Limit theorems for a strongly irreducible product of independent random matrices with optimal moment conditions

Axel Péneau

February 1, 2024

Abstract

Let ν be a probability distribution over the linear semi-group $\mathrm{End}(E)$ for E a finite dimensional vector space over a locally compact field. We assume that ν is proximal, strongly irreducible and that $\nu^{*n}\{0\} = 0$ for all integers $n \in \mathbb{N}$. We consider the sequence $\overline{\gamma}_n := \gamma_0 \cdots \gamma_{n-1}$ for (γ_k) independents of distribution law ν. We define the logarithmic singular gap as $\log\left(\frac{\mu_1(\overline{\gamma}_n)}{\mu_2(\overline{\gamma}_n)}\right)$, where μ_1 and μ_2 are the two largest singular values. We show that $(\mathrm{sqz}(\overline{\gamma}_n))_{n\in\mathbb{N}}$ escapes to infinity linearly and satisfies exponential large deviations estimates below its escape rate. Using this escape speed, we also show that the image of a generic line by $\overline{\gamma}_n$ as well as its eigenspace of maximal eigenvalue both converge to the same random line l_∞ at an exponential speed. This is an extension of results by Guivarc'h and Raugi.

If we moreover assume that the push-forward distribution $N(\nu)$ is L^p for $N : g \mapsto \log\left(\|g\|\|g^{-1}\|\right)$ and for some $p \geq 1$, then we show that $\log|w(l_\infty)|$ is L^p for all unitary linear form w and the logarithm of each coefficient of $\overline{\gamma}_n$ is almost surely equivalent to the logarithm of the norm. This is an extension of results by Furstenberg and Kesten.

To prove these results, we do not rely on any classical results for random products of invertible matrices with L^1 moment assumption. Instead we describe an effective way to group the i.i.d factors into i.i.d random words that are aligned in the Cartan projection. We moreover have an explicit control over the moments.

Contents

1 Introduction

1.1 Results

Let E be a vector space of finite dimension $d \geq 2$ over a local field $\mathbb{K}$ and let $\Gamma := \mathrm{End}(E) \simeq \mathrm{Mat}_{d\times d}(\mathbb{K})$ be the monoid of linear maps on E. We endow Γ with the operator norm $\|\gamma\| = \max_{x\in E\setminus\{0\}} \frac{\|\gamma x\|}{\|x\|}$ where $\|\cdot\|$ is a Euclidean or ultra-metric norm on E. Note that Γ is a finite dimensional $\mathbb{K}$ vector space and $\mathbb{K}$ is a metric space, we endow Γ with the Borel σ-algebra associated to this structure. We also define:

$$N: \begin{array}{ccc} \mathrm{GL}(E) & \longrightarrow & [0,+\infty) \\ \gamma & \longmapsto & \log\|\gamma\| + \log\|\gamma^{-1}\|. \end{array}$$

We use the convention $N(\gamma) = +\infty$ when γ is not invertible, that way N is a sub-additive map $\Gamma \to [0, +\infty]$.

Let ν be a probability distribution on Γ , let $\mathbb{N} := \mathbb{N}_0$ be the set of non-negative integers. Draw $(\gamma_n) \in \Gamma^{\mathbb{N}}$ a random sequence of independent random variables of respective distribution ν. We write[1] $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$. We consider the sequence $(\overline{\gamma}_n)_{n\in\mathbb{N}}$ defined by $\overline{\gamma}_n := \gamma_0 \cdots \gamma_{n-1}$. We call it the random walk of step of law ν.

Our first object of interest will be the behaviour of the singular gap of $\overline{\gamma}_n$ as n goes to ∞. We define the singular gap as follows:

$$\mathrm{sqz}: \begin{array}{ccc} \Gamma\setminus\{0\} & \longrightarrow & [0,+\infty] \\ g & \longmapsto & \log\left(\dfrac{\|g\|\cdot\|g\|}{\|g\wedge g\|}\right) = \log\left(\dfrac{\mu_1(g)}{\mu_2(g)}\right). \end{array} \tag{1}$$

Here the norm on the exterior algebra is as in Definition A.37 and the family $(\mu_i(g))_{1\leq i\leq d}$ is the family of singular values as defined in A.42. Note that given $g \in \Gamma\setminus\{0\}$, we have $\mathrm{sqz}(g) = +\infty$ if and only if g has rank one. The singular gap of 0 is not well defined by (1) so we take the convention $\mathrm{sqz}(0) := +\infty$. We are also interested in the spectral gap. We define the logarithmic spectral gap of an endomorphism g as:

$$\mathrm{prox}(g) := \lim_{n\to\infty} \frac{\mathrm{sqz}(g^n)}{n} = \log\left(\frac{\rho_1(g)}{\rho_2(g)}\right). \tag{2}$$

[1] Through this paper, the symbol $\sim$ means "has distribution law" when used on a known random variable or "that has distribution law" when used to introduce a new random variable. We use the symbol $\otimes$ for the independent coupling of distributions

Here $\rho_1(g)$ and $\rho_2(g)$ stand for the absolute values of the first and second largest eigenvalues of g. Note that $\mathrm{prox}(g)$ is well defined as long as g is not nilpotent. We take the same convention as before and write $\mathrm{prox}(g) := +\infty$ when g is nilpotent.

We say that ν is proximal if we have $\mathbb{P}(\mathrm{prox}(\overline{\gamma}_n) > 0) > 0$ for at least one integer $n \in \mathbb{N}$.

We say that a probability distribution ν is irreducible if there is no proper non-trivial subspace V of E that is ν-stable *i.e,* stable by the action of the closed semi-group generated by the support of ν. We say that ν is strongly irreducible if there is no non-trivial finite union of proper subspaces of E that is ν-stable.

Using the pivotal method, we prove the following central result. Consider two probability distributions $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ over the monoid $\widetilde{\Gamma}$ of words[2] with letters in Γ. We write κ_i for the distributions of the left-to-right product of the letters of a word drawn with law $\tilde{\kappa}_i$. Write $\odot$ for the concatenation and write $\widetilde{\nu}$ for the distribution of the one-letter word whose single letter has law ν. We say that a distribution ζ on $\mathbb{R}_{\geq 0}$ dominates a distribution η when there is a constant C such that $\eta(t, +\infty) \leq C\zeta(t - C, +\infty)$.

Theorem 1.1 (Pivotal extraction)**.** *Let ν be a strongly irreducible and proximal probability distribution over $\Gamma = \mathrm{End}(E)$. Let $\rho < 1$, $\varepsilon > 0$ and $\lambda \in \mathbb{R}$. There exists a triple $(\tilde{\kappa}_s, \tilde{\kappa}_a, \tilde{\kappa}_b)$ that satisfies all of the following conditions.*

1. *We have $\tilde{\kappa}_s \odot (\tilde{\kappa}_a \odot \tilde{\kappa}_b)^{\odot \mathbb{N}} = \nu^{\otimes \mathbb{N}}$.*
2. *The lengths $L(\tilde{\kappa}_s)$ and $L(\tilde{\kappa}_b)$ have finite exponential moment and $L(\tilde{\kappa}_a)$ is a constant.*
3. *For all $i \in \{s, a, b\}$, one has $\tilde{\kappa}_i\{\gamma \in \Gamma | \mathrm{sqz}(\gamma) \geq \lambda\} = 1$.*
4. *For all pair $(i, j) \in \{(s, a), (a, b), (b, a)\}$, we have $\|gh\| \geq \varepsilon\|g\|\|h\|$ for all g in the support of κ_i and all h in the support of κ_j.*
5. *For all $g \in \Gamma$, we have $\kappa_a\{\gamma \in \Gamma \,|\, g\mathbb{A}^{\varepsilon}\gamma\} \geq 1 - \rho$ and $\kappa_a\{\gamma \in \Gamma | \gamma\mathbb{A}^{\varepsilon} g\} \geq 1 - \rho$.*
6. *For $i \in \{s, a, b\}$ and a word $\tilde{g} \sim \tilde{\kappa}_x$, conditionally to the length $L(\tilde{g})$, for all $k < L(\tilde{g})$ and for χ_k the k-th letter of $\tilde{g}$, the distribution of $N(\gamma_k)$ is dominated by $N(\nu)$.*

In other words it means that there is a way to randomly group the factors into consecutive groups of size $w_0, w_1, w_2, \ldots$ so that we have an aligned product in the sense that for g, h the products of consecutive groups, we have $\|gh\| \geq \varepsilon\|g\|\|h\|$. Moreover, the grouping is so that, knowing the size of each groups, we have some slack on each odd indexed factors to align them, the groups are independent and we have good conditional moment assumptions on the norms of the letters *i.e,* if $N(\gamma_k)$ is L^p (resp. weakly L^p , resp. exponentially integrable) by itself, then it still satisfies the same moment assumptions conditionally to the way of grouping the factors. Points (1) to (5) are used in the proofs of Theorems 1.2 and 1.3 and point (6) is more technical and is only used in the proof of Theorem 1.5.

The main result is the following.

Theorem 1.2 (Quantitative estimate of the escape speed)**.** *Let (γ_n) be a random i.i.d sequence of matrices that follow a strongly irreducible probability distribution. Write $\overline{\gamma}_n := \gamma_0 \cdots \gamma_{n-1}$ for all n. Then there exists a constant $\sigma(\nu) \in [0, +\infty]$ (with $\sigma(\nu) > 0$ if and only if ν is proximal) such that almost surely $\frac{\mathrm{sqz}(\overline{\gamma}_n)}{n} \to \sigma(\nu)$. Moreover, we have the following large deviations inequalities:*

$$\forall \alpha < \sigma(\nu), \exists C, \beta > 0, \forall n \in \mathbb{N}, \mathbb{P}(\mathrm{sqz}(\overline{\gamma}_n) \leq \alpha n) \leq C \exp(-\beta n). \tag{3}$$

$$\forall \alpha < \sigma(\nu), \exists C, \beta > 0, \forall n \in \mathbb{N}, \mathbb{P}(\mathrm{prox}(\overline{\gamma}_n) \leq \alpha n) \leq C \exp(-\beta n). \tag{4}$$

One can conjecture that $\frac{\mathrm{prox}(\overline{\gamma}_n)}{n} \to \sigma(\nu)$ almost surely. It is true in the specific setting described in Theorem 5.15.

Define $\mathbf{P}(E)$ the set of vector lines in E , that we endow with the metric

$$\mathrm{d} : (\mathbb{K}x, \mathbb{K}y) \mapsto \frac{\|x \wedge y\|}{\|x\| \cdot \|y\|}. \tag{5}$$

[2]Through this paper, we will mark with a ˜ on top the objects that are words *i.e,* tuples of "letters" that we can concatenate with each other.

Here $\|x \wedge y\|$ is the area of the parallelogram $(0, x, x+y, y)$, see appendix A.4 for more details.

A consequence of the fact that ν is not supported on $\mathrm{GL}(E)$ is that there may be some non-zero vectors $x \in E$ such that $\overline{\gamma}_n(x) = 0$ with non-zero probability for some integer n. However, we can show (see Proposition 3.31) that either $\overline{\gamma}_n = 0$ for all n larger than some random integer n_0 that has finite exponential moment or for all x outside of a countable union of proper subspaces of E , we have $\overline{\gamma}_n(x) \neq 0$ almost surely and for all $n \in \mathbb{N}$. We write $\ker(\nu)$ for the essential kernel of ν , *i.e,* the set of vectors $x \in E$ such that we have $\mathbb{P}(\overline{\gamma}_n(x) = 0) > 0$ for some integer $n \in \mathbb{N}$.

Theorem 1.3 (Quantitative convergence of the image). *Let ν be a strongly irreducible and proximal probability distribution on $\mathrm{End}(E)$. Let $\ker(\nu)$ be a σ-closed set with empty interior as in Definition 3.30. Consider a random sequence $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$. There is a random variable $u_\infty \in \mathbf{P}(E)$ that is the image of $(\gamma_n)_{n\in\mathbb{N}}$ by a shift-equivariant measurable map U_∞ and such that for all $\alpha < \sigma(\nu)$, we have some constants $C, \beta > 0$ such that:*

$$\forall u \in \mathbf{P}(E) \setminus \ker(\nu), \forall n \in \mathbb{N}, \mathbb{P}\left(\mathrm{d}(\overline{\gamma}_n u, u_\infty) \geq \exp(-\alpha n)\right) \leq C \exp(-\beta n). \tag{6}$$

The above theorem associates a distribution $\xi_\infty := U_\infty\left(\nu^{\otimes\mathbb{N}}\right)$ to any given strongly irreducible and proximal probability distribution ν on $\mathrm{End}(E)$. In terms of dynamics, the action of ν on the projective space endowed with the probability measure ξ_∞ is measure preserving (because U_∞ is shift-invariant) and ergodic and has a spectral gap on the space of Lipschitz functions in the following sense.

Corollary 1.4 (Strong irreducibility implies exponential mixing). *Let ν be any strongly irreducible and proximal distribution of positive rank on $\mathrm{End}(E)$. There is a unique ν-invariant probability distribution ξ_ν on $\mathbf{P}(E)$. Moreover, there are constants C, β such that for all distribution ξ on $\mathbf{P}(E) \setminus \ker(\nu)$ and for all Lipschitz function $f : \mathbf{P}(E) \to \mathbb{R}$ with Lipschitz constant $\lambda(f)$, we have:*

$$\forall n \in \mathbb{N}, \left| \int_{\mathbf{P}(E)} f \mathrm{d}\xi_\nu - \int_{\mathbf{P}(E)} f \mathrm{d}\nu^{*n} * \xi \right| \leq \lambda(f) C \exp(-\beta n). \tag{7}$$

In Section 5.3 we use the above results to show that under some algebraic conditions, we have an analogue of Oseledets Theorem. The interesting part is that this proof does not rely on usual ergodic theoretic tools.

In Section 5.4, we assume that ν is supported on $\mathrm{GL}(E)$. It means that for $N : \gamma \mapsto \log\|\gamma\| + \log\|\gamma^{-1}\|$, the distribution $N(\nu)$ is supported on $\mathbb{R}_{\geq 0}$.

Theorem 1.5 (Strong law of large numbers for the coefficients). *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible and proximal. There are some constants C, β such that for all $u \in E^* \setminus \{0\}$, all $v \in E \setminus \{0\}$, for all n and for $\overline{\gamma}_n \sim \nu^{*n}$, we have for all $t \in \mathbb{R}_{\geq 0}$:*

$$\mathbb{P}\left(\log \frac{\|\overline{\gamma}_n\|}{|u\overline{\gamma}_n v|} \geq t\right) \leq C\exp(-\beta n) + \sum_{k=1}^{+\infty} C \exp(-\beta k) \mathbb{P}\left(N(\gamma_0) \geq \frac{t}{k}\right). \tag{8}$$

See 5.18 for the proof. We write ζ_ν for probability distribution on $\mathbb{R}_{\geq 0}$ obtained by trunking $\sum_{k=1}^{+\infty} C \exp(-\beta k) \mathbb{P}\left(N(\gamma_0) \geq \frac{t}{k}\right.$
Then we have the following.

Corollary 1.6 (Regularity of the invariant distribution). *With the notations of Theorem 1.5 and with u_∞ as in Theorem 1.3, for all $x \in \mathbf{P}(E)$ and all $r > 0$, we have:*

$$\mathbb{P}(\mathrm{d}(u_\infty, x) < r) \leq \zeta_\nu(\log(2r), +\infty). \tag{9}$$

In particular if $N(\gamma_0)$ is weakly L^p , then we have a constant C the depends only on ν such that:

$$\mathbb{P}(\mathrm{d}(u_\infty, x) < r) \leq \frac{C}{\log(r)^p}. \tag{10}$$

See 5.23 for the proof.

In Proposition 4.5 of the article [BQ16b, p. 20], Y.Benoist and J-F.Quint have shown that if we assume that $N(\nu)$ has a finite L^p moment then for all $w \in E^*$ unitary and for u_∞ a random unitary vector in l_∞ (as in Theorem 1.3), the random variable $-\log|w(u_\infty)|$ has a finite L^{p-1} moment. A direct consequence of Theorems 1.5 and 1.3 implies that $-\log|w(l_\infty)|$ actually has a finite L^p moment and this is optimal.

Corollary 1.7 (Almost sure convergence of coefficients and central limit theorem)**.** *Let ν be a strongly irreducible and proximal probability measure supported on* $\mathrm{GL}(E)$ *and such that* $\mathbb{E}(\log \|\gamma\|)$ *and* $\mathbb{E}(\log \|\gamma^{-1}\|)$ *are both finite for $\gamma \sim \nu$. Then for all $w \in E^* \setminus \{0\}$ and all non-trivial $u \in E$, we have almost surely:*

$$\lim_{n\to\infty} \frac{\log |w\overline{\gamma}_n u|}{n} = \lim_{n\to\infty} \frac{\log \|\overline{\gamma}_n\|}{n} =: \rho_1(\nu) \in \mathbb{R}.$$

If we moreover assume that $\mathbb{E}(\log^2 \|\gamma\|)$ *and* $\mathbb{E}(\log^2 \|\gamma^{-1}\|)$ *are both finite, then we have a central limit theorem for the sequence* $(\log |w\overline{\gamma}_n u|)_n$*.*

See 5.21 and 5.22 for the proof.

1.2 Background

The study or products of random matrices bloomed with the eponym article [FK60] where Furstenberg and Kesten construct an escape speed for the logarithm of the norm using the sub-additivity. This proof was generalized by Kingmann's sub-additive ergodic Theorem [Kin68]. This article followed the works of Bellman [Bel54] who showed the almost sure convergence of coefficient for one specific example. In [FK60] Furstenberg and Kesten show a weaker version of Theorem 1.3. These works on matrices inspired the theory of measurable boundary theory for random walks on groups [Fur73]. In [BL85], Bougerol and Lacroix give an overview of the field of study with applications to quantum physics. In [GR86], Guivarc'h and Raugi show a weaker version of Theorem 1.3 in the case when ν is proximal and strongly irreducible. In [GR89] the same authors show that we have almost sure convergence of the limit flag for absolutely strongly irreducible distributions. In [GM89] Goldshied and Margulis show that the distribution ν is proximal and absolutely strongly irreducible when the Zariski support of ν is $\mathrm{SL}(E)$. In [BQ16a] Yves Benoist and Jean-François Quint give an extensive state of the art overview of the field of study with an emphasis on the algebraic properties of semi-groups. Later, in [XGL21] Xiao, Grama and Liu use [BQ16b] to show that coefficients satisfy a law of large numbers under some technical L^2 moment assumption. Similar results were proven in [Aou20] and [AG20], with some technical assumptions on the distribution. We can also mention [GQX20] and [XGL22] that give other probabilistic estimates for the distribution of the coefficients. The strong law of large numbers is already known for the norm from [AS21].

The importance of alignment of matrices was first noted in [AMS95] along with the importance of Schottky sets. Those notions were then used by Aoun in [Aou11] where he uses it to show that independent draws of an irreducible random walk that has finite exponential moment generate a free group outside of an exponentially rare event (note that the pivotal allows us to drop the finite exponential moment assumption). In [CDJ16] and [CDM17], Cuny, Dedeker, Jan and Merlevède give KMT estimates for the behaviour of $(\log \|\overline{\gamma}_n\|)_{n\in\mathbb{N}}$ under some L^p moment assumptions for $p > 2$. Note that, using Borel-Cantelli's Lemma and Theorem 1.5, we have the same estimates for the coefficients with the same hypotheses, the proof is similar to the proof of Corollary 5.22.

The main difference between these previous works and this paper is that the measure ν has to be supported on The General Linear group $\mathrm{GL}(E)$ for the above methods to work. Some work has been done to study non-invertible matrices in the specific case of matrices that have real positive coefficients. In [Fur63], Furstenberg and Kesten show limit laws for the coefficients under an L^∞ moment assumption, in [Muk87] and [KS84] Mukherjea, Kesten and Spitzer show some limit theorems for matrices with non-negative entries that are later improved by Henion in [Hen97] and more recently improved by Cuny, Dedeker and Merlevède in [CDM23].

In [GP16], Guivarc'h and Lepage show the exponential mixing property by exhibiting a spectral gap for the action of ν on the projective space under some moments assumptions on ν. The large deviations inequalities were already known for the norm in the specific case of distributions having finite exponential moment by the works of Sert [Ser18].

1.3 Method used

To prove the results, we use Markovian extractions. The idea is to adapt the following "toy model" construction to the case of matrices.

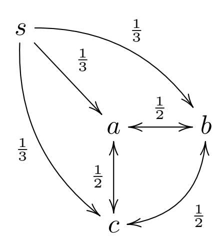
Let $G = \langle a, b, c | a^2 = b^2 = c^2 = 1 \rangle$ be the free right angle Coxeter group with 3 generators. One can see the elements of G as simple (or irreducible) words in $\{a, b, c\}$, *i.e,* finite sequences of letters without double letters. We write $\Sigma := \{a, b, c\}^{(\mathbb{N})}$ for the set of words in the alphabet $\{a, b, c\}$. We say that a word $w = (l_1, \dots, l_w) \in \Sigma$ can be reduced if it contains at least a double letter *i.e,* $l_i = l_{i+1}$ for some $i = 1, \dots, n-1$. We then say that the word $w' = (l_1, \dots, l_{i-1}, l_{i+2}, \dots, l_n)$ obtained by removing a double letter is reduced from w and write $w' \leq_r w$. Then one can show that every word in Σ can be reduced to a unique word in G in finitely many steps, so the partial order relation generated by $\leq_r$ has the elements of G as minimal elements. We write $\emptyset$ for the empty word, which is the identity element of G and of Σ. We write $\odot$ for the concatenation product on Σ and $\Pi : \Sigma \to G$ the word reduction map, note that Π is a monoid morphism $(\Sigma, \odot) \xrightarrow{\Pi} (G, \cdot)$. Indeed if we take three words $p, w, s \in \Sigma$ and $w' \leq_r w$, then $pws \leq_r pw's$. However the lift map $G \hookrightarrow \Sigma$ is not a morphism since a concatenation of two reduced word is not necessarily reduced. We also define the word length function $|\cdot| : G, \Sigma \to \mathbb{N}$ that is a monoid morphism from $(\Sigma, \odot)$ to $\mathbb{N}$ and a group norm on G.

