CONDITIONED STOCHASTIC STABILITY OF EQUILIBRIUM STATES ON UNIFORMLY EXPANDING REPELLERS

BERNAT BASSOLS-CORNUDELLA[1], MATHEUS M. CASTRO[1], AND JEROEN S.W. LAMB[1,2,3]

Abstract. We propose a notion of conditioned stochastic stability of invariant measures on repellers: we consider whether quasi-ergodic measures of absorbing Markov processes, generated by random perturbations of the deterministic dynamics and conditioned upon survival in a neighbourhood of a repeller, converge to an invariant measure in the zero-noise limit. Under suitable choices of the random perturbation, we find that equilibrium states on uniformly expanding repellers are conditioned stochastically stable. In the process, we establish a rigorous foundation for the existence of "natural measures", which were proposed by Kantz and Grassberger in 1984 to aid the understanding of chaotic transients.

Contents

[1]Department of Mathematics, Imperial College London, London SW7 2AZ, UK
[2]International Research Center for Neurointelligence, The University of Tokyo, Tokyo 113-0033, Japan
[3]Centre for Applied Mathematics and Bioinformatics, Department of Mathematics and Natural Sciences, Gulf University for Science and Technology, Halwally 32093, Kuwait
E-mail addresses: bernat.bassols-cornudella20@imperial.ac.uk, m.de-castro20@imperial.ac.uk, jeroen.lamb@imperial.ac.uk.

Date: May 3, 2024.
2020 *Mathematics Subject Classification.* 47D08; 60J05; 37D20; 37D35; 37D45.
Key words and phrases. Stochastic stability; absorbing Markov processes; conditioned random dynamics; thermodynamic formalism; transient dynamics.

1. Introduction

Understanding how typical trajectories evolve in a dynamical system and describing its relevant statistics is a central topic in Dynamical Systems theory. This question is commonly addressed from an ergodic theoretical point of view, stating that each (ergodic) invariant measures μ provides the distribution of the trajectory starting at a point x, μ-almost surely. Dynamical systems often admit infinitely many ergodic invariant measures, so it is natural to ask which ones are the most meaningful or relevant to study. To tackle this, Kolmogorov and Sinai, proposed the notion of *stochastic stability* of invariant measures [35, 1].

Stochastic stability concerns the stationary measures of Markov processes generated by small bounded random perturbations of a deterministic dynamical system and their limit as the amplitude of the perturbation vanishes [35, 1]. When a stationary measure converges to an invariant measure of the original deterministic system we say that the limiting measure is stochastically stable. These measures have been recognised to highlight the statistics of (Lebesgue) typical trajectories [60]. Note that stochastically stable invariant measures sit on attractors.

In transient dynamics [39], trajectories that remain for a long time near a repeller have been observed to have well-defined statistics. While there is also an abundance of invariant ergodic measures on repellers, so-called *natural measures* have been heuristically identified as the relevant invariant measures that represent observed long time behaviour of trajectories near a repeller, and provide important information regarding the statistics of transient dynamics [34]. Despite the fact that such measures feature at the heart of the intuitive understanding of transient dynamics, their existence and mathematical properties remain to be rigorously established.

Like stochastic stability successfully provides relevant measures on attractors, we seek a strategy to establish persistence of measures on repellers under random perturbations. The strategy of Kolmogorov and Sinai fails since stationary measures of the Markov process generated by random perturbations of the original system do not converge to invariant measures supported on repellers in the deterministic limit.

In this paper, we propose a novel notion of stochastic stability for repellers referring to *quasi-ergodic measures* rather than stationary measures. Quasi-ergodic measures originate from the theory of absorbing Markov processes and capture the typical average behaviour of trajectories conditioned upon remaining in a certain region of the state space for asymptotically long times. By conditioning the Markov process generated by random bounded perturbations of the original map upon survival in a suitable neighbourhood of the repeller, the associated quasi-ergodic measure provides the conditioned statistics of (Lebesgue) typical trajectories that stay close to the repeller for asymptotically long times. When these quasi-ergodic measures converge to an invariant measure of the deterministic system, we say that the limiting measure is *conditioned stochastically stable*. Note that while stochastically stable invariant measures are supported on attractors, conditioned stochastically stable invariant measures may be supported on repellers.

We show that uniformly expanding repellers admit a unique conditioned stochastically stable invariant measure, which corresponds to the equilibrium state associated with the geometric potential [20, Section 1.2.2] in the framework of thermodynamic formalism [50]. More generally, we establish that any equilibrium state from the thermodynamic formalism on repellers[1] is approximated by quasi-ergodic measures of so-called *weighted Markov processes*, which originate from the theory of Feynman-Kac path distributions (see [30, 18, 16, 36] and references therein), and thus show that equilibrium states are conditioned stochastically stable in a broader sense.

[1]This result also applies to attractors.

1.1. **Conditioned stochastic stability.** The notion of conditioned stochastic stability that we propose is based on ideas from the theory of absorbing Markov processes [22] and conditioned random dynamical systems [62, 31, 14, 12, 13]. As mentioned above, the statistical behaviour of a Markov process X_n on a state space M conditioned upon remaining outside of a subset $\partial \subset M$ is captured by its quasi-ergodic measure ν on $M \setminus \partial$ [26, 10, 61, 23]. This object describes the limiting distribution of the conditioned Birkhoff averages of X_n, i.e. given an observable $h : M \to \mathbb{R}$ it holds that for ν-almost every $x \in M \setminus \partial$,

$$\mathbb{E}_x\left[\frac{1}{n}\sum_{i=0}^{n-1} h\circ X_i \,\middle|\, \tau > n\right] \coloneqq \frac{1}{\mathbb{P}_x\left[\tau > n\right]}\mathbb{E}_x\left[\mathbb{1}_{\{\tau>n\}}\frac{1}{n}\sum_{i=0}^{n-1} h\circ X_i\right] \xrightarrow{n\to\infty} \int h(x)\nu(\mathrm{d}x),$$

where $\tau \coloneqq \min\{n\in \mathbb{N}; X_n \in \partial\}$.

Given a map $T : M \to M$ on a manifold M and a subset $\partial \subset M$, consider the Markov process X_n^ε on M generated by ε-bounded random perturbations of T. Conditioned stochastic stability concerns the quasi-ergodic measures of X_n^ε on $M \setminus \partial$, and their limit as the amplitude of the perturbation ε goes to 0. When these quasi-ergodic measures converge to a T-invariant measure ν_0 (in the weak* topology), we say that the limiting measure is conditioned stochastically stable on $M \setminus \partial$. Observe that this notion depends on the choice of random perturbation generating X_n^ε, which is also true for (classical) stochastic stability. As is common in the study of (classical) stochastic stability, we only consider random bounded diffusive perturbations [7, 4, 2, 6, 1] (see Section 2 for the precise details).

In this paper, we first consider the case where T admits a topologically mixing[2] uniformly expanding repeller R and establish the following result (see Theorem 2.10):

Theorem A1. *There exists a unique T-invariant measure ν_0 on R which is conditioned stochastically stable on every sufficiently small neighbourhood of R.*

It turns out that ν_0 is a well-known object in the thermodynamic formalism theory [20] and corresponds to the unique equilibrium state on R associated with the potential $-\log|\det \mathrm{d}T|$, i.e. ν_0 is the unique T-invariant measure satisfying

$$h_{\nu_0}(T) - \int \log|\det \mathrm{d}T|\,\mathrm{d}\nu_0 = \sup_{\mu\in\mathcal{I}(T,R)}\left(h_\mu(T) - \int \log|\det \mathrm{d}T|\,\mathrm{d}\mu\right),$$

where h_μ is the Kolmogorov-Sinai (or metric) entropy [37, 54] and $\mathcal{I}(T, R)$ is the set of T-invariant probability measures on R. This result has its parallel in the (classical) theory of stochastic stability. Indeed, given a uniformly hyperbolic transformation $T : M \to M$ on a compact metric space M, it is well known that stochastically stable invariant measures on attractors correspond to the equilibrium states from the thermodynamic formalism associated with the potential $-\log|\det \mathrm{d}T\,|_{E^u}|$, where E^u denotes the unstable expanding direction of T [60].

In this paper, we uncover a stronger connection between conditioned stochastic stability and the thermodynamic formalism, establishing the approximation of any equilibrium state by quasi-ergodic measures of *weighted Markov processes*.

1.2. **Thermodynamic formalism and weighted Markov processes.** The thermodynamic formalism is a powerful framework for the analysis of statistical properties of dynamical systems. Pioneered by Sinai, Ruelle and Bowen [55, 8, 9, 49, 50] and motivated by the field of statistical physics, this theory aims to describe properties of equilibrium states, such as the measure of maximal entropy and other invariant Gibbs measures [20, 3].

[2] This condition is relaxed in the main theorem but assumed here for the sake of simplicity.

Given a T-invariant set $\Lambda \subset M$, an equilibrium state on Λ is defined for each given potential $\psi : \Lambda \to \mathbb{R}$ as an invariant measure ν^{ψ} on Λ whose *metric pressure* is equal to the *topological pressure* $P(T, \psi, \Lambda)$ of the system on Λ, i.e. ν^{ψ} satisfies

$$h_{\nu_\psi}(T) + \int \psi \, \mathrm{d}\nu_\psi = \sup_{\mu \in \mathcal{I}(T,\Lambda)} \left(h_\mu(T) + \int \psi \, \mathrm{d}\mu \right) =: P(T, \psi, \Lambda). \tag{1}$$

In particular, observe that when $\psi = 0$ the equilibrium states associated with this potential correspond to the measures of maximal entropy [57, Section 10.5]. Moreover, a classical result of Ruelle (see [51, Lemma 1.4] or Lemma 2.7 below) provides the existence and uniqueness of equilibrium states for Hölder potentials on uniformly expanding repellers [57, Theorem 11.2.15].

It is natural to ask whether the definition of conditioned stochastic stability can be extended to approximate other equilibrium states of T. This question appears not to have been raised in the literature, even for stochastic stability of equilibrium states on attractors. Here, we show that equilibrium states on uniformly expanding repellers are approximated by quasi-ergodic measures of weighted Markov processes [30, 18, 16, 36], providing a general notion of conditioned stochastic stability.

Given a Markov process X_n on M, consider a non-positive *weight* function[3] $\phi : M \to \mathbb{R}_{\leq 0}$ and define the new process X_n^{ϕ} by

$$X_{n+1}^{\phi} = \begin{cases} X_{n+1}, & \text{with probability } e^{\phi(X_n)}, \\ \partial, & \text{with probability } 1 - e^{\phi(X_n)}, \end{cases} \tag{2}$$

where ∂ is a cemetery state. If X_n is already an absorbing Markov process killed at ∂', we may (and do) set $\partial = \partial'$. We refer to the new Markov process X_n^{ϕ} as a *ϕ-weighted Markov process*. As before, a quasi-ergodic measure ν^{ϕ} provides the statistical behaviour of the process when conditioned upon survival, i.e. for any observable $h : M \to \mathbb{R}$,

$$\mathbb{E}_x^{\phi} \left[\frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i^{\phi} \,\middle|\, \tau^{\phi} > n \right] \xrightarrow{n\to\infty} \int h(x) \nu^{\phi}(\mathrm{d}x),$$

for ν^{ϕ}-almost every $x \in M \setminus \partial$, where $\tau^{\phi} := \min\{n \in \mathbb{N};\, X_n^{\phi} \in \partial\}$ and $\mathbb{E}^{\phi}$ is the expectation with respect to the weighted process X_n^{ϕ}.

The random variable τ^{ϕ} denotes the time at which the process is killed, either by dynamically entering ∂ (hard killing) or due to the weight ϕ (soft killing). When both are present, the conditioned Birkhoff averages simplify to (see Section 2 for precise details)

$$\begin{aligned} \mathbb{E}_x^{\phi} \left[\frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i^{\phi} \,\middle|\, \tau^{\phi} > n \right] &:= \frac{1}{\mathbb{E}_x[e^{S_n\phi} \mathbb{1}_{\{\tau > n\}}]} \mathbb{E}_x \left[e^{S_n \phi} \mathbb{1}_{\{\tau > n\}} \frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i \right] \\ &\xrightarrow{n\to\infty} \int h(x) \nu^{\phi}(\mathrm{d}x), \end{aligned} \tag{3}$$

where $\tau = \min\{n;\, X_n \in \partial\}$ relates to hard killing and $S_n\phi := \sum_{i=0}^{n-1} \phi \circ X_i$ relates to soft killing. Note that when $\phi = 0$, we recover the setting introduced in the previous section.

Observe that the right-hand side of equation (3) is also well-defined as long as ϕ is measurable and bounded, even if it is occasionally positive. Indeed,

$$\frac{\mathbb{E}_x \left[e^{S_n\phi} \mathbb{1}_{\{\tau > n\}} \frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i \right]}{\mathbb{E}_x[e^{S_n\phi} \mathbb{1}_{\{\tau > n\}}]} = \frac{\mathbb{E}_x \left[e^{S_n\bar{\phi}} \mathbb{1}_{\{\tau > n\}} \frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i \right]}{\mathbb{E}_x[e^{S_n\bar{\phi}} \mathbb{1}_{\{\tau > n\}}]}$$

[3]The weight function is sometimes referred to as a "potential" in the literature [30, 61]. Here, we only use this term when referring to ψ in equation (1).

where $\bar{\phi} = \phi - \sup \phi_+$, with $\phi_+(x) := \max\{\phi(x), 0\}$. Defining the ϕ-weighted process X_n^ϕ to be equal to $X_n^{\bar{\phi}}$, we recover the interpretation from equation (2).

Recall that X_n^ε is a Markov process on M generated by ε-bounded random perturbations of the map T and absorbed on $\partial \subset M$ and denote by $X_n^{\varepsilon,\phi}$ the ϕ-weighted of X_n^ε. We say that a T-invariant measure ν_0^ϕ is *conditioned ϕ-weighted stochastically stable* if the quasi-ergodic measures ν_ε^ϕ on $M \setminus \partial$ of the weighted process $X_n^{\varepsilon,\phi}$ converge to ν_0^ϕ in the weak* topology as ε goes to 0.

The following theorem generalises Theorem A1 when T has a topologically mixing uniformly expanding repeller R (see Theorem 2.10 for the precise details):

Theorem A2. *Given a Hölder potential ϕ, there exists a unique T-invariant measure ν^ψ on R which is conditioned ϕ-weighted stochastically stable on every sufficiently small neighbourhood of R. Moreover, ν^ψ is the unique equilibrium state associated with the potential $\psi = \phi - \log|\det \mathrm{d}T|$ on R, i.e. $\nu^\psi = \nu_0^\phi$.*

In general, repellers of a given transformation T are not necessarily transitive, let alone topologically mixing. In the case of uniformly expanding repellers, it is natural to assume that the repelling set is characterised by

$$\Lambda = \bigcap_{n\geq 0} T^{-n}(M \setminus \partial), \tag{4}$$

where ∂ could be, for example, a small open neighbourhood of the attractors of T (see Section 5.1), or the complement of a neighbourhood of the repeller (see Section 5.2). In this setting, we prove the following result (see Theorem 2.11 for a more precise and more general result):

Theorem B. *Given a $\mathcal{C}^2$ map T, a Hölder potential ϕ, and a suitable open set $\partial \subset M$, with Λ as in equation (4), assume that*

(1) $T|_\Lambda : \Lambda \to \Lambda$ is uniformly expanding,
(2) $\Lambda \subset \mathrm{Int}(M \setminus \partial)$, and
(3) $T : \Lambda \to \Lambda$ admits a unique equilibrium state ν^ψ associated with the potential $\psi = \phi - \log|\det \mathrm{d}T|$, which is mixing (see e.g. [57, Section 7.1]).

Then ν^ψ is conditioned ϕ-weighted stochastically stable on $M \setminus \partial$, i.e. $\nu^\psi = \nu_0^\phi$.

The proof of Theorems A1 and A2 are based on classical techniques of hyperbolic dynamics. In particular, we adapt the arguments presented in the seminal paper of Pianigiani and Yorke [46] to the context of absorbing Markov processes. To prove Theorem B, we identify a graph structure representing the dynamical behaviour of $X_n^{\varepsilon,\phi}$ conditioned upon staying on $M \setminus U$. This construction resembles the graphs built via chain recurrence and filtration methods [24, 28, 27] and allows us to recover the setting of Theorem A2.

1.3. **Outline.** This paper is organised as follows. In Section 2, we introduce the objects of interest from the theory of conditioned random dynamics. We also lay out the required technical conditions (Hypotheses H1 and H2), explore their direct implications and present the two main theorems (Theorems 2.10 and 2.11). In Section 3, we analyse the local problem (i.e. conditioning the random dynamics on a small neighbourhood of a repeller) and prove Theorem 2.10. In Section 4, we consider the global picture (i.e. conditioning upon not escaping from a general neighbourhood of the repeller) and prove Theorem 2.11. We provide examples in Section 5 where these theorems are applicable. Finally, we devote Appendix A to a general proof for the existence of (weighted) quasi-ergodic measures, simplifying previous techniques.

2. Abstract setup and notation

We begin with a brief recollection of the basic concepts in the theory of conditioned random dynamics as introduced in [14, 13]. Consider a Markov chain X_n evolving in a metric space (E, d) and let $Y \subset E$ be a compact subset. We are interested in studying the behaviour of a Markov chain as it evolves in Y, we condition upon remaining in Y, and kill the process as soon as it leaves this subset. We thus identify $E \setminus Y$ with a "cemetery state" ∂ and consider the space $E_Y := Y \sqcup \partial$ with the induced topology. Throughout this paper, we assume that

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

is a Markov chain with state space E_Y, in the sense of [48, Definition III.1.1]. Hard killing, or absorption, on ∂ means that $\mathbf{P}(\partial, \partial) = 1$. We define the (dynamical) stopping time $\tau := \inf\{n \in \mathbb{N};\ X_n \in \partial\}$.

Consider an α-Hölder weight function $\phi : Y \to \mathbb{R}$ and define the weighted process X_n^ϕ as in Section 1.2. For this process, we define the stopping time $\tau^\phi := \min\{n \in \mathbb{N};\ X_n^\phi \in \partial\}$, providing the time at which X_n^ϕ enters ∂ either dynamically (hard killing) or due to the weight ϕ (soft killing).

Observe that the weighted process X_n^ϕ has transition probabilities given by $\mathbf{P}^\phi(x, \mathrm{d}y) = e^{\bar{\phi}(x)}\mathbf{P}(x, \mathrm{d}y)$ for all $x \in Y$, recall that $\bar{\phi} = \phi - \sup \phi_+$. Moreover, (2) naturally induces a filtered space $(\Omega^\phi, \{\mathcal{F}_n^\phi\}_{n\in\mathbb{N}_0})$ and a family of probability measures $\{\mathbb{P}_x^\phi\}_{x\in E_Y}$ which makes X_n^ϕ a Markov process (see [48, Section III.7] for such a construction). Finally, we denote by $\mathbb{E}_x$ and $\mathbb{E}_x^\phi$ the expectation with respect to $\mathbb{P}_x$ and $\mathbb{P}_x^\phi$, respectively.

Under an irreducibility condition of X_n on Y [14], the process almost surely escapes this set, implying that the system's long-term behaviour is characterised by a stationary delta measure sitting on the cemetery state. To understand the dynamics of the process before escaping from Y one generalises the notion of stationary measures to that of quasi-stationary measures [26, 10, 22, 15].

Definition 2.1. Given a bounded and measurable function $\phi : Y \to \mathbb{R}$, we say that a Borel probability measure μ on Y is a *quasi-stationary measure* of the weighted Markov process X_n^ϕ if

$$\int_Y e^{\phi(y)}\mathbf{P}(y, \mathrm{d}x)\mu(\mathrm{d}x) = \lambda^\phi \mu(\mathrm{d}x)$$

and $\lambda^\phi = \int_Y e^{\phi(x)}\mathbf{P}(x, Y)\mu(\mathrm{d}x) > 0$ is the *growth rate* of μ for X_n^ϕ on Y. Observe that when $\phi = 0$ we recover the classical definition of quasi-stationary measure [21, Definition 2.1].

Remark 2.2. Note that in the usual setting of absorbed Markov processes with no weight function, i.e. $\phi = 0$, and only hard killing, $\lambda^\phi \leq 1$ is called the *survival rate* and denotes the probability that the process is not killed in the next iterate when distributed according to μ.

We recall that quasi-stationary measures are not the relevant measures to consider when studying conditioned Birkhoff averages [26, 10, 17, 14], as these measures do not perceive how likely it is for a point to remain indefinitely in Y. Instead, this information is provided by the so-called quasi-ergodic measure.

Definition 2.3. A probability measure ν on Y is a *quasi-ergodic measure* of the ϕ-weighted Markov process X_n^ϕ if for any bounded measurable function $f : Y \to \mathbb{R}$ it holds that

$$\lim_{n\to\infty} \mathbb{E}_x^\phi\left[\frac{1}{n}\sum_{i=0}^{n-1} f \circ X_i^\phi \,\middle|\, \tau^\phi > n\right] = \int_Y f(y)\nu(\mathrm{d}y) \quad \text{for } \nu\text{-almost every } x \in Y.$$

If X_n^ϕ has both hard and soft killing, then for every $n \in \mathbb{N}$

$$\mathbb{E}_x^\phi\left[\frac{1}{n}\sum_{i=0}^{n-1} h \circ X_i^\phi \,\middle|\, \tau^\phi > n\right] = \frac{1}{\mathbb{E}_x[e^{S_n\phi}\mathbb{1}_{\{\tau>n\}}]}\mathbb{E}_x\left[e^{S_n\phi}\mathbb{1}_{\{\tau>n\}}\frac{1}{n}\sum_{i=0}^{n-1} h \circ X_i\right],$$

where $S_n\phi \coloneqq \sum_{i=0}^{n-1}\phi \circ X_i$ is the Birkhoff sum.

While showing the existence of quasi-stationary measures relates to solving an eigen-functional equation and can be approached using fixed point arguments (see [44, Theorem 4] and [21, Proposition 2.10]), this is not the case for quasi-ergodic measures and proving their existence and uniqueness is not straightforward. Indeed, this involves characterising the non-trivial limit of a conditional expectation that requires rigorous techniques in functional analysis and probability theory [17, 61, 14]. We devote the Appendix A to address this question in our setup.

From here onwards, let $(M, \langle\cdot,\cdot\rangle)$ be an orientable Riemannian compact manifold, possibly with boundary and let $U \subset M$ be an open subset. Without loss of generality, we may assume that M is embedded in an orientable boundaryless compact manifold E of the same dimension and endowed with a Riemannian metric whose restriction to M coincides with $\langle\cdot,\cdot\rangle$ (in the case that M is without boundary we assume that $E = M$). Since this will be clear by context, we may also write the Riemannian metric of E as $\langle\cdot,\cdot\rangle$. The manifold E should be thought of as an ambient space for M and a mere theoretical artefact since it does not play a major role in applications, while U may be interpreted as an open hole in the system.

Notation 2.4. Throughout this paper, we use the following notation:

(i) Given $x \in E$ and $v \in T_xE$, define $\|v\|_x \coloneqq \sqrt{\langle v,v\rangle_x}$ as the natural norm on T_xM.
(ii) We denote by $\mathrm{dist}(\cdot,\cdot)$ the distance on E induced by the Riemannian metric $\langle\cdot,\cdot\rangle$.
(iii) As usual, we write ρ for a Borel measure on E induced by a smooth volume form V_E compatible with $\langle\cdot,\cdot\rangle$.
(iv) We denote by $\mathcal{C}^k(E)$ the space of continuous functions with k continuous derivatives on E and use $\mathcal{M}(E)$ to denote the space of signed Borel finite measures on a E. Given a non-negative measure $\rho \in \mathcal{M}(E)$, we denote by $L^k(E,\rho)$ the space of functions with finite k-th ρ-moment (although ρ may be omitted when it is the reference measure). $\mathcal{C}^k_+(E), L^k_+(E)$ and $\mathcal{M}_+(E)$ denote the respective subsets of non-negative functions and measures on E.
(v) Given a $\mathcal{C}^1$ function $G : E \to E$ and $x \in E$, we denote its determinant by

$$\det \mathrm{d}G(x) = \frac{V_E(G(x))(\mathrm{d}G(x)v_1,\ldots,\mathrm{d}G(x)v_{\dim E})}{V_E(x)(v_1,\ldots,v_{\dim E})},$$

for any (and therefore all) orthonormal basis $\{v_1,\ldots,v_m\}$ of T_xM.
(vi) Given a set $A \subset E$ we denote its closed neighbourhood of radius $\delta > 0$ by $A_\delta = \overline{B_\delta(A)} \coloneqq \{x \in E;\ \mathrm{dist}(x,a) \le \delta \text{ for some } a \in A\}$.

2.1. **Main results.** Let $T : E \to E$ be a map such that $T|_{E\setminus U}$ is $\mathcal{C}^2$, and let $\phi : E \to \mathbb{R}$ be α-Hölder. The following two hypotheses contain the main assumptions in this paper:

Hypothesis H1. *There exists a compact T-invariant set $\Lambda \subset E$ that is uniformly hyperbolic expanding, i.e. there exist $C, r > 0$ such that for all $x \in \Lambda$,*

$$\|\mathrm{d}T^n(x)^{-1}\| < C\frac{1}{(1+r)^n} \quad \textit{for every } n \ge 1,$$

and there exists a neighbourhood V of Λ in E such that $T^{-1}(\Lambda) \cap V = \Lambda$.

Remark 2.5. Observe that the manifold M does not play a role in Hypothesis H1. Moreover, we use the term *uniformly hyperbolic expanding* to refer to a set that is *eventually* uniformly expanding, i.e. the equation above is equivalent to the following statement: there exist $N \in \mathbb{N}$ and $r > 0$ such that for all $x \in \Lambda$,

$$\|\mathrm{d}T^n(x)^{-1}\| < \frac{1}{(1+r)^n} \quad \text{for every } n \geq N.$$

Hypothesis H2. *We say that T satisfies Hypothesis H2 for the open hole U if*

(1) $\Lambda := \bigcap_{n\geq 0} T^{-n}(M \setminus U)$ is a uniformly hyperbolic expanding set, and
(2) T admits a unique equilibrium state for the potential $\phi - \log|\det \mathrm{d}T|$ on Λ.

In the case that $\dim(M) > 1$, we additionally assume that there exists $\delta > 0$ such that $T^{-1}(\Lambda_\delta) \cap M_\delta \subset \Lambda_\delta$ and $M_\delta \setminus M$ has no T-invariant subsets.

Remark 2.6. If $M = [0,1]$, items (1) and (2) of Hypothesis H2 are equivalent to the Axiom *A* (see [29, Chapter 3.2.b]). In the case that M is a manifold without boundary, then $M_\delta = M = E$ and $\Lambda_\delta \subset M$. Moreover, observe that Hypothesis H2 implies Hypothesis H1.

