

Random 2D linear cocycles II: Statistical properties

Pedro Duarte, Marcelo Durães, Tomé Graxinha, and Silvius Klein

Abstract. Consider the space of two-dimensional random linear cocycles over a shift in finitely many symbols, with at least one singular and one invertible matrix. We provide an explicit formula for the unique stationary measure associated to such cocycles and establish a Furstenberg-type formula characterizing the Lyapunov exponent. Using the spectral properties of the corresponding Markov operator and a parameter elimination argument, we prove that Lebesgue almost every cocycle in this space satisfies large deviations estimates and a central limit theorem.

1. Introduction and statements

Let ν be a compactly supported measure on the space of matrices $\text{Mat}_m(\mathbb{R})$, $m \geq 2$. Consider an i.i.d. sequence of matrices $\{A_n\}_{n \in \mathbb{Z}}$ in $\text{Mat}_m(\mathbb{R})$, with common distribution ν , and the corresponding multiplicative process $\Pi_n := A_n \cdots A_2 A_1$. Under a general integrability assumptions, namely $\int \log^+ \|A\| d\nu(A) < \infty$, by the Furstenberg–Kesten theorem, we have that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|\Pi_n\| =: L_1(\nu) \quad (1.1)$$

exists almost surely and it is a constant. We call this constant $L_1(\nu)$ the *first Lyapunov exponent* of the multiplicative process determined by the measure ν .

Some of the central problems regarding Lyapunov exponents of such processes concern the regularity of the map $\nu \mapsto L_1(\nu)$, as well as the availability of other probability laws. More specifically, the almost sure convergence (which implies convergence in probability) in (1.1) is the multiplicative analogue of the law of large numbers. It is then natural to ask if there is an explicit rate of convergence in probability, that is, if a large-deviations-type (LDT) estimate holds. Moreover, it is also natural to ask if a central limit theorem (CLT) holds as well.

Mathematics Subject Classification 2020: 37H15 (primary); 37H05, 37D25, 47B39, 37H30, 81Q10, 82B44 (secondary).

Keywords: random linear cocycles, Lyapunov exponents, large deviations, central limit theorem, discrete Schrödinger operators, the Anderson model.

Multiplicative processes as described above appear naturally in the study of the discrete Schrödinger operator with random potential on the integer lattice (or, more generally, on a band-integer lattice).

More precisely, let $\{v_n\}_{n \in \mathbb{Z}}$ be an i.i.d. sequence of real-valued random variables and consider the operator H on $l^2(\mathbb{Z}, \mathbb{R})$ given by

$$(Hu)_n := -(u_{n+1} + u_{n-1}) + v_n u_n \quad (1.2)$$

for $u = \{u_n\}_{n \in \mathbb{Z}} \in l^2(\mathbb{Z}, \mathbb{R})$.

The Schrödinger (or eigenvalue) equation $Hu = tu$ becomes a first-order matrix recurrence. Its (formal) solution is given by

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} v_n - t & -1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} v_1 - t & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}$$

thus leading to a multiplicative process in $SL_2(\mathbb{R})$ generated by the i.i.d. sequence of random matrices $\{A_n\}_{n \in \mathbb{Z}}$ where $A_n = \begin{pmatrix} v_n - t & -1 \\ 1 & 0 \end{pmatrix}$.

More general operators than (1.2) (e.g., Jacobi operators) or band lattice versions thereof lead to multiplicative processes in other, more general semigroups of matrices. Moreover, the properties of their Lyapunov exponents as discussed above (especially their Hölder continuity and the availability of large deviations) are directly relevant in the study of the spectral properties of the corresponding operator. See the survey [9] for a nice review of these connections.

Multiplicative processes of this or a more general type (i.e., non-independent) can be studied in the abstract framework of linear cocycles in ergodic theory. A base measure preserving dynamical system (X, μ, f) , where X is a compact metric space, $f: X \rightarrow X$ is continuous transformation and μ is an f -invariant, ergodic Borel probability measure on X , together with a bounded, measurable fiber map $A: X \rightarrow \text{Mat}_m(\mathbb{R})$ determine the skew-product transformation $F: X \times \mathbb{R}^m \rightarrow X \times \mathbb{R}^m$,

$$F(x, v) = (f(x), A(fx)v).$$

The new dynamical system F is called a *linear cocycle*. Its iterates are $F^n(x, v) = (f^n(x), A^n(x)v)$, where for all $n \in \mathbb{N}$,

$$A^n(x) := A(f^n x) \cdots A(f^2 x)A(fx).$$

The first Lyapunov exponent of a linear cocycle F measures the asymptotic exponential growth of its fiber iterates. It is defined via the Furstenberg–Kesten theorem as the μ -a.e. limit

$$L_1(F) = L_1(A) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A^n(x)\|$$

provided A satisfies the integrability condition $\int_X \log^+ \|A\| d\mu < \infty$.

Moreover, if the norm (that is, the first singular value) of the fiber iterates is replaced by the second singular value, the corresponding μ -a.e. limit above still exists, it is called the *second Lyapunov exponent* and it is denoted by $L_2(F) = L_2(A)$.

An i.i.d. multiplicative process in $\text{Mat}_m(\mathbb{R})$ can be identified with a locally constant linear cocycle over a Bernoulli shift. Indeed, let $\mathcal{A}$, the space of symbols (or alphabet), be a compact metric space and let p be a probability measure on $\mathcal{A}$. Denote by $X := \mathcal{A}^{\mathbb{Z}}$ the space of bi-infinite sequences $\omega = \{\omega_n\}_{n \in \mathbb{Z}}$ in this alphabet, which we endow with the product measure $\mu = p^{\mathbb{Z}}$. Let $\sigma: X \rightarrow X$ be the corresponding forward shift $\sigma\omega = \{\omega_{n+1}\}_{n \in \mathbb{Z}}$. Then (X, μ, σ) is a measure preserving, ergodic dynamical system called a *Bernoulli shift*. A continuous function $\mathcal{A} \ni \omega_0 \mapsto A_{\omega_0} \in \text{Mat}_m(\mathbb{R})$ determines the locally constant fiber map $A: X \rightarrow \text{Mat}_m(\mathbb{R})$, $A(\omega) = A_{\omega_0}$, which in turn determines a linear cocycle over the Bernoulli shift, referred to as a (random) Bernoulli cocycle.

In this paper we restrict to the case of a *finite* alphabet and two-dimensional matrices. That is, $\mathcal{A} = \{1, \dots, k\}$ for some $k \geq 2$, $p = (p_1, \dots, p_k)$ is a probability vector with $p_i > 0$ for all i , while the fiber map A is determined by a k -tuple $\underline{A} = (A_1, \dots, A_k) \in \text{Mat}_2(\mathbb{R})^k$. We identify the corresponding Bernoulli cocycle with the tuple $\underline{A}$ and denote by $L_i(\underline{A}) = L_i(\underline{A}, p)$, $i \in \{1, 2\}$, its Lyapunov exponents.

More generally, we also consider linear cocycles over a Markov shift. That is, let P be a (left) stochastic matrix, i.e., $P = (p_{ij})_{1 \leq i, j \leq k}$ with $p_{ij} \geq 0$ and $\sum_{i=1}^k p_{ij} = 1$ for all $1 \leq j \leq k$. Given a P -stationary probability vector q , i.e., $q = Pq$, the pair (P, q) determines a probability measure μ on X for which the process $\xi_n: X \rightarrow \mathcal{A}$, $\xi_n(\omega) := \omega_n$ is a stationary Markov chain in $\mathcal{A}$ with probability transition matrix P and initial distribution law q . Then (X, σ, μ) is a measure preserving dynamical system called a *Markov shift*.

Given a finite word $(i_0, i_1, \dots, i_n) \in \mathcal{A}^{n+1}$ and $k \in \mathbb{Z}$, the set

$$[k; i_0, i_1, \dots, i_n] := \{\omega \in X: \omega_{k+l} = i_l \text{ for all } 0 \leq l \leq n\}$$

is called a *cylinder of length $n + 1$ in X* . Its (Markov) measure is then

$$\mu([k; i_0, i_1, \dots, i_n]) = q_{i_0} p_{i_i, i_0} \cdots p_{i_n, i_{n-1}}.$$

We will assume that the matrix P is primitive, i.e., there exists $n \geq 1$ such that $p_{ij}^n > 0$ for all entries of the power matrix P^n . Then $\lim_{n \rightarrow \infty} p_{ij}^n = q_i > 0$ for all $1 \leq i, j \leq k$ and the corresponding Markov shift (X, μ, σ) is ergodic and mixing, see [25, Theorems 7.2.8 and 7.2.11].

As before, a k -tuple $\underline{A} = (A_1, \dots, A_k) \in \text{Mat}_2(\mathbb{R})^k$ determines a locally constant linear cocycle over this base dynamics, which we refer to as a (random) Markov cocycle. Its Lyapunov exponents are denoted by $L_i(\underline{A}) = L_i(\underline{A}, P, q)$.

When we restrict to *invertible* random linear cocycles, that is, when $\underline{A} = (A_1, \dots, A_k) \in \text{GL}_2(\mathbb{R})^k$, the problems on Lyapunov exponents formulated above are well understood, having been the subject of many works throughout the years. See Table 1 below for a more detailed picture of available results. The case when all matrices $A_1, \dots, A_k$ defining the cocycle $\underline{A}$ are singular (non-invertible) is treated easily.

We thus consider the case of random cocycles with *both singular and invertible* components, and assume, moreover, that the invertible components have positive determinant. In a recent work, see [10], we show that the first Lyapunov exponent of such cocycles behaves rather differently from the invertible case: when it comes to its regularity, it exhibits a dichotomy in the spirit of Bochi–Mañé’s, from being analytic to being discontinuous, see [10, Theorem 1.1 and Corollary 1.1].

As explained in [10], random cocycles with both singular and invertible components can be used to provide a formal model of a one-parameter family of Bernoulli cocycles $\underline{A}_t \in \text{Mat}_2(\mathbb{R})^2$ associated to the random Schrödinger operator H whose potential takes two values: ∞ and a finite value a , as in Craig and Simon [8, Example 3].¹

Indeed, given a probability vector $(p, 1 - p)$ with $p \in (0, 1)$ and $a \in \mathbb{R}$, define the random cocycle $\underline{A}_t = (A_t(1), A_t(2)) \in \text{Mat}_2(\mathbb{R})^2$ by $A_t(1) \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $A_t(2) = \begin{pmatrix} a^{-t} & -1 \\ 1 & 0 \end{pmatrix}$.

Moreover, let H_λ be the Schrödinger operator with random potential $\{v_j\}_{j \in \mathbb{Z}}$ taking the value λ with probability p and the value a with probability $1 - p$, and denote by $S_{t,\lambda}(j) = \begin{pmatrix} v_j^{-t} & -1 \\ 1 & 0 \end{pmatrix}$ its corresponding random Schrödinger cocycle.

Then $\underline{A}_t$ is the asymptotic limit as $\lambda \rightarrow \infty$ of the *rescaled* random Schrödinger cocycle $\underline{A}_{t,\lambda} = (\underline{A}_{t,\lambda}(1), \underline{A}_{t,\lambda}(2))$, where $\underline{A}_{t,\lambda}(j) = \lambda_j^{-1} S_{t,\lambda}(j)$ and $\lambda_j = \lambda$ or 1 with probability $p, 1 - p$. Note the following relation between the Lyapunov exponents: $L_1(S_{t,\lambda}) = p \log \lambda + L_1(\underline{A}_{t,\lambda})$.

For each of the limiting cocycles $\underline{A}_t$, one component is invertible and the other one is singular. The map $t \mapsto L_1(\underline{A}_t)$ is a.e. discontinuous on $[a - 2, a + 2]$, the limit (as $\lambda \rightarrow \infty$) of the spectrum of the operator H_λ . This correlates well with the fact (proven by Craig and Simon) that the integrated density of states of the operator H is discontinuous. See [10, Example 5.4] for more details.

The main purpose of this paper is to establish statistical properties (large-deviations-type estimates and a central limit theorem) for random (Bernoulli and Markov) cocycles with both singular and invertible components. Let $\mathcal{A} = \mathcal{A}_{\text{sing}} \sqcup \mathcal{A}_{\text{inv}}$ be a partition into two non-empty sets of the alphabet and let

$$\mathcal{M} := \{\underline{A} = (A_1, \dots, A_k) \in \text{Mat}_2(\mathbb{R})^{\mathcal{A}} : \text{rank } A_i = 1 \text{ for all } i \in \mathcal{A}_{\text{sing}}, \\ \det(A_j) > 0 \text{ for all } j \in \mathcal{A}_{\text{inv}}\}.$$

¹Other versions of such an operator (with more values) can similarly be considered.

The set $\mathcal{M}$ is an analytic manifold (see [10, end of Section 4]), thus it is equipped with a Lebesgue measure.

Theorem 1.1. *For Lebesgue almost every random cocycle $\underline{A} \in \mathcal{M}$, $L_1(\underline{A}) > -\infty$ and for every $\varepsilon > 0$ we have*

$$\mu\left\{\omega \in X: \left|\frac{1}{n} \log \|A^n(\omega)\| - L_1(\underline{A})\right| > \varepsilon\right\} \leq C e^{-c_0(\varepsilon)n^{1/3}}$$

where $C < \infty$, $c_0(\varepsilon) > 0$ is an explicit function of ε and μ is the Bernoulli (or Markov) measure on X defined above.

For the next result, in the Markov case only, we assume that the subshift of finite type $\underline{A}_{\text{inv}}$, defined by the same underlying graph but retaining only edges labelled by invertible matrices in $\underline{A}$, is topologically mixing. This technical condition allows the use of Avila–Bochi–Yoccoz theory via the work in [10]. As in that work, the assumption is purely technical: it can be removed here, in [10], as well as in Avila–Bochi–Yoccoz, where uniform hyperbolicity is characterised by invariant multi-cones. For simplicity, we impose it and leave the general Markov case for future work. In the Bernoulli case, this assumption is unnecessary.

Theorem 1.2. *For Lebesgue almost every random cocycle $\underline{A} \in \mathcal{M}$, $L_1(\underline{A}) > -\infty$ and there exists $\sigma > 0$ such that the following convergence in distribution to the normalised Gaussian holds:*

$$\frac{\log \|A^n\| - nL_1(\underline{A})}{\sigma \sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1).$$

We will actually prove something stronger. Given any random cocycle $\underline{A} \in \mathcal{M}$ with no null words, if $t \mapsto \underline{A}_t$ is a one-parameter family of such cocycles passing through $\underline{A}$ and satisfying a certain monotonicity property (positive winding), then for almost every parameter t , the cocycle $\underline{A}_t$ satisfies LDT estimates and a CLT.

The main tools in deriving such limit laws for random linear cocycles and other types of dynamical systems are the spectral properties of either the Markov transition operator or the Ruelle transfer operator associated to the system. This spectral method goes back to the work of Nagaev [22] on CLTs for stationary Markov chains and it is described in an abstract setting in the book [18] of Hennion and Hervé. The main ingredient required by the method is the quasi-compactness (which is equivalent to the spectral gap property) of the transition/transfer operator on an appropriate space of observables.

Recently, Cai, Duarte, and Klein [7] obtained an *effective* abstract LDT estimate for Markov chains under a much weaker hypothesis, a certain mixing property. Theorem 1.1 will be established using this result, its effectiveness being crucial in our argument. To establish Theorem 1.2 we will use an abstract CLT due to Gordin and Livšic [17].

	R-LE	LDT	CLT
rank = 2	(weak) Hölder ³	yes ⁴	yes ⁵
rank = 1	C^ω ([10, Corollary 4.4])	yes (Remark 3.4)	yes (Remark 3.4)
rank = 1&2	disc ([10, Corollary 1.1])	yes (Theorem 1.1)	yes (Theorem 1.2)

Table 1. Random (Bernoulli) two-dimensional cocycles.

It turns out that these abstract results are not directly applicable to cocycles $\underline{A} \in \mathcal{M}$, precisely because of their singularities. We will employ some rather involved parameter elimination arguments that eventually lead to limit laws for almost every parameter.

Moreover, we will also derive an explicit formula for the (unique) stationary measure of the transition operator (see Proposition 2.1) and a Furstenberg-type formula for the first Lyapunov exponent (see Theorem 2.2). These results are crucial in the arguments used for establishing the limit laws and also interesting in themselves. We note, moreover, that standard results and methods of Furstenberg’s theory (such as the formula describing the first Lyapunov exponent via stationary measures) are *not* applicable in our setting of singular cocycles, and will be studied independently in this work.

Together with the results in our first paper [10] on this topic, as well as the previously available works on invertible matrix components, we are now able to present an almost complete picture regarding the crucial problems on Lyapunov exponents described above, in the setting of random cocycles driven by a finitely supported measure on $\text{Mat}_2(\mathbb{R})$.

The table below thus summarises what it is known regarding the minimal regularity of the Lyapunov exponent (R-LE), namely its modulus of continuity² or whether it is discontinuous (Disc.) as well as the availability of large-deviations-type (LDT) estimates and of a central limit theorem (CLT) for $\text{Mat}_2(\mathbb{R})$ -valued Bernoulli cocycles $\underline{A} \in \text{Mat}_2(\mathbb{R})^k$, $k \geq 2$. For the purpose of this table, we assume that $L_1 > L_2 \geq -\infty$. There are three possibilities for such a given cocycle $\underline{A}$: rank = 2, meaning its components are all invertible; rank = 1, meaning its components are all singular; rank = 1&2, the case treated in this paper, where some components are invertible and some singular.

²Given a metric space (M, d) , a function $\varphi: M \rightarrow \mathbb{R}$ is said to be *weak-Hölder continuous* if $|\varphi(x) - \varphi(y)| \leq C \exp(-\alpha \log^b \frac{1}{d(x,y)})$ for some $C < \infty$, $\alpha, b \in (0, 1]$ and all $x, y \in M$. When $b = 1$, this corresponds to α -Hölder continuity. Moreover, if $|\varphi(a) - \varphi(x)| \leq C d(a, x)^\alpha$ holds for a given point a and all x , we call φ *pointwise Hölder at a* .

The concept of projectively uniformly hyperbolic cocycle (or dominated splitting property) is defined and characterised for this setting in [10, Section 2], extending results of Avila, Bochi, and Yoccoz [1] for $SL_2(\mathbb{R})$ -valued cocycles. Recall that in this case the Lyapunov exponent is analytic and the statistical properties hold trivially (they are easily reduced to their classical analogues for additive i.i.d. processes).