We consider the simple random walk on the 3-tree, seen as the Cayley graph of G. Draw a random independent uniformly distributed sequence of letters $(l_n)_{n \in \mathbb{N}} \in \{a, b, c\}^{\mathbb{N}}$. Then for every $n \in \mathbb{N}$, write $g_n := l_0 \cdots l_{n-1} \in G$ for the position of the random walk at step n and $\tilde{g}_n := (l_0, \dots, l_{n-1})$ the word encoding the trajectory of the random walk up to step n. Then we know that (g_n) almost surely escapes to a point in ∂G , the set of infinite simple words. To prove it, we can show, using Markov's inequality, that $\mathbb{P}(g_n = \emptyset) \leq \left(\frac{8}{9}\right)^{n/2}$ so (g_n) visits $\emptyset$ only finitely many times, and then gets trapped in a branch (the set of simple words starting with a given letter $x_0 \in \{a, b, c\}$). Then using the same argument, (g_n) visits the first node of this branch only finitely many times and then escapes along the branch starting with $x_0 x_1$ for some $x_1 \neq x_0$ and by induction, one can show that (g_n) escapes along a branch $(x_0, x_1, x_2, \dots)$ (*i.e,* an infinite reduced word).

By symmetry, one can show that for all $k > 1$, the distribution of the letter x_k knowing $x_1, \dots, x_{k-1}$ is the uniform distribution on $\{a, b, c\} \setminus \{x_{k-1}\}$. We call pivotal times for the sequence (l_n) the times $t \in \mathbb{N}$ such that for every $k \geq t$, we have $|g_k| \geq |g_t|$ with $|\cdot|$ the reduced word length function. For example the first pivotal time t_1 is the first time after the last visit in $\emptyset$ and we use the convention $t_0 = 0$. An interesting observation is that if we write t_k for the k-th pivotal time then $x_k = l_{t_k}$.

Then instead of drawing the sequence $(l_n)_{n \in \mathbb{N}}$ of letters, we can draw the limit $(x_n)_{n \in \mathbb{N}}$ first and then the letters $(l_n)_{n \in \mathbb{N}}$ as follows.

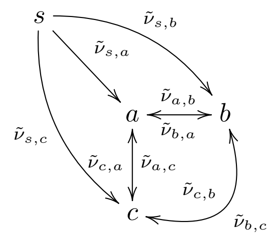
Write $X = \{a, b, c, s\}$, (s like "start") endow X with a transition kernel p such that $p(i, j) = \frac{1}{2}$ for all $i \neq j \in \{a, b, c\}$ and $p(s, i) = \frac{1}{3}$ for $i \in \{a, b, c\}$.

$(X, p) =$ 

Then take $x_0 = s$ and draw a Markov chain $(x_n)_{n \in \mathbb{N}}$ in (X, p). That means that:

$$\forall n \in \mathbb{N}, \forall l \in X, \mathbb{P}(x_n = l | x_0, \dots, x_{n-1}) = p(x_{n-1}, l).$$

Then the sequence $(x_k)_{k \geq 1}$ has the same distribution as the sequence $l_{t_k - 1}$ defined above. Moreover, the distribution of the word $(l_{t_k}, \dots, l_{t_{k+1}-1})$ only depends on x_k and x_{k+1} and not on the time $k \geq 1$. Write $\tilde{\nu}_{a,b}$ for the distribution of $(l_{t_k}, \dots, l_{t_{k+1}-1})$ knowing that $l_{t_k} = a$ and $l_{t_{k+1}} = b$ and write $\tilde{\nu}_{s,a}$ for the distribution of the word $(l_0, \dots, l_{t_1})$ knowing that $l_{t_1} = a$. Both are probability distributions on Σ. In the same fashion, we define the whole decoration:

$(X, p, \tilde{\nu}) =$ 

Then instead of drawing the (l_n) 's uniformly and independently, one can simply draw some random words $(\tilde{w}_k)$ with distribution $\bigotimes \tilde{\nu}_{x_k,x_{k+1}}$. Then for every $k \in \mathbb{N}$, $\tilde{w}_k$ has the distribution of $(l_{t_k}, \dots, l_{t_{k+1}-1})$ and the infinite word $W = \bigodot_{k=0}^{\infty} w_k \in \{a, b, c\}^{\mathbb{N}}$ has the distribution of the infinite word $L = (l_0, l_1, l_2, \dots)$.

Now, if we take a filtration $(\mathcal{F}_k)_{k\geq 0}$ such that x_k and w_{k-1} are $\mathcal{F}_k$-measurable for all $k \geq 1$, the distribution of x_{k+1} knowing $\mathcal{F}_k$ is $p(x_k, \cdot)$ and the distribution of w_k knowing $\mathcal{F}_k$ and x_{k+1} is $\tilde{\nu}_{x_k,x_{k+1}}$. Now the fact that a time t is pivotal or not is decided as soon as $w_0 \odot \cdots \odot w_{k-1}$ has length at least t. In particular the event (t is a pivotal time) is $\mathcal{F}_t$-measurable. However for $(\mathcal{C}_n)_{n\in\mathbb{N}}$ the cylinder filtration associated to the random sequence $(l_n)_{n\in\mathbb{N}}$, the event (t is a pivotal time) is never $\mathcal{C}_n$-measurable whatever the choice of $n, t \in \mathbb{N}$.

This construction gives a proof of the exponential large deviations inequalities for the random walk (g_n). This is not the simplest proof but it shows how and why we want to use the setting of Markovian extractions.

$$\exists \sigma > 0, \forall \varepsilon > 0, \exists C, \beta > 0, \forall n \in \mathbb{N}, \mathbb{P}\left(||g_n| - n\sigma| \geq \varepsilon n\right) \leq C \exp(-\beta n). \tag{11}$$

Proof. Let $(l_0, l_1, l_2, \dots) = \tilde{w}_0 \odot \tilde{w}_1 \odot \tilde{w}_2 \odot \cdots$ be as above. We can associate to every integer $n \in \mathbb{N}$ a pair of indices $k \in \mathbb{N}$, $r \in \{0, \dots, |w_k| - 1\}$ such that $n = |\tilde{w}_0| + \cdots + |\tilde{w}_{k-1}| + r$. This means that l_{n-1} is the r-th letter of $\tilde{w}_k$ and then by triangular inequality, we have $k - r \leq |g_n| \leq k + r$ because $k = |x_0 \cdots x_k|$ and $r \geq |l_{n-r} \cdots l_{n-1}|$. Then note that the lengths $(|w_k|)_{k\geq 1}$ are independent, identically distributed random variables that have finite exponential moment and $|w_0|$ has a finite exponential moment. As a consequence, if we write $\sigma := \frac{1}{\mathbb{E}(|w_1|)} = \frac{1}{3}$ and take some $\varepsilon > 0$, by the classical large deviations inequalities (see Lemma B.26 and Lemma B.29 (6)), we have: $\mathbb{P}(|k - n\sigma| \geq n\varepsilon/2) \leq C \exp(-\beta' n)$ for some $C, \beta' > 0$ and for all n. Now note that $||g_n| - n\sigma| \leq |k - n\sigma| + r$ so we have (11). □

This proof is not really useful in our case because in this case $(|g_n|)_n$ is already a martingale with bounded steps so Lemma B.26 applies. However it shows the importance of Markovian extractions.

1.4 About the pivotal method

In the second part of this article we mainly use the tools used in [Gou22], some of them having been introduced or used in former works like [BMSS20] where Adrien Boulanger, Pierre Mathieu, Cagri Sert and Alessandro Sisto state large deviations inequalities from below for random walks in discrete hyperbolic groups or [MS20] where Mathieu and Sisto show some bi-lateral large deviations inequalities in the context of distributions that have a finite exponential moment. In [Gou22] Sébastien Gouëzel uses the pivotal method in the setting of hyperbolic groups. The most interesting part of [Gou22] is the "toy model" described in section 2. From which Section 4 is inspired. In [Cho22] Inhyeok Choi applies the pivotal method to show results that are analogous to the ones presented in this article but for Γ the mapping class group of an hyperbolic surface. In [CFFT22], Chawla, Forghani, Frisch and Tiozzo use another view of the pivotal method and the results of [Gou22] to show that the Poisson boundary of random walk with finite entropy on a group that has an acylindrical action on an hyperbolic space is in fact the Gromov Boundary. I believe that similar method can be used to describe the Poisson boundary of an absolutely strongly irreducible random walk that has finite entropy 5.17.

1.5 Structure of this paper

In Section 2 of this article, we define a notion of alignment of linear maps between Euclidean vector spaces without considering any measure theoretic object. Basics in Euclidean Geometry are detailed in appendix A. The main result of this section is the heredity of the alignment in Lemma 2.15. In Section 3, we define Markov bundles *i.e,* measure theoretic objects that encode random extractions of random walks and state some basic results about them. The main result of this section is the transitivity of the extraction defined in Definition 3.22. In Section 4, we use all of this vocabulary to describe the pivotal method and to prove the main technical lemma of this paper: Theorem 4.9. Then in Section 5 we give complete proofs of Theorems 1.2, 1.3 and 5.18 using pivotal times methods and Theorem 4.9.

2 About the Cartan decomposition of rectangular matrices

In this section, we describe the geometry of the monoid $\Gamma := \mathrm{End}(E)$ for E a Euclidean space or a standard ultra-metric vector space over $\mathbb{K}$, a locally compact field. For that, we use classical results in Euclidean and ultra-metric geometry that are proven in Appendix A.

2.1 Alignment and squeezing coefficients

First we define the singular gap, it is the analogue of the reduced length of words in Section 1.3.

Definition 2.1 (Singular gap). *Let E, F be standard vector spaces and $h \in \mathrm{Hom}(E,F) \setminus \{0\}$. We define the first (logarithmic) singular gap, or squeeze of h as:*

$$\mathrm{sqz}(h) := \log\left(\frac{\mu_1(h)}{\mu_2(h)}\right) = \log\left(\max_{v_1 \in E}\min_{v_2 \perp v_1} \frac{\|h(v_1)\| \cdot \|v_2\|}{\|v_1\| \cdot \|h(v_2)\|}\right) \in [0, +\infty]$$

Where the μ_i's are the singular values as defined in Definition A.42. Given $j \geq 1$, we define the j-th singular gap of h as:

$$\mathrm{sqz}_j(h) := \log\left(\frac{\mu_j(h)}{\mu_{j+1}(h)}\right) = \mathrm{sqz}\left(\bigwedge^j h\right) \in [0, +\infty].$$

Remark 2.2. *The sum of the singular gaps is a group semi-norm on $\mathrm{GL}(E)$. Indeed, for $\gamma \in \mathrm{GL}(E)$, we have $\sum \mathrm{sqz}_i(\gamma) = \log\|\gamma\| + \log\|\gamma^{-1}\| =: N(\gamma)$. Note also that for $K := \mathrm{O}(E)$ the group of isometries of E and $\gamma \in \mathrm{End}(E)$, the equivalence class $\mathbb{K}^* K\gamma K := \{rU\gamma W \mid r \in \mathbb{K}^*, U, W \in K\}$ is completely determined by the vector $\widetilde{\mathrm{sqz}}(\gamma) := (\mathrm{sqz}_j(\gamma))_{1 \leq j < \dim(E)}$ and the rank of γ.*

In this part, we often use unitary vectors in the proofs of theorems. We write $\mathbf{S}(E) := \{x \in E \mid \|x\| = 1\}$ for the set of unitary vectors. in the proof, we often renormalize vectors, it means that we consider a unitary vector proportional to our original vector. We can do it because we made sure in Definition A.18 that all lines contain a unitary vector.

Definition 2.3 (Projective space). *Let E be a standard $\mathbb{K}$-vector space. We write $\mathbf{P}(E)$ the projective space of E i.e, the set of lines in E, endowed with the distance map:*

$$\mathrm{d}(\mathbb{K}x, \mathbb{K}y) := \frac{\|x \wedge y\|}{\|x\| \cdot \|y\|}. \tag{12}$$

Lemma 2.4. *The distance* d *is indeed a distance, moreover the normalized product:*

$$\begin{array}{rcl} \mathbf{P}(E^*) \times \mathbf{P}(E) & \longrightarrow & \mathbb{R}_{\geq 0} \\ (\mathbb{K}u, \mathbb{K}v) & \longmapsto & \frac{|uv|}{\|u\| \cdot \|v\|} \end{array}$$

is a contracting map for the distance d *on $\mathbf{P}(E^*)$ and on $\mathbf{P}(E)$.*

Proof. Up to a renormalization, one may assume that both u and v are unitary and then this is simply a reformulation of Lemmas A.45 and A.46. $\square$

Definition 2.5 (Dominant spaces)**.** *Let E, F be standard vector spaces and $h \in \mathrm{Hom}(E,F) \setminus \{0\}$. We write $V^+(h) := \{x \in E; \|h(x)\| = \mu_1(h)\|x\|\}$ and $U^+(h) := h(V^+) \subset F$ and write $W^+(h) := U^+(h^*) \subset E^*$. We use the convention $U^+(0) := F$ for convenience.*

Note that in the Euclidean case, the space $W^+(h)$ is simply the image of $V^+(h)$ by $v \mapsto v^\top$. We will see that in the ultra-metric case $U^+(h)$ is the right space to look at because its diameter in the projective space[3] is $\exp(-\mathrm{sqz}(h))$ while $V^+(h)$ has diameter at least $\rho_{\mathbb{K}}$ where $\rho_{\mathbb{K}} = \max\{|x|; x \in \mathbb{K}, |x| < 1\}$. Note also that for h an endomorphism of rank one, $U^+(h)$ is the image of h. For an intuitive understanding of this article, most reader want to think of Γ as $\mathrm{End}(\mathbb{R}^3)$. This is a good intuition because we can not directly apply what we know about hyperbolic groups and we can have some rank 2 random walks. However the dominant space of a squeezing matrix is always a line, that is not a good intuition, one should rather think of $U^+(h)$ as a bundle of lines whose thickness is proportional to $\frac{\mu_2(h)}{\mu_1(h)}$.

Lemma 2.6. *Let E, F be standard vector spaces and $h \in \mathrm{Hom}(E,F)\setminus\{0\}$. Then for every $w \in W^+(h)$, there exists a unitary vector $v \in V^+(h)$ such that $|w(v)| = 1$ and for all $x \in \ker(w)$, we have $\|h(x)\| \leq \mu_2(h)\|x\|$.*

Proof. Take $w \in W^+(h)$. By definition, there is a linear form $v' \in F^*$ unitary such that $\|v'h\| = \mu_1(h)$ and $v'h \in \mathbb{K}w$. Since $\|v'h\| = \mu_1(h)$, we may take a unitary vector $v \in E$ such that $|v'hv| = \mu_1(h)$. Since v' is unitary and $\|hv\| \geq \mu_1(h)$, we have $v \in V^+(h)$ and $|v'hv| = \|v'h\|$ so $v'h$ reaches its norm on $V^+(h)$ and w also does because they are co-linear. Now take $x \in \ker(w)$, unitary, one has $x \in \ker(v')$ and v' reaches its norm on $\mathbb{K}h(v)$ so $h(x) \perp h(v)$ so $\|h(x \wedge v)\| = \|h(x)\|\|h(v)\| \leq \mu_2(h)\mu_1(h)\|x \wedge v\|$ but $\|h(v)\| = \mu_1(h)$ and $\|x \wedge v\| = 1$ so $\|h(x)\| \leq \mu_2(h)$. ☐

Now we want to define a notion of alignment that allows us to define a notion of aligned sequences analogous to the notion of chains and chain-shadows used in [Gou22]. In the toy model of section 1.3, an aligned sequence is a sequence of non-empty reduced words whose concatenation is a geodesic, *i.e,* such the last letter of the n-th word is not equal to the first letter of the $(1+n)$-th word. The important property that differentiate random walks on a tree or on a Gromov hyperbolic group from random walks on an abelian lattice $\mathbb{Z}^d$ is that one can see whether a sequence of non-trivial words is along a quasi-geodesic line or not only by looking at local conditions. We want to construct a similar notion for products in $\mathrm{End}(E)$.

Definition 2.7 (Alignment of matrices)**.** *Let E, F, G be standard vector spaces, $f \in \mathrm{Hom}(E,F)$, $g \in \mathrm{Hom}(F,G)$ and $\varepsilon > 0$. We say that g is ε-aligned with f and write $g\mathbb{A}^\varepsilon f$ whenever there is a unitary linear form $w \in W^+(g) \subset F^*$ and a unitary vector $u \in U^+(f) \subset F$ such that $|w(u)| \geq \varepsilon$.*

Remark 2.8. *Definition 2.7 implies that any matrix is aligned with the 0 matrix and with the identity matrix.*

Proposition 2.9. *Let E, F, G be standard vector spaces, $f \in \mathrm{Hom}(E,F)$, $g \in \mathrm{Hom}(F,G)$ and $\varepsilon > 0$. Then we have $g\mathbb{A}^\varepsilon h$ if and only if $f^*\mathbb{A}^\varepsilon g^*$.*

Proof. This is obvious since $W^+(f^*) = U^+(f)$ and $U^+(g^*) = W^+(g)$ by definition. ☐

Proposition 2.10 (Alignment in term of the norm)**.** *Let E be a standard vector space, let $\varepsilon \geq 0$ and $g, h \in \mathrm{End}(E)$. If $g\mathbb{A}^\varepsilon h$ then we have*

$$\|gh\| \geq \|g\|\|h\|\varepsilon. \tag{13}$$

Moreover, if we assume (13)*, then for $\varepsilon' := \varepsilon - \exp(-\mathrm{sqz}(g)) - \exp(-\mathrm{sqz}(h))$, we have $g\mathbb{A}^{\varepsilon'}h$. In particular, when we have $\mathrm{sqz}(g), \mathrm{sqz}(h) \geq 2|\log(\varepsilon)| + 2\log(2)$ then* (13) *implies $g\mathbb{A}^{\frac{\varepsilon}{2}}h$.*

Proof. Lemma 2.10 is trivial when $g = 0$ or $h = 0$ so we may assume that $\mu_1(g)$ and $\mu_1(h)$ are positive. We first assume that we have $g\mathbb{A}^\varepsilon h$. Then, there is a unitary vector $v \in V^+(h)$ and a unitary linear form $w \in V^+(g^*)$ such that $|wghv| \geq \varepsilon\|wg\|\|hv\|$ but by definition of V^+, we have $\|wg\| = \|g\|$ and $\|hv\| = h$. This proves (13). Now assume that we do not have $g\mathbb{A}^{\varepsilon'}h$. Then write $g = g_1 + g_2$ where $g_1 := \mu_1(g)u_1^g w_1^g$

[3]By definition $U^+(h)$ and $V^+(h)$ are homogeneous and contain 0 so they are characterized by their image in the protective space

has rank 1 and $g_2 := \sum_{i\geq 2} \mu_i(g) u_i^g w_i^g$, that way $\|g_2\| = \mu_2(g)$. In the same fashion, write $h = h_1 + h_2$. Saying that we do not have $g\mathbb{A}^{\varepsilon'} h$ is equivalent to saying that $\|g_1 h_1\| < \varepsilon' \mu_1(g)\mu_1(h)$ so we have:

$$\begin{aligned}
\|gh\| &\leq \|g_1 h_1\| + \|g_2 h\| + \|h_2 g_1\| \\
&< \varepsilon' \mu_1(g)\mu_1(h) + \|g_2\|\|h\| + \|h_2\|\mu_1(g) \\
&< \left(\varepsilon - \frac{\mu_2(g)}{\mu_1(g)} - \frac{\mu_2(h)}{\mu_1(h)}\right) \|g\|\|h\| + \mu_2(g)\|h\| + \|g\|\mu_2(h) \\
&< \varepsilon \|g\|\|h\|.
\end{aligned}$$

□

Definition 2.11 (Weakly and strongly aligned sequences)**.** *Let $\mathbb{A}$ be a binary relation on* $\mathrm{End}(E)$ *and* $(g_n)_{n\in\mathbb{Z}} \in \mathrm{End}(E)^{\mathbb{Z}}$. *We say that (g_n) is $\mathbb{A}$-aligned if for every $n \in \mathbb{Z}$, we have $g_n \mathbb{A} g_{n+1}$. We say that (g_n) is strongly $\mathbb{A}$-aligned if for every $a \leq b \leq c$, we have $(g_a \cdots g_{b-1})\mathbb{A}(g_b \cdots g_{c-1})$.*

Definition 2.12 (Squeezing sequences)**.** *Let $(g_i)_{i\in I} \in \mathrm{End}(E)^I$ for I any set and $\lambda \geq 0$. We say that the family (g_i) is λ-squeezing if for all $i \in I$, $\mathrm{sqz}(g_i) \geq \lambda$.*

Lemma 2.13 (Contraction property)**.** *Let E be a Euclidean vector space or a standard ultra-metric vector space and $g, h \in \mathrm{End}(E) \setminus \{0\}$. Assume that $g\mathbb{A}^{\varepsilon} h$ for some $\varepsilon > 0$. Then one has:*

$$\mathrm{sqz}(gh) \geq \mathrm{sqz}(g) + \mathrm{sqz}(h) - 2|\log(\varepsilon)|. \tag{14}$$

Moreover, for every unitary vectors $u \in U^+(g)$, $u' \in U^+(gh)$, we have:

$$\mathrm{d}(u, u') \leq \frac{\mu_2(g)}{\varepsilon \mu_1(g)} \tag{15}$$

Proof. We use (13) in Lemma 2.10. Note also that norm of the $\wedge$ product is sub-multiplicative by definition so:

$$\mu_1(gh)\mu_2(gh) \leq \mu_1(g)\mu_2(g)\mu_1(h)\mu_2(h). \tag{16}$$

So if we do $2\log(13) - \log(16)$ we find (14).

