Stochastic Processes and their Applications 173 (2024) 104364

Contents lists available at ScienceDirect

Stochastic Processes and their Applications

journal homepage: www.elsevier.com/locate/spa

Existence and uniqueness of quasi-stationary and quasi-ergodic measures for absorbing Markov chains: A Banach lattice approach

Matheus M. Castro [a,*], Jeroen S.W. Lamb [a,b,c], Guillermo Olicón-Méndez [d], Martin Rasmussen [a]

[a] Department of Mathematics, Imperial College London, Queen's Gate, London, SW7 2AZ, United Kingdom
[b] International Research Center for Neurointelligence, The University of Tokyo, Tokyo, 113-0033, Japan
[c] Centre for Applied Mathematics and Bioinformatics, Department of Mathematics and Natural Sciences, Gulf University for Science and Technology, Halwally, Kuwait
[d] Institut für Mathematik, Universität Potsdam, Karl-Liebknecht-Str. 24-25, Potsdam, 14476, Germany

ARTICLE INFO

MSC:
37H05
47B65
60J05

Keywords:
Markov chains with absorption
Banach lattice
Quasi-stationary measure
Quasi-ergodic measure
Yaglom limit

ABSTRACT

We establish the existence and uniqueness of quasi-stationary and quasi-ergodic measures for almost surely absorbed discrete-time Markov chains under weak conditions. We obtain our results by exploiting Banach lattice properties of transition functions under natural regularity assumptions.

1. Introduction and motivation

The existence and uniqueness of quasi-stationary and quasi-ergodic measures is a central question in the theory of absorbing Markov processes, but sufficient conditions have been tailored to specific contexts such as stochastic differential equations [1–6]. In this paper, we prove the existence and uniqueness of quasi-stationary and quasi-ergodic measures for absorbing Markov chains under weak continuity and irreducibility assumptions. Our results substantially extend the settings in which quasi-stationary and quasi-ergodic measures are known to exist, including in particular random systems with bounded noise.

We use Banach lattices [7–9] to address the problem of the existence and uniqueness of quasi-stationary and quasi-ergodic measures from a functional-analytic point of view by considering transition probabilities as a bounded linear operator. Banach lattices allow us to study its spectrum and subsequently construct the desired measures. Previously, Banach lattice theory has been employed to construct stationary measures and obtain statistical results for conservative Markov processes without absorption (see [10–13]). In this paper, we extend the use of this theory to absorbing Markov chains.

We consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ on a metric space E, and study the behaviour of this chain conditioned on survival in some compact set $M \subset E$. Recall that a Borel measure μ on M is called *quasi-stationary measure* for X_n if for every $n \in \mathbb{N}$ and measurable set $A \subset M$, we have

$$\mathbb{P}_\mu[X_n \in A \mid \tau > n] := \frac{\int_M \mathbb{P}_x[X_n \in A]\mu(\mathrm{d}x)}{\int_M \mathbb{P}_x[X_n \in M]\mu(\mathrm{d}x)} = \mu(A),$$

* Corresponding author.
E-mail address: mmd20@ic.ac.uk (M.M. Castro).

https://doi.org/10.1016/j.spa.2024.104364
Received 20 May 2022; Received in revised form 19 October 2023; Accepted 21 April 2024
Available online 24 April 2024
0304-4149/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

where τ is the stopping time for X_n conditioned upon survival in M.

Quasi-stationary measures generalise stationary measures [6, Section 5] in the sense that they allow for escape from the state space. They have received increased attention in recent years [2–4]. We refer to [14] for a bibliography on the topic. Quasi-stationary measures may be used for a variety of purposes, including determining average escape times (see [6, Lemma 5.6]). In addition, in several contexts, it can be shown that the so-called *Yaglom limit*,

$$\lim_{n\to\infty} \mathbb{P}_x\left[X_n \in A \mid \tau > n\right] := \lim_{n\to\infty} \frac{\mathbb{P}_x\left[X_n \in A\right]}{\mathbb{P}_x\left[X_n \in M\right]}, \tag{1}$$

converges to the unique quasi-stationary measure [15–17]. A commonly employed technique for establishing the existence and uniqueness of quasi-stationary measures for X_n, as well as the existence of the Yaglom limit, relies on showing that the transition kernel of the process can be interpreted as a compact operator on a suitable space, together with the presence of a spectral gap (see [1,18–20]). In this paper, we provide easy-to-check sufficient conditions that guarantee these properties.

In contrast to the ergodic stationary case, the quasi-stationary measure is not the correct object to employ when studying Birkhoff averages for absorbed processes. Although quasi-stationary measures are relevant for statistics in terms of the limit (1), one needs to introduce the notion of quasi-ergodic measures in order to compute conditioned Birkhoff averages. Quasi-ergodic measures are probability measures η on M that satisfy

$$\lim_{n\to\infty} \mathbb{E}_x\left[\frac{1}{n}\sum_{i=0}^{n-1} g\circ X_i \;\middle|\; \tau > n\right] = \int_M g(y)\eta(\mathrm{d}y) \text{ for } \eta\text{-a.e. } x \in M \tag{2}$$

for every bounded measurable function $g : M \to \mathbb{R}$.

While quasi-stationary measures have been well-studied, less is known about quasi-ergodic measures. Quasi-ergodic limits of the form (2) were first considered by Darroch and Seneta [21], who established their existence for finite irreducible Markov chains. Their study was followed by Breyer and Roberts [22] and later by Champagnat and Villemonais [2,3] to obtain conditions that guarantee the existence of a quasi-ergodic measure for Markov processes defined in a general state space, also providing a description of the quasi-ergodic measure in relation to the quasi-stationary measure. Explicit formulas for quasi-ergodic measures for reducible finite absorbing Markov chains have been obtained in [23]. Quasi-ergodic measures are promising tools for analysing random dynamical systems; for instance, they are crucial for the existence of so-called conditioned Lyapunov exponents [24].

Recently, numerous contributions have been made to understand the Yaglom limit [25–29]. In particular, in [2], necessary and sufficient conditions for quasi-stationary measures have been established, addressing the issue of uniform exponential convergence. Although these conditions are sharp in the context of uniform exponential convergence, verifying them in applications outside the stochastic differential equations framework is complicated due to their abstract formulation. Moreover, they rely on the standing assumption that the probability of remaining in the allowed state space M is positive for all positive times, and all initial conditions $x \in M$. Notice that typically, this assumption is not fulfilled on *repelling sets* of random maps with bounded noise (see Example 2.6).

Additionally to the uniform convergence scenario, [3,30–32], have provided valuable insights. These studies have examined cases where the (1) converges to a unique quasi-stationary measure in the total variation norm, although this convergence is not uniform in x. This is relevant since uniform convergence of (1) in the total variation norm is often absent in random dynamical systems with bounded noise; see Section 2.1 below. The existence of quasi-stationary and quasi-ergodic measures in such settings is established in this paper.

As in [2,3], alternative conditions for the existence of quasi-stationary measures found in the literature are also aimed at such stochastic differential equations [33–36]. These conditions are based either on lower-bound estimations of $\mathbb{P}_x[X_n \in \cdot \mid \tau > n]$ (such as the Doeblin condition) or on the existence of a Lyapunov function for the transition kernel. Even in elementary examples, these conditions are usually difficult to verify, and their dynamical interpretation is often unclear. In light of this, we establish the existence of both quasi-stationary and quasi-ergodic measures under natural, weak and easily verifiable dynamical conditions.

This paper is divided into six sections and one appendix. In Section 2, we briefly recall the basic concepts of the theory of absorbing Markov chains. We introduce the main underlying hypothesis of this paper (Hypothesis (H)), we state the main results of this paper (Theorems A, B and C), and apply them to some fundamental examples. In Section 3, we derive some direct consequences of Hypothesis (H). Section 4 is dedicated to a brief presentation of Banach lattice theory, and the proof of Theorem 4.4. This is central to the proof of our main results. In Section 5, we show the existence and uniqueness of quasi-stationary measures for Markov chains that fulfil Hypothesis (H), and prove Theorem A. Finally, in Section 6, we prove the existence and uniqueness of quasi-ergodic measures under Hypothesis (H), and we also prove Theorem B and C. Appendix contains the proof of Lemma 6.4, a technical yet essential result for the existence and uniqueness of quasi-ergodic measures.

2. Main results

Let (E, d) be a metric space and consider a compact subset $M \subset E$ endowed with the induced topology. We assume that $(M, \mathscr{B}(M), \rho)$ is a Borel finite measure space, where $\mathscr{B}(M)$ denotes the Borel σ-algebra of M. Throughout this paper, we aim to study Markov chains that are killed when the process escapes the region M. Since the behaviour of X_n in the set $E \setminus M$ is not relevant for the desired analysis, we identify $E \setminus M$ as a single point ∂ and consider $E_M = M \sqcup \{\partial\}$ as the topological space generated by the topological basis

$$\beta = \{B_r(x);\ x \in M \text{ and } r \in \mathbb{R}\} \cup \{\partial\},$$

where $B_r(x) := \{y \in M; d(x,y) < r\}$, and $\sqcup$ denotes disjoint union.

In this paper, we assume that

$$X := \left(\Omega, \{\mathcal{F}_n\}_{n\in\mathbb{N}_0}, \{X_n\}_{n\in\mathbb{N}_0}, \{\mathcal{P}^n\}_{n\in\mathbb{N}_0}, \{\mathbb{P}_x\}_{x\in E_M}\right)$$

is a Markov chain with state space E_M, in the sense of [37, Definition III.1.1]. This means that the pair $(\Omega, \{\mathcal{F}_n\}_{n\in\mathbb{N}})$ is a filtered space; X_n is an $\mathcal{F}_n$-adapted process with state space E_M; $\mathcal{P}^n$ a time-homogeneous transition probability function of the process X_n satisfying the usual measurability assumptions and Chapman–Kolmogorov equation; $\{\mathbb{P}_x\}_{x\in E_M}$ is a family of probability function satisfying $\mathbb{P}_x[X_0 = x] = 1$ for every $x \in E_M$; and for all $m, n \in \mathbb{N}_0$, $x \in E_M$, and every bounded measurable function f on E_M,

$$\mathbb{E}_x\left[f \circ X_{m+n} \mid \mathcal{F}_n\right] = (\mathcal{P}^m f)(X_n)\ \mathbb{P}_x\text{-almost surely.}$$

As mentioned before, we assume that X_n is a Markov chain that is absorbed at ∂, i.e. $\mathcal{P}(\partial, \{\partial\}) = 1$. Given the nature of the process X_n it is natural to define the stopping time $\tau(\omega) := \inf\{n \in \mathbb{N}; X_n(\omega) \notin M\}$.

We introduce some notation that is used throughout the paper.

Notation 2.1. *Given a measure μ on M, we set $\mathbb{P}_\mu(\cdot) := \int_M \mathbb{P}_x(\cdot)\mu(\mathrm{d}x)$.*

We consider the set $\mathcal{F}_b(M)$ as the set of bounded Borel measurable functions on M. Given $f \in \mathcal{F}_b(M)$ write

$$\mathcal{P}^n(f)(x) := \mathcal{P}^n\left(\mathbb{1}_M f\right)(x) = \int_M f(y)\mathcal{P}^n(x, \mathrm{d}y),$$

and $f \circ X_n := \mathbb{1}_M \circ X_n\ f \circ X_n$.

For every $p \in [1, \infty]$ we denote the space $L^p(M, \mathscr{B}(M), \rho)$ by $L^p(M)$; $C^0(M)$ refers to the set of continuous functions $f : M \to M$; and $\mathcal{M}(M)$ the set of Borel signed-measures on M. Finally, we define the sets $C^0_+(M) := \{f \in C^0(M); f \geq 0\}$ and $\mathcal{M}_+(M) := \{\mu \in \mathcal{M}(M);\ \mu(A) \geq 0,$ for every $A \in \mathscr{B}(M)\}$.

In the following, we recall the definition of a quasi-stationary measure.

Definition 2.2. A Borel measure μ on M is said to be a *quasi-stationary measure* for the Markov chain X_n if $\mathbb{P}_\mu\left[X_n \in \cdot \mid \tau > n\right] = \mu(\cdot),$ for all $n \in \mathbb{N}$. We call $\lambda = \int_M \mathcal{P}(x, M)\mu(\mathrm{d}x)$ the *survival rate* of μ.

Remark 2.3. Note that since $\{\partial\}$ is absorbing

$$\mathbb{P}_\mu\left[X_n \in \cdot \mid \tau > n\right] = \frac{\int_M \mathcal{P}^n(x, \cdot)\mu(\mathrm{d}x)}{\int_M \mathcal{P}^n(x, M)\mu(\mathrm{d}x)},$$

for every $\mu \in \mathcal{M}_+(M)$.

Our initial goal is to establish the existence of quasi-stationary and quasi-ergodic measures for a Markov chain X_n. We mention that our results also cover the case where X_n has almost surely escaping points, i.e. points $x \in M$, such that $\mathcal{P}^k(x, M) = 0$ for some $k \in \mathbb{N}$. This occurs naturally in random iterated functions with bounded noise (see Section 2.1).

The sufficient conditions to reach our results are introduced in the form of Hypothesis (H). Briefly, we say that a Markov chain X_n satisfies Hypothesis (H) if its transition kernels $\mathcal{P}(x, dy)$ are well behaved with respect to a fixed Borel measure ρ on M, and if X_n eventually reach any open set of M with positive probability provided $X_0 = x$ is a non-escaping point.

Hypothesis (H). We say that the Markov chain X_n on E_M absorbed at ∂ fulfils *Hypothesis (H)* if the following two properties hold:

(H1) For all $x \in M$, $\mathcal{P}(x, \mathrm{d}y) \ll \rho(\mathrm{d}y)$; and for every $\varepsilon > 0$, there exists $\delta > 0$, such that

$$d(x,z) < \delta \Rightarrow \|g(x,\cdot) - g(z,\cdot)\|_1 := \int_M |g(x,y) - g(z,y)|\,\rho(\mathrm{d}y) < \varepsilon,$$

where g is the Radon–Nikodym derivative

$$g(x,\cdot) := \frac{\mathcal{P}(x, \mathrm{d}y)}{\rho(\mathrm{d}y)} \in L^1(M, \rho).$$

(H2) Let $Z := \left\{x \in M;\ \mathcal{P}^k(x, M) = 0 \text{ for some } k \in \mathbb{N}\right\}$. Then $0 < \rho(M \setminus Z)$ and for any $x \in M \setminus Z$ and open set $A \subset M \setminus Z$ in the induced topology of $M \setminus Z$, there exists a natural number $n = n(x, A)$ such that

$$\mathbb{P}_x\left[X_n \in A\right] = \mathcal{P}^n(x, A) > 0.$$

We now state the first main result of this paper, asserting that Hypothesis (H) implies the existence and uniqueness of a quasi-stationary measure for X_n on M.