Lemma 2.7. *Let T satisfy Hypothesis H1 and $R := \overline{\{p \in \Lambda;\ p \text{ is a } T\text{-periodic point}\}}$. Then, there exists a (finite) partition of R in non-empty compact sets $R^{i,j}$, with $1 \leq i \leq k$ and $1 \leq j \leq m(i)$, such that*

(1) $R^i = \cup_{j=1}^{m(i)} R^{i,j}$ is a T-invariant set for every i,
(2) $T(R^{i,j}) = R^{i,j+1 \pmod{m(i)}}$ for every i,j,
(3) $T : R^i \to R^i$ is uniformly hyperbolic and topologically transitive, and
(4) each $T^{m(i)} : R^{i,j} \to R^{i,j}$ is uniformly hyperbolic and topologically exact.

Furthermore, the number k, the numbers $m(i)$ and the sets $R^{i,j}$ are unique up to renumbering.

Proof. This result follows directly for uniformly expanding maps, i.e. $C = 1$ in Hypothesis H1 (see e.g. [57, Theorem 11.2.15]), so the proof is concluded by changing the Riemannian metric on M in a way that T becomes uniformly expanding on Λ under Hypothesis H1 (see [53, Proposition 4.2]). $\square$

When T satisfies Hypothesis H1, given a finite $\varepsilon > 0$ we consider the random perturbation of the form $F_\varepsilon : [-\varepsilon, \varepsilon]^m \times E \to E$, where $F_\varepsilon(\omega, \cdot) \in \mathcal{C}^2(E \setminus U; E)$ and $\partial_\omega F_\varepsilon(\omega, x)$ is surjective for all $\omega \in [-\varepsilon, \varepsilon]^m$. Moreover, we assume that $\mathrm{dist}_{\mathcal{C}^2}(F_\varepsilon(\omega, \cdot), T) \leq C\|\omega\|$ for some $C > 0$, where $\mathrm{dist}_{\mathcal{C}^2}$ denotes the metric on $\mathcal{C}^2(E \setminus U, E)$ which generates the $\mathcal{C}^2$-Whitney topology [45, Chapter 1.2]. In particular, surjectivity of $\partial_\omega F_\varepsilon(\omega, x)$ implies $m \geq \dim E$. We note that this type of random perturbation is natural and commonly considered [7, 4, 2, 6, 1].

Notation 2.8. Let $\Omega_\varepsilon := ([-\varepsilon, \varepsilon]^m)^{\mathbb{N}}$ be the space of semi-infinite sequences of elements in $[-\varepsilon, \varepsilon]^m$ endowed with the probability measure $\mathbb{P}_\varepsilon := (\mathrm{Leb}|_{[-\varepsilon,\varepsilon]^m} / (2\varepsilon)^m)^{\otimes \mathbb{N}}$, and let $\mathbb{E}_\varepsilon$ denote the corresponding expectation with respect to $\mathbb{P}_\varepsilon$. For every $\omega \in \Omega_\varepsilon$, $\omega = (\omega_0 \omega_1 \ldots)$, we define $T_\omega(x) := T_{\omega_0}(x) := F_\varepsilon(\omega_0, x)$ and $T^n_\omega(x) := T_{\omega_{n-1}} \circ \cdots \circ T_{\omega_n}(x)$ for every $n \in \mathbb{N}$.

For each $i \in \{1, \ldots, k\}$ and every α-Hölder function $\phi : R^i_\delta \to \mathbb{R}$ (see Notation 2.4 item (vi)), we define the annealed Koopman operator

$$\mathcal{P}_\varepsilon : f \mapsto e^{\phi(x)} \mathbb{E}_\varepsilon[f \circ T_\omega(x) \cdot \mathbb{1}_{R^i_\delta} \circ T_\omega(x)],$$

for f in a suitable domain.

Theorems A1 and A2 above apply to suitable δ-neighbourhoods of each repeller R^i, $1 \leq i \leq k$, in the dynamical decomposition of Lemma 2.7. In particular, we choose δ and ε_0 such that the following holds true.

Lemma 2.9. *Assume Hypothesis H1, for every $\delta > 0$ small enough there exists $\varepsilon_0 := \varepsilon_0(\delta) > 0$ such that for every $0 \leq \varepsilon \leq \varepsilon_0$ we have that:*

(1) $R_\delta = R_\delta^1 \sqcup \ldots \sqcup R_\delta^k$, *and*
(2) $\sup_{x\in R_\delta^i} \mathbb{P}_\varepsilon[\omega \in \Omega_\varepsilon;\ T_\omega(x) \in R_\delta^j] = 0$ *for every* $i \neq j \in \{1, \ldots, k\}$.

The main results of this paper are as follows:

Theorem 2.10. *Assume Hypothesis H1 and let* $\delta > 0$ *be small enough. Given* $1 \leq i \leq k$ *and an* α*-Hölder function* $\phi : R_\delta^i \to \mathbb{R}$*, the following properties hold for* $\varepsilon > 0$ *sufficiently small:*

(1) *the* ϕ*-weighted Markov process* $X_n^{\varepsilon,\phi}$ *admits a unique quasi-stationary measure* μ_ε *on* R_δ^i*,*
(2) *let* λ_ε *be the growth rate of* $X_n^{\varepsilon,\phi}$ *on* R_δ^i*, then* λ_ε *is equal to the spectral radius of* $\mathcal{P}_\varepsilon : L^\infty(R_\delta^i, \rho) \to L^\infty(R_\delta^i, \rho)$*, and* $\log(\lambda_\varepsilon) \to P(T, \phi - \log|\det \mathrm{d}T|, R^i)$ *as* $\varepsilon \to 0$*,*
(3) *there exists a unique positive eigenfunction* $g_\varepsilon \in L^\infty(R_\delta^i, \rho)$ *for the operator* $\mathcal{P}_\varepsilon : L^\infty(R_\delta^i, \rho) \to L^\infty(R_\delta^i, \rho)$*, associated with the eigenvalue* λ_ε*,*
(4) *let* $\nu_\varepsilon(\mathrm{d}x)$ *be the unique quasi-ergodic measure of the* ϕ*-weighted process* $X_n^{\varepsilon,\phi}$ *on* $\{g_\varepsilon > 0\}$*. Then,* $\nu_\varepsilon(\mathrm{d}x) \to \nu_0(\mathrm{d}x)$ *in the weak* topology as* $\varepsilon \to 0$*, and*
(5) ν_0 *is the unique* T*-invariant equilibrium state for the potential* $\phi - \log|\det \mathrm{d}T|$ *on* R^i*.*

If the measure ν_0 *is mixing for the map* $T : R^i \to R^i$*, then the measure* ν_ε *is also a quasi-ergodic measure of the* ϕ*-weighted Markov process* X_n^ε *on* R_δ^i*.*

Theorem 2.11. *Assume Hypothesis H2 and let* $\delta > 0$ *be small enough. Given an* α*-Hölder function* $\phi : M_\delta \setminus U \to \mathbb{R}$*, the following properties hold for* $\varepsilon > 0$ *sufficiently small:*

(1) *the* ϕ*-weighted Markov process* $X_n^{\varepsilon,\phi}$ *admits a unique quasi-stationary measure* μ_ε *on* $M_\delta \setminus U$*,*
(2) *let* λ_ε *be growth rate of* $X_n^{\varepsilon,\phi}$ *on* $M_\delta \setminus U$*, then* λ_ε *is equal the spectral radius of* $\mathcal{P}_\varepsilon : L^\infty(M_\delta \setminus U) \to L^\infty(M_\delta \setminus U)$*, and* $\log(\lambda_\varepsilon) \to P(T, \phi - \log|\det \mathrm{d}T|, \Lambda)$ *as* $\varepsilon \to 0$*,*
(3) *there exists a unique positive eigenfunction* $g_\varepsilon \in L^\infty(M_\delta \setminus U, \rho)$ *for the operator* $\mathcal{P}_\varepsilon : L^\infty(M_\delta \setminus U, \rho) \to L^\infty(M_\delta \setminus U, \rho)$*, associated with the eigenvalue* λ_ε*,*
(4) *let* $\nu_\varepsilon(\mathrm{d}x)$ *be the unique quasi-ergodic measure of the* ϕ*-weighted Markov process* $X_n^{\varepsilon,\phi}$ *on* $\{g_\varepsilon > 0\} \cap \operatorname{supp} \mu_\varepsilon$*. Then,* $\nu_\varepsilon(\mathrm{d}x) \to \nu_0(\mathrm{d}x)$ *in the weak* topology as* $\varepsilon \to 0$*, and*
(5) ν_0 *is the unique* T*-invariant equilibrium state for the potential* $\phi - \log|\det \mathrm{d}T|$ *on* Λ*.*

If ν_0 *is mixing for the map* $T : R \to R$*, then the conclusions of the above theorem remain true when changing the set* $\{g_\varepsilon > 0\} \cap \operatorname{supp} \mu_\varepsilon$ *by* $M_\delta \setminus U$*. Additionally, if* ν_0 *is mixing and* $\operatorname{supp} \nu_0 \subset \operatorname{Int}(M \setminus U)$*, then (4) is also true on the set* $M \setminus U$*.*

2.2. **Some direct consequences of Hypothesis H1.** This section contains several dynamical and topological results that follow from Hypothesis H1 rather immediately and are exploited later in the paper. We also introduce the equilibrium states we shall approximate and present the transfer operators $\mathcal{P}_\varepsilon$ and $\mathcal{L}_\varepsilon$ that reappear throughout the text.

Lemma 2.12. *Let* T *satisfy Hypothesis H1 and* $\|\mathrm{d}T(x)^{-1}\| < 1/(1+r)$ *for every* $x \in \Lambda$*. Consider* $\delta > 0$ *small enough. Then there exists* $\varepsilon_0 \coloneqq \varepsilon_0(\delta)$ *and* $\sigma_1 \coloneqq \sigma_1(\delta) < 1$ *such that for every* $x, y \in \Lambda_\delta$ *satisfying* $T(y) = x$ *and for every* $0 < \varepsilon < \varepsilon_0$*, there exists a* $\mathcal{C}^2$ *function* $h : [-\varepsilon, \varepsilon]^m \times C_x \to E$*, where* C_x *is the connected component of* x *in* Λ_δ*, with the following properties holding for every* $\omega \in \Omega_\varepsilon$*:*

(1) *the map* $z \mapsto h(\omega, z)$ *is a diffeomorphism onto its image,*
(2) $T_\omega \circ h(\omega, z) = z$ *for every* $z \in C_x$ *and* $h(0, x) = y$*,*

(3) $\mathrm{dist}(h(\omega,x_1),h(\omega,x_2)) \le \sigma_1 \mathrm{dist}(x_1,x_2)$ *for every* $x_1,x_2 \in C_x$*, and*

(4) there exists $K_0 = K_0(\delta) > 0$ *uniform on* $\varepsilon \in (0,\varepsilon_0)$, $x \in \Lambda_\delta$ *and* $y \in T^{-1}(x)\cap\Lambda_\delta$*, such that* $\sup\{\|\partial_\omega h(\omega,z)\|;\ \omega\in\Omega_\varepsilon,\ z\in C_x\} \le K_0$.

All statements in this lemma also hold true replacing Λ_δ *by* R^i_δ, $1 \le i \le k$.

Proof. Take $\delta_0, \varepsilon_0 > 0$ small enough such that

$$\sigma_1 \coloneqq \sup\left\{\|\mathrm{d}T_\omega(x)^{-1}\|;\ x\in\Lambda^i_{\delta_0},\ \omega\in\Omega_{\varepsilon_0}\right\} < 1, \tag{5}$$

and such that the exponential map $\exp_z : B_{\delta_0}(0) \subset T_zE \to E$ is well defined for every $z \in \Lambda_{\delta_0}$.

Observe that there exists $\delta_2 > 0$ such that for $\varepsilon_0 > 0$ small enough and for every $\omega \in [-\varepsilon_0,\varepsilon_0]^m$, if $\mathrm{dist}(x_1,x_2) < \delta_2$ then we obtain that $\mathrm{dist}(T_\omega(x_1),T(x_2)) < \delta_0$. Consider the map

$$G = G_{x,y} : [-\varepsilon,\varepsilon]^m \times B_\delta(x) \times B_{\delta_2}(y) \to T_xE$$
$$(\omega,z_1,z_2) \mapsto \exp_x^{-1}(z_1) - \exp_x^{-1}(T_\omega(z_2)).$$

Observe that $G(0,x,y) = 0$. Since $\partial_y G(0,x,y)$ is surjective, by means of the implicit function theorem, there exists a $\mathcal{C}^2$ function $h : [-\varepsilon_0(y),\varepsilon_0(y)]^m \times B_{\tau(y)}(x) \to E$ such that $T_\omega(h(\omega,z)) = z$ for every $z \in B_{\tau(y)}(x)$ and $h(0,x) = y$. Notice, as well, that $\varepsilon_0(y), \tau(y)$ can be taken uniformly since Λ_δ is compact and therefore we can $\mathcal{C}^2$-extend h to the domain $[-\varepsilon,\varepsilon]^m \times C_x$. Finally, from (5) we obtain that the function h satisfies all the desirable properties. Replacing Λ_δ by R^i_δ follows from Lemma 2.9 item (2). □

Lemma 2.13. *Assume Hypothesis H1 and* $\|\mathrm{d}T(x)^{-1}\| < 1/(1+r)$ *for every* $x \in R$*. There exists* $\delta_0 > 0$ *small enough satisfying Lemma 2.9 such that*

(1) there exists $\sigma_0 \coloneqq \sigma_0(\delta_0) \in (0,1)$ *such that* $T^{-1}(\Lambda_\delta)\cap\Lambda_\delta \subset \Lambda_{\sigma_0\delta}$ *for all* $0 < \delta < \delta_0$.

Moreover, there exists $\varepsilon_0 \coloneqq \varepsilon_0(\delta)$ *satisfying Lemma 2.12 such that for every* $0 < \varepsilon < \varepsilon_0$ *we have that:*

(2) there exists $\sigma \coloneqq \sigma(\delta,\varepsilon) \in (0,1)$*, such that* $T_\omega^{-1}(\Lambda_\delta)\cap\Lambda_\delta \subset \Lambda_{\sigma\delta}$*, for every* $\omega \in \Omega_\varepsilon$*, and*

(3) for all x, y *lying in the same connected component of* Λ_δ *and* $\omega \in \Omega_\varepsilon$ *we have* $\#\{T^{-1}(x)\cap\Lambda_\delta\} = \#\{T_\omega^{-1}(y)\cap\Lambda_\delta\}$.

All statements in this lemma also hold true replacing Λ_δ *by* R^i_δ, $1 \le i \le k$.

Proof. We prove (1). First of all, take δ_0 and ε_0 small enough such that $\|\mathrm{d}T_{\omega_0}(x)^{-1}|_{\Lambda_{\delta_0}}\| < 1$ for all $\omega_0 \in [-\varepsilon_0,\varepsilon_0]^m$ and $\Lambda_{2\delta_0} \subset V$, where V is as in Hypothesis H1. Let $0 < \delta < \delta_0$ and $0 < \varepsilon < \varepsilon_0$.

Given $x \in E$ and $v \in T_xE$ such that $\|v\|_x = 1$, let $\gamma_{x,v} : (-\delta_0,\delta_0) \to E$ be a geodesic on E such that $\gamma_{x,v}(0) = x$ and $\gamma'_{x,v}(0) = v \in T_xE$. From Hypothesis H1, the fact that T is a $\mathcal{C}^2$ function and Λ is compact, we have that $\gamma_{y,w}(\delta)$ is well defined for every $y \in \Lambda$ and $w \in T_yE$, and that there exists $r_0 > 0$ such that

$$\mathrm{dist}(T\circ\gamma_{y,w}(t), T(y)) \ge |t|(1+r_0) \quad \text{for every } |t| \le \delta_0. \tag{6}$$

From the above equation and the fact that $\mathrm{d}T(x)$ is a surjective linear operator, we obtain that $T(B_{\delta/(1+r_0)}(y)) \supset B_\delta(T(y))$. Take $y \in \Lambda_\delta$, then there exits $x \in \Lambda$, $v \in T_xE$ and $h \in [-\delta,\delta]$ such that $y = \gamma_{x,v}(h)$. Let $x_1,\dots,x_\ell \in \Lambda$ be all pre-images of x. From (6) there exist $h_1,\dots,h_\ell \in [-\delta/(1+r_0),\delta/(1+r_0)]$, and unit vectors $v_1 \in T_{x_1}E,\dots,v_\ell \in T_{x_\ell}E$ such that all $y_i = \gamma_{x_i,v_i}(h_i), 1 \le i \le \ell$, are pre-images of y. Note that $y_i \in \Lambda_{\delta/(1+r_0)}$. We claim that these are precisely the only pre-images of y in Λ_δ. Suppose there exists $y' \in \Lambda_\delta \setminus \Lambda_{\delta/(1+r_0)}$ such that $T(y') = y$. Since $x \in B_\delta(y)$, from (6) there is $h' \in [-\delta/(1+r_0),\delta/(1+r_0)]$ and $v' \in T_xE$ such that $T(\gamma_{y',v'}(h')) = x$. This contradicts Hypothesis H1 as $\gamma_{y,v}(h') \in \Lambda_{2\delta}\setminus\Lambda \subset V\setminus\Lambda$. Therefore, $T^{-1}(\Lambda_\delta)\cap\Lambda_\delta \subset \Lambda_{\delta/(1+r_0)}$. Set $\sigma_0 \coloneqq 1/(1+r_0)$.

We prove (2). For every $y \in \Lambda_\delta$, let $\{y_1, \dots, y_\ell\} := T^{-1}(y) \cap \Lambda_\delta$. Let $h_1, \dots, h_\ell : [-\varepsilon, \varepsilon]^m \times C_y \to E$ be the inverse branch functions defined in Lemma 2.12, such that $h_i(0, y) = y_i$. Since $\mathrm{dist}(h_i(\omega, y), y_i) = \mathrm{dist}(h_i(\omega, y), h_i(0, y)) < K_0\|\omega\| \le K_0\varepsilon$, and $y_i \in \Lambda_{\delta/(1+r_0)}$ from item (1), we obtain that $h_i(\omega, y) \in \Lambda_{\delta(K_0\varepsilon+1/(1+r_0))}$. Choosing ε small enough, there exists $\sigma \in (0, 1)$ for which $h_i(\omega, y) \in \Lambda_{\sigma\delta}$ for all $\omega \in \Omega_\varepsilon$.

To finish the proof, we show that for $\varepsilon > 0$ small enough, $\#\{T^{-1}(x) \cap \Lambda_\delta\} = \#\{T_\omega^{-1}(x) \cap \Lambda_\delta\}$, for every $x \in \Lambda_\delta$ and $\omega \in \Omega_\varepsilon$. From the construction above, we obtain that $\#\{T^{-1}(x) \cap \Lambda_\delta\} \le \#\{T_\omega^{-1}(x) \cap \Lambda_\delta\}$. Suppose by contradiction that there exist sequences $\{x_n\}_{n\in\mathbb{N}} \subset \Lambda_\delta$ and $\{\omega_n\}_{n\in\mathbb{N}} \subset [-\varepsilon_0, \varepsilon_0]^m$, such that $\#\{T^{-1}(x_n) \cap \Lambda_\delta\} < \#\{T_{\omega_n}^{-1}(x_n) \cap \Lambda_\delta\}$ and $\omega_n \to 0$. From the compactness of Λ_δ and the pigeonhole principle, the above assumption implies that there exist sequences $\{y_n^1\}_{n\in\mathbb{N}}$ and $\{y_n^2\}_{n\in\mathbb{N}}$ such that: (a) $y_n^1 \neq y_n^2$ and $T_{\omega_n}(y_n^1) = T_{\omega_n}(y_n^2)$ for every $n \in \mathbb{N}$; and (b) $y_n^1, y_n^2 \xrightarrow{n\to\infty} y^* \in \Lambda_\delta$. From the continuity of $(\omega, x) \mapsto T_\omega(x)$, we obtain that $\mathrm{dist}(T(y_n^1), T(y_n^2)) \xrightarrow{n\to\infty} 0$, which contradicts the fact that $\mathrm{d}T(y^*)$ is invertible and completes the proof.

We prove (3). From the last part in the proof of item (2) we obtain that $\#\{T^{-1}(x) \cap \Lambda_\delta\} = \#\{T_\omega^{-1}(x) \cap \Lambda_\delta\}$, for every $x \in \Lambda_\delta$ and $\omega \in \Omega_\varepsilon$ for ε sufficiently small. Therefore, it is sufficient to show that the map $x \in \Lambda_\delta \mapsto \#\{T^{-1}(x) \cap \Lambda_\delta\}$ is locally constant. This is a direct consequence of $\|\mathrm{d}T^{-1}(x)\| < 1$ for all $x \in \Lambda_\delta$ and the inverse function theorem (see e.g. the proof of [57, Lemma 11.1.4]).

The last statement follows from replacing Λ_δ by R^i_δ in every argument above, and from item (2) in Lemma 2.9. Note that for (1), we have that R^i is open in $T^{-1}(R^i)$ (see the proof of [57, Corollary 11.2.16]). ☐

As mentioned in the introduction, stochastic stability has been previously studied in the context of trajectories accumulating on attractors. Instead, in this paper, we are interested in characterising the stochastic stability of general equilibrium states on uniformly hyperbolic expanding repellers, for which no such notion exists in the literature. The existence and uniqueness of equilibrium states on uniformly hyperbolic expanding repellers is guaranteed by the following classical result of Ruelle (see [51, Lemma 1.4] or [50, Chapters 7.26-7.31]).

Theorem 2.14 (Ruelle). *Let T satisfy Hypothesis H1. Let $R^1, \dots, R^k$ be as in Lemma 2.7 and fix $i \in \{1, \dots, k\}$. For every α-Hölder potential $\phi : R^i \to \mathbb{R}$ consider the operator*

$$\mathcal{L} : \mathcal{C}^0(R^i) \to \mathcal{C}^0(R^i)$$
$$f \mapsto \sum_{T(y)=x} e^{\phi(y)} f(y).$$

Then, there exist unique $m \in \mathcal{C}^0_+(R^i)$, $\gamma \in \mathcal{M}_+(R^i)$ and $\lambda > 0$ satisfying

- $\ker(\mathcal{L} - \lambda) = \mathrm{span}(m)$,
- $\ker(\mathcal{L}^* - \lambda) = \mathrm{span}(\gamma)$ *and* $\int_{R^i} m(x)\gamma(\mathrm{d}x) = 1$, *and*
- $\log \lambda = \log r(\mathcal{L}) = h_\nu + \int \phi(x)\nu(\mathrm{d}x)$, *where* $\nu(\mathrm{d}x) = m(x)\mu(\mathrm{d}x)$.

In this context, ν is the (unique) T-invariant equilibrium state for the potential ϕ on R^i.

To approximate these equilibrium states, we propose using quasi-ergodic measures, which we construct from the principal eigenfunctions of the following annealed transfer operators.

Notation 2.15. For each $i \in \{1, \dots, k\}$ and every α-Hölder function $\phi : R^i_\delta \to \mathbb{R}$, we define the annealed Ruelle-Perron-Frobenius operator

$$\mathcal{L}_\varepsilon : L^1(R^i_\delta, \rho) \to L^1(R^i_\delta, \rho)$$
$$f \mapsto \mathbb{E}_\varepsilon\left[\sum_{T_\omega(y)=x} \frac{e^{\phi(y)} f(y) \mathbb{1}_{R^i_\delta}(y)}{|\det \mathrm{d}T_\omega(y)|}\right]$$

and the annealed Koopman operator

$$\begin{aligned}\mathcal{P}_\varepsilon : L^\infty(R^i_\delta,\rho) &\to L^\infty(R^i_\delta,\rho)\\ f &\mapsto e^{\phi(x)}\mathbb{E}_\varepsilon[f\circ T_\omega(x)\cdot \mathbb{1}_{R^i_\delta}\circ T_\omega(x)],\end{aligned}$$

which are well-posed from Lemmas 2.9, 2.12 and 2.13. Moreover, given $x \in R^i_\delta$ and $n \in \mathbb{N}$ we refer to the measure $\mathcal{P}^n_\varepsilon(x,\cdot)$ as the unique measure on R^i_δ such that $\mathcal{P}^n_\varepsilon(x,A) = \mathcal{P}^n_\varepsilon \mathbb{1}_A(x)$ for every measurable subset A of R^i_δ.

Remark 2.16. Note that the Ruelle-Perron-Frobenius operator $\mathcal{L}$ introduced in Theorem 2.14 differs from the operator $\mathcal{L}_\varepsilon$ since the latter is divided by $|\det \mathrm{d}T_\omega|$. This causes the correction $-\log|\det \mathrm{d}T|$ for the limiting potential in Theorems 2.10 and 2.11. This choice provides a more interpretable expression for $\mathcal{P}_\varepsilon$ and its eigenfunctions as quasi-stationary measures for the ϕ-weighted Markov process $X^{\varepsilon,\phi}_n$.

The following proposition establishes that $X^{\varepsilon,\phi}_n$ is a strong Feller absorbing Markov process. This is constantly exploited throughout the paper.