Now to prove (15) take $y \in V^+(gh)$ and $z \in V^+(g)$ unitary, by (13), we have $\|gh(y)\| \geq \mu_1(g)\mu_1(h)\varepsilon$ and by definition of the singular values, we have $\|gh(y) \wedge g(z)\| \leq \mu_1(g)\mu_1(h)\mu_2(g)$ so:

$$\frac{\|gh(y) \wedge g(z)\|}{\|gh(y)\|\|g(z)\|} \leq \frac{\mu_1(g)\mu_1(h)\mu_2(g)}{\mu_1(g)\mu_1(g)\mu_1(h)\varepsilon} = \frac{\mu_2(g)}{\varepsilon\mu_1(g)}.$$

□

Lemma 2.14. *Let $f, g, h \in \mathrm{End}(E)$. Let $\lambda, \varepsilon_1, \varepsilon_2 > 0$. Write $\varepsilon_3 := \varepsilon_1 - \frac{\exp(-\lambda)}{\varepsilon_2}$ and assume that $\varepsilon_3 > 0$. If we assume that $f\mathbb{A}^{\varepsilon_1} g \mathbb{A}^{\varepsilon_2} h$ and $\mathrm{sqz}(g) \geq \lambda$, then we have $f\mathbb{A}^{\varepsilon_3}(gh)$. If we assume that that $f\mathbb{A}^{\varepsilon_1}(gh)$ and $\mathrm{sqz}(g) \geq \lambda$ and $g\mathbb{A}^{\varepsilon_2} h$, then $f\mathbb{A}^{\varepsilon_3} g$*

Proof. Let $u' \in U^+(gh)$ be unitary. Let $u \in U^+(g)$ and $w \in W^+(h)$ be unitary and such that $|w(u)| \geq \varepsilon_1$. By Lemma 2.13, we have $\mathrm{d}(u', u) \leq \frac{\exp(-\lambda)}{\varepsilon_2}$ so by Lemma 2.4, we have $|w(u')| \geq |w(u)| - \frac{\exp(-\lambda)}{\varepsilon_2} \geq \varepsilon_3$. □

Now is a good time to get a visual intuition of what is going on here. In Figure 1, we illustrate the first part of Lemma 2.15

When we look at hyperbolic groups in the sense of Gromov, the intuition is that everything is defined up to an additive constant δ. In our case, the intuition is that everything is defined up to a multiplicative constant. Moreover saying that two endomorphisms are aligned is useless if not at least one of them has a large squeeze coefficient. In that sense saying that $\frac{\|gh\|}{\|g\|\|h\|} \geq \varepsilon$ is up-to-a-constant-equivalent to saying that $g\mathbb{A}^{\varepsilon} h$. Now we will see that the partial products of an aligned sequence are up-to-a-constant aligned. This will allow us to treat $\mathrm{End}(E)$ like if it were a free group up-to-a-constant and use our intuition on hyperbolic groups.

Lemma 2.15 (Soft transmission of the alignment)**.** *Let E be a standard vector space and $f, g, h \in \mathrm{End}(E) \setminus \{0\}$. Let $0 < \varepsilon \leq 1$. Assume that $\mathrm{sqz}(g) \geq 2|\log(\varepsilon)| + \log(4)$. If $f\mathbb{A}^{\varepsilon} g \mathbb{A}^{\frac{\varepsilon}{2}} h$ then $f\mathbb{A}^{\frac{\varepsilon}{2}}(gh)$. Conversely, if $g\mathbb{A}^{\frac{\varepsilon}{2}} h$ and $f\mathbb{A}^{\varepsilon}(gh)$, then $f\mathbb{A}^{\frac{\varepsilon}{2}} g$.*

Let $e, f, g, h \in \mathrm{End}(E)$. Assume that $\mathrm{sqz}(g), \mathrm{sqz}(f) \geq 2|\log(\varepsilon)| + 2\log(4)$. If $e\mathbb{A}^{\frac{\varepsilon}{2}} f \mathbb{A}^{\varepsilon} g \mathbb{A}^{\frac{\varepsilon}{2}} h$, then $(ef)\mathbb{A}^{\frac{\varepsilon}{2}}(gh)$. Conversely, if $(ef)\mathbb{A}\varepsilon(gh)$ and $g\mathbb{A}^{\frac{\varepsilon}{2}} h$ and $e\mathbb{A}^{\frac{\varepsilon}{2}} f$, then $f\mathbb{A}^{\frac{\varepsilon}{2}} g$.

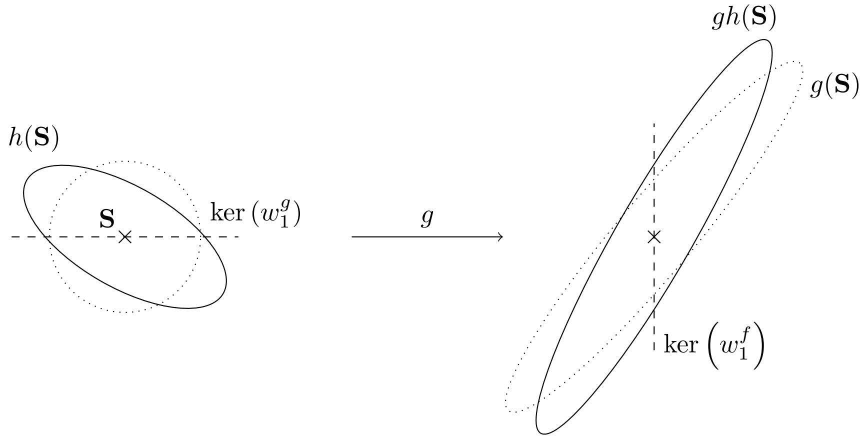


Figure 1: Illustration Of Lemma 2.14 in the case $E = \mathbb{R}^2$, with ε_2 being the sinus of the angle between the main axis of the ellipse $h(\mathbf{S})$ and the dashed line $\ker(w_1^g)$, on the left of the arrow with ε_2 being the angle between the dotted ellipse $g(\mathbf{S})$ and the dashed line $\ker(w_1^f)$ and with ε_3 being the angle between the plain ellipse $gh(\mathbf{S})$ and the dashed line. We write $\mathbf{S}$ for the unit sphere.

Proof. For the first part, we use Lemma 2.14 with $\lambda := 2|\log(\varepsilon)| + \log(4)$ and $\varepsilon_1 := \varepsilon$ and $\varepsilon_2 := \frac{\varepsilon}{2}$ Then we have $\varepsilon_3 = \frac{\varepsilon}{2}$. For the second part, we apply Lemma 2.14 twice. First to the product f, g, h with $\lambda := 2|\log(\varepsilon)| + 2\log(4)$ and $\varepsilon_1 := \varepsilon$ and $\varepsilon_2 := \varepsilon/2$ and we get $\varepsilon_3 = \frac{3\varepsilon}{4}$. Then to the product $(gh)^*, f^*, e^*$ with the same λ, ε_2 and with $\varepsilon_1 := \frac{3\varepsilon}{4}$ and we get $\varepsilon_3 = \frac{\varepsilon}{2}$. □

Note that for standard ultra-metric vector spaces Lemma 2.15 is a consequence of Lemma 2.16. Note also that it also holds trivially for endomorphisms that have rank one because U^+ would be the image.

Lemma 2.16 (Hard transmission of the alignment)**.** *Let E be a standard ultra-metric vector space and $f, g, h \in \mathrm{End}(E) \setminus \{0\}$. Assume that for some $\varepsilon > 0$ we have $\mathrm{sqz}(g) > 2|\log(\varepsilon)|$ and that $g\mathbb{A}^\varepsilon h$. Then $f\mathbb{A}^\varepsilon(gh)$ if and only if $f\mathbb{A}^\varepsilon g$.*

Proof. By (15) in Lemma 2.13, we have $\mathrm{d}(U^+(gh), U^+(g)) \leq \varepsilon^{-1}\exp(-\mathrm{sqz}(g)) < \varepsilon$ so given $w \in W^+(f)$ and $u \in U^+(g)$ and $u' \in U^+(gh)$ unitary, we have $\|u - u'\| < \varepsilon$, so since w is unitary, $|w(u) - w(u')| < \varepsilon$ but by alignment property, we may assume that $|w(u)| \geq \varepsilon$ and by ultra-metric inequality we have $|w(u')| = |w(u)| \geq \varepsilon$. This means that $f\mathbb{A}^\varepsilon(gh)$. □

A direct consequence is the following:

Corollary 2.17 (Rigidity of the alignment in the ultra-metric case)**.** *Let E be a standard ultra-metric vector space and $\varepsilon > 0$. Let (γ_n) be a sequence in $\mathrm{End}(E)$ that is $2|\log(\varepsilon)|$ squeezing. Then (f_n) is $\mathbb{A}^\varepsilon$-strongly aligned if and only if it is $\mathbb{A}^\varepsilon$-aligned.*

Proof. It is trivial that the strong alignment implies weak alignment. Now take (f_n) an $\mathbb{A}^\varepsilon$ aligned and $2|\log(\varepsilon)|$-squeezing sequence and an index $b \in \mathbb{N}$. then we show by induction on $k \in \mathbb{N}$, that for all $a \leq b \leq c$ such that $bc - b, b - a \leq k$, we have $(\gamma_a \cdots \gamma_{b-1})\mathbb{A}(\gamma_b \cdots \gamma_{c-1})$, and $(\gamma_a \cdots \gamma_{b-1}), \mathrm{sqz}(\gamma_b \cdots \gamma_{c-1}) \geq 2|\log(\varepsilon)|$. Then we have $(\gamma_b \cdots \gamma_{c-1})\mathbb{A}\gamma_c$ so $(\gamma_a \cdots \gamma_{b-1})\mathbb{A}(\gamma_b \cdots \gamma_c)$ by Lemma 2.16 and $\mathrm{sqz}(\gamma_b \cdots \gamma_c) \geq \mathrm{sqz}(\gamma_b \cdots \gamma_{c-1}) + \mathrm{sqz}(\gamma_c) - 2|\log(\varepsilon)$ by Formula (14) in Lemma 2.13. □

Note that in the toy model of section 1.3, the heredity of the alignment is even more robust. Indeed, if a sequence $(\gamma_1, \dots, \gamma_n)$ of non-empty reduced words is aligned then for all $1 \leq i \leq n$, the product $\gamma_1 \odot \cdots \odot \gamma_i$ is aligned with $\gamma_{i+1} \odot \cdots \odot \gamma_n$ because the last letter of $\gamma_1 \odot \cdots \odot \gamma_i$ is the last letter of γ_i and the first letter of $\gamma_{i+1} \odot \cdots \odot \gamma_n$ is the first letter of γ_{i+1}. In the case of matrices we use the following trick

Lemma 2.18 (Alignment of partial products)**.** *Let E be a standard metric vector space and $f, g_1, \dots, g_n, h \in \mathrm{End}(E) \setminus \{0\}$. Assume that for every $i \in \{1, \dots, n\}$ and for some $\varepsilon > 0$, we have $\mathrm{sqz}(g_i) \geq 2|\log(\varepsilon)| + \log(4)$ and $f \mathbb{A}^\varepsilon g_1 \mathbb{A}^\varepsilon \dots \mathbb{A}^\varepsilon g_n \mathbb{A}^\varepsilon h$. Then we have $f g_1 \cdots g_n h \neq 0$ and:*

$$\mathrm{sqz}(f g_1 \cdots g_n h) \geq \mathrm{sqz}(f) + \sum_{i=1}^{m} \mathrm{sqz}(g_i) + \mathrm{sqz}(h) - (n+1)(2|\log(\varepsilon)| + 2\log(2)). \tag{17}$$

If we now assume that $\mathrm{sqz}(g_i) \geq 2|\log(\varepsilon)| + 2\log(4)$ for every index $i \in \{1, \dots, n\}$. Now assume that for some specific index i, we have that $f \mathbb{A}^\varepsilon g_1 \mathbb{A}^\varepsilon \dots \mathbb{A}^\varepsilon g_i$ and $g_{i+1} \mathbb{A}^\varepsilon \dots \mathbb{A}^\varepsilon g_n \mathbb{A}^\varepsilon h$, then:

$$g_i \mathbb{A}^\varepsilon g_{i+1} \Rightarrow (f g_1 \cdots g_i) \mathbb{A}^{\frac{\varepsilon}{2}} (g_{i+1} \cdots g_n h), \tag{18}$$

$$(f g_1 \cdots g_i) \mathbb{A}^\varepsilon g_{i+1} \Rightarrow g_i \mathbb{A}^{\frac{\varepsilon}{2}} (g_{i+1} \cdots g_n h), \tag{19}$$

$$(f g_1 \cdots g_i) \mathbb{A}^\varepsilon (g_{i+1} \cdots g_n h) \Rightarrow g_i \mathbb{A}^{\frac{\varepsilon}{2}} g_{i+1}. \tag{20}$$

Proof. We show (17) by induction using Lemma 2.15 and Formula (14). We use the convention $g_0 := f$ and $g_{n+1} := h$. We will show that for every $i = 0, \dots, n$, we have:

$$\begin{cases} (f g_1 \cdots g_i) \mathbb{A}^{\frac{\varepsilon}{2}} g_{i+1} \quad \text{and} \\ \mathrm{sqz}(f g_1 \dots g_i) \geq \sum_{j=0}^{i} \mathrm{sqz}(g_j) - i(2|\log(\varepsilon)| + 2\log(2)). \end{cases} \tag{$A(i)$}$$

For $i = 0$, $f \mathbb{A}^{\frac{\varepsilon}{2}} g_1$ is a consequence of Definition 2.7 and $\mathrm{sqz}(f) \geq \mathrm{sqz}(f)$ is trivial. Then for $i = 1, \dots, n$, assume $(A(i-1))$. Apply Corollary 2.19 to $(f g_1 \cdots g_{i-1}) \mathbb{A}^{\frac{\varepsilon}{2}} g_i \mathbb{A}^\varepsilon g_{i+1}$ with $\mathrm{sqz}(g_i) \geq 2|\log(\varepsilon)| + \log(4)$ to get that $(f g_1 \cdots g_i) \mathbb{A}^{\frac{\varepsilon}{2}} g_{i+1}$. Then apply (14) to $(f g_1 \cdots g_{i-1}) \mathbb{A}^{\frac{\varepsilon}{2}} g_i$ to get that

$$\begin{aligned} \mathrm{sqz}(f g_1 \dots g_i) &\geq \mathrm{sqz}(f g_1 \dots g_{i-1}) + \mathrm{sqz}(g_i) - (2|\log(\varepsilon)| + \log(4)) \\ &\geq \sum_{j=0}^{i} \mathrm{sqz}(g_j) - i(2|\log(\varepsilon)| + \log(4)). \end{aligned}$$

This proves $(A(i))$ for all integer $i \in \{0, \dots, n\}$. Then take the transpose and change the order of factors in $(A(n-i))$ to get $g_i \mathbb{A}^{\frac{\varepsilon}{2}} (g_{i+1} \dots h)$ for all $i \leq n$. To prove (18) for $i \in \{0, n\}$, for $0 < i < n$, we simply apply Lemma 2.15 to $(f g_1 \cdots g_{i-1}) \mathbb{A}^{\frac{\varepsilon}{2}} g_i \mathbb{A}^\varepsilon g_{i+1} \mathbb{A}^{\frac{\varepsilon}{2}} (g_{i+1} \dots h)$. To prove (19) and (20), we use Lemma 2.13 to show that $\mathrm{d}(U^+(g_{i+1}), U^+(g_{i+1} \cdots g_n h) \leq \frac{\varepsilon}{4}$ and the transpose of Lemma 2.13 to show that $\mathrm{d}(W^+(g_i), W^+(g_1 \cdots g_i)) \leq \frac{\varepsilon}{4}$ and conclude using Lemma 2.4. ☐

The main result to keep in mind is the following.

Corollary 2.19 (Semi-rigidity of the alignment)**.** *Let $\varepsilon > 0$, let $\lambda \geq 2|\log(\varepsilon)| + 2\log(4)$ and let I be an interval of $\mathbb{Z}$. Consider E a standard vector space. Any sequence $(g_i)_{i \in I} \in \mathrm{End}(E)^I$ that is $\mathbb{A}^\varepsilon$-aligned and λ-squeezing (i.e, $\mathrm{sqz}(g_i) \geq \lambda$ for all i) is strongly $\mathbb{A}^{\frac{\varepsilon}{2}}$-aligned.*

Proof. This is a reformulation of the first half of Lemma 2.18. ☐

Lemma 2.20. *Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $\mathrm{End}(E)$ that is $\mathbb{A}^\varepsilon$-aligned and λ-squeezing. There is a limit u_∞ such that:*

$$\forall n \in \mathbb{N},\ \forall u \in U^+(\gamma_0 \cdots \gamma_{n-1}),\ \mathrm{d}(u, u_\infty) \leq \frac{2}{\varepsilon} \exp(-\mathrm{sqz}(\gamma_0 \cdots \gamma_{n-1}))$$

2.2 Discretisation of the alignment

Now we want to describe alignments with finite partitions that way knowing a finite number of alignment conditions for a random family of endomorphisms amounts to knowing the position of a random point in a finite set. This will be convenient for the construction of the pivotal extraction in Section 4. Indeed conditional probabilities behave much better when we condition with respect to finite algebras. Note that in the ultra-metric case we do not have anything to do due to the following result.

Proposition 2.21 (Discreteness of ultra-metric alignments)**.** *Let E be a standard ultra-metric vector space and $\varepsilon > 0$. There is a measurable partition*

$$\mathrm{End}(E) \setminus \{0\} = \bigsqcup_{i=1}^{M} L_i = \bigsqcup_{j=1}^{M} R_j$$

and a set $A \subset \{1, \ldots, M\} \times \{1, \ldots, M\}$ such that:

$$\mathbb{A}^{\varepsilon} = \bigsqcup_{i,j \in A} L_i \times R_j.$$

Proof. First by Lemma A.47, there are two partitions:

$$\mathbf{P}(E) = \bigsqcup_{j=1}^{N} S_j \quad ; \quad \mathbf{P}(E^*) = \bigsqcup_{i=1}^{N} S'_i,$$

where the S_i's and S'_i's are the balls of radius ε. Then for every pair of indices (i, j), we write $L_i := \{\gamma \in \mathrm{End}(E) | W^+(\gamma) \cap S'_i \neq \emptyset\}$ and $R_j := \{\gamma \in \mathrm{End}(E) | U^+(\gamma) \cap S_j \neq \emptyset\}$. Then write A for the set of indices (i, j) such that there are two unitary $w \in S'_i$ and $u \in S_j$ such that $|w(u)| \geq \varepsilon$. Then the quantity $|w(u)|$ is a constant for all $w \in S'_i$ and $u \in S_j$ by ultra-metric inequality and as a consequence, for $(\gamma', \gamma) \in L_i \times R_j$, we have two unitary $w \in W^+(\gamma')$ and $u \in U^+(\gamma)$ such that $|w(u)| \geq \varepsilon$ and then $\gamma' \mathbb{A}^{\varepsilon} \gamma$. Conversely, if $\gamma' \mathbb{A}^{\varepsilon} \gamma$, then there are two unitary $w \in W^+(\gamma')$ and $u \in U^+(\gamma)$ such that $|w(u)| \geq \varepsilon$ and if we write S'_i for the ball centred at w and S_j for the ball centred at u then $(\gamma', \gamma) \in L_i \times R_j$. Then to have a disjoint union, we use the following trick. Write $F = \mathcal{P}(\{1, \ldots, M\})$ the set of subsets of $\{1, \ldots, M\}$, it is a finite set. Then write $A' \subset F \times F$ the set of pairs (I, J) such that $A \cap I \times J \neq \emptyset$ and write for all $I, J \in F$:

$$L'_I := \bigcap_{i \in I} L_i \setminus \bigcup_{i' \notin I} L_{i'}$$
$$R'_J := \bigcap_{j \in J} R_j \setminus \bigcup_{j' \notin J} R_{j'}$$

Then the family $(L'_I)_{I \in F}$ is a disjoint covering of $\mathrm{End}(E)$ and so is $(R'_J)_{J \in F}$. Moreover the relations $\mathbb{A} = \bigcup_{(i,j) \in A} L_i \times R_j$ and $\mathbb{A}' = \bigcup_{(I,J) \in A'} L'_I \times R'_J$ are the same. The inclusion $\mathbb{A}' \subset \mathbb{A}$ comes from the fact that for every $(I, J) \in A'$ there is a pair $(i, j) \in A \cap I \times J$ and as a consequence $L'_I \times R'_J \subset L_i \times R_j$ and the inclusion $\mathbb{A} \subset \mathbb{A}'$ comes from the fact that if one takes a pair $g \mathbb{A} h$ and write $I := \{i | g \in L_i\}$ and $J := \{j | h \in R_j\}$ then $(g, h) \in L'_I \times R'_J$ and $I \times J$ contains a pair $(i, j) \in A$ because $g \mathbb{A} h$. □

Lemma 2.22 (Fine partition of the Euclidean projective plane)**.** *Let E be a Euclidean or Hermitian vector space and $0 < \varepsilon_1 < \varepsilon_2 < 1$. There exists a measurable partition:*

$$\mathbf{P}(E) = \bigsqcup_{i=1}^{M} S_i$$

and a set $A \subset \{1, \ldots, M\} \times \{1, \ldots, M\}$ such that:

$$\forall (i, j) \in A, \forall u \in S_i, \forall v \in S_j, |\langle u, v \rangle| \geq \varepsilon_1 \tag{21}$$
$$\forall (i, j) \notin A, \forall u \in S_i, \forall v \in S_j, |\langle u, v \rangle| < \varepsilon_2. \tag{22}$$

Proof. Let $\varepsilon = \frac{\varepsilon_2 - \varepsilon_1}{4}$ and let $(t_1, \ldots, t_M)$ be an ε-dense covering of $\mathbf{P}(E)$ for the geodesic distance on $\mathrm{S}(E)$. Define $A := \left\{ i, j \middle| |\langle t_i, t_j \rangle| \geq \frac{\varepsilon_2 + \varepsilon_1}{2} \right\}$ and define by induction $S_i := \mathcal{B}(t_i, \varepsilon) \setminus \bigcup_{j<i} S_j$. Now take two indices i, j and $x \in S_i, y \in S_j$. One has $||\langle x, t_j \rangle| - |\langle t_i, t_j \rangle|| \leq \mathrm{d}(t_i, x)$ and $||\langle x, y \rangle| - |\langle x, t_j \rangle|| \leq \mathrm{d}(t_j, y)$ by Lemma 2.4 so $||\langle x, y \rangle| - |\langle t_i, t_j \rangle|| \leq 2\varepsilon$. In conclusion, one has:

$$|\langle x, y \rangle| \geq \varepsilon_2 \Rightarrow |\langle t_i, t_j \rangle| \geq \frac{\varepsilon_2 + \varepsilon_1}{2} \Rightarrow |\langle x, y \rangle| \geq \varepsilon_1.$$
□

If we reformulate Lemma 2.22 in term of alignments, we get the following corollary. We can read it as a result about density of the set of finitely described alignments among $\{\mathbb{A}^\varepsilon | \varepsilon > 0\}$.