Theorem A. *Let X_n be a Markov chain on $E_M = M \sqcup \{\partial\}$ absorbed at ∂ satisfying Hypothesis (H), then*

(a) if $\mathcal{P}(x, M) = 1$ for every $x \in M$, then X_n admits a unique stationary probability measure μ and $\mathrm{supp}(\mu) = M$;

(*b*) *if there exists* $x \in M \setminus Z$, *such that* $\mathcal{P}(x, M) < 1$, *then* $\lim_{n\to\infty} \mathcal{P}^n(y, M) = 0$, *for all* $y \in M$, *and the process* X_n *admits a unique quasi-stationary measure* μ *with*

$$\operatorname{supp}(\mu) \supset \bigcup_{k\in\mathbb{N}} \bigcup_{x\in M\setminus Z} \operatorname{supp}(\mathcal{P}^k(x, \mathrm{d}y)) \supset M \setminus Z \tag{3}$$

and survival rate $\lambda > 0$.

Theorem A is proved in Section 5.

Results, analogous to Theorem A, were previously obtained for families of diffeomorphisms [38,39]. Theorem A extend these, as we also cover families of endomorphisms.

The main techniques presented in this paper rely on the analysis of the spectral properties of the transition function $\mathcal{P}$, seen as a linear operator (see Section 3). We summarise the main properties of $\mathcal{P}$ in Theorem B.

Theorem B. *Let* X_n *be a Markov chain on* E_M *absorbed at* ∂ *satisfying Hypothesis (H) and* λ *be the survival rate given by Theorem A. Then,*

(*a*) *the stochastic Koopman operator*

$$\begin{aligned} \mathcal{P} : (C^0(M), \|\cdot\|_\infty) &\to (C^0(M), \|\cdot\|_\infty) \\ f &\mapsto \int_M f(y)\mathcal{P}(x, \mathrm{d}y) \end{aligned}$$

is a compact linear operator with spectral radius $r(\mathcal{P}) = \lambda$;

(*b*) *the set of eigenvalues of* $\mathcal{P}$ *with modulus* λ *is given by* $\{\lambda e^{\frac{2\pi i j}{m}}; j = \{0, \dots, m-1\}\}$ *for some* $m \le \#\{$*connected components of* $M \setminus Z\}$;

(*c*) *for every* $j \in \{0, \dots, m-1\}$, $\dim(\ker(\mathcal{P} - \lambda e^{\frac{2\pi i j}{m}})) = 1$; *and*

(*d*) *there exists a non-negative continuous function* f *such that* $\mathcal{P}f = \lambda f$ *and* $\{x \in M; f(x) > 0\} = M \setminus Z$.

Theorem B is proved in Section 6.2.

Remark 2.4. The inequality $m \le \#\{$connected components of $M \setminus Z\}$, in the above theorem, shows that the spectrum of $\mathcal{P}$ presents topological obstructions. Moreover, it is shown in Example 2.9 that it is possible for m to be strictly smaller than the number of connected components of $M \setminus Z$.

We conclude this section by stating the final main result of this paper, which concerns the existence and characterisation of quasi-ergodic measures for a Markov chain X_n satisfying Hypothesis (H). Recall the formal definition of a quasi-ergodic measure.

Definition 2.5. A measure η is called a *quasi-ergodic measure* on M if, for every $h \in \mathcal{F}_b(M)$,

$$\lim_{n\to\infty} \mathbb{E}_x\left[\frac{1}{n}\sum_{i=0}^{n-1} h\circ X_i \;\middle|\; \tau > n\right] = \int_M h(y)\eta(\mathrm{d}y) \tag{4}$$

holds for η-almost every $x \in M$.

Theorem C. *Let* X_n *be a Markov chain on* E_M *absorbed at* ∂ *satisfying Hypothesis (H). Let* μ *denote the unique quasi-stationary measure for* X_n *and* λ *its survival rate, as in Theorem A. Let* $f \in C^0_+(M)$ *be a non-negative continuous function such that* $\mathcal{P}f = \lambda f$ *and* m *be the number of eigenvalues of* $\mathcal{P}$ *in the circle of radius* λ, *as defined in Theorem B. Assume additionally that either* $m = 1$ *or* $\rho(Z) = 0$. *Then* X_n *admits a unique quasi-ergodic measure on* M *given by*

$$\eta(\mathrm{d}x) = \frac{f(x)\mu(\mathrm{d}x)}{\int_M f(y)\mu(\mathrm{d}y)},$$

and Eq. (4) holds for every $x \in M \setminus Z$. *Moreover,*

(*M1*) *if* $m = 1$, *given for every* $\nu \in \mathcal{M}_+(M)$, *such that* $\int f\,\mathrm{d}\nu > 0$, *there exist constants* $K(\nu)$, $\alpha > 0$, *such that*

$$\left\|\mathbb{P}_\nu\left[X_n \in \cdot \mid \tau > n\right] - \mu\right\|_{TV} \le K(\nu)e^{-\alpha n}, \text{ for all } n \in \mathbb{N}. \tag{5}$$

(*M2*) *if* $m > 1$ *and* $\rho(Z) = 0$, *there exist open sets (on the induced topology of* M*)* C_0, C_1, …, C_{m-1}, *such that* $M \setminus Z = C_0 \sqcup C_1 \sqcup \dots \sqcup C_{m-1}$, *satisfying* $\{\mathcal{P}(\cdot, C_i) \neq 0\} = C_{i-1}$, *for all* $i \in \{0, 1, \dots, m-1\}$, *with* $C_{-1} = C_{m-1}$.

Theorem C is proved in Section 6.2.

Observe that items (M1) and (M2) of Theorem C provide important information on the expected behaviour of X_n before being absorbed into ∂.

In particular, if $m = 1$, then for an $x \in M \setminus Z$, we have that

$$\frac{\mathbb{P}\left[X_n \in A \mid X_0 = x\right]}{\mathbb{P}\left[X_n \in M \mid X_0 = x\right]} = \frac{\mathcal{P}^n(x, A)}{\mathcal{P}^n(x, M)} \to \mu(A),$$

exponentially fast, as $n \to \infty$. This limit means that the expected long-time behaviour of the noise realisations that stay in M is described by the measure μ. On the other hand, in case (M2), the process X_n displays cyclic behaviour.

2.1. Applications

In this subsection, we discuss some concrete applications of Theorems A, B and C. The main purpose of this subsection is to illustrate that the above theorems can be applied to a wide class of Markov chains.

Example 2.6. Consider $T : \mathbb{R} \to \mathbb{R}$ and $\varepsilon > 0$, such that:

- T is an increasing homeomorphism;
- There exist unique x_- and x_+ satisfying $T(x_-) + \varepsilon = x_-$ and $T(x_+) - \varepsilon = x_+$; and
- $T(x) + \varepsilon \leq x$ for every $x < x_-$ and $x \leq T(x) - \varepsilon$ for every $x > x_+$.

Set $M = [x_-, x_+]$ and consider the Markov chain $X_{n+1} = T(X_n) + \omega_n$ where $\{\omega_i\}_{i\in\mathbb{N}_0}$ is an i.i.d. sequence of random variable uniformly distributed in $[-\varepsilon, \varepsilon]$. From this construction, it follows that X_n is absorbed at $\partial = \mathbb{R} \setminus M$.

Given $A \in \mathscr{B}(M)$,

$$\mathcal{P}(x, A) = \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} \mathbb{1}_A(T(x) + \omega)\mathrm{d}\omega = \frac{1}{2|x_+ - x_-|\varepsilon}\int_A \mathbb{1}_M(y - T(x))\mathrm{d}y.$$

It is readily verified that X_n fulfils Hypothesis (H) with $Z = \{x_-, x_+\}$, the only two points that immediately escape M almost surely. Since $M \setminus Z$ is connected. Theorem B guarantees that $m = 1$ and Theorems A and C provide the existence of a unique quasi-stationary and quasi-ergodic measure on M. Moreover, Theorem C also asserts that for every Borel measure ν on M, such that $\mathrm{supp}(\nu) \not\subset Z$,

$$\left\|\mathbb{P}_\nu\left[X_n \in \cdot \mid \tau > n\right] - \mu\right\|_{TV} \to 0 \text{ when } n \to \infty,$$

exponentially fast.

Natural choices of T in the above example are $T(x) = ax^{2n+1}$ for $n \in \mathbb{N}_0$, $a \in \mathbb{R}$ and $\varepsilon > 0$ large enough. We note that in the case that $|T(x) - x| < 1$ for every $x \in M$, the existence of quasi-stationary and quasi-ergodic measures also follow from [3, Proposition 7.4] and [36, Corollary 6.5]. However, our results are applicable to a broader class of functions. In contrast, we would like to point out that the existence of quasi-stationary measures in the case where M is not compact and $T(x) = ax$ was already investigated in [32].

In the next two examples, we explore natural extensions of Example 2.6. Unlike the previous scenario, which dealt with Markov chains generated by bounded additive perturbations of homeomorphisms in $\mathbb{R}$, Example 2.7 focuses on bounded geodesic-induced perturbations of the antipodal map in $\mathbb{S}^2$, and Example 2.8 studies absorbing Markov chains generated by perturbations of quadratic polynomials conditioned upon staying in a neighbourhood of its Julia set.

Example 2.7. Let $E = \mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$ and let $M = \{(x, y, z) \in \mathbb{S}^2; x \geq 0\}$ be a hemisphere of E. Fix $\varepsilon > 0$ and consider the Markov chain

$$X_{n+1} = \begin{cases} -Q(\varphi_n, \theta_n)X_n & \text{, if } X_n \in M \\ X_n & \text{, if } X_n \in E \setminus M, \end{cases}$$

where

$$Q(\varphi_n, \theta_n) := \begin{bmatrix} \cos(\varphi_n) & -\sin(\varphi_n) & 0 \\ \sin(\varphi_n) & \cos(\varphi_n) & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_n) & -\sin(\theta_n) \\ 0 & \sin(\theta_n) & \cos(\theta_n) \end{bmatrix},$$

and $\{(\varphi_n, \theta_n)\}_{i\in\mathbb{N}_0}$ is an i.i.d. sequence of random variable uniformly distributed in $[-\pi/2, \pi/2]\times[-\varepsilon, \varepsilon]$. By construction, this process effectively stops as soon as $X_n \in \partial = E \setminus M$.

It is clear from the above construction that $Z = \{(1, 0, 0)\}$. For every $A \in \mathscr{B}(M)$, we have

$$\mathcal{P}(x, B) = \frac{1}{2\pi\varepsilon}\int_{-\pi/2}^{\pi/2}\int_{-\varepsilon}^{\varepsilon} \mathbb{1}_A(-Q(\varphi, \theta)x)\mathrm{d}\varphi\mathrm{d}\theta.$$

By employing the coordinate transformation $v = -Q(\varphi, \theta)x$ and making use of the rank theorem [40, Theorem 4.1], it follows that X_n satisfies Hypothesis (H) when considering ρ as the volume form in $\mathbb{S}^2$. Furthermore, from the definition of X_n and the aforementioned observations, it can be shown that X_n satisfies Hypothesis (H).

Once again, since $M \setminus Z$ is connected, Theorem B ensures that $m = 1$. Moreover, Theorems A and C imply that on M, X_n admits a unique quasi-stationary measure which is fully supported, and also a unique quasi-ergodic measure. Furthermore, (5) holds.

Example 2.8. Given $c \in \mathbb{C}$, define $p_c : \mathbb{C} \to \mathbb{C}$ as $p_c(z) = z^2 + c$. Denote by $K_c := \{z : \|f^n(x)\| \not\to \infty \text{ as } n \to \infty\}$ and let $J_c := \partial K_c$ be the so-called *Julia set*. J_c is non-empty compact and totally invariant, i.e. $J_c = p_c(J_c) = p_c^{-1}(J_c)$ ([41, Lemma 3.1.]).

Let $c \in \mathbb{C}$ be such that J_c is connected, i.e. c lies in the Mandelbrot set. Let $\delta > 0$ and define $J_c^\delta = \{z \in J_c; \|z - c\| \leq \delta\}$. Given $\varepsilon > 0$, define the Markov chain $X_{n+1} = p_c(X_n) + \mathbb{1}_M(X_n)\omega_n$ on $E := \mathbb{C}$, where $\{\omega_i\}_{i\in\mathbb{N}_0}$ is an i.i.d. sequence of random variables uniformly distributed in $B_\varepsilon(0) := \{z \in \mathbb{C}; \|z\| \leq \varepsilon\}$, and M is the closure of the maximal subset of $\mathbb{C}$ such that:

- $J_c \subset M \subset J_c^{\delta}$; and
- X_n is transitive in the interior of M.

Observe that M is well-defined since $p_c : J_c \to J_c$ has a dense orbit [41, Corollary 3.11.], moreover $J_c^{\delta_1} \subset M$ for $\delta_1 > 0$ small enough.

Given $A \in \mathscr{B}(M)$ we obtain that

$$\mathcal{P}(x, A) = \frac{1}{\pi\varepsilon^2} \int_{B_\varepsilon(0)} \mathbb{1}_A(p_c(x) + \omega)\mathrm{d}\omega = \frac{1}{\pi\varepsilon^2} \int_A \mathbb{1}_{B_0(\varepsilon)}(z - p_c(x))\mathrm{d}z.$$

The above equation and the definition of M imply that X_n satisfies (H). Since $p_c : J_c \to J_c$ is topologically exact [41, Lemma 3.7.], we obtain that X_n does not display cyclic behaviour in M, so that Theorem C implies $m = 1$. Finally, it follows from Theorems A and C that, X_n admit unique quasi-stationary measure and quasi-ergodic measures on M. As before, (5) holds.

In the last example, we show that the constant m given by Theorem B can be strictly less than the number of connected components of $M \setminus Z$.

Example 2.9. Let $E = \mathbb{R}$ and M be a compact regular subset of E, i.e. $M = \overline{\mathrm{Int}(M)}$. Consider X_n be the Markov chain $X_{n+1} = X_n + \mathbb{1}_M(X_n)\omega_n$, where $\{\omega_i\}_{i\in\mathbb{N}_0}$ is an i.i.d. sequence of random variables normally distributed with mean 0 and variance 1. Let $\mathbb{P}$ denote the standard Gaussian distribution.

Then, for every $A \in \mathscr{B}(M)$,

$$\mathcal{P}(x, A) = \int_{\mathbb{R}} \mathbb{1}_A(x + \omega)\,\mathbb{P}(\mathrm{d}\omega) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{(y-x)^2}{2}}\mathrm{d}y$$

and Hypothesis (H) is fulfilled with $Z = \emptyset$. Since X_n does not display cyclic behaviour in M, Theorem C implies that $m = 1$. Therefore, by Theorems A and C we conclude that X_n admits a unique quasi-stationary stationary measure which is fully supported on M, and a unique quasi-ergodic measure.

Note that, in this case, $m = 1$ while although M may have an infinite number of connected components.

3. Some direct consequences of Hypothesis (H)

The purpose of this section is to present three fundamental results that are extensively used throughout this paper. The first two are direct consequences of Hypothesis (H), while the third one is functional-analytic.

We begin by stating some properties of the map $\mathcal{P}$, which follow from standard arguments that can be found in the literature [6,42].