Proposition 2.17. *For every α-Hölder function $\phi : R^i_\delta \to \mathbb{R}$, the operator $\mathcal{P}_\varepsilon : L^\infty(R^i_\delta,\rho) \to L^\infty(R^i_\delta,\rho)$ is strong Feller, i.e. given a bounded measurable function $h : M \to \mathbb{R}$ we have $\mathcal{P}_\varepsilon h \in \mathcal{C}^0(M)$. In particular, $\mathcal{P}^2_\varepsilon$ is a compact operator.*

Proof. Let $\{x_n\}_{n\in\mathbb{N}} \subset R^i_\delta$ be a sequence converging to $x \in R^i_\delta$. Write $[-\varepsilon,\varepsilon]^m = [-\varepsilon,\varepsilon]^e \times [-\varepsilon,\varepsilon]^{m-e}$, where $e = \dim E$, and let $F := F_\varepsilon : [-\varepsilon,\varepsilon]^e \times [-\varepsilon,\varepsilon]^{m-e} \times R^i_\delta \to E$. Assume without loss of generality that $\partial_{\omega_0}F(\omega_0,\omega_1,x_n)$ is surjective for every $n \in \mathbb{N}$. Let $F^{-1}_{(\omega_1,x)}$ denote the inverse of F for fixed (ω_1,x). Then, for any bounded and measurable function $h : M \to \mathbb{R}$ we obtain that

$$\begin{aligned}\mathcal{P}_\varepsilon h(x_n) &= \frac{e^{\phi(x_n)}}{(2\varepsilon)^m}\int_{[-\varepsilon,\varepsilon]^{m-e}\times[-\varepsilon,\varepsilon]^e} (\mathbb{1}_{R^i_\delta}h)\circ F(\omega_0,\omega_1,x_n)\mathrm{d}\omega_0\mathrm{d}\omega_1\\ &= \frac{e^{\phi(x_n)}}{(2\varepsilon)^m}\int_{[-\varepsilon,\varepsilon]^{m-e}}\int_{F([-\varepsilon,\varepsilon]^e,\omega_1,x_n)\cap R^i_\delta} h(y)\left|\det \mathrm{d}F^{-1}_{(\omega_1,x)}(y)\right|\rho(\mathrm{d}y)\mathrm{d}\omega_1\\ &= \int_{R^i_\delta}\left[\frac{e^{\phi(x_n)}}{(2\varepsilon)^m}\int_{[-\varepsilon,\varepsilon]^{m-e}} \mathbb{1}_{F([-\varepsilon,\varepsilon]^e,\omega_1,x_n)}(y)\left|\det \mathrm{d}F^{-1}_{(\omega_1,x_n)}(y)\right|\mathrm{d}\omega_1\right]h(y)\rho(\mathrm{d}y).\end{aligned}$$

Defining κ as

$$\kappa(x_n,y) := \frac{e^{\phi(x_n)}}{(2\varepsilon)^m}\int_{[-\varepsilon,\varepsilon]^{m-e}} \mathbb{1}_{F([-\varepsilon,\varepsilon]^e,\omega_1,x_n)}(y)\left|\det \mathrm{d}F^{-1}_{(\omega_1,x_n)}(y)\right|\mathrm{d}\omega_1,$$

it is clear that $\kappa(x_n,y) \xrightarrow{n\to\infty} \kappa(x,y)$ for ρ-a.e. $y \in R^i_\delta$. We obtain from the Lebesgue dominated convergence theorem that $\lim_{n\to\infty}\mathcal{P}_\varepsilon h(x_n) = \mathcal{P}_\varepsilon h(x)$, so $\mathcal{P}_\varepsilon$ is strong Feller. From [47, Chapter 1, Theorem 5.11] (which we recall in Lemma A.2), we have that $\mathcal{P}^2_\varepsilon$ is a compact operator. ∎

Observe that given an α-Hölder-potential $\phi : R^i_\delta \to \mathbb{R}$, then $\mathcal{L}^*_\varepsilon = \mathcal{P}_\varepsilon$. Indeed, for any $f \in L^1(R^i_\delta)$ and $g \in L^\infty(R^i_\delta)$,

$$\begin{aligned}\int_{R^i_\delta} f(x)\mathcal{P}_\varepsilon g(x)\rho(\mathrm{d}x) &= \mathbb{E}_\varepsilon\left[\int_{R^i_\delta} e^{\phi(x)}f(x)g\circ T_\omega(x)\mathbb{1}_{R^i_\delta}\circ T_\omega(x)\rho(\mathrm{d}x)\right]\\ &= \int_{R^i_\delta}\mathbb{E}_\varepsilon\left[\sum_{T_\omega(y)=x}\frac{e^{\phi(y)}f(y)\mathbb{1}_{R^i_\delta}(y)}{|\det \mathrm{d}T_\omega(y)|}\right]g(x)\rho(\mathrm{d}x) \qquad (7)\\ &= \int_{R^i_\delta}\mathcal{L}_\varepsilon f(x)g(x)\rho(\mathrm{d}x),\end{aligned}$$

where (if needed) we assume that $f : E \to \mathbb{R}$ vanishes outside of R^i_δ. From Proposition 2.17 and equation (7), we obtain that $\mathcal{L}^2_\varepsilon : L^1(R^i_\delta) \to L^1(R^i_\delta)$ is also a compact operator.

3. The local problem

In this section, we focus on a single repeller R^i of T from the dynamical decomposition of Lemma 2.7 and establish the stochastic stability of equilibrium states associated with the restricted transformation $T|_{R^i}$. To achieve this, we condition the process X^ε_n upon remaining within a δ-neighbourhood of the repeller R^i. For a given α-Hölder potential ϕ, we begin by showing that there exists a unique quasi-stationary measure μ_ε for the ϕ-weighted Markov process $X^{\varepsilon,\phi}_n$ on R^i_δ absorbed in $\partial := M_\delta \setminus R^i_\delta$. To do so, we adapt the analysis of conditionally invariant probability measures provided by Pianigiani and Yorke [46] as fixed points of the (normalised) Ruelle-Perron-Frobenius operator, $\widehat{\mathcal{L}_\varepsilon}$. We continue with a detailed study of the operator $\mathcal{P}_\varepsilon$ to obtain the (unique) eigenfunction g_ε of maximal eigenvalue λ_ε. Finally, we prove the existence and uniqueness of a quasi-ergodic measure of the ϕ-weighted Markov process X^ε_n conditioned upon not escaping the support of g_ε, and characterise its limiting behaviour as the noise strength ε vanishes. This measure follows from the pointwise product of μ_ε and g_ε. As previously mentioned, we show that the limiting object as $\varepsilon \to 0$ corresponds to an ergodic invariant measure sitting on the repelling set R^i that corresponds to the unique equilibrium state for the potential $\phi - \log|\det \mathrm{d}T|$.

Throughout this section, we assume Hypothesis H1 holds true and employ the notation introduced in Section 2. In particular, we use "ε small enough" and "δ small enough" to refer to ε and δ as in Lemmas 2.9, 2.12 and 2.13. All arguments in this section hold for each $1 \le i \le k$ and every α-Hölder potential $\phi : R^i_\delta \to \mathbb{R}$, which we fix once and for all. To improve readability we drop the super-index ϕ of the weighted Markov process $X^{\varepsilon,\phi}_n$ and simply write X^ε_n.

3.1. Quasi-stationary measures on R^i_δ.

Denote by $\widehat{\mathcal{L}_\varepsilon}$ the L^1-normalised operator $\mathcal{L}_\varepsilon$, i.e.

$$\widehat{\mathcal{L}_\varepsilon} f = \frac{\mathcal{L}_\varepsilon f}{\|\mathcal{L}_\varepsilon f\|_1}.$$

Notation 3.1. Given a compact metric space (N, d) and $0 < \alpha < 1$ we denote by $\mathcal{C}^\alpha(N)$ the set of α-Hölder functions $f : N \to \mathbb{R}$ and consider the α-Hölder norm

$$\|f\|_{\mathcal{C}^\alpha} = \sup_{x \in N} |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x,y)^\alpha}.$$

To obtain a quasi-stationary density for the conditioned process on each component R^i_δ we apply the Schauder-Tychonoff fixed point theorem (see, e.g. [57, Theorem 2.2.3]) to the operator $\widehat{\mathcal{L}_\varepsilon}$ acting on a suitable space C_β.

Theorem 3.2. *Consider an α-Hölder potential $\phi : R^i_\delta \to \mathbb{R}$ and suppose that T satisfies Hypothesis H1 and $\|\mathrm{d}T(x)^{-1}\| < 1/(1+r)$. Let $\delta > 0$ be small enough. Then, for every $\varepsilon > 0$ small enough there exists a measure μ_ε on R^i_δ such that:*

1. *μ_ε is the unique quasi-stationary measure of the ϕ-weighted Markov process X^ε_n on R^i_δ,*
2. *μ_ε is absolutely continuous with respect to ρ, and*
3. *defining $m_\varepsilon := \mu_\varepsilon(\mathrm{d}x)/\rho(\mathrm{d}x)$, there exists $C > 0$ such that $\|m_\varepsilon\|_{\mathcal{C}^\alpha} \le C$ and $m_\varepsilon(x) > 0$ for every $x \in R^i_\delta$.*

Proof. Given $\beta > 0$ consider the set

$$C_\beta := \left\{ g \in L^1(R^i_\delta, \rho) \;\middle|\; \begin{array}{l} \int g \, \mathrm{d}\rho = 1, \ g > 0, \text{ and } \frac{g(x)}{g(y)} \le e^{\beta d(x,y)^\alpha} \text{ if } x, y \\ \text{lie in the same connected component of } R^i_\delta \end{array} \right\}.$$

We divide the proof into 3 steps.

Step 1. *There exists $\beta > 0$ such that $\widehat{\mathcal{L}_\varepsilon}(C_\beta) \subset C_\beta$.*

First of all, observe that if $\delta > 0$ is small enough $f > 0, f \in C_\beta$, implies $\mathcal{L}_\varepsilon f > 0$ for every $\varepsilon > 0$. Define $\psi : R^i_\delta \to \mathbb{R}$ as $\psi := \phi - \log |\det \mathrm{d}T|$ and let

$$D := \sup_{x \neq y} \frac{|\psi(x) - \psi(y)|}{d(x,y)^\alpha} < \infty,$$

Recall from Lemma 2.13 (3) that if $\varepsilon > 0$ is small enough, then given x, y in the same connected component C of R^i_δ we have

$$\#\{T^{-1}(x) \cap \Lambda_\delta\} = \#\{T^{-1}_\omega(y) \cap \Lambda_\delta\}$$

for every $\omega \in \Omega_\varepsilon$. Suppose that $\#\{T^{-1}(x) \cap \Lambda_\delta\} = \ell$. Let $h_1, \ldots, h_\ell : [-\varepsilon, \varepsilon]^m \times C \to R^i_\delta$ be the pre-image functions (inverse branches) defined in Lemma 2.12. Given $f \in C_\beta$, we have that

$$\begin{aligned}
\mathcal{L}_\varepsilon f(x) &= \mathbb{E}_\varepsilon \left[\sum_{T_\omega(z) = x} \frac{e^{\phi(z)} f(z)}{|\det \mathrm{d}T_\omega(z)|} \right] = \mathbb{E}_\varepsilon \left[\sum_{i=1}^{\ell} e^{\psi \circ h_i(\omega, x)} f \circ h_i(\omega, x) \right] \\
&= \mathbb{E}_\varepsilon \left[\sum_{i=1}^{\ell} e^{\psi \circ h_i(\omega,x) - \psi \circ h_i(\omega,y)} \frac{f \circ h_i(\omega, x)}{f \circ h_i(\omega, y)} f \circ h_i(\omega, y) e^{\psi \circ h_i(\omega, y)} \right] \\
&\le \left(\sup_{i \in \{1, \ldots, \ell\}} e^{(\beta + D) \mathrm{dist}(h_i(\omega,x), h_i(\omega,y))^\alpha} \right) \mathcal{L}_\varepsilon f(y) \\
&\le e^{\sigma_1^\alpha (\beta + D) \mathrm{dist}(x,y)^\alpha} \mathcal{L}_\varepsilon f(y).
\end{aligned}$$

Therefore, if $f \in C_\beta$ then $\widehat{\mathcal{L}_\varepsilon} f \in C_{\sigma_1^\alpha(\beta + D)}$. Taking $\beta > D\sigma_1^\alpha/(1 - \sigma_1^\alpha) > 0$ we conclude Step 1.

Step 2. *For every $\varepsilon > 0$ small, there exists $m_\varepsilon \in C_\beta$ such that $\widehat{\mathcal{L}_\varepsilon} m_\varepsilon = m_\varepsilon$.*

Observe that C_β is pre-compact and convex in $L^1(R^i_\delta, \rho)$. From the Schauder fixed-point theorem, there exists m_ε lying in the closure of C_β such that $\widehat{\mathcal{L}_\varepsilon} m_\varepsilon = m_\varepsilon$, which implies that $\mathcal{L}_\varepsilon m_\varepsilon = \lambda_\varepsilon m_\varepsilon$ for $\lambda_\varepsilon = \|\mathcal{L}_\varepsilon m_\varepsilon\| > 0$. We claim that $m_\varepsilon \in C_\beta$. Suppose by contradiction that $m_\beta \in \overline{C_\beta}^{L^1(R^i_\delta, \rho)} \setminus C_\beta$. Then, there exists $x \in R^i_\delta$ such that $m_\varepsilon(x) = 0$. Therefore, $m_\varepsilon(y) = 0$ for every y in the same connected component C_x of x in R^i_δ. Hence, for every $y \in C_x$

$$0 = m_\varepsilon(y) = \frac{1}{\lambda_\varepsilon^n} \mathbb{E}_\varepsilon \left[\sum_{T^n_\omega(z) = y} \frac{e^{S_n \phi(\omega, z)} \mathbb{1}_{R^i_\delta}(z) m_\varepsilon(z)}{|\det \mathrm{d}T^n_\omega(z)|} \right],$$

where $S_n \phi(\omega, z) = \sum_{i=0}^{n-1} \phi \circ T^i_\omega(z)$.

This implies that m_ε vanishes in the connected components of points in $T^{-n}(y) \cap R^i_\delta$, for every $y \in C_x$. Since there exists $z \in R^i$ such that $\{T^n(z)\}_{n \in \mathbb{N}}$ is dense in R^i, it follows that $m_\varepsilon \equiv 0$, which is a contradiction.

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

Item (1) follows from the same arguments in the proof of [46, Theorem 2]. Items (2) and (3) are readily verified since $m_\varepsilon \in C_\beta$ for a uniform β. ∎

For the remainder of this section, let $m_\varepsilon \in C_\beta$ denote the unique function such that $\mathcal{L}_\varepsilon m_\varepsilon = \lambda_\varepsilon m_\varepsilon$ and let μ_ε be the unique quasi-stationary measure, i.e. be such that $m_\varepsilon = \mu_\varepsilon(\mathrm{d}x)/\rho(\mathrm{d}x)$, as of Theorem 3.2.

3.2. **Analysis of the operator $\mathcal{P}_\varepsilon : L^\infty(R^i_\delta) \to L^\infty(R^i_\delta)$.** We now study the adjoint operator of $\mathcal{L}_\varepsilon$ to obtain the eigenfunction g_ε, associated with the maximal eigenvalue λ_ε from the previous result, and its properties. We also construct the unique quasi-ergodic measure of the ϕ-weighted Markov process X^ε_n on $\{g_\varepsilon > 0\}$.

Lemma 3.3. *Assume Hypothesis H1, let $\delta, \epsilon > 0$ be small enough and let λ_ε be the eigenvalue associated with m_ε from Theorem 3.2. Let $g_\varepsilon \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) \cap \mathcal{C}^0_+(R^i_\delta)$, then $R^i \subset \{g_\varepsilon > 0\}$.*

Proof. We divide the proof into two steps. First, we check that g_ε is positive on dense orbits of T and, second, construct a neighbourhood of R^i where g_ε is positive. It is clear that a dense orbit exists since $T : R^i \to R^i$ is topologically transitive by Lemma 2.7. Let $K_0 = \min_{x\in R^i_\delta} e^{\phi(x)} > 0$.

Step 1. *If $\{T^i(x_0)\}_{i\in\mathbb{N}}$ is dense in R^i, then $g_\varepsilon(x_0) > 0$.*

Recall that $g_\varepsilon \in \mathcal{C}^0_+(R^i_\delta)$. Assume that $g_\varepsilon(x_0) = 0$, then for every $n \in \mathbb{N}$,

$$0 = g_\varepsilon(x_0) = \frac{1}{\lambda^n_\varepsilon}\mathcal{P}^n_\varepsilon g_\varepsilon(x_0) \geq \frac{K^n_0}{\lambda^n_\varepsilon}\mathbb{E}_\varepsilon[g_\varepsilon \circ T^n_\omega(x_0) \cdot \mathbb{1}_{R^i_\delta} \circ T^n_\omega(x_0)].$$

Combining this with the submersion theorem applied to $\partial_\omega T_\omega$ (see [41, Theorem 4.12]) and the fact that R^i_δ is compact, we obtain that there exists $r_0 > 0$ such that $g_\varepsilon|_{B_{r_0}(T^n(x_0))} = 0$ for every $n \in \mathbb{N}$. Since $\{T^i(x_0)\}_{i\in\mathbb{N}}$ is dense in R^i, there exists a neighbourhood $U \supset R^i$ such that $g_\varepsilon|_U = 0$.

Let $T_{\cap R^i_\delta}(U) := T(U \cap R^i_\delta)$. Recall that, from (6), there exists $N \in \mathbb{N}$ such that $T^N_{\cap R^i_\delta}(U) \supset R^i_\delta$. Take $y \in R^i_\delta$. Then, there exists $z \in U$ such that $T^N(z) = y$ and $T^i(z) \in U$ for every $i \in \{1, \ldots, N\}$. Since

$$0 = g_\varepsilon(z) = \frac{1}{\lambda^N_\varepsilon}\mathcal{P}^N_\varepsilon g_\varepsilon(z) \geq \frac{K^N_0}{\lambda^N_\varepsilon}\mathbb{E}_\varepsilon[g_\varepsilon \circ T^N_\omega(z) \cdot \mathbb{1}_{R^i_\delta} \circ T^N_\omega(z)],$$

continuity of g_ε and the submersion theorem applied to $\partial_\omega T_\omega$ yields $g_\varepsilon(y) = 0$. This contradicts $g_\varepsilon \neq 0$.

Step 2. *There exists an open set $B \supset R^i$, such that $g_\varepsilon(x) > 0$ for every $x \in B$.*

Set $B := \{x \in R^i_\delta;\ \exists\ \omega_0 \in (-\varepsilon/2, \varepsilon/2)^m \text{ s.t. } T_{\omega_0}(x) \in R^i\}$. From the submersion theorem, we have that given $x \in B$ and $\omega_0 \in (-\varepsilon/2, \varepsilon/2)$ such that $T_{\omega_0}(x) \in R^i$ we obtain that there exists $r_1 > 0$ such that

$$\bigcup_{\omega\in\Omega_\varepsilon} T_\omega(x) \supset B_{r_1}(T_{\omega_0}(x)) \text{ for some } r_1 > 0.$$

Let $x_0 \in R^i$ be such that $\{T^i(x_0)\}_{i\in\mathbb{N}}$ is dense in R^i, then there exists N_0 such that $T^{N_0}(x_0) \in B_{r_1}(T_{\omega_0}(x))$. From Step 1, $g_\varepsilon(T^{N_0}(x_0)) > 0$. Continuity of g_ε then implies

$$0 < \frac{K_0}{\lambda_\varepsilon}\mathbb{E}_\varepsilon[g_\varepsilon \circ T_\omega(x) \cdot \mathbb{1}_{R^i_\delta} \circ T_\omega(x)] \leq \frac{1}{\lambda_\varepsilon}\mathcal{P}_\varepsilon g_\varepsilon(x) = g_\varepsilon(x).$$

This concludes Step 2 and proves the lemma. □

Theorem 3.4. *Consider the operator $\mathcal{P}_\varepsilon : L^\infty(R^i_\delta) \to L^\infty(R^i_\delta)$. Then, $\ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) = \operatorname{span}(g_\varepsilon)$ for some $g_\varepsilon \in \mathcal{C}^0_+(R^i_\delta)$.*

Proof. Let $g_\varepsilon \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$. Since $\mathcal{P}_\varepsilon$ is strong Feller, then $g_\varepsilon \in \mathcal{C}^0(R^i_\delta, \mathbb{C})$. Moreover, since $\mathcal{P}(\mathcal{C}^0(R^i_\delta)) \subset \mathcal{C}^0(R^i_\delta)$, it is clear that $\mathrm{Re}(g_\varepsilon), \mathrm{Im}(g_\varepsilon) \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$. Given $g_\varepsilon \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) \cap \mathcal{C}^0(R^i_\delta)$, we claim that $g_\varepsilon^\pm \in \ker(\mathcal{P} - \lambda_\varepsilon)$. Indeed, observe that (see [40, Propositions 3.1.1 and 3.1.3])

$$g_\varepsilon^\pm = \left(\frac{1}{\lambda_\varepsilon}\mathcal{P}_\varepsilon g_\varepsilon\right)^\pm \leq \frac{1}{\lambda_\varepsilon}\mathcal{P}_\varepsilon g_\varepsilon^\pm,$$

where $g_\varepsilon^\pm = \max\{0, \pm g_\varepsilon\}$. Therefore,

$$\begin{aligned} 0 &= \int_{R^i_\delta} \left(\frac{1}{\lambda_\varepsilon}\mathcal{P}_\varepsilon |g_\varepsilon|(x) - |g_\varepsilon|(x)\right)\mu_\varepsilon(\mathrm{d}x) \\ &= \int_{R^i_\delta} \left(\frac{1}{\lambda_\varepsilon}\mathcal{P}_\varepsilon g_\varepsilon^+(x) - g_\varepsilon^+(x)\right)\mu_\varepsilon(\mathrm{d}x) + \int_{R^i_\delta} \left(\frac{1}{\lambda_\varepsilon}\mathcal{P}^\varepsilon g_\varepsilon^-(x) - g_\varepsilon^-(x)\right)\mu_\varepsilon(\mathrm{d}x). \end{aligned}$$

From Theorem 3.2, $\operatorname{supp}\mu_\varepsilon = R^i_\delta$, therefore $\mathcal{P}_\varepsilon g_\varepsilon^\pm = \lambda_\varepsilon g_\varepsilon^\pm$.

Take $g_1, g_2 \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) \cap \mathcal{C}^0_+(R^i_\delta)$. From Lemma 3.3, we have $g_1, g_2 > 0$ on R^i. Choose $t_0 > 0$ such that $t_0 = \inf\{t;\, g_1(x) - tg_2(x) < 0 \text{ for some } x \in R^i\}$.

Since $g_1 - t_0 g_2 \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$, then $(g_1 - t_0g_2)^+ \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$. However, from the choice of t_0 and Lemma 3.3 we obtain that $(g_1 - t_0 g_2)^+ = 0$. From the minimality of t_0, it follows that $g_1(x) = t_0 g_2(x)$ for every $x \in R^i$. Observe that $(g_1 - t_0 g_2)^+ = 0$ yields that $t_0 g_2 \geq g_1$. Therefore $t_0 g_2 - g_1 \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) \cap \mathcal{C}^0_+(R^i_\delta)$ and $(t_0g_2 - g_1)|_{R^i} = 0$, implying that $t_0 g_2 - g_1 = 0$. □

For the remainder of this section, let $g_\varepsilon \in C^0_+(R^i_\delta)$ denote the unique function such that $\mathcal{P}_\varepsilon g_\varepsilon = \lambda_\varepsilon g_\varepsilon$ and normalised so that $\int g_\varepsilon \mathrm{d}\mu_\varepsilon = 1$, as of Theorem 3.4. We summarise some relevant properties of $\mathcal{P}_\varepsilon : L^\infty(R^i_\delta) \to L^\infty(R^i_\delta)$ that have been shown above:

(1) $\mathcal{P}_\varepsilon : L^\infty(R^i_\delta, \rho) \to L^\infty(R^i_\delta, \rho)$ is a strong Feller operator,
(2) $\dim\ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) = 1$, where $\lambda_\varepsilon = r(\mathcal{P}_\varepsilon)$ is the spectral radius,
(3) there exists $\mu_\varepsilon \in \mathcal{M}_+(R^i_\delta)$ and $g_\varepsilon \in \mathcal{C}^0_+(R^i_\delta)$, such that $\mathcal{P}^*_\varepsilon \mu_\varepsilon = \lambda_\varepsilon \mu_\varepsilon$ and $\mathcal{P} g_\varepsilon = \lambda_\varepsilon g_\varepsilon$, and
(4) $\mu_\varepsilon \ll \rho$ and $\operatorname{supp}\mu_\varepsilon = R^i_\delta$.

In particular, this implies that $\mathcal{P}_\varepsilon$ satisfies Hypothesis HA in Appendix A. The lemma below is a consequence of the properties just listed and Theorems A.13 and A.14, whose proof is deferred to the appendix in order not to break the flow of the text.

Lemma 3.5. *The measure $\nu_\varepsilon(\mathrm{d}x) := g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x)$ is the unique quasi-ergodic measure of the ϕ-weighted Markov process X^ε_n on $\{g_\varepsilon > 0\}$. If we further assume that $T : R^i \to R^i$ is topologically mixing, then ν_ε is a ϕ-weighted quasi-ergodic measure on R^i_δ.*

Proof. It is clear from the properties of $\mathcal{P}_\varepsilon$ listed above that it satisfies Hypothesis HA (see Appendix A). Hence, Theorem A.13 implies that ν_ε is the unique ϕ-weighted quasi-ergodic measure for X^ε_n on $\{g_\varepsilon > 0\}$.

To finish the proof of the theorem, it remains to be shown that if T is topologically mixing, then ν_ε is a ϕ-weighted quasi-ergodic measure for X^ε_n on R^i_δ. Since T is topologically mixing, then X^ε_n is aperiodic in R^i_δ and $\{g_\varepsilon > 0\}$. Let $k_\varepsilon := \#(\sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P}_\varepsilon) \cap \mathbb{S}^1)$. From Lemma A.3, $k_\varepsilon < \infty$. Moreover, from Proposition A.6 and Lemma A.7 we obtain that there exist sets $C_i \subset \{g_\varepsilon > 0\}$, $i \in \{0, 1, \ldots, k_\varepsilon - 1\}$ such that $C_0 \sqcup C_1 \sqcup \ldots \sqcup C_{k_\varepsilon - 1} = \{g_\varepsilon > 0\}$, and $\{\mathcal{P}_\varepsilon \mathbb{1}_{C_i} > 0\} \subset C_{i-1 \pmod{k_\varepsilon}}$, for every $i \in \{0, 1, \ldots, k_\varepsilon - 1\}$. Since T is assumed to be topologically mixing on R^i, X^ε_n is aperiodic, thus $k_\varepsilon = 1$. Finally, from Theorem A.14 we obtain that ν_ε is a quasi-ergodic measure of the ϕ-weighted Markov process X^ε_n on R^i_δ. □

3.3. **Proof of the main (local) result.** We conclude this section with the main results concerning the stochastic stability of equilibrium states on each repeller R^i_δ and their limiting behaviour as $\varepsilon \to 0$.

Notation 3.6. Recall that given a suitable function f we denote the action of the (deterministic) Ruelle-Perron-Frobenius operator $\mathcal{L}$ for the potential $\phi - \log|\det \mathrm{d}T|$ [57, Chapter 12] by

$$\mathcal{L} : f \mapsto \sum_{T(y)=x} \frac{e^{\phi(y)} f(y) \mathbb{1}_{R^i_\delta}(y)}{|\det \mathrm{d}T(y)|},$$

when this is well-posed (see Theorem 2.14). In particular, this is the case for any function supported on R^i_δ.