Definition 2.23. *Let $\mathbb{A}$ be a measurable binary relation on a measurable set Γ i.e, an $\mathcal{A}_\Gamma \otimes \mathcal{A}_\Gamma$ measurable subset of $\Gamma \times \Gamma$. We say that $\mathbb{A}$ is discrete or finitely described if there are two measurable partitions:*

$$\Gamma = \bigsqcup_{i=1}^{M} L_i = \bigsqcup_{j=1}^{M'} R_j$$

and a subset $A \subset \{1, \dots, M\} \times \{1, \dots, M'\}$ such that for all $(i,j) \in \{1, \dots, M\} \times \{1, \dots, M'\}$, all $g \in L_i$ and all $h \in R_j$, we have $g\mathbb{A}h$ if and only if $(i,j) \in A$. We then call $(L_i)_{1 \le i \le M}$ the family of left tiles of $\mathbb{A}$ and call the R_j's the right tiles of $\mathbb{A}$.

Note that in Definition 2.23, one may assume that the families (L_i) and (R_j) are the equivalence classes of the equivalence relations $\sim_L := \{(g, g') \,|\, \forall h \in \Gamma, g\mathbb{A}h \Leftrightarrow g'\mathbb{A}h\}$ and $\sim_R := \{(h, h') \,|\, \forall g \in \Gamma, g\mathbb{A}h \Leftrightarrow g\mathbb{A}h'\}$ respectively. Then saying that an alignment relation is discrete simply means that $\sim_L$ and $\sim_R$ have finitely many equivalence classes.

Corollary 2.24 (Discrete descriptions of alignment relations)**.** *Let E be a standard Archimedean vector space and $0 < \varepsilon_1 < \varepsilon_2 < 1$. There exist a discrete binary relation $\mathbb{A}$ that satisfies the inclusions $\mathbb{A}^{\varepsilon_2} \subset \mathbb{A} \subset \mathbb{A}^{\varepsilon_1}$ i.e, for any given $g, h \in \mathrm{End}(E)$, we have $g\mathbb{A}^{\varepsilon_2}h \Rightarrow g\mathbb{A}h \Rightarrow g\mathbb{A}^{\varepsilon_1}h$.*

Proof. In the Ultra-metric case, this is a consequence of Proposition2.21. Now assume that E is Euclidean. Let $\mathbf{P}(E) = \bigsqcup_{i=1}^{M} S_i$ and A be as in Lemma 2.22. For every $1 \le i, j \le M$, write $L_i := \{h \in \mathrm{End}(E) | W^+(h) \cap S_i^* \neq \emptyset\}$ and $R_j := \{h \in \mathrm{End}(E) | U^+(h) \cap S_j \neq \emptyset\}$. Then the inclusion is a direct consequence of the definition of $\mathbb{A}^\varepsilon$. Then we use the same trick as in the proof of Proposition 2.21 to get a disjoint partition. ☐

2.3 Link between singular values and eigenvalues

Definition 2.25. *Let $h \in \mathrm{End}(E)$ and $1 \le j < \dim(E)$. Let $\rho_1(h) \ge \cdots \ge \rho_d(h)$ be the absolute values of the eigenvalues of h as defined in Definition A.53. We define:*

$$\mathrm{prox}(h)_j := -\log\left(\frac{\rho_1(h)}{\rho_2(h)}\right)$$

For convenience, we will simply write $\mathrm{prox}(h)$ *for* $\mathrm{prox}_1(h)$*. We write $E^+(h)$ for the eigenspace of h associated to the eigenvalues whose absolute value is $\rho_1(h)$.*

Lemma 2.26. *Let E be a Euclidean vector space and $h \in \mathrm{End}(E)$. One has:*

$$\mathrm{prox}(h) = \lim_{n \to \infty} \frac{\mathrm{sqz}(h^n)}{n} \tag{23}$$

$$E^+(h) = \lim_{n \to \infty} U^+(h^n). \tag{24}$$

Proof. Formula (23) is a direct consequence of Definition A.53. Then up to taking the exterior product $\bigwedge^k h$, we may assume that h is proximal, then for $n \in \mathbb{N}$ large enough, $E^+(h^n)$ is a line (because a sub-vector space of dimension k in E can be identified with a line in $\bigwedge^k E$). Take $e_n \in E^+(h^n)$ unitary, then we have $h(e_n) = \lambda_1(h)e_n$, write $e_n = a_n + b_n$ with $a_n \in V^+(h^n)$ and $b_n \in \ker(w)$ for some $w \in W^+(h^n)$, then we may assume that $a_n \perp b_n$ and $h^n(a_n) \perp h^n(b_n)$ by Lemma A.41. Then we have $h^n(b_n) \le \mu_2(h^n)$ and $h^n(a_n) \le \mu_1(h^n)\|a_n\|$ so $\|a_n\| \ge \frac{|\lambda_1(h)^n| - \mu_2(h^n)}{\mu_1(h^n)}$ so by Lemma A.52, we have $\frac{\log\|a_n\|}{n} \to 0$ so $\frac{\log\|h(a_n)\|}{n} \to \log|\lambda_1(h)|$ and $\frac{\log\|h(b_n)\|}{n} \to \log|\lambda_2(h)|$. Then we have $\mathrm{d}(E^+(h), U^+(h^n)) = \mathrm{d}(\mathbb{K}e_n, \mathbb{K}h(a_n)) = \frac{\|h(b_n)\|}{\|h(e_n)\|} \to 0$, which proves (24). ☐

Corollary 2.27. *Let E be a standard vector space, let $\varepsilon > 0$ and take $\lambda > 2|\log(\varepsilon)| + 2\log(2)$. Let $h \in \mathrm{End}(E) \setminus \{0\}$ be such that $h \mathbb{A}^\varepsilon h$ and $\mathrm{sqz}(h) \geq \lambda$. Then we have:*

$$|\lambda_1(h)| \geq \mu_1(h)\frac{\varepsilon}{2} \tag{25}$$

$$\mathrm{prox}(h) \geq \mathrm{sqz}(h) - 2|\log(\varepsilon)| - 2\log(2). \tag{26}$$

Moreover $E^+(h)$ is a line and we have:

$$\forall u \in U^+(h),\ \mathrm{d}(E^+(h), \mathbb{K}u) \leq \frac{2\exp(-\lambda)}{\varepsilon}. \tag{27}$$

Proof. We have $h\mathbb{A}^\varepsilon h$, so we can apply Corollary 2.19 to a sequence $(h, \ldots, h)$ of $n+m$ copies of h, for some $n, m \in \mathbb{N}$ we get that $h^m \mathbb{A}^{\frac{\varepsilon}{2}} h^n$. Then we apply Lemma 2.13 and we have:

$$\forall m, n \in \mathbb{N}, \mathrm{sqz}(h^n) \geq \mathrm{sqz}(h^n) + \mathrm{sqz}(h^m) - 2|\log(\varepsilon)| - 2\log(2). \tag{28}$$

So going to the limit, and by Lemma 2.26, we have (26). With the same reasoning, we show (25). Now consider a unitary $x \in E^+(h)$, and $v \in V^+(h)$. Consider some vector subspace V^- that is complementary of v and such that for all $v^- \in V^-$ unitary, we have $|h(v)| \leq \mu_2(h)$ as constructed in Proposition A.43. We can write $x = av + bv'$ with $v' \in V^-$ unitary and $a, b \in \mathbb{K}$, then by orthogonality, we have $|a|, |b| \leq 1$. Then we have:

$$\begin{aligned} h(x) \wedge h(v) &= ah(v) \wedge h(v) + bh(v') \wedge h(v) \\ \|h(x) \wedge h(v)\| &\leq b\mu_1(h)\mu_2(h) \\ \frac{\|h(x) \wedge h(v)\|}{\|h(x)\| \cdot \|h(v)\|} &\leq \frac{\mu_1(h)\mu_2(h)}{|\lambda_1(h)|\mu_1(h)} \\ \mathrm{d}(\mathbb{K}h(x), \mathbb{K}h(v)) &\leq \frac{2\exp(-\lambda)}{\varepsilon} \end{aligned}$$

Moreover $\mathbb{K}h(x) = \mathbb{K}x \in \mathbf{P}(E^+(h))$ and $U^+(h) = h(V^+(h))$ by definition so for all line $u \in \mathbf{P}(U^+(h))$, we have a unitary vector $v \in V^+(h)$ such that $u = \mathbb{K}h(v)$. ☐

3 Random products and extractions

In this part we construct a theory for Markovian extractions, we extensively use the notations and results of B.

3.1 Markov bundles

To define what it means to look at a random sub-sequence of a random sequence, we define integrable extractions. First we will use the convenient language of category theory to define Markov bundles. However, one does not need any base knowledge in category theory to understand the following. We only use the formalism to compose matrices and concatenate oriented pairs of points (also called edges) to make paths. Whenever we speak of category, the example to have in mind is the following.

Definition 3.1 (Category of paths)**.** *Let X be a finite set. We define the category $\mathrm{Paths}(X)$ as the set of finite paths $(x_0, \ldots, x_n) \in X^{n+1}$, endowed with the partially defined and associative concatenation:*

$$(x_0, \ldots, x_n = y_0) \odot (y_0, \ldots, y_m) := (x_0, \ldots, x_n, y_1, \ldots, y_m).$$

There is a length functor $L : \mathrm{Paths}(X) \to \mathbb{N}; (x_0, \ldots, x_n) \mapsto n$ and we write $\mathrm{Paths}^n(X)$ the set of paths that have length n. There is also a functor:

$$\begin{array}{cccc} \theta : & \mathrm{Paths}(X) & \longrightarrow & \mathcal{E}(X) \\ & (x_0, \ldots, x_n) & \longmapsto & (x_0 : x_n) \end{array}$$

where $\mathcal{E}(X)$ is the set of oriented pairs $X \times X$ endowed with the partially defined associative product $(x:y)\cdot(y:z) := (x:z)$ we call it the category of edges of X or the trivial category of base X. We write $\mathrm{Paths}_{x-}(X)$ the set of paths that start in x, $\mathrm{Paths}_{-y}(X)$ the set of paths that end in y, and $\mathrm{Paths}_{x,y}(X) = \theta^{-1}(x,y)$ the set of paths that go from x to y. Note that when seen as sets $\mathcal{E}(X) = \mathrm{Paths}^1(X)$. Given $e \in \mathcal{E}(X)$ we write $\widetilde{e} \in \mathrm{Paths}^1(X)$ the lift of e. Note that the lift is not a functor, however it is a right inverse of θ. For every path $\tilde{\gamma} = (x_0, \dots, x_n) \in \mathrm{Paths}(X)$, we have $(x_0) \odot \tilde{\gamma} = \gamma \odot (x_n) = \tilde{\gamma}$. For every $x \in X$, we say that the trivial path (x) is the identity element of $\mathrm{Paths}(X)$ at the base point x.

Note that with our notations, the trivial path $(x) \in \mathrm{Paths}^0(X)$ is not the same as the loop $(x,x) \in \mathrm{Paths}^1(X)$. Then one can easily check that the category of edges and the category of paths are categories in the following sense.

Definition 3.2 (Category)**.**

- *We call semi-category a set Γ endowed with a partially defined associative product $\cdot$ i.e, such that for all $f,g,h \in \Gamma$, if the products $f\cdot g$ and $g \cdot h$ are both defined then $(f\cdot g)\cdot h$ and $f \cdot (g \cdot h)$ are both defined and equal.*
- *Given Σ and Γ two semi-categories, we call functor a map $\phi : \Gamma \to \Sigma$ such that for every $g,h \in \Gamma$, the product $\phi(g)\cdot\phi(h)$ is well defined when the product $g \cdot h$ is and then $\phi(g\cdot h) = \phi(g)\cdot\phi(h)$.*
- *We say that a semi-category Γ has base B if there is a canonical[4] functor $\theta = (\theta_0 : \theta_1) : \Gamma \to \mathcal{E}(B)$ such that for every g,h in Γ, the product $g\cdot h$ is defined if and only if $\theta(g)\cdot\theta(h)$ is, which means that $\theta_1(g) = \theta_0(h)$. In this case we call θ the base edge functor.*
- *We say that a semi-category Γ with base B is a category if there are some identity elements $(\mathbf{e}_x)_{x\in X}$ such that $\theta(\mathbf{e}_x) = (x:x)$ and for every $g \in \gamma$, one has $\mathbf{e}_{\theta_0(g)} \cdot g = g \cdot \mathbf{e}_{\theta_1(g)} = g$ for $\theta_0(g), \theta_1(g)$ the ends of the oriented pair $\theta(g)$.*

Definition 3.3 (Change of base)**.** *Let Γ be a category of base B and X a set together with a map[5] $\phi : X \to B$. Write Γ_X the bundle category of Γ over X:*

$$\Gamma_X = \Gamma_{X,\phi} := \Gamma \times_{\mathcal{E}(B)} \mathcal{E}(X) \simeq \{(\gamma, e) \in \Gamma \times \mathcal{E}(X) \,|\, \theta(\gamma) = \phi_:(e)\}.$$

We may also write $\Gamma = \Gamma_B$ to specify the base of Γ. We also write $\Gamma_{x,y} = \Gamma_{x-} \cap \Gamma_{-y}$ for $\theta^{-1}(x:y) = \theta_0^{-1}(x) \cap \theta_1^{-1}(y)$ respectively. For an edge $e = (x:y)$ and an element $\gamma \in \Gamma$ such that $\theta(\gamma) = \phi_:(e) = (\phi(x) : \phi(y))$, we write $(x:\gamma:y) \in \Gamma_X$ for the element represented by the pair (γ, e) and call it a decorated edge.

Lemma 3.4 (Composition of bundles)**.** *Let Γ_B be a category and $Y \xrightarrow{\phi} X \xrightarrow{\psi} B$, we have a canonical identification $\Gamma_{Y,\psi\circ\phi} \simeq (\Gamma_{X,\psi})_{Y,\phi}$.*

Proof. Saying that $\Gamma_{Y,\psi\circ\phi} \simeq (\Gamma_{X,\psi})_{Y,\phi}$ simply means that, knowing ϕ and ψ, the data $(\gamma, e) \in \Gamma_Y$ of an element $\gamma \in \Gamma$ and an edge $e \in \mathcal{E}(Y)$ such that $\psi \circ \phi(e) = \theta(\gamma)$ is the same as giving the data $(\gamma, \phi(e), e) \in (\Gamma_X)_Y$. ☐

Definition 3.5 (Step distribution on a measurable category)**.** *Let X be a finite set and Γ be a category of base X. We say that Γ is a measurable category if it is endowed with a σ-algebra $\mathcal{A}_\Gamma$ such that θ is a measurable function (for the discrete σ-algebra on X) and for every $x,y,z \in X$, the product map $\cdot : \theta^{-1}(x,y) \times \theta^{-1}(y,z) \to \theta^{-1}(x,z)$ is measurable. We call step distribution on Γ a family of probability distributions $(\nu_x)_{x\in X}$ such that ν_x is supported on $\Gamma_{x-} := \theta_0^{-1}(x)$ for all x. We write $\mathrm{Step}(\Gamma)$ the set of step distributions on Γ.*

Definition 3.6 (Convolution)**.** *Let Γ be a measurable category of finite base X. Define $* : \mathrm{Step}(\Gamma) \times \mathrm{Step}(\Gamma) \to \mathrm{Step}(\Gamma)$ as follows. Let $(\nu_x)_{x\in X}$ and $(\kappa_x)_{x\in X}$ be step distributions on Γ. Draw two families of random variables $f_x \sim \nu_x$ and $g_x \sim \kappa_x$ for all $x \in X$ such that for all $x, y \in X$, the random variables h_x and g_y are independent. For all $x \in X$, we write $(\nu * \kappa)_x$ for the distribution of the product $f_x \cdot g_{\theta_1(f_x)}$.*

Proposition 3.7. *The convolution product is associative.*

[4]This means that the data of a category contains the data of the base edge functor.
[5]Note that a map $\phi : X \to B$ induces a functor $\phi_: : \mathcal{E}(X) \to \mathcal{E}(B)$.

Proof. Let $(\nu_x)_{x\in X}$, $(\kappa_x)_{x\in X}$ and $(\eta_x)_{x\in X}$ be step distributions on a category Γ. Let $(f_x, g_y, h_z) \sim \nu_x \otimes \kappa_y \otimes \eta_z$ for all x, y, z. Then we have $(\nu * \kappa * \eta)_x = f_x \cdot g_{\theta_1(f_x)} \cdot h_{\theta_1(g_{\theta_1(f_x)})}$. ☐

Remark 3.8. *We will write $\star$ for the convolution according to the composition law $+$, to be coherent with literature and to avoid confusion with the multiplication convolution $*$. Note that since $+$ is commutative, $\star$ also is. On $\mathbb{R}$, the convolution product is given by the well known formula:*

$$\forall t \in \mathbb{R}, \kappa \star \eta(t, +\infty) = \int_{\mathbb{R}} \kappa(t-u, +\infty)\mathrm{d}\eta(u). \tag{29}$$

On a measurable category, this formula becomes:

$$\forall A \subset \Gamma, \forall x \in X, (\nu * \kappa)_x(A) = \sum_{y\in X} \int \kappa_x(g^{-1}(A))d\nu_y(g). \tag{30}$$

Here $g^{-1}(A)$ is the set of elements $h \in \Gamma$ such that $gh \in A$. That way, we have $gg^{-1}(A) \subset A$.

Definition 3.9 (Alternative definition for a Markov kernel)**.** *A Markov kernel (or stochastic matrix) over a finite set X is a step distribution over $\mathcal{E}(X)$.*

Remark 3.10 (The lift encodes the history)**.** *Given (X, p) a Markov space, write $p^n := p \overset{n \text{ times}}{* \cdots *} p \in \mathrm{Step}(X)$ the n-th power of p and $\widetilde{p} \in \mathrm{Step}(\mathrm{Paths}(X))$ the lift of p, supported on $\mathrm{Paths}^1(X)$. Then we have for all $n \in \mathbb{N}$, $p^n = \theta(\widetilde{p}^n)$ and:*

$$\forall (x_0, \ldots, x_n) \in \mathrm{Paths}^n(X), \, \widetilde{p}^n_{x_0}(x_0, \ldots, x_n) = \prod_{k=0}^{n-1} p(x_k, x_{k+1}).$$

Now we can define the notion of Markov bundle. Intuitively this is a way to draw almost i.i.d sequences. In practice it is a nice object because it allows us to prove Theorem 1.1.

