a similar result has been established concurrently using different methods in [10].

Theorem 8.1. Suppose that $\boldsymbol{\Phi}$ is strongly aperiodic, that (DD5) is satisfied, and that $f^2 + PV_0$ is bounded on A. Then there exists a constant $R < \infty$ and $\rho < 1$ such that

$$\sup |\boldsymbol{E}_x[g(\Phi_k)h(\Phi_{n+k})] - \boldsymbol{E}_x[g(\Phi_k)]\boldsymbol{E}_x[h(\Phi_{n+k})]| \leqq R(PV_0(x) + f^2(x) + 1)\rho^n, \quad n \in \mathbb{Z}_+$$

where the supremum is taken over all $k \in \mathbb{Z}_+$, and all measurable functions g, h satisfying $|g|$, $|h| \leqq f$. In particular, $\boldsymbol{\Phi}$ is strong mixing for each initial condition.

Proof. Fix $|g|$, $|h| \leqq f$, $x \in \boldsymbol{X}$, and with $\mu_0 \triangleq \delta_x$, let μ_k denote the probability $\mu_0 P^k$. Recall that $\bar{g}, \bar{h}$ denote the centered functions $g - \int g\, d\pi$, $h - \int h\, d\pi$ respectively. It is easily shown that for all $n, k \in \mathbb{Z}_+$,

$$\begin{aligned}\boldsymbol{E}_{\mu_0}[g(\Phi_k)h(\Phi_{k+n})] &- \boldsymbol{E}_{\mu_0}[g(\Phi_k)]\boldsymbol{E}_{\mu_0}[h(\Phi_{k+n})] \\ &= \boldsymbol{E}_{\mu_k}[\bar{g}(\Phi_0)\bar{h}(\Phi_n)] - \boldsymbol{E}_{\mu_k}[\bar{g}(\Phi_0)]\boldsymbol{E}_{\mu_k}[\bar{h}(\Phi_n)].\end{aligned}$$

By the definitions, the corollary to Theorem 6.2, and Jensen's inequality we have for some $\rho < 1$,

$$\begin{aligned}|E_{\mu_k}[\bar{g}(\Phi_0)]E_{\mu_k}[\bar{h}(\Phi_n)]| &\leqq \|\mu_0 P^k - \pi\|_f \; \|\mu_0 P^{k+n} - \pi\|_f \\ &\leqq \|\mu_0 P^k - \pi\|_f^2 + \|\mu_0 P^{k+n} - \pi\|_f^2 \\ &\leqq \|\mu_0 P^k - \pi\|_{f^2+1} + \|\mu_0 P^{k+n} - \pi\|_{f^2+1} \\ &= O((PV_0(x) + f^2(x) + 1)\rho^k).\end{aligned}$$

Hence to prove the theorem it is enough to bound $E_{\mu_k}[\bar{g}(\Phi_0)\bar{h}(\Phi_n)]$ by a geometric series.

We prove here that the function

$$M_k(z) = \sum_{n=0}^{\infty} E_{\mu_k}[\bar{g}(\Phi_0)\bar{h}(\Phi_n)]z^n \tag{27}$$

is analytic and appropriately bounded in an open ball containing the closed unit disk $\bar{U} \subset \mathbb{C}$.

Let $\boldsymbol{\alpha}$ denote an atom on the enlarged state space. By conditioning at time $\tau_{\boldsymbol{\alpha}}$, the function M_k may be written as the sum of an analytic function, and the product of two analytic functions, where the interchange of summation and expectation is justified by Fubini's theorem:

$$M_k(z) = \sum_{n=0}^{\infty} E_{\mu_k}[\bar{g}(\Phi_0)\bar{h}(\Phi_n)\mathbf{1}(n < \tau_{\boldsymbol{\alpha}})]z^n \tag{28}$$

$$+ \left(\sum_{n=0}^{\infty} E_{\mu_k}[\bar{g}(\Phi_0)\mathbf{1}(n = \tau_{\boldsymbol{\alpha}})]z^n\right) \tag{29}$$

$$\times \left(\sum_{n=1}^{\infty} E_{\boldsymbol{\alpha}}[\bar{h}(\Phi_n)]z^n\right). \tag{30}$$