Lemma 3.7. *Consider a sequence $\{m_\varepsilon\}_{\varepsilon>0} \subset \mathcal{C}^\alpha(R^i_\delta)$, with $\|m_\varepsilon\|_{\mathcal{C}^\alpha} \le C$ for every $\varepsilon > 0$. Then, $\|\mathcal{L}_\varepsilon m_\varepsilon - \mathcal{L} m_\varepsilon\|_{L^\infty} \xrightarrow{\varepsilon\to 0} 0$.*

Proof. Using the usual bounds, we have that

$$\begin{aligned}
|\mathcal{L}_\varepsilon m_\varepsilon(x) - \mathcal{L} m_\varepsilon(x)| &= \left| \mathbb{E}_\varepsilon \left[\sum_{T_\omega(y)=x} \frac{e^{\phi(y)} m_\varepsilon(y)}{|\det \mathrm{d}T_\omega(y)|} - \sum_{T(y)=x} \frac{e^{\phi(y)} m_\varepsilon(y)}{|\det \mathrm{d}T(y)|} \right] \right| \\
&= \left| \mathbb{E}_\varepsilon \left[\sum_i \frac{e^{\phi\circ h_i(\omega,x)} m_\varepsilon \circ h_i(\omega,x)}{|\det \mathrm{d}T_\omega \circ h_i(\omega,x)|} - \frac{e^{\phi\circ h_i(0,x)} m_\varepsilon \circ h_i(0,x)}{|\det \mathrm{d}T \circ h_i(0,x)|} \right] \right| \\
&\le \sum_i \mathbb{E}_\varepsilon \Bigg[\left| \frac{e^{\phi\circ h_i(\omega,x)} m_\varepsilon(h_i(\omega,x)) - e^{\phi\circ h_i(0,x)} m_\varepsilon(h_i(0,x)}{|\det \mathrm{d}T_\omega(h_i(\omega,x))|} \right| \\
&\qquad + |e^{\phi\circ h_i(\omega,x)} m_\varepsilon \circ h_i(0,x)| \left| \frac{1}{|\det \mathrm{d}T_\omega(h_i(\omega,x))|} - \frac{1}{|\det \mathrm{d}T(h_i(0,x))|} \right| \Bigg] \\
&\le N \max_i K_i C (\sup|D_\omega h_i|\varepsilon)^\alpha + C K_i' \sup|D_\omega h_i|\varepsilon \longrightarrow 0, \quad \text{as } \varepsilon \to 0,
\end{aligned}$$

where $N = \sup_{(x,\omega)\in R^i_\delta\times\Omega_\varepsilon} \#(T_\omega^{-1}(\{x\} \cap R^i_\delta) < \infty$, the K_i provide a bound for the term $|\det \mathrm{d}T_\omega(h_i(\omega,x))^{-1}|$, $C \sup|D_\omega h_i|^\alpha \varepsilon^\alpha$ are a Hölder-like bound for the difference

$$|e^{\phi\circ h_i(\omega,x)} m_\varepsilon \circ h_i(\omega,x) - e^{\phi\circ h_i(0,x)} m_\varepsilon \circ h_i(0,x)|,$$

and so is $K_i' \sup|D_\omega h_i|\varepsilon$ for $(|\det \mathrm{d}T_\omega(h_i(\omega,x))|^{-1} - |\det \mathrm{d}T(h_i(0,x))|^{-1}|)^{-1}$. □

Proposition 3.8. *Let $m_\varepsilon : R^i_\delta \to \mathbb{R}$ be the functions given by Theorem 3.2. There exist $\lambda_0 > 0$ and $m_0 \in \mathcal{C}^0(R^i_\delta)$, such that $\lambda_\varepsilon \xrightarrow{\varepsilon\to 0} \lambda_0$ and $m_\varepsilon|_{R^i} \xrightarrow{\varepsilon\to 0} m_0$ in $\mathcal{C}^0(R^i)$, with $\mathcal{L} m_0 = \lambda_0 m_0$.*

Proof. Since $\|m_\varepsilon\|_{\mathcal{C}^\alpha} \le C$, there exists $\{\varepsilon_n\}_{n\in\mathbb{N}}$, such that $\varepsilon_n \to 0$ and $m_0 \in \mathcal{C}^0(R^i_\delta)$, such that $\|m_{\varepsilon_n} - m_0\|_{\mathcal{C}^0(R^i_\delta)} \to 0$. We can assume without loss of generality, by restricting to a subsequence if necessarily, that $\lambda_{\varepsilon_n} \to \lambda_0 \ge 0$.

From Lemma 3.7 we obtain that

$$\lambda_0 m_0 = \lim_{n\to\infty} \lambda_{\varepsilon_n} m_{\varepsilon_n} = \lim_{n\to\infty} \mathcal{L}_{\varepsilon_n} m_{\varepsilon_n} = \lim_{n\to\infty} \mathcal{L} m_{\varepsilon_n} = \mathcal{L} m_0.$$

In the following, we show that $\lambda_0 > 0$. Since $\int_{R^i_\delta} m_{\varepsilon_n}(x)\rho(\mathrm{d}x) = 1$ for every $n \in \mathbb{N}$, then by the Lebesgue-dominated convergence theorem $\int_{R^\delta_i} m_0(x)\rho(\mathrm{d}x) = 1$. Therefore, there exists $x_0 \in R^i_\delta$ such that $m_0(x) > 0$. Let $C_{x_0} \subset R^i_\delta$ be the connect component of x_0 in R^i_δ. From the proof of Theorem 3.2 (Step 2), we obtain that for every $n \in \mathbb{N}$, $e^{-\beta d(x_0,y)} m_{\varepsilon_n}(x_0) \le m_{\varepsilon_n}(y)$, for every $y \in C_{x_0}$. Therefore, taking $n \to \infty$ we obtain that

$0 < m_0(y)$, for every $y \in C_{x_0}$. In particular $m_0|_{R^i} \neq 0$. Assume by contradiction that $\lambda_0 = 0$. Then, for every $x \in R^i$

$$\sum_{T(y)=x} \frac{e^{\phi(y)} m_0(y)}{|\det \mathrm{d}T(y)|} = \mathcal{L}m_0 = 0.$$

Since $R^i \subset T^{-1}(R^i)$ the above equation implies that $m_0|_{R^i} = 0$, which is a contradiction. Therefore $\lambda_0 > 0$.

Since there exists a unique $m_0 \in \mathcal{C}^0(R^i)$ such that $\mathcal{L}m_0(x) = \lambda_0 m_0(x)$ for every $x \in R^i$ and $\lambda_0 > 0$ [57, Chapter 12], the proposition follows. □

Proposition 3.9. *Let $g_\varepsilon \in \mathcal{C}^0_+(R^i_\delta)$ be the functions given by Theorem 3.4 and consider $\lambda_0 > 0$ as in Proposition 3.8. There exists a probability measure γ on R^i_δ such that $g_\varepsilon(x)\mathrm{d}x \xrightarrow{\varepsilon\to 0} \gamma(\mathrm{d}x)$ in the weak* topology of $\mathcal{M}(R^i_\delta)$. Moreover, γ is the unique conformal measure for T on R^i for the potential $\phi - \log|\det \mathrm{d}T|$, i.e. γ is the unique probability measure on R^i such that $\mathcal{L}^*\gamma = \lambda_0\gamma$.*

Proof. Let γ be an accumulation point of $\{g_\varepsilon(x)\mathrm{d}x\}_{\varepsilon>0}$ in the weak* topology of $\mathcal{M}(R^i_\delta)$, i.e. there exists a sequence $\{g_{\varepsilon_n}(x)\mathrm{d}x\}_{n\in\mathbb{N}}$ such that $\varepsilon_n \xrightarrow{n\to\infty} 0$ and $g_{\varepsilon_n}(x)\mathrm{d}x \xrightarrow{n\to\infty} \gamma(\mathrm{d}x)$ in the weak* topology. We first check that γ is a conformal measure on R^i_δ. Indeed, for a test function $f \in \mathcal{C}^\alpha(R^i_\delta)$ we have:

$$\begin{aligned}
(\mathcal{L}^*\gamma)(f) &= \int_{R^i_\delta} \mathcal{L}f \,\mathrm{d}\gamma = \lim_{n\to\infty} \int_{R^i_\delta} \mathcal{L}f(x) g_{\varepsilon_n}(x)\mathrm{d}x \\
(\text{Lem. } 3.7) &= \lim_{n\to\infty} \int_{R^i_\delta} \mathcal{L}_{\varepsilon_n} f(x) g_{\varepsilon_n}(x)\mathrm{d}x = \lim_{n\to\infty} \int_{R^i_\delta} f(x)\mathcal{P}_{\varepsilon_n} g_{\varepsilon_n}(x)\mathrm{d}x \\
&= \lim_{n\to\infty} \lambda_{\varepsilon_n} \int_{R^i_\delta} f(x) g_{\varepsilon_n}(x)\mathrm{d}x = \lambda_0 \int_{R^i_\delta} f(x)\gamma(\mathrm{d}x) = \lambda_0\gamma(f).
\end{aligned}$$

We claim that $\operatorname{supp}\gamma \subset R^i$. From Lemma 2.9 item (2), we obtain that

$$\begin{aligned}
1 = \gamma(R^i_\delta) &= \frac{1}{\lambda_0}\int_{R^i_\delta} \mathcal{L}\mathbb{1}_{R^i_\delta}(x)\gamma(\mathrm{d}x) = \frac{1}{\lambda_0}\int_{R^i_\delta} \sum_{T(y)=x} \frac{e^{\phi(y)}\mathbb{1}_{R^i_\delta}(y)}{|\det \mathrm{d}T(y)|}\gamma(\mathrm{d}x) \\
&= \frac{1}{\lambda_0}\int_{R^i_\delta} \sum_{T(y)=x} \frac{e^{\phi(y)}\mathbb{1}_{R^i_{\sigma_0\delta}}(y)}{|\det \mathrm{d}T(y)|}\gamma(\mathrm{d}x) = \frac{1}{\lambda_\varepsilon}\int_{R^i_\delta} \mathcal{L}\mathbb{1}_{R^i_{\sigma_0\delta}}(x)\gamma(\mathrm{d}x) = \gamma(R^i_{\sigma_0\delta}).
\end{aligned}$$

Repeating this argument n times we obtain that $\gamma(R^i_{\sigma_0^n\delta}) = 1$ and the claim follows by taking $n\to\infty$. Since there exists a unique measure γ in R^i such that $\mathcal{L}^*\gamma = \lambda_0\gamma$ (see [57, Chapter 12]), we conclude that $g_\varepsilon(x)\mathrm{d}x \xrightarrow{\varepsilon\to 0} \gamma(\mathrm{d}x)$ in the weak* topology. □

Proposition 3.10. *Assume Hypothesis H1 and that $\|\mathrm{d}T(x)^{-1}\| < 1/(1+r)$ for every $x \in R$. Let ν_ε be the unique quasi-ergodic measure of the ϕ-weighted Markov process X^ε_n on $\{g_\varepsilon > 0\}$. Then, $\nu_\varepsilon \xrightarrow{\varepsilon\to 0} \nu_0(x) := m_0(x)\gamma(\mathrm{d}x)$, in the weak* topology. Moreover, ν_0 is the unique T-invariant equilibrium state for the potential $\phi - \log|\det \mathrm{d}T|$ in R^i.*

Proof. From Lemma 3.5, we have that $\nu_\varepsilon(\mathrm{d}x) = g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x)$ is the unique quasi-ergodic measure of the ϕ-weighted Markov process X^ε_n on $\{g_\varepsilon > 0\}$ such that $R^i \subset \operatorname{supp} g_\varepsilon$. Since $m_\varepsilon \xrightarrow{\varepsilon\to 0} m_0$ in $\mathcal{C}^0(R^i_\delta)$ and $g_\varepsilon(x)\mathrm{d}x \xrightarrow{\varepsilon\to 0} \gamma(\mathrm{d}x)$ in the weak* topology then, $\nu_\varepsilon \xrightarrow{\varepsilon\to 0} \nu_0$ in the weak* topology. The final part of the proposition follows from well-known results in the thermodynamic formalism for expanding maps (see [57, Chapter 12]). □

We close this section proving Theorem 2.10.

Proof of Theorem 2.10. Let $n_0 \in \mathbb{N}$ be large enough such that T^{n_0} has the same dynamical decomposition as T and $|(T^{n_0})'(x)| > 1+r_0$ for every $x \in R$. Then, Proposition 3.10 holds for T^n_ω. It is clear that:

(1) $\sigma_{\text{per}}\left(\frac{1}{\lambda_\varepsilon^{n_0}}\mathcal{P}_\varepsilon^{n_0}\right) = \sigma_{\text{per}}\left(\frac{1}{\lambda_\varepsilon}\mathcal{P}_\varepsilon\right)$ and $\sigma_{\text{per}}\left(\frac{1}{\lambda_\varepsilon^{n_0}}\mathcal{L}_\varepsilon^{n_0}\right) = \sigma_{\text{per}}\left(\frac{1}{\lambda_\varepsilon}\mathcal{L}_\varepsilon\right)$,

(2) $\sigma_{\text{per}}\left(\frac{1}{\lambda_0^{n_0}}\mathcal{L}^{n_0} : \mathcal{C}^0(R^i) \to \mathcal{C}^0(R^i)\right) = \sigma_{\text{per}}\left(\frac{1}{\lambda_0}\mathcal{L} : \mathcal{C}^0(R^i) \to \mathcal{C}^0(R^i)\right)$, and

(3) $\sigma_{\text{per}}\left(\frac{1}{\lambda_0^{n_0}}(\mathcal{L}^*)^{n_0} : \mathcal{M}(R^i) \to \mathcal{M}(R^i)\right) = \sigma_{\text{per}}\left(\frac{1}{\lambda_0}\mathcal{L}^* : \mathcal{M}(R^i) \to \mathcal{M}(R^i)\right)$,

where σ_{per} denotes the point peripheral spectrum, i.e. $\sigma_{\text{per}}(\mathcal{P}_\varepsilon) = \{\alpha \in \mathbb{C};\ |\alpha| = r(\mathcal{P}_\varepsilon) \text{ and } \ker(\mathcal{P}_\varepsilon - \alpha) \neq \emptyset\}$.

Therefore $\nu_\varepsilon(\mathrm{d}x) = g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x)$ is the unique quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on $\{g_\varepsilon > 0\}$ such that $R^i \subset \operatorname{supp}\nu_\varepsilon$. Since ν_0 is a equilibrium state of T^{m_0} for the potential $S_{n_0}\phi - \log|\det \mathrm{d}T^{n_0}|$ on R^i and an equilibrium state of T for the potential $\phi - \log|\det \mathrm{d}T|$ on R^i, the proof is finished.

If ν_0 is mixing for the map $T : R^i \to R^i$, then $T : R^i \to R^i$ is topologically mixing since $\operatorname{supp}\nu_0 = R^i$. From Lemma 3.5 we obtain that ν_ε is a quasi-ergodic measure ϕ-weighted Markov process X_n^ε on R^i_δ and the result follows. □

Corollary 3.11. *Assume Hypothesis H1 and that $T|_{R^i}$ is topologically mixing. Let $\delta > 0$ be small enough. For every $\varepsilon > 0$ sufficiently small, let $\nu_\varepsilon(\mathrm{d}x)$ be the unique quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on R^i_δ such that $R^i \subset \operatorname{supp}\nu_\varepsilon$. Then, $\nu_\varepsilon(\mathrm{d}x) \xrightarrow{\varepsilon\to 0} \nu_0(\mathrm{d}x)$ in the weak* topology. Finally, ν_0 is the unique T-invariant equilibrium state for the potential $\phi - \log|\det \mathrm{d}T|$ on R^i.*

Proof. From Lemma 3.5 we have that $g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x)$ is a quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on R^i_δ. Combining this observation with Theorem 2.10 we obtain the result. □

4. The global problem

In this section, we prove conditioned stochastic stability of equilibrium states on the global repeller Λ by studying the quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on Λ_δ and absorbed in $\partial := U \cup (E \setminus M_\delta)$. As in Section 3, let us fix once and for all an α-Hölder potential $\phi : M_\delta \setminus U \to \mathbb{R}$. Moreover, we assume that T satisfies Hypothesis H2.

We start by arguing that restricting the study of quasi-ergodic measures on Λ_δ is sufficient to characterise those on $M_\delta \setminus U$. Then, we decompose Λ_δ into transient and recurrent subsets, the latter being those that contain the original repellers R^i. In particular, we show that all the relevant information for the global dynamics follows from the recurrent subset containing the repeller R^0 of maximal growth rate. The stochastic stability of global equilibrium states is then inferred via the stochastic stability of equilibrium states around R^0.

Proposition 4.1. *Assume that T satisfies Hypothesis H2. Let $\delta > 0$ be sufficiently small and μ_ε be a quasi-stationary measure of the ϕ-weighted Markov process X_n^ε on $M_\delta \setminus U$. Then, for sufficiently small $\varepsilon > 0$,*

(1) $\mu_\varepsilon \ll \rho$,

(2) $\operatorname{supp}\mu_\varepsilon \cap \Lambda_\delta \neq \emptyset$, *and*

(3) $\mu_\varepsilon|_{\Lambda_\delta}$, *after normalisation, is a quasi-stationary measure of the ϕ-weighted Markov process X_n^ε on Λ_δ.*

Proof. Observe that (1) follows directly from the fact that $\mathcal{P}_\varepsilon(x, \mathrm{d}y) \ll \rho(\mathrm{d}y)$ for every $x \in M_\delta \setminus U$.

To show item (2), arguing by contradiction, suppose that $\operatorname{supp}\mu_\varepsilon \cap \Lambda_\delta = \emptyset$. We claim that there exists $N \in \mathbb{N}$ and $\varepsilon > 0$ small enough such that for every $x \in \overline{M_\delta \setminus \Lambda_\delta}$ there exists $i \in \{0, 1, \dots, N\}$ such that $T^i_\omega(x) \in U \cup (E \setminus M_\delta)$ for every $\omega \in \Omega_\varepsilon$, or in other words, $\tau(x, \omega) \le N$ for every $\omega \in \Omega_\varepsilon$. This is sufficient to prove (2) since, if true, any measurable set $A \subset M_\delta \setminus U$ would be assigned measure zero:

$$\begin{aligned}
\mu_\varepsilon(A) &= \frac{1}{\lambda_\varepsilon^N} \int_{M_\delta \setminus U} \mathcal{P}_\varepsilon^N(x, A)\mu_\varepsilon(\mathrm{d}x) \\
\left({\scriptstyle \text{assumed} \atop \operatorname{supp}\mu_\varepsilon \cap \Lambda_\delta = \emptyset}\right) &= \frac{1}{\lambda_\varepsilon^N} \int_{M_\delta \setminus \Lambda_\delta} \mathcal{P}_\varepsilon^N(x, A)\mu_\varepsilon(\mathrm{d}x) \\
&= \frac{1}{\lambda_\varepsilon^N} \int_{M_\delta \setminus \Lambda_\delta} \mathbb{E}_\varepsilon\left[e^{\sum_{i=0}^{N-1} \phi \circ T^i_\omega(x)} \mathbb{1}_A \circ T^N_\omega(x) \mathbb{1}_{\{\tau(\omega, x) > N\}}\right] \mu_\varepsilon(\mathrm{d}x) = 0,
\end{aligned}$$

which is a contradiction. To verify the claim, choose $y \in \overline{M_\delta \setminus \Lambda_\delta}$. Then, there exists $n(y) \in \mathbb{N}$ such that $T^n(y) \in U \cup (E \setminus M_\delta)$. Since this is an open set, by continuity of T and $\mathcal{C}^2$ closeness of the perturbation, there exists $r(y) > 0$ and $\varepsilon(y) > 0$ such that $T^{n(y)}_\omega(B_{r(y)}(y)) \subset U \cup (E \setminus M_\delta)$ for all $\omega \in \Omega_{\varepsilon(y)}$. Consider a finite open cover of $M_\delta \setminus \Lambda_\delta$ with such balls around n points $y_1, \dots, y_n$ with respective radius $r(y_1), \dots, r(y_n)$. Setting $N = \max\{n(y_1), \dots, n(y_n)\}$, and $\varepsilon = \min\{\varepsilon(y_1), \dots, \varepsilon(y_n)\}$ the claim follows.

Finally we show (3). Since $T^{-1}(\Lambda) \cap M = \Lambda$, from the same proof of Lemma 2.13 items (2) and (3) we obtain that $T^{-1}_\omega(\Lambda_\delta) \cap M_\delta \subset \Lambda_\delta$ for every $\omega \in \Omega_\varepsilon$. Let A be a measurable subset of Λ_δ, then

$$\begin{aligned}
\int_{\Lambda_\delta} \mathcal{P}_\varepsilon(x, A)\mu_\varepsilon(\mathrm{d}x) &= \int_{\Lambda_\delta} e^{\phi(x)} \mathbb{E}_\varepsilon[\mathbb{1}_A \circ T_\omega(x)]\mu_\varepsilon(\mathrm{d}x) = \int_{M_\delta \setminus U} e^{\phi(x)} \mathbb{E}_\varepsilon[\mathbb{1}_A \circ T_\omega(x)]\mu_\varepsilon(\mathrm{d}x) \\
&= \int_{M_\delta \setminus U} \mathcal{P}_\varepsilon(x, A)\mu_\varepsilon(\mathrm{d}x) = \lambda_\varepsilon \mu_\varepsilon(A),
\end{aligned}$$

so $\mu_\varepsilon|_{\Lambda_\delta}$ normalised is a quasi-stationary measure of the ϕ-weighted Markov process X^ε_n on Λ_δ. □

Proposition 4.2. *Assume that T satisfies Hypothesis H2. Consider the operator $\mathcal{P}_\varepsilon : L^\infty(M_\delta \setminus U) \to L^\infty(M_\delta \setminus U)$. If $g \in L^\infty_+(M_\delta \setminus U)$ is such that $\mathcal{P}_\varepsilon g = \lambda g$, then $g|_{M_\delta \setminus \Lambda_\delta} = 0$.*

Proof. The proof of Proposition 4.1 yields that for $\varepsilon > 0$ small enough there exists N such that $T^N_\omega(x) \in U$ for every $x \in M_\delta \setminus \Lambda_\delta$ and $\omega \in \Omega_\varepsilon$. Therefore,

$$\mathcal{P}_\varepsilon^N(x, M_\delta \cap U) = 0 \text{ for every } x \in M_\delta \setminus \Lambda_\delta.$$

It follows that for every $x \in M_\delta \setminus \Lambda_\delta$,

$$0 \le g(x) = \frac{1}{\lambda^N} \mathcal{P}_\varepsilon^n g(x) \le \frac{\|g\|_\infty}{\lambda^N} \mathcal{P}_\varepsilon(x, M_\delta \setminus U) = 0,$$

verifying the claim. □

As a result of Propositions 4.1 and 4.2, it is natural to redefine the operator $\mathcal{P}_\varepsilon$ as

$$\begin{aligned}
\mathcal{P}_\varepsilon : L^\infty(\Lambda_\delta) &\to L^\infty(\Lambda_\delta) \\
f &\mapsto e^\phi \mathbb{E}_\varepsilon[f \circ T_\omega \cdot \mathbb{1}_{\Lambda_\delta} \circ T_\omega],
\end{aligned}$$

and denote by $\lambda_\varepsilon = r(\mathcal{P}_\varepsilon)$ its spectral radius. Moreover, observe that

$$\begin{aligned}
\mathcal{L}_\varepsilon : L^1(\Lambda_\delta) &\to L^1(\Lambda_\delta) \\
f &\mapsto \mathbb{E}_\varepsilon\left[\sum_{T_\omega(y) = x} \frac{e^{\phi(y)} f(y) \mathbb{1}_{\Lambda_\delta}(y)}{|\det \mathrm{d}T_\omega(y)|}\right]
\end{aligned}$$

is well defined and that $\mathcal{L}_\varepsilon^* = \mathcal{P}_\varepsilon$.

4.1. **Recurrent and transient regions.** In this section, we represent the relevant dynamical behaviour of the absorbing Markov process X_n^ε for every $\varepsilon > 0$ via a graph whose vertices are the connected components of Λ_δ. This approach resembles the graphs constructed via chain recurrence and filtration methods for classical dynamical systems (see [24, 28, 27]). Later, we use this construction to characterise the support of the relevant quasi-stationary measure of the ϕ-weighted Markov process X_n^ε.

Given $\varepsilon > 0$, we define an equivalence relation $\sim_\varepsilon$ on the set of connected components

$$\Gamma_\delta \coloneqq \{C \subset \Lambda_\delta;\ C \text{ is a connected component of } \Lambda_\delta\}$$

as follows: for any $C_1, C_2 \in \Gamma_\delta$, we say that $C_1 \sim_\varepsilon C_2$ if

- $C_1 = C_2$, or
- both sets are reachable from each other, i.e. for every $i, j \in \{1, 2\}$, there exist $W_0, W_1, \ldots, W_n, W_{n+1} \in \Gamma_\delta$ such that $\min_{\ell\in\{0,\ldots,n\}} \sup_{x\in W_\ell} \mathcal{P}_\varepsilon(x, W_{\ell+1}) > 0$, with $W_0 = C_i$ and $W_{n+1} = C_j$.

Proposition 4.3. *Assume that T satisfies Hypothesis H2. The set of equivalence classes $\Gamma_\delta/\sim_\varepsilon$ stabilises as $\varepsilon \to 0$, i.e. there exist $C_1, \ldots, C_n \in \Gamma_\delta$ such that for every ε small enough we have that*

$$\Gamma_\delta/\sim_\varepsilon = \{[C_1], \ldots, [C_n]\},$$

where $[C_i]$ represents the equivalence class of the element C_i.

Proof. Observe that if $0 < \varepsilon_1 < \varepsilon_2$, then $C_1 \sim_{\varepsilon_1} C_2$ implies $C_1 \sim_{\varepsilon_2} C_2$. Since

The amount of elements of Γ_δ is finite, and observe that if $0 < \varepsilon_1 < \varepsilon_2$, then $C_1 \sim_{\varepsilon_1} C_2$ implies $C_1 \sim_{\varepsilon_2} C_2$. This ensures that $\Gamma_\delta/\sim_\varepsilon$ stabilises as $\varepsilon \to 0$. □

Definition 4.4. Given $\delta > 0$ small enough, let $C_1, \ldots, C_n \in \Gamma_\delta$ be the sets given in Proposition 4.3. Define

$$M_i \coloneqq \bigcup_{C\in[C_i]} C,$$

i.e. M_i is the (disconnected) region spanned by all elements in the class $[C_i]$. Then:

- If there exists $j \in \{1, \ldots, k\}$ such that $R_\delta^j \subset M_i$, we say that M_i is a *recurrent region*.
- If there are no sets R_δ^j intersecting M_i, we say that M_i is a *transient region*.

Lemma 4.5. *All regions M_i can be classified as either recurrent or transient.*

Proof. Assume that M_i is not a transient region so that there exists R_δ^j such that such that $R_\delta^j \cap M_i \neq \emptyset$. Then, there exists a connected component $C \Subset R_\delta^j$ such that $C \subset M_i$. Since T is topologically transitive on R^j we obtain that $R_\delta^j \subset M_i$ and therefore M_i is recurrent. □

Proposition 4.6. *Let M_t be a transient region, then there exists $N \in \mathbb{N}$ such that for all $x \in M_t$, $\mathcal{P}_\varepsilon^n(x, M_t) = 0$ for all $n \geq N$.*

Proof. We begin by showing that there exists $N \in \mathbb{N}$ such that for every $x \in M_t$, either:

- $T^n(x) \in \operatorname{Int}(\bigcup M_r)$ for some $n \leq N$, where $\bigcup M_r$ is the union of all recurrent regions, or
- $T^n(x) \notin \Lambda_\delta$ for some $n \leq N$.

Let $x \in \widetilde{\Lambda} \cap M_t$, where

$$\widetilde{\Lambda} \coloneqq \{x \in M_\delta;\ \text{there exists } n \in \mathbb{N} \text{ such that } T^n(x) \in R\}.$$

There exists an open neighbourhood U_x of x, such that $T^{n_x}(U_x) \subset \operatorname{Int}(\bigcup M_r)$, the union of recurrent regions. Since $\widetilde{\Lambda} \cap M_t$ is compact, there exist points $x_1, \ldots, x_s$ with respective open neighbourhoods $U_{x_1}, \dots, U_{x_s}$ such that

$$\widetilde{\Lambda} \cap M_t \subset \bigcup_{j=1}^{s} U_{x_j}.$$

Set $N = \max\{n_{x_1}, \dots, n_{x_s}\}$.