Proposition 3.1. *If X_n fulfils Hypothesis (H), then the map $\mathcal{P} : L^\infty(M) \to L^\infty(M)$, $\mathcal{P}f = \mathbb{E}_x[f \circ X_1] = \int_M f(y)\mathcal{P}(x, \mathrm{d}y)$ has the following properties:*

(a) For all $n \in \mathbb{N}$ and $f \in L^\infty(M)$, $(\mathcal{P})^n(f) = \mathbb{E}_x[f \circ X_n]$;
(b) If $f \in L^\infty(M)$, then $\mathcal{P}f \in C^0(M)$;
(c) $\mathcal{P}|_{C^0(M)} : \left(C^0(M), \|\cdot\|_\infty\right) \to \left(C^0(M), \|\cdot\|_\infty\right)$ is a positive compact operator; and
(d) If $z \in Z$ then $\mathcal{P}(z, M \setminus Z) = 0$.

Proof. Item (a) follows from the Markov property of the process X_n. To prove (b), let $x \in M$ and $\varepsilon > 0$. By Hypothesis (H2) there exists $\delta > 0$ such that $d(x, z) < \delta$ then $\|g(x, \cdot) - g(z, \cdot)\|_1 < \varepsilon$. Therefore, given $z \in M$, such that $d(x, z) < \delta$, we obtain

$$|\mathcal{P}f(z) - \mathcal{P}f(x)| \le \|f\|_\infty \|g(x, \cdot) - g(z, \cdot)\|_1 < \varepsilon\|f\|_\infty \tag{6}$$

implying that $\mathcal{P}f \in C^0(M)$.

Item (c) follows from inequality (6) and the Arzelá–Ascoli Theorem [6, Proposition 5.3.].

Finally, we prove (d). Suppose by contradiction that there exists $z \in Z$, such that $\mathcal{P}(z, M \setminus Z) > 0$. Therefore, we can find a compact set $K \subset M \setminus Z$, satisfying $\mathcal{P}(z, K) > 0$. In this way, given $k \in \mathbb{N}$, we obtain that

$$\mathcal{P}^k(z, M) = \int_M \mathcal{P}^{k-1}(x, M)\mathcal{P}(z, \mathrm{d}y) \ge \mathcal{P}(z, K) \min_{x\in K} \mathcal{P}^{k-1}(x, M) > 0,$$

since, for every $k \in \mathbb{N}$, $\mathcal{P}^{k-1}(\cdot, M)$ is a continuous function and $\mathcal{P}^k(x, M) > 0$ for every $x \in K \subset M \setminus Z$. This contradicts $z \in Z$. □

Proposition 3.2. *Let the absorbing Markov chain X_n satisfy Hypothesis (H). Then the following assertions hold:*

(a) If there exists $x_0 \in M \setminus Z$, such that $\mathcal{P}(x_0, M) < 1$, then there exist $n_0 \in \mathbb{N}$ and $\alpha \in (0, 1)$ such that

$$\mathcal{P}^n(x, M) < \alpha^{\left\lfloor \frac{n}{n_0} \right\rfloor}, \quad \text{for all } x \in M.$$

In particular, $\lim_{n\to\infty} \mathbb{P}_x[\tau > n] = \lim_{n\to\infty} \mathcal{P}^n(x, M) = 0$, for all $x \in M$, and $r(\mathcal{P}) < 1$.

(b) Let $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{B}(M)$ *such that* $\lim_{n\to\infty} \mathbb{1}_{A_n} = 0$, *ρ-almost surely. Then,* $\lim_{i\to\infty} \|\mathcal{P}(x, A_i)\|_\infty = 0$.

Proof. Since $\mathcal{P}(\cdot, M)$ is continuous, there exists an open neighbourhood B of x_0, such that $\sup_{y\in B} \mathcal{P}(y, M) < 1$.

To prove (*a*), note that given $x \in M \setminus Z$, by Hypothesis (H2), there exists $n_x = n(x, B)$, such that $\mathcal{P}^{n_x}(x, B) > 0$. Therefore

$$\begin{aligned}\mathcal{P}^{n_x+1}(x, M) &= \int_B \mathcal{P}(y, M)\mathcal{P}^{n_x}(x, \mathrm{d}y) + \int_{M\setminus B} \mathcal{P}(y, M)\mathcal{P}^{n_x}(x, \mathrm{d}y)\\ &\le \mathcal{P}^{n_x}(x, B) \sup_{y\in B} \mathcal{P}(y, M) + \mathcal{P}^{n_x}(x, M \setminus B) < 1,\end{aligned}$$

which implies that $\mathcal{P}^{n_x+1}(x, M) < 1$. Hence, given any $x \in M$, there exists $m_x = n_x + 1$, such that $\mathcal{P}^{m_x}(x, M) < 1$. Since $\mathcal{P}^{m_x}(\cdot, M)$ is continuous, there exists an open neighbourhood U_x of x such that $\mathcal{P}^{m_x}(y, M) < 1$, for all $y \in U_x$.

Thus $M = \bigcup_{x\in M} U_x$, and since M is compact, there exist $x_1, x_2, \dots, x_n \in M$ such that $M = U_{x_1} \cup U_{x_2} \cup \cdots \cup U_{x_n}$. It is readily verified that taking $n_0 = m_{x_1} \cdot m_{x_2} \cdot \ldots \cdot m_{x_n}$ we have that $\mathcal{P}^{n_0}(y, M) < 1$. Hence, defining $\alpha = \sup_{x\in M} \mathcal{P}^{n_0}(x, M) < 1$, it follows that

$$\mathcal{P}^{nn_0}(x, M) = \int_M \mathcal{P}^{n_0}(y, M)\mathcal{P}^{n_0(n-1)}(x, \mathrm{d}y) \le \alpha\mathcal{P}^{(n-1)n_0}(x, M) \le \cdots \le \alpha^n.$$

It remains to prove (*b*). From the Arzelá–Ascoli, $\{\mathcal{P}(\cdot, A_i)\}_{i\in\mathbb{N}}$ is pre-compact in $(C^0(M), \|\cdot\|_\infty)$. (H1) implies that for a fixed $x \in M$, $\lim_{n\to\infty} \mathcal{P}(x, A_i) = 0$. Therefore, if f is an accumulation point of $\{\mathcal{P}(\cdot, A_i)\}_{i\in\mathbb{N}}$ in $(C^0(M), \|\cdot\|_\infty)$, $f(x) = 0$ for every $x \in M$. We can therefore conclude that $\mathcal{P}(\cdot, A_i)$ converges to 0 in $C^0(M)$. □

The following proposition states a functional-analytic result that is extensively used throughout this paper.

Proposition 3.3. *Let* E *be a Banach space,* λ *a positive real number, and* $T : E \to E$ *a bounded linear operator such that* $r(T) < \lambda$. *Then there exist constants* $K > 0$ *and* $\delta \in (0, \lambda)$ *such that* $\|T^n\| \le K(\lambda - \delta)^n$, *for all* $n \in \mathbb{N}$.

We do not include a proof of this result as it follows from standard arguments in the literature and the formula of the spectral radius $r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$ for bounded linear operators $T : E \to E$.

4. Banach lattices

In this section, we introduce Banach lattices, which are essential for the proof of the main theorems of this paper. We show that the operator $\mathcal{P}$ is well-behaved from a Banach lattice point of view, allowing us to deduce important properties of the spectrum of the operator $\mathcal{P}$. We start this section with some basic definitions from Banach lattice theory and state some key results from this field, which we then apply to the operator $\mathcal{P}$.

Given $(L, \le)$ a partial ordered set and a set $B \subset L$, we define, if they exists

$$\sup B = \min\{\ell \in L;\ b \le \ell,\ \text{for all } b \in B\} \text{ and } \inf B = \max\{\ell \in L;\ \ell \le b,\ \text{for all } b \in B\}.$$

We say that L is a *lattice*, if for every $f_1, f_2 \in L$,

$$f_1 \vee f_2 := \sup\{f_1, f_2\},\ f_1 \wedge f_2 := \inf\{f_1, f_2\}$$

exist. Additionally, in the case that L is a vector space and the lattice $(L, \le)$ satisfies

$$f_1 \le f_2 \Rightarrow f_1 + f_3 \le f_2 + f_3,\ \text{ for all } f_3 \in L,\ \text{and}$$

$$f_1 \le f_2 \Rightarrow \alpha f_1 \le \alpha f_2,\ \text{ for all } \alpha > 0,$$

then $(L, \le)$ is called *vector lattice*. Finally, if $(L, \|\cdot\|)$ is a Banach space and the vector lattice $(L, \le)$ satisfies

$$|f_1| \le |f_2| \Rightarrow \|f_1\| \le \|f_2\|,$$

where $|f_1| := f_1 \vee (-f_1)$, then the triple $(L, \le, \|\cdot\|)$ is called a *Banach lattice*. When the context is clear, we denote the Banach lattice $(L, \le, \|\cdot\|)$ by L.

This paper uses two concepts from the Banach lattice theory: an *ideal* of a Banach lattice, and an *irreducible operator* acting on a Banach lattice. A vector subspace $I \subset L$, is called an *ideal* if, for every $f_1, f_2 \in L$ such that $f_2 \in I$ and $|f_1| \le |f_2|$, we have $f_1 \in I$. A positive linear operator $T : L \to L$ is called *irreducible* if, $\{0\}$ and L are the only T-invariant closed ideals of T.

The following two results give us tools to analyse the spectrum of a compact positive irreducible operator. In Section 4.1, we show that these results apply to the operator $\mathcal{P}$ when restricted to a specific subspace of $C^0(M)$. This procedure allows us to analyse the spectrum of the operator $\mathcal{P}$.

Proposition 4.1 (*[9, Proposition 2.1.9 (iii)]*)**.** *Let* M *be a compact Hausdorff space. Consider the Banach lattice* $C^0(M)$. *Then,* I *is an ideal of* $C^0(M)$, *if and only if* $I = \{f \in C^0(M);\ f|_A = 0\}$, *for some closed set* A.

Definition 4.2. Let L be a Banach lattice. We denote by $L_+ := \{f \in L;\ 0 \leq f\}$, and write

$$L_+^* := \{\varphi \in L^*;\ \varphi \text{ is a continuous positive linear operator}\},$$

where L^* is the topological dual space of L.

A point $f \in L$ is called *quasi-interior* if $f \in L_+$ and, for every $\varphi \in L_+^* \setminus \{0\}$ we have $\varphi(f) > 0$.

The main results from the theory of Banach lattices used in the upcoming sections are summarised in the following theorem.

Theorem 4.3 (*[8, Proposition 5.2 and Theorem 5.3]*)**.** *Let L be a Banach lattice and let T be an irreducible operator. If T^k is compact for some $k \in \mathbb{N}$, then $r(T) > 0$, and if $\lambda_1, \lambda_2, \dots, \lambda_m$ are the different eigenvalues of T satisfying $|\lambda_j| = r = r(T)$, for $j = 1, \dots, m$, every λ_j is a root of the equation $\lambda^m - r^m = 0$. All these eigenvalues have algebraic multiplicity one and the spectrum of T is invariant under a rotation of the complex plane under the angle $2\pi/m$, multiplicities included. Moreover, the eigenspace $\ker(T - r(T))$ is spanned by a quasi-interior point $u \in L_+$.*

4.1. Banach lattice properties of $\mathcal{P}$

In this section we exploit the Banach lattice of $(C^0(M), \leq, \|\cdot\|_\infty)$ as the stochastic Koopman operator $\mathcal{P}$ acts from $C^0(M)$ to itself, i.e. $\mathcal{P} : C^0(M) \to C^0(M)$. It is shown in Sections 5 and 6 that the existence and uniqueness of quasi-stationary and quasi-ergodic measures are closely related to the spectrum of $\mathcal{P}$, $\mathrm{Spec}(\mathcal{P})$. To our advantage, the tools described above allow us to give a detailed characterisation of $\mathrm{Spec}(\mathcal{P})$, which we present in the following theorem.

Theorem 4.4. *Let X_n be a Markov chain on E_M absorbed at ∂ satisfying Hypothesis (H) and consider the operator $\mathcal{P} : C^0(M) \to C^0(M)$. Then, defining $\lambda := r(\mathcal{P}) > 0$, there exists $m \in \mathbb{N}$ such that $\left\{\lambda e^{\frac{2\pi i j}{m}}\right\}_{j=0}^{m-1}$ are the unique eigenvalues of modulus equal to λ, and each one has algebraic multiplicity one.*

Moreover, $\dim\left(\ker\left(\mathcal{P} - \lambda e^{\frac{2\pi i j}{m}}\right)\right) = 1$, *for all $j \in \{0, 1, \dots, m-1\}$, and there exists $f \in C_+^0(M)$, such that $f(x) = 0$ if and only if $x \in Z$.*

Proof. From Proposition 3.1 (d) and Hypothesis (H), we obtain that $Z = \{x \in M, \mathcal{P}(x, M \setminus Z) = 0\}$. Since $x \mapsto \mathcal{P}(x, M \setminus Z)$ is continuous, then the set Z is compact. Moreover, defining the set $C_Z := \{f \in C^0(M);\ f(z) = 0, \text{ for all } z \in Z\}$, it is clear that C_Z is a closed invariant subspace of $C^0(M)$ and therefore C_Z admits a Banach lattice structure induced by $C^0(M)$. From Hypothesis (H) and Propositions 3.1 and 4.1 it is readily verified that $\mathcal{P}|_{C_Z} : C_Z \to C_Z$ is an irreducible compact positive operator. Therefore, by application of Theorem 4.3 to the operator $\mathcal{P}|_{C_Z}$, we have that $\mathcal{P}|_{C_Z}$ fulfils all the conditions stated in Theorem 4.4. The proof is finished noticing that $\mathrm{Spec}(\mathcal{P}) = \mathrm{Spec}\left(\mathcal{P}|_{C_Z}\right)$, since f if is an eigenvector of $\mathcal{P}$ then $f(z) = 0$ for every $z \in Z$. □

5. Existence and uniqueness of a quasi-stationary measures

In this section, we show that Hypothesis (H) is sufficient to obtain existence and uniqueness of a quasi-stationary measure for X_n on M.

Recall that $\mathcal{M}(M) := \{\mu;\ \mu \text{ is a Borel signed measure on } M\}$ has a Banach space structure when endowed with the total variation norm. It is well known, from the Riesz–Markov–Kakutani representation theorem [43, Theorem 6.19], that $(C^0(M), \|\cdot\|_\infty)^* = \left(\mathcal{M}(M), \|\cdot\|_{TV}\right)$. Thus we may identify any $\mu \in \mathcal{M}(M)$, with an element of $(C^0(M), \|\cdot\|_\infty)^*$, by $\mu(f) := \int f(x)\mu(\mathrm{d}x)$, for every $f \in C^0(M)$.