Moreover, in this paper we show that the statistical properties on the second and third lines of Table 1 also hold for mixing *Markov* cocycles. Furthermore, in this Markov setting, the results on the first line are available only in the generic (irreducible) case (see [11, Chapter 5]), but we expect them to still hold without the generic assumption, as is the case in the Bernoulli setting, see [12].

The rest of the paper is organised as follows. In Section 2 we show that given a random cocycle with both invertible and singular components, the induced projective Markov chain is a renewal process with a unique and explicitly defined stationary measure supported on a discrete set. Moreover, the transition stochastic kernel is uniformly ergodic in the sense that its iterated kernel converges to the stationary measure in the total variation norm. Furthermore, we establish a Furstenberg-type formula for the Lyapunov exponent. Using parameter elimination arguments, in Section 3 we establish statistical properties (large deviations and a central limit theorem) for one-parameter families of such cocycles satisfying a positive winding condition. This in particular implies the availability of these limit laws for Lebesgue almost every random cocycle in finite symbols, i.e., Theorems 1.1 and 1.2.

Dedication. The last named author would like to dedicate this paper to the memory of Michael Goldstein, whose work has deeply influenced his own. My interest in problems concerning the regularity of Lyapunov exponents, large deviations estimates, spectral theory of Schrödinger operators and the relations between these topics, as well as the subtleties involved in performing delicate parameter elimination arguments, all stem from my exposure to Michael's profound work (e.g. [6, 15, 16]), which has permeated this and indeed most of my research alone or in collaboration with others, throughout the years.

³Locally Hölder for quasi-irreducible cocycles [11, 20]; locally weak-Hölder in the remaining case [12]; pointwise Hölder always [23].

⁴Locally uniform LDT of exponential type in the quasi-irreducible case [11]; locally uniform LDT of sub-exponential type in the remaining case [12]; non-uniform LDT of exponential type holds always [12].

⁵See [3, 4, 19].

2. Stationary measures

In this section we establish formulas for the stationary measure and the Lyapunov exponent of a random linear cocycle $F: X \times \mathbb{R}^2 \rightarrow X \times \mathbb{R}^2$ with both singular and invertible components.

Given a finite alphabet $\mathcal{A} = \{1, \dots, k\}$ and a partition into two non-empty sets $\mathcal{A} = \mathcal{A}_{\text{sing}} \sqcup \mathcal{A}_{\text{inv}}$, let $\underline{A} := (A_i)_{i \in \mathcal{A}} \in \text{Mat}_2(\mathbb{R})^k$ be a k -tuple such that $\text{rank } A_i = 1$ if $i \in \mathcal{A}_{\text{sing}}$ and $\text{rank } A_i = 2$ if $i \in \mathcal{A}_{\text{inv}}$. Note that in this section we do not make any other assumptions on $\underline{A}$.

Let $P = \{p_{ij}\}_{1 \leq i, j \leq k}$ be a primitive left stochastic matrix (that is, p_{ij} encodes the transition probability from state j to state i) and let $q = \{q_i\}_{1 \leq i \leq k}$ be its unique stationary probability vector. The triplet $(\underline{A}, P, q)$ determines the (Markov) random linear cocycle $F: X \times \mathbb{R}^2 \rightarrow X \times \mathbb{R}^2$, where $X = \mathcal{A}^{\mathbb{Z}}$ is endowed with the Markov measure μ with initial distribution q and transition probability given by P and $F(\omega, v) = (\sigma\omega, A(\sigma\omega)v) = (\sigma\omega, A_{\omega_1}v)$.

Let $\mathbb{P}^1 := \mathbb{P}(\mathbb{R}^2)$ denote the projective line. An element of $\mathbb{P}^1$ will be denoted by $\hat{v}$, where v is a non-zero vector (or a one-dimensional subspace) in $\mathbb{R}^2$. Given an invertible matrix $A \in \text{Mat}_2(\mathbb{R})$, its induced projection action $\hat{A}: \mathbb{P}^1 \rightarrow \mathbb{P}^1$ is given by $\hat{A}\hat{v} := \widehat{Av}$. If A has rank 1 (that is, if it is non-zero and non-invertible), we define its projective action as the constant map $\hat{A}\hat{v} := \hat{r}$, where $r = \text{Range}(A)$. Note that we are formally removing the discontinuity of the projective action of a singular matrix at the kernel.

For a random cocycle $\underline{A} = (A_i)_{i \in \mathcal{A}} \in \text{Mat}_2(\mathbb{R})^{\mathcal{A}}$, we will use the notations r_i and k_i to represent, respectively, the range and the kernel of A_i , as well as, when convenient, a unit vector belonging to these one-dimensional subspaces.

Remark 2.1. In the Bernoulli case, if we consider an i.i.d. sequence of $\mathcal{A}$ -valued random variables $\{\xi_n\}_n$, then the sequence of return times to $\mathcal{A}_{\text{sing}}$ forms a discrete time renewal process, see [2, Definition 2.1] for the meaning of this concept. Note that at these return times, when $\xi_n = i \in \mathcal{A}_{\text{sing}}$, the projective action induced by $\underline{A}$ is reset to the range of A_i . This simple observation lies at the heart of all the explicit formulas presented below. Since in the Markov case the reduction to the setting of [2] is less evident, the results of this section will be presented for this case. The Bernoulli case can be treated similarly, but in a much more straightforward way; moreover, all relevant quantities have much simpler expressions.

Let $L^\infty(\mathcal{A} \times \mathbb{P}^1)$ be the Banach space of bounded and measurable functions $\varphi: \mathcal{A} \times \mathbb{P}^1 \rightarrow \mathbb{R}$ endowed with the sup norm denoted by $\|\cdot\|_\infty$.

The cocycle $(\underline{A}, P, q)$ determines the operator $\mathcal{Q}: L^\infty(\mathcal{A} \times \mathbb{P}^1) \rightarrow L^\infty(\mathcal{A} \times \mathbb{P}^1)$ defined by

$$\begin{aligned} (\mathcal{Q}\varphi)(j, \hat{v}) &:= \sum_{i \in \mathcal{A}} \varphi(i, \hat{A}_i \hat{v}) p_{ij} \\ &= \sum_{i \in \mathcal{A}_{\text{inv}}} \varphi(i, \hat{A}_i \hat{v}) p_{ij} + \sum_{i \in \mathcal{A}_{\text{sing}}} \varphi(i, \hat{r}_i) p_{ij}. \end{aligned}$$

Moreover, we write $\mathcal{Q} = \mathcal{Q}_{\text{inv}} + \mathcal{Q}_{\text{sing}}$, where the operators $\mathcal{Q}_{\text{inv}}$ and $\mathcal{Q}_{\text{sing}}$ are given respectively by the two terms above.

The operator $\mathcal{Q}$ is clearly linear, positive and it takes the constant function $\mathbb{1}$ to itself; in other words, it is a *Markov operator*. Let us recall the concept of stationary measure for such an operator.

Definition 2.1. A measure $\eta \in \text{Prob}(\mathcal{A} \times \mathbb{P}^1)$ is $\mathcal{Q}$ -stationary if for all observables $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$,

$$\int \mathcal{Q}\varphi \, d\eta = \int \varphi \, d\eta.$$

In this case, we also call η *stationary relative to the cocycle* $(\underline{A}, P, q)$.

By general principles, a stationary measure always exists. In our setting it will be shown to be unique.

Let $\pi: \mathcal{A} \times \mathbb{P}^1 \rightarrow \mathcal{A}$ denote the canonical projection in the first coordinate. If $\eta \in \text{Prob}(\mathcal{A} \times \mathbb{P}^1)$ is $\mathcal{Q}$ -stationary, then its push-forward measure via π is the P -stationary measure q on $\mathcal{A}$, that is, $\pi_*\eta = q$.

Indeed, the left stochastic matrix P induces the Markov operator $u \mapsto uP$ on $L^\infty(\mathcal{A}) \simeq \mathbb{R}^k$. For a (column) vector q' , $Pq' = q'$ if and only if $uPq' = uq'$ for every (row) vector $u \in \mathbb{R}^k$, so q' is P -stationary if and only if it is stationary relative to its corresponding Markov operator. Moreover, given any $u \in L^\infty(\mathcal{A})$ and applying the fact that η is $\mathcal{Q}$ -stationary to the observable $\varphi := u \circ \pi$ it follows immediately that $\int uP \, d(\pi_*\eta) = \int u \, d(\pi_*\eta)$, showing that $\pi_*\eta$ is P -stationary. But since P is primitive, it has a unique stationary probability vector, so $\pi_*\eta = q$.

Since η projects down via π to q , we can consider its disintegration $\{\eta_i\}_{i \in \mathcal{A}} \subset \text{Prob}(\mathbb{P}^1)$, which is characterised by

$$\int_{\mathcal{A} \times \mathbb{P}^1} \varphi(i, \hat{v}) \, d\eta(i, \hat{v}) = \sum_{i \in \mathcal{A}} q_i \int_{\mathbb{P}^1} \varphi(i, \hat{v}) \, d\eta_i(\hat{v}) \quad \text{for all } \varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1).$$

Then

$$\eta = \sum_{i \in \mathcal{A}} q_i \delta_i \times \eta_i \quad \text{and} \quad \eta_i(E) = \frac{1}{q_i} \eta(\{i\} \times E) \quad \text{for all } i \in \mathcal{A}, E \subset \mathbb{P}^1 \text{ Borel.}$$

Given $n \geq 1$, $s \in \mathcal{A}_{\text{sing}}$ and $l \in \mathcal{A}$, let $\mathcal{B}_n(s, l)$ denote the set of words $\underline{\omega} = (\omega_0, \dots, \omega_n)$ of length $n + 1$ such that $\omega_0 = s$, $\omega_n = l$ and $\omega_i \in \mathcal{A}_{\text{inv}}$ for all $i = 1, \dots, n - 1$. For such a word we write $A^n(\underline{\omega}) := A_{\omega_n} \dots A_{\omega_2} A_{\omega_1}$ and also put $p(\underline{\omega}) := p_{\omega_n \omega_{n-1}} \dots p_{\omega_1 \omega_0}$.

With these notations, we have the following explicit formula for a (or, a-posteriori, *the*) $\mathcal{Q}$ -stationary measure.

Proposition 2.1. *Let $\eta = \sum_{j \in \mathcal{A}} q_j \delta_j \times \eta_j$ where for all $j \in \mathcal{A}$,*

$$\eta_j := \frac{1}{q_j} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) \delta_{\hat{A}^n(\underline{\omega}) \hat{r}_s}.$$

Then η is an $(\underline{A}, P, q)$ -stationary probability measure on $\mathcal{A} \times \mathbb{P}^1$.

Proof. We first show that for all $j \in \mathcal{A}$,

$$\sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) = q_j. \quad (2.1)$$

This will then imply that η_j is a probability measure on $\mathbb{P}^1$ for all $j \in \mathcal{A}$, hence η is a probability measure on $\mathcal{A} \times \mathbb{P}^1$.

Indeed, given $j \in \mathcal{A}$, consider the cylinder $[0; j]$ determined by the word (j) , so $\mu([0; j]) = q_j$. This cylinder is partitioned, modulo a zero measure set (mod 0), into cylinders $[-n; s\underline{\omega}'j]$, determined by words $(s, \underline{\omega}', j)$ where $s \in \mathcal{A}_{\text{sing}}$ and $\underline{\omega}' = (\omega_1, \dots, \omega_{n-1})$ with $\omega_i \in \mathcal{A}_{\text{inv}}$ for all $1 \leq i \leq n - 1$, $n \geq 0$ (when $n = 1$, the word $(s, \underline{\omega}', j)$ is simply (s, j)). That is,

$$[0; j] = \bigsqcup_{s \in \mathcal{A}_{\text{sing}}} \bigsqcup_{n=1}^{\infty} \bigsqcup_{\underline{\omega} = (s, \underline{\omega}', j) \in \mathcal{B}_n(s, j)} [-n; s\underline{\omega}'j] \pmod{0}.$$

The cylinders $[-n; s\underline{\omega}'j]$ are mutually disjoint because $s \in \mathcal{A}_{\text{sing}}$ and $\omega_i \in \mathcal{A}_{\text{inv}}$ for all i . Modulo a zero measure set, their union is $[0; j]$ since, by the mixing of the Markov shift determined by (P, q) , almost every infinite word starting with the symbol j contains in its past a symbol $s \in \mathcal{A}_{\text{sing}}$. Moreover, $\mu([-n; s\underline{\omega}'j]) = q_s p(s, \underline{\omega}', j)$ which establishes the identity (2.1).

It remains to verify that $\eta = \sum_{j \in \mathcal{A}} q_j \delta_j \times \eta_j$ is $\mathcal{Q}$ -stationary, that is, that

$$\int \mathcal{Q}\varphi d\eta = \int \varphi d\eta$$

for all $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$.

Firstly, note that if $j \in \mathcal{A}_{\text{sing}}$, then $\eta_j = \delta_{\hat{r}_j}$. Indeed, for all $n \geq 1$, $s \in \mathcal{A}_{\text{sing}}$, $\underline{\omega} \in \mathcal{B}_n(s, j)$ we have $\hat{A}^n(\underline{\omega})\hat{r}_s = \hat{A}_j \cdots \hat{A}_{\omega_1}\hat{r}_s = \hat{r}_j$, so using the identity (2.1),

$$\begin{aligned} \eta_j &= \frac{1}{q_j} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) \delta_{\hat{A}^n(\underline{\omega})\hat{r}_s} \\ &= \frac{1}{q_j} \left(\sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) \right) \delta_{\hat{r}_j} = \delta_{\hat{r}_j}. \end{aligned}$$

Secondly, for any observable $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$, we write

$$\begin{aligned} \int \varphi \, d\eta &= \sum_{j \in \mathcal{A}} q_j \int \varphi(j, \hat{v}) \, d\eta_j(\hat{v}) \\ &= \Sigma_{\text{inv}}(\varphi) + \Sigma_{\text{sing}}(\varphi), \end{aligned}$$

where

$$\begin{aligned} \Sigma_{\text{inv}}(\varphi) &:= \sum_{j \in \mathcal{A}_{\text{inv}}} q_j \int \varphi(j, \hat{v}) \, d\eta_j(\hat{v}) \\ &= \sum_{j \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) \varphi(j, \hat{A}^n(\underline{\omega})\hat{r}_s) \end{aligned}$$

and

$$\Sigma_{\text{sing}}(\varphi) := \sum_{j \in \mathcal{A}_{\text{sing}}} q_j \int \varphi(j, \hat{v}) \, d\eta_j(\hat{v}) = \sum_{j \in \mathcal{A}_{\text{sing}}} q_j \varphi(j, \hat{r}_j).$$

Recall that

$$\mathcal{Q}\varphi = \mathcal{Q}_{\text{inv}}\varphi + \mathcal{Q}_{\text{sing}}\varphi,$$

where

$$\begin{aligned} \mathcal{Q}_{\text{inv}}\varphi(j, \hat{v}) &= \sum_{i \in \mathcal{A}_{\text{inv}}} p_{ij} \varphi(i, \hat{A}_i \hat{v}), \\ \mathcal{Q}_{\text{sing}}\varphi(j, \hat{v}) &= \sum_{i \in \mathcal{A}_{\text{sing}}} p_{ij} \varphi(i, \hat{r}_i). \end{aligned}$$

Then,

$$\begin{aligned} \int \mathcal{Q}\varphi \, d\eta &= \Sigma_{\text{inv}}(\mathcal{Q}\varphi) + \Sigma_{\text{sing}}(\mathcal{Q}\varphi) \\ &= \Sigma_{\text{inv}}(\mathcal{Q}_{\text{inv}}\varphi) + \Sigma_{\text{inv}}(\mathcal{Q}_{\text{sing}}\varphi) + \Sigma_{\text{sing}}(\mathcal{Q}_{\text{inv}}\varphi) + \Sigma_{\text{sing}}(\mathcal{Q}_{\text{sing}}\varphi). \end{aligned}$$

We evaluate each of the four terms above separately. We have

$$\begin{aligned}
 \Sigma_{\text{inv}}(\mathcal{Q}_{\text{inv}}\varphi) &= \sum_{j \in \mathcal{A}_{\text{inv}}} \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) p_{ij} \varphi(i, \hat{A}_i \hat{A}^n(\underline{\omega}) \hat{r}_s) \\
 &= \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega}' \in \mathcal{B}_{n+1}(s, i)} p(\underline{\omega}') \varphi(i, \hat{A}^n(\underline{\omega}') \hat{r}_s) \\
 &= \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=2}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, i)} p(\underline{\omega}) \varphi(i, \hat{A}^n(\underline{\omega}) \hat{r}_s).
 \end{aligned}$$

The first line turned into the second by adding the letter i to the end of the word $\underline{\omega} \in \mathcal{B}_n(s, j)$ with $j \in \mathcal{A}_{\text{inv}}$, so that $\underline{\omega}' := (\underline{\omega}, i) \in \mathcal{B}_{n+1}(s, i)$; note that $p(\underline{\omega}') = p_{ij} p(\underline{\omega})$. Similarly,

$$\begin{aligned}
 \Sigma_{\text{inv}}(\mathcal{Q}_{\text{sing}}\varphi) &= \sum_{j \in \mathcal{A}_{\text{inv}}} \sum_{i \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, j)} p(\underline{\omega}) p_{ij} \varphi(i, \hat{r}_i) \\
 &= \sum_{i \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega}' \in \mathcal{B}_{n+1}(s, i)} p(\underline{\omega}') \varphi(i, \hat{r}_i) \\
 &= \sum_{i \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=2}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, i)} p(\underline{\omega}) \varphi(i, \hat{r}_i).
 \end{aligned}$$

Then,

$$\begin{aligned}
 \Sigma_{\text{sing}}(\mathcal{Q}_{\text{inv}}\varphi) &= \sum_{j \in \mathcal{A}_{\text{sing}}} q_j \sum_{i \in \mathcal{A}_{\text{inv}}} p_{ij} \varphi(i, \hat{A}_i \hat{r}_j) \\
 &= \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{\underline{\omega} \in \mathcal{B}_1(s, i)} p(\underline{\omega}) \varphi(i, \hat{A}^1(\underline{\omega}) \hat{r}_s)
 \end{aligned}$$

where in passing from the second to the last line we changed the index j for s and used the fact that (s, i) is the only word $\underline{\omega} \in \mathcal{B}_1(s, i)$ and in this case $p(\underline{\omega}) = p_{is}$ and $A^1(\underline{\omega}) = A_i$. Finally,

$$\begin{aligned}
 \Sigma_{\text{sing}}(\mathcal{Q}_{\text{sing}}\varphi) &= \sum_{j \in \mathcal{A}_{\text{sing}}} q_j \sum_{i \in \mathcal{A}_{\text{sing}}} p_{ij} \varphi(i, \hat{r}_i) \\
 &= \sum_{i \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{\underline{\omega} \in \mathcal{B}_1(s, i)} p(\underline{\omega}) \varphi(i, \hat{r}_i).
 \end{aligned}$$

Note that

$$\Sigma_{\text{inv}}(\mathcal{Q}_{\text{inv}}\varphi) + \Sigma_{\text{sing}}(\mathcal{Q}_{\text{inv}}\varphi) = \Sigma_{\text{inv}}(\varphi)$$

and using again (2.1),

$$\begin{aligned}\Sigma_{\text{inv}}(\mathcal{Q}_{\text{sing}}\varphi) + \Sigma_{\text{sing}}(\mathcal{Q}_{\text{sing}}\varphi) &= \sum_{i \in \mathcal{A}_{\text{sing}}} \left(\sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\omega \in \mathcal{B}_n(s,i)} p(\omega) \right) \varphi(i, \hat{r}_i) \\ &= \sum_{i \in \mathcal{A}_{\text{sing}}} q_i \varphi(i, \hat{r}_i) = \Sigma_{\text{sing}}(\varphi),\end{aligned}$$

which concludes the proof. $\blacksquare$

Theorem 2.1. *If the cocycle (A, P, q) has both singular and invertible components, then the corresponding Markov operator $\mathcal{Q}$ is uniformly ergodic, i.e., there exist constants $C < \infty$ and $a > 0$ such that*

$$\left\| \mathcal{Q}^n \varphi - \int \varphi d\eta \right\|_{\infty} \leq C e^{-an} \|\varphi\|_{\infty}$$

for all $n \in \mathbb{N}$ and $\varphi \in L^{\infty}(\mathcal{A} \times \mathbb{P}^1)$, where η is any $\mathcal{Q}$ -stationary measure.