On the other hand, observe that for every $y \in M_t \setminus \widetilde{\Lambda} \subset M_\delta \setminus \widetilde{\Lambda}$, it follows from T satisfying Hypothesis H2 and [57, Theorem 11.2.14] that there exists n_y such that $T^{n_y}(y) \notin \Lambda_\delta$. From continuity there exists an open neighbourhood V_y of y such that $T^{n_y}(V_y) \cap \Lambda_\delta = \emptyset$. Since $M_t \setminus B$ is compact, there exist $y_1, \dots, y_m$ with respective open neighbourhoods $V_{y_1}, \dots, V_{y_m}$ such that

$$M_t \setminus B \subset \bigcup_{i=1}^{m} V_{y_i}.$$

Set $N = \max\{n_{y_1}, \dots, n_{y_m}\}$. From continuity of $(x, \omega) \to T_\omega(x)$ we obtain that for every $x \in M_t$, either $T_\omega^n(x) \in \bigcup M_r$, for every $\omega \in \Omega_\varepsilon$ and some $n \leq N$; or $T_\omega^n(x) \notin \Lambda_\delta$, for every $\omega \in \Omega_\varepsilon$ and some $n \leq N$. In the first case, allowing return to M_t would join the equivalence classes $[C_t]$ of M_t with $[C_r]$ for some recurrent region M_r, contradicting transience. In the second case, once the process escapes Λ_δ it is killed. Thus, $\mathcal{P}_\varepsilon^n(x, M_t) = 0$ for every $n > N$. $\square$

Proposition 4.6 naturally motivates the following definition.

Definition 4.7. Fix $\varepsilon > 0$ such that the conclusions of Proposition 4.3 hold. Let $M_1, \dots, M_n$ be the sets introduced in Definition 4.4. We define the directed graph $\mathscr{G}_\delta = (V_\delta, E)$ in the following way:

- the set of vertices V_δ is given by $V_\delta := \{M_1, \dots, M_n\}$,
- given $M_i, M_j \in V_\delta$ we say that the edge $M_i \to M_j$ is in E_ε if $M_i \neq M_j$ and there exists $x \in M_i$ such that $\mathcal{P}_\varepsilon(x, M_j) > 0$.

Observe that using the same argument as in Proposition 4.3, the set of edges E does not depend on ε as long as this parameter is small enough.

Proposition 4.8. *Given a transient region $M_t \in V_\delta$ there exists a path in $\mathscr{G}_\delta$ connecting M_t to a recurrent region M_r. Moreover, the graph $\mathscr{G}_\delta$ is acyclic.*

Proof. To see the first part of the proposition, observe that there exists $x \in M_t \cap (\Lambda \setminus R)$. In this way, there exists $n \in \mathbb{N}$, such that $T^n(x) \in R$. Defining M_r as the unique recurrent region such that $T^n(x) \in M_r$, we obtain that there exists a path from M_t to M_r in the graph $\mathscr{G}_\delta$.

Finally, observe that if $\mathscr{G}_\delta$ had a cycle then this would contradict the maximality of the equivalence classes $[C_1], \dots, [C_n]$. $\square$

4.2. **Proof of the main (global) result.** Recall from Lemma 2.7 that $R = \sqcup_{i=1}^k R^i$. For every $i \in \{1, \dots, k\}$, consider the (deterministic) operator

$$\mathcal{L}_i : \mathcal{C}^0(R^i) \to \mathcal{C}^0(R^i)$$
$$f \mapsto \sum_{T(y)=x} \frac{e^{\phi(y)} f(y)}{|\det \mathrm{d}T(y)|},$$

and set $\lambda_i = r(\mathcal{L}_i)$.

Notation 4.9. Assume Hypothesis H2. Given a closed set $A \subset \Lambda_\delta$ we write:

- $\mathcal{P}_{A,\varepsilon} : L^\infty(A, \rho) \to L^\infty(A, \rho)$, $\mathcal{P}_{A,\varepsilon} f = \mathcal{P}_\varepsilon(\mathbb{1}_A \cdot f)$,
- $\mathcal{L}_{A,\varepsilon} : L^1(A, \rho) \to L^1(A, \rho)$, $\mathcal{L}_{A,\varepsilon} f = \mathcal{L}_\varepsilon(\mathbb{1}_A \cdot f)$, and

- for each vertex M_v of the graph $\mathscr{G}_\delta$ we define

$$\mathcal{L}_{M_v} : \mathcal{C}^0(M_v) \to \mathcal{C}^0(M_v)$$
$$f \mapsto \sum_{T(y)=x} \frac{e^{\phi(y)} f(y) \mathbb{1}_{M_v}(y)}{|\det \mathrm{d}T(x)|}.$$

  Note from Lemma 2.13 that this linear operator is well-defined.

Lemma 4.10. *Given a recurrent region M_v we have that*

$$r(\mathcal{P}_{M_v,\varepsilon}) \xrightarrow{\varepsilon\to 0} \lambda_{M_v} := \max\{\lambda_i;\ i \in \mathcal{I}_{M_v}\},$$

where $\mathcal{I}_{M_v} := \{i \in \{1, \dots, k\};\ R^i \subset M_v\}$.

Proof. We divide the proof into two steps.

Step 1. $\lambda_{M_v} \le \liminf_{\varepsilon\to 0} r(\mathcal{P}_{M_v,\varepsilon})$.

Observe that for every $i \in \mathcal{I}_{M_v}$ and every $f \in L^\infty(M_v)$, $\mathbb{1}_{R^i_\delta}\mathcal{P}_\varepsilon(\mathbb{1}_{R^i_\delta} \cdot f) \le \mathcal{P}_{M_v,\varepsilon} f$. From Theorem 2.10 and the above equation we obtain

$$\lambda_i = \lim_{\varepsilon\to 0} r(\mathcal{P}_{R^i_\delta,\varepsilon}) \le \liminf_{\varepsilon\to 0} r(\mathcal{P}_{M_v,\varepsilon}),$$

for every $i \in \mathcal{I}_{M_v}$.

Step 2. $\limsup_{\varepsilon\to\infty} r(\mathcal{P}_{M_v,\varepsilon}) \le \lambda_{M_v}$.

Repeating the same argumentation of Section 3, we obtain that:

(1) there exists $g_\varepsilon \in \ker(\mathcal{P}_{M_v,\varepsilon} - r(\mathcal{P}_{M_v,\varepsilon}))$ for some $g_\varepsilon(x) \in \mathcal{C}^0_+(M_v)$ and $\int g_\varepsilon \mathrm{d}\rho = 1$,
(2) $\ker(\mathcal{L}_{M_v,\varepsilon} - r(\mathcal{P}_{M_v,\varepsilon})) = \mathrm{span}(m_\varepsilon)$ for some $m_\varepsilon \in \mathcal{C}^\alpha(M_v)$ and $m_\varepsilon(x) > 0$ for every $x \in M_v$, and
(3) there exists a sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ satisfying $\varepsilon_n \to 0$, such that:
  - $r(\mathcal{P}_{M_v,\varepsilon_n}) \xrightarrow{n\to\infty} \lambda_0 = \limsup_{\varepsilon\to 0} r(\mathcal{P}_{M_v,\varepsilon})$,
  - $g_{\varepsilon_n}(x)\mathrm{d}x \xrightarrow{n\to\infty} \gamma(\mathrm{d}x)$ in the weak* topology and $\mathcal{L}^*_{M_v}\gamma = \lambda_0\gamma$, and
  - $m_{\varepsilon_n} \xrightarrow{n\to\infty} m$ in $\mathcal{C}^0(M_v)$ and $\mathcal{L}_{M_v} m = \lambda_0 m$.

It is clear that $\gamma(M_v \cap \Lambda) = 1$. Since $\Lambda = \bigcup_{n\in\mathbb{N}} T^{-n}(R)$, there exists $N \in \mathbb{N}$ such that $\gamma(M_v \cap T^{-N}(R)) > 0$. This implies that

$$\begin{aligned}
0 < \gamma(M_v \cap T^{-N}(R)) &= \frac{1}{\lambda_0^N}\int_{M_v\cap\Lambda} \mathcal{L}^N_{M_v}\mathbb{1}_{T^{-N}(R)}(x)\gamma(\mathrm{d}x)\\
&= \frac{1}{\lambda_0^N}\int_{M_v\cap\Lambda} \sum_{T^N(y)=x} \frac{e^{S_N\phi(y)}\mathbb{1}_R \circ T^N(y)}{|\det \mathrm{d}T^N(y)|}\gamma(\mathrm{d}x)\\
&= \frac{1}{\lambda_0^N}\int_{M_v\cap R} \sum_{T^N(y)=x} \frac{e^{S_N\phi(y)}\mathbb{1}_R \circ T^N(y)}{|\det \mathrm{d}T^N(y)|}\gamma(\mathrm{d}x),
\end{aligned}$$

where $S_N\phi(x) = \sum_{i=0}^{N-1} \phi \circ T^i(x)$, therefore $\gamma(M_v \cap R) > 0$. In this way, there exists $R^j \subset M_v$ such that $\gamma(R^j) > 0$. Define $\gamma_j(\mathrm{d}x) := \gamma(R^j \cap \mathrm{d}x)$. Given $f \in \mathcal{C}^0(R^j)$, we obtain that

$$\begin{aligned}
\mathcal{L}^*_j\gamma_j(f) = \gamma_j(\mathcal{L}_j f) &= \int_{R^j} \sum_{T(y)=x} \frac{e^{\phi(y)}\mathbb{1}_{R^j}(y) f(y)}{|\det \mathrm{d}T(y)|}\gamma(\mathrm{d}x)\\
&= \int_\Lambda \mathcal{L}(\mathbb{1}_{R^j} f)\gamma(\mathrm{d}x) = \lambda_0\gamma(\mathbb{1}_{R^j} f) = \lambda_0\gamma_j(f) \le r(\mathcal{L}_j)\gamma_j(f).
\end{aligned}$$

Since $r(\mathcal{L}_j) = \lambda_j$, this implies that

$$\lambda_0 = \limsup_{\varepsilon\to 0} r(\mathcal{P}_{M_v,\varepsilon}) \le \lambda_j \le \lambda_{M_v},$$

and conclude the proof. $\square$

Remark 4.11. Observe that from Theorem 2.14, item (3) of Hypothesis H2 is equivalent to the existence of $i \in \{1, \dots, k\}$ such that $\lambda_i > \max_{j \neq i} \lambda_j$.

Notation 4.12. If T satisfies Hypothesis H2, we define $\lambda_0 := \max\{\lambda_i;\ i \in \{1, \dots, k\}\}$. Let $i_0 \in \{1, \dots, k\}$ be the unique natural number such that $\lambda_{i_0} = \lambda_0$. We denote by M_0 the unique recurrent region such that $R^0 := R^{i_0} \subset M_0$.

Proposition 4.13. *Assume Hypothesis H2 and let ε be small enough. If $g \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$, then $\mathcal{P}_{M_0,\varepsilon}(\mathbb{1}_{M_0} g) = \lambda_\varepsilon \mathbb{1}_{M_0} g$. Moreover, for every vertex M_v of $\mathscr{G}_\delta$ such that there exists a path from M_0 to M_v, we have that $g|_{M_v} = 0$. Also, if $g|_{M_0} = 0$, then $g(x) = 0$ for every $x \in \Lambda_\delta$.*

Proof. First of all, observe that such a g exists from the Krein-Rutman Theorem (see e.g. [42, Theorem 4.1.4]).

Let
$$V_g := \{M_i;\ M_i \text{ is a vertex of } \mathscr{G}_\delta \text{ and } M_i \cap \{g \neq 0\} \neq \emptyset\}$$
and define $\mathscr{G}_g := (V_g, E_g) \subset \mathscr{G}_\delta$ as the maximal subgraph which contains the vertices V_g. Since $\mathscr{G}_g$ is acyclic, there exists a terminal vertex $M_f \in V_g$, i.e. no edge in $\mathscr{G}_g$ exits from M_f. We claim that $M_f = M_0$.

Observe that if $x \in M_f$ and $T_\omega(x) \in \{g \neq 0\}$ for some $\omega \in \Omega_\varepsilon$, then $T_\omega(x) \in M_f$. Indeed, if there exists $M_v \in V_g$ such that $T_\omega(x) \in M_v$, then $M_f \to M_v \in E_g$ but M_f is a terminal vertex. This shows the second part of the proposition for M_f. It remains to verify that $M_f = M_0$.

We claim that $\mathcal{P}_{M_f,\varepsilon}(\mathbb{1}_{M_f} g) = \lambda_\varepsilon \mathbb{1}_{M_f} g$. Indeed, for every $x \in M_f$ we obtain that
$$\begin{aligned}
\mathcal{P}_{M_f,\varepsilon}(\mathbb{1}_{M_f} g)(x) &= e^{\phi(x)} \mathbb{E}_\varepsilon[\mathbb{1}_{M_f} \circ T_\omega(x) \cdot g \circ T_\omega(x)] \\
&= e^{\phi(x)} \mathbb{E}_\varepsilon[\mathbb{1}_{M_f \cap \{g \neq 0\}} \circ T_\omega(x) \cdot g \circ T_\omega(x)] \\
&= e^{\phi(x)} \mathbb{E}_\varepsilon[\mathbb{1}_{\{g \neq 0\}} \circ T_\omega(x) \cdot g \circ T_\omega(x)] = \mathcal{P}_\varepsilon g(x) = \lambda_\varepsilon g(x).
\end{aligned}$$
Taking $\varepsilon \to 0$, from Lemma 4.10 and item (3) of Hypothesis H2 we obtain that $M_f = M_0$. □

Proposition 4.14. *Assume Hypothesis H2 and let ε be small enough. We have that, if $m \in \ker(\mathcal{L}_\varepsilon - \lambda_\varepsilon) \cap L^1_+(\Lambda_\delta)$, then $\mathcal{L}_{M_0,\varepsilon}(\mathbb{1}_{M_0} m) = \lambda_\varepsilon \mathbb{1}_{M_0} m$. Moreover, for every vertex M_v of $\mathscr{G}_\delta$ such that there exists a path from M_v to M_0, we have that $m|_{M_v} = 0$.*

Proof. Again, such an m exists from the Krein-Rutman Theorem [42, Theorem 4.1.4].

Analogous to the previous proof, let
$$V_m := \{M_i;\ M_i \text{ is a vertex of } \mathscr{G}_\delta \text{ and } M_i \cap \{m > 0\} \neq \emptyset\}$$
and define $\mathscr{G}_m \subset \mathscr{G}_\delta$ as the maximal subgraph which contains the vertices V_m. Since $\mathscr{G}_m$ is acyclic, there exists an initial vertex $M_s \in V_m$, i.e. no edge in $\mathscr{G}_m$ ends in M_s. We claim that $M_s = M_0$.

Observe that for every $x \in M_s$ and $\omega \in \Omega_\varepsilon$,
$$T_\omega^{-1}(M_s) \cap \{m > 0\} = T_\omega^{-1}(M_s) \cap M_s \cap \{m > 0\}.$$
This shows the second part of the proposition for M_s. It remains to show that $M_s = M_0$.

We claim that $\mathcal{L}_{M_s,\varepsilon}(\mathbb{1}_{M_s} m) = \lambda_\varepsilon \mathbb{1}_{M_s} m$. In fact, observe that for every $x \in M_s$
$$\mathcal{L}_{M_s,\varepsilon}(\mathbb{1}_{M_s} m)(x) = \mathbb{E}_\varepsilon\left[\sum_{T_\omega(y)=x} \frac{e^{\phi(y)} \mathbb{1}_{M_s}(y) m(y)}{|\det \mathrm{d}T_\omega(y)|}\right] = \lambda_\varepsilon \mathbb{1}_{M_s}(x) m(x).$$
Hence, from the choice of ε we obtain that $M_s = M_0$ and the result follows. □

Proposition 4.15. *Assume Hypothesis H2 and let $\varepsilon > 0$ be small enough. There exists $g_\varepsilon \in \mathcal{C}^0_+(\Lambda_\delta)$ and $m_\varepsilon \in L^1_+(\Lambda_\delta)$ such that:*

(1) $\ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) = \operatorname{span}(g_\varepsilon)$,

(2) $\ker(\mathcal{L}_\varepsilon - \lambda_\varepsilon) = \operatorname{span}(m_\varepsilon)$,
(3) $g_\varepsilon(x) > 0$ *for every* $x \in R^0$, *and*
(4) $\mathbb{1}_{M_0} m_\varepsilon \in \mathcal{C}^\alpha(M_0)$ *and* $m_\varepsilon(x) > 0$ *for every* $x \in M_0$.

Proof. From the same method provided in Theorem 3.2, there exists $\widetilde{m}_\varepsilon \in \mathcal{C}^\alpha(M_0)$ such that $\mathcal{L}_{M_0,\varepsilon}\widetilde{m}_\varepsilon = \lambda_\varepsilon \widetilde{m}_\varepsilon$ and $M_0 = \{\widetilde{m}_\varepsilon > 0\}$.

Given $g_\varepsilon \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$, from Proposition 4.13, we have that $\mathcal{P}_{M_0,\varepsilon}(\mathbb{1}_{M_0} g_\varepsilon) = \lambda_\varepsilon \mathbb{1}_{M_0} g_\varepsilon$. Since $M_0 = \{\widetilde{m}_\varepsilon > 0\}$, repeating the same argument as in Theorem 3.4, we obtain that

$$\mathbb{1}_{M_0} g_\varepsilon^{\pm} \in \ker(\mathcal{P}_{M_0,\varepsilon} - \lambda_\varepsilon). \tag{8}$$

We divide what is left of the proof into three steps.

Step 1. *For every* $\varepsilon > 0$ *sufficiently small, if* $\widetilde{g}_\varepsilon \in \ker(\mathcal{P}_{M_0,\varepsilon} - \lambda_\varepsilon)$, *then* $\widetilde{g}_\varepsilon \in \mathcal{C}^0(\Lambda_\delta)$ *and* $\widetilde{g}_\varepsilon(x) > 0$ *for every* $x \in R^0$.

Using the fact that $\mathcal{P}_{M_0,\varepsilon}$ is strong Feller and equation (8), assume by contradiction that there exists a sequence of positive numbers $\{\varepsilon_n\}_{n\in\mathbb{N}}$ such that $\varepsilon_n \to 0$, and for every $n \in \mathbb{N}$, there exists a non-negative function $\widetilde{g}_{\varepsilon_n} \in \ker(\mathcal{P}_{M_0,\varepsilon_n} - \lambda_{\varepsilon_n})$ such that $\widetilde{g}_{\varepsilon_n}(x_n) = 0$ for some $x_n \in R^0$.

From the same arguments presented in Steps 1 and 2 of Lemma 3.3 we have that if $\widetilde{g}_{\varepsilon_n}(x) = 0$ for some $x \in R^0$ then $\widetilde{g}_{\varepsilon_n}|_{R^0_\delta} = 0$. Again, as in the proof of Lemma 4.10, up to taking a subsequence of $\{\varepsilon_n\}_{n\in\mathbb{N}}$ we can assume that

(a) $r(\mathcal{P}_{M_0,\varepsilon_n}) \xrightarrow{n\to\infty} \lambda_0$,
(b) $\widetilde{g}_{\varepsilon_n}(x)\mathrm{d}x \xrightarrow{n\to\infty} \gamma(\mathrm{d}x)$ in the weak* topology and $\mathcal{L}^*_{M_0}\gamma = \lambda_0\gamma$, and
(c) $\widetilde{m}_{\varepsilon_n} \xrightarrow{n\to\infty} m_0$ in $\mathcal{C}^0(M_0)$ and $\mathcal{L}_{M_0} m_0 = \lambda_0 m_0$.

Observe that $\gamma(R^0_\delta) = 0$ by construction. Repeating the same computations in Step 2 of Lemma 4.10 (now with Λ instead of M) we obtain that there exists $R^j \subset M_0$ such that $\gamma(R^j) > 0$ and $\mathcal{L}^*_j\gamma(R^j \cap \mathrm{d}x) = \lambda_0\gamma(R^j \cap \mathrm{d}x)$, contradicting Hypothesis H2 since $r(\mathcal{L}_j) < \lambda_0$. Therefore, $\widetilde{g}_{\varepsilon_n}(x) > 0$ for every $x \in R^0$ and $n \in \mathbb{N}$.

Step 2. *We show that* $\dim\ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon) = 1$.

Let $g_1, g_2 \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$. Observe that from the same proof of Theorem 3.4, we obtain that there exists t_0 such that $(g_1 - t_0 g_2)|_{R^0} = 0$. Since $g_1 - t_0 g_2 \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$, we have from Step 1 that $\mathbb{1}_{M_0}(g_1 - t_0 g_2) = 0$. Finally, from Proposition 4.13 we obtain that $g_1 - t_0 g_2 = 0$.

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

From the Krein-Rutman theorem (see [42, Theorem 4.1.4]) and the fact that $\lambda_\varepsilon > 0$, we obtain that there exists $g_\varepsilon \in L^\infty_+(\Lambda_\delta)$ such that $\mathcal{P}_\varepsilon g_\varepsilon = \lambda_\varepsilon g_\varepsilon$, and since $\mathcal{P}_\varepsilon$ is strong Feller we obtain that $g_\varepsilon \in \mathcal{C}^0_+(\Lambda_\delta)$. Combining Steps 1 and 2, the fact that $\mathcal{L}^*_\varepsilon = \mathcal{P}_\varepsilon$ and choosing $m_\varepsilon \in L^1(\Lambda_\delta)$ such that $\mathbb{1}_{M_0} m_\varepsilon = \widetilde{m}_\varepsilon$ the result follows. □

Theorem 4.16. *Assume Hypothesis H2 and let* $\varepsilon > 0$ *be small enough. Let* $g_\varepsilon \in \ker(\mathcal{P}_\varepsilon - \lambda_\varepsilon)$ *and* $m_\varepsilon \in \ker(\mathcal{L}_\varepsilon - \lambda_\varepsilon)$ *be non-negative functions. Then,*

$$\nu_\varepsilon(\mathrm{d}x) = \frac{m_\varepsilon(x) g_\varepsilon(x)\rho(\mathrm{d}x)}{\int_{\Lambda_\delta} m_\varepsilon(y) g_\varepsilon(y)\rho(\mathrm{d}y)}$$

is the unique quasi-ergodic measure of the ϕ*-weighted Markov process* X^ε_n *on* $\{m_\varepsilon > 0\} \cap \{g_\varepsilon > 0\}$.

Moreover $\nu_\varepsilon \xrightarrow{\varepsilon\to 0} \nu_0$ *in the week* topology, where* ν_0 *is the unique equilibrium state for* T *for the potential* $\phi - \log|\det \mathrm{d}T|$ *supported on* Λ.

Proof. For every $\varepsilon > 0$ small enough, choose $g_\varepsilon \in \mathcal{C}^0_+(\Lambda_\delta)$ and $m_\varepsilon \in L^1_+(\Lambda_\delta)$ satisfying the conclusions of Proposition 4.15. Following the same strategy as in the proof of Lemma 3.5 we obtain that

$$\nu_\varepsilon(\mathrm{d}x) = \frac{g_\varepsilon(x)m_\varepsilon(x)\rho(\mathrm{d}x)}{\int_{M_0} g_\varepsilon(x)m_\varepsilon(x)\rho(\mathrm{d}x)},$$

is a quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on $\{g_\varepsilon m_\varepsilon > 0\}$. From Propositions 4.13 and 4.14 we obtain that $R^0 \subset \{g_\varepsilon m_\varepsilon > 0\} \subset M_0$, $\mathcal{P}_{M_0,\varepsilon}\mathbb{1}_{M_0} g_\varepsilon = \lambda_\varepsilon g_\varepsilon$ and $\mathcal{L}_{M_0,\varepsilon}\mathbb{1}_{M_0,\varepsilon} m_\varepsilon = \lambda_\varepsilon \mathbb{1}_{M_0} m_\varepsilon$. Repeating the proof of Proposition 3.10 and Theorem 2.10 changing R^i_δ to M_0 we obtain the last part of the result. □

We close this section proving Theorem 2.11.

Proof of Theorem 2.11. Items (1) to (5) follow directly from Propositions 4.1 and 4.15 and Theorem 4.16.

We divide the remaining of the proof into six steps.

Step 1. *If ν_0 is topologically mixing, then for every $\varepsilon > 0$ small enough, the operator*

$$\begin{aligned}\overline{\mathcal{P}}_\varepsilon : \mathcal{C}^0(M_\delta \setminus U) &\to \mathcal{C}^0(M_\delta \setminus U)\\ f &\mapsto e^{\phi(x)}\mathbb{E}_\varepsilon[f \circ T_\omega(x)\mathbb{1}_{M_\delta\setminus U} \circ T_\omega(x)]\end{aligned}$$

satisfies the following properties:

(1) $\overline{\mathcal{P}}_\varepsilon$ is a strong Feller operator, therefore $\overline{\mathcal{P}}^2_\varepsilon$ is a compact operator,
(2) $r(\overline{\mathcal{P}}_\varepsilon) = r(\mathcal{P}_\varepsilon) = \lambda_\varepsilon$,
(3) there exists a probability measure $\overline{\mu}_\varepsilon$ on $M_\delta\setminus U$ such that $\operatorname{span}\{\overline{\mu}_\varepsilon\} = \ker(\overline{\mathcal{P}}^*_\varepsilon - \lambda_\varepsilon)$ and $\overline{\mu}_\varepsilon|_{\Lambda_\delta}/\overline{\mu}_\varepsilon(\Lambda_\delta) = \mu_\varepsilon$, with $\mu_\varepsilon(\mathrm{d}x) := m_\varepsilon(x)\mathrm{d}x$ given by Proposition 4.15, and
(4) $\operatorname{span}\{\overline{g}_\varepsilon\} = \ker(\overline{\mathcal{P}}_\varepsilon - \lambda_\varepsilon)$ where $\overline{g}_\varepsilon := \mathbb{1}_{\Lambda_\delta} g_\varepsilon \in \mathcal{C}^0(M_\delta \setminus U)$, with g_ε given by Proposition 4.15 and $\int \overline{g}_\varepsilon \mathrm{d}\overline{\mu}_\varepsilon = 1$.

Observe that the strong Feller property of $\overline{\mathcal{P}}_\varepsilon$ follows by the same computations provided in Theorem 2.17, showing (1). Item (2) follows since $\overline{\mathcal{P}}_\varepsilon$ is strong Feller, then

$$r(\overline{\mathcal{P}}_\varepsilon : \mathcal{C}^0(M_\delta \setminus U) \to \mathcal{C}^0(M_\delta \setminus U)) = r(\mathcal{P}_\varepsilon : L^\infty(\Lambda_\delta, \rho) \to L^\infty(\Lambda_\delta, \rho)) = \lambda_\varepsilon.$$

Finally, (3) and (4) are direct consequences of Propositions 4.1, 4.2 and 4.15.