In order to prove the existence and uniqueness of a quasi-stationary measure for X_n we study the spectrum of the operator $\mathcal{L} : \mathcal{M}(M) \to \mathcal{M}(M)$, $\mathcal{L}\mu := \int_M \mathcal{P}(x, \cdot)\mu(\mathrm{d}x)$. However, it is first necessary to linearly extend the action of $\mathcal{P}$ and $\mathcal{L}$ to their complex domains

$$C^0(M, \mathbb{C}) := C^0(M) \oplus iC^0(M) \text{ and } \mathcal{M}(M, \mathbb{C}) := \mathcal{M}(M) \oplus i\mathcal{M}(M).$$

Definition 5.1 (*[44]*)**.** Let E be a Banach space. Then, the *scalar product for the duality* E^*, E is the bilinear form $\langle\cdot,\cdot\rangle : E^* \times E \to \mathbb{C}$, defined by $\langle\varphi, v\rangle := \varphi(v)$.

Given a bounded linear operator $T : E \to E$, we denote $T^* : E^* \to E^*$ as the linear operator $T^*\varphi(v) = \varphi(Tv)$, for all $v \in E$.

Using this definition, it follows that $\mathcal{P}^* = \mathcal{L}$. The following lemma shows a connection between the spectrum of the operators $\mathcal{P}$ and $\mathcal{L}$. The proof follows from standard functional analytical techniques.

Lemma 5.2. *The operators $\mathcal{P}$ and $\mathcal{L}$ have the same eigenvalues and*

$$\dim(\ker(\mathcal{L} - \beta)) = \dim(\ker(\mathcal{P} - \beta)), \quad \text{for all } \beta \in \mathrm{Spec}(\mathcal{P}) = \mathrm{Spec}(\mathcal{L}).$$

Moreover, if $f_0 \in C_+^0(M)$ is an eigenfunction of $\mathcal{P}$ with eigenvalue $\lambda = r(\mathcal{P}) = r(\mathcal{L})$, then there exists an eigenmeasure $\mu_{f_0} \in \mathcal{M}_+(M)$ of $\mathcal{L}$ with respect to the eigenvalue λ such that $\mu_{f_0}(f_0) = 1$.

We can finally state and prove the main result of this section.

Theorem 5.3. *If a Markov chain X_n satisfies Hypothesis (H), then X_n admits a unique quasi-stationary measure μ on M, and* $\mathrm{supp}(\mu)$ *satisfies (3).*

Proof. We divide the proof into three steps.

Step 1. *We show that if $\mu \in \mathcal{M}_+(M)$ is an eigenmeasure of $\mathcal{L}$, then*

$$\mathrm{supp}\,(\mu) \supset \bigcup_{k\in\mathbb{N}} \bigcup_{x\in M\setminus Z} \mathrm{supp}(\mathcal{P}^k(x,\mathrm{d}y)) \supset M\setminus Z.$$

First, observe that there exists $x_0 \in \mathrm{supp}(\mu) \cap M \setminus Z$. In fact, if $\mu(Z) = 1$, define $Z_k := \{x \in Z; \mathcal{P}^k(x, M) = 0\}$. It is clear that $Z = \bigcup_{k\in\mathbb{N}} Z_k$. Therefore for every $k \in \mathbb{N}$

$$\begin{aligned} 1 = \mu(Z) &= \frac{1}{\lambda^k}\int_Z \mathcal{P}^k(z,Z)\mu(\mathrm{d}z) = \int_{Z\setminus Z^k} \frac{1}{\lambda^k}\mathcal{P}^k(z,Z)\mu(\mathrm{d}z) \\ &\le \sup_{z\in Z}\left|\frac{1}{\lambda^k}\mathcal{P}^k(z,Z)\right| \mu(Z\setminus Z_k) \xrightarrow{k\to\infty} 0. \end{aligned}$$

The last inequality holds since $r(\mathcal{P}) = \lambda$ and $\mu(Z_k) \to \mu(Z)$ as $k \to \infty$.

Suppose by contradiction that there exists an open set $A \subset M \setminus Z$, such that $\mu(A) = 0$. Therefore, for every open neighbourhood U of x, $\mu(U) > 0$.

Using Hypothesis (H2), there exists $n > 0$ such that $\mathcal{P}^n(x_0, A) > 0$. Since $\mathcal{P}^n(\cdot, A)$ is continuous there exists an open neighbourhood B of x such that $\mathcal{P}^n(y, A) > \mathcal{P}^n(x_0, A)/2 > 0$, for all $y \in B$.

On the other hand,

$$0 = \mu(A) = \frac{1}{\lambda^n}\int_M \mathcal{P}^n(y,A)\mu(\mathrm{d}y) \ge \frac{1}{\lambda^n}\frac{\mathcal{P}^n(x_0,A)}{2}\mu(B) > 0,$$

which is a contradiction since $x_0 \in \mathrm{supp}(\mu)$ and $x_0 \in B$. Thus $\mathrm{supp}(\mu) = M\setminus Z$. The proof is finished by observing that

$$\mu(\mathrm{d}y) \ge \frac{1}{\lambda^k}\int_{M\setminus Z} \mathcal{P}^k(x,\mathrm{d}y)\mu(\mathrm{d}x) \text{ for every } k \in \mathbb{N}.$$

Step 2. *We show that the operator $\mathcal{L}$ admits a unique eigenmeasure that lies in $\mathcal{M}_+(M)$.*

Define $\lambda := r(\mathcal{P}) = r(\mathcal{L})$. Combining Theorem 4.4 *and Lemma* 5.2, we have that

$$\dim(\ker(\mathcal{L} - \lambda)) = \dim(\ker(\mathcal{P} - \lambda)) = 1 \tag{7}$$

and there exists a probability measure $\mu \in \mathcal{M}_+(M)$, such that $\mathcal{L}\mu = \lambda\mu$.

Suppose by contradiction that there exists a Borel probability measure ν on M linearly independent from μ such that $\mathcal{L}\nu = \lambda_0\nu$. From Theorem 4.3, the linear operator $\mathcal{P}$ admits a eigenfunction $f \in C^0_+(M)$, such that $f(x) > 0$ for every $x \in M \setminus Z$ with eigenvalue $\lambda = r(\mathcal{P}) = r(\mathcal{L})$. Therefore, $0 < \nu(f) \le \|f\|_\infty \nu(M) < \infty$. On the other hand,

$$\int_M f(x)\nu(\mathrm{d}x) = \frac{1}{\lambda_0}\langle \mathcal{L}\nu, f\rangle = \frac{1}{\lambda_0}\langle \nu, \mathcal{P}f\rangle = \frac{\lambda}{\lambda_0}\int_M f(x)\nu(\mathrm{d}x).$$

Therefore $\lambda_0 = \lambda$, contradiction (7). This concludes Step 2.

Step 3. *We show that the Markov chain X_n admits a unique quasi-stationary measure μ, and* $\mathrm{supp}(\mu)$ *satisfies (3).*

Let μ be the unique probability eigenmeasure of $\mathcal{L}$, given by Step 2. We claim that μ is a quasi-stationary measure. Note that, for every $A \in \mathscr{B}(M)$ and $n \in \mathbb{N}$,

$$\mathbb{P}_\mu\left[X_n \in A \mid \tau > n\right] = \frac{\int_M \mathcal{P}^n(x,A)\mu(\mathrm{d}x)}{\int_M \mathcal{P}^n(x,M)\mu(\mathrm{d}x)} = \frac{\mathcal{L}^n\mu(A)}{\mathcal{L}^n\mu(M)} = \frac{\lambda^n\mu(A)}{\lambda^n\mu(M)} = \mu(A),$$

showing that μ is a quasi-stationary measure for X_n.

Reciprocally, if ν is a quasi-stationary measure for X, then $\nu \in \mathcal{M}_+(M)$. Therefore, defining $\lambda_0 = \int_M \mathcal{P}(x,M)\nu(\mathrm{d}x)$, we obtain $\int_M \mathcal{P}(x,\cdot)\nu(\mathrm{d}x) = \lambda_0\nu(\cdot)$.

Hence, there is a 1-1 correspondence between probability eigenmeasures of $\mathcal{L}$ lying in $\mathcal{M}_+(M)$ and quasi-stationary measures of X_n.

This concludes the proof of the theorem. □

We close this section with the proof of Theorem A.

Proof of Theorem A. Note that from Theorem 5.3, X_n admits a unique quasi-stationary measure μ and (3) holds. We first prove part (a). Since $\mathcal{P}(x, M) = 1$ for all $x \in M$, the constant function $x \mapsto 1$ is an eigenfunction of $\mathcal{P}$. Therefore, $r(\mathcal{P}) = r(\mathcal{L}) = 1$ and μ corresponds to a stationary measure. In case (b), since there exists $x_0 \in M \setminus Z$ such that $\mathcal{P}(x_0, M) < 1$, Proposition 3.2 (a) guarantees that $\lim_{n\to\infty} \mathcal{P}^n(y, M) = 0$, for all $y \in M$, implying that μ is the unique quasi-stationary measure for X_n with survival rate $\lambda < 1$. □

6. Existence of a quasi-ergodic measure

The proof of the existence of a quasi-ergodic measure is much more intricate than that of a quasi-stationary measure. Our technique is inspired by the seminal Ref. [22], which focuses on time-continuous absorbing Harris' chains, and [5], where time-discrete Markov chains were considered in the context of moving boundaries. In this section, we extend the results presented in [22] to include discrete-time Markov processes, which may not necessarily be aperiodic.

As in [5], our method depends on the number of eigenvalues of $\mathcal{P}$ in the circle $r(\mathcal{P})\mathbb{S}^1 \subset \mathbb{C}$. For this reason, we define the following quantity.

Notation 6.1. *Given a Markov chain X_n satisfying Hypothesis (H), from Theorem 4.4 it follows that the number of eigenvalues in the peripheral spectrum of $\mathcal{P}$, i.e. $\mathrm{Spec}(\mathcal{P}) \cap r(\mathcal{P})\mathbb{S}^1$, is finite. For the remaining of this section we denote $\lambda := r(\mathcal{P}) = r(\mathcal{L})$ and $m := \#\{\alpha \in \mathrm{Spec}(\mathcal{P});\ \|\alpha\| = \lambda\}$.*

Proposition 6.2. *Let X_n be a Markov chain on E_M absorbed at ∂ satisfying Hypothesis (H), and $\lambda = r(\mathcal{P}) = r(\mathcal{L})$. Consider f_0, f_1, …, $f_{m-1} \in C^0(M, \mathbb{C})$ and μ_0, μ_1, …, $\mu_{m-1} \in \mathcal{M}(M, \mathbb{C})$ such that $\mathcal{P} f_j = \lambda e^{\frac{2i\pi j}{m}} f_j$ and $\mathcal{L}\mu_j = \lambda e^{\frac{2i\pi j}{m}} \mu_j$, for all $j \in \{0, 1, \dots, m-1\}$. Then $\langle \mu_j, f_r\rangle = 0$, if $j \neq r$. In particular, normalise f_i with respect to μ_i such that $\langle \mu_j, f_k\rangle = \delta_{jk}$, where δ is the Kronecker delta.*

Proof. Note that

$$\langle \mu_j, f_r\rangle = \frac{1}{\lambda e^{\frac{2\pi i j}{m}}}\langle \mathcal{L}\mu_j, f_r\rangle = \frac{1}{\lambda e^{\frac{2\pi i j}{m}}}\langle \mu_j, \mathcal{P} f_r\rangle = \frac{\lambda e^{\frac{2\pi i r}{m}}}{\lambda e^{\frac{2\pi i j}{m}}}\langle \mu_j, f_r\rangle = e^{\frac{2\pi i (r-j)}{m}}\langle \mu_j, f_r\rangle.$$

Given $j, r \in \{0, 1, \dots, m-1\}$, if $r \neq j$ we have that $e^{\frac{2\pi i (r-j)}{m}} \neq 1$, and therefore $\langle \mu_j, f_r\rangle = 0$. In the following we argue that $\langle f_i, \mu_i\rangle > 0$ for every $i \in \{0, 1, \dots, m-1\}$.

From [45, Theorem 8.4-5], we can decompose $C^0(M, \mathbb{C}) = W_0 \oplus W$, where $W_0 := \mathrm{span}_{\mathbb{C}}(f_0, f_1, \dots, f_{m-1})$ and W is a $\mathcal{P}$-invariant subspace of $C^0(M, \mathbb{C})$ satisfying $r\left(\mathcal{P}|_W\right) < \lambda$. Observe that the above decomposition implies that $\langle \mu_j, f_j\rangle \neq 0$, some $j \in \{0, 1, \dots, m-1\}$, otherwise we would have that $\mu_j = 0$. Redefining f_j as $f_j / \langle \mu_j, f_j\rangle$, the proof is concluded. □

Until the end of this section, we set $\{f_j\}_{j=0}^{m-1} \subset C^0(M, \mathbb{C})$ and $\{\mu_j\}_{j=0}^{m-1} \subset \mathcal{M}(M, \mathbb{C})$ as in Proposition 6.2 assuming that $\langle \mu_j, f_i\rangle = \delta_{ij}$ for every $i, j \in \{0, \dots, m-1\}$.

Furthermore, from Theorem 4.4 and [45, Theorem 8.4-5], we have

$$C^0(M, \mathbb{C}) = \mathrm{span}_{\mathbb{C}}(f_0, f_1, \dots, f_{m-1}) \oplus W, \tag{8}$$

where W is $\mathcal{P}$-invariant subspace of $\mathbb{C}$ and $r\left(\mathcal{P}|_W\right) < \lambda$, and

$$\mathcal{M}(M, \mathbb{C}) = \mathrm{span}_{\mathbb{C}}(\mu_0, \mu_1, \dots, \mu_{m-1}) \oplus V, \tag{9}$$

where V is $\mathcal{L}$-invariant subspace of $\mathcal{M}(M, \mathbb{C})$ and $r\left(\mathcal{L}|_V\right) < \lambda$.

Note that the decompositions (8) and (9) imply that $\|\mathcal{P}^n\| = \mathcal{O}(\lambda^n)$ and $\|\mathcal{L}^n\| = \mathcal{O}(\lambda^n)$. Indeed, writing $\mathbb{1}_M = \alpha_0 f_0 + \cdots + \alpha_{m-1} f_{m-1} + w$, where $\alpha_0, \dots, \alpha_{m-1} \in \mathbb{C}$ and $w \in W$. Since $\mathcal{P}$ is a positive operator,

$$\|\mathcal{P}^n\| = \left\| \sum_{i=0}^{m-1} \alpha_i \mathcal{P}^n f_i + \mathcal{P}^n w \right\|_\infty \leq \lambda^n \sum_{i=0}^{m-1} |\alpha_i| \|f_i\|_\infty + o(\lambda^n) = \mathcal{O}(\lambda^n). \tag{10}$$

Finally, since $\|\mathcal{L}^n\| = \|\mathcal{P}^n\|$, for all $n \in \mathbb{N}$ we find that $\|\mathcal{L}^n\| = \mathcal{O}(\lambda^n)$. In the following lemma, we discuss the behaviour of the Dirac measures δ_x under the decomposition (9).