Proof. Note that $L^{\infty}(\mathcal{A}) \simeq \mathbb{R}^k$, seen as the subspace of observables φ that do not depend on the projective variable, is invariant under the Markov operator $\mathcal{Q}$; when restricted to this subspace, $\mathcal{Q}$ becomes the operator $u \mapsto uP$ induced by the left stochastic matrix P . Recall also that η projects down to q , the stationary probability vector of P . Since P is primitive, for some constants $C < \infty$ and $\sigma \in (0, 1)$ (these constants may change a few times throughout the argument)

$$\|uP^n - u \cdot q\|_{\infty} \leq C\sigma^n \|u\|_{\infty} \quad (2.2)$$

for all $n \in \mathbb{N}$ and $u \in \mathbb{R}^k \simeq L^{\infty}(\mathcal{A})$.

We show that for all $n \in \mathbb{N}$ there is a linear operator $T_n: L^{\infty}(\mathcal{A} \times \mathbb{P}^1) \rightarrow L^{\infty}(\mathcal{A})$ such that for all $\varphi \in L^{\infty}(\mathcal{A} \times \mathbb{P}^1)$,

$$\mathcal{Q}^n \varphi = \mathcal{Q}_{\text{inv}}^n \varphi + T_n \varphi. \quad (2.3)$$

In other words, $\mathcal{Q}^n \varphi - \mathcal{Q}_{\text{inv}}^n \varphi$ does not depend on the projective variable.

Indeed, when $n = 1$,

$$\mathcal{Q}\varphi = \mathcal{Q}_{\text{inv}}\varphi + \mathcal{Q}_{\text{sing}}\varphi,$$

where

$$\mathcal{Q}_{\text{sing}}\varphi(j, \hat{v}) = \sum_{i \in \mathcal{A}_{\text{sing}}} \varphi(i, \hat{r}_i) p_{ij} =: T_1 \varphi(j)$$

does not depend on $\hat{v}$. By induction, if we assume that for all observables φ we can write $\mathcal{Q}^n \varphi = \mathcal{Q}_{\text{inv}}^n \varphi + T_n \varphi$ with $T_n \varphi$ not depending on the projective variable, then

$$\begin{aligned}\mathcal{Q}^{n+1} \varphi &= \mathcal{Q}(\mathcal{Q}^n \varphi) = \mathcal{Q}(\mathcal{Q}_{\text{inv}}^n \varphi + T_n \varphi) = \mathcal{Q}(\mathcal{Q}_{\text{inv}}^n \varphi) + \mathcal{Q}(T_n \varphi) \\ &= \mathcal{Q}_{\text{inv}}(\mathcal{Q}_{\text{inv}}^n \varphi) + \mathcal{Q}_{\text{sing}}(\mathcal{Q}_{\text{inv}}^n \varphi) + (T_n \varphi)P = \mathcal{Q}_{\text{inv}}^{n+1} \varphi + T_{n+1} \varphi,\end{aligned}$$

where

$$T_{n+1}\varphi := \mathcal{Q}_{\text{sing}}(\mathcal{Q}_{\text{inv}}^n\varphi) + (T_n\varphi)P \quad (2.4)$$

does not depend on the projective variable, thus establishing (2.3).

We show that $\mathcal{Q}_{\text{inv}}^n$ eventually contracts, so $\mathcal{Q}_{\text{inv}}^n\varphi$ converges exponentially fast to 0, while $T_n\varphi$ converges exponentially fast to a constant (in the sup norm), which immediately imply the conclusion of the theorem.

Since the stochastic matrix P is primitive, there is $N \geq 1$ such that $P_{ij}^N > 0$ for all $i, j \in \mathcal{A}$. In particular, since P^N is also stochastic,

$$\sigma_0 := \max_{j \in \mathcal{A}} \sum_{i \in \mathcal{A}_{\text{inv}}} P_{ij}^N < 1.$$

We show that $\mathcal{Q}_{\text{inv}}^N$ is a contraction on $L^\infty(\mathcal{A} \times \mathbb{P}^1)$. Indeed, given any observable φ and $(j, \hat{v}) \in \mathcal{A} \times \mathbb{P}^1$, clearly

$$\mathcal{Q}_{\text{inv}}^N\varphi(j, \hat{v}) = \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{\underline{\omega} \in \mathcal{B}_N(j, i)} \varphi(i, A^N(\underline{\omega})\hat{v})p(\underline{\omega}).$$

Then,

$$\begin{aligned} \|\mathcal{Q}_{\text{inv}}^N\varphi\|_\infty &\leq \|\varphi\|_\infty \max_{j \in \mathcal{A}} \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{\underline{\omega} \in \mathcal{B}_N(j, i)} p(\underline{\omega}) \\ &= \|\varphi\|_\infty \max_{j \in \mathcal{A}} \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{\underline{\omega}' \in \mathcal{A}_{\text{inv}}^{N-1}} p(j \underline{\omega}' i) \\ &\leq \|\varphi\|_\infty \max_{j \in \mathcal{A}} \sum_{i \in \mathcal{A}_{\text{inv}}} \sum_{\underline{\omega}' \in \mathcal{A}^{N-1}} p(j \underline{\omega}' i) \\ &= \|\varphi\|_\infty \max_{j \in \mathcal{A}} \sum_{i \in \mathcal{A}_{\text{inv}}} P_{ij}^N = \sigma_0 \|\varphi\|_\infty. \end{aligned}$$

Given any integer n , we write $n = mN + r$ with $0 \leq r < N$ and use the fact that the operator norm of $\mathcal{Q}_{\text{inv}}$ is equal to 1 to conclude that there are $C < \infty$, $\sigma \in (0, 1)$ so that for all $n \in \mathbb{N}$ and $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$,

$$\|\mathcal{Q}_{\text{inv}}^n\varphi\|_\infty \leq C\sigma^n\|\varphi\|_\infty. \quad (2.5)$$

From (2.4) and (2.5), we get that for all $n \in \mathbb{N}$,

$$\|T_{n+1}\varphi - (T_n\varphi)P\|_\infty = \|\mathcal{Q}_{\text{sing}}(\mathcal{Q}_{\text{inv}}^n\varphi)\|_\infty \leq \|\mathcal{Q}_{\text{inv}}^n\varphi\|_\infty \leq C\sigma^n\|\varphi\|_\infty,$$

$$\begin{aligned} \|T_{n+2}\varphi - (T_n\varphi)P^2\|_\infty &\leq \|T_{n+2}\varphi - (T_{n+1}\varphi)P\|_\infty + \|(T_{n+1}\varphi - (T_n\varphi)P)P\|_\infty \\ &\leq C(\sigma^n + \sigma^{n+1})\|\varphi\|_\infty, \end{aligned}$$

and proceeding inductively, for all $n, k \in \mathbb{N}$,

$$\|T_{n+k}\varphi - (T_n\varphi)P^k\|_\infty \leq C'\sigma^n\|\varphi\|_\infty.$$

In particular,

$$\|T_{2n}\varphi - (T_n\varphi)P^n\|_\infty \leq C'\sigma^n\|\varphi\|_\infty. \quad (2.6)$$

By (2.2),

$$\|(T_n\varphi)P^n - (T_n\varphi) \cdot q\|_\infty \leq C\sigma^n\|(T_n\varphi)\|_\infty \leq 2C\sigma^n\|\varphi\|_\infty. \quad (2.7)$$

We have $\mathcal{Q}^n\varphi = \mathcal{Q}_{\text{inv}}^n\varphi + T_n\varphi$. Since the measure η is $\mathcal{Q}$ -stationary and it projects down to q , it follows that

$$\begin{aligned} \left| \int \varphi d\eta - (T_n\varphi) \cdot q \right| &= \left| \int \mathcal{Q}^n\varphi d\eta - \int T_n\varphi d\eta \right| \\ &= \left| \int \mathcal{Q}_{\text{inv}}^n\varphi d\eta \right| \leq \|\mathcal{Q}_{\text{inv}}^n\varphi\|_\infty \leq C\sigma^n\|\varphi\|_\infty. \end{aligned} \quad (2.8)$$

Combining (2.6), (2.7), and (2.8), we conclude that

$$\left\| T_{2n}\varphi - \int \varphi d\eta \right\|_\infty \leq C''\sigma^n\|\varphi\|_\infty,$$

which, together with (2.5) establishes our result (after choosing the constants C, a appropriately). ■

Corollary 2.1. *If the cocycle (A, P, q) has both singular and invertible components, then it admits a unique stationary measure.*

Proof. If η_1 and η_2 are $\mathcal{Q}$ -stationary, then Theorem 2.1 above applies to each of them, so for any $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$, $\mathcal{Q}^n\varphi$ converges uniformly to $\int \varphi d\eta_1$ and to $\int \varphi d\eta_2$. Thus, $\int \varphi d\eta_1 = \int \varphi d\eta_2$ for all observables φ , showing that $\eta_1 = \eta_2$. ■

Lemma 2.1. *Given rank one matrices $B_1, B_2, \dots, B_n$ and a unit vector r_0 , we have that*

$$\|B_n \cdots B_1 r_0\| = \prod_{l=1}^n \|B_l r_{l-1}\|,$$

where r_{l-1} is a unit vector in the range of B_{l-1} .

Proof. We write $B_1 r_0 = \lambda_1 r_1$, thus $|\lambda_1| = \|B_1 r_0\|$. It follows that $B_2 B_1 r_0 = \lambda_1 B_2 r_1$, so

$$\|B_2 B_1 r_0\| = |\lambda_1| \|B_2 r_1\| = \|B_2 r_1\| \|B_1 r_0\|.$$

From here,

$$\|B_{n+1}B_n(B_{n-1}\cdots B_1r_0)\| = \|B_{n+1}r_n\| \|B_nB_{n-1}\cdots B_1r_0\|$$

and the conclusion follows by induction. $\blacksquare$

Next we derive a closed-form expression for the Lyapunov exponent which, together with the explicit formula for the stationary measure in Proposition 2.1 will eventually lead to a Furstenberg-type formula.

Lemma 2.2. *If the random cocycle $(\underline{A}, P, q)$ has both singular and invertible components, then*

$$L_1(\underline{A}) = \sum_{i \in \mathcal{A}_{\text{sing}}} \sum_{j \in \mathcal{A}_{\text{sing}}} \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(i,j)} q_i p(\underline{\omega}) \log \|A^n(\underline{\omega})r_i\|.$$

Proof. Consider the cylinders

$$C_i := [0; i]$$

with $i \in \mathcal{A}_{\text{sing}}$ and their union

$$\mathbb{C} := \bigcup_{i \in \mathcal{A}_{\text{sing}}} C_i,$$

the set of all (bi-infinite) words with a singular symbol at the zeroth position. Then

$$q_0 := \mu(C) = \sum_{i \in \mathcal{A}_{\text{sing}}} q_i.$$

Moreover, let

$$C_{i,j} := \bigcup_{n=1}^{\infty} \bigcup_{\underline{\omega} \in \mathcal{B}_n(i,j)} [0; \underline{\omega}].$$

By the mixing of the Markov shift, these sets give rise to the following partitions modulo a zero measure set (mod 0):

$$C_i = \bigcup_{j \in \mathcal{A}_{\text{sing}}} C_{i,j} = \bigcup_{j \in \mathcal{A}_{\text{sing}}} \bigcup_{n=1}^{\infty} \bigcup_{\underline{\omega} \in \mathcal{B}_n(i,j)} [0; \underline{\omega}]$$

and

$$C = \bigcup_{j \in \mathcal{A}_{\text{sing}}} \bigcup_{i \in \mathcal{A}_{\text{sing}}} C_{i,j} = \bigcup_{i \in \mathcal{A}_{\text{sing}}} \bigcup_{j \in \mathcal{A}_{\text{sing}}} \bigcup_{n=1}^{\infty} \bigcup_{\underline{\omega} \in \mathcal{B}_n(i,j)} [0; \underline{\omega}].$$

Let $g: C \rightarrow C$ be the first return map to the cylinder C , given by $g(\omega) = \sigma^{\tau(\omega)}(\omega)$, where $\tau(\omega) = \min\{k \geq 1: \sigma^k(\omega) \in C\}$. The map g preserves the induced measure $\bar{\mu}_C = \mu(C)^{-1} \mu|_C = \frac{1}{q_0} \mu|_C$.

We define the induced cocycle $F_C: C \times \mathbb{R}^2 \rightarrow C \times \mathbb{R}^2$, given by $F_C(\omega, v) := (g(\omega), \mathcal{C}(\omega)v)$, where $\mathcal{C}(\omega) := A^{\tau(\omega)}(\omega)$.

By [24, Proposition 4.18 and Exercise 4.8] (it is easy to verify that these statements also hold for $\text{Mat}_2(\mathbb{R})$ -valued fiber maps) its Lyapunov exponent is related to that of the original cocycle (A, P, q) via the expression $L_1(F_C) = \frac{1}{q_0} L_1(A)$; thus it is enough to compute $L_1(F_C)$.

The induced cocycle F_C leaves invariant the 1-dimensional sub-bundle $X \ni \omega \mapsto R(\omega) := \text{Range}(A_{\omega_0})$. Then using Oseledets' theorem, its first Lyapunov exponent is the rate of exponential growth of the fiber iterates of F_C along this sub-bundle. Thus, for μ -a.e. $\omega \in C$ and for a unit vector $r_0 \in R(\omega) = \text{Range}(A_{\omega_0})$,

$$L_1(F_C) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|\mathcal{C}^n(\omega)r_0\| = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left\| \prod_{l=1}^n \mathcal{C}(g^l \omega)r_0 \right\|. \quad (2.9)$$

Given $\omega \in C$, let $0 = k_1 < k_2 < \dots$ denote all future entries to the singular part of the alphabet, that is, $k \in \mathbb{N}$ is such that $\omega_k \in \mathcal{A}_{\text{sing}}$ if and only if $k = k_l$ for some $l \in \mathbb{N}$. Then clearly $g^l(\omega) = \sigma^{k_l}(\omega)$ (whose zeroth entry is ω_{k_l}) and $\tau(g^l \omega) = k_{l+1} - k_l$ for all $l \in \mathbb{N}$. Moreover,

$$B_l := \mathcal{C}(g^l \omega) = A^{\tau(g^l \omega)}(g^l \omega) = A^{k_{l+1} - k_l}(\sigma^{k_l} \omega) = A_{\omega_{k_{l+1}}} \cdots A_{\omega_{k_l+1}},$$

which is a rank one matrix whose range $r_l := \text{Range}(B_l) = r_{\omega_{k_{l+1}}}$.

By Lemma 2.1,

$$\begin{aligned} \|\mathcal{C}^n(\omega)r_0\| &= \left\| \prod_{l=n}^1 \mathcal{C}(g^l \omega)r_0 \right\| = \left\| \prod_{l=n}^1 B_l r_0 \right\| = \prod_{l=1}^n \|B_l r_{l-1}\| \\ &= \prod_{l=1}^n \|\mathcal{C}(g^l \omega)r_{\omega_{k_l}}\|. \end{aligned} \quad (2.10)$$

Consider the observable $\varphi: C \rightarrow \mathbb{R}$,

$$\varphi(\omega) := \log \|\mathcal{C}(\omega)r_{\omega_0}\| = \log \|A^{\tau(\omega)}(\omega)r_{\omega_0}\|.$$

By (2.9), (2.10) and Birkhoff's ergodic theorem,

$$\begin{aligned} L_1(F_C) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n \log \|\mathcal{C}(g^l \omega)r_{\omega_{k_l}}\| = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n \varphi(g^l \omega) \\ &= \int_C \varphi(\omega) d\bar{\mu}|_C(\omega) = \frac{1}{q_0} \int_C \log \|\mathcal{C}(\omega)r_{\omega_0}\| d\mu(\omega). \end{aligned}$$

Note that on each given cylinder $[0; \underline{\omega}]$ with $\underline{\omega} \in \mathcal{B}_n(i, j)$ for some $i, j \in \mathcal{A}_{\text{sing}}$ and $n \geq 1$, the first return map τ is constant and equal to n , while the observable φ is

equal to $\log \|A^n(\omega)r_{\omega_0}\| = \log \|A^n(\underline{\omega})r_i\|$, thus it is constant. Moreover, the Markov measure of this cylinder is $\mu[0; \underline{\omega}] = q_i p(\underline{\omega})$. Since these cylinders partition (mod 0) the set C , we conclude that

$$\int_C \varphi(\omega) d\mu(\omega) = \sum_{i,j \in \mathcal{A}_{\text{sing}}} \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(i,j)} q_i p(\underline{\omega}) \log \|A^n(\underline{\omega})r_i\|,$$

which completes the proof of the lemma. $\blacksquare$

Consider the observable $\Psi: \mathcal{A} \times \mathbb{P}^1 \rightarrow [-\infty, \infty)$,

$$\Psi(j, \hat{v}) = \sum_{i \in \mathcal{A}} p_{ij} \log \frac{\|A_i v\|}{\|v\|}.$$

We are ready to phrase and prove a Furstenberg-type formula relating the first Lyapunov exponent of a random cocycle $(\underline{A}, P, q)$ to the integral of the observable Ψ with respect to the stationary measure η of its associated Markov operator.