By Lemma 6.3, to prove the theorem it is sufficient to demonstrate that each of these functions is analytic and bounded, uniformly in k, and g and h satisfying the conditions of the theorem, in a fixed neighborhood of the closed unit disk in $\mathbb{C}$. To bound the functions on lines (28) and (29) we use the estimate $xy \leqq \varepsilon x^2 + \varepsilon^{-1}y^2$:

$$\sum_{n=0}^{\infty} |E_{\mu_k}[\bar{g}(\Phi_0)\bar{h}(\Phi_n)\mathbf{1}(n < \tau_{\boldsymbol{\alpha}})]|r^n \leqq \frac{E_{\mu_k}[\bar{g}^2(\Phi_0)]}{1 - r^{-1}} + E_{\mu_k}\left[\sum_{n=0}^{\tau_{\boldsymbol{\alpha}}} \bar{h}^2(\Phi_n)r^{2^n}\right]$$

$$\sum_{n=0}^{\infty} |E_{\mu_k}[\bar{g}(\Phi_0)\mathbf{1}(n = \tau_{\boldsymbol{\alpha}})]|r^n \leqq \frac{E_{\mu_k}[\bar{g}^2(\Phi_0)]}{1 - r^{-1}} + E_{\mu_k}[r^{2\tau_{\boldsymbol{\alpha}}}].$$

In both instances, the right-hand side is bounded by a fixed constant times $(PV_0(x) + f^2(x) + 1)$ for some $r > 1$, where the constant is independent of k, g, and h, by Lemma 6.1, Lemma 6.2, and Theorem 6.2.

The function on line (30) may be bounded as follows:

$$\sum_{n=1}^{\infty} |E_{\boldsymbol{\alpha}}[\bar{h}(\Phi_n)]|r^n \leqq \sum_{n=1}^{\infty} \|P^n(\boldsymbol{\alpha}, \cdot) - \pi\|_f r^n$$

which is finite for some r by the corollary to Theorem 6.2.

9. The central limit theorem and law of the iterated logarithm

We conclude by presenting versions of the central limit theorem and law of the iterated logarithm, two limit theorems which describe the asymptotic behaviour of normalized sample path averages of the chain $\{f(\Phi_k)\}$.

An obvious necessary condition for either of these results is the existence of a second moment for the random variable $f(\Phi_k)$, with respect to an invariant probability. The results of [8], [11], [5] show that these and related results may be obtained under a mixing condition on the chain. From these considerations, Theorem 6.2 and Theorem 8.1, we see that (DD5) allows us to meet each of these criteria, and in fact when (DD5) holds we can prove both a central limit theorem and law of the iterated logarithm.

We again consider the strongly aperiodic case and refer to [22] for extensions to general irreducible chains.

Theorem 9.1. Suppose that $\mathbf{\Phi}$ is strongly aperiodic, that $f: \mathbf{X} \to \mathbb{R}$ is a measurable function, and that (DD5) holds, with $f^2 + PV_0$ bounded on A. Then $\mathbf{\Phi}$ is positive Harris recurrent with invariant probability measure π, and the limit

$$\gamma_f^2 = \lim_{n\to\infty} \frac{1}{n} \boldsymbol{E}_\pi\left[\left(\sum_{k=1}^{n} \bar{f}(\Phi_k)\right)^2\right] \tag{31}$$

exists, and is finite. We also have

$$\gamma_f^2 = \sum_{k=-\infty}^{\infty} \boldsymbol{E}_\pi[\bar{f}(\Phi_0)\bar{f}(\Phi_k)] \tag{32}$$

where the sum converges absolutely.

If $\gamma_f^2 > 0$, then the following limits hold for each initial condition:

(i) $\displaystyle \lim_{n\to\infty} \boldsymbol{P}_x\left\{\frac{1}{\sqrt{n\gamma_f^2}} \sum_{k=1}^{n} \bar{f}(\Phi_k) \leq t\right\} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\, dx$

(ii) the limit infimum and limit supremum of the sum

$$\frac{1}{\sqrt{2\gamma_f^2 n \log\log(n)}} \sum_{k=1}^{n} \bar{f}(\Phi_k)$$

are respectively -1 and $+1$ with probability 1.

To prove these results we will again consider the split chain. We first use (DD5) to obtain hypothesis (7) on page 102 of [7].

Lemma 9.1. If condition (DD5) holds, and if $f^2 + PV_0$ is bounded on A, then an atom $\boldsymbol{\alpha}$ exists with

$$\boldsymbol{E}_\mu\left[\left(\sum_{k=1}^{\tau_\alpha} |f(\Phi_k)|\right)^2\right] < \infty, \qquad \boldsymbol{E}_\mu[\tau_\alpha^2] < \infty,$$

for any of the initial distributions $\mu = \delta_\alpha$, $\mu = \delta_x$ for $x \in \mathbf{X}$, or $\mu = \pi$.