Step 2. *The operator $\frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon$ is power-bounded, i.e. $\sup_{n\in\mathbb{N}} \|\frac{1}{\lambda^n_\varepsilon}\overline{\mathcal{P}}^n_\varepsilon\| < \infty$.*

Repeating the same argumentation of the proof Proposition 4.1 (2). There exists $N > 0$ such that $\mathcal{P}^N_\varepsilon f(x) = 0$ for every $x \in M_\delta \setminus \Lambda_\delta$ and $f \in \mathcal{C}^0(M_\delta \setminus U)$. In this way, for every $n > 0$, we obtain that for every

$$\frac{1}{\lambda^{n+N}_\varepsilon}\overline{\mathcal{P}}^{N+n}_\varepsilon f = \mathbb{1}_{\Lambda_\delta}\frac{1}{\lambda^{n+N}_\varepsilon}\mathcal{P}^n_\varepsilon\left(\mathbb{1}_{\Lambda_\varepsilon}\overline{\mathcal{P}}^N_\varepsilon f\right).$$

Since $\frac{1}{\lambda^n_\varepsilon}\mathcal{P}^n_\varepsilon$ is power-bounded we obtain the result.

Step 3. *Given a function $f \in \mathcal{C}^0(M_\delta\setminus U, \mathbb{C})$ let us define $|f| \in \mathcal{C}^0(M_\delta\setminus U)$ as the function $x \mapsto \|f(x)\|_{\mathbb{C}}$. Let $\alpha \geq 0$ and $f_\varepsilon \in \mathcal{C}^0(M_\delta \setminus U, \mathbb{C})$ be such that $\frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon f_\varepsilon = e^{i\alpha} f_\varepsilon$. Then for every $x \in \operatorname{supp}\overline{\mu}_\varepsilon$, we have that $|f_\varepsilon|(x) = \overline{g}_\varepsilon(x)\int |f_\varepsilon|\mathrm{d}\overline{\mu}_\varepsilon$ and $\int |f_\varepsilon|\mathrm{d}\overline{\mu}_\varepsilon > 0$.*

It follows that

$$|f_\varepsilon| = |e^{i\alpha} f_\varepsilon| = \left|\frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon f_\varepsilon\right| \leq \frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon|f_\varepsilon|,$$

therefore, for every $n \in \mathbb{N}$ we obtain

$$|f_\varepsilon| \leq \frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon|f_\varepsilon| \leq \frac{1}{\lambda^2_\varepsilon}\overline{\mathcal{P}}^2_\varepsilon|f_\varepsilon| \leq \dots \leq \frac{1}{\lambda^n_\varepsilon}\overline{\mathcal{P}}^n_\varepsilon|f_\varepsilon|.$$

Since $\overline{\mathcal{P}}_\varepsilon^2$ is a compact operator and $\frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon$ is power-bounded from Step 2, the above sequence is monotone and bounded. Hence, there exists $g \in \mathcal{C}^0(M_\delta \setminus U)$ such that $\frac{1}{\lambda_\varepsilon^n}\overline{\mathcal{P}}_\varepsilon^n|f_\varepsilon| \xrightarrow{n\to\infty} g$ in $\mathcal{C}^0(M_\delta \setminus U)$. It follows that $g \in \ker(\overline{\mathcal{P}}_\varepsilon - \lambda_\varepsilon) = \operatorname{span}\{\overline{g}_\varepsilon\}$. From $0 \le |f_\varepsilon| \ne 0$, we obtain that there exists $a > 0$ such that $g = a\overline{g}_\varepsilon$. Finally, since $|f_\varepsilon| \le g$, both functions are continuous, and their integrals with respect to $\overline{\mu}_\varepsilon$ coincide, i.e.

$$\int_M |f_\varepsilon|\, \mathrm{d}\overline{\mu}_\varepsilon = \int_M g\, \mathrm{d}\overline{\mu}_\varepsilon = \int_M a\overline{g}_\varepsilon\, \mathrm{d}\overline{\mu}_\varepsilon,$$

it follows that $|f_\varepsilon|(x) = \overline{g}_\varepsilon(x)\int_M |f_\varepsilon| \mathrm{d}\overline{\mu}_\varepsilon$ for every $x \in \operatorname{supp}\overline{\mu}_\varepsilon$.

Step 4. *The operator $\overline{\mathcal{P}}_\varepsilon$ has the spectral gap property, i.e. there exist a $\overline{\mathcal{P}}_\varepsilon$-invariant closed space $W \subset \mathcal{C}^0(M_\delta \setminus U)$ such that $\mathcal{C}^0(M_\delta \setminus U) = \operatorname{span}\{g_\varepsilon\} \oplus W$ and $r(\overline{\mathcal{P}}_\varepsilon|_W) < \lambda_\varepsilon$.*

Since $\overline{\mathcal{P}}_\varepsilon^2$ is a compact operator and $\frac{1}{\lambda_\varepsilon^n}\overline{\mathcal{P}}_\varepsilon^n$ is power-bounded, it is enough to show that $\sigma_{\mathrm{per}}(\overline{\mathcal{P}}_\varepsilon) \cap \lambda_\varepsilon\mathbb{S}^1 = \{\lambda_\varepsilon\}$ (see details in the proof of Lemma A.5). Choose $\alpha \in [0, 2\pi)$ such that $e^{i\alpha}\lambda_\varepsilon \in \sigma(\overline{\mathcal{P}}_\varepsilon)$. Then, there exists $f_\varepsilon \in \mathcal{C}^0(M_\delta \setminus U, \mathbb{C})$ such that $\frac{1}{\lambda_\varepsilon}\overline{\mathcal{P}}_\varepsilon f_\varepsilon = e^{i\alpha} f_\varepsilon$. From Step 3, we can assume without loss of generality that $\int |f_\varepsilon| \mathrm{d}\overline{\mu}_\varepsilon = 1$. Using again Step 3 and Propositions 4.13 and 4.14 we have that, there exists a continuous function $\theta : \{g_\varepsilon > 0\} \to \mathbb{R}$ such that $f_\varepsilon(x) = \overline{g}_\varepsilon(x)e^{i\theta(x)} = g_\varepsilon(x)e^{i\theta(x)}$ for every $x \in M_0$ and

$$\frac{1}{\lambda_\varepsilon}\mathcal{P}_{M_0,\varepsilon}(\mathbb{1}_{M_0} f_\varepsilon) = e^{i\alpha}\mathbb{1}_{M_0} f_\varepsilon.$$

In this way, for every $n \in \mathbb{N}$ and $x \in M_0$

$$e^{i(\theta(x)+n\alpha)}g_\varepsilon(x) = \int_{M_0} e^{i\theta(y)} g_\varepsilon(y)\frac{1}{\lambda_\varepsilon^n}(\mathcal{P}_{M_0,\varepsilon}^n)^*(\delta_x)(\mathrm{d}y),$$

which implies that

$$g_\varepsilon(x) = \int_{M_0} e^{i(\theta(y)-\theta(x)-n\alpha)} g_\varepsilon(y)\frac{1}{\lambda_\varepsilon^n}(\mathcal{P}_{M_0,\varepsilon}^n)^*(\delta_x)(\mathrm{d}y).$$

Since

$$g_\varepsilon(x) = \int_{M_0} g_\varepsilon(y)\frac{1}{\lambda_\varepsilon^n}(\mathcal{P}_{M_0,\varepsilon}^n)^*(\delta_x)(\mathrm{d}y),$$

we obtain that $e^{i(\theta(y)-\theta(x)-n\alpha)} = 1$ for every $y \in \operatorname{supp}\{(\mathcal{P}_{M_0,\varepsilon}^n)^*(\delta_x)\} \cap \{g_\varepsilon > 0\}$. By hypothesis, the measure ν_0 is mixing for the map $T : R^0 \to R^0$, so $T : R^0 \to R^0$ is topologically mixing and therefore topologically exact. Hence, there exists $n_0 \in \mathbb{N}$ such that

$$R^0 \subset \operatorname{supp}\{(\mathcal{P}_{M_0,\varepsilon}^n)^*(\delta_x)\} \cap \{g_\varepsilon > 0\}, \text{ for every } n > n_0.$$

This implies that $e^{i(\theta(x)-\theta(y)-n\alpha)} = 1$ for every $n > n_0$ and $x, y \in R^0$, so $\alpha = 0$.

Step 5. *We show that $\nu_\varepsilon(\mathrm{d}x) = g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x) = \overline{g}_\varepsilon(x)\overline{\mu}_\varepsilon(\mathrm{d}x)/\int \overline{g}_\varepsilon(y)\overline{\mu}_\varepsilon(\mathrm{d}y)$ is a quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on $M_\delta \setminus U$.*

From Step 1 and Propositions 4.1, 4.2 and 4.15 it is clear that $\nu_\varepsilon(\mathrm{d}x) = g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x) = \overline{g}_\varepsilon(x)\overline{\mu}_\varepsilon(\mathrm{d}x)/\int \overline{g}_\varepsilon(y)\overline{\mu}_\varepsilon(\mathrm{d}y)$. From Steps 3 and 4 we obtain that for every bounded and measurable function $h : M_\delta \setminus U \to \mathbb{R}$,

$$\frac{1}{\lambda^n}\overline{\mathcal{P}}_\varepsilon^n h \xrightarrow{n\to\infty} \overline{g}_\varepsilon \int_{M_\delta\setminus U} h(y)\overline{\mu}_\varepsilon(\mathrm{d}y) \text{ in } \mathcal{C}^0(M_\delta \setminus U),$$

since $\overline{\mathcal{P}}_\varepsilon h \in \mathcal{C}^0(M_\delta \setminus U)$.

Recall that $\tau^\phi = \min\{n; X_n^\varepsilon \in (E \setminus M_\delta) \cup U\}$ and by construction of the operator $\overline{\mathcal{P}}_\varepsilon$, for every $x \in \{g_\varepsilon > 0\} \cap \operatorname{supp}\mu_\varepsilon = \{\overline{g}_\varepsilon > 0\} \cap \operatorname{supp}\overline{\mu}_\varepsilon$ and for every $n \in \mathbb{N}$

$$\mathbb{E}_x^\phi\left[\frac{1}{n}\sum_{i=0}^{n-1} h \circ X_i^\varepsilon \,\middle|\, \tau^\phi > n\right] = \frac{\lambda_\varepsilon^n}{\overline{\mathcal{P}}_\varepsilon^n \mathbb{1}_{M_\delta\setminus U}(x)}\frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_\varepsilon^i}\overline{\mathcal{P}}_\varepsilon^i\left(h\frac{1}{\lambda_\varepsilon^{n-i}}\overline{\mathcal{P}}_\varepsilon^{n-i}\mathbb{1}_{M_\delta\setminus U}\right)(x).$$

Since $\frac{1}{\lambda_\varepsilon^n}\overline{\mathcal{P}}_\varepsilon^n \mathbb{1}_{M_\delta\setminus U}(x) \xrightarrow{n\to\infty} \overline{g}_\varepsilon(x)$, it is enough to show that

$$\frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_\varepsilon^i}\overline{\mathcal{P}}_\varepsilon^i\left(h\frac{1}{\lambda_\varepsilon^{n-i}}\overline{\mathcal{P}}_\varepsilon^{n-i}\mathbb{1}_{M_\delta\setminus U}\right)(x) \xrightarrow{n\to\infty} \overline{g}_\varepsilon(x)\int h(y)\overline{g}_\varepsilon(y)\overline{\mu}_\varepsilon(\mathrm{d}y).$$

This holds true since

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_\varepsilon^i}\overline{\mathcal{P}}_\varepsilon^i\left(h\frac{1}{\lambda_\varepsilon^{n-i}}\overline{\mathcal{P}}_\varepsilon^{n-i}\mathbb{1}_{M_\delta\setminus U}\right)(x) =&\frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_\varepsilon^i}\overline{\mathcal{P}}_\varepsilon^i\left(h\left(\frac{1}{\lambda_\varepsilon^{n-i}}\overline{\mathcal{P}}_\varepsilon^{n-i}\mathbb{1}_{M_\delta\setminus U}-\overline{g}_\varepsilon\right)\right)(x)\\ &+\frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_\varepsilon^i}\overline{\mathcal{P}}_\varepsilon^i\left(h\overline{g}_\varepsilon\right)(x),\end{aligned}$$

and

$$\frac{1}{\lambda_\varepsilon^i}\overline{\mathcal{P}}_\varepsilon^i\left(h\overline{g}_\varepsilon\right)(x) \xrightarrow{i\to\infty} \overline{g}_\varepsilon(x)\int h(y)\overline{g}_\varepsilon(y)\overline{\mu}_\varepsilon(\mathrm{d}y).$$

Step 6. *We conclude the proof of the theorem.*

To conclude, we need to show that if $\operatorname{supp}\nu_0 = R^0 \subset \operatorname{Int}(M\setminus U)$, then ν_ε is a quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on $M\setminus U$. Redefine the operator $\overline{\mathcal{P}}_\varepsilon$ as

$$\begin{aligned}\overline{\mathcal{P}}_\varepsilon : \mathcal{C}^0(M\setminus U) &\to \mathcal{C}^0(M\setminus U)\\ f &\mapsto e^{\phi(x)}\mathbb{E}_\varepsilon[f\circ T_\omega(x)\cdot \mathbb{1}_{M\setminus U}\circ T_\omega(x)].\end{aligned}$$

Observe that since $R^0 \subset \operatorname{Int}(M\setminus U)$, we can choose $\delta > 0$ small enough such that $M_0 \subset M\setminus U$. Repeating Steps 1, 2, 3 and 4 we obtain that

(1) $\overline{\mathcal{P}}_\varepsilon$ is a strong Feller operator,
(2) $r(\overline{\mathcal{P}}_\varepsilon) = r(\mathcal{P}_\varepsilon) = \lambda_\varepsilon$,
(3) there exists a probability measure $\overline{\mu}_\varepsilon$ on $M_\delta\setminus U$ such that $\operatorname{span}\{\overline{\mu}_\varepsilon\} = \ker(\overline{\mathcal{P}}_\varepsilon^*-\lambda_\varepsilon)$ and $\overline{\mu}_\varepsilon|_{M_0}/\overline{\mu}_\varepsilon(M_0) = \mu_\varepsilon|_{M_0}/\mu_\varepsilon(M_0)$.
(4) $\operatorname{span}\{\overline{g}_\varepsilon\} = \ker(\overline{\mathcal{P}}_\varepsilon - \lambda_\varepsilon)$ and $\mathbb{1}_{M_0}\overline{g}_\varepsilon = \mathbb{1}_{M_0}g_\varepsilon$, with g_ε given by Proposition 4.15 and $\int \overline{g}_\varepsilon \mathrm{d}\overline{\mu}_\varepsilon = 1$.
(5) $\overline{\mathcal{P}}_\varepsilon : \mathcal{C}^0(M\setminus U) \to \mathcal{C}^0(M\setminus U)$ has the spectral gap property.

As in Step 5, we obtain that $\nu_\varepsilon(\mathrm{d}x) = g_\varepsilon(x)\mu_\varepsilon(\mathrm{d}x) = \overline{g}_\varepsilon(x)\overline{\mu}_\varepsilon(\mathrm{d}x)$ is a quasi-ergodic measure of the ϕ-weighted Markov process X_n^ε on $M\setminus U$. □

5. Examples

5.1. **The logistic map.** Consider the Markov process $X_{n+1}^\varepsilon = T(X_n^\varepsilon)+\omega_n$, $n\in\mathbb{N}$, with $T(x) = ax(1-x)$ and $\omega_n \sim \operatorname{Unif}(-\varepsilon,\varepsilon)$. Fix $a = 3.83$ so that the deterministic dynamical system (with $\varepsilon = 0$) has an almost sure global three-periodic attractor [25, 52], i.e. Lebesgue almost every initial condition in $[0,1]$ is attracted to the unique three-periodic hyperbolic attractor $\mathcal{A} = \{p, T(p), T^2(p)\}$, with $p \approx 0.1456149$ (see Remark 2.6).

The dynamical decomposition of Lemma 2.7 yields two invariant sets: the origin $R^1 = \{0\}$, and a hyperbolic Cantor set R^2 consisting of the closure of all periodic points in $(0,1)$ that are not in the basin of attraction $B(T)$ of $\mathcal{A}$ [56]. Let $\Lambda := [0,1]\setminus B(T)$ and $U \supset \mathcal{A}$ be a small enough neighbourhood of the attractor such that $U\cap\Lambda = \emptyset$. We consider the family of α-Hölder potentials $\phi_t : [0,1] \to \mathbb{R}, x \mapsto (-t+1)\log|T'(x)|$ for $t \geq 0$. Recall that an equilibrium state ν_i associated with the potential $\phi_t - \log|a(1-2x)|$ for T on R^i is a measure maximising

$$\mu \mapsto h_\mu(T) + \int(\phi_t - \log|T'|)\mathrm{d}\mu = h_\mu(T) - t\int \log|T'|\mathrm{d}\mu,$$

where h_μ is the metric entropy and $\mu \in \mathcal{I}(T, R^i)$, the set of T-invariant measures on R^i.

It is well known that Λ is a hyperbolic (uniformly expanding) invariant set [29] and T admits a unique equilibrium state associated with the potential $\phi_t(x) - \log|T'(x)|$ on Λ (see e.g. [57, Chapters 11 and 12]). Therefore, Hypothesis H2 is satisfied and we can apply the theory developed above. For R^1, it is clear that $\nu_1 = \delta_0$ and $P(T, \phi_t - \log|T'|, R^1) = -t\log|a|$. For R^2,

$$P(T, \phi_t - \log|T'|, R^2) = h_{\nu_2}(T) - t\int \log|a(1-2x)|\nu_2(\mathrm{d}x) > -t\log|a|,$$

since $-\log|a(1-2x)|$ reaches its minimum at 0 and $h_{\nu_2} = (1+\sqrt{5})/2$ (see [56] for precise details). Therefore,

$$\log\lambda_\varepsilon = \log r(\mathcal{P}_\varepsilon) \xrightarrow{\varepsilon\to 0} P(T, \phi_t - \log|T'|, \Lambda) = P(T, \phi_t - \log|T'|, R^2) = \log\lambda_2,$$

with $\mathcal{P}_\varepsilon : L^\infty([0,1]\setminus U) \to L^\infty([0,1]\setminus U)$ the global annealed Koopman operator. It follows from Theorem 2.11 that the unique equilibrium state sits on the invariant Cantor set repeller R^2 and can be approximated by quasi-ergodic measure ν_ε of the ϕ_t-weighted Markov process X_n^ε on a neighbourhood of R^2, as $\varepsilon \to 0$.

For the particular choice of $t = 1$, i.e. $\phi_{t=1} = 0$, the systems is no longer spatially weighted and we recover the so-called "natural measure" of the repeller [34]. We note that the relationship between limiting quasi-ergodic measures and natural measures in the case of the zero weighting was previously discussed in [5] for this example.

Finally, consider the potential $\phi_0(x) = \log|T'|$. The topological pressure of the deterministic system on Λ is given by $P(T, 0, \Lambda) = h_\nu(T)$, where ν is the unique equilibrium state. Since this measure maximises $P(T, 0, \Lambda)$, it coincides with the measure of maximal entropy of the system.

5.2. **The complex quadratic map.** Similarly to the previous example, let us consider random perturbations of iterates of the complex quadratic map $p_c(z) = z^2 + c$, $c \in \mathbb{C}$, acting on the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \sqcup \{\infty\}$. As before, we study the Markov process $X_{n+1}^\varepsilon = p_c(X_n^\varepsilon) + \omega_n$, where $\{\omega_n\}_n$ are i.i.d. random variables uniformly distributed on $\{a + ib \in \mathbb{C};\ (a,b) \in [-\varepsilon, \varepsilon]^2\}$, with $\varepsilon > 0$ small enough.

Consider the Julia set $J \subset \widehat{\mathbb{C}}$ associated with the polynomial p_c. Recall that J is the closure of the set of repelling periodic points [43, Theorem 11.1]. The set J is non-empty, compact, and totally invariant, meaning that $J = p_c(J) = p_c^{-1}(J)$ (see [43, Lemma 3.1]). Now, let c be a hyperbolic complex number within the Mandelbrot set, which ensures that J is hyperbolic, i.e., J is connected and satisfies $\|p_c'(z)\| = \|2z\| > 1$ for every $z \in J$.

In this context, it is readily verified that p_c admits a finite attractor $\mathcal{A} \subset \mathbb{C}$. Moreover, for any α-Hölder potential $\phi : \mathbb{C} \to \mathbb{R}$, p_c satisfies Hypothesis H2 with $T = p_c$, $\Lambda = J$ and $E = M = \widehat{\mathbb{C}}$. Furthermore, notice that the unique equilibrium state of p_c for the potential $\phi - \log|\det \mathrm{d}p_c|$ on J is mixing.

Finally, from Theorem 2.11, for any α-Hölder function $\psi : \widehat{\mathbb{C}} \to \mathbb{R}$, the unique p_c-invariant equilibrium state for the potential ψ on J, can be approximated in the weak* topology by quasi-ergodic measures of the $(\psi + \log|\det \mathrm{d}p_c|)$-weighted Markov process X_n^ε on $\widehat{\mathbb{C}} \setminus U$, where U is a neighbourhood of $\mathcal{A}$ such that $U \cap J = \emptyset$.

Acknowledgements

We would like to thank Matteo Tabaro and Martin Rasmussen for valuable comments and insightful discussions, as well as Mark Demers and Alex Blumenthal for their feedback and suggestions. The authors gratefully acknowledge support from the EPSRC Centre for Doctoral Training in Mathematics of Random Systems: Analysis, Modelling and Simulation (EP/S023925/1). MMC has been supported by an Imperial College President's PhD scholarship. JSWL has been supported by the EPSRC (EP/Y020669/1) and thanks IRCN (Tokyo) and GUST (Kuwait) for their research support.

References

[1] J. F. Alves, V. Araújo, and C. H. Vásquez. Stochastic stability of non-uniformly hyperbolic diffeomorphisms. *Stoch. Dyn.*, 7(3):299–333, 2007.

[2] V. Araújo. Attractors and time averages for random maps. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 17(3):307–369, 2000.

[3] V. Baladi. *Dynamical zeta functions and dynamical determinants for hyperbolic maps. A functional approach.*, volume 68 of *Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics*. Springer, Cham, 2018.

[4] V. Baladi and M. Viana. Strong stochastic stability and rate of mixing for unimodal maps. *Ann. Sci. École Norm. Sup.*, 29(4):483–517, 1996.

[5] B. Bassols-Cornudella and J. S. W. Lamb. Noise-induced chaos: a conditioned random dynamics perspective. *Chaos*, 33(12):Paper No. 121102, 7, 2023.

[6] M. Benedicks and M. Viana. Random perturbations and statistical properties of Hénon-like maps. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 23(5):713–752, 2006.

[7] M. Benedicks and L.-S. Young. Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. *Ergodic Theory Dynam. Systems*, 12(1):13–37, 1992.

[8] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1975.

[9] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math.*, 29(3):181–202, 1975.

[10] L. A. Breyer and G. O. Roberts. A quasi-ergodic theorem for evanescent processes. *Stochastic Process. Appl.*, 84(2):177–186, 1999.

[11] A. Brunel and D. Revuz. Quelques applications probabilistes de la quasi-compacité. *Ann. Inst. H. Poincaré Sect. B*, 10(3):301–337, 1974.

[12] M. M. Castro, D. Chemnitz, H. Chu, M. Engel, J. S. W. Lamb, and M. Rasmussen. The Lyapunov spectrum for conditioned random dynamical systems, 2022. arXiv:2204.04129v1. To appear in Annales de l'Institut Henri Poincaré.

[13] M. M. Castro, V. P. H. Goverse, J. S. W. Lamb, and M. Rasmussen. On the quasi-ergodicity of absorbing markov chains with unbounded transition densities, including random logistic maps with escape. *Ergodic Theory and Dynamical Systems*, pages 1–38, 2023.

[14] M. M. Castro, J. S. W. Lamb, G. Olicón Méndez, and M. Rasmussen. Existence and uniqueness of quasi-stationary and quasi-ergodic measures for absorbing Markov processes: a Banach lattice approach, 2021. arXiv:2111.13791v6. To appear in Stochastic Processes and their Applications.

[15] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5):1926–1969, 2009.

[16] N. Champagnat, E. Strickler, and D. Villemonais. Uniform Wasserstein convergence of penalized Markov processes, 2023. arXiv:2306.16051v1.

[17] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields*, 164(1-2):243–283, 2016.

[18] N. Champagnat and D. Villemonais. Uniform convergence of penalized time-inhomogeneous Markov processes. *ESAIM Probab. Stat.*, 22:129–162, 2018.

[19] N. Champagnat and D. Villemonais. Quasi-limiting estimates for periodic absorbed markov chains, 2022. arXiv:2211.02706v1.

[20] V. Climenhaga and Y. Pesin. Building thermodynamics for non-uniformly hyperbolic maps. *Arnold Math. J.*, 3(1):37–82, 2017.

[21] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions. Markov chains, diffusions and dynamical systems*. Probability and its Applications (New York). Springer, Heidelberg, 2013.

[22] P. Collet, S. Martínez, and B. Schmitt. The Yorke-Pianigiani measure and the asymptotic law on the limit Cantor set of expanding systems. *Nonlinearity*, 7(5):1437–1443, 1994.

[23] F. Colonius and M. Rasmussen. Quasi-ergodic limits for finite absorbing Markov chains. *Linear Algebra Appl.*, 609:253–288, 2021.

[24] S. Crovisier and R. Potrie. Introduction to partial hyperbolicity. *Trieste Lecture notes (2015), available at* http://www.cmat.edu.uy/rpotrie/documentos/pdfs/. Last accessed: 17/02/2024.

[25] J. P. Crutchfield, J. D. Farmer, and B. A. Huberman. Fluctuations and simple chaotic dynamics. *Phys. Rep.*, 92(2):45–82, 1982.

[26] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probability*, 2:88–100, 1965.

[27] R. De Leo and J. A. Yorke. The graph of the logistic map is a tower. *Discrete and Continuous Dynamical Systems*, 41(11):5243–5269, 2021.

[28] R. De Leo and J. A. Yorke. Infinite towers in the graphs of many dynamical systems. *Nonlinear Dynamics*, 105(1):813–835, 2021.

[29] W. de Melo and S. van Strien. *One-dimensional dynamics*, volume 25 of *Results in Mathematics and Related Areas (3)*. Springer-Verlag, Berlin, 1993.