Theorem 2.2 (Furstenberg's Formula). *If the random cocycle $(\underline{A}, P, q)$ has both singular and invertible components, then*

$$L_1(\underline{A}) = \int \Psi d\eta.$$

Proof. Using the explicit formula of the stationary measure η derived in Proposition 2.1,

$$\int \Psi d\eta = \sum_{l \in \mathcal{A}} \sum_{i \in \mathcal{A}} p_{il} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s,l)} p(\underline{\omega}) \log \frac{\|A_i A^n(\underline{\omega})r_s\|}{\|A^n(\underline{\omega})r_s\|}.$$

For $\underline{\omega} \in \mathcal{B}_n(s, l)$ with $l \in \mathcal{A}_{\text{sing}}$, $\hat{A}^n(\underline{\omega})\hat{r}_s = \hat{A}_l \cdots \hat{A}_{\omega_1} \hat{r}_s = \hat{r}_l$. The expression above can then be split into

$$\begin{aligned} & \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{i \in \mathcal{A}} p_{il} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s,l)} p(\underline{\omega}) \log \|A_i r_l\| \\ & + \sum_{l \in \mathcal{A}_{\text{inv}}} \sum_{i \in \mathcal{A}} p_{il} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s,l)} p(\underline{\omega}) \log \frac{\|A_i A^n(\underline{\omega})r_s\|}{\|A^n(\underline{\omega})r_s\|}. \end{aligned}$$

By (2.1), we have

$$\sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s,l)} p(\underline{\omega}) = q_l$$

which, when used in the first term above, turns the last expressions into

$$\begin{aligned}
& \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{i \in \mathcal{A}} p_{il} q_l \log \|A_i r_l\| \\
& + \sum_{l \in \mathcal{A}_{\text{inv}}} \sum_{i \in \mathcal{A}} p_{il} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, l)} p(\underline{\omega}) \log \|A_i A^n(\underline{\omega}) r_s\| \\
& - \sum_{l \in \mathcal{A}_{\text{inv}}} \sum_{i \in \mathcal{A}} p_{il} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, l)} p(\underline{\omega}) \log \|A^n(\underline{\omega}) r_s\| \\
& = \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{i \in \mathcal{A}} p_{il} q_l \log \|A_i r_l\| \\
& + \sum_{i \in \mathcal{A}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=2}^{\infty} \sum_{\underline{\omega}' \in \mathcal{B}_n(s, i)} p(\underline{\omega}') \log \|A^n(\underline{\omega}') r_s\| \\
& - \sum_{l \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, l)} p(\underline{\omega}) \log \|A^n(\underline{\omega}) r_s\|.
\end{aligned}$$

The third line above turned into the sixth because $\sum_{i \in \mathcal{A}} p_{il} = 1$. The second line turned into the fifth by adding the letter i to the end of the word $\underline{\omega} \in \mathcal{B}_n(s, l)$ with $l \in \mathcal{A}_{\text{inv}}$, so that $\underline{\omega}' := (\underline{\omega}, i) \in \mathcal{B}_{n+1}(s, i)$; note that $p(\underline{\omega}') = p_{il} p(\underline{\omega})$.

The last two lines above cancel each other out partially (note the slight difference in the ranges of the index n and the fact that the final letter in the words appearing in the penultimate term ranges over the whole alphabet, while in the ultimate term it ranges only over the invertible part).

The whole expression is then equal to

$$\begin{aligned}
& \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{i \in \mathcal{A}} p_{il} q_l \log \|A_i r_l\| \\
& + \sum_{i \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=2}^{\infty} \sum_{\underline{\omega}' \in \mathcal{B}_n(s, i)} p(\underline{\omega}') \log \|A^n(\underline{\omega}') r_s\| \\
& - \sum_{l \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{\underline{\omega} \in \mathcal{B}_1(s, l)} p(\underline{\omega}) \log \|A^1(\underline{\omega}) r_s\| \\
& = \sum_{l \in \mathcal{A}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s p_{ls} \log \|A_l r_s\| \\
& + \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=2}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, l)} p(\underline{\omega}) \log \|A^n(\underline{\omega}) r_s\| \\
& - \sum_{l \in \mathcal{A}_{\text{inv}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s p_{ls} \log \|A_l r_s\|
\end{aligned}$$

$$\begin{aligned}
 &= \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s p_{ls} \log \|A_l r_s\| \\
 &\quad + \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=2}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, l)} p(\underline{\omega}) \log \|A^n(\underline{\omega}) r_s\| \\
 &= \sum_{l \in \mathcal{A}_{\text{sing}}} \sum_{s \in \mathcal{A}_{\text{sing}}} q_s \sum_{n=1}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s, l)} p(\underline{\omega}) \log \|A^n(\underline{\omega}) r_s\|.
 \end{aligned}$$

The first line turned into the fourth after re-denoting the index i by l and the index l by s ; the final equality follows since (s, l) is the only word $\underline{\omega} \in \mathcal{B}_1(s, l)$ and in this case $p(\underline{\omega}) = p_{ls}$ and $A^1(\underline{\omega}) = A_l$. The expression in the last line is exactly the formula obtained in Lemma 2.2 for $L_1(\underline{A})$ so the proof of Furstenberg's formula is complete. ■

Remark 2.2. We note that the classical Furstenberg's formula (i.e. [13, Theorem 8.5]) for random cocycles in $\text{GL}_m(\mathbb{R})$, or other more recent versions thereof, are obviously not applicable to this singular setting. Moreover, the probabilistic approach used by Furstenberg and Kifer to establish such results (see e.g. [14, Theorem 1.4]) is not immediately applicable either, since the observable Ψ is not continuous, not even bounded on $\mathcal{A} \times \mathbb{P}^1$. That is why we had to employ an ad-hoc argument which uses the explicit formula of the stationary measure.

Remark 2.3. The uniform ergodicity established in Theorem 2.1 is the strongest form of mixing of a transition operator, see [21, Section 16.2] for a discussion of this and other types of mixing. It is equivalent to the quasi-compactness of the operator $\mathcal{Q}$ on the entire space $L^\infty(\mathcal{A} \times \mathbb{P}^1)$.

We note that in the classical case of $\text{GL}_m(\mathbb{R})$ -valued cocycles satisfying an irreducibility assumption, the corresponding Markov operator was shown to be quasi-compact on the smaller subspace of observables that are Hölder continuous in the projective variable, endowed with the stronger Hölder norm, see for instance [11, Proposition 5.22]. This was then used to establish statistical properties (and the Hölder continuity) of the corresponding Lyapunov exponent.

In the setting of this paper, however, the observable Ψ appearing in Furstenberg's formula does *not* belong to the space $L^\infty(\mathcal{A} \times \mathbb{P}^1)$, since it is not bounded. Thus, the mixing property (i.e., the convergence $\mathcal{Q}^n \varphi \rightarrow \int \varphi d\eta$) does not apply to Ψ , which is what makes the establishing of statistical properties in the next section an extremely technical matter.

Remark 2.4. The stochastic dynamical system over the alphabet $\mathcal{A}$, governed by the stochastic matrix P , induces a singular (partial) random walk on $\mathcal{A} \times \mathbb{P}^1$, where some

paths terminate upon reaching the kernel of the subsequent transition matrix. In fact, termination occurs with probability 1 whenever the cocycle $\underline{A}$ contains null words.

Nevertheless, the transition operator $\mathcal{Q}: L^\infty(\mathcal{A} \times \mathbb{P}^1) \rightarrow L^\infty(\mathcal{A} \times \mathbb{P}^1)$, introduced at the beginning of this section to reflect the projective action of the Markov kernel, is desingularised. Specifically, if $i \in \mathcal{A}_{\text{sing}}$, the term $\varphi(i, \hat{A}_i \hat{v})$, which would otherwise exhibit a discontinuity at the kernel of A_i , is replaced in the definition of the operator by the constant function $\varphi(i, \hat{r}_i)$.

This desingularization explains why the Markov operator $\mathcal{Q}$ always admits stationary measures (in the sense of Definition 2.1) and it is uniformly ergodic (Theorem 2.1). It also clarifies why Furstenberg's formula in Theorem 2.2 remains valid for cocycles with both singular and invertible components.

3. Statistical Properties

For simplicity of exposition, all statements in this section will be proven in the Bernoulli setting, but it will be clear that the arguments easily extend to the mixing Markov shift setting. The main difference between the two models, when it comes to these results, is the form of the transition operator and the explicit formula of its stationary measure, which is more cumbersome in the Markov case (as seen in Proposition 2.1).

Let us summarise the relevant concepts in the Bernoulli setting. Given the alphabet $\mathcal{A} = \{1, \dots, k\}$ and a partition $\mathcal{A} = \mathcal{A}_{\text{sing}} \sqcup \mathcal{A}_{\text{inv}}$ into two non-empty sets, consider the set $\mathcal{M}^* \subset \mathcal{M}$ of all k -tuples $\underline{A} = (A_1, \dots, A_k) \in \text{Mat}_2(\mathbb{R})^k$ such that $\text{rank } A_i = 1$ if $i \in \mathcal{A}_{\text{sing}}$, $\det A_j > 0$ if $j \in \mathcal{A}_{\text{inv}}$ and, additionally, $\hat{r}_i \neq \hat{k}_j$ for all $i, j \in \mathcal{A}_{\text{sing}}$.

Recall that we identify such a k -tuple $\underline{A}$ with the locally constant map $A: X \rightarrow \text{Mat}_2(\mathbb{R})$, $A(\omega) = A_{\omega_0}$, where $\omega = \{\omega_n\}_{n \in \mathbb{Z}} \in X = \mathcal{A}^{\mathbb{Z}}$.

Given a probability vector $p = (p_1, \dots, p_k)$ with $p_i > 0$ for all i , the pair $(\underline{A}, p)$ determines the random linear cocycle

$$F: X \times \mathbb{R}^2 \rightarrow X \times \mathbb{R}^2, \quad F(\omega, v) = (\sigma\omega, A(\omega)v) = (\sigma\omega, A_{\omega_0}v),$$

where X is endowed with the product probability $\mu = p^{\mathbb{Z}}$. As before, we also identify the cocycle F with the tuple $\underline{A} = (A_1, \dots, A_k)$.

The Markov operator in this setting is $\mathcal{Q}: L^\infty(\mathbb{P}^1) \rightarrow L^\infty(\mathbb{P}^1)$,

$$\mathcal{Q}\varphi(\hat{v}) := \sum_{i \in \mathcal{A}} \varphi(\hat{A}_i \hat{v}) p_i.$$

Given $n \geq 0$ and $s \in \mathcal{A}_{\text{sing}}$, let $\mathcal{B}_n(s)$ denote the set of $(n+1)$ -tuples $(s\underline{\omega})$ where $\underline{\omega} = (\omega_1, \dots, \omega_n) \in \mathcal{A}_{\text{inv}}^n$. For such a word we write, as before,

$$A^n(s\underline{\omega}) := A_{\omega_n} \dots A_{\omega_2} A_{\omega_1}$$

(with the convention that $A^0(s\omega)$ is the identity matrix) and denote the quantity $p(\underline{\omega}) := p_{\omega_1} \cdots p_{\omega_n}$.

Let $q := \sum_{i \in \mathcal{A}_{\text{sing}}} p_i < 1$ and note that

$$\sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{n=0}^{\infty} \sum_{\underline{\omega} \in \mathcal{A}_{\text{inv}}^n} p(\underline{\omega}) = \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{n=0}^{\infty} (1-q)^n = q \frac{1}{q} = 1. \quad (3.1)$$

With these notations, we have the following explicit formula for the unique $\mathcal{Q}$ -stationary measure $\eta \in \text{Prob}(\mathbb{P}^1)$ (its derivation is quite straightforward compared to the Markov case, so we omit it):

$$\eta = \sum_{s \in \mathcal{A}_{\text{sing}}} \sum_{n=0}^{\infty} \sum_{\underline{\omega} \in \mathcal{B}_n(s)} p(\underline{\omega}) \delta_{\hat{A}^n(s\omega)\hat{r}_s}.$$

We write $\mathcal{Q} = \mathcal{Q}_{\text{inv}} + \mathcal{Q}_{\text{sing}}$, with $\mathcal{Q}_{\text{inv}}\varphi(\hat{v}) := \sum_{i \in \mathcal{A}_{\text{inv}}} \varphi(\hat{A}_i \hat{v}) p_i$ and $\mathcal{Q}_{\text{sing}}\varphi(\hat{v}) := \sum_{i \in \mathcal{A}_{\text{sing}}} \varphi(\hat{r}_i) p_i$. Note that $\mathcal{Q}_{\text{sing}}$ is a projection to the constant functions. This then easily implies by induction that for all $n \in \mathbb{N}$,

$$\mathcal{Q}^n = \mathcal{Q}_{\text{inv}}^n + \mathcal{Q}_{\text{sing}} \circ \left(\sum_{i=0}^{n-1} \mathcal{Q}_{\text{inv}}^i \right).$$

In particular, $\mathcal{Q}^n - \mathcal{Q}_{\text{inv}}^n$ takes any observable φ to a constant function, which then implies the uniform ergodicity of the Markov operator.

That is, for all $n \in \mathbb{N}$ and $\varphi \in L^\infty(\mathbb{P}^1)$,

$$\left\| \mathcal{Q}^n \varphi - \int \varphi d\eta \right\|_\infty \leq C e^{-an} \|\varphi\|_\infty. \quad (3.2)$$

Finally, defining the observable $\psi: \mathbb{P}^1 \rightarrow [-\infty, \infty)$,

$$\psi(\hat{v}) := \sum_{i \in \mathcal{A}} p_i \log \frac{\|A_i v\|}{\|v\|},$$

the Lyapunov exponent of the random cocycle $\underline{A}$ is given by

$$L_1(\underline{A}) = \int \psi(\hat{v}) d\eta(\hat{v}), \quad (3.3)$$

that is, a Furstenberg-type formula holds.

The goal of this section is to establish an LDT estimate and a CLT for Lebesgue a.e. cocycle $\underline{A} \in \mathcal{M}$. This will be obtained as a consequence of a stronger result essentially saying that if $t \mapsto \underline{A}_t \in \mathcal{M}^*$ is a one-parameter family of cocycles satisfying a positive winding condition, then an LDT estimate and a CLT hold for Lebesgue a.e.

parameter t . By general results, see Cai, Duarte, and Klein [7], such limit laws hold as soon as an associated Markov operator satisfies an appropriate mixing condition. We will use a parameter elimination argument to show that for Lebesgue a.e. parameter t , the Markov operator corresponding to $\underline{A}_t$ satisfies the desired condition. In order to apply the general results in [7], we need to associate to a given cocycle $\underline{A} \in \mathcal{M}^*$ a Markov operator on a slightly larger space, namely $\bar{\mathcal{Q}}: L^\infty(\mathcal{A} \times \mathbb{P}^1) \rightarrow L^\infty(\mathcal{A} \times \mathbb{P}^1)$ defined by

$$\bar{\mathcal{Q}}\varphi(j, \hat{v}) = \sum_{i=1}^k \varphi(i, \hat{A}_j \hat{v}) p_i.$$

Consider the projection $\pi: L^\infty(\mathcal{A} \times \mathbb{P}^1) \rightarrow L^\infty(\mathcal{A})$ given by

$$\pi\varphi(\hat{v}) = \sum_{i=1}^k p_i \varphi(i, \hat{v}) = \int_{\mathcal{A}} \varphi(i, \hat{v}) dp(i).$$

Then the Markov operators $\bar{\mathcal{Q}}$ and $\mathcal{Q}$ are related as follows.

Lemma 3.1. *The following diagram is commutative:*

$$\begin{array}{ccc} L^\infty(\mathcal{A} \times \mathbb{P}^1) & \xrightarrow{\bar{\mathcal{Q}}} & L^\infty(\mathcal{A} \times \mathbb{P}^1) \\ \pi \downarrow & & \downarrow \pi \\ L^\infty(\mathbb{P}^1) & \xrightarrow{\mathcal{Q}} & L^\infty(\mathbb{P}^1) \end{array}$$

That is, $\pi \circ \bar{\mathcal{Q}} = \mathcal{Q} \circ \pi$ and inductively, $\pi \circ \bar{\mathcal{Q}}^n = \mathcal{Q}^n \circ \pi$ for all $n \in \mathbb{N}$.

Proof. Let $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$. Then,

$$\pi \circ \bar{\mathcal{Q}}(\varphi)(\hat{v}) = \sum_{j=1}^k p_j \bar{\mathcal{Q}}\varphi(j, \hat{v}) = \sum_{j=1}^k \sum_{i=1}^k p_j p_i \varphi(i, \hat{A}_j \hat{v}),$$

while

$$\mathcal{Q} \circ \pi(\varphi)(\hat{v}) = \sum_{j=1}^k p_j \pi\varphi(\hat{A}_j, \hat{v}) = \sum_{j=1}^k \sum_{i=1}^k p_j p_i \varphi(i, \hat{A}_j \hat{v})$$

which establishes the claim. ■

Corollary 3.1. *The measure $p \times \eta$ is $\bar{\mathcal{Q}}$ -stationary and $\bar{\mathcal{Q}}$ is uniformly ergodic. More precisely, there exist constants c and C such that*

$$\left\| \bar{\mathcal{Q}}^n \varphi - \int \varphi d\mu d\eta \right\|_\infty \leq C e^{-cn} \|\varphi\|_\infty \quad \text{for all } n \in \mathbb{N}, \varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1).$$

In particular, $p \times \eta$ is the unique $\bar{\mathcal{Q}}$ -stationary measure.