Definition 3.11 (Markov bundle)**.** *Given Γ a measurable category of base B, we call Markov bundle of base X and fibre Γ the data of a map $\phi : X \to B$ and a step distribution $(\nu_x)_{x\in X}$ in the bundle category Γ_X as well as a starting set $X_s \subset X$ such that for all $x \in X$, there is a point $x_s \in X_s$ and a path from x_s to x that is ν-adapted i.e, a sequence $x_s = x_0, \ldots, x_n = x$ such that $\nu_{x_k}\{\gamma \,|\, \theta_1(\gamma) = x_{k+1}\} > 0$.*

Definition 3.12 (Notations for Markov bundles)**.** *Given (X, ν) a Markov bundle over a category Γ:*

- *write $p_\nu := \theta(\nu)$ the Markov kernel on X induced by ν,*
- *for every x, y such that $p_\nu(x, y) > 0$ write $\nu_{x,y} := \frac{(\mathbb{1}_{\theta^{-1}(x:y)})}{p_\nu(x,y)}\nu_x$ and note that it is the distribution of a random variable $g \sim \nu_x$ conditionally to $(\theta_1(g) = y)$,*
- *for every probability distribution ξ on X and every subset $A \subset X$ such that $p(\xi, A) := \sum_{x\in X, y\in A} \xi(x)p_\nu(x, y) > 0$, we write:*

$$\nu_{\xi,A} = \frac{1}{p(\xi, A)} \sum_{x\in X, y\in A} \xi(x)p_\nu(x, y)\nu_{x,y}.$$

Given $(\Omega, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ a filtered probability space, we say that a random sequence $(\gamma_n)_{n\geq 0} \in \Gamma_X^{\mathbb{N}}$ is an ornamented Markov chain in (X, ν) that respects $\mathcal{F}$ if:

- *$\theta_1(\gamma_n) = \theta_0(\gamma_{n+1}) =: x_{n+1}$ for all n,*
- *γ_n is $\mathcal{F}_{n+1}$-measurable for all n,*
- *γ_n has distribution ν_{x_n} conditionally to $\mathcal{F}_n$.*

We call $x_0 := \theta_0(\gamma_0)$ the starting state of $(\gamma_n)_{n\in\mathbb{N}}$ and call its distribution the starting distribution of (γ_n), we then write $\widetilde{\nu}^\infty_{\xi_0}$ for the distribution of $(\gamma_n)_{n\in\mathbb{N}}$ when ξ_0 is the starting distribution of γ_n. We say that (γ_n) is adapted if ξ_0 is supported on the starting set X_s.

Definition 3.13 (Integrable bundles)**.** *Let (X, ν) be a Markov bundle over $(\mathbb{R}_{\geq 0}, +)$. Given $p > 0$, we say that (X, ν) is L^p integrable if $\|\nu_x\|_{\mathrm{L}^p}$ is finite for all $x \in X$.*

Remark 3.14. *Given (X, ν) a Markov bundle over a measurable category Γ and (γ_n) an ornamented Markov chain in (X, ν), the sequence $(x_n := \theta_0(\gamma_n))_{n\in\mathbb{N}}$ is a Markov chain in (X, p_ν). We call the distribution of $x_0 = \theta_0(\gamma_0)$ the starting distribution of the ornamented Markov chain (γ_n) Moreover, just like for Markov chains, it fully determines the distribution of the whole decorated Markov chain.*

Lemma 3.15. *Let Γ be a category of finite base B. Let (γ_n) be a random adapted sequence (i.e, such that $\theta_1(\gamma_n) = \theta_0(\gamma_{n+1})$ for all n) defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that for all $n \in \mathbb{N}$, γ_n is $\mathcal{F}_{n+1}$-measurable, and the distribution of γ_n conditionally to $\mathcal{F}_n$ only depends on $\theta_0(\gamma_n)$ (which is $\mathcal{F}_n$-measurable) and not on n. Then (γ_n) is an ornamented Markov chain in a Markov bundle (X, ν).*

Proof. Given $x \in B$ and an integer n such that $\mathbb{P}(\theta_0(\gamma_n) = x) > 0$, we write $\nu_x^{(n)}$ for the distribution of γ_n knowing $\theta_0(\gamma_n) = x$. By assumption, $\nu_x^{(n)}$ does not depend on n when it is defined. Then take X_s to be the support of the distribution of $\theta_0(g_0)$ and X to be the range of $(\theta_0(\gamma_n))$ *i.e,* $X_s := \{x \in B \mid \mathbb{P}(\theta_0(\gamma_0) = x) > 0\}$ and $X := \{x \in B \mid \exists n \in \mathbb{N}, \mathbb{P}(\theta_0(\gamma_n) = x) > 0\}$. Then one can check that $X_s \subseteq X$, that ν_x is well defined on X, that all points of X are reachable from X_s following a (ν_x)-adapted path and that the distribution of γ_n knowing $\mathcal{F}_n$ is indeed $\nu_{\theta_0(\gamma_n)}$ for all n. □

3.2 Extractions

Think of the example in Section 1.3, (g_n) is the simple random walk in the 3-tree identified with the Cayley graph of G, the free Coxeter group with 3-generators, so that $g_n = l_0 \cdots l_n$ where the l_i's are chosen uniformly and independently among the three generators $\{a, b, c\}$ of G. We defined a sequence of pivotal times $0 = p_0 < p_1 < p_2 < \dots$ such that $(g_{p_n})_{n\in\mathbb{N}}$ is a half geodesic in G. The interesting thing is that we had a way to create the sequence of words $(l_{p_n}, \dots, l_{p_{n+1}-1})_{n\in\mathbb{N}}$ as the decoration of a ornamented Markov chain in a given Markov bundle $(X, p, \tilde{\nu})$. We say that $(X, \tilde{\nu})$ is an extraction of the Markov bundle $\left(\{*\}, \frac{\delta_a+\delta_b+\delta_c}{3}\right)$ with $\{*\}$ the trivial Markov space with only one state and $\frac{\delta_a+\delta_b+\delta_c}{3}$ the uniform probability measure over the generators of G. First, we need a convenient notation for decomposing a sequence into words.

Definition 3.16 (Category of words)**.** *Let Γ be a category of base X. We define the category of ornamented paths or adapted words in Γ as $\mathrm{Paths}(\Gamma) := \bigsqcup_{n\in\mathbb{N}} \mathrm{Paths}^n(\Gamma)$ where:*

$$\mathrm{Paths}^n(\Gamma) := \{(\gamma_0, \dots, \gamma_{n-1}) \in \Gamma^n \mid \forall k \in \{1, \dots, n-1\}, \theta_1(\gamma_{k-1}) = \theta_0(\gamma_k)\}$$

for all $n \geq 1$ and $\mathrm{Paths}^0(\Gamma) := X$. The base of $\mathrm{Paths}(\Gamma)$ is still X, the elements of $\mathrm{Paths}^0(\Gamma)$ are the identity elements and there is a natural coordinate by coordinate projection morphism $\theta : \mathrm{Paths}(\Gamma) \to \mathrm{Paths}(X)$. We also define a product morphism:

$$\Pi : \begin{array}{ccc} \mathrm{Paths}(\Gamma) & \longrightarrow & \Gamma \\ (\gamma_0, \dots, \gamma_{n-1}) & \longmapsto & \gamma_0 \cdots \gamma_{n-1} \end{array}$$

When there is an ambiguity, we write Π_Γ instead of Π to specify on which category and $\Pi_\odot$ to specify for which composition law the product morphism is defined. Given an adapted word $\tilde{g} = (\gamma_0, \dots, \gamma_{n-1}) \in \mathrm{Paths}(\Gamma)$, we define:

- *$L(\tilde{g}) := n$ the length of $\tilde{g}$,*
- *for all integer $k \in \{0, \dots, n\}$, and for $i \in \{0, 1\}$, write $s_k(\tilde{g}) := \theta_i(\gamma_{k-i}) \in B$ for the k-th state of $\tilde{g}$ note that the fact that the word $\tilde{g}$ is adapted means that $s_k(\tilde{g})$ does not depend on the choice of $i \in \{0, 1\}$ when $k \in \{1, \dots, n-1\}$,*
- *for $k = 0, \dots, n-1$, write $\chi_k(\tilde{g}) := \gamma_k$ for the $(k+1)$-th letter of $\tilde{g}$.*

Definition 3.17 (Lift)**.** *Let Γ be a measurable category of base X. Given $(\nu_x)_{x\in X}$ a step distribution over Γ, we write $(\tilde{\nu}_x)_{x\in X}$ for the lift of ν which is a step distribution over $\mathrm{Paths}(\Gamma)$. In the same fashion, given Σ a category of base B and $\pi : \Sigma \to \Gamma$ a morphism, we write $\widetilde{\pi} : \mathrm{Paths}(\Sigma) \to \mathrm{Paths}(\Gamma)$ for the letter by letter* *morphism defined by* $\widetilde{\pi}(\sigma_0, \ldots, \sigma_{n-1}) := (\pi(\sigma_0), \ldots, \pi(\sigma_{n-1}))$*, that way, we have the following commutative diagram:*

$$\begin{array}{ccc} \mathrm{Paths}(\Sigma) & \xrightarrow{\widetilde{\pi}} & \mathrm{Paths}(\Gamma) \\ \downarrow \Pi_\Sigma & & \downarrow \Pi_\Gamma \\ \Sigma & \xrightarrow{\pi} & \Gamma \end{array} \quad (31)$$

With the notations of Definition 3.6, we can define for all $n \in \mathbb{N}$ the n-th power of a Markov bundle (X, ν) as (X, ν^n) where $\nu^n := \nu * \cdots * \nu$. That way ν_x^n is the distribution of $\overline{\gamma}_n$ where (γ_k) is an ornamented Markov chain in (X, ν) that starts at the point x.

Definition 3.18 (Iteration of a Markov bundle)**.** *To a Markov bundle* (X, ν) *on a category* Γ*, we associate the lift* $(X, \widetilde{\nu})$ *that is a step distribution over* $\mathrm{Paths}\,(\Gamma_X)$ *and we have* $\Pi\widetilde{\nu}^n = \nu^n$ *for all* $n \in \mathbb{N}$*. We write* $\widetilde{\nu}_x^\infty$ *for the distribution of the ornamented Markov chain in* (X, ν) *starting at the point* x*.*

Definition 3.19. *Given* Γ *a category, we define* $\mathrm{Paths}^\infty(\Gamma)$ *as the set of infinite paths, or infinite adapted sequences* $(\gamma_n)_{n\in\mathbb{N}}$*, this is the set on which ornamented Markov chains are valued. Moreover, one can extend the concatenation convolution* $\odot$ *as a monoid action:*

$$\odot : \mathrm{Step}(\mathrm{Paths}(\gamma)) \times \prod_{x\in X} \mathrm{Prob}\left(\mathrm{Paths}_x^\infty(\Gamma)\right) \to \prod_{x\in X} \mathrm{Prob}\left(\mathrm{Paths}_x^\infty(\Gamma)\right).$$

It means that as step distribution is a way to add a random letter to an infinite random adapted word. Note that that way we have $\widetilde{\nu}^n \odot \widetilde{\nu}^\infty = \widetilde{\nu}^\infty$ *for all* n*, which justifies the notation. We say that* $\widetilde{\nu}^\infty$ *is* $\widetilde{\nu}$*-stable or equivariant. Write* $\mathrm{Paths}^*(\Gamma) := \mathrm{Paths}(\Gamma) \cup \mathrm{Paths}^\infty(\Gamma)$*. Then one can extend the concatenation product functor* $\Pi_\odot \mathrm{Paths}(\mathrm{Paths}(\Gamma)) \to \mathrm{Paths}(\Gamma)$ *as follows:*

$$\begin{array}{rccl} \Pi_\odot^* : & \mathrm{Paths}^*(\mathrm{Paths}(\Gamma)) & \longrightarrow & \mathrm{Paths}^*(\Gamma) \\ & (\tilde{g}_m)_{m\in\mathbb{N}} & \longmapsto & \left(L(\tilde{g}_0) + \cdots + L(\tilde{g}_{k-1}) + r \mapsto \chi_r(\tilde{g}_k)\right)_{k\in\mathbb{N}, 0\le r< L(\tilde{g}_k)}. \end{array}$$

For the definition of extractions we will use the following non-probabilistic notations.

Definition 3.20 (Convenient notation)**.** *Given* $(w_n)_{n\geq 0}$ *a sequence of integers, we write for all* $n \in \mathbb{N}$, $\overline{w}_n := w_0 + \cdots + w_{n-1}$*. In general, for* $(\gamma_n)_{n\in\mathbb{N}}$ *an adapted sequence in a category* Γ*, we write* $\overline{\gamma}_n := \gamma_0 \cdots \gamma_{n-1}$*. We also write for all* $n \in \mathbb{N}$*:*

$$\begin{aligned} \gamma_n^w &:= \gamma_{\overline{w}_n} \cdots \gamma_{\overline{w}_{n+1}-1} \in \Gamma \\ \widetilde{\gamma}_n^w &:= (\gamma_{\overline{w}_n}, \ldots, \gamma_{\overline{w}_{n+1}-1}) \in \mathrm{Paths}(\Gamma). \end{aligned}$$

We say that $\widetilde{\gamma}^w$ *is the word sequence extracted from* γ *with waiting time* w *and* $\overline{\gamma}^w$ *is the product sequence extracted from* Γ*. That way, one has formally:*

- $\Pi\left(\widetilde{\gamma}_n^w\right) = \gamma_n^w$ *for all* $n \in \mathbb{N}$
- $\Pi_\odot^*((\widetilde{\gamma}_n^w)_{n\in\mathbb{N}}) = (\gamma_n)_{n\in\mathbb{N}}$ *for* $\Pi_\odot^*$ *as in Definition 3.19,*
- $\gamma_{\overline{w}_k+r} = \chi_r\left(\widetilde{\gamma}_k^w\right)$ *as long as* $0 \le r < w_k$*,*
- $\overline{\gamma}_{\overline{w}_n} = \overline{\gamma}_n^w$ *for all* $n \in \mathbb{N}$*,*
- *for all sequence* $(v_n) \in \mathbb{N}^\mathbb{N}$*, we have* $\left(\widetilde{\gamma}^w\right)_n^v = \widetilde{\gamma}_n^{(w^v)}$ *for all* $n \in \mathbb{N}$*.*

Remark 3.21. *Note that the product* $\overline{\gamma}_{\overline{w}_n} \cdots \overline{\gamma}_{\overline{w}_n+w_n-1}$ *has no reason to be well defined so we will not need it, therefore we write* $\overline{\gamma}_n^w$ *for* $\overline{\gamma^w}_n$ *because it is clear that taking the partial product is the last operation to do in the order of priorities.*

Now we can define extractions

Definition 3.22 (Extraction)**.** *Given (X,ν) a Markov bundle over a category Γ, we call extraction of (X,ν) the data of a Markov bundle $(Y,\tilde{\kappa})$ over $\mathrm{Paths}(\Gamma_X)_Y$ with a base map $\phi : Y \to X$ such that $Y_s = \phi^{-1}(X_s)$ and a family of probability distributions $(\Phi_x)_{x\in X_s}$ on Y_s such that such that for all $x \in X_s$ and $y \in Y_s$, we have $\Phi_x(y) > 0$ if and only if $\phi(y) = x$. We moreover impose that :*

$$\widetilde{\nu}_{x_0}^{\infty} = \sum_{y\in Y} \Phi_{x_0}(y)\phi\left(\Pi_{\odot}^{\infty}\widetilde{\kappa}_y^{\infty}\right).$$

This means that for $(\tilde{g}_n)_n$ the ornamented Markov chain in $(Y,\tilde{\kappa})$ with starting distribution Φ_{x_0}, the sequence $\Pi_{\odot}^{\infty}(\tilde{g})$ is the ornamented Markov chain in (X,ν) with starting point x_0.

In other words, saying that a bundle $(Y,\tilde{\kappa})$ is an extraction of (X,ν) means that the distribution of the ornamented Markov chain $(\gamma_n)_{n\in\mathbb{N}}$ in (X,ν) and of starting distribution ξ_0 as defined in Definition 3.12 is the distribution of the letters of the ornamented Markov chain $(\tilde{g}_m)_{m\in\mathbb{N}}$ of starting distribution $\Phi(\xi_0)$ in $(Y,\tilde{\kappa})$ where we have forgotten the data of the spacing between words (represented by $\Pi_{\odot}^{\infty}$ in the formula) and have also forgotten the data of the exact position of each state y_m in the fiber $\phi^{-1}(x_n)$ for $x_n = \theta_0(\gamma_n) \in X$ and $y_m = \theta_0(\tilde{g}_m) \in Y$ and $n = \sum_{k=0}^{m-1} L(\tilde{g}_k)$.

Lemma 3.23 (Composition of extractions)**.** *Let (X,ν) be a Markov bundle. Let $(Y,\tilde{\kappa}),\phi,\Phi$ be an extraction of (X,ν) and $\left(Z,\tilde{\tilde{\eta}}\right),\psi,\Psi$ be an extraction of $(Y,\tilde{\kappa})$. Write $\tilde{\eta} := \Pi_{\odot}\tilde{\tilde{\eta}}$ for the distribution of the random word*

$$\left(z_0 : h_0^0,\dots,h_{N_0-1}^0,h_0^1,\dots,h_{N_1-1}^1,\dots,h_0^M,\dots,h_{N_1-1}^M : z_1\right) \in \mathrm{Paths}(\Gamma_X)_Z$$

when the random word of words

$$\left(z_0 : \left(y_0 : h_0^0,\dots,h_{N_0-1}^0 : y_1\right),\dots,\left(y_M : h_0^M,\dots,h_{N_1-1}^M : y_{M+1}\right) : z_1\right) \in \mathrm{Paths}(\mathrm{Paths}(\Gamma_X)_Y)_Z$$

has distribution $\tilde{\tilde{\eta}}$. Then $(Z,\tilde{\eta}),\phi\circ\psi,\Psi\circ\Phi$ is an extraction of (X,ν).

Proof. First note that $\Pi_{\odot} : \mathrm{Paths}(\mathrm{Paths}(\Gamma)_Y)_Z \to (\mathrm{Paths}(\Gamma)_Y)\,Z = \mathrm{Paths}(\Gamma)_Z$ so $\tilde{\eta}$ is indeed a Markov bundle over $\mathrm{Paths}(\Gamma)_Z$. Now we need to check that the distributions match take $(\gamma_n)_{n\in\mathbb{N}}$ an ornamented Markov chain in (X,ν) with starting distribution ξ_0, take $(\tilde{g}_m)_{m\in\mathbb{N}}$ an ornamented Markov chain in $(Y,\tilde{\kappa})$ with starting distribution $\Phi(\xi_0)$ and take $\left(\tilde{\tilde{h}}_l\right)_{l\in\mathbb{N}}$ an ornamented Markov chain in $\left(Z,\tilde{\tilde{\eta}}\right)$ with starting distribution $\Psi\circ\Phi(\xi_0)$. Write $w_m := L\left(\tilde{g}_m\right)$ for all $m\in\mathbb{N}$ and write $v_l := L\left(\tilde{\tilde{h}}_l\right)$ and $\tilde{h}_l := \Pi_{\odot}\tilde{\tilde{h}}_l$ for all $l\in\mathbb{N}$. Then one has $L\left(\tilde{h}_l\right) = w_l^v$ for all $l\in\mathbb{N}$ because L is a category functor and $\tilde{h}_l = \tilde{g}_l^v = \left(\widetilde{\gamma}^w\right)_l^v = \widetilde{\gamma}_l^{(w^v)}$, so $\Psi\circ\Phi$ is an extraction of (X,ν). □

Now note that a generic extraction is not necessarily meaningful because some basic properties are not preserved by general extractions. For example, a recurrent random walk may have an extraction that escapes to infinity at quadratic or exponential or whatever speed.

Lemma 3.24 (Meaningless extraction)**.** *Let Γ be a discrete group and ν be a probability distribution on Γ that is recurrent and irreducible. Then for all Markov bundle (Y,κ) on Γ, there is an extraction $(Y,\tilde{\kappa})$ of $(\{*\},\nu)$ such that $\Pi\tilde{\kappa} = \kappa$.*

Proof. Consider an ornamented Markov chain (g_m) in (Y,κ) and a sequence $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$ independent of (g_m). We define a sequence (w_m) of integers by induction: assume $\overline{w}_k$ to be defined for some integer $k\geq 0$ and define w_k to be the smallest integer such that $\gamma_{\overline{w}_k}\cdots\gamma_{\overline{w}_k+w_k-1} = g_k$. Note that since ν is recurrent, w_k is almost surely finite for all k. Then write $\tilde{g}_m := \widetilde{\gamma}_m^w$ for all $m\in\mathbb{N}$. Note that for all $m\in\mathbb{N}$, the distribution of $\tilde{g}_m$ knowing $y_m := \theta_0(g_m)$ is independent of $(w_k)_{k<m}$, of $(\gamma_n)_{n\leq w_k}$ and of $(g_k)_{k<m}$, so it is independent of $(\tilde{g}_k)_{k<m}$. By Lemma 3.15 $(\tilde{g_m})$ is an ornamented Markov chain on a Markov bundle $(Y,\tilde{\kappa})$. Moreover, we have $g_m = \Pi\tilde{g}_m$ for all m so $\Pi\tilde{\kappa} = \kappa$ and $\Pi_{\odot}^{\infty}(\tilde{g}_m) = (\gamma_n)$ so $(Y,\tilde{\kappa})$ is an extraction of $(\{*\},\nu)$. □

Given that we want to show some large deviations inequalities in Theorem 1.2, a relevant notion of a good extraction would be the following.