Proposition 6.3 (*Decomposition of Dirac measures*). *Let $x \in M$, then there exists $\nu_x \in V$ such that*

$$\delta_x = f_0(x)\mu_0 + f_1(x)\mu_1 + \cdots + f_{m-1}(x)\mu_{m-1} + \nu_x. \tag{11}$$

Moreover, the family of function $\{\nu_x\}_{x\in M}$ satisfies

$$\sup_{x\in M} \|\nu_x\|_{TV} \leq 1 + \sum_{i=0}^{m-1} \|f_i\|_\infty \|\mu_i\|_{TV} < \infty, \text{ and } \sup_{x\in M} \|\mathcal{L}^n \nu_x\|_{TV} = o(\lambda^n). \tag{12}$$

Proof. By decomposition (9), there exist $\alpha_0, \ldots, \alpha_{m-1} \in \mathbb{C}$, and $\nu_x \in V$ such that $\delta_x = \sum_{i=0}^{m-1} \alpha_i \mu_i + \nu_x$. Since $\frac{1}{\lambda^n}\mathcal{L}^n\delta_x = \sum_{k=0}^{m-1} \alpha_k \mu_k e^{\frac{2\pi i k n}{m}} + \frac{1}{\lambda^n}\mathcal{L}^n\nu_x$, and $\langle f_i, \mu_j\rangle = \delta_{ij}$ for every $j \in \{0, \ldots, m-1\}$, we obtain that

$$\left\langle \frac{1}{\lambda^n}\mathcal{L}^n\delta_x, f_j \right\rangle = \alpha_j e^{\frac{2\pi i j n}{m}} + \left\langle \frac{1}{\lambda^n}\mathcal{L}^n\nu_x, f_j \right\rangle. \tag{13}$$

On the other hand,

$$\left\langle \frac{1}{\lambda^n}\mathcal{L}^n\delta_x, f_j \right\rangle = \frac{1}{\lambda^n}\left\langle \delta_x, \mathcal{P}^n(f_j)\right\rangle = \frac{\lambda^n e^{\frac{2\pi i j n}{m}}}{\lambda^n}\left\langle \delta_x, f_j\right\rangle = e^{\frac{2\pi i j n}{m}} f_j(x). \tag{14}$$

From (13)*and* (14) it follows that

$$f_j(x) = \left\langle \frac{1}{\lambda^{nm}}\mathcal{L}^{nm}\delta_x, f_j \right\rangle = \alpha_j + \left\langle \frac{1}{\lambda^{nm}}\mathcal{L}^{nm}\nu_x, f_j \right\rangle, \quad \text{for all } n \in \mathbb{N}.$$

From Proposition 3.3 we obtain that $f_j(x) = \alpha_j + \lim_{n\to\infty}\left\langle \frac{1}{\lambda^{nm}}\mathcal{L}^{nm}\nu_x, f_j\right\rangle = \alpha_j$.

The last part of the proposition follows from the fact that

$$\begin{aligned}\sup_{x\in M}\|\nu_x\|_{TV} &= \sup_{x\in M}\left\|\delta_x - \left(f_0(x)\mu_0 + \cdots + f_{m-1}(x)\mu_j\right)\right\|_{TV} \\ &\le \sup_{x\in M}\|\delta_x\|_{TV} + \sum_{j=0}^{m-1}\|f_j\|_\infty\|\mu_j\|_{TV} = 1 + \sum_{j=0}^{m-1}\|f_j\|_\infty\|\mu_j\|_{TV} < \infty,\end{aligned}$$

and $\frac{1}{\lambda^n}\sup_{x\in M}\left\|\mathcal{L}^n\nu_x\right\|_{TV} \le \frac{1}{\lambda^n}\left\|\mathcal{L}^n|_V\right\|_{TV}\sup_{y\in M}\|\nu_y\| \longrightarrow 0$, when $n \to \infty$, due to Proposition 3.3. □

The following lemma lies at the basis of our proof for the existence of a quasi-ergodic measure. Since the proof of this result is long and technical, we include it in Appendix.

Lemma 6.4. *Let $x \in M \setminus Z$, and $h : M \to \mathbb{R}$ a bounded measurable function. Then, for every $n \in \mathbb{N}$.*

$$\mathbb{E}_x\left[\sum_{k=0}^{n-1} h(X_k)\mathbb{1}_M(X_n)\right] = n\lambda^n\sum_{\ell=0}^{m-1} e^{\frac{2\pi i n\ell}{m}} f_\ell(x)\langle\mu_\ell, h\cdot f_\ell\rangle\mu_\ell(M) \ + o(n\lambda^n).$$

In the following two subsections, we analyse the cases when the number of eigenvalues of the operator $\mathcal{P}$ in $\lambda\mathbb{S}^1$, m, is either one or greater than one. The first case is much simpler compared to the second one. Indeed, if $m = 1$, the operator $\mathcal{P}$ has the spectral gap property, simplifying the situation. Meanwhile, in the case $m > 1$, the process X_n admits a cyclic property, which requires a more sophisticated analysis.

6.1. Analysis of the case $m = 1$

In this section, we use Lemma 6.4 in order to prove item (M1) of Theorem C.

Theorem 6.5. *Let X_n be a Markov chain satisfying Hypothesis (H), and assume that $m = 1$. Let μ be the unique quasi-stationary measure on M, given by Theorem A, and f be the unique function $C^0_+(M)$ satisfying $\mathcal{P}f = \lambda f$ and $\int_M f(x)\mu(\mathrm{d}x) = 1$. Then, the measure $\eta(\mathrm{d}x) = f(x)\mu(\mathrm{d}x)$ is a quasi-ergodic measure for X_n on the set M and (4) holds for every $x \in M \setminus Z$.*

Moreover, given $\nu \in \mathcal{M}_+(M)$, such that $\int_M f(y)\nu(\mathrm{d}y) > 0$, there exist $K(\nu) > 0$ and $\alpha > 0$, such that $\left\|\mathbb{P}_\nu\left[X_n \in \cdot \mid \tau > n\right] - \mu\right\|_{TV} < K(\nu)e^{-\alpha n}$, for all $n \in \mathbb{N}$.

Proof. Let $x \in M$ and $h \in \mathcal{F}_b(M)$. Recall from Definition 2.5 that we need to show that

$$\lim_{n\to\infty}\mathbb{E}_x\left[\frac{1}{n}\sum_{i=0}^{n-1} h\circ X_i \;\middle|\; \tau > n\right] = \int_M h(y)\eta(\mathrm{d}y),$$

for η-almost every $x \in M$. Since $\eta(Z) = 0$, let us consider $x \in M \setminus Z$.

From Lemma 6.4 we obtain that

$$\mathbb{E}_x\left[\sum_{k=0}^{n-1} h(X_k)\mathbb{1}_M(X_n)\right] = n\lambda^n f(x)\int_M h(y)f(y)\mu(\mathrm{d}y) + o(n\lambda^n). \tag{15}$$

On the other hand, from Proposition 6.3, there exists $\nu_x \in V$ such that $\delta_x = f(x)\mu + \nu_x$ and $r\left(\mathcal{L}|_V\right) < \lambda$. Since $\mathcal{L}^n(\nu_x)(M) = o(\lambda^n)$, by Proposition 3.3, we obtain

$$\mathcal{P}^n(x, M) = \langle\mathcal{L}^n\delta_x, \mathbb{1}_M\rangle = \lambda^n f(x)\mu(M) + \mathcal{L}^n\nu_x(M) = \lambda^n f(x) + o(\lambda^n). \tag{16}$$

Hence, from (15) and (16)

$$\mathbb{E}_x\left[\frac{1}{n}\sum_{i=0}^{n-1} h\circ X_i \;\middle|\; \tau > n\right] = \frac{1}{n}\frac{\sum_{i=0}^{n-1}\mathbb{E}_x\left[h(X_i)\mathbb{1}_m\right]}{\mathcal{P}^n(x, M)} = \frac{1}{n}\frac{n\lambda^n f(x)\mu(hf) + o(n\lambda^n)}{\lambda^n f(x) + o(\lambda^n)}$$

$$\xrightarrow{n\to\infty} \mu(hf)\mu(\mathrm{d}y) = \int_M h(y)\eta(\mathrm{d}y).$$

Now, we prove the second part of the theorem. Let $\nu \in \mathcal{M}_+(M)$ be such that $\nu(f) = \int_M f(y)\nu(\mathrm{d}y) > 0$. Given any measurable set A, by Proposition 6.3

$$\mathcal{L}^n(\nu)(A) = \int_M \mathcal{L}^n\delta_x(A)\nu(\mathrm{d}x) = \lambda^n\mu(A)\nu(f) + \int_M \mathcal{L}^n\nu_x(A)\nu(\mathrm{d}x). \tag{17}$$

From Proposition 3.3, there exist $\widetilde{K} > 0$ and $\delta \in (0,\lambda)$, such that $\|\mathcal{L}^n|_V\| < \widetilde{K}(\lambda-\delta)^n$, for every $n \in \mathbb{N}_0$ and therefore we can define

$$\alpha := \left|\log\left(\frac{\lambda-\delta}{\lambda}\right)\right|, \text{ and } K(\nu) := \widetilde{K}\frac{\nu(M)}{\nu(f)}\left(\sup_{x\in M}\|\nu_x\|_{TV}\right)(1+\|\mu\|_{TV}) < \infty.$$

Thus, from (17), we obtain that for every $n \in \mathbb{N}$,

$$\begin{aligned}
\left\|\mathbb{P}_\nu\left[X_n \in \cdot \mid \tau > n\right] - \mu\right\|_{TV} &= \left\|\frac{\int_M \mathcal{P}^n(x,\cdot)\nu(\mathrm{d}x)}{\int_M \mathcal{P}^n(x,M)\nu(\mathrm{d}x)} - \mu\right\|_{TV} \\
&= \left\|\frac{\lambda^n\nu(f)\mu + \int_M \mathcal{L}^n\nu_x(\cdot)\nu(\mathrm{d}x)}{\lambda^n\nu(f) + \int_M \mathcal{L}^n\nu_x(M)} - \mu\right\|_{TV} \\
&\le \frac{1}{\lambda^n\nu(f)}\left\|\int_M \mathcal{L}^n\nu_x(\cdot)\nu(\mathrm{d}x) - \mu\int_M \mathcal{L}^n\nu_x(M)\nu(\mathrm{d}x)\right\|_{TV} \\
&\le \frac{\nu(M)}{\nu(f)}\left(\frac{\|\mathcal{L}^n|_V\|}{\lambda^n}\right)\left(\sup_{x\in M}\|\nu_x\|_{TV}\right)\left(1+\|\mu\|_{TV}\right) \\
&\le K(\nu)e^{-\alpha n}. \quad \square
\end{aligned}$$

At this stage, a significant portion of Theorems B and C has already been proven. Theorem B still requires us to show that m is less or equal to the number of connected components of $M \setminus Z$, and Theorem C requires the proof of condition (M2). We will address these remaining aspects in the upcoming subsection.

6.2. Analysis of the case $m > 1$

In this section we show that $m > 1$ implies that the Markov chain X_n admits a cyclic behaviour and a quasi-ergodic measure on M. Also, in order to prove the second part of Theorem C, we always assume that $\rho(Z) = 0$.

To conduct the desired analysis we study linear transformations $\mathcal{P}^m$ and $\mathcal{L}^m$. From Proposition 6.2, it is clear that $r(\mathcal{P}^m) = r(\mathcal{L}^m) = \lambda^m$ and $\mathrm{Spec}(\mathcal{P}^m) \cap \lambda^m\mathbb{S}^1 = \mathrm{Spec}(\mathcal{L}^m) \cap \lambda^m\mathbb{S}^1 = \{\lambda^m\}$. Moreover,it also follows that $\dim(\ker(\mathcal{P}^m - \lambda^m)) = \dim(\ker(\mathcal{L}^m - \lambda^m)) = m$.

Throughout this section, we denote the *support of a function* $f : M \to \mathbb{R}$, as $\mathrm{supp}(f) = \{x \in M;\ f(x) \neq 0\}$.

In the following proposition, we study the eigenfunctions of $\mathcal{P}^m$ associated to the eigenvalue λ^m. A consequence of the proposition below is that

$$m \le \#\{\text{number of connected components of } M \setminus Z\}.$$

Proposition 6.6. *There exist eigenfunctions $g_1, \dots, g_{m-1} \in C^0_+(M)$ of the operator $\mathcal{P}^m$ such that $\int_M g_j \mathrm{d}\mu = 1$ for every $j \in \{0,1,\dots,m-1\}$ and $\mathrm{span}_{\mathbb{C}}(\{g_i\}_{i=0}^{m-1}) = \ker(\mathcal{P}^m - \lambda^m)$.*

Moreover, the eigenfunctions g_0, g_1, …, g_{m-1} can be chosen to have disjoint support, i.e., with $C_i := \mathrm{supp}(g_i)$ for all $i \in \{0,\dots,m-1\}$, satisfying $C_i \cap C_j = \emptyset$ for all $i \neq j$.

Furthermore, the family of sets $\{C_i\}_{i=0}^{m-1}$ satisfies $M \setminus Z = C_0 \sqcup C_1 \sqcup \dots \sqcup C_{m-1}$. In particular, $m \le \#\{$connected components of $M \setminus Z\}$.

Proof. Observe that since $\lambda^m \in \mathbb{R}$ and $\mathcal{P}(C^0(M)) \subset C^0(M)$, it follows that if $f \in C^0(M,\mathbb{C})$ satisfies $\mathcal{P}^m f = \lambda^m f$, then $\mathcal{P}^m\mathrm{Re}(f) = \lambda^m\mathrm{Re}(f)$ and $\mathcal{P}^m\mathrm{Im}(f) = \lambda^m\mathrm{Im}(f)$.

Let μ be the unique quasi-stationary measure of X_n given by Theorem A. Note that the operator $\mathcal{P}^m$ satisfies $\int_M \frac{1}{\lambda^m}\mathcal{P}^m f(x)\mu(\mathrm{d}x) = \int_M f(x)\mathrm{d}\mu$ for every $f \in C^0(M)$. By the same techniques of [46, Theorems 3.1.1 and 3.1.3] it follows that if $f \in C^0(M)$ is an eigenfunction of $\mathcal{P}^m$ with eigenvalue λ^m, then $f^+(x) = \max\{0, f(x)\}$ and $f^-(x) = \max\{0, -f(x)\}$ are also eigenfunctions of $\mathcal{P}^m$ with λ^m.

We will divide the proof into two steps.

Step 1. *We prove that if $h_1, h_2 \in C^0_+(M)$ are eigenfunctions of $\mathcal{P}^m$ with eigenvalue λ^m such that $G := \{h_1 > 0\} \setminus \{h_2 > 0\} \neq \emptyset$, then $\mathbb{1}_G h_1$ is an eigenfunction of $\mathcal{P}^m$ with eigenvalue λ^m.*

By the observations at the beginning of the proof, for all $t \in \mathbb{R}_{\ge 0}$, $h_t = (h_1 - th_2)^+$ is an eigenfunction for $\mathcal{P}^m$, with eigenvalue λ^m.