Proof. For any observable $\varphi \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$, we have

$$\begin{aligned} \iint \bar{\mathcal{Q}}\varphi(i, \hat{v}) dp(i) d\eta(\hat{v}) &= \int \pi(\bar{\mathcal{Q}}\varphi)(\hat{v}) d\eta(\hat{v}) = \int \mathcal{Q}(\pi\varphi)(\hat{v}) d\eta(\hat{v}) \\ &= \int \pi\varphi(\hat{v}) d\eta(\hat{v}) = \iint \varphi(i, \hat{v}) dp(i) d\eta(\hat{v}), \end{aligned}$$

which shows that $p \times \eta$ is $\bar{\mathcal{Q}}$ -stationary. Moreover,

$$\begin{aligned} \bar{\mathcal{Q}}^{n+1}\varphi(j, \hat{v}) &= \bar{\mathcal{Q}}\bar{\mathcal{Q}}^n\varphi(j, \hat{v}) = \sum_{i=1}^k p_i \bar{\mathcal{Q}}^n\varphi(i, \hat{A}_j \hat{v}) \\ &= \pi(\bar{\mathcal{Q}}^n\varphi)(\hat{A}_j \hat{v}) = \mathcal{Q}^n(\pi\varphi)(\hat{A}_j \hat{v}). \end{aligned}$$

Since by (3.2), the operator $\mathcal{Q}$ is strongly mixing, the last term converges uniformly and exponentially fast to $\int \pi\varphi d\eta = \iint \varphi dp d\eta$, thus showing the strong mixing of $\bar{\mathcal{Q}}$. ■

Let $\Lambda \subset \mathbb{R}$ be a compact interval (normalised to have length 1) and let $A: \Lambda \rightarrow \mathcal{M}^*$ be a smooth map. We think of this map as a one-parameter family of random cocycles so we use the subscript notation $t \mapsto A_t$. For its components, we write $A_t(i)$ instead of $(A_t)_i$, while the fiber iterates are denoted by

$$A_t^n(\omega) = A_t(\omega_n) \cdots A_t(\omega_1).$$

For every parameter $t \in \Lambda$, denote by $\mathcal{Q}_t$ and $\bar{\mathcal{Q}}_t$ the Markov operators corresponding to the cocycle A_t (defined as above). Moreover, let $\eta_t \in \text{Prob}(\mathbb{P}^1)$ be the unique $\mathcal{Q}_t$ -stationary measure.

All the results proven above for a given cocycle $\underline{A} \in \mathcal{M}$, namely the explicit formula of the stationary measure, Furstenberg's formula and the strong mixing of the Markov operator, apply to A_t for all $t \in \Lambda$; since the probability vector $p = (p_1, \dots, p_k)$ and the singular/invertible symbols do not change, all the mixing parameters are uniform in t . Finally, recall that the stationary measure is given by

$$\eta_t = \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{n=0}^{\infty} \sum_{\omega \in \mathcal{A}_{\text{inv}}^n} p(\underline{\omega}) \delta_{\hat{A}_t^n(s\underline{\omega})\hat{r}_s}.$$

3.1. Assumptions

We assume that the family of cocycles $t \mapsto \underline{A}_t$ is positively winding (see [10, Definition 4.1]) and that its singular components are constant. More precisely, we impose the following conditions on the smooth map $\Lambda \ni t \mapsto \underline{A}_t \in \mathcal{M}^*$:

(A1) for all $i \in \mathcal{A}_{\text{sing}}$, $A_t(i) \equiv A_i$;

(A2) there is $c_0 > 0$ such that for all $t \in \Lambda$, $j \in \mathcal{A}_{\text{inv}}$, and $\hat{v} \in \mathbb{P}^1$, we have

$$\frac{A_t(j)v \wedge \frac{d}{dt}A_t(j)v}{\|A_t(j)v\|^2} \geq c_0.$$

Remark 3.1. By [5, Proposition 3.1], the quantity in (A2) above, which we refer to as the rotation speed of the map $t \mapsto A_t$, can be characterised by

$$\frac{d}{dt}\hat{A}_t(j)\hat{v} = \frac{d}{dt} \frac{A_t(j)v}{\|A_t(j)v\|} = \frac{A_t(j)v \wedge \frac{d}{dt}A_t(j)v}{\|A_t(j)v\|^2}.$$

Remark 3.2. A more general version of the winding condition requires that the inequality in assumption (A2) above holds for some iterate $A_t^{n_0}$ of the cocycle. For simplicity, we assume that $n_0 = 1$. Moreover, if the assumption (A2) holds, then there exists $c_1 > 0$ that only depends on the family A_t , such that for all $n \in \mathbb{N}$, $\omega \in \mathcal{A}_{\text{inv}}^n$, and $\hat{v} \in \mathbb{P}^1$, we have

$$\frac{d}{dt} \frac{A_t^n(\omega)v}{\|A_t^n(\omega)v\|} \geq c_1.$$

In other words, if A_t is a family of positively winding cocycles, then so is A_t^n for every $n \in \mathbb{N}$, with rotation speed uniformly (in n) bounded from below. For a proof of this statement, see [5, Section 3.1, specifically Propositions 3.3 and 3.4.].

Example 3.1. Given any tuple $\underline{A} = (A_1, \dots, A_k) \in \mathcal{M}^*$, the one-parameter family $[-\pi, \pi] \ni t \mapsto A_t \in \text{Mat}_2(\mathbb{R})^k$, $A_t(i) = A_i$ if $i \in \mathcal{A}_{\text{sing}}$ and $A_t(j) = A_j R_t = A_j \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$ if $j \in \mathcal{A}_{\text{inv}}$ satisfies the assumptions (A1) and (A2) above, passes through $\underline{A}$ and $\underline{A}_t \in \mathcal{M}^*$ for all $t \in \Lambda$.

3.2. Some technical estimates

In order to simplify the exposition in the estimates to follow, we will write $a \lesssim b$ if there is some absolute or at least context universal constant $C < \infty$ such that $a \leq Cb$. Moreover, for an arc $\hat{I} \subset \mathbb{P}^1$, $m(\hat{I})$ denotes its length, while $\text{Leb}(E)$ stands for the Lebesgue measure of a subset E of the real line.

Lemma 3.2. *Given any $\varepsilon > 0$, the set*

$$\hat{I}_\varepsilon := \left\{ \hat{v} \in \mathbb{P}^1 : \left\| A_i \frac{v}{\|v\|} \right\| < \varepsilon \text{ for some } i \in \mathcal{A}_{\text{sing}} \right\}$$

is a finite union of arcs with $m(\hat{I}_\varepsilon) \lesssim \varepsilon$.

Proof. Let $i \in \mathcal{A}_{\text{sing}}$ and consider the normalised directions r_i and k_i of the range and kernel of A_i , respectively, such that $\|k_i\| = \|r_i\| = 1$. Given any vector $v \in \mathbb{R}^2$, we write $v = \alpha_1 r_i + \alpha_2 k_i$. Then

$$\left\| A_i \frac{v}{\|v\|} \right\| = \frac{|\alpha_1|}{\|v\|}.$$

Thus, $\|A_i \frac{v}{\|v\|}\| < \varepsilon$ in a small interval around each kernel k_i . Moreover, $\hat{I}_\varepsilon$ is a finite union of arcs of length $\asymp 2\varepsilon$ each. Hence, $m(\hat{I}_\varepsilon) \lesssim \varepsilon$. ■

Lemma 3.3. *Given any arc $\hat{I} \subset \mathbb{P}^1$, $n \in \mathbb{N}$, $\omega \in \mathcal{A}_{\text{inv}}^n$ and $\hat{v} \in \mathbb{P}^1$ we have*

$$\text{Leb}\{t \in \Lambda : \hat{A}_t^n(\omega)\hat{v} \in \hat{I}\} \leq \frac{m(\hat{I})}{c_1}.$$

Proof. Since the projective line $\mathbb{P}^1$ is one-dimensional, we may regard the map $\Lambda \ni t \mapsto \hat{A}_t^n(\omega)\hat{v}$ as a real valued map, whose derivative (because of the winding condition) is bounded from below by c_1 . The result then follows using basic calculus (the mean value theorem). ■

Consider the maps $\varphi_t: \mathcal{A} \times \mathbb{P}^1 \rightarrow [-\infty, \infty)$ and $\psi_t: \mathbb{P}^1 \rightarrow [-\infty, \infty)$ given by

$$\varphi_t(i, \hat{v}) := \log \left\| A_t(i) \frac{v}{\|v\|} \right\| \quad \text{and} \quad \psi_t(\hat{v}) = \sum_{i \in \mathcal{A}} p_i \varphi_t(i, \hat{v}) = \int_{\mathcal{A}} \varphi_t(i, \hat{v}) dp(i).$$

Let $c = c(A)$ be a constant that satisfies

- (i) $e^{-c} \leq \|A_t(j)\| \leq e^c$ for all $t \in \Lambda$ and $j \in \mathcal{A}_{\text{inv}}$,
- (ii) $\|A_i\| \leq e^c$ for all $i \in \mathcal{A}_{\text{sing}}$,
- (iii) $\|A_i r_l\| \geq e^{-c}$ for all $i, l \in \mathcal{A}_{\text{sing}}$.

Such a constant exists by the compactness of Λ and the fact that $\underline{A} \in \mathcal{M}^*$, so $\hat{k}_i \neq \hat{r}_l$, thus $A_i r_l \neq 0$ for all $i, l \in \mathcal{A}_{\text{sing}}$. It follows that for all $j \in \mathcal{A}_{\text{inv}}$ and $\hat{v} \in \mathbb{P}^1$, we have $|\varphi_t(j, \hat{v})| \leq c$. Moreover, the upper bound $\varphi_t(j, \hat{v}) \leq c$ holds for every $j \in \mathcal{A}$ and $\hat{v} \in \mathbb{P}^1$.

Lemma 3.4. *There is $C < \infty$ which depends on the constants c, c_1 , so only on $\underline{A}$, such that the following hold.*

- (i) *For all $n \in \mathbb{N}$, $\omega \in \mathcal{A}_{\text{inv}}^n$, $i \in \mathcal{A}_{\text{sing}}$, $\hat{v} \in \mathbb{P}^1$ and $N \geq 0$, we have*

$$\text{Leb}\{t \in \Lambda : \varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) < -N\} \leq C e^{-N}.$$

Moreover,

$$\int_{\Lambda} \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) dt \leq C.$$

(ii) For all $n \in \mathbb{N}$, $\omega \in \mathcal{A}^n$, $i, s \in \mathcal{A}_{\text{sing}}$, and $N \geq 0$, we have

$$\text{Leb}\{t \in \Lambda: \varphi_t(i, \hat{A}_t^n(\omega)\hat{r}_s) < -N\} \leq C e^{-N}.$$

Moreover,

$$\int_{\Lambda} \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{r}_s) dt \leq C.$$

(iii) Furthermore, for all $i \in \mathcal{A}$, $\omega \in X$, and $n \in \mathbb{N}$,

$$\int_{\Lambda} \int_{\mathbb{P}^1} \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) d\eta_t(\hat{v}) dt \leq C.$$

In particular,

$$\int_{\Lambda} \int_{\mathbb{P}^1} \psi_t^2(\hat{A}_t^n(\omega)\hat{v}) d\eta_t(\hat{v}) dt \leq C.$$

Proof. (i) If $i \in \mathcal{A}_{\text{sing}}$, then $\varphi_t(i, \hat{w}) < -N$ holds if and only if $\|A_i \frac{w}{\|w\|}\| < e^{-N}$. By Lemma 3.2 the set of such points $\hat{w}$ is a finite union of arcs with total measure of order $\varepsilon := e^{-N}$, and by Lemma 3.3 the measure of the set of parameters t for which $\hat{A}_t^n(\omega)\hat{v}$ belongs to these arcs is $\lesssim \varepsilon$. This proves the first statement in item (i).

For the second statement, note that

$$\begin{aligned} \int_{\Lambda} \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) dt &= \int_0^{\infty} \text{Leb}\{t \in \Lambda: \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) \geq x\} dx \\ &= \int_0^{2c^2} \text{Leb}\{t \in \Lambda: \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) \geq x\} dx \\ &\quad + \int_{2c^2}^{\infty} \text{Leb}\{t \in \Lambda: \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) \geq x\} dx \\ &\leq 2c^2 + \int_{2c^2}^{\infty} \text{Leb}\{t \in \Lambda: \varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) \leq -\sqrt{x}\} dx \\ &\lesssim c^2 + \int_{2c^2}^{\infty} e^{-\sqrt{x}} dx \lesssim 1. \end{aligned}$$

The bound in the penultimate line above holds because Λ has length one and, moreover, the upper bound $\varphi_t(i, \hat{w}) \leq c$ is valid for all $\hat{w} \in \mathbb{P}^1$, so that

$$\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) \leq c < \sqrt{x} \quad \text{when } x \geq 2c^2.$$

(ii) Note that if $\omega \in \mathcal{A}_{\text{inv}}^n$, then the statement follows from (i). The difference here is that we consider any word ω , invertible or not, but begin with a singular vector, namely r_s .

Let $\omega = (\omega_0, \dots, \omega_{n-1}) \notin \mathcal{A}_{\text{inv}}^n$. We split the argument into two cases: either $\omega_{n-1} \in \mathcal{A}_{\text{sing}}$ or else $\omega_{n-1} \in \mathcal{A}_{\text{inv}}$, but there is $0 \leq j \leq n-2$ such that $\omega_j \in \mathcal{A}_{\text{sing}}$.

If $\omega_{n-1} \in \mathcal{A}_{\text{sing}}$, then

$$\hat{A}_t^n(\omega)\hat{r}_s = \hat{A}_t(\omega_{n-1})\hat{A}_t^{n-1}(\omega)\hat{r}_s = \hat{A}_{\omega_{n-1}}\hat{A}_t^{n-1}(\omega)\hat{r}_s = \hat{r}_{\omega_{n-1}}.$$

Then by the choice of the constant c , for all parameters $t \in \Lambda$,

$$\varphi_t(i, \hat{A}_t^n(\omega)\hat{r}_s) = \log\|\hat{A}_i\hat{r}_{\omega_{n-1}}\| \geq -c,$$

so the set $\{t \in \Lambda : \varphi_t(i, \hat{A}_t^n(\omega)\hat{r}_s) < -N\}$ becomes empty for $N \geq c$ and the statement follows trivially.

If $\omega_{n-1} \in \mathcal{A}_{\text{inv}}$ and $\omega_j \in \mathcal{A}_{\text{sing}}$ for some index $0 \leq j \leq n-2$, let n' be the largest such index and let $\omega' := (\omega_{n'+1}, \dots, \omega_{n-1}) \in \mathcal{A}_{\text{inv}}^{n-n'-1}$. Then

$$\hat{A}_t^n(\omega)\hat{r}_s = \hat{A}_t^{n-n'-1}(\omega')\hat{A}_{\omega_n'}\hat{A}_t^{n'}(\omega)\hat{r}_s = \hat{A}_t^{n-n'-1}(\omega')\hat{r}_{\omega_n'}.$$

Thus, $\varphi_t(i, \hat{A}_t^n(\omega)\hat{r}_s) = \varphi_t(i, \hat{A}_t^{n-n'-1}(\omega')\hat{r}_{\omega_n'})$, with $\omega' \in \mathcal{A}_{\text{inv}}^{n-n'-1}$, hence (i) is applicable and the conclusion follows.

(iii) If $i \in \mathcal{A}_{\text{inv}}$ the statement is obvious since $\varphi_t^2(i, \hat{w}) \leq c^2$ for all $(t, \hat{w}) \in \Lambda \times \mathbb{P}^1$. Let us then fix $i \in \mathcal{A}_{\text{sing}}$. Using the explicit formula of η_t we have

$$\begin{aligned} & \int_{\Lambda} \int_{\mathbb{P}^1} \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{v}) d\eta_t(\hat{v}) dt \\ &= \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=0}^{\infty} \sum_{\underline{\omega} \in \mathcal{A}_{\text{inv}}^j} p(\underline{\omega}) \int_{\Lambda} \varphi_t^2(i, \hat{A}_t^n(\omega)\hat{A}_t^j(s\underline{\omega})\hat{r}_s) dt \\ &= \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=0}^{\infty} \sum_{\underline{\omega} \in \mathcal{A}_{\text{inv}}^j} p(\underline{\omega}) \int_{\Lambda} \varphi_t^2(i, \hat{A}_t^{n+j}(s\underline{\omega}\omega)\hat{r}_s) dt \\ &\leq \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=0}^{\infty} \sum_{\underline{\omega} \in \mathcal{A}_{\text{inv}}^j} p(\underline{\omega}) C = C, \end{aligned}$$

where the inequality in the last line follows from (ii) above and the last equality comes from (3.1). ■

Let $\tilde{\eta} \in \text{Prob}(\Lambda \times \mathbb{P}^1)$ be the probability measure whose disintegration relative to the Lebesgue measure on Λ is $\{\eta_t\}$, that is, the measure characterised by

$$\int \varphi(t, \hat{v}) d\tilde{\eta}(t, \hat{v}) = \int_{\Lambda} \int_{\mathbb{P}^1} \varphi(t, \hat{v}) d\eta_t(\hat{v}) dt \quad \text{for all } \varphi \in L^\infty(\Lambda \times \mathbb{P}^1).$$

Lemma 3.5. *Given $i \in \mathcal{A}$, $n \in \mathbb{N}$, $N > c$ and $\omega \in X$, let*

$$P_{i,n,N}(\omega) := \{(t, \hat{v}) : \varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) < -N\}.$$

Then $\tilde{\eta}(P_{i,n,N}(\omega)) \leq Ce^{-N}$.

Proof. If $i \in \mathcal{A}_{\text{inv}}$, then $\varphi_t(i, \hat{w}) \geq -c > -N$ for all $\hat{w} \in \mathbb{P}^1$, so $P_{i,n,N}(\omega)$ is empty. We then fix $i \in \mathcal{A}_{\text{sing}}$ and note that

$$\begin{aligned} \tilde{\eta}(P_{i,n,N}(\omega)) &= \int_{\Lambda} \int_{\mathbb{P}^1} \mathbb{1}_{P_{i,n,N}(\omega)}(t, \hat{v}) d\eta_t(\hat{v}) dt \\ &= \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=0}^{\infty} \sum_{\underline{\omega} \in \mathcal{A}_{\text{inv}}^j} p(\underline{\omega}) \int_{\Lambda} \mathbb{1}_{P_{i,n,N}(\omega)}(t, \hat{A}_t^j(s\underline{\omega})\hat{r}_s) dt. \end{aligned}$$

Moreover,

$$\begin{aligned} (t, \hat{A}_t^j(s\underline{\omega})\hat{r}_s) \in P_{i,n,N}(\omega) &\iff \varphi_t(i, \hat{A}_t^n(\omega)\hat{A}_t^j(s\underline{\omega})\hat{r}_s) < -N \\ &\iff \varphi_t(i, \hat{A}_t^{n+j}(s\underline{\omega})\hat{r}_s) < -N. \end{aligned}$$

By Lemma 3.4(ii), for all $j \in \mathbb{N}$,

$$\int_{\Lambda} \mathbb{1}_{P_{i,n,N}(\omega)}(t, \hat{A}_t^j(s\underline{\omega})\hat{r}_s) dt = \text{Leb}\{t : \varphi_t(i, \hat{A}_t^{n+j}(s\underline{\omega})\hat{r}_s) < -N\} \lesssim e^{-N}$$

and the conclusion follows from (3.1). ■

Given any $N > c$, consider the truncation

$$\varphi_{t,N} := \max\{\varphi_t, -N\}.$$

Note that $\varphi_{t,N} \in L^\infty(\mathcal{A} \times \mathbb{P}^1)$ (a property that does not hold for φ_t) and $\|\varphi_{t,N}\|_\infty \leq N$. Moreover, $\varphi_t(i, \hat{v}) = \varphi_{t,N}(i, \hat{v})$ if and only if $i \in \mathcal{A}_{\text{inv}}$ or else $i \in \mathcal{A}_{\text{sing}}$ and $\varphi_t(i, \hat{v}) < -N$.