Definition 3.25 (Exponentially integrable extraction)**.** *Let $(Y, \tilde{\kappa})$ be an extraction of a Markov bundle (X, ν) on a monoid Γ. We say that the extraction $(Y, \tilde{\kappa})$ is exponentially integrable if there are constants $C, \beta > 0$ such that:*

$$\forall y \in Y, \mathbb{E}\left(\exp\left(\beta L(\tilde{\kappa}_y)\right)\right) \leq C. \tag{32}$$

Lemma 3.26 (Stability under composition)**.** *With the same notations as in Lemma 3.23, if moreover the extractions are exponentially integrable, then the composed extraction also is.*

Proof. We use the same notations as for the proof of Lemma 3.23. Let $\left(Z, \tilde{\tilde{\eta}}\right)$ be an exponentially integrable extraction of $(Y, \tilde{\kappa})$ itself being an exponentially integrable extraction of a Markov bundle (X, ν). Take $z \in Z$ a point, there is a point $z_0 \in Z_s$ that depends on z and an integer l_z such that $p_{\tilde{\tilde{\eta}}}^{l_z}(z_0, z) > 0$. Then take $(\tilde{g}_m)$ a decorated Markov chain in $(Y, \tilde{\kappa})$ starting at $\Phi(z_0)$ and write (v_n) and (z_n) for the waiting time and Markov chain such that $\left(z_n : \tilde{\tilde{g}}_n^v : z_{n+1}\right)_{n \in \mathbb{N}}$ is an ornamented Markov chain in $\left(Z, \tilde{\tilde{\eta}}\right)$. Then $L(\tilde{\eta}_z)$ is the distribution of $L\left(\tilde{g}_{\overline{v}_{l_z}}\right) + \cdots + L\left(\tilde{g}_{\overline{v}_{l_z} + v_{l_z} - 1}\right)$ knowing that $(z_{l_z} = z)$, which is a non-negligible event. Therefore, we can apply Lemma B.24 to the sequence $(L(\tilde{g}_m))_{\overline{v}_{l_z} \leq m < \overline{v}_{l_z} + v_{l_z} - 1}$ where the random integers $\overline{v}_{l_z}$ has finite exponential moment and the $(L(\tilde{g}_m))$'s all have bounded relative exponential moment by assumption. Doing this gives that $L(\tilde{\eta}_z)$ has a finite exponential moment. ☐

Lemma 3.27 (Motivation for Definition 3.22)**.** *Let ν be a probability distribution on $\mathbb{R}$ whose support has a lower bound $-M$. Then the following assertions are equivalent:*

1. *There is an extraction $(Y, \tilde{\kappa})$ of $(\{*\}, \nu)$ that is exponentially integrable and such that κ is supported on $[1, +\infty)$.*
2. *There are constants $\alpha > 0$, $\beta > 0$ and C such that for $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$:*

$$\mathbb{P}(\overline{\gamma}_n \leq \alpha n) \leq C \exp(-\beta n).$$

Proof. Consider a sequence $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$. We first show the interesting[6] implication, namely that 1 implies 2. Consider a sequence (w_m) such that $(\tilde{\gamma}_m^w)$ is an ornamented Markov chain in $(Y, \tilde{\kappa})$. Then since Y is finite, by Lemma B.26, there is a constant A and some $C, \beta > 0$ such that for all $k \in \mathbb{N}$, we have $\mathbb{P}(\overline{w}_k \geq Ak) \leq C \exp(-\beta k)$. By taking the sum and writing $C' := \frac{C}{\beta}$, we get $\mathbb{P}(\exists k \geq m, \overline{w}_k \geq Ak) \leq C' \exp(-\beta m)$ for all $m \in \mathbb{N}$. Now for all $k \in \mathbb{N}$, write r_k for the largest integer such that $\overline{w}_{r_k} \leq k$ and write $\beta' := \frac{\beta}{A}$. Then for all $n \in \mathbb{N}$ we have $\mathbb{P}(\exists k \geq n, Ar_k \leq k) \leq C' \exp(-\beta' n)$. Now for all $k \in \mathbb{N}$, we have $\overline{\gamma}_k \geq r_k - M(k - \overline{w}_{r_k})$. The by Lemma B.23, the random variable $k - \overline{w}_{r_k}$ has a bounded exponential moment, this means that we have some constants $C'', \beta'' > 0$ such that $\mathbb{P}(k - \overline{w}_{r_k} \geq \varepsilon k) \leq C'' \exp(-\beta'' \varepsilon k)$ for all $k \in \mathbb{N}$ and all $\varepsilon > 0$. Now take $\varepsilon := \frac{1}{2MA}$ and we get

$$\begin{aligned}
\mathbb{P}(\exists k \geq n, 2A\overline{\gamma}_k \leq k) &\leq \mathbb{P}(\exists k \geq n, Ar_k \leq k) + \mathbb{P}(\exists k \geq n, k - \overline{w}_{r_k} \geq \varepsilon k) \\
&\leq C' \exp(-\beta' n) + \frac{C''}{\varepsilon \beta''} \exp(-\beta'' \varepsilon n) \\
&\leq \left(C' + \frac{C''}{\varepsilon \beta''}\right) \exp\left(-\min\{\beta', \beta'' \varepsilon\} n\right).
\end{aligned}$$

We now prove that 2 implies 1. Consider some constants $\alpha > 0$, $\beta > 0$ and C such that $\mathbb{P}(\overline{\gamma}_n \leq \alpha n) \leq C \exp(-\beta n)$ for all n. We define a sequence $(w_k)_{k \in \mathbb{N}}$ by induction. Consider some integer $k \geq 0$ and assume $\overline{w}_k := w_0 + \cdots w_{k-1}$ to be defined and define w_k to be the minimal integer such that $\gamma_k^w := \gamma_{\overline{w}_k} + \cdots \gamma_{\overline{w}_k + w_k - 1} \geq 1$. Then $\overline{w}_k$ is a stopping time for the cylinder filtration $(\mathcal{F}_n)$ associated to the sequence (γ_n) and w_k only depends on the sequence $(\gamma_{n + \overline{w}_k})$ that has law $\nu^{\otimes \mathbb{N}}$ relatively to $\mathcal{F}_{\overline{w}_k}$ so the sequence $(\tilde{\gamma}_k^w)$ is i.i.d, write $\tilde{\kappa}^{\otimes \mathbb{N}}$ for its distribution. Now we only need to show that w_0 has finite exponential moment, consider $n \geq \frac{1}{\alpha}$ an integer, we have:

$$\begin{aligned}
\mathbb{P}(w_0 > n) &\leq \mathbb{P}(\overline{\gamma}_n < 1) \\
&\leq \mathbb{P}(\overline{\gamma}_n \leq \alpha n) \\
&\leq C \exp(-\beta n).
\end{aligned}$$

[6]This part of the proof is interesting because it is similar to the proof of Theorem 1.2.

This means that w_0 has a finite exponential moment and therefore $(\{*\}, \tilde{\kappa})$ is an exponentially integrable extraction of ν. □

3.3 Rank, kernel and boundary of a probability distribution

Now we will consider a Markov bundle with trivial basis $(\{*\}, \nu)$ over $\Gamma = \mathrm{End}(E)$ for E a standard vector space. This amounts to considering an i.i.d sequence of random matrices.

Definition 3.28 (Rank of a distribution)**.** *Let ν be a step distribution on $\mathrm{End}(E)$. We define the rank of ν as the largest integer $\mathbf{rk}(\nu)$ such that:*

$$\forall n \geq 0, \nu^{*n}\left\{\gamma \in \Gamma \mid \mathbf{rk}(\gamma) < \mathbf{rk}(\nu)\right\} = 0. \tag{33}$$

Lemma 3.29 (Eventual rank of a distribution)**.** *Let ν be a probability distribution on Γ. There exists an exponentially integrable extraction $(\{*\}, \tilde{\kappa})$ of ν such that the product $\kappa = \Pi\tilde{\kappa}$ is supported on the set of rank $\mathbf{rk}(\nu)$ endomorphisms. In other word, for $(\gamma_n) \sim \nu^{\otimes n}$, there is a stopping time w such that $\overline{\gamma}_w$ has rank $\mathbf{rk}(\nu)$ and w has finite exponential moment.*

Proof. Let (γ_n) be a sequence of independent random matrices of distribution law ν defined over a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $(\overline{\gamma}_n)$ be the right partial product associated to γ. The random sequence $\mathbf{rk}(\overline{\gamma}_n)$ is a non-increasing sequence of non-negative integers, so it converges to a random limit r_0. We want to show that r_0 is almost surely constant. Define w_0 as the first integer such that $\mathbf{rk}(\overline{\gamma}_n) = r_0$, then write r_1 for the limit rank of $\gamma_{w_0} \cdots \gamma_{n-1}$ as n goes to infinity. Since w_0 is a stopping time, the distribution of r_1 is the same as the distribution of r_0 and they are independent but $r_0 \leq r_1$ almost surely so r_0 is constant equal to $\mathbf{rk}(\nu)$. Then we write w_1 for the smallest integer such that $\gamma_{w_0} \cdots \gamma_{w_0+w_1-1}$ has rank $\mathbf{rk}(\nu)$ and define a sequence (w_n) by induction such that γ_n^w has rank $\mathbf{rk}(\nu)$ and w_{n+1} is taken minimal for this property. Then $\overline{w}_n$ is a stopping time so the sequence $(\tilde{\gamma}_n^w)$ is i.i.d and as a consequence is an ornamented Markov chain for a Markov bundle of trivial base that we call $(\{*\}, \tilde{\kappa})$. Now to show that w_0 has finite exponential moment, we consider n_0 the smallest integer such that $\mathbb{P}(w_0 = n_0) > 0$. It means that $\mathbb{P}(\overline{\gamma}_{n_0} = 0) = \alpha_0 > 0$. Then since (γ_n) is i.i.d, we have $\mathbb{P}(\gamma_{kn_0} \cdots \gamma_{(k+1)n_0-1} = 0) = \alpha_0$ for all $k \in \mathbb{N}$ and these events are independent so for all $k \in \mathbb{N}$, we have:

$$\begin{aligned}
\mathbb{P}(\forall k' < k, \gamma_{k'n_0} \cdots \gamma_{(k'+1)n_0-1} \neq 0) &= (1-\alpha_0)^k \\
\mathbb{P}(\gamma_0 \cdots \gamma_{kn_0-1} \neq 0) &\leq (1-\alpha_0)^k \\
\forall n \geq kn_0,\, \mathbb{P}(\overline{\gamma}_n \neq 0) &\leq \sqrt[n_0]{1-\alpha_0}^{kn_0} \\
\forall kn_0 \leq n < (k+1)n_0,\, \mathbb{P}(w_0 \geq n) &\leq \frac{1}{1-\alpha_0} \sqrt[n_0]{1-\alpha_0}^{n} \\
\forall n \in \mathbb{N},\, \mathbb{P}(w_0 \geq n) &\leq C \exp(-\beta n).
\end{aligned}$$

For $C = \frac{1}{1-\alpha_0}$ and $\beta = |\log(1-\alpha_0)|/n_0 > 0$. □

Definition 3.30 (Essential kernel)**.** *Let ν be a probability distribution on $\mathrm{End}(E)$. We write $\ker(\nu)$ for the essential kernel of ν defined by:*

$$\ker(\nu) := \left\{x \in E \,|\, \exists n \geq 0, \nu^{*n}\{\gamma | \gamma x = 0\} \neq 0\right\}. \tag{34}$$

Proposition 3.31. *Let ν be a probability distribution on $\mathrm{End}(E)$. The set $\ker(\nu)$ is included in a finite union of subset of E that each have dimension at most $\dim(E) - \mathbf{rk}(\nu)$.*

Proof. Consider a sequence $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$ and write w_0 for the first integer such that $g := \gamma_{w_0-1} \cdots \gamma_0$ has rank $\mathbf{rk}(\nu)$. Then for all x, we have $g(x) = 0$ if and only if there is an integer n such that $\gamma_{n-1} \cdots \gamma_n(x) = 0$. So we have:

$$\begin{aligned}
\ker(\nu) &= \{x \in E \,|\, \mathbb{P}(g(x) = 0) > 0\} \\
&= \bigcup_{\varepsilon > 0} \{x \in E \,|\, \mathbb{P}(g(x) = 0) \geq \varepsilon\} \\
&= \bigcup_{n \in \mathbb{N}} \{x \in E | \nu(\gamma x = 0) \geq 2^{-n}\}.
\end{aligned}$$

Now given some $\varepsilon > 0$, we will show that $K_\varepsilon := \{x \in E \,|\, \mathbb{P}(g(x) = 0) \geq \varepsilon\}$ is included in a finite union of subspaces of dimension $d' = \dim(E) - \mathbf{rk}(\nu)$. Consider a family $(x_1, \dots, x_N)$ in general position in K_ε in the sense that for all $\leq i_1 < \cdots < i_k \leq N$ with $k \leq d' + 1$, the space $\bigoplus_{j=1}^k \mathbb{K}x_{i_j}$ has dimension exactly k. Then for all $i \in \{1, \dots, N\}$, we define a_i to be 1 if $g(x_i) = 0$ and 0 otherwise. Then since $\ker(g)$ has dimension at most d' almost surely, we have $\sum_{i=1}^N a_i \leq d'$ almost surely. So if we loot at the expectation, we get $d' \geq N\varepsilon$. Now take N_ε to be maximal so that there is a family $(x_1, \dots, x_N)$ in general position in K_ε. Then for all $x_0 \in K_\varepsilon$, the family $(x_0, \dots, x_N)$ is not in general position so there is some $k \leq d'$ and there are indices $0 \leq i_0 < \cdots < i_k < N$ such that $\bigoplus_{j=0}^k \mathbb{K}x_{i_j}$ has dimension at most k. We also know that $i_0 = 0$ because $(x_1, \dots, x_N)$ is in general position. As a consequence $x_0 \in \bigoplus_{j=1}^k \mathbb{K}x_{i_j}$. Therefore, we have:

$$K_\varepsilon \subset \bigcup_{1 \leq i_1 \leq \cdots \leq i_{d'} \leq N} \left(\bigoplus_{j=1}^k \mathbb{K}x_{i_j} \right).$$

So $K_{2^{-n}}$ is included in a union of at most $\binom{d'2^n}{d'}$ subspaces of dimension d'. Now that is true for all $n \in \mathbb{N}$. Therefore $\ker(\nu)$ is included in a union of countably many subspaces of dimension d'. □

Proposition 3.32. *Let ν be a probability distribution over $\mathrm{End}(E)$ and (γ_n) be a random sequence of distribution $\nu^{\otimes \mathbb{N}}$. Then for every $x \in E$, the random sequence $\overline{\gamma}_n x$ is almost surely never zero if and only if $x \notin \ker(\nu)$.*

Proof. First note that if $x \notin \ker(\nu)$, then the event $(\exists n \in \mathbb{N}, \overline{\gamma}_n(x) = 0)$ is a countable union of negligible events so it is negligible. Then if $x \in \ker(\nu)$, then there is an integer k such that $\nu^{*k}\{g \in \Gamma \,|\, g(x) = 0\} = \varepsilon > 0$ so we have $\mathbb{P}(\gamma_{kn} \cdots \gamma_{k(n+1)-1}(x) = 0) = \varepsilon > 0$ for all n and these events are independent so we have $\mathbb{P}(\exists n, \overline{\gamma}_n(x) = 0) \geq \mathbb{P}(\exists n, \gamma_{kn} \cdots \gamma_{k(n+1)-1}(x) = 0) = 1$ □

Definition 3.33. *For a given probability distribution ν on $\mathrm{End}(E)$, we define $\bigwedge^j \nu$ to be the distribution law of $\bigwedge^j \gamma$ when $\gamma \sim \nu$. It is a distribution on $\mathrm{End}\left(\bigwedge^j E\right)$.*

Definition 3.34 (Irreducible distributions)**.** *Let E be a standard vector space of dimension $d \geq 2$ and ν be a probability distribution over $\mathrm{End}(E)$. We say that ν is irreducible if there is no proper subspace $V \subset E$ such that $\nu(\mathbf{stab}(V)) = 1$. We say that ν is strongly irreducible if there is no non-trivial finite union of proper subspaces $A = \bigcup V_i$ such that $\nu(\mathbf{stab}(A)) = 1$. We say that ν is absolutely strongly irreducible if for every $1 \leq j \leq d-1$, the distribution $\bigwedge^j \nu$ is strongly irreducible.*

Definition 3.35 (Proximal distributions)**.** *Let ν be a probability distribution on Γ and $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$. We say that ν is proximal if $\mathbf{rk}(\nu) \neq 0$ and there exists an integer n such that the partial product $\overline{\gamma}_n$ is proximal with non-zero probability.*

Definition 3.36 (Proximal boundary)**.** *Let E be a standard vector space and ν be a probability distribution over $\mathrm{End}(E)$ of non-zero rank.*

- *We define $\mathbf{rg}(\nu)$ the range of ν as the set $\bigcup_{n \in \mathbb{N}} \mathbf{supp}\,(\nu^{*n})$.*
- *We define the conical range $\mathbf{C}(\nu)$ as the closure of $\mathbb{K}\mathbf{rg}(\nu)$.*
- *We define the rank r cone of ν as the intersection between $\mathbf{C}(\nu)$ and the set of rank r matrices and write $\mathbf{C}^r(\nu)$ for the rank r cone.*
- *We write $\partial^r(\nu) \subset \mathbf{P}(\mathrm{End}(E))$ for the quotient of $\mathbf{C}^r(\nu) \setminus \{0\}$ by the multiplicative action of $\mathbb{K} \setminus \{0\}$.*
- *We call proximal boundary of ν the set $\partial(\nu) := \partial^1(\nu)$.*
- *We define the Furstenberg boundaries of ν as:*

$$\partial_u(\nu) := \{\mathrm{im}(\pi) \,|\, \mathbb{K}\pi \in \partial(\nu)\} \subset \mathbf{P}(E) \tag{35}$$

$$\partial_w(\nu) := \{\mathrm{im}(\pi^*) \,|\, \mathbb{K}\pi \in \partial(\nu)\} \subset \mathbf{P}(E^*). \tag{36}$$

A direct consequence of Lemma 3.38 is that the proximal boundary of a given probability distribution ν is non-empty if and only if ν is proximal. Remark also that by definition, the Green function is monotonous for the extraction order and so is the boundary *i.e,* if a given distribution ν' is extracted from a distribution ν, then $\partial^r(\nu') \subset \partial^r(\nu)$ for all r.

Lemma 3.37. *With the notations of Definition 3.36. If we moreover assume that ν is irreducible then we can decompose $\partial^1(\nu)$ as a product:*

$$\partial(\nu) = \{\mathbb{K}ef \mid \mathbb{K}e \in \partial_u(\nu), \mathbb{K}f \in \partial_w(\nu)\}\,. \tag{37}$$

Proof. Let $e_1, e_2 \in E$ and $f_1, f_2 \in \mathbf{P}(E^*)$ be unitary. Let $\pi_1 := e_1 f_1$ and $\pi_2 := e_2 f_2$. We want to show that $e_1 f_2 \in \partial(\nu)$. Take two sequences $(g_m), (h_m) \in \mathbb{K}\mathbf{rg}(\nu)^{\mathbb{N}}$ such that $g_m \to \pi_1$ and $h_m \to \pi_2$. Since ν is irreducible, there exists an endomorphism $\gamma \in \mathbf{supp}(\nu^{*k})$ for some $k \geq 0$ such that $\gamma e_2 \notin \ker(f_1)$. Consider such a γ, then by continuity of the product, we have $g_m \gamma h_m \to e_1 f_1 \gamma e_2 f_2$ and $f_1 \gamma e_2$ is a non-zero scalar so $\lim_m g_m \gamma h_m \in \mathbb{K}^* e_1 f_2$. Now if we assume that for all m, there are integers i, j such that $g_m \in \mathbb{K}\mathbf{supp}(\nu^{*i})$ and $h_m \in \mathbb{K}\mathbf{supp}(\nu^{*j})$ then $g_m \gamma h_m \in \mathbf{supp}(\nu^{*i+j+k})$ so $\mathbb{K}e_1 f_2 \in \partial^1(\nu)$ which concludes the proof. ☐

Lemma 3.38 (Characterisation of proximality)**.** *Let ν be an irreducible probability distribution of positive rank. Then the following assertions are equivalent:*

1. *ν is not proximal,*
2. *there is a bound B such that $\mathrm{sqz}(\gamma) \leq B$ for every $\gamma \in \mathbf{rg}(\nu)$,*
3. *$\partial(\nu) = \emptyset$.*

Proof. We show $3 \Leftrightarrow 2$. Note that sqz is a continuous function from $\mathrm{End}(E) \setminus \{0\}$ to $[0, +\infty]$ and it is invariant by scalar multiplication so it is continuous on $\mathbf{P}(\mathrm{End}(E))$ which is compact. Moreover $\mathrm{sqz} = +\infty$ only on the set of rank one matrices so saying that $\mathrm{sqz}(\gamma)$ is bounded on $\mathbf{rg}(\nu)$ is equivalent to saying that $\partial(\nu) = \emptyset$.