Note that, $(h_1 - sh_2)^+ \le (h_1 - th_2)^+$, for all $s > t$. We claim, that $\{(h_1 - th_2)^+\}_{t\in\mathbb{R}_+}$ stabilises on t, i.e. there exists $t_0 \ge 0$, such that $(h_1 - t_0h_2)^+ = (h_1 - th_2)^+$, for every $t > t_0$.

Suppose by contradiction that the above statement is false. Then, we can find $t_1 < t_2 < \cdots < t_{m+1}$, such that

$$\operatorname{supp}\left((h_1 - t_{m+1}h_2)^+\right) \subsetneq \operatorname{supp}\left((h_1 - t_m h_2)^+\right) \subsetneq \ldots \subsetneq \operatorname{supp}\left((h_1 - t_1 h_2)^+\right).$$

The above equation implies that $\{(h_1 - t_1h_2)^+, \ldots, (h_1 - t_{m+1}h_2)^+\}$ is a linearly independent set in $C^0(M)$. This generates a contradiction since $\mathcal{P}^m$ admits only m eigenfunctions with eigenvalue λ^m.

Therefore, $\{(h_1 - th_2)^+\}_{t\in\mathbb{R}_{\geq 0}}$ stabilises at some t_0. Finally, $\lim_{t\to\infty}(h_1 - th_2)^+ = \mathbb{1}_G h_1 = (h_1 - t_0h_2)^+$, and $(h_1 - t_0h_2)^+$ is an eigenfunction of $\mathcal{P}^m$.

Step 2. *We complete the proof of the proposition.*

By Step 1 and the observations in the first two paragraphs, from linear combinations of the positive and negative parts of $\{f_i\}_{i=0}^{m-1}$ one can construct normalised eigenfunctions $g_0, \ldots, g_{m-1} \in C^0_+(M)$ of $\mathcal{P}^m$, such that $\operatorname{span}_{\mathbb{C}}(\{g_i\}_{i=0}^{m-1}) = \ker(\mathcal{P}^m - \lambda^m)$, and the family of sets $\{C_i := \operatorname{supp}(g_i)\}_{i=0}^{m-1}$ fulfils $C_i \cap C_j = \emptyset$ for all $i \neq j$.

To prove the last part of the proposition, let $f \in C^0_+(M)$ be an eigenfunction of $\mathcal{P}$ with eigenvalue λ. From Theorem 4.4, $\operatorname{supp}(f) = M \setminus Z$. Since $f \in \ker(\mathcal{P}^m - \lambda^m) = \operatorname{span}_{\mathbb{C}}\left(\{g_i\}_{i=0}^{m-1}\right)$, there exist $\alpha_0, \ldots, \alpha_{m-1} \geq 0$, such that $f = \alpha_0 g_0 + \cdots + \alpha_{m-1}g_{m-1}$, implying that $M \setminus Z = C_0 \sqcup \ldots \sqcup C_{m-1}$. Since each C_i is open and closed in the topology in induced by $M \setminus Z$, we have that $m \leq \#\{\text{connected components on } M \setminus Z\}$. □

Proof of Theorem B. The proof follows directly from Theorem 4.4 and Proposition 6.6. □

From now on, we consider $\{g_i\}_{i=0}^{m-1} \subset C^0_+(M)$ as in Proposition 6.6 and $\{C_i := \operatorname{supp}(g_i)\}_{i=0}^{m-1}$. We may decompose

$$C^0(M) = \operatorname{span}_{\mathbb{R}}(g_0, g_1, \ldots, g_{m-1}) \oplus V_0, \tag{18}$$

where $r\left(\mathcal{P}^m|_{V_0}\right) < \lambda^m$. For convenience we denote $C_i = C_{i \pmod m}$, for every $i \in \mathbb{N}_0$.

We proceed to address the cyclic property of the eigenvectors of $\mathcal{P}^m$ and $\mathcal{L}^m$, when $m > 1$. The following four results show that we may choose suitable eigenfunctions and eigenmeasures of the operators $\mathcal{P}^m$ and $\mathcal{L}^m$, respectively, so that these permute cyclically, by application of $\mathcal{P}$ and $\mathcal{L}$.

Proposition 6.7. *Let $j \in \{0, 1, \ldots, m-1\}$, then $\operatorname{supp}(\mathcal{P}^m(\cdot, A)) \subset C_j$, for all $A \in \mathscr{B}(C_j)$. Assume that $\rho(Z) = 0$ and let μ be the unique probability measure on $\mathcal{M}(M)$ such that $\mathcal{L}\mu = \lambda\mu$. Then, for every $i \in \{1, \ldots, m-1\}$, $\nu_i(\mathrm{d}x) := \mu(C_i \cap \mathrm{d}x)$ is a probability eigenmeasure of $\mathcal{L}^m$ with eigenvalue λ^m.*

Proof. Let $A \in \mathscr{B}(C_j)$. For every $n \in \mathbb{N}$, let us define $A_n := A \cap \{g_j > 1/n\}$. Since $\mathcal{P}^m$ is a positive operator, it follow that for every $n \in \mathbb{N}$, $\mathcal{P}^m(x, A_n) \leq n\lambda^m g_j(x)$. This implies that $\operatorname{supp}\left(\mathcal{P}^m(\cdot, A_m)\right) \subset C_j$ for every $n \in \mathbb{N}$. Since C_i is an open set and $\operatorname{supp}(\mathcal{P}(\cdot, A)) \subset M \setminus Z$, Proposition 3.2(b) completes the proof.

For the last part of the proposition, observe that the above paragraph implies that for every $B \in \mathscr{B}(M)$ and $i \in \{1, \ldots, m-1\}$,

$$\lambda^m \nu_i(B) = \lambda^m \mu(B \cap C_i) = \int_M \mathcal{P}^m(x, B \cap C_i)\mu(\mathrm{d}x) = \int_M \mathcal{P}^m(x, B)\nu_i(\mathrm{d}x).$$

In the last equality, we used that $\{\mathcal{P}^m(x, B \cap C_i) > 0\} \cap C_j = \emptyset$ if $i \neq j$, and that $\mathcal{P}^m(x, B \cap Z) = 0$ since $\rho(Z) = 0$. In the last equality, we used that $\{\mathcal{P}^m(x, B \cap C_i) > 0\} \cap C_j\} = \emptyset$ if $i \neq j$, and that $\mathcal{P}^m(x, B \cap Z) = 0$ since $\rho(Z) = 0$. □

From now on, we consider $\{\nu_i\}_{i=0}^{m-1}$ as in Proposition Proposition 6.7. The following two lemmas establish a combinatorial structure in the set of measures $\{\nu_i\}_{i=0}^{m-1}$ (and the set of functions $\{g_i\}_{i=0}^{m-1}$). We omit the proof of the first lemma since it follows directly from standard linear algebra techniques.

Lemma 6.8. *Let $A = [a_{ij}]_{i,j=1}^m$ be $m \times m$ matrix, such that the following properties hold:*

(i) $A^m = \mathrm{Id}$;
(ii) $\sum_{j=1}^m a_{ij} \leq 1$, *for all* $i \in \{1, \ldots, m\}$; *and*
(iii) $a_{ij} \geq 0$, *for all* $i, j \in \{1, \ldots, m\}$.

Then, the matrix A permutes the canonical basis. In other words, for every $k \in \{1, \ldots, m\}$ there exists $s \in \{1, \ldots, m\}$, such that $Ae_k = e_s$.

Lemma 6.9. *Let $\{g_i\}_{i=0}^{m-1} \subset C^0_+(M)$ as in Proposition Proposition 6.6 and $\{\nu_i\}_{i=0}^{m-1} \subset \mathcal{M}_+(M)$ as in Proposition Proposition 6.7. Then, there exists a cyclic permutation $\sigma : \{0, 1, \ldots, m-1\} \to \{0, 1, \ldots, m-1\}$ of order m such that for every $i \in \{0, 1, \ldots, m-1\}$, $\frac{1}{\lambda}\mathcal{L}\nu_i = \nu_{\sigma(i)}$ and $\frac{1}{\lambda}\mathcal{P}g_i = g_{\sigma^{-1}(i)}$. Moreover $\operatorname{supp}\mathcal{P}(\cdot, C_i) = C_{\sigma^{-1}(i)}$ for every $i \in \{0, 1, \ldots, m-1\}$.*

Proof. We divide the proof into four steps.

Step 1. *There exists a permutation σ on $\{0, 1, \ldots, m-1\}$ such for every $i \in \{0, 1, \ldots, m-1\}$, $\frac{1}{\lambda}\mathcal{P}g_i = g_{\sigma^{-1}(i)}$ and $\operatorname{supp}\mathcal{P}(\cdot, C_i) = C_{\sigma^{-1}(i)}$.*

Given $i \in \{0, 1, \dots, m-1\}$, note that $\mathcal{P}^m\left(\mathcal{P}g_i\right) = \mathcal{P}(\mathcal{P}^m g_i) = \lambda^m \mathcal{P}g_i$, implying that $\mathcal{P}g_i$ is an eigenmeasure of $\mathcal{P}^m$ with eigenvalue λ^m. Since $\mathcal{P}$ is a positive operator, $\mathcal{P}g_i \in C^0_+(M)$. This implies that, for every $i \in \{0, 1, \dots, m-1\}$ there exist $\alpha_{i0}, \dots, \alpha_{im-1} \in \mathbb{R}_+$, such that $\frac{1}{\lambda}\mathcal{P}g_i = \sum_{j=0}^{m-1} \alpha_{ij} g_j$.

Note that, since

$$1 = \int g_i(x)\ \mu(\mathrm{d}x) = \int \frac{1}{\lambda}\mathcal{P}g_i(x)\ \mu(\mathrm{d}x) = \sum_{j=0}^{m-1} \alpha_{ij} \int_M g_j(x)\mu(\mathrm{d}x) = \sum_{j=0}^{m-1} \alpha_{ij}$$

for all $i \in \{0, 1, \dots, m-1\}$, defining the $m \times m$ matrix $A := \left[\alpha_{ij}\right]_{i,j=0}^{m-1}$, we obtain that $A^m = \mathrm{Id}$. By Lemma Lemma 6.8, the matrix A permutes the canonical basis, and therefore, for every $i \in \{0, 1, \dots, m-1\}$, there exists $j_i \in \{0, 1, \dots, m-1\}$, such that $\frac{1}{\lambda}\mathcal{P}g_i = g_{j_i}$.

Setting $\sigma : \{0, 1, \dots, m-1\} \to \{0, 1, \dots, m-1\}$ such that $\sigma^{-1}(i) = j_i$ we conclude the first step.

Step 2. *σ is a cyclic permutation of order m.*

Suppose that the permutation σ admits a k-subcycle $\widetilde{\sigma}$. Without loss of generality assume that $\widetilde{\sigma} = (0, \widetilde{\sigma}(0), \dots, \widetilde{\sigma}^{k-1}(0))$. Defining $\widetilde{f} = \frac{1}{k}\sum_{i=0}^{k-1} g_{\widetilde{\sigma}^{-i}(0)}$, we have that $\mathcal{P}\widetilde{f} = \lambda\widetilde{f}$, implying that $\widetilde{f} \in \ker(\mathcal{P} - \lambda\mathrm{Id})$. Since $\dim\ker(\mathcal{P} - \lambda\mathrm{Id}) = 1$ we obtain that f and $\widetilde{f}$ are linearly dependent, therefore $k = m$.

Step 3. *We conclude the proof of the lemma.*

Let σ be the m-cycle constructed in Steps Step 1 and Step 2, such that

$$\frac{1}{\lambda}\mathcal{P}g_i = g_{\sigma^{-1}(i)}, \text{ for every } i \in \{0, 1, \dots, m-1\}. \tag{19}$$

It remains to show that $\frac{1}{\lambda}\mathcal{L}\nu_i = \nu_{\sigma(i)}$, for every $i \in \{0, 1, \dots, m-1\}$. As in Step 1, for every $j \in \{0, 1, \dots, m-1\}$, $\mathcal{L}\nu_j$ is an eigenfunction of $\mathcal{L}^m$. Therefore, there exists $\beta_{j0}, \dots, \beta_{jm-1} > 0$, such that $\frac{1}{\lambda}\mathcal{L}\nu_j = \sum_{i=0}^{m-1} \beta_{ji}\nu_i$. From equation 19 and duality, we obtain

$$\frac{1}{\lambda}\langle \mathcal{L}\nu_i, g_j\rangle = \frac{1}{\lambda}\langle \nu_i, \mathcal{P}g_j\rangle = \langle \nu_i, g_{\sigma^{-1}(j)}\rangle = \delta_{\sigma(i)j}\langle \nu_i, g_{\sigma^{-1}(j)}\rangle,$$

for every $i, j \in \{0, 1, \dots, m-1\}$. Since $\frac{1}{\lambda}\mathcal{L}\nu_j = \sum_{i=0}^{m-1} \beta_{ji}\nu_i$ and $\langle \nu_i, g_j\rangle = \delta_{ij}$, we obtain that $\mathcal{L}\nu_i = \lambda\nu_{\sigma(i)}$. □

The previous lemma establishes that $\mathrm{supp}(\mathcal{P}(\cdot, C_i)) = C_{\sigma^{-1}(i)}$, for all $i \in \{0, 1, \dots, m-1\}$, and $\mathrm{supp}(\mathcal{P}^k(\cdot, C_i)) \cap \mathrm{supp}(\mathcal{P}^s(\cdot, C_i)) = \emptyset$ for all $k, s \in \mathbb{N}$ with $k \neq s \pmod m$. From Lemma 6.9, we label the components C_i such that

$$\mathcal{P}(x, C_i) \begin{cases} = 0, \text{ if } x \notin C_{i-1} \\ > 0, \text{ if } x \in C_{i-1}, \end{cases} \tag{20}$$

and for every $i \in \mathbb{N}_0$,

$$\frac{1}{\lambda}\mathcal{P}g_i = g_{i-1 \pmod m} \text{ and } \frac{1}{\lambda}\mathcal{L}\nu_i = \frac{1}{\lambda}\int_M \mathcal{P}(x, \cdot)\nu_i(\mathrm{d}x) = \nu_{i+1 \pmod m}. \tag{21}$$

We proceed to characterise the eigenvectors of the operators $\mathcal{P}$ and $\mathcal{L}$ and show that their eigenvalues lie in $\lambda\mathbb{S}^1 \subset \mathbb{C}$.

Theorem 6.10. *Let X_n be a Markov chain on E_M absorbed at ∂ satisfying (H) and $\rho(Z) = 0$. Let $\{g_i\}_{i=0}^{m-1} \subset C^0_+(M)$ and $\{\nu_i\}_{i=0}^{m-1} \subset \mathcal{M}_+(M)$ be as in Lemma 6.9 and in (20)–(21). Then, for every $j \in \{0, 1, \dots, m-1\}$, the measure*

$$\mu_j = \frac{1}{m}\sum_{k=0}^{m-1} e^{\frac{-2\pi i k j}{m}}\nu_k, \text{ and the function } f_j = \sum_{k=0}^{m-1} e^{\frac{2\pi i k j}{m}} g_k,$$

satisfy

$$\mathcal{L}\mu_j = \lambda e^{\frac{2\pi i j}{m}}\mu_j \quad \text{and } \mathcal{P}f_j = \lambda e^{\frac{2\pi i j}{m}} f_j. \tag{22}$$

Moreover, the measures $\eta_j(\mathrm{d}x) := f_j(x)\mu_j(\mathrm{d}x)$ do not depend on $j \in \{0, 1, \dots, m-1\}$.