Lemma 3.6. *For all $i \in \mathcal{A}$, $n \in \mathbb{N}$, $N > c$, and $\omega \in X$, we have*

$$\int_{\Lambda} \int_{\mathbb{P}^1} |\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) - \varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})| d\eta_t(\hat{v}) dt \lesssim e^{-N/3}.$$

Proof. The statement is obvious when $i \in \mathcal{A}_{\text{inv}}$, since in this case $\varphi_t(i, \hat{w}) = \varphi_{t,N}(i, \hat{w})$ for all $\hat{w} \in \mathbb{P}^1$.

We fix $i \in \mathcal{A}_{\text{sing}}$ and recall that $\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) \neq \varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})$ if and only if $\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) < -N$, that is, if and only if $(t, \hat{v}) \in P_{i,n,N}(\omega) =: P_N$, which by Lemma 3.5, is a set of $\tilde{\eta}$ -measure $\lesssim e^{-N}$. Then

$$\begin{aligned} & \int_{\Lambda} \int_{\mathbb{P}^1} |\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) - \varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})| d\eta_t(\hat{v}) dt \\ &= \int_{\Lambda \times \mathbb{P}^1} |\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) - \varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})| \mathbb{1}_{P_N}(t, \hat{v}) d\tilde{\eta}(t, \hat{v}) \\ &\leq \|g(t, \hat{v})\|_{L^2(\tilde{\eta})} \tilde{\eta}(P_N)^{1/2} \end{aligned}$$

where $g(t, \hat{v}) := \varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) - \varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})$ and we used Cauchy–Schwarz in the last inequality above.

Moreover, by Lemma 3.4 (iii),

$$\|g(t, \hat{v})\|_{L^2(\tilde{\eta})} \leq \|\varphi_t(i, \hat{A}_t^n(\omega)\hat{v})\|_{L^2(\tilde{\eta})} + \|\varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})\|_{L^2(\tilde{\eta})} \leq \sqrt{C} + N.$$

Thus,

$$\int_{\Lambda} \int_{\mathbb{P}^1} |\varphi_t(i, \hat{A}_t^n(\omega)\hat{v}) - \varphi_{t,N}(i, \hat{A}_t^n(\omega)\hat{v})| d\eta_t(\hat{v}) dt \lesssim (\sqrt{C} + N)e^{-N/2} \lesssim e^{-N/3},$$

which completes the proof. $\blacksquare$

3.3. Large deviations estimates

In order to prove large deviations estimates for cocycles $\underline{A} \in \mathcal{M}^*$, we use the following abstract recent result of Cai, Duarte, and Klein [7, Theorem 1.1].

Let $\{Z_n\}_{n \geq 0}$ be a Markov chain with values in a compact space M , and transition kernel $M \ni x \mapsto K_x \in \text{Prob}(M)$. Let $Q: L^\infty(M) \rightarrow L^\infty(M)$ be the Markov operator corresponding to this kernel, namely $Q\varphi(x) = \int_M \varphi(y) dK_x(y)$. If the operator Q is strongly mixing (in an appropriate sense) on some subspace of observables $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$, relative to a stationary measure η , then every observable in this subspace $\mathcal{E}$ satisfies (exponential) large deviations. More precisely, for all $n \geq 1$, consider the n -th stochastic sum

$$S_n\varphi := \varphi \circ Z_{n-1} + \cdots + \varphi \circ Z_0.$$

Then for all $\varepsilon > 0$,

$$\mathbb{P} \left\{ \left| \frac{1}{n} S_n\varphi - \int_M \varphi d\eta \right| > \varepsilon \right\} \leq 8e^{-c(\varepsilon)n}, \quad (3.4)$$

where $c(\varepsilon)$ is an explicit rate function. This rate function is essentially of order $c(\varepsilon) \sim \varepsilon^2 \|\varphi\|_{\mathcal{G}}^{-2}$. See [7, Definition 1.1, Theorem 1.1, and Remark 1.1] for more details.

In this work, the Markov operator $\bar{\mathcal{Q}}$ is uniformly ergodic (hence strongly mixing) on the entire space of observables $L^\infty(\mathcal{A} \times \mathbb{P}^1)$ and relative to the L^∞ -norm. Moreover, for the remaining part of the argument, it is crucially important to note the precise dependence of the large deviations estimate on the norm (in our case the L^∞ -norm) of the observable.

We are ready to state and prove a stronger, parametric version of the large-deviations-type (LDT) estimate for cocycles in $\mathcal{M}^*$, which will also imply the result stated in the introduction.

Theorem 3.2. *Let $A: \Lambda \rightarrow \mathcal{M}^*$ be a smooth family of cocycles satisfying Assumptions 3.1. Then for every $\varepsilon > 0$ and $\tilde{\eta}$ -a.e. $(t, \hat{v})$, there exist $c_0(\varepsilon) > 0$ and $n_0(\varepsilon, t, \hat{v}) \in \mathbb{N}$ such that for every $n \geq n_0(\varepsilon, t, \hat{v})$,*

$$\mu \left\{ \omega \in X: \left| \frac{1}{n} \log \|A_t^n(\omega)v\| - L_1(A_t) \right| > \varepsilon \right\} < e^{-c_0(\varepsilon)n^{1/3}}.$$

Moreover, for Lebesgue a.e. $t \in \Lambda$, given any $\varepsilon > 0$, there are $c_0(\varepsilon) > 0$ and $n_0(\varepsilon, t) \in \mathbb{N}$ such that for all $n \geq n_0(\varepsilon, t)$,

$$\mu \left\{ \omega \in X: \left| \frac{1}{n} \log \|A_t^n(\omega)\| - L_1(A_t) \right| > \varepsilon \right\} < e^{-c_0(\varepsilon)n^{1/3}},$$

that is, a (sub-exponential) large-deviations-type estimate holds for Lebesgue almost every cocycle along the curve $t \mapsto A_t$.

Proof. Given a parameter $t \in \Lambda$, consider the Markov chain $Z_n: X \times \mathbb{P}^1 \rightarrow \mathcal{A} \times \mathbb{P}^1$,

$$Z_n(\omega, \hat{v}) := (\omega_{n+1}, \hat{A}_t^n(\omega)\hat{v}).$$

It is easy to see that its transition kernel is given by

$$\mathcal{A} \times \mathbb{P}^1 \ni (\omega_1, \hat{v}) \mapsto p \times \delta_{\hat{A}_t(\omega_1)\hat{v}},$$

so that its corresponding Markov operator is precisely the operator $\bar{\mathcal{Q}}_t$ defined above, whose stationary measure is $p \times \eta_t$. Recall that by Corollary 3.1, $\bar{\mathcal{Q}}_t$ is strongly mixing on $(L^\infty(\mathcal{A} \times \mathbb{P}^1), \|\cdot\|_\infty)$.

For the observable $\varphi_t(i, \hat{v}) := \log \|A_t(i) \frac{v}{\|v\|}\|$, we have that

$$\begin{aligned} \varphi_t \circ Z_{n-1}(\omega, \hat{v}) &= \varphi_t(\omega_n, \hat{A}_t^{n-1}(\omega)\hat{v}) = \log \left\| A_t(\omega_n) \frac{\hat{A}_t^{n-1}(\omega)\hat{v}}{\|\hat{A}_t^{n-1}(\omega)\hat{v}\|} \right\| \\ &= \log \|\hat{A}_t^n(\omega)\hat{v}\| - \log \|\hat{A}_t^{n-1}(\omega)\hat{v}\|. \end{aligned}$$

Thus,

$$\frac{1}{n} S_n \varphi_t(\omega, \hat{v}) = \frac{1}{n} \log \|A_t^n(\omega)v\|,$$

where v is a unit vector representative of the point $\hat{v}$.

Moreover, by Furstenberg's formula (3.3),

$$L_1(A_t) = \int \psi_t(\hat{v}) d\eta_t(\hat{v}) = \int \varphi_t(i, \hat{v}) d(p \times \eta_t)(i, \hat{v}).$$

It is not difficult to verify that Lemma 3.4 (iii) implies that for Lebesgue a.e. $t \in \Lambda$, $L_1(A_t) > -\infty$.

Our statement (in fact, a stronger version thereof) would then immediately follow from (3.4) if the observable φ_t were bounded. That is of course not the case, precisely because of the singular matrices. The idea is then to use the truncations $\varphi_{t,N}$ (for which estimate (3.4) is applicable) as a substitute for φ_t , where the order N of the truncation is adapted to the scale n at which we prove the LDT estimate.

More precisely, given any large enough scale $n \in \mathbb{N}$, let $N_n \in \mathbb{N}$ be a truncation order to be chosen later (anticipating, the choice that optimises the estimates will turn out to be $N_n \sim n^{1/3}$). We will transfer the LDT estimate at scale n from φ_{t,N_n} to φ_t by eliminating an $\tilde{\eta}$ -negligible set of parameters $(t, \hat{v})$.

For all $n \in \mathbb{N}$, define the following real-valued functions on $X \times \Lambda \times \mathbb{P}^1$:

$$\begin{aligned} \Delta_n(\omega, t, \hat{v}) &:= \frac{1}{n} S_n \varphi_t(\omega, \hat{v}) - \int \varphi_t dp \times \eta_t \\ E_n(\omega, t, \hat{v}) &:= \frac{1}{n} S_n \varphi_{t,N_n}(\omega, \hat{v}) - \int \varphi_{t,N_n} dp \times \eta_t \\ g_n(\omega, t, \hat{v}) &:= \frac{1}{n} S_n \varphi_t(\omega, \hat{v}) - \frac{1}{n} S_n \varphi_{t,N_n}(\omega, \hat{v}). \end{aligned}$$

Then

$$|\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| \leq |g_n(\omega, t, \hat{v})| + \int |\varphi_t - \varphi_{t,N_n}| dp \times \eta_t. \quad (3.5)$$

Note that

$$\Delta_n(\omega, t, \hat{v}) = \frac{1}{n} \log \|A_t^n(\omega)v\| - L_1(A_t).$$

Moreover, given $\varepsilon > 0$, the general LDT estimate (3.4) and the comments following it show that *for every* $(t, \hat{v}) \in \Lambda \times \mathbb{P}^1$,

$$\mu\left\{\omega \in X: |E_n(\omega, t, \hat{v})| > \frac{\varepsilon}{2}\right\} \leq 8e^{-c_0(\varepsilon)N_n^{-2n}}, \quad (3.6)$$

where $c_0(\varepsilon)$ is essentially of order ε^2 . Using Lemma 3.6, we have

$$\begin{aligned} & \int \left(\int |\varphi_t - \varphi_{t,N_n}| dp \times \eta_t \right) dt \\ &= \sum_{i \in \mathcal{A}} p_i \int_{\Lambda} \int_{\mathbb{P}^1} |\varphi_t(i, \hat{v}) - \varphi_{t,N_n}(i, \hat{v})| d\eta_t(\hat{v}) dt \lesssim e^{-N_n/3}. \end{aligned}$$

Moreover, we will also show that

$$\iint |g_n(\omega, t, \hat{v})| d\mu(\omega) d\tilde{\eta}(t, \hat{v}) \lesssim \sqrt{n} e^{-N_n/3}. \quad (3.7)$$

Indeed, note that

$$g_n(\omega, t, \hat{v}) = \frac{1}{n} \sum_{j=1}^n (\varphi_t(\omega_j, \hat{A}_t^{j-1}(\omega)\hat{v}) - \varphi_{t,N_n}(\omega_j, \hat{A}_t^{j-1}(\omega)\hat{v})).$$

Given $\omega \in X$, if $g_n(\omega, t, \hat{v}) \neq 0$, then there is $1 \leq j \leq n$ such that

$$\varphi_t(\omega_j, \hat{A}_t^{j-1}(\omega)\hat{v}) \neq \varphi_{t,N_n}(\omega_j, \hat{A}_t^{j-1}(\omega)\hat{v}),$$

that is, $\varphi_t(\omega_j, \hat{A}_t^{j-1}(\omega)\hat{v}) < -N_n$, or $(t, \hat{v}) \in P_{\omega_j, j-1, N_n}(\omega)$.

Thus, for $(t, \hat{v}) \notin B_n(\omega) := \bigcup_{j=1}^n P_{\omega_j, j-1, N_n}(\omega)$, where $\tilde{\eta}(B_n(\omega)) \lesssim n e^{-N_n}$, we have that $g_n(\omega, t, \hat{v}) = 0$.

Using Fubini and Cauchy–Schwarz, it follows that

$$\begin{aligned} \iint |g_n(\omega, t, \hat{v})| d\mu(\omega) d\tilde{\eta}(t, \hat{v}) &= \iint |g_n(\omega, t, \hat{v})| d\tilde{\eta}(t, \hat{v}) d\mu(\omega) \\ &= \iint |g_n(\omega, t, \hat{v})| \mathbb{1}_{B_n(\omega)}(t, \hat{v}) d\tilde{\eta}(t, \hat{v}) d\mu(\omega) \\ &\leq \int \|g_n(\omega, \cdot)\|_{L^2(\tilde{\eta})} \tilde{\eta}(B_n(\omega))^{1/2} d\mu(\omega) \\ &\lesssim \sqrt{n} e^{-N_n/2} \int \|g_n(\omega, \cdot)\|_{L^2(\tilde{\eta})} d\mu(\omega). \end{aligned}$$

By Lemma 3.4 (iii), for all $\omega \in X$,

$$\int \varphi_t^2(\omega_j, \hat{A}_t^{j-1}(\omega)\hat{v}) d\eta_t(\hat{v}) dt \leq C,$$

so

$$\left\| \frac{1}{n} S_n \varphi_t(\omega, \hat{v}) \right\|_{L^2(\tilde{\eta})} \leq \sqrt{C},$$

while $|\varphi_{t,N_n}(i, \hat{w})| \leq N$, so

$$\left\| \frac{1}{n} S_n \varphi_{t,N_n}(\omega, \hat{v}) \right\|_{L^2(\tilde{\eta})} \leq N_n.$$

Thus, for all $\omega \in X$, $\|g_n(\omega, \cdot)\|_{L^2(\tilde{\eta})} \leq \sqrt{C} + N_n$, which then easily implies (3.7).

Integrating inequality (3.5) with respect to μ and $\tilde{\eta}$, we obtain

$$\iint |\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| d\mu(\omega) d\tilde{\eta}(t, \hat{v}) \lesssim \sqrt{n}e^{-N_n/3}.$$

Applying Chebyshev's inequality to the function

$$(t, \hat{v}) \mapsto \int |\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| d\mu(\omega),$$

there is a set $\mathcal{B}_n \subset \Lambda \times \mathbb{P}^1$ such that $\tilde{\eta}(\mathcal{B}_n) \lesssim \sqrt{n}e^{-N_n/6}$ and if $(t, \hat{v}) \notin \mathcal{B}_n$, then

$$\int |\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| d\mu(\omega) < e^{-N_n/6}.$$

Since $\sum_{n \geq 1} \tilde{\eta}(\mathcal{B}_n) < \infty$, by Borel–Cantelli, $\tilde{\eta}$ -almost every $(t, \hat{v})$ belongs to a finite number of sets $\mathcal{B}_n$. That is, for $\tilde{\eta}$ -a.e. $(t, \hat{v})$, there is $n_0(t, \hat{v}) \in \mathbb{N}$ such that for all $n \geq n_0(t, \hat{v})$, we have $(t, \hat{v}) \notin \mathcal{B}_n$, so

$$\int |\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| d\mu(\omega) < e^{-N_n/6}.$$

Applying again Chebyshev, we have that

$$\mu\{\omega \in X : |\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| > e^{-N_n/12}\} \leq e^{-N_n/12}.$$

We have shown that there is a set $\Omega_n \subset X$ with $\mu(\Omega_n) < e^{-N_n/12}$ such that if $\omega \notin \Omega_n$, then

$$|\Delta_n(\omega, t, \hat{v}) - E_n(\omega, t, \hat{v})| < e^{-N_n/12} \leq \frac{\varepsilon}{2},$$

provided n is large enough depending on $(t, \hat{v})$ as above and on ε .

Thus, if $|\Delta_n(\omega, t, \hat{v})| > \varepsilon$, then either $\omega \in \Omega_n$ or $|E_n(\omega, t, \hat{v})| > \frac{\varepsilon}{2}$, the latter holding for ω in a set of measure $\leq 8e^{-c_0(\varepsilon)N_n^{-2n}}$ by (3.6).

We conclude that for $\tilde{\eta}$ -a.e. $(t, \hat{v})$ and for all $\varepsilon > 0$ there is $n_0 = n_0(\varepsilon, t, \hat{v}) \in \mathbb{N}$ such that for all $n \geq n_0$ we have

$$\mu\{\omega : |\Delta_n(\omega, t, \hat{v})| > \varepsilon\} < e^{-N_n/12} + 8e^{-c_0(\varepsilon)N_n^{-2n}} < e^{-c_0(\varepsilon)n^{1/3}},$$

provided we choose $N_n \sim n^{1/3}$. This establishes the first statement of the theorem.

In particular, by the definition of the measure $\tilde{\eta}$, for Lebesgue almost every parameter $t \in \Lambda$ and for η_t -a.e. point $\hat{v}$ this first statement of the theorem holds, namely

$$\mu\left\{\omega \in X : \left| \frac{1}{n} \log \|A_t^n(\omega)v\| - L_1(A_t) \right| > \varepsilon \right\} < e^{-c_0(\varepsilon)n^{1/3}}. \quad (3.8)$$

The support of the measure η_t is the set

$$\{\hat{A}_t^n(s\omega)\hat{r}_s : s \in \mathcal{A}_{\text{sing}}, \omega \in \mathcal{A}_{\text{inv}}^n, n \in \mathbb{N}\}.$$

There is at least one singular symbol, $s \in \mathcal{A}_{\text{sing}}$ and at least one invertible one $i \in \mathcal{A}_{\text{inv}}$. By the positive winding condition,

$$\frac{d}{dt} \hat{A}_t(i) \hat{r}_s \geq c_0 > 0,$$

so $t \mapsto \hat{A}_t(i) \hat{r}_s$ cannot be constant on a set of positive Lebesgue measure. Therefore, $\hat{A}_t(i) \hat{r}_s \neq \hat{r}_s$ for Lebesgue almost every t .