[30] P. Del Moral. *Feynman-Kac formulae. Genealogical and interacting particle systems with applications.* Probability and its Applications (New York). Springer-Verlag, New York, 2004.
[31] M. Engel, J. S. W. Lamb, and M. Rasmussen. Conditioned Lyapunov exponents for random dynamical systems. *Trans. Amer. Math. Soc.*, 372(9):6343–6370, 2019.
[32] J. Glück. On the peripheral spectrum of positive operators. *Positivity*, 20(2):307–336, 2016.
[33] J. J. Grobler. Spectral theory in Banach lattices. In *Operator theory in function spaces and Banach lattices*, volume 75 of *Oper. Theory Adv. Appl.*, pages 133–172. Birkhäuser, Basel, 1995.
[34] H. Kantz and P. Grassberger. Repellers, semi-attractors, and long-lived chaotic transients. *Phys. D*, 17(1):75–86, 1985.
[35] J. I. Kifer. Some theorems on small random perturbations of dynamical systems. *Uspehi Mat. Nauk*, 29(3(177)):205–206, 1974.
[36] D. Kim, T. Tagawa, and A. Velleret. Quasi-ergodic theorems for Feynman-Kac semigroups and large deviation for additive functionals, 2024. arXiv:2401.17997v1.
[37] A. N. Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk SSSR (N.S.)*, 119:861–864, 1958.
[38] E. Kreyszig. *Introductory functional analysis with applications.* John Wiley & Sons, New York-London-Sydney, 1978.
[39] Y.-C. Lai and T. Tél. *Transient chaos. Complex dynamics on finite-time scales*, volume 173 of *Applied Mathematical Sciences.* Springer, New York, 2011.
[40] A. Lasota and M. C. Mackey. *Chaos, fractals, and noise. Stochastic aspects of dynamics*, volume 97 of *Applied Mathematical Sciences.* Springer-Verlag, New York, second edition, 1994.
[41] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics.* Springer, New York, second edition, 2013.
[42] P. Meyer-Nieberg. *Banach lattices.* Universitext. Springer-Verlag, Berlin, 1991.
[43] J. Milnor. *Dynamics in one complex variable*, volume 160 of *Annals of Mathematics Studies.* Princeton University Press, Princeton, NJ, third edition, 2006.
[44] T. Oikhberg and V. G. Troitsky. A theorem of Krein revisited. *Rocky Mountain J. Math.*, 35(1):195–210, 2005.
[45] J. Palis and W. de Melo. *Geometric theory of dynamical systems. An introduction.* Springer-Verlag, New York-Berlin, 1982.
[46] G. Pianigiani and J. A. Yorke. Expanding maps on sets which are almost invariant. Decay and chaos. *Trans. Amer. Math. Soc.*, 252:351–366, 1979.
[47] D. Revuz. *Markov chains*, volume 11 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam, second edition, 1984.
[48] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales*, volume 1 of *Wiley Series in Probability and Mathematical Statistics.* John Wiley & Sons, Ltd., Chichester, second edition, 1994.
[49] D. Ruelle. A measure associated with axiom-A attractors. *Amer. J. Math.*, 98(3):619–654, 1976.
[50] D. Ruelle. *Thermodynamic formalism*, volume 5 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Co., Reading, MA, 1978.
[51] D. Ruelle. The thermodynamic formalism for expanding maps. *Comm. Math. Phys.*, 125(2):239–262, 1989.
[52] Y. Sato, T. S. Doan, J. S. W. Lamb, and M. Rasmussen. Dynamical characterization of stochastic bifurcations in a random logistic map, 2018. arXiv:1811.03994v1.
[53] M. Shub. *Global stability of dynamical systems.* Springer-Verlag, New York, 1987.
[54] J. G. Sinaĭ. On the notion of entropy for a dynamic system. *Dokl. Akad. Nauk SSSR*, 124:768–771, 1959.
[55] J. G. Sinaĭ. Gibbs measures in ergodic theory. *Uspehi Mat. Nauk*, 27(4(166)):21–64, 1972.
[56] S. Smale and R. F. Williams. The qualitative analysis of a difference equation of population growth. *J. Math. Biol.*, 3(1):1–4, 1976.
[57] M. Viana and K. Oliveira. *Foundations of ergodic theory*, volume 151 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2016.
[58] K. Yosida. Asymptotic almost periodicities and ergodic theorems. *Proc. Imp. Acad. Tokyo*, 15:255–259, 1939.
[59] K. Yosida. Quasi-completely-continuous linear functional operations. *Japan. J. Math.*, 15:297–301, 1939.
[60] L.-S. Young. What are SRB measures, and which dynamical systems have them? *J. Statist. Phys.*, 108(5-6):733–754, 2002.
[61] J. Zhang, S. Li, and R. Song. Quasi-stationarity and quasi-ergodicity of general Markov processes. *Sci. China Math.*, 57(10):2013–2024, 2014.
[62] H. Zmarrou and A. J. Homburg. Bifurcations of stationary measures of random diffeomorphisms. *Ergodic Theory Dynam. Systems*, 27(5):1651–1692, 2007.

Appendix A. Quasi-ergodic measures for a class of strong Feller Markov chains

In this appendix, we provide sufficient conditions for the existence and uniqueness of quasi-ergodic measures of ϕ-weighted Markov processes. We prove Theorems A.13 and A.14, which are essential for the proof Lemma 3.5. The results below employ techniques of absorbing Markov processes theory [14, 19, 13] and Banach Lattice theory [42].

Let M be a compact metric space, and consider an absorbing Markov process X_n on $E = M \sqcup \partial$ absorbed at ∂. For every $x \in M$ and function $f \in L^1(M, \mu)$, we denote by $\mathbb{E}_x[f \circ X_1]$ the expected value of the observable f after one iterate of the process starting at $X_0 = x$. Define the annealed Koopman operator as

$$\begin{aligned} \mathcal{P} : L^\infty(M, \mu) &\to L^\infty(M, \mu) \\ f &\mapsto e^{\phi(x)} \mathbb{E}_x[f \circ X_1 \cdot \mathbb{1}_M \circ X_1]. \end{aligned}$$

Throughout this section, we assume that μ is a probability measure on M and $\phi : M \to \mathbb{R}$ is a continuous function. The assumptions on $\mathcal{P}$ exploited in this appendix are:

Hypothesis HA.

(1) $\mathcal{P}$ *is strong Feller, i.e. given* $f \in L^\infty(M, \mu)$ *then* $\mathcal{P}f \in \mathcal{C}^0(M)$*,*
(2) $\dim \ker(\mathcal{P} - \lambda) = 1$*, where* $\lambda = r(\mathcal{P})$*,*
(3) there exists $\mu \in \mathcal{M}_+(M)$ *and* $g \in \mathcal{C}^0_+(M)$*, such that* $\mathcal{P}^*\mu = \lambda\mu$ *and* $\mathcal{P}g = \lambda g$ *and* $\int g \, \mathrm{d}\mu = \mu(\{g > 0\})$*, and*
(4) $\operatorname{supp} \mu = M$*.*

Notation A.1. Given $n \in \mathbb{N}$ and $x \in E$ we write $\mathcal{P}^n(x, \mathrm{d}y)$ for the unique measure on M such that $\mathcal{P}^n(x, A) = \mathcal{P}^n \mathbb{1}_A(x)$ for every measurable set $A \subset M$. Observe that $\mathcal{P}^n(x, \mathrm{d}y)$ is well defined since $\mathcal{P}(L^\infty(M, \mu)) \subset \mathcal{C}^0(M)$.

A.1. Spectral properties of $\mathcal{P}$.

We begin by recalling a classical lemma in the theory of Markov processes and prove a series of results which characterises the spectrum of $\mathcal{P}$.

Lemma A.2 ([47, Chapter 1, Lemma 5.11])**.** *The operator* $\mathcal{P}^n : L^\infty(M, \mu) \to L^\infty(M, \mu)$ *is compact for every* $n > 1$*.*

Lemma A.3. *Let* $\lambda = r(\mathcal{P})$ *denote the spectral radius of* $\mathcal{P}$*. Then, there exists* $k \in \mathbb{N}$ *such that* $\sigma_{\mathrm{per}}(\mathcal{P}) = \left\{\lambda e^{2\pi i j/k}\right\}_{j=0}^{k-1}$ *where* $\sigma_{\mathrm{per}}(\mathcal{P}) := \{\alpha \in \mathbb{C};\ \|\alpha\|_{\mathbb{C}} = r(\mathcal{P})$ *and* $\ker(\mathcal{P} - \alpha) \neq \{0\}\}$ *denotes the point peripheral spectrum of* $\mathcal{P}$*.*

Proof. We divide the proof into three steps:

Step 1. *If* $f \in \ker(\mathcal{P} - \lambda e^{i\beta})$ *for some* $\beta > 0$ *then* $|f| \in \operatorname{span}\{g\}$*, where* $|f| : M \to \mathbb{R}_+$*,* $|f|(x) = \|f(x)\|_{\mathbb{C}}$ *and* $\|\cdot\|_{\mathbb{C}}$ *denotes the complex norm.*

Since $\mathcal{P}$ is a positive operator $|f| = |e^{i\beta} f| = \frac{1}{\lambda}|\mathcal{P}f| \leq \frac{1}{\lambda}\mathcal{P}|f|$. Moreover,

$$0 \leq \int_M \frac{1}{\lambda}\mathcal{P}|f| - |f| \, \mathrm{d}\mu = \int_M |f| \mathrm{d}\mu - \int_M |f| \mathrm{d}\mu = 0.$$

Since $\operatorname{supp} \mu = M$ and $|f|$ is continuous, then $|f| \in \ker(\mathcal{P} - \lambda) = \operatorname{span}\{g\}$.

Step 2. *If* $e^{i\beta_1}, e^{i\beta_2} \in \sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P})$ *for some for some* $\beta_1, \beta_2 > 0$ *then* $e^{i(\beta_1 + \beta_2)} \in \sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P})$*.*

Given $j \in \{0, 1\}$, let $f_j \in \ker(\mathcal{P} - \lambda e^{i\beta_j})$. From Step 1 and rescaling f_j, if necessary, there exists a measurable function $\theta_j : M \to \mathbb{R}$ such that $f(x) = e^{i\theta_j(x)} g(x)$.

Hence, for every $x \in M$

$$e^{i\beta_j} f(x) = e^{i\beta_j} \left(e^{i\theta_j(x)} g(x)\right) = \frac{1}{\lambda}\mathcal{P}\left(e^{i\theta_j} g\right)(x) = \frac{1}{\lambda} \int_M e^{i\theta(y)} g(y) \mathcal{P}(x, \mathrm{d}y),$$

implying that

$$g(x) = \int_M e^{i(\theta_j(y)-\theta_j(x)-\beta_j)} g(y)\mathcal{P}(x, \mathrm{d}y).$$

Since $g(x) \geq 0$ and $g(x) = \int_M g(y)\mathcal{P}(x, \mathrm{d}y)$, we obtain that $e^{i(\theta_j(y)-\theta_j(x)-\beta_j)} = 1$, for $\mathcal{P}(x, \cdot)$-almost every $y \in \{g > 0\}$.

Finally, observe that by defining $h(x) := e^{i(\theta_1(x)+\theta_2(x))} g(x)$ we obtain that

$$\begin{aligned}\mathcal{P}h(x) &= \int_M e^{i(\theta_1(y)+\theta_2(y))} g(y)\mathcal{P}(x, \mathrm{d}y)\\ &= \int_M e^{i\theta_1(x)+i\theta_2(x)+i(\beta_1+\beta_2)} g(y)\mathcal{P}(x, \mathrm{d}y) = e^{i(\beta_1+\beta_2)} h(x),\end{aligned}$$

which implies that $e^{i(\beta_1+\beta_2)} \in \sigma_{\mathrm{per}}(\mathcal{P})$.

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

From Step 2, it is enough to show that $\sigma_{\mathrm{per}}(\mathcal{P})$ is finite. Theorem A.2 implies that $\mathcal{P}^2$ is a compact operator and therefore $\sigma_{\mathrm{per}}(\mathcal{P}^2)$ is finite. Finally, since $\{\lambda^2; \lambda \in \sigma_{\mathrm{per}}(\mathcal{P})\} \subset \sigma_{\mathrm{per}}(\mathcal{P}^2)$, we obtain that $\sigma_{\mathrm{per}}(\mathcal{P})$ is also finite. □

Let $k \in \mathbb{N}$, be fixed as in Lemma A.3.

Lemma A.4. *The sequence* $\{\frac{1}{\lambda^n}\mathcal{P}^n : \mathcal{C}^0(M) \to \mathcal{C}^0(M)\}_{n\in\mathbb{N}}$ *is power bounded, i.e.* $\sup_{n\in\mathbb{N}} \|\frac{1}{\lambda^n}\mathcal{P}^n\| < \infty$.

Proof. We dive the proof into three steps.

Step 1. *We show that* $\sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P}) = \sigma_{\mathrm{per}}(\frac{1}{\lambda^{k+1}}\mathcal{P}^{k+1})$.

Let us consider $\beta \in (0, 2\pi)$ such that $e^{i\beta} \in \sigma_{\mathrm{per}}(\frac{1}{\lambda^{k+1}}\mathcal{P}^{k+1}) \setminus \sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P})$. Since $\mathcal{P}^{k+1}$ is a compact operator, there exists $f \in \ker(\mathcal{P}^{k+1} - \lambda^{k+1}e^{i\beta})$. Observe that for every $j \in \{0, 1, \dots, k\}$, we obtain that

$$0 = \left(\frac{1}{\lambda^{k+1}}\mathcal{P}^{k+1} - e^{i\beta}\right) f = \left(\frac{1}{\lambda}\mathcal{P} - e^{\frac{i\beta}{k+1}+\frac{2\pi ij}{k+1}}\right) \sum_{\ell=0}^{k} \frac{e^{\frac{i\beta(k-\ell)}{k+1}+\frac{2\pi ij(k-\ell)}{k+1}}}{\lambda^\ell}\mathcal{P}^\ell f.$$

From Step 2 of Lemma A.3, we have $\gamma := e^{\frac{i\beta}{k+1}+\frac{2\pi ij}{k+1}} \notin \sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P})$ for every $j \in \{0, 1, \dots, k\}$, as otherwise we would have $\gamma^{k+1} = e^{i\beta} \in \sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P})$. Hence, we obtain that the sum above must be zero or, multiplying by the appropriate phase, that

$$0 = e^{-\frac{i\beta k}{k+1}-\frac{2\pi ijk}{k+1}} \sum_{\ell=0}^{k} \frac{e^{\frac{i\beta(k-\ell)}{k+1}+\frac{2\pi ij(k-\ell)}{k+1}}}{\lambda^\ell}\mathcal{P}^\ell f = \sum_{\ell=0}^{k} \frac{e^{\frac{-i\beta\ell}{k+1}+\frac{-2\pi ij\ell}{k+1}}}{\lambda^\ell}\mathcal{P}^\ell f, \tag{A.9}$$

for every $j \in \{0, 1, \dots, k\}$. Finally,

$$\begin{aligned}f &= \frac{1}{k+1}\sum_{j=0}^{k} f = \frac{1}{k+1}\left[\sum_{j=0}^{k} f + \sum_{\ell=1}^{k}\left(\sum_{j=0}^{k} e^{\frac{-2\pi ij\ell}{k+1}}\right)\frac{e^{\frac{-i\beta\ell}{k+1}}}{\lambda^\ell}\mathcal{P}^\ell f\right]\\ &= \frac{1}{k+1}\sum_{j=0}^{k}\sum_{\ell=0}^{k} \frac{e^{\frac{-i\beta\ell}{k+1}+\frac{-2\pi ij\ell}{k+1}}}{\lambda^\ell}\mathcal{P}^\ell f = 0.\end{aligned}$$

where the last equality follows form (A.9). This yields a contradiction, so such a β cannot exist.

Step 2. *We show that* $\ker(\mathcal{P}^{k+1} - \lambda^{k+1}) = \ker(\mathcal{P} - \lambda) = \mathrm{span}\{g\}$.

It is clear that $\ker(\mathcal{P}-\lambda) \subset \ker(\mathcal{P}^{k+1}-\lambda^{k+1})$. In the following, we show the reverse inclusion. Let $f \in \ker(\mathcal{P}^{k+1}-\lambda^{k+1})$. For every $j \in \{1,\dots,k\}$, consider the functions

$$h_j := \sum_{\ell=0}^{k} e^{\frac{-2\pi i j \ell}{k+1}} \frac{1}{\lambda^\ell} \mathcal{P}^\ell f.$$

Since $\mathcal{P}^{k+1} f = \lambda^{k+1} f$, we obtain that $\mathcal{P} h_j = \lambda e^{2\pi i j/(k+1)} h_j$ for every $j \in \{1,\dots,k\}$. From Lemma A.3, we have that $\lambda e^{2\pi i j/(k+1)} \notin \sigma_{\mathrm{per}}(\mathcal{P})$, therefore $h_j = 0$ for every $j \in \{1,\dots,k\}$. Thus,

$$\begin{aligned}
\left(\frac{1}{\lambda}\mathcal{P}-1\right) f &= \left(\frac{1}{\lambda}\mathcal{P}-1\right) \frac{1}{k+1} \sum_{j=0}^{k} f \\
&= \left(\frac{1}{\lambda}\mathcal{P}-1\right) \frac{1}{k+1} \left[\sum_{j=0}^{k} f + \sum_{\ell=1}^{k} \left(\sum_{j=0}^{k} e^{\frac{-2\pi i j \ell}{k+1}} \right) \frac{1}{\lambda^\ell} \mathcal{P}^\ell f \right] \\
&= \left(\frac{1}{\lambda}\mathcal{P}-1\right) \sum_{j=0}^{k} \sum_{\ell=0}^{k} \frac{e^{\frac{-2\pi i j \ell}{k+1}}}{\lambda^\ell} \mathcal{P}^\ell f = \left(\frac{1}{\lambda}\mathcal{P}-1\right) \left(\sum_{\ell=0}^{k} \frac{1}{\lambda^\ell} \mathcal{P}^\ell \right) f \\
&= \left(\frac{1}{\lambda^{k+1}} \mathcal{P}^{k+1} - 1 \right) f = 0,
\end{aligned}$$

which implies that $f \in \ker(\mathcal{P}-\lambda)$.

Step 3. *There exists a decomposition* $\mathcal{C}^0(M) = \bigoplus_{j=0}^{k-1} \ker\left(\mathcal{P}^{k+1} - \lambda^{k+1} e^{2\pi i j/k}\right) \oplus W_0$, *where* W_0 *is* $\mathcal{P}^{k+1}$*-invariant subspace of* $\mathcal{C}^0(M)$ *and* $r(\mathcal{P}^{k+1}|_{W_0}) < \lambda^{k+1}$. *In particular,* $\{\frac{1}{\lambda^n}\mathcal{P}^n\}_{n\in\mathbb{N}}$ *is power bounded.*

Recall from Lemma A.2 that, $\mathcal{P}^{k+1}$ is a compact linear operator. Moreover, from Step 1 we obtain that $\sigma_{\mathrm{per}}(\frac{1}{\lambda}\mathcal{P}) = \sigma_{\mathrm{per}}(\frac{1}{\lambda^{k+1}}\mathcal{P}^{k+1})$. From [38, Theorems 8.4-3 and 8.4-5] and Lemma A.3 we obtain that there exist non-zero $r_0, r_1, \dots, r_{k-1} \in \mathbb{N}$ such that

$$\mathcal{C}^0(M) = \bigoplus_{j=0}^{k-1} \ker(\mathcal{P}^{k+1} - \lambda^{k+1} e^{2\pi i j/k})^{r_j} \oplus W_0,$$

where $r_j = \inf\{m;\ \ker(\mathcal{P}^{k+1} - \lambda^{2\pi i j/k})^{m+n} = \ker(\mathcal{P}^{k+1} - \lambda^{2\pi i j/k})^m,\ \text{all } n \in \mathbb{N}\}$, and W_0 is $\mathcal{P}^{k+1}$-invariant satisfying $r(\mathcal{P}^{k+1}|_{W_0}) < \lambda^{k+1}$. We show that $r_0 = r_1 = \ldots = r_{k-1} = 1$. Using once again that $\mathcal{P}^{k+1}$ is a compact operator, we obtain from the Krein-Rutman theorem [33, Theorem 4.1] that the spectral radius $\lambda^{k+1} = r(\mathcal{P}^{k+1})$ is a pole of maximal order in the spectral circle, i.e. $r_0 \geq \max\{r_1, \dots, r_{k-1}\}$. Suppose that $r_0 > 1$, then there exists $f \in \mathcal{C}^0(M)$ such that $g = (\mathcal{P}^{k+1} - \lambda^{k+1}) f$. Therefore,

$$0 < \int g \, \mathrm{d}\mu = \int \mathcal{P}^{k+1} f - \lambda^{k+1} f \, \mathrm{d}\mu = \int f \, \mathrm{d}(\mathcal{P}^{k+1})^* \mu - \int \lambda^{k+1} f \, \mathrm{d}\mu = 0,$$

implying that $r_0 = 1$ and therefore $r_0 = r_1 = \ldots = r_{k-1} = 1$. □

Lemma A.5. *There exists a decomposition* $\mathcal{C}^0(M) = \bigoplus_{j=0}^{k-1} \ker(\mathcal{P} - \lambda e^{\frac{2\pi i j}{k}}) \oplus W$, *where* W *is a* $\mathcal{P}$*-invariant space* $r\left(\mathcal{P}|_W\right) < \lambda$, *and* $\dim \ker(\mathcal{P} - \lambda e^{\frac{2\pi i j}{k}}) = 1$ *for every* $j \in \{0, 1 \dots, k-1\}$.

Proof. From Lemmas A.2 and A.4 we obtain that $\mathcal{P}^2$ is a compact linear operator and $\sup_{n\geq 0} \|\frac{1}{\lambda^n}\mathcal{P}^n\| < \infty$. Then, from [59, *An extension of Frechet-Kryloff-Bogoliouboff's theorem*] (see also [11, Théorèm above Définition 1.5] and [58, Equation (8) in the proof of Theorem 4]), there exists a $\mathcal{P}$-invariant space $W \subset \mathcal{C}^0(M)$ such that $r\left(\mathcal{P}|_W\right) < \lambda$ and

$$\mathcal{C}^0(M) = \bigoplus_{j=0}^{k-1} \ker\left(\mathcal{P} - \lambda e^{\frac{2\pi i j}{k}}\right) \oplus W.$$

Finally, we show that $\dim\ker(\mathcal{P} - \lambda e^{\frac{2\pi i j}{k}}) = 1$ for every $j \in \{0, 1\dots, k-1\}$. Since $\sup_{n\geq 0}\left\|\frac{1}{\lambda^n}\mathcal{P}^n\right\| < \infty$, [32, Theorem 5.1] implies that

$$\begin{aligned}\dim\ker(\mathcal{P} - \lambda e^{2\pi i/k}) &\leq \dim\ker(\mathcal{P} - \lambda e^{2\pi i 2/k}) \leq \dots \\ &\leq \dim\ker(\mathcal{P} - \lambda e^{2\pi i(k-1)/k}) \leq \dim\ker(\mathcal{P} - \lambda) = 1,\end{aligned}$$

which concludes the proof. □

A.2. **Cyclic properties of $\mathcal{P}$.** Consider $\mathcal{P}$ acting only on continuous functions $\mathcal{P} : \mathcal{C}^0(M) \to \mathcal{C}^0(M)$. From Lemma A.5, we know that $\mathcal{C}^0(M) = \ker(\mathcal{P}^k - \lambda^k) \oplus W$, where W is a $\mathcal{P}^k$-invariant Banach space such that $r\left(\mathcal{P}^k|_W\right) < \lambda^k$ and $\dim\ker(\mathcal{P}^k - \lambda^k) = k$.

Proposition A.6. *There exist k non-negative linearly independent eigenfunctions g_0, …, $g_{k-1} \in \mathcal{C}^0_+(M)\cap\ker(\mathcal{P}^k - \lambda^k)$ such that $\mathrm{span}_{\mathbb{C}}(\{g_i\}_{i=0}^{k-1}) = \ker(\mathcal{P}^k - \lambda^k)$ and $\int g_i \mathrm{d}\mu = \mu(\{g > 0\})$ for every $i \in \{0, 1, \dots, k-1\}$. Moreover, these can be chosen such that the sets $C_i := \{g_i > 0\}$ are pairwise disjoint.*

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

Recall that μ is a measure on M satisfying $\mathcal{P}^*\mu = \lambda\mu$ and $\mathrm{supp}\,\mu = M$. Note that the operator $\mathcal{P}^k$ satisfies $\int_M \frac{1}{\lambda^k}\mathcal{P}^k f(x)\mu(\mathrm{d}x) = \int_M f(x)\mathrm{d}\mu$, for every $f \in \mathcal{C}^0(M)$. By the same techniques of Theorem 3.4 (see also [40, Propositions 3.1.1 and 3.1.3]), it follows that if $f \in \mathcal{C}^0(M)$ is an eigenfunction of $\mathcal{P}^k$ associated with the eigenvalue λ^k, then $f^+(x) := \max\{0, f(x)\}$ and $f^-(x) = \max\{0, -f(x)\}$ are also eigenfunctions of $\mathcal{P}^k$ associated with the eigenvalue λ^k. This provides a set of k linearly independent non-negative continuous functions that span $\ker(\mathcal{P}^k - \lambda^k)$. We are left to check that these can be chosen with pair-wise disjoint support.

Define $G := \{h_1 > 0\} \setminus \{h_2 > 0\} \neq \emptyset$ and $H := \{h_1 > 0\} \cap \{h_2 > 0\}$. We claim that if $h_1, h_2 \in \mathcal{C}^0_+(M)\cap\ker(\mathcal{P}^k - \lambda^k)$, then $\mathbb{1}_G h_1$ and $\mathbb{1}_H h_1$ are also a eigenfunctions of $\mathcal{P}^k$ associated with the eigenvalue λ^k. Observe that this is enough to conclude the proof since we can choose k functions of the set below which have disjoint supports

$$\left\{h_i \mathbb{1}_{\{(\sum_{j=0}^{k-1} t_j h_j)^{\pm} > 0\}};\ t_0, \dots, t_{k-1} \geq 0 \text{ and } i \in \{0, \dots, k-1\}\right\} \subset \ker(\mathcal{P}^k - \lambda^k).$$

We organise the remainder of the proof into three steps:

Step 1. $\mathbb{1}_G\mathcal{P}^k\mathbb{1}_H = 0$.

Let $x \in G$ and assume that $\mathcal{P}^k\mathbb{1}_H > 0$. Then, $0 < \mathcal{P}^k\mathbb{1}_H = \int \mathbb{1}_H(y)\mathcal{P}^k(x, dy)$, and since $h_2 > 0$ on H, $0 < \int \mathbb{1}_H(y)h_2(y)(\mathcal{P}^*)^k\delta_x(dy)$, implying that $\mathcal{P}^k\mathbb{1}_H h_2(x) > 0$. Moreover,

$$h_2(x) = \frac{1}{\lambda^k}\mathcal{P}^k h_2(x) = \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_2 + (1 - \mathbb{1}_H)h_2)(x) \geq \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_2)(x) > 0,$$

which contradicts $x \notin \{h_2 > 0\}$. In particular, $\mathbb{1}_H\mathcal{P}^k\mathbb{1}_H = \mathcal{P}^k\mathbb{1}_H$.

Step 2. $\mathbb{1}_H\mathcal{P}^k\mathbb{1}_G = 0$.

From Step 1, it follows that

$$\mathbb{1}_H h_1 = \mathbb{1}_H\frac{1}{\lambda^k}\mathcal{P}^k h_1 = \mathbb{1}_H\frac{1}{\lambda}\mathcal{P}^k(\mathbb{1}_H h_1 + \mathbb{1}_G h_1) = \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_1) + \mathbb{1}_H\frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_G h_1).$$

Integrating either side and using $\mu \in \ker((\mathcal{P}^k)^* - \lambda^k)$, we obtain

$$\begin{aligned}\int \mathbb{1}_H h_1 \mathrm{d}\mu &= \int \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_1)\mathrm{d}\mu + \int \mathbb{1}_H \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_G h_1)\mathrm{d}\mu \\ &= \int \mathbb{1}_H h_1 \mathrm{d}\mu + \int \mathbb{1}_H \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_G h_1)\mathrm{d}\mu,\end{aligned}$$

implying that $\int \mathbb{1}_H \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_G h_1)\mathrm{d}\mu = 0$, which yields $\mathbb{1}_H \mathcal{P}^k \mathbb{1}_G = 0$.