We show $1 \Rightarrow 3$. Assume that ν is not proximal. We want to show that $\partial(\nu) = \emptyset$. Take $e \in E$ and $f \in E^*$ unitary and write $h := ef$, then one has $\lambda_1(h) = f(e)$. Now assume that $\mathbb{K}h \in \partial(\nu)$ and that $\lambda_1(h) \neq 0$. Then consider a sequence $g_n \in \Gamma$ such that $g_n \to h$. Write $\varepsilon := |\lambda_1(h)|/2$ and $\lambda := 2|\log(\varepsilon)| + 4\log(2)$. Then for n large enough, we may assume that $\mathrm{sqz}(g_n) \geq \lambda$ and that $g_n \mathbb{A}^{\varepsilon} g_n$. Then by Corollary 2.27, we have $\mathrm{prox}(g_n) > 0$ which is a contradiction. This means that $h^2 = 0$ so we have $e \in \ker(f)$ and this is true for all e such that $\mathbb{K}e \in \partial_u(\nu)$ which contradicts the irreducibility of ν so $\partial^1(\nu)$ is empty.

Now we show $2 \Rightarrow 1$. Assume that ν is proximal. Then there is an integer k and an element $\gamma \in \mathbf{supp}(\nu^{*k})$ such that $\mathrm{prox}(\gamma) > 0$. Then by the spectral theorem, we have $\mathrm{prox}(\gamma) = \lim \frac{\mathrm{sqz}(\gamma^n)}{n}$ so we have $\mathrm{sqz}(\gamma^n) \to \infty$. Moreover $\gamma^n \in \mathbf{supp}(\nu^{*kn}) \subset \mathbf{rg}(\nu)$ for all n so sqz is not bounded on $\mathbf{rg}(\nu)$. ☐

Note that Theorem 1.2 is not true for a general irreducible and proximal probability distribution. A simple couter can be created by taking $E := \mathbb{R}^2$ and $\nu := \frac{1}{2}(\delta_A + \delta_B)$ with:

$$A := \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Then ν is irreducible because the action of B alone is. However the group generated by A and B is isomorphic to $\mathbb{Z}/4 \ltimes \mathbb{Z}$ (because $B^4 = \mathrm{Id}$ and $BAB^{-1} = A^{-1}$) and one can see that the random walk $(\bar{\gamma}_n)$ associated to ν is in fact recurrent. Take $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$. Write w_0 for the first time such that $\gamma_{w_0} = B$ and then take (w_k) such that $\overline{w}_k$ is the k-th time such that $\gamma_{\overline{w}_k} = B$. Then for all integer $k \geq 0$, we have $\gamma^w_{2k}\gamma^w_{2k+1} = -A^{n_k}$ with n_k a symmetric $\mathbb{Z}$-valued random variable that has finite exponential moment. Then we know from basic probability theory that the random walk $\overline{n}_k$ is recurrent.

4 Pivotal time method

4.1 Statement of the result

In this section, we want to show that every strongly irreducible and proximal Markov bundle has a pivotal extraction *i.e,* an extraction that is convenient and allows us to have nice lower bounds on the singular and spectral gaps using the aligned sequences that we described in section 2. All the statements of the Lemmas mention an abstract measurable function N. This is only used in the proof of Theorem 1.5, for the proofs of Theorems 1.2 and 1.3, we may take N to be the constant function equal to 0 and make all the conditions on N trivial.

Definition 4.1. *We say that a Markov bundle (X, ν) over Γ is $\mathbb{A}$-aligned for a given binary relation $\mathbb{A}$ if for every p_ν-admissible path (x, y, z), we have $\nu_{x,y} \otimes \nu_{y,z}(\mathbb{A}) = 1$. We say that (X, ν) is strongly $\mathbb{A}$-aligned if for every $x_0 \in X$, the ornamented Markov chain $(\gamma_n) \sim \widetilde{\nu}_{x_0}^{\infty}$ almost surely projects to a strongly $\mathbb{A}$-aligned sequence in Γ.*

Proposition 4.2. *Given (X, ν) a Markov bundle that is strongly $\mathbb{A}$-aligned for some binary relation $\mathbb{A}$ (such that the identity is $\mathbb{A}$-aligned with everyone) and $(Y, \tilde{\kappa})$ an extraction of (X, ν), then $(Y, \tilde{\kappa})$ is strongly $\mathbb{A}$-aligned.*

Proof. Take $(\gamma_n)_{n\in\mathbb{N}}$ a strongly aligned sequence, then take (w_n) a sequence of non-negative integers. Then take $a \leq b \leq c$ three natural integers, we have $\overline{w}_a \leq \overline{w}_b \leq \overline{w}_c$ so $(\gamma_{\overline{w}_a} \cdots \gamma_{\overline{w}_b - 1}) \,\mathbb{A}\, (\gamma_{\overline{w}_b} \cdots \gamma_{\overline{w}_c - 1})$ because (γ_n) is strongly aligned, and if we rewrite this using the notation of Definition 3.20, we have $\left(\gamma_a^w \cdots \gamma_{b-1}^w\right) \mathbb{A} \left(\gamma_b^w \cdots \gamma_{c-1}^w\right)$so γ^w is strongly aligned. Then up to equivalence, one may assume that for all $y_0 \in Y$ the ornamented Markov chain $(\tilde{g}_n)$ in $(Y, \tilde{\kappa})$ starting from y_0 can be written as $(\widetilde{\gamma}_n^w)$ for $(\gamma_n) = (\gamma'_{n_0+n})$ with n_0 a random integer and (γ'_k) an ornamented Markov chain in (X, ν). □

Definition 4.3. *We say that a Markov bundle (X, ν) over Γ is λ-squeezing for a given constant λ if for every $x \in X$, we have $\nu_x\{\gamma | \mathrm{sqz}(\gamma) \geq \lambda = 1\}$.*

Proposition 4.4. *Let $\varepsilon > 0$ and $\lambda \geq 2|\log(\varepsilon)|$. Then every Markov bundle (X, ν) that is $\mathbb{A}^{2\varepsilon}$ aligned and λ-proximal is $\mathbb{A}^{\varepsilon}$-strongly aligned.*

In this part we want to control the L^p norms of the extractions, however e do not want to assume that ν is L^p for some $p \geq 1$ so we will consider an abstract function N that may be the one defined in the introduction or that may be 0 in the case when ν does not satisfy any moment condition.

Definition 4.5 (Word-norm)**.** *Let Γ be a measurable category and $N : \Gamma \to [0, +\infty]$. We write $\overline{N}$: $\mathrm{Paths}(\Gamma) \to [0, +\infty]$ for the measurable map:*

$$\overline{N} : (\gamma_1, \ldots, \gamma_L) \longmapsto N(\gamma_1) + \cdots + N(\gamma_L).$$

Note that the map N does not need to be a functor but $\overline{N}$ always is. Note also that saying that N is sub-additive is equivalent to saying that $N \circ \Pi \leq \overline{N}$.

Definition 4.6 (Pivotal bundle)**.** *We say that a step distribution ν_s over Γ is ρ-Schottky for the alignment relation $\mathbb{A}$ if*

1. *for all $g \in \Gamma$, we have $\nu_s\{\gamma | g\mathbb{A}\gamma\} \geq 1 - \rho$.*
2. *for all $h \in \Gamma$, we have $\nu_s\{\gamma | \gamma\mathbb{A}h\} \geq 1 - \rho$.*

We say that a Markov bundle (X, ν) over Γ is ρ-pivotal for some $\rho \in [0, 1]$ if there is an edge $(a : b)$ such that the distribution $\nu_{a:b}$ is ρ-Schottky, and such that a Markov chain in (X, p_ν) goes through $(a : b)$ with probability one. We call such an edge a pivotal edge or a ρ-pivotal edge.

Definition 4.7 (Ping-pong)**.** *We call ping pong base a Markov chain $(\{s, a, b\}, p)$ where $p(s, a) = p(a, b) = p(b, a) = 1$. The points s, a, b do not have to be distinct.*

Definition 4.8. *Let Γ be a measurable category endowed with a binary relation $\mathbb{A}$ and $\rho \in [0,1]$. We say that a Markov bundle (X, ν) is ρ-ping-pong if (X, p_ν) is a ping pong base with $X_s = \{s\}$ and $\nu_{a,b}$ is ρ-Schottky.*

The main theorem of this section is the following

Theorem 4.9. *Let ν be a strongly irreducible and proximal distribution over a linear monoid $\Gamma := \mathrm{End}(E)$. Let $0 < \rho < \frac{1}{3}$ and $p \geq 1$ and $N : \Gamma \to \mathbb{R}_{\geq 0}$ a measurable function. Assume that $N(\nu)$ is almost surely finite. There exists a constant $\varepsilon > 0$ such that for all $\lambda > 0$, there is an exponentially integrable and eventually irreducible extraction $(\{s,a,b\}, \tilde{\kappa})$ of $(\{*\}, \nu)$ and three constants $l_a \in \mathbb{N}$ and $B, C \in \mathbb{R}_{\geq 0}$ such that:*

1. *For $s = b$ and $\kappa = \Pi\tilde{\kappa}$, The Markov bundle $(\{s,a,b\}, \kappa)$ is ρ-ping-pong, strongly $\mathbb{A}^\varepsilon$-aligned and λ-squeezing.*
2. *Given a random word $\tilde{g} \sim \tilde{\kappa}_a$, the length $L(\tilde{g}) = l_a$ and for all $0 \leq k < l_a$, we have $N(\chi_k(\tilde{g})) \leq B$.*
3. *Given a word $\tilde{g} \sim \tilde{\kappa}_s$ or $\tilde{g} \sim \tilde{\kappa}_b$, we have:*

$$\forall k \in \mathbb{N}, \forall t > B\ \forall l > k,\ \mathbb{E}\left(N\left(\chi_k(\tilde{g})\right) > t \,|\, L\left(\tilde{g}\right) = l\right) \leq CN(\nu)(t, +\infty). \tag{38}$$

This theorem says two things, the first one is that there is always an aligned and squeezing pivotal extraction, without any moment conditions on ν. The second is that when we have a moment condition on ν then we can have it on the pivotal extraction, it is not a direct consequence of the existence of the extraction. Indeed a sum of exponentially many L^p integrable random variables is not necessarily L^p. Note also that one cans show that C does not actually depend on λ but this is not useful to prove the main results because we may as well take $\lambda = 100|\log(\varepsilon)| + 100$ once and for all.

4.2 Construction of the ping-pong extraction

Definition 4.10. *Let E be a Euclidean vector space, let $0 < \varepsilon, \rho, \alpha \leq 1$. We say that a probability distribution ν on $\mathrm{End}(E)$ is $(\varepsilon, \rho, \alpha, \lambda)$-mixing if there is a probability distribution ν_A that is ρ-Schottky for $\mathbb{A}^\varepsilon$, and λ-squeezing and such that $\alpha\nu_A \leq \nu$.*

Remark 4.11. *We define the projective Wasserstein distance between two probability distributions η and ν over $\Gamma \setminus \{0\}$ as the minimum of $\mathbb{E}(\|ag - bh\|)$ for a coupling of two random variables $g \sim \nu$ and $h \sim \eta$ and a, b two random scalars such that $\|ag\| = \|bh\| = 1$. Let $\varepsilon > 0, \rho > 0, \alpha > 0$ and $\lambda \geq |\log(\varepsilon)| + \log(2)$. For all $\rho' < \rho$, all $\varepsilon' < \varepsilon$, all $\alpha' < \alpha$ and $\lambda' < \lambda$, there is a constant $\delta > 0$ such that for all distributions ν that is $(\varepsilon, \rho, \alpha, \lambda)$-mixing, all distribution η at Wasserstein distance at most δ is $(\varepsilon', \rho', \alpha', \lambda')$-mixing. In other words, being mixing is a robust notion that can be guaranteed by a Las Vegas type algorithm. Indeed, drawing a random matrix with distribution law ν is not possible with a computer in general but on can model a distribution arbitrarily Wasserstein-close to ν this would give an approximation of the constants defined in the introduction.*

Lemma 4.12. *Let ν be a strongly irreducible and proximal distribution over $\Gamma := \mathrm{End}(E)$. For all $\rho > 0$, there is a constant $\varepsilon > 0$ such that for all λ, there is an integer m and $\alpha > 0$ such that ν^{*m} is $(\varepsilon, \rho, \alpha, \lambda)$-mixing.*

Proof. Since ν is proximal, the set $\partial(\nu)$ defined in 3.36 is not empty. Write $\partial_u(\nu)$ and $\partial_w(\nu)$ for the left and right boundaries as defined in Lemma 3.37. Then $\partial_u(\nu)$ in invariant by the left action of ν and $\partial_w(\nu)$ is invariant by the right action of ν. The claim is that for every N, one can take two families $(w_1, \cdots, w_N) \in \partial_w(\nu)^N$ and $(u_1, \ldots, u_N) \in \partial_u(\nu)^N$ in general position[7].

We prove this claim by induction: For $N = 0$ this is trivial. Then if we assume the claim for some $N - 1 \geq 0$, take $x_1, \cdots, x_{N-1} \in \partial_v(\nu)$ in general position and write:

$$F := \bigcup_{1 \leq i_1 \leq \cdots \leq i_{d-1} \leq N-1} \mathbf{span}\left\{u_{i_1}, \ldots, u_{i_{d-1}}\right\} \subset E,$$

[7]we say that a family (x_i) of vectors is in general position if for every subset $\{i_1, \ldots, i_k\}$ of indices with $k \leq d$, the family $(x_{i_1}, \ldots, x_{i_d})$ is linearly free.

Since F is a finite union of proper subspaces of E, then F does not contain $\partial_u(\nu)$ so one can take a line $u_N \in \partial_u(\nu) \setminus F$ and then the family $u_1, \cdots, u_N \in \partial_w(\nu)$ is in general position. The proof of the claim is the same for $\partial_w(\nu)$ by duality. Now take $N := \left\lceil \frac{d-1}{\rho} \right\rceil$ and two families $w_1, \cdots, w_N \in \partial_w(\nu)$ and $u_1, \ldots, u_N \in \partial_u(\nu)$ in general position and write:

$$\varepsilon_u := \min_{\substack{f\in E^*,\\ \|f\|=1}} \left(\max_{i_1<\cdots<id-1} \left(\min_{i\notin\{i_1,\ldots,i_{d-1}\}} |f(u_i)| \right) \right); \tag{39}$$

$$\varepsilon_w := \min_{\substack{x\in E,\\ \|x\|=1}} \left(\max_{i_1<\cdots<i_{d-1}} \left(\min_{i\notin\{i_1,\ldots,i_{d-1}\}} |w_i(x)| \right) \right); \tag{40}$$

$$\varepsilon := \frac{\min\{\varepsilon_u, \varepsilon_w\}}{2} \tag{41}$$

The constant ε is not 0 because for every x, the hyperplane $x^\perp$ contains at most $d-1$ w_i's so:

$$\forall x \in \mathbf{S}(E), \max_{1\leq i_1<\cdots<id-1\leq M} \left(\min_{i\notin\{i_1,\ldots,i_{d-1}\}} |w_i(x)| \right) > 0.$$

Moreover each $|w_(x)|$ is 1-Lipschitz so the max min also is and $\mathbf{P}(E)$ is compact so this is indeed a minimum.

Now take any $\lambda \geq 2|\log(\varepsilon)| + 2\log(2)$ and for every pair of indices $(i,j) \in \{1,\ldots,N\}^2$, write:

$$S_{i,j} := \{\gamma \in \Gamma \setminus \{0\} | \, \mathrm{d}(W^+(\gamma), w_i) \leq \varepsilon, \mathrm{d}(U^+(\gamma), u_j) \leq \varepsilon, \mathrm{sqz}(\gamma) \geq \lambda\}.$$

Now we claim that all probability distribution κ, that satisfies $\kappa(S_{i,j}) \geq \frac{1}{N^2}$ for all i,j, is ρ-Schottky. Indeed, consider such a κ, and $g \in \Gamma$. Take some unitary $u_g \in U^+(g)$ and $w_h \in W^+(g)$. There is a subset $I \subset \{1,\ldots,N\}$ of size $N-d+1$ such that for all $i \in I$, we have $|w_g(u_i)| \geq 2\varepsilon$. Then consider some indices $(i,j) \in I \times \{1,\ldots,N\}$ and an endomorphism $\gamma \in S_{i,j}$, take a unitary $u_\gamma \in U^+(\gamma)$, we have $\mathrm{d}(u_\gamma, u_i) \leq \varepsilon$ so by Lemma 2.4, we have $|w_g(u_\gamma)| \geq 2\varepsilon - \varepsilon \; g\mathbb{A}^\varepsilon\gamma$ so $\kappa\{\gamma \,|\, g\mathbb{A}^\varepsilon\gamma\} \geq \frac{N-d+1}{N} \geq 1-\rho$. Then with the same reasoning, we get that $\kappa\{\gamma \,|\, \gamma\mathbb{A}^\varepsilon g\} \geq 1-\rho$ so κ is ρ-Schottky

Moreover each $S_{i,j}$ is invariant by scalar multiplication and is a neighbourhood of the point $u_j w_i \in \partial(\nu)$ so by definition, there is an integer $n_{i,j}$ such that $\nu^{*n_{i,j}}(S_{i,j}) > 0$ We want to show that actually there is an integer n such that $\nu^{*n}(S_{i,j}) > 0$ for all $1 \leq i,j \leq N$.

For all $1 \leq i,j \leq N$, write:

$$S'_{i,j} := \left\{ \gamma \in \Gamma \setminus \{0\} \,\middle|\, \mathrm{d}(W^+(\gamma), w_i) < \frac{\varepsilon}{2}, \mathrm{d}(U^+(\gamma), u_j) < \frac{\varepsilon}{2}, \mathrm{sqz}(\gamma) > \lambda \right\}.$$

Then for all i,j we have some $n'_{i,j}$ such that $\nu^{*n'_{i,j}}(S'_{i,j}) > 0$. Then write:

$$A := \left\{(i,j) \in \{1,\ldots,N\}^2 \,|\, |u_i(w^j)| \geq 2\varepsilon \right\}.$$

By Lemma 2.18, for all sequence $(i_k, j_k)_{1\leq k\leq M}$ such that $(i_k, j_{k+1}) \in A$ for all $1 \leq k < M$, we have $S'_{i_1,j_1} \cdots S'_{i_M,j_M} \in S_{i_1,j_M}$. To all pair (i,j) of indices, we associate a pair $(i',j') \in A$ such that $(i,j') \in A$ and $(i',j) \in A$. Such a pair exists because if we take $(i',j') \in \{1,\ldots,N\}^2$ uniformly at random, we have $\mathbb{P}((i,j') \in A) \geq 1-\rho$ and $\mathbb{P}((i',j) \in A) \geq 1-\rho$ and $\mathbb{P}((i',j') \in A) \geq 1-\rho$ and $\rho < \frac{1}{3}$ so $\mathbb{P}(\{(i,j'),(i',j),(i',j')\} \subset A) \geq 1-3\rho > 0$. Therefore, there is at least one pair of indices (i',j') such that $\{(i,j'),(i',j),(i',j')\} \subset A$. Now for all pair (i,j), write $p_{i,j} := n'_{i,j} + n'_{i',j'}$ and write M for the smallest common multiple of the family $(p_{i,j})_{1\leq i,j\leq N}$. Then for all pair (i,j), write $q_{i,j}$ for the integer such that $p_{i,j}q_{i,j} = M$. Then for all pair (i,j), the product $\left(S'_{i,j} \cdot S'_{i',j'}\right)^{q_{i,j}} \cdot \left(S'_{i',j'} \cdot S'_{i,j}\right)^{q_{i,j}}$ has positive ν^{*2M}-measure and is included in $S_{i,j}$ by Lemma 2.18. Therefore $\nu^{*2M}(S_{i,j}) > 0$ for all pair i,j. Then we write $m := 2M$ and $\alpha' := \min_{i,j} \nu^{*m}(S_{i,j})$. For all i,j, write $\tilde{\eta}_{i,j}$ for the normalised restriction of $\nu^{\otimes m}$ to $\widetilde{S}_{i,j}$. Then $\alpha'\tilde{\eta}_{i,j} \leq \nu^{\otimes m}$ for all i,j (if we see them as functions $\mathcal{A}_\Gamma^{\otimes m} \to [0,1]$) so we may write $\alpha \mathrm{d}\tilde{\eta}_{i,j} = f_{i,j}\mathrm{d}\nu^{\otimes m}$ with $f_{i,j} : \Gamma^m \to [0,1]$ a measurable density function. Then write:

$$f := \max_{1\leq i\leq j\leq N} f_{i,j}$$
$$\mathrm{d}\tilde{\eta}_a := \frac{f\mathrm{d}\nu^{\otimes m}}{\int_{\Gamma^m} f\mathrm{d}\nu^{\otimes m}}$$
$$\mathrm{d}\tilde{\eta}_b := \frac{(1-f)\mathrm{d}\nu^{\otimes m}}{\int_{\Gamma^m}(1-f)\mathrm{d}\nu^{\otimes m}}.$$

Then we have $\int f\mathrm{d}\nu^{\otimes m} \leq N^2\alpha$ because $\max f_{i,j} \leq \sum f_{i,j}$ but $\int_{S_{i,j}} f\mathrm{d}\nu^{\otimes m} \geq \alpha'$ for all i,j so $\tilde{\eta}_a(\widetilde{S}_{i,j}) \geq \frac{1}{N^2}$ for all $1 \leq i,j \leq N$, which means that the product η_a is ρ-Schottky for $\mathbb{A}^{\varepsilon}$. Write $\alpha := \int f\mathrm{d}\nu^{\otimes m}$ then we have $\alpha\eta_a \leq \nu^{*m}$ by definition. □