It follows that for Lebesgue almost every $t \in \Lambda$, the support of the measure η_t has at least two points, $\hat{e}_t^1$ and $\hat{e}_t^2$. Thus, (3.8) holds for two linearly independent vectors e_t^1 and e_t^2 so it must hold with the matrix norm $\|A_t^n(\omega)\|$, which establishes the second statement of the theorem. ■

Let us now derive the version of the LDT estimate stated in the introduction.

Proof of Theorem 1.1. Recall that given the partition of the alphabet $\mathcal{A} = \mathcal{A}_{\text{sing}} \sqcup \mathcal{A}_{\text{inv}}$ into two non-empty sets, we denote by $\mathcal{M}$ the set of all cocycles $\underline{A} = (A_1, \dots, A_k) \in \text{Mat}_2(\mathbb{R})^{\mathcal{A}}$ with $\text{rank } A_i = 1$ if $i \in \mathcal{A}_{\text{sing}}$ and $\det(A_j) > 0$ if $j \in \mathcal{A}_{\text{inv}}$. Moreover, $\mathcal{M}^*$ denotes the subset of cocycles in $\mathcal{M}$ for which $\hat{r}_i \neq \hat{k}_j$ for all $i, j \in \mathcal{A}_{\text{sing}}$.

Firstly, note that the set $\mathcal{M} \setminus \mathcal{M}^*$, consisting of cocycles for which a kernel $\hat{r}_i$ coincides with a range $\hat{k}_j$ for some symbols $i, j \in \mathcal{A}_{\text{sing}}$, has zero Lebesgue measure in the analytic manifold $\mathcal{M}$. Secondly, by [10, Corollary 4.5], $L_1(\underline{A}) > -\infty$ for Lebesgue almost every $\underline{A} \in \mathcal{M}$.⁶ We may then neglect cocycles in these negligible sets.

Given any cocycle $\underline{A} \in \mathcal{M}$, for all parameters $t \in [-\pi, \pi]$ define

$$\underline{A}_t := (A_t(1), \dots, A_t(k)),$$

where

$$A_t(i) := \begin{cases} A_i & \text{if } i \in \mathcal{A}_{\text{sing}}, \\ A_i R_t & \text{if } i \in \mathcal{A}_{\text{inv}}, \end{cases} \quad \text{and} \quad R_t := \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

This is an analytic family of cocycles. If $\underline{A} \in \mathcal{M}^*$, then $\underline{A}_t$ takes values in $\mathcal{M}^*$ and it satisfies the Assumptions 3.1; then by Theorem 3.2, $\underline{A}(t)$ satisfies large deviations of sub-exponential type for Lebesgue almost every $t \in \mathbb{R}$. Therefore, since the map $\mathcal{M} \times [-\pi, \pi] \rightarrow \mathcal{M}$, $(\underline{A}, t) \mapsto \underline{A}_t$ is a submersion, Lebesgue almost every $\underline{A} \in \mathcal{M}$ satisfies similar large deviations estimates. ■

⁶Note that the partition $\mathcal{A} = \mathcal{A}_{\text{sing}} \sqcup \mathcal{A}_{\text{inv}}$ in this paper is fixed, and we use the notation $\mathcal{M}$ to refer to the space that in [10] was denoted by $\mathcal{M}_I$ with $I = \mathcal{A}_{\text{sing}}$.

3.4. Central limit theorem

We now establish a central limit theorem for cocycles with singular components. The proof uses the following abstract central limit theorem due to Gordin and Livšic [17].

Let M be a compact metric space, let $M \ni x \mapsto K_x \in \text{Prob}(M)$ be an ergodic⁷ transition kernel, let Q be the corresponding Markov operator and let η be a Q -stationary measure. Given an observable $\varphi \in L^2(M, \eta)$ with $\int \varphi d\eta = 0$, if we have $\sum_{n=0}^{\infty} \|Q^n \varphi\|_2 < \infty$, we can define $g \in L^2(M, \eta)$ by $g := \sum_{n=0}^{\infty} Q^n \varphi$. Then $\varphi = g - Qg$. Let $\sigma^2(\varphi) := \|g\|_2^2 - \|Qg\|_2^2 \in [0, \infty)$.

Theorem 3.3 (Gordin–Livšic). *Let M, Q, η, φ be as above and assume that*

$$\sum_{n=0}^{\infty} \|Q^n \varphi\|_2 < \infty \quad \text{and} \quad \sigma^2(\varphi) > 0.$$

Then we have the following convergence to the standard normal distribution:

$$\frac{S_n \varphi}{\sigma \sqrt{n}} \rightarrow \mathcal{N}(0, 1).$$

Moreover, this holds not only relative to the Markov probability with initial distribution η , but also with initial distribution δ_x for η -a.e. x .

We consider as in the previous subsection a smooth family of cocycles $\Lambda \ni t \mapsto \underline{A}_t \in \mathcal{M}^*$ satisfying Assumptions 3.1. We also consider the special observables ψ_t and φ_t from before.

We introduced two related (families of) Markov operators acting on measurable bounded observables. They can actually be defined for arbitrary observables; namely, if $\psi: \mathbb{P}^1 \rightarrow [-\infty, \infty)$, we put

$$\mathcal{Q}_t \psi(\hat{v}) = \sum_{i \in \mathcal{A}} p_i \psi(\hat{A}_t(i) \hat{v})$$

and if $\varphi: \mathcal{A} \times \mathbb{P}^1 \rightarrow [-\infty, \infty)$, we put

$$\bar{\mathcal{Q}}_t \varphi(j, \hat{v}) = \sum_{i \in \mathcal{A}} p_i \varphi(i, \hat{A}_t(j) \hat{v}).$$

Consider the projection

$$\begin{aligned} \pi: L^2(\mathcal{A} \times \mathbb{P}^1) &\rightarrow L^2(\mathbb{P}^1) \\ \pi \varphi(\hat{v}) &:= \sum_{i \in \mathcal{A}} p_i \varphi(i, \hat{v}). \end{aligned}$$

Using this notation, $\pi \varphi_t = \psi_t$. Moreover, $\bar{\mathcal{Q}}_t^{n+1} \varphi(i, \hat{v}) = \mathcal{Q}_t^n(\pi \varphi)(\hat{A}_t(i) \hat{v})$.

⁷This automatically holds when there is a unique stationary measure.

Lemma 3.7. *For Lebesgue almost every parameter $t \in \Lambda$, we have that*

$$\sum_{n=0}^{\infty} \left\| \bar{\mathcal{Q}}_t^{n+1} \varphi_t - \int \varphi_t dp d\eta_t \right\|_{L^2(p \times \eta_t)} < \infty.$$

Proof. By the previous observations,

$$\begin{aligned} & \sum_{n=0}^{\infty} \left\| \bar{\mathcal{Q}}_t^{n+1} \varphi_t - \int \varphi_t dp d\eta_t \right\|_{L^2} \\ &= \sum_{n=0}^{\infty} \left\| \mathcal{Q}_t^n \psi_t(\hat{A}_t(i)\hat{v}) - \int \psi_t d\eta_t \right\|_{L^2} \\ &= \sum_{n=0}^{\infty} \left(\sum_{i \in \mathcal{A}} p_i \int \left| \mathcal{Q}_t^n \psi_t(\hat{A}_t(i)\hat{v}) - \int \psi_t d\eta_t \right|^2 d\eta_t(\hat{v}) \right)^{1/2}. \end{aligned}$$

We start by showing that for every $i \in \mathcal{A}$, there exists $\sigma \in (0, 1)$ such that for every $n \in \mathbb{N}$,

$$\int_{\Lambda} \int_{\mathbb{P}^1} \left| \mathcal{Q}_t^n \psi_t(\hat{A}_t(i)\hat{v}) - \int \psi_t d\eta_t \right|^2 d\eta_t(\hat{v}) dt \lesssim \sigma^n.$$

Fix $i \in \mathcal{A}$ and note that by Lemma 3.4 (iii) and Jensen's inequality,

$$\begin{aligned} & \int_{\Lambda} \int_{\mathbb{P}^1} \left| \mathcal{Q}_t^n \psi_t(\hat{A}_t(i)\hat{v}) - \int \psi_t d\eta_t \right|^2 d\eta_t(\hat{v}) dt \\ &= \int_{\Lambda} \int_{\mathbb{P}^1} \left| \sum_{\omega \in \mathcal{A}_{\text{inv}}^n} p(\omega) \psi_t(\hat{A}_t^n(\omega) \hat{A}_t(i)\hat{v}) \right. \\ & \quad \left. - \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=n}^{\infty} \sum_{\omega \in \mathcal{A}_{\text{inv}}^j} p(\omega) \psi_t(\hat{A}_t^j(\omega) r_s) \right|^2 d\eta_t(\hat{v}) dt \\ &\leq 2 \int_{\Lambda} \int_{\mathbb{P}^1} \sum_{\omega \in \mathcal{A}_{\text{inv}}^n} p(\omega) |\psi_t(\hat{A}_t^n(\omega) \hat{A}_t(i)\hat{v})|^2 d\tilde{\eta}(t, \hat{v}) \\ & \quad + 2 \int_{\Lambda} \int_{\mathbb{P}^1} \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=n}^{\infty} \sum_{\omega \in \mathcal{A}_{\text{inv}}^j} p(\omega) |\psi_t(\hat{A}_t^j(\omega) r_s)|^2 d\tilde{\eta}(t, \hat{v}) \\ &\lesssim \sum_{\omega \in \mathcal{A}_{\text{inv}}^n} p(\omega) + \sum_{s \in \mathcal{A}_{\text{sing}}} p_s \sum_{j=n}^{\infty} \sum_{\omega \in \mathcal{A}_{\text{inv}}^j} p(\omega) \lesssim (1-q)^n, \end{aligned}$$

so the claim holds with $\sigma = 1 - q$.

By Chebyshev's inequality,

$$\text{Leb} \left\{ t \in \Lambda : \int_{\mathbb{P}^1} \left| \mathcal{Q}_t^n \psi_t(\hat{A}_t(i)\hat{v}) - \int \psi_t d\eta_t \right|^2 d\eta_t(\hat{v}) > \sigma^{n/2} \right\} \lesssim \sigma^{n/2}.$$

Therefore, for every $i \in \mathcal{A}$ and for each $n \in \mathbb{N}$, there exists a set $B_n(i) \subset \Lambda$, with $\text{Leb}(B_n(i)) \lesssim \sigma^{n/2}$ such that for every $t \notin B_n(i)$,

$$\left\| \bar{\mathcal{Q}}_t^{n+1} \varphi_t - \int \varphi_t dp d\eta_t \right\|_{L^2(p \times \eta_t)} \lesssim \sigma^{n/2}. \quad (3.9)$$

Moreover, since there is a finite number of symbols, there exists a set B_n that satisfies the same properties for all symbols $i \in \mathcal{A}$ simultaneously. Furthermore, since $\sum_{n=0}^{\infty} \text{Leb}(B_n) < \infty$, by the Borel–Cantelli lemma, for almost every $t \in \Lambda$, there exists $n_0(t) \in \mathbb{N}$ such that for every $n \geq n_0(t)$, $t \notin B_n$. Thus, for almost every t , the inequality (3.9) holds, hence

$$\sum_{n=0}^{\infty} \left\| \bar{\mathcal{Q}}_t^{n+1} \varphi_t - \int \varphi_t dp d\eta_t \right\|_{L^2(p \times \eta_t)} < \infty,$$

which proves the lemma. ■

We are ready to formulate and to prove a parametric version of the CLT for cocycles with singular components.

Theorem 3.4. *Let $A: \Lambda \rightarrow \mathcal{M}^*$ be a smooth family of cocycles satisfying Assumptions 3.1. Then, for almost every $t \in \Lambda$, there exists $\sigma = \sigma(t) > 0$ such that*

$$\frac{\log \|A_t^n v\| - nL_1(A_t)}{\sigma \sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

holds for η_t -a.e. $\hat{v} \in \mathbb{P}^1$ and, moreover,

$$\frac{\log \|A_t^n\| - nL_1(A_t)}{\sigma \sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Proof. If A_t is projectively uniformly hyperbolic for every $t \in \Lambda$, then the cocycle reduces to a 1-dimensional cocycle to which the classical CLT applies. Hence, we assume from now on that the cocycle family A_t is not projectively uniformly hyperbolic for some $t_0 \in \Lambda$. By Theorem 3.5, there exist parameters $\Lambda \ni t_n \rightarrow t_0$ for which A_{t_n} admits a null word, which implies a singularity $L_1(A_{t_n}) = -\infty$.

Consider the special observable φ_t defined before and recall from Lemma 3.7 that for Lebesgue almost every parameter $t \in \Lambda$, we have

$$\sum_{n=0}^{\infty} \left\| \bar{\mathcal{Q}}_t^n \varphi_t - \int \varphi_t dp d\eta_t \right\|_{L^2(p \times \eta_t)} < \infty.$$

This in particular allows us to define

$$g_t := \sum_{n=0}^{\infty} \bar{\mathcal{Q}}_t^n \left(\varphi_t - \int \varphi_t dp d\eta_t \right) \in L^2(p \times \eta_t).$$

Then for almost every parameter t ,

$$\varphi_t - \int \varphi_t dp d\eta_t = g_t - \bar{\mathcal{Q}}_t g_t.$$

Let $\sigma^2(t) = \sigma_t^2(\varphi_t) := \|g_t\|_2^2 - \|\bar{\mathcal{Q}}_t g_t\|_2^2 \in [0, \infty)$.

In order to apply Theorem 3.3, it remains to prove that $\sigma^2(t) > 0$ for almost every $t \in \Lambda$. We accomplish this in several steps.

We recall the following definitions and results. A cocycle is said to be *projectively uniformly hyperbolic* if it admits a dominated splitting [10, Definition 3.1].

In the Bernoulli case, an *invariant multi-cone* is a non-empty, non-dense open subset $M \subset \mathbb{P}^1$ such that $\overline{A_i M} \subset M$ for every $i \in \mathcal{A}$. In [10, Theorem 3.1] we proved that a cocycle $\underline{A}$ is projectively uniformly hyperbolic if and only if it admits an invariant multi-cone.

Assume now that the invertible cocycle $\underline{A}_{\text{inv}} = (A_i)_{i \in \mathcal{A}_{\text{inv}}}$ is projectively uniformly hyperbolic, and let M be an invariant multi-cone. Define

$$K_{\text{inv}}^u = \bigcap_{n=0}^{\infty} \bigcup_{i_1, \dots, i_n \in \mathcal{A}_{\text{inv}}} A_{i_n} \cdots A_{i_1}(M),$$

$$K_{\text{inv}}^s = \bigcap_{n=0}^{\infty} \bigcup_{i_1, \dots, i_n \in \mathcal{A}_{\text{inv}}} (A_{i_n} \cdots A_{i_1})^{-1}(\mathbb{P}^1 \setminus \bar{M}).$$

These sets are, respectively, the closures of the unstable and stable Oseledets directions of $\underline{A}_{\text{inv}}$ (see [10, Proposition 3.1]).

Definition 3.1. A smooth family of cocycles $\underline{A}: \Lambda \rightarrow \text{Mat}_2^+(\mathbb{R})^k$ is called *rich* if for every $p_0 \in \Lambda$ there exists a smooth curve $t \mapsto \alpha(t) \in \Lambda$ through p_0 such that the one-parameter family $\underline{A}(\alpha(t))$ is positively winding.

The following theorem was proved in [10, Theorem 4.1].

Theorem 3.5. Let $\underline{A}: \Lambda \rightarrow \text{Mat}_2^+(\mathbb{R})^k$ be a rich smooth family of cocycles on an open set $\Lambda \subseteq \mathbb{R}^m$, and suppose there exists $i \in \{1, \dots, k\}$ such that $\text{rank}(A_i(t)) = 1$ for every $t \in \Lambda$. Then, for each $t \in \Lambda$, either $\underline{A}(t)$ is projectively uniformly hyperbolic, or there exists a sequence $t_n \rightarrow t$ in Λ such that $\underline{A}(t_n)$ has a periodic null word, hence $L_1(\underline{A}(t_n)) = -\infty$ for all $n \in \mathbb{N}$.

For the Markov case, we refer to the companion paper [10, Definitions 3.1 and 3.5, Proposition 3.1, and Theorem 1.1].

Lemma 3.8. *Let $\mathcal{H}$ be the set of parameters $t \in \Lambda$ such that the corresponding cocycle $\underline{A}_t$ is projectively uniformly hyperbolic. For Lebesgue, almost every $t \in \mathcal{H}$, the corresponding stationary measure $p \times \eta_t$ has infinite support.*

Proof. Firstly, note that if the parameter $t \in \mathcal{H}$ is such that $\underline{A}_t$ is diagonalizable, then the ranges of the singular matrices and the unstable directions of the invertible matrices are aligned. Therefore, the support of the corresponding stationary measure has only one point. However, this case only happens with zero measure, since this alignment will be undone by the winding property as the parameter t varies. Note that singular matrices remain constant in the process.

Hence, we only need to consider the case in which the cocycle $\underline{A}_t$ is not diagonalizable. Note that if $\underline{A}_t$ is projectively uniformly hyperbolic, then so is its invertible part $\underline{A}_{\text{inv}}(t) := (A_t(i))_{i \in \mathcal{A}_{\text{inv}}}$. Hence, we can define K_{inv}^u , the set of unstable Oseledets directions $E^u(\omega)$ of the cocycle $(A_i)_{i \in \mathcal{A}_{\text{inv}}}$ over the set of points $\omega \in \mathcal{A}_{\text{inv}}^{\mathbb{Z}}$. We divide the proof in two cases: $\#K_{\text{inv}}^u = 1$ and $\#K_{\text{inv}}^u > 1$.

First, we suppose that $\#K_{\text{inv}}^u = 1$. Since we assume that $\underline{A}_t$ is not diagonalizable, there is no range in K_{inv}^u . Therefore, the iterates of the ranges are going to spread into infinitely many different points in the projective space by the projective uniform hyperbolic dynamics (in fact, the iterates will converge to K_{inv}^u). Since the corresponding stationary measure $p \times \eta_t$ is discrete and gives positive weight to every pair of the form $(j, \hat{A}_t^n(\underline{\omega})\hat{r}_s)$, we conclude that the support of η_t is infinite.