Proof. Observe that for any measurable function $g: \boldsymbol{X} \to \mathbb{R}_+$,

$$\boldsymbol{E}_\mu\Big[\Big(\sum_{k=1}^{\tau_\alpha} g(\Phi_k)\Big)^2\Big] = \boldsymbol{E}_\mu\Big[\sum_{j=1}^{\tau_\alpha}\sum_{i=1}^{\tau_\alpha} g(\Phi_j)g(\Phi_i)\Big]$$

$$\leqq 1/2\boldsymbol{E}_\mu\Big[\sum_{j=1}^{\tau_\alpha}\sum_{i=1}^{\tau_\alpha} r^{j-i}g^2(\Phi_j) + r^{i-j}g^2(\Phi_i)\Big]$$

$$\leqq \frac{1}{r-1}\boldsymbol{E}_\mu\Big[\sum_{j=1}^{\tau_\alpha} r^j g^2(\Phi_j)\Big]$$

where we have used the inequality $2xy \leqq \varepsilon x^2 + \varepsilon^{-1}y^2$. Applying Lemma 6.1 with $g = |f|$ and $g \equiv 1$ gives the desired result.

Proof of Theorem 9.1. Since an atom exists for the split chain, and since the proofs of the central limit theorem and law of the iterated logarithm are based on the existence of, and mean return times to, atoms, we may adapt the proofs of [7]. Lemma 9.1 shows that hypothesis (7) on page 102 of [7] is satisfied. This implies that the hypotheses of both the central limit theorem (Theorem 1 on page 99 of [7]), and the law of the iterated logarithm (Theorem 5 on page 106 [7]) are satisfied. Similarly, (31) holds by Theorem 3 on page 102 of [7]. It is assumed in this result that the initial distribution has finite support, which is used only to obtain the bound

$$\boldsymbol{E}_\pi[(Y')^2] = \boldsymbol{E}_\pi\Big[\Big(\sum_{i=1}^{\min(n,\tau_\alpha-1)} f(\Phi_i)\Big)^2\Big] < \infty.$$

This bound is obtained for the stationary version of the chain by Lemma 9.1. When $\boldsymbol{\Phi}$ is aperiodic we have by a simple calculation [8],

$$\frac{1}{n}\boldsymbol{E}_\pi\Big[\Big(\sum_{k=0}^{n} \bar{f}(\Phi_k)\Big)^2\Big] = r(0) + \frac{2}{n}\sum_{k=1}^{n-1} R(k)$$

where $r(i) \triangleq \boldsymbol{E}_\pi[\bar{f}(\Phi_k)\bar{f}(\Phi_{k+i})]$, which is independent of k by stationarity, and $R(k) \triangleq \sum_{i=1}^{k} r(i)$. By Theorem 8.1 the series $\sum_{i=1}^{\infty} r(i)$ is absolutely convergent, and hence the expression (32) holds.

References

[1] Athreya, K. B. and Ney, P. (1978) A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245**, 493–501.

[2] Athreya, K. B. and Ney, P. (1980) Some aspects of ergodic theory and laws of large numbers for Harris recurrent Markov chains. *Colloq. Math. Soc. J. Bolyai* **32**, 41–56.

[3] Athreya, K. B. and Pantula, S. G. (1986) Mixing properties of Harris chains and autoregressive processes. *J. Appl. Prob.* **23**, 880–892.

[4] Bhatia, N. P. and Szegö, G. P. (1970) *Stability Theory of Dynamical Systems.* Springer-Verlag, Berlin.

[5] Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

[6] Chan, K. S. (1989) A note on the geometric ergodicity of a Markov chain. *Adv. Appl. Prob.* **21,** 702–704.

[7] Chung, K. L. (1967) *Markov Chains with Stationary Transition Probabilities,* 2nd edn. Springer-Verlag, Berlin.

[8] Cogburn, R. (1972) The central limit theory for Markov processes. *Proc. 6th Berkeley Symp. Math. Statist. Prob.,* pp. 485–512.

[9] Cogburn, R. (1975) A uniform theory for sums of Markov chain transition probabilities. *Ann. Prob.* **3,** 191–214.

[10] Diebolt, J. and Guégan, D. (1990) Probability properties of the general non-linear Markovian process of order one and applications to time series modelling. Technical 125, Laboratoire de Statistique Théorique et Appliquée, CNRS-URA, Novembre 1990.