Proof. First, we verify that (22) holds. Consider $j \in \{0, 1, \dots, m-1\}$. From Lemma 6.9 and re-ordering of $\{\nu_i\}_{i=0}^{m-1}$ and $\{g_i\}_{i=0}^{m-1}$ we have that

$$\mathcal{L}\mu_j = \mathcal{L}\left(\frac{1}{m}\sum_{k=0}^{m-1} e^{\frac{-2\pi i k j}{m}}\nu_k\right) = \frac{1}{m}\sum_{k=0}^{m-1} \lambda e^{\frac{-2\pi i k j}{m}}\nu_{k+1} = \lambda e^{\frac{2\pi i j}{m}}\mu_j$$

and

$$\mathcal{P}f_j = \mathcal{P}\left(\sum_{k=0}^{m-1} e^{\frac{2\pi i k j}{m}} g_k\right) = \sum_{k=0}^{m-1} \lambda e^{\frac{2\pi i k j}{m}} g_{k-1} = \lambda e^{\frac{2\pi i j}{m}} f_j.$$

From Lemma 6.9 (*b*), we have for every $j \in \{0, 1, \dots, m-1\}$ and $A \in \mathscr{B}(M)$

$$\begin{aligned}\eta_j(A) &= \int_A f_j(x)\mu_j(\mathrm{d}x) = \frac{1}{m}\int_A \left(\sum_{k=0}^{m-1} e^{\frac{2\pi i k j}{m}} g_k(x)\right)\left(\sum_{s=0}^{m-1} e^{\frac{-2\pi i j s}{m}} \nu_s(\mathrm{d}x)\right)\\ &= \frac{1}{m}\sum_{k=0}^{m-1}\int_A g_k(x)\nu_k(\mathrm{d}x) + \frac{1}{m}\sum_{k\neq s} e^{\frac{2\pi i j(k-s)}{m}} \underbrace{\int_A g_k(x)\nu_s(\mathrm{d}x)}_{=0}\\ &= \int_A f_0(x)\mu_0(\mathrm{d}x) = \eta_0(A),\end{aligned}$$

which completes the proof. □

Remark 6.11. In particular, Theorem 6.10 establishes that $\mu_0 = \frac{1}{m}\sum_{j=0}^{m-1}\nu_j$, is the unique quasi-stationary measure for X_n, given by Theorem A.

Corollary 6.12. *In the context of Theorem 6.10, if the Markov chain X_n satisfies (H) and $\rho(Z) = 0$, then for every $x \in M \setminus Z$ and $h \in \mathcal{F}_b(M)$,*

$$\mathbb{E}_x\left[\sum_{k=0}^{n-1} h(X_k)\mathbb{1}_M(X_n)\right] = n\lambda^n\langle\mu_0, f_0\cdot h\rangle\sum_{\ell=0}^{m-1} e^{\frac{2\pi i n\ell}{m}} f_\ell(x)\langle\mu_\ell, \mathbb{1}_M\rangle + o(n\lambda^n). \tag{23}$$

Proof. Combining Lemma 6.4, Theorem 6.10, and since $\langle\mu_\ell, hf_\ell\rangle = \langle\mu_0, hf_0\rangle$ for every $\ell \in \{0, 1, \dots, m-1\}$ we obtain (23). □

The following result establishes the existence of a quasi-ergodic measure for X_n on M, in the case that X_n satisfies Hypothesis (H) and $\rho(Z) = 0$.

Theorem 6.13. *Let X_n be a Markov chain on E_M absorbed at ∂ that satisfies Hypothesis (H) and $\rho(Z) = 0$. Then X_n admits a quasi-ergodic measure η on M and (4) holds for every $x \in M \setminus Z$.*

Moreover, $\eta(\mathrm{d}x) = f_0(x)\mu_0(\mathrm{d}x)$, where $f_0 \in C^0_+(M)$ and $\mu_0 \in \mathcal{M}_+(M)$, are such that $\mathcal{P}f_0 = \lambda f_0$, $\mathcal{L}\mu_0 = \lambda\mu_0$, and $\langle\mu_0, f_0\rangle = 1$.

Proof. Let h be a bounded measurable function. Recall from Definition 2.5 that we need to show that

$$\lim_{n\to\infty}\mathbb{E}_x\left[\frac{1}{n}\sum_{i=0}^{n-1} h\circ X_i \;\middle|\; \tau > n\right] = \int_M h(y)\eta(\mathrm{d}y),$$

for η-almost every $x \in M$. Observe that $\eta(Z) = 0$ and let $x \in X \setminus Z$.

From Proposition 6.3 we have

$$0 < \mathcal{P}^n(x, M) = \langle\mathcal{L}^n\delta_x, \mathbb{1}_M\rangle = \sum_{j=0}^{m-1}\lambda^n f_j(x)e^{\frac{2\pi i n j}{m}}\langle\mu_j, \mathbb{1}_M\rangle + o(\lambda^n). \tag{24}$$

Moreover, for every $n \in \mathbb{N}$

$$\mathbb{E}_x\left[\frac{1}{n}\sum_{k=0}^{n-1} h(X_i)\middle|\tau > n\right] = \frac{\mathbb{E}_x\left[\frac{1}{n}\sum_{k=0}^{n-1} h(X_k)\mathbb{1}_M(X_n)\right]}{\mathcal{P}^n(x, M)}.$$

Applying Corollary 6.12 for the numerator of the above equation, and (24) for the denominator, we obtain

$$\begin{aligned}\mathbb{E}_x\left[\frac{1}{n}\sum_{k=0}^{n-1} h(X_i)\middle|\tau > n\right] &= \frac{\lambda^n\langle\mu_0, h\cdot f_0\rangle\sum_{\ell=0}^{m-1} e^{\frac{2\pi i n\ell}{m}} f_\ell(x)\langle\mu_\ell, \mathbb{1}_M\rangle + o(n\lambda^n)}{\lambda^n\sum_{j=0}^{m-1} f_j(x)e^{\frac{2\pi i n j}{m}}\langle\mu_j, \mathbb{1}_M\rangle + o(\lambda^n)}\\ &= \langle\mu_0, h\cdot f_0\rangle\frac{\sum_{\ell=0}^{m-1} e^{\frac{2\pi i n\ell}{m}} f_\ell(x)\langle\mu_\ell, \mathbb{1}_M\rangle + o(1)_{n\to\infty}}{\sum_{j=0}^{m-1} e^{\frac{2\pi i n j}{m}} f_j(x)\langle\mu_j, \mathbb{1}_M\rangle + o(1)_{n\to\infty}}.\end{aligned}$$

The proof concludes by taking limit $n \to \infty$. □

We proceed to prove Theorem C.

Proof of Theorem C. Observe that Theorems Theorem 6.5 and Theorem 6.13 imply that $\eta(\mathrm{d}x) = f(x)\mu(\mathrm{d}x)$ is a quasi-ergodic measure for X_n on M and Eq. (4) holds for every $x \in M \setminus Z$. Moreover, (M1) follows directly from Theorem Theorem 6.5 and (M2) is a consequence of Lemma Lemma 6.9. □

Declaration of competing interest

There is no conflict of interest.

Acknowledgements

The authors gratefully thank Bernat Bassols-Cornudella for his valuable comments and suggestions. MC's research has been supported by an Imperial College President's PhD scholarship and FAPESP (process 2019/06873-2). MC and JL are also supported by the EPSRC Centre for Doctoral Training in Mathematics of Random Systems: Analysis, Modelling and Simulation (EP/S023925/1). JL gratefully acknowledges research support from IRCN, University of Tokyo and CAMB, Gulf University of Science and Technology, as well as from the London Mathematical Laboratory. GOM's research has been supported by a PhD scholarship from CONAHCYT (formerly CONACYT) and the Department of Mathematics of Imperial College, the Deutsche Forschungsgemeinschaft (DFG) through the CRC 1114 (Project A02) and under Germany's Excellence Strategy — The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

Appendix. Proof of Lemma 6.4

This appendix contains a proof of Lemma 6.4. The proof below is inspired by [5, Proposition 4]. Although the results in [5] are focused on finite state spaces, it is possible to extend these results to our setting by making several adaptations. Since these are not straightforward, we present them here in detail.

The results of [5] extend classical results by Darroch and Senata [21], where similar bounds are found for irreducible finite Markov chains.

Proof of Lemma 6.4. Fix $x \in M \setminus Z$ and $h \in \mathcal{F}_b(M)$. We divide the proof into three steps.

Step 1. *We prove that for every $n \in \mathbb{N}$,*

$$\begin{aligned}\sum_{k=0}^{n-1} \mathbb{E}_x\left[h(X_k)\mathbb{1}_M(X_n)\right] &= \sum_{k=0}^{n-1}\sum_{\ell=0}^{m-1}\sum_{j=0}^{m-1} \lambda^n e^{\frac{2\pi i \ell k}{m}+\frac{2\pi i (n-k) j}{m}} f_\ell(x)\langle \mu_\ell, h f_j\rangle \langle \mu_j, \mathbb{1}_M\rangle \\ &\quad + \sum_{k=0}^{n-1}\left\langle \sum_{\ell=0}^{m-1} \lambda^k e^{\frac{2\pi i \ell k}{m}} f_\ell(x)\mu_\ell, h\langle \mathcal{L}^{n-k}\nu_\cdot, \mathbb{1}_M\rangle \right\rangle \\ &\quad + \sum_{k=0}^{n-1}\left\langle \mathcal{L}^k \nu_x, h\cdot \mathcal{P}^{n-k}\mathbb{1}_M\right\rangle. \end{aligned} \tag{A.1}$$

First, observe that

$$\mathbb{E}_x\left[h(X_n)\mathbb{1}_m\right] = \mathcal{P}^n h(x) = \langle \delta_x, \mathcal{P}^n h\rangle = \langle \mathcal{L}^n \delta_x, h\rangle. \tag{A.2}$$

From the Markov property of X_n, for every $k, n \in \mathbb{N}$, with $k \le n$, we have

$$\begin{aligned}\mathbb{E}_x\left[h(X_k)\mathbb{1}_M(X_n)\right] &= \mathbb{E}_x\left[\mathbb{E}_x\left[h(X_k)\mathbb{1}_M(X_n)\mid \mathcal{F}_k\right]\right] = \mathbb{E}_x\left[h(X_k)\mathcal{P}^{n-k}(X_k, M)\right] \\ &= \mathcal{P}^k(h\cdot \mathcal{P}^{n-k}\mathbb{1}_M)(x).\end{aligned}$$

Using (A.2) and Proposition 6.3 this implies that for every $y \in M$,

$$h(y)\mathcal{P}^{n-k}(y, M) = h(y)\langle \delta_y, \mathcal{P}^{n-k}\mathbb{1}_M\rangle = \sum_{j=0}^{m-1} \lambda^{n-k} e^{\frac{2\pi i (n-k) j}{m}} h(y) f_j(y)\langle \mu_j, \mathbb{1}_M\rangle + h(y)\langle \mathcal{L}^{n-k}\nu_y, \mathbb{1}_M\rangle.$$

Recall from Proposition 6.3 that $\sup_{y\in M} \|\mathcal{L}^n \nu_y\|_{TV} = \mathcal{O}(\lambda^n)$. Hence,

$$\begin{aligned}\mathbb{E}_x\left[h(X_k)\mathbb{1}_m\right] &= \mathcal{P}^k(h\cdot \mathcal{P}^{n-k}\mathbb{1}_M)(x) = \left\langle \mathcal{L}^k\delta_x, h\cdot \mathcal{P}^{n-k}\mathbb{1}_M\right\rangle \\ &= \left\langle \sum_{\ell=0}^{m-1} \lambda^k e^{\frac{2\pi i \ell k}{m}} f_\ell(x)\mu_\ell, h\cdot \mathcal{P}^{n-k}\mathbb{1}_M\right\rangle + \left\langle \mathcal{L}^k\nu_x, h\cdot \mathcal{P}^{n-k}\mathbb{1}_M\right\rangle \\ &= \left\langle \sum_{\ell=0}^{m-1} \lambda^k e^{\frac{2\pi i \ell k}{m}} f_\ell(x)\mu_\ell, \sum_{j=0}^{m-1} \lambda^{n-k} e^{\frac{2\pi i (n-k) j}{m}} h f_j\langle \mu_j, \mathbb{1}_M\rangle + h\langle \mathcal{L}^{n-k}\nu_y, \mathbb{1}_M\rangle\right\rangle \\ &\quad + \langle \mathcal{L}^k\nu_x, h\cdot \mathcal{P}^{n-k}\mathbb{1}_M\rangle \\ &= \sum_{\ell=0}^{m-1}\sum_{j=0}^{m-1} \lambda^n e^{\frac{2\pi i \ell k}{m}+\frac{2\pi i (n-k) j}{m}} f_\ell(x)\langle \mu_\ell, h f_j\rangle \langle \mu_j, \mathbb{1}_M\rangle \\ &\quad + \left\langle \sum_{\ell=0}^{m-1} \lambda^k e^{\frac{2\pi i \ell k}{m}} f_\ell(x)\mu_\ell, h\langle \mathcal{L}^{n-k}\nu_\cdot, \mathbb{1}_M\rangle\right\rangle + \langle \mathcal{L}^k\nu_x, h\cdot \mathcal{P}^{n-k}\mathbb{1}_M\rangle.\end{aligned}$$

Then, for each $n \in \mathbb{N}$, (A.1) holds. This completes Step 1.