Step 3. *$\mathbb{1}_G h_1$ and $\mathbb{1}_H h_1$ are eigenfunctions of $\mathcal{P}^k$ with eigenvalue λ^k.*

From Steps 1 and 2, it follows that

$$\begin{aligned}\mathbb{1}_H h_1 + \mathbb{1}_G h_1 = h_1 &= \frac{1}{\lambda^k}\mathcal{P}^k h_1 = \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_1 + \mathbb{1}_G h_1) \\ &= \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_1) + \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_G h_1) = \mathbb{1}_H \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_H h_1) + \mathbb{1}_G \frac{1}{\lambda^k}\mathcal{P}^k(\mathbb{1}_G h_1).\end{aligned}$$

Since $G \cap H = \emptyset$, the claim is verified and we finish the proof. □

Lemma A.7. *Let $\{g_i\}_{i=0}^{k-1} \subset \mathcal{C}_+^0(M)$ be as in Proposition A.6. Then, these can be relabelled so that $\frac{1}{\lambda}\mathcal{P}g_i = g_{i-1\,(mod\ k)}$, for $i \in \{0, 1, \ldots, k-1\}$. In particular, we have that $g = \frac{1}{k}\sum_{i=0}^{k-1} g_i$.*

Proof. We divide the proof into two steps.

Step 1. *There exists a continuous function $\theta : \{g > 0\} \to \{0, 1/k, 2/k, \ldots, (k-1)/k\}$, such that*

(1) for every $j \in \{0, 1, \ldots, k-1\}$, $\theta|_{\{g_j>0\}} = \theta_j$ is constant,

(2) the set $\{g_j\}_{j=0}^{k-1}$ can be relabelled so that $\theta_j = j/k$.

From Step 1 of Lemma A.3, there exists a function $\theta : \{g > 0\} \to \mathbb{R}$ such that $e^{2\pi i\theta(x)}g \in \ker(\mathcal{P} - \lambda e^{2\pi i/k})$. Observe that by multiplying θ by a complex constant, we can assume without loss of generality that there exists $x \in M$ such that $\theta(x) = 0$. Since $e^{2\pi i\theta(x)}g, g \in \ker(\mathcal{P}^k - \lambda^k) = \operatorname{span}\{g_0, \ldots, g_{k-1}\}$, there exist $\alpha_0, \ldots, \alpha_{k-1} \geq 0$ and $\theta_0, \ldots, \theta_{k-1} \geq 0$ such that

$$g = \sum_{j=0}^{k-1} \alpha_j g_j \text{ and } e^{2\pi i\theta}g = \sum_{j=0}^{k-1} \alpha_j e^{2\pi i\theta_j} g_j.$$

Since $\{g_{j_1} > 0\} \cap \{g_{j_2} > 0\} = \emptyset$ if $j_1 \neq j_2$, then $\theta(x) = \theta_j$ for every $x \in \{g_j > 0\}$. This proves (1).

Without loss of generality, we may assume that $\alpha_0 \neq$ and $\theta_0 = 0$. Let us fix $x \in \{g_0 > 0\}$. Then,

$$e^{2\pi i/k}g(x) = \frac{1}{\lambda}\mathcal{P}(e^{2\pi i\theta}g)(x) = \int_M e^{2\pi i\theta(y)}g(y)\frac{1}{\lambda}\mathcal{P}(x, \mathrm{d}y),$$

and therefore

$$g(x) = \int_M e^{2\pi i(\theta(y)-1/k)}g(y)\frac{1}{\lambda}\mathcal{P}^*\delta_x(\mathrm{d}y).$$

Since $g(x) = \int g(y)\frac{1}{\lambda}\mathcal{P}(x, \mathrm{d}y)$, and θ is continuous we obtain

$$\theta(y) = \frac{1}{k} \text{ for every } y \in \operatorname{supp}\mathcal{P}(x, \mathrm{d}y) \cap \{g > 0\}.$$

The same argument for $\mathcal{P}^n$ yields

$$\theta(y) = \frac{n}{k} \text{ for every } y \in \operatorname{supp}\mathcal{P}^n(x, \mathrm{d}y) \cap \{g > 0\}. \tag{A.10}$$

Note that if $y \in \operatorname{supp}\mathcal{P}^n(x, \mathrm{d}y) \cap \{g_j > 0\}$ for some j, then $\theta_j = \theta(y) = n/k$. This implies that $\operatorname{supp}\mathcal{P}^m(x, \mathrm{d}y) \cap \{g_j > 0\} = \emptyset$ for any $m \neq n\,(\mathrm{mod}\ k)$. Since $\operatorname{supp}\mathcal{P}^k(x, \mathrm{d}y) \cap \{g_0\} \neq$

$\emptyset$ and there are exactly k functions $g_0, \dots, g_k$, each must have a different phase θ_j. After relabelling, we may assume that $\theta_j = j/k$, for $j \in \{0, 1, \dots, k\}$, showing (2).

Step 2. *We conclude the proof of the lemma.*

It follows immediately from equation (A.10) that $\{\mathcal{P}g_j > 0\} \subset \{g_{j-1\,(\mathrm{mod}\ k)} > 0\}$. Moreover, since $e^{2\pi i\theta} g \in \ker(\mathcal{P} - \lambda e^{2\pi i/k})$, we have

$$\lambda e^{2\pi i/k} e^{2\pi i\theta} g = \lambda e^{2\pi i/k} \sum_{j=0}^{k-1} \alpha_j e^{2\pi i j/k} g_j = \mathcal{P}(e^{2\pi i\theta} g) = \sum_{j=0}^{k-1} \alpha_j e^{2\pi i j/k} \mathcal{P} g_j,$$

so $\lambda\alpha_{j-1} g_{j-1} = \alpha_j \mathcal{P} g_j$, with $-1 = k-1$. Integrating both sides with respect to μ yields $\alpha_{j-1\,(\mathrm{mod}\ k)} = \alpha_j$, from which we conclude that $\alpha_0 = \alpha_1 = \dots = \alpha_{k-1}$ and $\frac{1}{\lambda}\mathcal{P} g_j = g_{j-1\,(\mathrm{mod}\ k)}$, for every $j \in \{0, 1, \dots, k-1\}$. □

The following corollary follows directly from Lemma A.7.

Corollary A.8. *Every function* $f_\ell := \frac{1}{k}\sum_{j=0}^{k-1} e^{2\pi i j\ell/k} g_j$ *satisfies* $\mathcal{P} f_\ell = \lambda e^{2\pi i\ell/k} f_\ell$*, i.e.* $\ker(\mathcal{P} - \lambda e^{2\pi i\ell/k}) = \mathrm{span}(f_\ell)$*, for* $\ell \in \{0, 1, \dots, k-1\}$.

A.3. **Existence of quasi-ergodic measures.** Recall that $g \in \mathcal{C}^0_+(M)$ is the unique function satisfying $\mathcal{P} g = \lambda g$.

Lemma A.9. $\mathcal{P}\mathbb{1}_{\{g>0\}} \le c\mathbb{1}_{\{g>0\}}$*, for some constant* $c > 0$.

Proof. Observe that for every $a > 0$ we have $\mathcal{P}\mathbb{1}_{\{g>a\}} \le \frac{1}{a}\mathcal{P} g = \frac{\lambda}{a} g$. Hence, $\{\mathcal{P}\mathbb{1}_{\{g>a\}} > 0\} \subset \{g > 0\}$. Since $\mathbb{1}_{\{g>0\}} = \sum_{n=1}^{\infty} \mathbb{1}_{\{\|g\|_\infty/n \ge g > \|g\|_\infty/(n+1)\}}$, we obtain that

$$\{\mathcal{P}\mathbb{1}_{\{g>0\}} > 0\} \subset \bigcup_{n\in\mathbb{N}} \{\mathcal{P}\mathbb{1}_{\{g>1/n\}} > 0\} \subset \{g > 0\}.$$

It follows that $\mathcal{P}\mathbb{1}_{\{g>0\}} \le \|\mathcal{P}\|\mathbb{1}_{\{g>0\}}$. □

Notation A.10. We define the operator $\mathcal{P}_g : L^\infty(\{g > 0\}, \mu) \to L^\infty(\{g > 0\}, \mu)$ as $\mathcal{P}_g f := \mathcal{P}(\mathbb{1}_{\{g>0\}} f)$.

Corollary A.11. *The measure* $\widetilde{\mu}(\mathrm{d}x) := \mu(\mathrm{d}x \cap \{g > 0\})/\mu(\{g > 0\})$ *satisfies* $\mathcal{P}_g^*\widetilde{\mu} = \lambda\widetilde{\mu}$.

Proof. From Lemma A.9 we have that for every $h \in L^\infty(\{g > 0\})$, $\mathcal{P}_g h = \mathcal{P}(\mathbb{1}_{\{g>0\}} h) = \mathbb{1}_{\{g>0\}}\mathcal{P}_g h$. Therefore,

$$\begin{aligned}\int h\mathrm{d}(\mathcal{P}_g^*\widetilde{\mu}) &= \int \mathcal{P}_g h\mathrm{d}\widetilde{\mu} = \frac{1}{\mu(\{g>0\})}\int \mathbb{1}_{\{g>0\}}\mathcal{P}_g h\mathrm{d}\mu = \frac{1}{\mu(\{g>0\})}\int \mathcal{P}(\mathbb{1}_{\{g>0\}} h)\mathrm{d}\mu \\ &= \frac{1}{\mu(\{g>0\})}\int \mathbb{1}_{\{g>0\}} h\mathrm{d}(\mathcal{P}^*\mu) = \frac{\lambda}{\mu(\{g>0\})}\int \mathbb{1}_{\{g>0\}} h\mathrm{d}\mu = \lambda\int h\mathrm{d}\widetilde{\mu}.\end{aligned}$$

□

Observe that since $\int g\,\mathrm{d}\mu = \mu(\{g > 0\})$, $\int g\,\mathrm{d}\widetilde{\mu} = 1$. The above corollary implies that $\sigma_{\mathrm{per}}(\mathcal{P}) = \sigma_{\mathrm{per}}(\mathcal{P}_g)$ and

$$\mathbb{1}_{\{g>0\}}\ker(\mathcal{P} - \lambda e^{2\pi i j/k}) = \ker(\mathcal{P}_g - \lambda e^{2\pi i j/k}),\ \text{for every } j \in \{0, \dots, k-1\}.$$

Since each g_i defined in Lemma A.5 satisfies $C_i = \{g_i > 0\} \subset \{g > 0\}$, we can assume by abuse of notation that $g_i \in L^\infty(\{g > 0\}, \widetilde{\mu})$. Moreover,

$$L^\infty(\{g > 0\}, \widetilde{\mu}) = \mathrm{span}(g_0, \dots, g_{k-1}) \oplus V,$$

where V is $\mathcal{P}$-invariant and $r(\mathcal{P}|_V) < \lambda$.

Lemma A.12. *For every* $i \in \{0, 1, \dots, k-1\}$ *define* $\widetilde{\mu}_i(\mathrm{d}x) = \widetilde{\mu}(C_i \cap \mathrm{d}x)$*, where* $C_i = \{g_i > 0\}$*. Then* $\mathcal{P}_g^*\widetilde{\mu}_i = \lambda\widetilde{\mu}_{i+1\,(\mathrm{mod}\ k)}$.

Proof. We divide the proof into two steps

Step 1. *$v \in V$ if and only if $\int_{C_i} v \, \mathrm{d}\widetilde{\mu} = 0$ for every $i \in \{0, 1, \dots, k-1\}$.*

Suppose first that $v \in V$. We claim that $\mathbb{1}_{C_i} v \in V$ for all $i \in \{0, 1, \dots, k-1\}$. Indeed, if $\mathbb{1}_{C_i} v \notin V$, then $v = \alpha_i g_i + w + \sum_{j \neq i} \mathbb{1}_{C_j} v$ with $\alpha_i \neq 0$ and $w \in V$. Since, $C_i \cap C_j = \emptyset$ for all $j \neq i$, we get that $v \notin V$. If follows that

$$\left| \int_{C_i} v \, \mathrm{d}\widetilde{\mu} \right| = \left| \int_M \mathbb{1}_{C_i} v \, \mathrm{d}\widetilde{\mu} \right| = \left| \int_M \frac{1}{\lambda^n} \mathcal{P}^n (\mathbb{1}_{C_i} v) \, \mathrm{d}\widetilde{\mu} \right| \leq \left\| \frac{1}{\lambda^n} \mathcal{P}^n \Big|_V \right\| \|v\| \xrightarrow{n \to \infty} 0.$$

Suppose now that $\int_{C_i} v \, \mathrm{d}\widetilde{\mu} = 0$ for every $i \in \{0, 1, \dots, k-1\}$. Write $v = \sum_{i=0}^{k-1} \alpha_i g_i + w$, with $w \in V$. Observing that $\int g_i \mathrm{d}\widetilde{\mu} = 1$, we have

$$\alpha_i = \int_{C_i} \alpha_i g_i \, \mathrm{d}\widetilde{\mu} = \int_{C_i} \left(\sum_{j=0}^{k-1} \alpha_j g_j + w \right) \mathrm{d}\widetilde{\mu} = \int_{C_i} v \, \mathrm{d}\widetilde{\mu} = 0.$$

We obtain that $\alpha_i = 0$ for every $i \in \{0, 1, \dots, k-1\}$, which implies $v \in V$.

Step 2. *We conclude the proof of the lemma.*

Take $f \in L^\infty(\{g > 0\}, \mu)$. Therefore $f = \sum_{i=0}^{k-1} \alpha_i g_i + v$, $v \in V$. From Step 1, it follows that

$$\alpha_i = \int_{C_i} f \, \mathrm{d}\widetilde{\mu} = \int f \, \mathrm{d}\widetilde{\mu}_i,$$

and

$$\begin{aligned} \int f \, \mathrm{d}\mathcal{P}^* \widetilde{\mu}_i = \int \mathcal{P} f \, \mathrm{d}\widetilde{\mu}_i &= \sum_{j=0}^{k-1} \int_M \mathbb{1}_{C_i} \mathcal{P}(\alpha_j g_j) \, \mathrm{d}\widetilde{\mu} \\ &= \sum_{j=0}^{k-1} \int_M \mathbb{1}_{C_i} \lambda \alpha_j g_{j-1} \, \mathrm{d}\widetilde{\mu} = \lambda \alpha_{i+1} = \lambda \int_M f \, \mathrm{d}\widetilde{\mu}_{i+1 \pmod k}. \end{aligned}$$

□

Theorem A.13. *Assume that $\mathcal{P}$ satisfies Hypothesis HA. Given a bounded and measurable function $h : \{g > 0\} \to \mathbb{R}$ we have that for every $x \in \{g > 0\}$*

$$\frac{1}{\mathbb{E}_x[e^{S_n \phi} \mathbb{1}_{\{\tau > n\}}]} \mathbb{E}_x \left[e^{S_n \phi} \mathbb{1}_{\{\tau > n\}} \frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i \right] \xrightarrow{n \to \infty} \frac{\int h(x) g(x) \mu(\mathrm{d}x)}{\int g(x) \mu(\mathrm{d}x)},$$

where $\tau = \min\{n \in \mathbb{N}; X_n \notin \{g > 0\}\}$ and $S_n \phi = \sum_{i=0}^{n-1} \phi \circ X_i$. In other words, there exists a unique quasi-ergodic measure for the ϕ-weighted Markov process X_n^ϕ on $\{g > 0\}$.

Proof. In this proof, we adopt the notation $g_m := g_{m \pmod k}$, $\widetilde{\mu}_m := \widetilde{\mu}_{m \pmod k}$ and $C_m = C_{m \pmod k}$. Recall that

$$\frac{\int h(x) g(x) \mu(\mathrm{d}x)}{\int g(x) \mu(\mathrm{d}x)} = \int h(x) g(x) \, \widetilde{\mu}(\mathrm{d}x).$$

Given $n \in \mathbb{N}$ and $x \in \{g > 0\}$ define

$$Q_h^n(x) := \left| \frac{1}{\mathbb{E}_x[e^{S_n \phi} \mathbb{1}_{\{\tau > n\}}]} \mathbb{E}_x \left[e^{S_n \phi} \mathbb{1}_{\{\tau > n\}} \frac{1}{n} \sum_{i=0}^{n-1} h \circ X_i \right] - \int h(x) g(x) \widetilde{\mu}(\mathrm{d}x) \right|.$$

Observe that to prove the theorem, it suffices to show that for every measurable and bounded non-negative $h : \{g > 0\} \to \mathbb{R}$ we have

$$\max \left\{ Q_h^{nk+\ell}(x); \, \ell \in \{0, 1, \dots, k-1\} \right\} \xrightarrow{n \to \infty} 0,$$

for every $x \in C_s$ where $s \in \{0, 1, \dots, k-1\}$.

We divide the remainder of the proof into three steps.

Step 1. *For every bounded and measurable function $h : \{g > 0\} \to \mathbb{R}$, $\ell, s \in \{0, 1, \dots, k-1\}$ and $x \in C_s$ we have*

$$\lim_{n\to\infty} \frac{1}{\lambda^{nk+\ell}} \mathcal{P}_g^{nk+\ell} h(x) = g_s(x) \int_{C_{s+\ell}} h(y) \widetilde{\mu}_{s+\ell}(\mathrm{d}y).$$

From Step 1 of Lemma A.12 it is clear that

$$h = \sum_{j=0}^{k-1} g_j \int_{C_j} h \,\mathrm{d}\widetilde{\mu} + v,$$

with $v \in V$. Since $\mathcal{P}_g^{nk+\ell} g_j(x) = \lambda^{nk+\ell} g_{j-\ell}(x)$, we obtain that

$$\frac{1}{\lambda^{nk+\ell}} \mathcal{P}_g^{nk+\ell} h = \sum_{j=0}^{k-1} g_{j-\ell} \int_{C_j} h \,\mathrm{d}\widetilde{\mu} + \frac{1}{\lambda^{nk+\ell}} \mathcal{P}^{nk+\ell} v \xrightarrow{n\to\infty} \sum_{j=0}^{k-1} g_{j-\ell} \int_{C_j} h \,\mathrm{d}\widetilde{\mu}.$$

Finally, if $x \in C_s$, then

$$\lim_{n\to\infty} \frac{1}{\lambda^{nk+\ell}} \mathcal{P}_g^{nk+\ell} h(x) = g_s(x) \int_{C_{s+\ell}} h \,\mathrm{d}\widetilde{\mu}.$$

Step 2. *For every non-negative bounded and measurable function $h : \{g > 0\} \to \mathbb{R}$, $\ell, s \in \{0, 1, \dots, k-1\}$ and $x \in C_s$ we have*

$$\lim_{n\to\infty} \frac{1}{nk+\ell} \sum_{i=0}^{nk+\ell-1} \frac{1}{\lambda^i} \mathcal{P}_g^i \left(h \frac{1}{\lambda^{nk+\ell-i}} \mathcal{P}_g^{nk+\ell-i} \mathbb{1}_{\{g>0\}} \right)(x) = g_s(x) \widetilde{\mu}(C_{s+\ell}) \int hg \,\widetilde{\mu}(\mathrm{d}x).$$

We denote $\mathcal{G} := \mathcal{P}_g / \lambda$ to simplify the notation and improve readability. Recall that $\mathbb{1}_{\{g>0\}} = \sum_{j=0}^{k-1} \widetilde{\mu}(C_j) g_j + v$, where $v \in V$. It follows that,

$$\begin{aligned}
&\sum_{i=0}^{nk+\ell-1} \mathcal{G}^i \left(h \mathcal{G}^{nk+\ell-i} \mathbb{1}_{\{g>0\}} \right)(x) = \\
&\qquad = \sum_{j=0}^{k-1} \widetilde{\mu}(C_j) \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i \left(h \mathcal{G}^{nk+\ell-i} g_j \right)(x) + \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i (h \mathcal{G}^{nk+\ell-i} v)(x) \\
&\qquad = \sum_{j=0}^{k-1} \widetilde{\mu}(C_j) \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i \left(h g_{j-\ell+i} \right)(x) + \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i (h \mathcal{G}^{nk+\ell-i} v)(x).
\end{aligned}$$

Observe that

$$\begin{aligned}
\left| \frac{1}{nk+\ell} \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i (h \mathcal{G}^{nk+\ell-i} v)(x) \right| &\le \sup_{i\ge 0} \|\mathcal{G}^i\| \frac{1}{nk+\ell} \sum_{i=0}^{nk+\ell-1} \|h \mathcal{G}^{nk+\ell-i} v\|_\infty \\
&\le \sup_{i\ge 0} \|\mathcal{G}^i\| \frac{1}{nk+\ell} \sum_{i=0}^{nk+\ell-1} \|h\|_\infty \left\| \mathcal{G}^{nk+\ell-i} v \right\|_\infty \xrightarrow{n\to\infty} 0.
\end{aligned}$$

Moreover, since $0 \le h g_{j-\ell+i} \le \|h\|_\infty g_{i-\ell+i}$, then

$$0 \le \mathcal{G}^i (h g_{j-\ell+i}) \le \|h\|_\infty g_{j-\ell} \le \|g\|_\infty \|g_{j-\ell}\| \mathbb{1}_{C_{i-j}}.$$

It follows that for every $x \in C_s$,

$$\begin{aligned}
\frac{1}{nk+\ell} \sum_{j=0}^{k-1} \widetilde{\mu}(C_j) \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i \left(h g_{j-\ell+i} \right)(x) &= \frac{\widetilde{\mu}(C_{s+\ell})}{nk+\ell} \sum_{i=0}^{nk+\ell-1} \mathcal{G}^i \left(h g_{s+i} \right)(x) \\
&= \frac{\widetilde{\mu}(C_{s+\ell})}{nk+\ell} \sum_{i=0}^{n-1} \mathcal{G}^{ik} \left(\sum_{r=0}^{k-1} \mathcal{G}^r \left(h g_{s+r} \right) \right)(x) + \frac{\widetilde{\mu}(C_{s+\ell})}{nk+\ell} \mathcal{G}^{nk} \left(\sum_{r=0}^{\ell} \mathcal{G}^r \left(h g_{s+r} \right) \right)(x).
\end{aligned}$$

From Step 1, we obtain that

$$\lim_{n\to\infty}\frac{1}{nk+\ell-1}\sum_{j=0}^{k-1}\widetilde{\mu}(C_j)\sum_{i=0}^{nk+\ell}\mathcal{G}^i\left(hg_{s+i}\right)(x) = g_s(x)\frac{\widetilde{\mu}(C_{s+\ell})}{k}\sum_{j=0}^{k-1}\int \mathcal{G}^j(hg_{s+j})(y)\widetilde{\mu}_s(\mathrm{d}y)$$

$$\left(\substack{\text{Lem. A.7 and}\\ \text{Lem. A.12}}\right) = g_s(x)\frac{\widetilde{\mu}(C_{s+\ell})}{k}\sum_{j=0}^{k-1}\int h(y)g_{s+j}(y)\widetilde{\mu}_{s+j}(\mathrm{d}y)$$

$$= g_s(x)\widetilde{\mu}(C_{s+\ell})\int h(y)g(y)\widetilde{\mu}(\mathrm{d}y).$$

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

Given a non-negative bounded and measurable function $h:\{g>0\}\to\mathbb{R}$, $\ell,s\in\{0,1,\dots,k-1\}$ and $x\in C_s$.

$$I_h^{n,\ell}(x) := \frac{\lambda^{nk+\ell}}{\mathcal{P}_g^{nk+\ell}\mathbb{1}_{\{g>0\}}(x)}\frac{1}{nk+\ell}\sum_{i=0}^{nk+\ell-1}\frac{1}{\lambda^i}\mathcal{P}_g^i\left(h\frac{1}{\lambda^{nk+\ell-i}}\mathcal{P}_g^{nk+\ell-i}\mathbb{1}_{\{g>0\}}\right)(x).$$

From Steps 1 and 2, we obtain that

$$\lim_{n\to\infty}\frac{\lambda^{nk+\ell}}{\mathcal{P}_g^{nk+\ell}\mathbb{1}_{\{g>0\}}(x)} = \frac{1}{g_s(x)\mu(C_{s+\ell})},$$

and

$$\lim_{n\to\infty}\frac{1}{nk+\ell}\sum_{i=0}^{nk+\ell-1}\frac{1}{\lambda^i}\mathcal{P}_g^i\left(h\frac{1}{\lambda^{nk+\ell-i}}\mathcal{P}_g^{nk+\ell-i}\mathbb{1}_{\{g>0\}}\right)(x) = g_s(x)\mu(C_{s+\ell})\int h(y)g(y)\widetilde{\mu}(\mathrm{d}y).$$

Therefore $Q_h^{nk+\ell}(x)\xrightarrow{n\to\infty}0$ for all $s,\ell\in\{0,1,\dots,k-1\}$ and $x\in C_s$, which concludes the proof of the theorem. □

Theorem A.14. *Assume that $\mathcal{P}$ satisfies Hypothesis HA and $\sigma(\frac{1}{\lambda}\mathcal{P})\cap\mathbb{S}^1=\{1\}$. Then, given a bounded measurable function $h:M\to\mathbb{R}$, for every $x\in\{g>0\}$,*

$$\frac{1}{\mathbb{E}_x[e^{S_n\phi}\mathbb{1}_{\{\tau>n\}}]}\mathbb{E}_x\left[e^{S_n\phi}\mathbb{1}_{\{\tau>n\}}\frac{1}{n}\sum_{i=0}^{n-1}h\circ X_i\right]\xrightarrow{n\to\infty}\frac{\int h(x)g(x)\mu(\mathrm{d}x)}{\int g(x)\mu(\mathrm{d}x)},$$

where $\tau := \min\{n; X_n\notin M\}$ and $S_n\phi=\sum_{i=0}^{n-1}\phi\circ X_i$. In other words, there exists a unique quasi-ergodic of the ϕ-weighted Markov process X_n^ϕ on M.

Proof. Note that the spectral gap in the operator $\frac{1}{\lambda}\mathcal{P}$, along with its strong Feller property, ensures that for any bounded and measurable function $h:M\to\mathbb{R}$, it holds that

$$\sup_{x\in M}\left|\frac{1}{\lambda^n}\mathcal{P}^n h(x)-g(x)\int h\mathrm{d}\mu\right|\xrightarrow{n\to\infty}0. \tag{A.11}$$

Repeating the proof of Step 2 of Theorem A.13 we obtain that

$$\sup_{x\in M}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\lambda^i}\mathcal{P}^i\left(h\frac{1}{\lambda^{n-i}}\mathcal{P}^{n-i}\mathbb{1}_M\right)(x)-g(x)\int h(y)g(y)\mu(\mathrm{d}y)\right|\xrightarrow{n\to\infty}0. \tag{A.12}$$

Combining equations (A.11)-(A.12) and the same computations in the proof of Theorem A.13 we obtain the result. □