[11] Doob, J. L. (1953) *Stochastic Processes.* Wiley, New York.

[12] Foguel, S. R. (1969) Positive operators on $C(X)$. *Proc. Amer. Math. Soc.* **22,** 295–297.

[13] Foster, F. G. (1953) On the stochastic matrices associated with certain queueing processes. *Ann. Math. Statist.* **24,** 355–360.

[14] Has'minskii, R. Z. (1980) *Stochastic Stability of Differential Equations.* Sijthoff & Noordhoff, Alphen an den Rijn, The Netherlands.

[15] Kelley, J. L. (1955) *General Topology.* Van Nostrand, Princeton, NJ.

[16] Lamperti, J. (1960) Criteria for the recurrence or transience of stochastic processes I. *J. Math. Anal. Appl.* **1,** 314–330.

[17] Meyn, S. P. (1989) Ergodic theorems for discrete time stochastic systems using a stochastic Lyapunov function. *SIAM J. Control Optim.* **27,** 1409–1439.

[18] Meyn, S. P. and Caines, P. E. (1991) Asymptotic behavior of stochastic systems possessing Markovian realizations. *SIAM J. Control Optim.* **29,** 535–561.

[19] Meyn, S. P. and Guo, L. (1990) Adaptive control of time varying stochastic systems. In *Proc. 11th IFAC World Cong.* Tallinn, Estonia, ed. V. Utkin and O. Jaaksoo, Vol. 3, pp. 198–202.

[20] Meyn, S. P. and Guo, L. (1992) Geometric ergodicity of a bilinear time series model. *J. Time Series Anal.*

[21] Meyn, S. P. and Guo, L. (1992) Stability, convergence, and performance of an adaptive control algorithm applied to a randomly varying system. *IEEE Trans. Autom. Control.* **3,** 535–540.

[22] Meyn, S. P. and Tweedie, R. L. (1992) *Markov Chains and Stochastic Stability.* Control and Communication in Engineering. Springer-Verlag, Berlin.

[23] Meyn, S. P. and Tweedie, R. L. (1992) Stability of Markovian processes II: Continuous time processes and sampled chains.

[24] Meyn, S. P. and Tweedie, R. L. (1992) Stability of Markovian processes III: Foster–Lyapunov criteria for continuous time processes.

[25] Nummelin, E. (1978) A splitting technique for Harris recurrent chains. *Z. Wahrscheinlichkeitsth.* **43,** 309–318.

[26] Nummelin, E. (1984) *General Irreducible Markov Chains and Non-Negative Operators.* Cambridge University Press.

[27] Nummelin, E. and Tuominen, P. (1982) Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stoch. Proc. Appl.* **12,** 187–202.

[28] Nummelin, E. and Tweedie, R. L. (1978) Geometric ergodicity and R-positivity for general Markov chains. *Ann. Prob.* **6,** 404–420.

[29] Orey, S. (1971) *Limit Theorems for Markov Chain Transition Probabilities.* Van Nostrand Reinhold Mathematical Studies **34,** London.

[30] Pollard, D. B. and Tweedie, R. L. (1976) R-theory for Markov chains on a topological state space II. *Z. Wahrscheinlichkeitsth.* **34,** 269–278.

[31] Revuz, D. (1984) *Markov Chains.* North-Holland, Amsterdam.

[32] Spieksma, F. M. (1990) Geometrically Ergodic Markov Chains and the Optimal Control of Queues. Ph.D. Thesis, University of Leiden.

[33] Tuominen, P. and Tweedie, R. L. (1979) Markov chains with continuous components. *Proc. London Math. Soc* 3(**38**), 89–114.

[34] Tweedie, R. L. (1975) Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Proc. Appl.* **3,** 385–403.

[35] TWEEDIE, R. L. (1976) Criteria for classifying general Markov chains. *Adv. Appl. Prob.* **8,** 737–771.

[36] TWEEDIE, R. L. (1983) Criteria for rates of convergence of Markov chains with application to queueing and storage theory. In *Probability, Statistics and Analysis,* ed. J. F. C. Kingman and G. E. H. Reuter, London Mathematical Society Lecture Note Series, Cambridge University Press.

[37] TWEEDIE, R. L. (1988) Invariant measures for Markov chains with no irreducibility assumptions. *J. Appl. Prob.* **25A,** 275–285.

[38] VERE-JONES, D. (1969) Some limit theorems for evanescent processes. *Austral. J. Statist.* **11,** 67–78.