Step 2. *We prove that the following identity holds*

$$\sum_{k=0}^{n-1}\left\langle\sum_{\ell=0}^{m-1}\lambda^k e^{\frac{2\pi i\ell k}{m}} f_\ell(x)\mu_\ell, h\langle\mathcal{L}^{n-k}\nu_\cdot, \mathbb{1}_M\rangle\right\rangle + \sum_{k=0}^{n-1}\langle\mathcal{L}^k\nu_x, h\mathcal{P}^{n-k}\mathbb{1}_M\rangle = o(n\lambda^n). \tag{A.3}$$

Recall from Proposition 6.3 that $\sup_{y\in M}\|\mathcal{L}^n\nu_y\| = o(\lambda^n)$. Hence, for all $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $n > n_0$ implies

$$\frac{1}{\lambda^n}\|\mathcal{L}^n\nu_x\|_{TV} < \varepsilon. \tag{A.4}$$

On the other hand, recall from Eq. (10) that $\|\mathcal{P}^n\| = \mathcal{O}(\lambda^n)$. Thus, there exists $K \geq 0$ such that $\|\mathcal{P}^n\| \leq K\lambda^n$, for every $n \geq 0$. We first discuss the second term in (A.3). Note that for every $n > n_0 + 1$,

$$\begin{aligned}\frac{1}{\lambda^n n}\sum_{k=0}^{n-1}\langle\mathcal{L}^k\nu_x, h\cdot\mathcal{P}^{n-k}\mathbb{1}_M\rangle \leq& \frac{1}{\lambda^n n}\sum_{k=0}^{n-1}\|\mathcal{L}^k\nu_x\|_{TV}\|h\|_\infty\|\mathcal{P}^{n-k}\|\\ \leq& \frac{K\|h\|_\infty}{n}\sum_{k=0}^{n-1}\frac{\|\mathcal{L}^k\nu_x\|_{TV}}{\lambda^k} \leq \frac{K\|h\|_\infty}{n}\left(\sum_{k=0}^{n_0-1}\frac{\|\mathcal{L}^k\nu_x\|_{TV}}{\lambda^k} + \sum_{j=n_0}^{n-1}\varepsilon\right)\\ \longrightarrow& K\|h\|_\infty\varepsilon, \text{ as } n\to\infty.\end{aligned}$$

Hence, since ε is arbitrary,

$$\sum_{k=0}^{n-1}\langle\mathcal{L}^k\nu_x, h\cdot\mathcal{P}^{n-k}\mathbb{1}_M\rangle = o(\lambda^n n). \tag{A.5}$$

On the other hand, defining for each $n \in \mathbb{N}$

$$I_n := \sum_{k=0}^{n-1}\left\langle\sum_{\ell=0}^{m-1}\lambda^k e^{\frac{2\pi i\ell k}{m}} f_\ell(x)\mu_\ell, h\langle\mathcal{L}^{n-k}\nu_\cdot, \mathbb{1}_M\rangle\right\rangle,$$

we have

$$\begin{aligned}\frac{1}{n\lambda^n}I_n \leq& \frac{1}{n\lambda^n}\sum_{k=0}^{n-1}\sum_{\ell=0}^{m-1}\lambda^k\|f_\ell\|_\infty\|\mu_\ell\|_{TV}\|h\|_\infty\sup_{y\in M}\|\mathcal{L}^{n-k}\nu_y\|_{TV}\\ =& C\frac{1}{n}\sum_{k=0}^{n-1}\frac{\sup_{y\in M}\|\mathcal{L}^{n-k}\nu_y\|_{TV}}{\lambda^{n-k}}\end{aligned}$$

where $C := m\max_{l\in\{0,\dots,m-1\}}(\|h\|_\infty\|f_\ell\|_\infty\|\mu_\ell\|_{TV})$. Hence, by (A.4)

$$\begin{aligned}\frac{1}{n\lambda^n}I_n \leq& C\left(\frac{1}{n}\sum_{k=0}^{n_0-1}\frac{\sup_{y\in M}\|\mathcal{L}^k\nu_y\|_{TV}}{\lambda^k} + \frac{1}{n}\sum_{k=n_0}^{n-1}\frac{\sup_{y\in M}\|\mathcal{L}^k\nu_y\|_{TV}}{\lambda^k}\right)\\ \leq& C\left(\frac{1}{n}\sum_{k=0}^{n_0-1}\frac{\sup_{y\in M}\|\mathcal{L}^k\nu_y\|_{TV}}{\lambda^k} + \frac{1}{n}\sum_{k=n_0}^{n-1}\varepsilon\right)\\ \longrightarrow& C\varepsilon, \text{ when } n\to\infty,\end{aligned}$$

Once again, since ε is arbitrary

$$\sum_{k=0}^{n-1}\left\langle\sum_{\ell=0}^{m-1}\lambda^k e^{\frac{2\pi i\ell k}{m}} f_\ell(x)\mu_\ell, h\langle\mathcal{L}^{n-k}\nu_\cdot, \mathbb{1}_M\rangle\right\rangle = o(n\lambda^n). \tag{A.6}$$

The sought identity is implies by (A.5) and (A.6).

Step 3. *We prove that for every* $n \in \mathbb{N}$.

$$\mathbb{E}_x\left[\sum_{k=0}^{n-1}h(X_k)\mathbb{1}_M(X_n)\right] = n\lambda^n\sum_{\ell=0}^{m-1}e^{\frac{2\pi in\ell}{m}}f_\ell(x)\langle\mu_\ell, h\cdot f_\ell\rangle\mu_\ell(M) + o(n\lambda^n).$$

Note that by the previous two steps, $\sum_{k=0}^{n-1}\mathbb{E}_x\left[h(X_k)\mathbb{1}_M(X_n)\right] = A + o(n\lambda^n)$, where

$$A := \sum_{k=0}^{n-1}\sum_{\ell=0}^{m-1}\sum_{j=0}^{m-1}\lambda^n e^{\frac{2\pi i\ell k}{m}+\frac{2\pi i(n-k)j}{m}}f_\ell(x)\langle\mu_\ell, hf_j\rangle\langle\mu_j, \mathbb{1}_M\rangle.$$

By exchanging the order of the sums, we obtain that

$$A = \sum_{\ell=0}^{m-1}\sum_{j=0}^{m-1}\left(\sum_{k=0}^{n-1}e^{\frac{2\pi inj}{m}+\frac{2\pi i(\ell-j)k}{m}}\right)\lambda^n f_\ell(x)\langle\mu_\ell, hf_j\rangle\langle\mu_j, \mathbb{1}_M\rangle.$$

Splitting the above double sum into $\ell = j$ and $\ell \neq j$ we obtain

$$A = \sum_{\ell=0}^{m-1} n\lambda^n e^{\frac{2\pi i\ell n}{m}} f_\ell(x)\langle \mu_\ell, hf_j\rangle\langle \mu_\ell, \mathbb{1}_M\rangle + \sum_{\ell\neq j} \lambda^n e^{\frac{2\pi ij}{m}} \left(\frac{e^{\frac{2\pi i\ell n}{m}} - e^{\frac{2\pi ijn}{m}}}{e^{\frac{2\pi i\ell}{m}} - e^{\frac{2\pi ij}{m}}}\right) f_\ell(x)\langle \mu_\ell, hf_j\rangle\langle \mu_j.\mathbb{1}_M\rangle$$

Since $e^{\frac{2\pi ij}{m}} \left(\frac{e^{\frac{2\pi i\ell n}{m}} - e^{\frac{2\pi ijn}{m}}}{e^{\frac{2\pi i\ell}{m}} - e^{\frac{2\pi ij}{m}}}\right)$ is uniformly bounded in n for $\ell, j \in \{0, 1, \dots, m-1\}$ and $\ell \neq j$, we obtain that

$$\sum_{l\neq k} \lambda^n e^{\frac{2\pi ij}{m}} \left(\frac{e^{\frac{2\pi i\ell n}{m}} - e^{\frac{2\pi ijn}{m}}}{e^{\frac{2\pi i\ell}{m}} - e^{\frac{2\pi ij}{m}}}\right) f_\ell(x)\langle \mu_\ell, hf_j\rangle\langle \mu_j, \mathbb{1}_M\rangle = o(n\lambda^n).$$

The equation above implies

$$\mathbb{E}_x\left[\sum_{k=1}^{n} h(X_k)\mathbb{1}_M(X_n)\right] = n\lambda^n \sum_{\ell=0}^{m-1} e^{\frac{2\pi in\ell}{m}} f_\ell(x)\langle \mu_\ell, hf_\ell\rangle\mu_\ell(M) \ + o(n\lambda^n).$$

This proves the lemma. □

References

[1] M. Benaïm, N. Champagnat, W. Oçafrain, D. Villemonais, Transcritical bifurcation for the conditional distribution of a diffusion process, J. Theoret. Probab. 36 (3) (2023) 1555–1571.
[2] N. Champagnat, D. Villemonais, Exponential convergence to quasi-stationary distribution and Q-process, Probab. Theory Related Fields 164 (1–2) (2016) 243–283.
[3] N. Champagnat, D. Villemonais, General criteria for the study of quasi-stationarity, Electron. J. Probab. 28 (2023) Paper No. 22, 84.
[4] P. Collet, S. Martínez, J. San Martín, Quasi-stationary distributions: Markov chains, diffusions and dynamical systems, Springer Science & Business Media, 2012.
[5] W. Oçafrain, Quasi-stationarity and quasi-ergodicity for discrete-time Markov chains with absorbing boundaries moving periodically, ALEA Lat. Am. J. Probab. Math. Stat. 15 (1) (2018) 429–451.
[6] H. Zmarrou, A.J. Homburg, Bifurcations of stationary measures of random diffeomorphisms, Ergodic Theory Dynam. Systems 27 (5) (2007) 1651–1692.
[7] B. de Pagter, Irreducible compact operators, Math. Z. 192 (1) (1986) 149–153.
[8] J.J. Grobler, Spectral theory in Banach lattices, in: Operator Theory in Function Spaces and Latices, Springer, 1995, pp. 133–172.
[9] P. Meyer-Nieberg, Banach lattices, Springer Science & Business Media, 2012.
[10] M.F. Demers, N. Kiamari, C. Liverani, Transfer operators in hyperbolic dynamics—an introduction, in: 33 º Colóquio Brasileiro de Matemática, Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2021, p. 238.
[11] T. Eisner, B. Farkas, M. Haase, R. Nagel, Operator theoretic aspects of ergodic theory, in: Graduate Texts in Mathematics, vol. 272, Springer Cham, 2015, p. xviii+628.
[12] S.R. Foguel, The ergodic theory of Markov processes, in: Van Nostrand Mathematical Studies, No. 21, Van Nostrand Reinhold Co., New York-Toronto, Ont.-London, 1969, p. v+102.
[13] U. Krengel, Ergodic theorems, in: De Gruyter Studies in Mathematics, vol. 6, Walter de Gruyter & Co., Berlin, 1985, p. viii+357.
[14] P.K. Pollett, Quasi-stationary distributions: a bibliography, Citeseer, 2008, Available at https://people.smp.uq.edu.au/PhilipPollett/papers/qsds/qsds.pdf.
[15] B. Haas, V. Rivero, Quasi-stationary distributions and yaglom limits of self-similar Markov processes, Stochastic Process. Appl. 122 (12) (2012) 4054–4095.
[16] T.G. Kurtz, S. Wainger, The nonexistence of the yaglom limit for an age dependent subcritical branching process, Ann. Probab. (1973) 857–861.
[17] W. Oçafrain, Polynomial rate of convergence to the yaglom limit for Brownian motion with drift, Electron. Commun. Probab. 25 (2020) Paper No. 35, 12.
[18] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, J. San Martín, Quasi-stationary distributions and diffusion models in population dynamics, Ann. Probab. 37 (5) (2009) 1926–1969.
[19] Huasheng Li, H. Zhang, S. Liao, On quasi-stationaries for symmetric Markov processes, J. Math. Anal. Appl. 528 (1) (2023) Paper No. 127498.
[20] J. Zhang, S. Li, R. Song, Quasi-stationarity and quasi-ergodicity of general Markov processes, Sci. China Math. 57 (10) (2014) 2013–2024.
[21] J.N. Darroch, E. Seneta, On quasi-stationary distributions in absorbing discrete-time finite Markov chains, J. Appl. Probab. 2 (1) (1965) 88–100.
[22] L.A. Breyer, G.O. Roberts, A quasi-ergodic theorem for evanescent processes, Stochastic Process. Appl. 84 (2) (1999) 177–186.
[23] F. Colonius, M. Rasmussen, Quasi-ergodic limits for finite absorbing Markov chains, Linear Algebra Appl. 609 (2021) 253–288.
[24] M. Engel, J.S.W. Lamb, M. Rasmussen, Conditioned Lyapunov exponents for random dynamical systems, Trans. Amer. Math. Soc. 372 (9) (2019) 6343–6370.
[25] A. Lambert, Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct, Electron. J. Probab. 12 (14) (2007) 420–446.
[26] P. Mandl, Spectral theory of semi-groups connected with diffusion processes and its application, Czechoslovak Math. J. 11 (4) (1961) 558–569.
[27] R.G. Pinsky, On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes, Ann. Probab. 13 (2) (1985) 363–378.
[28] D. Steinsaltz, S. Evans, Quasistationary distributions for one-dimensional diffusions with killing, Trans. Amer. Math. Soc. 359 (3) (2007) 1285–1324.
[29] E.A. van Doorn, P.K. Pollett, Quasi-stationary distributions for reducible absorbing Markov chains in discrete time, Markov Process. Related Fields 15 (2) (2009) 191–204.
[30] G. Hinrichs, M. Kolb, V. Wachtel, Persistence of one-dimensional AR(1)-sequences, J. Theoret. Probab. 33 (1) (2020) 65–102.
[31] Erik A. van Doorn, Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes, Adv. in Appl. Probab. 23 (4) (1991) 683–700.
[32] M. Kolb, D. Steinsaltz, Quasilimiting behavior for one-dimensional diffusions with killing, Ann. Probab. 40 (1) (2012) 162–212.
[33] P. Del Moral, D. Villemonais, Exponential mixing properties for time inhomogeneous diffusion processes with killing, Bernoulli 24 (2) (2018) 1010–1032.
[34] A. Guillin, B. Nectoux, L. Wu, Quasi-stationary distribution for Hamiltonian dynamics with singular potentials, Probab. Theory Related Fields 185 (3–4) (2023) 921–959.
[35] J.B. Lasserre, C.E.M. Pearce, On the existence of a quasistationary measure for a Markov chain, Ann. Probab. 29 (1) (2001) 437–446.
[36] D. Villemonais, Exponential convergence to a quasi-stationary distribution and applications (Ph.D. thesis), Université de Lorraine (Nancy), 2019.
[37] L.C.G. Rogers, D. Williams, Diffusions, Markov Processes and Martingales, Volume 1: Foundations, vol. 7, John Wiley & Sons, Ltd., Chichester, 1994.

[38] V. Araújo, Attractors and time averages for random maps, Ann. Inst. H. Poincaré C Anal. Non Linéaire 17 (3) (2000) 307–369.
[39] M. Brin, Yu. Kifer, Dynamics of Markov chains and stable manifolds for random diffeomorphisms, Ergodic Theory Dynam. Systems 7 (3) (1987) 351–374.
[40] J.M. Lee, Introduction to smooth manifolds, second ed., in: Graduate Texts in Mathematics, vol. 218, Springer, 2013, p. xvi+708.
[41] J. Milnor, Dynamics in one complex variable, third ed., in: Annals of Mathematics Studies, vol. 160, Princeton University Press, Princeton, NJ, 2006, p. viii+304.
[42] S.P. Meyn, R.L. Tweedie, Markov chains and stochastic stability, Springer Science & Business Media, 2012.
[43] V.L. Shapiro, Book review: Walter rudin, real and complex analysis, Bull. Amer. Math. Soc. 74 (1) (1968) 79–83.
[44] H. Brezis, Functional analysis, Sobolev spaces and partial differential equations, Springer Science & Business Media, 2010.
[45] E. Kreyszig, Introductory functional analysis with applications, in: Wiley Classics Library, John Wiley & Sons, Inc., New York, 1989, p. xvi+688.
[46] A. Lasota, M.C. Mackey, Chaos, fractals, and noise: stochastic aspects of dynamics, vol. 97, Springer Science & Business Media, 2013.