Now, suppose that $\#K_{\text{inv}}^u > 1$. In this case, K_{inv}^u is in fact infinite, as it is the limit set of a contracting iterated function system; it may be a Cantor set, an interval, or a combination of both.

If any range lies outside K_{inv}^u , the proof proceeds as in the previous case, so we may assume that every range r_i is contained in K_{inv}^u .

Moreover, since $\underline{A}_{\text{inv}}$ is projectively uniformly hyperbolic (and topologically mixing in the Markov case), for any $\hat{v} \in K_{\text{inv}}^u$ and any point in an invariant multi-cone of $\underline{A}_{\text{inv}}$, in particular for any range $\hat{r}_i$, there exists an admissible word $\omega \in \mathcal{A}_{\text{inv}}^{\mathbb{N}}$ such that

$$\hat{v} = \lim_{n \rightarrow \infty} \hat{A}_t^n(\omega)\hat{r}_i.$$

Hence, by a suitable choice of ω , one can iterate any range r_i by $A_t^n(\omega)$ so as to converge to any prescribed element of K_{inv}^u . Since the corresponding stationary measure $p \times \eta_t$ is discrete and gives positive weight to every pair of the form $(j, \hat{A}_t^n(\underline{\omega})\hat{r}_s)$, we conclude that the support of η_t is infinite. ■

Lemma 3.9. *For Lebesgue almost every parameter $t \in \Lambda$, the observable φ_t cannot be $p \times \eta_t$ constant.*

Proof. By Theorem 3.5, for every $t \in \Lambda$, either $\underline{A}_t$ is projectively uniformly hyperbolic or it can be approximated by cocycles that admit null words. The proof is then divided into two cases.

First, let us consider the set $\mathcal{B}$ of parameters $t \in \Lambda$ such that $\underline{A}_t$ is not projectively uniformly hyperbolic. Note that there exists $c > 0$ such that for every $t \in \mathcal{B}$ and all $j \in \mathcal{A}$ and $i \in \mathcal{A}_{\text{sing}}$, it holds that $|\varphi_t(j, \hat{r}_i)| \leq c$. Moreover, by Theorem 3.5, there exists t' arbitrarily close to t such that $\underline{A}_{t'}$ has a null word. Hence, by continuity, there exists some range r_s and a finite word $\underline{\omega}$ such that $\|A_{t'}^n(\underline{\omega})r_s\|$ is arbitrarily small, thus $|\varphi_t| \gg c$. Therefore, we conclude that for every $t \in \mathcal{B}$, φ_t is not constant.

By Lemma 3.8, we have that for Lebesgue almost every $t \in \mathcal{H}$, there are infinitely many points in the support of the corresponding stationary measure $p \times \eta_t$. Moreover, φ_t is $p \times \eta_t$ -constant if and only if it takes the same value at every point in the support of $p \times \eta_t$. The only way for this to happen is if $\underline{A}_t$ had a conformal word (with a pair of non-real eigenvalues). In fact, the presence of a conformal word implies that this word is projectively elliptic, hence not projectively uniformly hyperbolic. Therefore, we conclude that for almost every $t \in \mathcal{H}$, the observable φ_t is not $p \times \eta_t$ constant, which finishes the proof. ■

The next lemma establishes the positivity of the variance-type quantity $\sigma^2(t)$ for almost every $t \in \Lambda$ by an adaptation of the proof of [7, Proposition 2.2]. We note that this proposition is a version of the abstract CLT of Gordin and Livšic, which is more applicable to dynamical systems.

Lemma 3.10. *For almost every $t \in \Lambda$ we have $\sigma^2(t) > 0$.*

Proof. Consider the family of Markov operators

$$\bar{\mathcal{Q}}_t: L^\infty(\mathcal{A} \times \mathbb{P}^1) \rightarrow L^\infty(\mathcal{A} \times \mathbb{P}^1),$$

given by

$$(\bar{\mathcal{Q}}_t \varphi)(j, \hat{v}) = \sum_{i \in \mathcal{A}} \varphi(i, \hat{A}_t(j) \hat{v}) p_i$$

and the corresponding family of Markov kernels

$$\bar{K}_t: \mathcal{A} \times \mathbb{P}^1 \rightarrow \text{Prob}(\mathcal{A} \times \mathbb{P}^1).$$

Recall that the measure $p \times \eta_t$ is $\bar{K}_t$ -stationary, in the sense that $(\bar{K}_t)_* p \times \eta_t = p \times \eta_t$.

Fix any $t \in \Lambda$. Assume by contradiction that $\sigma_t^2(\varphi) = 0$. Then

$$\begin{aligned} 0 &\leq \int (\bar{\mathcal{Q}}_t g_t(x) - g_t(y))^2 d\bar{K}_{t,x}(y) dp \times \eta_t(x) \\ &= \int \{(\bar{\mathcal{Q}}_t g_t(x))^2 + (g_t(y))^2 - 2g_t(y)\bar{\mathcal{Q}}_t g_t(x)\} d\bar{K}_{t,x}(y) dp \times \eta_t(x) \end{aligned}$$

$$\begin{aligned} &= \int \{(g_t(y))^2 - (\bar{\mathcal{Q}}_t g_t(x))^2\} d\bar{K}_{t,x}(y) dp \times \eta_t(x) \\ &= \|g_t\|_2^2 - \|\bar{\mathcal{Q}}_t g_t\|_2^2 = \sigma^2(t) = 0. \end{aligned}$$

We then conclude that $g_t(y) = \bar{\mathcal{Q}}_t g_t(x)$ for $p \times \eta_t$ -a.e x and $\bar{K}_{t,x}$ -a.e y . In particular, define

$$X_j(\omega) := \hat{A}_t^j(\omega) \hat{r}_i,$$

where $\omega \in \mathcal{A}^{\mathbb{N}}$ is a periodic admissible word with $\omega_0 = \omega_m = i \in \mathcal{A}_{\text{sing}}$. Then

$$\begin{aligned} \log \|A_t^n(\omega) r_i\| - nL_1(\underline{A}_t) &= \sum_{j=0}^{n-1} [\varphi_t(X_j(\omega)) - \int \varphi_t dp d\eta_t] \\ &= \sum_{j=0}^{n-1} [g_t(X_j(\omega)) - (\bar{\mathcal{Q}}_t g_t)(X_j(\omega))] \\ &= \sum_{j=0}^{n-1} [g_t(X_j(\omega)) - g_t(X_{j+1}(\omega))] \\ &= g_t(X_0(\omega)) - g_t(X_n(\omega)) = 0. \end{aligned}$$

This telescoping sum shows that for every n -periodic orbit ω containing the singular symbol $i \in \mathcal{A}_{\text{sing}}$ its dominant eigenvalue satisfies

$$|\lambda_{\max}(A_t^n(\omega))| = e^{nL_1(\underline{A}_t)}.$$

In particular, for the fixed point $\omega \equiv i$,

$$\beta := |\langle r_i, A_i r_i \rangle| = e^{L_1(\underline{A}_t)}$$

is independent of the parameter t . Consequently, for any admissible n -periodic orbit ω with $\omega_0 = \omega_n = i$,

$$\log |\langle r_i, A_t^n(\omega) r_i \rangle| = n \log \beta.$$

Observe that the left-hand side defines a non-constant harmonic function of t , except possibly at its singularities where vanishing words occur (see the beginning of this proof). Therefore, the above equality can only be satisfied for a discrete set of parameter values $t \in \Lambda$. For all other values of t , the equality fails, yielding a contradiction. It follows that for a.e. $t \in \Lambda$,

$$\sigma^2(t) > 0,$$

as claimed. ■

Theorem 3.3 is then applicable to the Markov kernel $\bar{K}_t$ on $\mathcal{A} \times \mathbb{P}^1$ and the observable φ_t , for Lebesgue almost every $t \in \Lambda$. We then conclude that for η_t -a.e. point $\hat{v} \in \mathbb{P}^1$,

$$\frac{\log \|A_t^n v\| - nL_1(A_t)}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1),$$

thus establishing the first statement of the theorem.

To prove the second statement, with $\|A_t^n v\|$ replaced by $\|A_t^n\|$, we choose a unit vector v for which the CLT holds and note that

$$\frac{\log \|A_t^n\| - nL_1(A_t)}{\sigma\sqrt{n}} = \frac{\log \|A_t^n v\| - nL_1(A_t)}{\sigma\sqrt{n}} + \frac{\log \|A_t^n\| - \log \|A_t^n v\|}{\sigma\sqrt{n}}.$$

We claim that the sequence $\log \|A_t^n\| - \log \|A_t^n v\|$ is almost surely bounded, hence the last term above converges to 0 almost surely, which would then conclude the argument via Slutsky's theorem.

The claim clearly follows from the a.s. boundedness from below by a positive constant of the sequence $\frac{\|A_t^n v\|}{\|A_t^n\|}$. Let $\bar{v}(A)$ denote the most expanding singular direction of the matrix A . By the Pythagorean's theorem,

$$\frac{\|A_t^n(\omega)v\|}{\|A_t^n(\omega)\|} \geq |v \cdot \bar{v}(A_t^n(\omega))| \rightarrow |v \cdot \bar{v}^\infty(A_t)(\omega)|.$$

The limit direction $\bar{v}^\infty(A_t)(\omega)$ exists almost surely by arguments in the proof of Oseledets' theorem (see for instance [11, Proposition 4.4]) and $\bar{v}^\infty(A_t)(\omega)^\perp = E^s(\omega)$, the stable subspace in the Oseledets theorem (see for instance the beginning of the proof of [11, Theorem 4.4]).

We must have that $|v \cdot \bar{v}^\infty(A_t)(\omega)| > 0$ almost surely (which then establishes the claim). Otherwise, on a set of positive measure we would have that $v \in \bar{v}^\infty(A_t)(\omega)^\perp = E^s(\omega)$, which is not possible. Indeed, $E^s(\omega)$ consists of the pre-images of the kernel k_{ω_n} via matrix products $A_t^n(\omega)$, for some $n \in \mathbb{N}$ such that $\omega_n \in \mathcal{A}_{\text{sing}}$ and $\omega_0, \dots, \omega_{n-1} \in \mathcal{A}_{\text{inv}}$. On the other hand, $\hat{v}$ is in the support of η_t , which consists of images of ranges $\hat{r}_s$ via matrix products $A_t^n(y)$. Thus, $v \in E^s(\omega)$ would imply the existence of null words, and in particular the fact that $L_1(\underline{A}_t) = -\infty$. But those parameters t form a zero measure set. ■

Theorem 1.2, the non-parametric version of the CLT stated in the introduction, can be derived from the parametric version exactly the same way the non-parametric LDT Theorem 1.1 was derived from the parametric LDT Theorem 3.2. Thus, Theorem 1.2 holds as well.

Remark 3.3. We note that the statistical properties derived above are sensitive to perturbations of the cocycle, that is, the parameters appearing in these estimates are not locally uniform.

Remark 3.4. Projectively uniformly hyperbolic cocycles automatically satisfy uniform LDT estimates and a CLT, since these properties can be immediately reduced to the corresponding limit laws for i.i.d. additive processes in classical probabilities. In particular, by [10, Theorem 4.2], if all components of $\underline{A} \in \text{Mat}_2(\mathbb{R})^k$ are singular and $L_1(\underline{A}) > -\infty$, then $\underline{A}$ satisfies uniform LDT estimates and a CLT.

Funding. P.D. was partially supported by FCT – Fundação para a Ciência e a Tecnologia, through the projects UIDB/04561/2020 and PTDC/MAT-PUR/29126/2017. M.D. was supported by a CNPq doctoral fellowship and also by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. T.G. was supported by FCT - Fundação para a Ciência e a Tecnologia, through the projects UI/BD/152275/2021 and by CEMS.UL Ref. UIDP/04561/2020, DOI 10.54499/UIDP/04561/2020. S.K. was supported by the CNPq research grant 313777/2020-9 and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

- [1] A. Avila, J. Bochi, and J.-C. Yoccoz, [Uniformly hyperbolic finite-valued \$\text{SL}\(2, \mathbb{R}\)\$ -cocycles](#). *Comment. Math. Helv.* **85** (2010), no. 4, 813–884 Zbl [1201.37032](#) MR [2718140](#)
- [2] V. Barbu and N. Limnios, [Discrete-time renewal processes](#). In *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications*, pp. 17–40, Springer, New York, 2008
- [3] P. G. Barrientos and D. Malicet, Mostly contracting random maps. [v1] 2024, [v2] 2025, [arXiv:2412.03729v2](#)
- [4] Y. Benoist and J.-F. Quint, [Central limit theorem for linear groups](#). *Ann. Probab.* **44** (2016), no. 2, 1308–1340 Zbl [1341.22006](#) MR [3474473](#)
- [5] J. Bezerra, A. Cai, P. Duarte, C. Freijo, and S. Klein, [A dynamical Thouless formula](#). *Adv. Math.* **438** (2024), article no. 109446 Zbl [1536.37052](#) MR [4674864](#)
- [6] J. Bourgain and M. Goldstein, [On nonperturbative localization with quasi-periodic potential](#). *Ann. of Math. (2)* **152** (2000), no. 3, 835–879 Zbl [1053.39035](#) MR [1815703](#)
- [7] A. Cai, P. Duarte, and S. Klein, [Statistical properties for mixing Markov chains with applications to dynamical systems](#). *Adv. Math.* **479** (2025), no. part B, article no. 110454 Zbl [08090559](#) MR [4939422](#)
- [8] W. Craig and B. Simon, [Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices](#). *Comm. Math. Phys.* **90** (1983), no. 2, 207–218 Zbl [0532.60057](#) MR [0714434](#)
- [9] D. Damanik, [Schrödinger operators with dynamically defined potentials](#). *Ergodic Theory Dynam. Systems* **37** (2017), no. 6, 1681–1764 Zbl [1541.81060](#) MR [3681983](#)
- [10] P. Duarte, M. Durães, T. Graxinha, and S. Klein, Random 2d linear cocycles I: dichotomic behavior. 2025, [arXiv:2503.21050v1](#)

- [11] P. Duarte and S. Klein, *Lyapunov exponents of linear cocycles*. Atlantis Stud. Dyn. Syst. 3, Atlantis Press, Paris, 2016 Zbl 1364.37001 MR 3468528
- [12] P. Duarte and S. Klein, *Large deviations for products of random two-dimensional matrices*. *Comm. Math. Phys.* **375** (2020), no. 3, 2191–2257 Zbl 1443.37039 MR 4091498
- [13] H. Furstenberg, *Noncommuting random products*. *Trans. Amer. Math. Soc.* **108** (1963), 377–428 Zbl 0203.19102 MR 0163345
- [14] H. Furstenberg and Y. Kifer, *Random matrix products and measures on projective spaces*. *Israel J. Math.* **46** (1983), no. 1-2, 12–32 Zbl 0528.60028 MR 0727020
- [15] M. Goldstein and W. Schlag, *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions*. *Ann. of Math. (2)* **154** (2001), no. 1, 155–203 Zbl 0990.39014 MR 1847592
- [16] M. Goldstein and W. Schlag, *Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues*. *Geom. Funct. Anal.* **18** (2008), no. 3, 755–869 Zbl 1171.82011 MR 2438997
- [17] M. I. Gordin and B. A. Livšic, *The central limit theorem for stationary Markov processes* (in Russian). *Dokl. Akad. Nauk SSSR* **239** (1978), 766–767. English translation: *Sov. Math., Dokl.* **19** (1978), 392–394 Zbl 0395.60057
- [18] H. Hennion and L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*. Lecture Notes in Math. 1766, Springer, Berlin, 2001 Zbl 0983.60005 MR 1862393
- [19] É. Le Page, *Théorèmes limites pour les produits de matrices aléatoires*. In *Probability measures on groups (Oberwolfach, 1981)*, pp. 258–303, Lecture Notes in Math. 928, Springer, Berlin ehg., 1982 Zbl 0506.60019 MR 0669072
- [20] É. Le Page, *Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications*. *Ann. Inst. H. Poincaré Probab. Statist.* **25** (1989), no. 2, 109–142 Zbl 0679.60010 MR 1001021
- [21] S. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*. Second edn., Cambridge University Press, Cambridge, 2009 Zbl 1165.60001 MR 2509253
- [22] S. V. Nagaev, *Some limit theorems for stationary Markov chains* (in Russian). *Teor. Veroyatnost. i Primenen.* **2** (1957), 389–416. English translation: *Theory Probab. Appl.* **2** (1957), no. 4, 378–406 Zbl 0078.31804 MR 0094846
- [23] E. H. Y. Tall and M. Viana, *Moduli of continuity for the Lyapunov exponents of random $GL(2)$ -cocycles*. *Trans. Amer. Math. Soc.* **373** (2020), no. 2, 1343–1383 Zbl 1437.37064 MR 4068266
- [24] M. Viana, *Lectures on Lyapunov exponents*. Cambridge Stud. Adv. Math. 145, Cambridge University Press, Cambridge, 2014 Zbl 1309.37001 MR 3289050
- [25] M. Viana and K. Oliveira, *Foundations of ergodic theory*. Cambridge Stud. Adv. Math. 151, Cambridge University Press, Cambridge, 2016 Zbl 1369.37001 MR 3558990

Pedro Duarte

Departamento de Ciências Matemáticas, Faculdade de Ciências, Universidade de Lisboa and CEMS.UL, Campo Grande, Edifício C6, Piso 2, 1749-016 Lisboa, Portugal;

pmduarte@fc.ul.pt

Author IDs: zbMATH [duarte.pedro](#) MR [352560](#) ORCID [0000-0002-1838-3770](#)

Marcelo Durães

Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rua Marquês de São Vicente, 225, Gávea, Cx. Postal: 38097, 22451-900 Rio de Janeiro, Brazil; Departamento de Ciências Matemática, Faculdade de Ciências, Universidade de Lisboa and CEMS.UL, Campo Grande, Edifício C6, Piso 2, 1749-016 Lisboa, Portugal;

accp95@gmail.com

Author IDs: zbMATH [duraes.marcelo](#) ORCID [0000-0003-3425-8491](#)

Tomé Graxinha

Departamento de Ciências Matemática, Faculdade de Ciências, Universidade de Lisboa and CEMS.UL, Campo Grande, Edifício C6, Piso 2, 1749-016 Lisboa, Portugal;

graxinhato@gmail.com

Author IDs: zbMATH [graxinha.tome](#) ORCID [0009-0007-2233-0990](#)

Silvius Klein

Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rua Marquês de São Vicente, 225, Gávea, Cx. Postal: 38097, 22451-900 Rio de Janeiro, Brazil; silviusk@puc-rio.br

Author IDs: zbMATH [klein.silvius](#) MR [750010](#) ORCID [0000-0001-6375-4561](#)