Lemma 4.13. *Let ν be a strongly irreducible and proximal probability distribution over $\Gamma := \mathrm{End}(E)$. Let $N : \Gamma \to [0,+\infty]$ be a measurable map such that $N(\nu)$ is almost surely finite. Let $0 < \rho < \frac{1}{3}$. There is a constant $\varepsilon > 0$ such that for all $\lambda < +\infty$, there is an exponentially integrable and eventually irreducible extraction $(\{a,b\},\tilde{\kappa})$ of $(\{*\},\nu)$ and two constants $l_a \in \mathbb{N}$ and $B, C \in \mathbb{R}_{\geq 0}$ such that:*

1. *For $s = b$ and $\kappa = \Pi\tilde{\kappa}$, The Markov bundle $(\{s,a,b\},\kappa)$ is ρ-ping-pong i.e, κ_a is ρ-Schottky for the alignment relation $\mathbb{A}^{\varepsilon}$.*
2. *Given a random word $\tilde{g} \sim \tilde{\kappa}_a$, the length $L(\tilde{g}) = l_a$ and for all $0 \leq k < l_a$, we have $N(\chi_k(\tilde{g})) \leq B$ and $\mathrm{sqz}(\Pi(\tilde{g})) \geq \lambda$ almost surely.*
3. *Given a word $\tilde{g} \sim \tilde{\kappa}_b$, we have for all $A \subset \Gamma$ measurable:*

$$\forall l > k,\ \mathbb{P}\left(\chi_k(\tilde{g}) \in A \,|\, L(\tilde{g}) = l\right) \leq 2N(\nu)(t,+\infty). \tag{42}$$

Proof. Consider $\varepsilon > 0$, $\lambda \geq 2|\log(\varepsilon)|$, $\alpha > 0$ and $l_a \in \mathbb{N}$ be such that there is a probability distribution η'_a that is $\frac{\rho}{2}$-Schottky, for $\mathbb{A}^{\varepsilon}$ and λ-proximal, and such that $\alpha'\eta'_a \leq \nu^{*l_a}$. Such constants exist by lemma 4.12. Consider $\tilde{\eta}_a$ a lift pf η'_a such that $\alpha'\tilde{\eta}'_a \leq \nu^{\otimes l_a}$. Since $N(\nu)$ is almost surely finite, we have a constant B such that $N(\nu)(B,+\infty) \to 0$ as B goes to infinity. Take b so that $\nu^{\otimes l_a}(N^{-1}(B,+\infty)^{l_a}) \geq 1-\alpha'/2$. Then the normalised restriction $\tilde{\eta}_a$ of $\tilde{\eta}'_a$ to $N^{-1}(B,+\infty)^{l_a}$ is bounded by $2\tilde{\eta}'_a$ so its product is ρ-Schottky and we have $\alpha'\tilde{\eta}_a \leq 2\nu^{\otimes l_a}$. Write $\alpha := \alpha'/2$ Draw $(\gamma_1,\dots,\gamma_{l_a},x) \sim \alpha\tilde{\eta}_a \otimes \delta_a + (1-\alpha')\tilde{\eta}_b \otimes \delta_b$. Then we have $(\gamma_1\dots,\gamma_m) \sim \nu^{\otimes m}$ so for all $k \in \{1,\dots,m\}$, and for all measurable $A \subset \Gamma$, we have:

$$\begin{aligned}\mathbb{P}\left(\gamma_k \in A\right) &\leq \nu(A)\\ \mathbb{P}\left(\gamma_k \in A \cap (x=b)\right) &\leq \nu(A)\\ \mathbb{P}\left(\gamma_k \in A \,|\, x=b\right) &\leq 2N(\nu)(t,+\infty).\end{aligned} \tag{43}$$

Then we define the Markov Bundle $(\{a,b\},\tilde{\kappa})$ as follows:

$$\tilde{\kappa}_{a,b} = \tilde{\eta}_a \qquad\qquad p_{\tilde{\kappa}}(a,b) = 1$$
$$\tilde{\kappa}_{b,a} = \sum_{k=0}^{\infty}\frac{(1-\alpha')^k}{\alpha'}\tilde{\eta}_b^{\odot k} \qquad\qquad p_{\tilde{\kappa}}(b,a) = 1.$$

Then $(\{a,b\},\tilde{\kappa})$ is ρ-ping-pong because η_a is ρ Schottky. Now if we draw:

$$(\gamma_{mk}\dots,\gamma_{m(k+1)-1},x_k)_{k\in\mathbb{N}} \sim \left(\alpha'\tilde{\eta}_a \otimes \delta_a + (1-\alpha')\tilde{\eta}_b \otimes \delta_b\right)^{\otimes\mathbb{N}}. \tag{44}$$

Then we have $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$ and if we write t_0 for the first time such that $x_{t_0} = a$ then we have $(\gamma_0,\cdots,\gamma_{mt_0+1-1}) \sim \tilde{\kappa}_{b,a} \odot \tilde{\kappa}_{a,b}$. Moreover, t_0 is a stopping time so $(\gamma_{m(k+t_0+1)},\dots,\gamma_{m(k+t_0+2)-1},x_{k+t_0+1})_{k\in\mathbb{N}}$ has the same distribution as $(\gamma_{mk},\dots,\gamma_{m(k+1)-1},x_k)_{k\in\mathbb{N}}$. In particular, $(\gamma_{n+mt_0+1}) \sim \nu^{\otimes\mathbb{N}}$, so we have:

$$\tilde{\kappa}_{b,a} \odot \tilde{\kappa}_{a,b} \odot \nu^{\otimes\mathbb{N}} = \nu^{\otimes\mathbb{N}}$$
$$(\tilde{\kappa}_{b,a} \odot \tilde{\kappa}_{a,b})^{\odot\mathbb{N}} = \nu^{\otimes\mathbb{N}},$$

because $L(\tilde{\kappa}_{b,a} \odot \tilde{\kappa}_{a,b})$ is not supported on $\{0\}$. It means that the Markov Bundle $(\{a,b\},\tilde{\kappa})$ is an extraction of $(\{*\},\nu)$. Then by definition, the length $L(\tilde{\kappa}_{b,a})$ has finite exponential moment. Moreover, for all k, the distribution of the k-th letter of $\tilde{g}$ knowing $L(\tilde{g})$ is simply the $k - m\left\lfloor \frac{k}{m} \right\rfloor$-th marginal of $\tilde{\eta}_b$ so (43) implies (42). □

Lemma 4.14. *Let ν be a strongly irreducible and proximal distribution over $\Gamma := \mathrm{End}(E)$. Let $N : \Gamma \to [0,+\infty]$ be a measurable map such that $N(\nu)$ is almost surely finite. Let $0 < \rho < \frac{1}{3}$. There is a constant $\varepsilon > 0$ such that for all $\lambda < +\infty$, there is an exponentially integrable and eventually irreducible extraction $(\{a,b\},\tilde{\eta})$ of $(\{*\},\nu)$ and two constants $l_a \in \mathbb{N}$ and $B \in \mathbb{R}_{\geq 0}$ such that:*

1. *For $s = b$ and $\eta = \Pi\tilde{\eta}$, The Markov bundle $(\{s,a,b\},\kappa)$ is λ-squeezing and ρ-ping-pong i.e, η_a is ρ-Schottky for the alignment relation $\mathbb{A}^\varepsilon$.*
2. *Given a random word $\tilde{g} \sim \tilde{\eta}_b$, the length $L(\tilde{g}) = l_a$ and for all $0 \leq k < l_a$, we have $N(\chi_k(\tilde{g})) \leq B$.*
3. *Given a word $\tilde{g} \sim \tilde{\eta}_b$, we have:*

$$\forall t \geq B,\ \forall l > k,\ \mathbb{P}\left(N\left(\chi_k(\tilde{g})\right) > t \,|\, L(\tilde{g}) = l\right) \leq 2N(\nu)(t,+\infty). \tag{45}$$

Proof. Let ε, B be as in Lemma 4.13 and let $\lambda' = \lambda + 4|\log(\varepsilon)| + 4\log(2)$. Consider $(\{a,b\},\tilde{\kappa})$ the extraction constructed in Lemma 4.13. Now we build a new extraction $(\{a,b\},\tilde{\eta})$. Write m for the length of $\tilde{\kappa}_a$. Draw three random sequences: $(\tilde{g}_{2k})_{k\in\mathbb{N}} \sim \tilde{\kappa}_b^{\otimes\mathbb{N}}$ and $(\tilde{g}_{2k+1})_{k\in\mathbb{N}} \sim \tilde{\kappa}_a^{\otimes\mathbb{N}}$ and $(\tau_k)_{k\in\mathbb{N}} \in [0,1]^{\mathbb{N}}$ uniformly. Given $f,g,h \in \mathrm{End}(E)$, and $\gamma \sim \kappa_a$, we define:

$$F(f,g,h) := \frac{\mathbb{1}(f\mathbb{A}^\varepsilon g\mathbb{A}^\varepsilon h)}{\mathbb{P}\left(f\mathbb{A}^\varepsilon\gamma\mathbb{A}^\varepsilon h\right)}(1-2\rho) \tag{46}$$

For all k, write $g_k := \Pi(\tilde{g}_k)$ and $w_k := L(\tilde{g}_k)$ and write $(\gamma_n)_{n\in\mathbb{N}} := \bigodot_{k=0}^{+\infty}\tilde{g}_k$. We define k_0 to be the smallest random integer such that:

$$\tau_{k_0} < F(\overline{g}_{2k_0+1}, g_{2k_0+1}, g_{2k_0+2}).$$

It implies that $F(\overline{g}_{2k_0+1}, g_{2k_0+1}, g_{2k_0+2}) > 0$, so we have:

$$\overline{g}_{2k_0+1}\mathbb{A}^\varepsilon g_{2k_0+1}\mathbb{A}^\varepsilon g_{2k_0+2}.$$

Then by Lemma 2.18, we have:

$$\mathrm{sqz}\left(\overline{g}_{2k_0+3}\right) \geq \lambda' - 4|\log(\varepsilon)| - 4\log(2) \geq \lambda. \tag{47}$$

Moreover, F takes values in $[0,1]$, so we have for all $k \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{P}\left(k_0 > k \,|\, (\tilde{g}_{2n})_{n\in\mathbb{N}}, (\tilde{g}_{2k'+1})_{k'<k}\right) &= \mathbb{P}\left(\tau_k \geq F_k \,|\, (\tilde{g}_{2n})_{n\in\mathbb{N}}, (\tilde{g}_{2k'+1}, \tau_{k'})_{k'<k}\right) \\
&= \mathbb{E}\left(1 - F_k \,|\, (\tilde{g}_{2n})_{n\in\mathbb{N}}, (\tilde{g}_{2k'+1}, \tau_{k'})_{k'<k}\right) \\
&= 2\rho
\end{aligned}$$

As a consequence, k_0 has a geometric distribution of scale factor 2ρ and is independent of $(\tilde{g}_{2n})_{n\in\mathbb{N}}$. Now for all $n \in \mathbb{N}$, write t_n for the only integer such that $\overline{w}_{t_n} \leq n < \overline{w}_{t_n+1}$. Then by (42), we have:

$$\forall A \subset \Gamma,\ \mathbb{P}\left(N(\gamma_n) > t \,|\, (w_{2k})_{k\in\mathbb{N}}, t_n \in 2\mathbb{N}\right) \leq 2\nu(A). \tag{48}$$

Moreover k_0 and $(\tilde{g}_{2k})_{k\in\mathbb{N}}$ are independent so conditionally to $t_n \in 2\mathbb{N}$), the random variables k_0 and γ_n are independent. Then (48) becomes:

$$\forall A \subset \Gamma,\ \mathbb{P}\left(N(\gamma_n) > t \,|\, (w_{2k})_{k\in\mathbb{N}}, k_0, t_n \in 2\mathbb{N}\right) \leq 2\nu(A). \tag{49}$$

Moreover, when t_n is odd, $N(\gamma_n) \leq B$, so for all $t \geq B$, we have:

$$\mathbb{P}(N(\gamma_n) > t \,|\, t_n \notin 2\mathbb{N}) = 0. \tag{50}$$

If we combine (50) with (49) and take $A = \{\gamma \in \Gamma \,|\, N(\gamma) > t\}$, then we get:

$$\forall t \geq B, \mathbb{P}(N(\gamma_n) > t \,|\, (w_{2k})_{k\in\mathbb{N}}, k_0) \leq 2N(\nu)(t, +\infty). \tag{51}$$

Then $\overline{w}_{k_0+3}$ is a function of $((w_{2k})_{k\in\mathbb{N}}, k_0)$ so we have:

$$\forall t \geq B, \mathbb{P}\left(N(\gamma_n) > t \,|\, \overline{w}_{k_0+3}\right) \leq 2N(\nu)(t, +\infty).$$

Which proves (45) for $\tilde{\eta}_b$ the distribution of $\tilde{g}_0 \odot \cdots \odot \tilde{g}_{2k_0+2}$. Write $\tilde{\eta}_a := \tilde{\kappa}_a$. Then note that $2k_0 + 2$ is a stopping time for the sequence $(\tilde{g}_k)$ and it is non-zero so we have:

$$\begin{aligned} \tilde{\eta}_b \odot \tilde{\eta}_a \odot (\tilde{\kappa}_b \odot \tilde{\kappa}_a)^{\odot\mathbb{N}} &= (\tilde{\kappa}_b \odot \tilde{\kappa}_a)^{\odot\mathbb{N}} \\ (\tilde{\eta}_b \odot \tilde{\eta}_a)^{\odot\mathbb{N}} &= (\tilde{\kappa}_b \odot \tilde{\kappa}_a)^{\odot\mathbb{N}} \\ (\tilde{\eta}_b \odot \tilde{\eta}_a)^{\odot\mathbb{N}} &= \nu^{\otimes\mathbb{N}}. \end{aligned}$$

This proves that $(\{a, b\}, \tilde{\eta})$ is indeed an extraction of $(\{*\}, \nu)$. □

4.3 Pivot algorithm

Now we show how to turn a ρ-ping pong and λ-squeezing extraction into a pivotal and aligned extraction.

Definition 4.15 (Naive pivot algorithm)**.** *Given (γ_n) an acceptable sequence in a category Γ, and $\mathbb{A}$ a binary relation on Γ. we define a sequence of integers $(w_k) \in \mathbb{N}^{\mathbb{N}} \sqcup \mathbb{N}^{(\mathbb{N})} \times \{+\infty\}$ as follows: create $n_0 := 0$ and $m_0 := 0$ two integer variables and create a variable sequence $(w^0_k)_k \in \mathbb{N}^{\mathbb{N}}$ that we initialize with the value $w^0_k := 1$ for all k. Then we define inductively for $j \geq 0$:*

1. *if $\gamma^w_{m_j}\mathbb{A}\gamma_{j+1}$, set $w^{j+1}_k := w^{j+1}_k$ for all k, and change $m_{j+1} := m_j + 1$ and $n_{j+1} := j + 1$,*
2. *otherwise, and if there is an integer $l \in \{0, \dots, m\}$ such that:*

   $$\gamma^{w^j}_{l-1}\mathbb{A}\left(\gamma^{w^j}_l \cdots \gamma^{w^j}_{m_j}\gamma_{j+1}\right)$$

   then take l the largest such integer and change $w^{j+1}_l := w^j_l + \cdots + w^j_{m_j+1}$, $m_{j+1} := l$ and $n_{j+1} := j+1$
3. *otherwise write $t \geq j + 2$ the smallest integer such that $\overline{\gamma}_{t-1}\mathbb{A}\gamma_{t-1}\mathbb{A}\gamma_t$ and set $w^{j+1}_0 := t$ then change $m_{j+1} := 0$ and $n_{j+1} := j + t$. If there is no such t, the algorithm ends and we set $w_0 := +\infty$ and $m = 0$.*

If there is an index k such that $(w^j_k)_j$ does not converge then write m for the smallest such integer and set $w_m := +\infty$, otherwise, we say that the pivot algorithm converges and write $m = +\infty$. For all $k < m$, the sequence $(w^j_k)_j$ converges (is stationary) and we set w_k to be its limit. This creates a sequence of integers $(w_k)_{0\leq k<m}$ with $m = \liminf m_j$ and $w_k = \lim w^j_k$.

Lemma 4.16. *Write $\Gamma = \langle a, b, c | a^2, b^2, c^2\rangle$, consider the binary relation $\mathbb{A}$ defined by $(g\mathbb{A}h) :\Leftrightarrow (\chi_{L(g)}(g) \neq \chi_1(h))$ or equivalently $(g\mathbb{A}h) :\Leftrightarrow (L(gh) = L(g) + L(h))$. Then for $(\gamma_n) \in \{a, b, c\}^{\mathbb{N}}$ uniformly distributed and independent, the naive pivot algorithm converges almost surely and the extracted sequence $(\widetilde{\gamma}^w_k)$ is the pivotal extraction described in the introduction.*

Sketch of proof. One can show by induction that at every step of the algorithm, $(\overline{\gamma}^w_0, \dots, \overline{\gamma}^w_m)$ is a geodesic segment. Moreover the average length gain in case 1 is $\frac{2}{3}$ and the average loss in case2 is $\frac{1}{3}$ because $l = m-1$ always works. As a consequence one has $m \sim \frac{n}{3}$ with exponential rare linear deviations so $\overline{\gamma}_n$ travels along a geodesic with exponentially rare linear deviations *i.e,* for all $\varepsilon > 0$, the probability of large deviations $\mathbb{P}\left(\mathrm{d}\left(\overline{\gamma}_n, \overline{\gamma}^w_{\lfloor n/3\rfloor}\right) \geq n\varepsilon\right)$ decreases exponentially fast with n. □

Now the problem with the naive pivot algorithm is that it is not easy to show that it converges in general, therefore we will use the following pivot algorithm, for which we can show that there is convergence only assuming that the odd elements of the sequence have a good behaviour, namely they are ρ-Schottky for ρ small enough.

Definition 4.17 (Alternate pivot algorithm)**.** *Let (γ_n) be an acceptable sequence in a category Γ, and $\mathbb{A}$ a binary relation on Γ. We define a sequence of integers $(w_k) \in \mathbb{N}^{\mathbb{N}} \sqcup \mathbb{N}^{(\mathbb{N})} \times \{+\infty\}$ as follows:*
Create by induction over $j \in \mathbb{N}$, a family of sequences $(w_k^j)_{0 \leq k} \in \mathbb{N}^{\mathbb{N}}$ that we initialise with the value $w_k^0 := 1$ for all k and a family of integers that we initialise with the value $m_0 = 0$. For all $j \in \mathbb{N}$, we do as follows:

1. *if $\gamma_{2m_j}^{w^j} \mathbb{A} \gamma_{2j+1} \mathbb{A} \gamma_{j+2}$, change $m_{j+1} := m_j + 1$ and keep $w_k^{j+1} = w_k^j$ for all k,*
2. *otherwise, and if there is an even integer $l \in \{1, \dots, m_j - 1\}$ such that:*

$$\gamma_{2l+1}^{w^j} \mathbb{A} \left(\gamma_{2l+2}^{w^j} \cdots \gamma_{2m_j}^{w^j} \gamma_{2j+1} \gamma_{2j+2} \right),$$

then take l the largest such integer or $l = 0$ if there is no such integer and change $w_{2l}^{j+1} := w_{2l}^j + \cdots + w_{m_j}^j + 2$, $m_{j+1} := l$, for all $k > 2m_{j+1}$, set $w_k^{j+1} := 1$ and for $k < 2m_{j+1}$, we keep $w_k^{j+1} := w_k^j$.

If there is an index k such that $(w_k^j)_j$ does not converge then write m for the smallest such integer and set $w_m := +\infty$, otherwise, we say that the pivot algorithm converges and write $m = +\infty$. For all $k < m$, the sequence $(w_k^j)_j$ converges (is stationary) and we set w_k to be its limit. This creates a sequence of integers $(w_k)_{0 \leq k < m}$ with $m = \liminf m_j$ and $w_k = \lim w_k^j$.

Lemma 4.18. *Take a constant $\varepsilon > 0$, define $\lambda = 4|\log(\varepsilon)| + 4\log(2)$ and take a relation $\mathbb{A} \subset \mathbb{A}^{\varepsilon}$. Take a sequence $(\gamma_k) \in \mathrm{End}(E)^{\mathbb{N}}$ that is λ-squeezing. Then the families $(w_k^j)_{j,k}$ and $(m_j)_j$ defined by the alternate pivot algorithm as in Definition 4.17 satisfy the following properties:*

1. *at each iteration j of the algorithm, we have $\overline{w}_{2m_j+1}^j = 2j$,*
2. *for all $k \in \mathbb{N}$, the sequence $(\overline{w}_k^j)_{j \in \mathbb{N}}$ is non-decreasing and for all odd k and all j, we have $\overline{w}_k^j = 1$,*
3. *at each iteration j of the algorithm, for all $k < 2m_j$, we have $\gamma_k^{w^j} \mathbb{A}^{\frac{\varepsilon}{2}} \gamma_{k+1}^{w^j}$,*
4. *at each iteration j of the algorithm, for all $1 \leq k \leq 2m_j$, we have $\mathrm{sqz}\left(\gamma_k^{w^j}\right) \geq \lambda$.*

Proof. Items (1) and (2) are direct properties of the construction. Note also that for all odd k, we have $\gamma_k^{w^j} = \gamma_{\overline{w}_k^j}$ so we have $\mathrm{sqz}(\gamma_k^{w^j}) \geq \lambda$ by hypothesis, which proves the odd part of (4). Now we show by induction on j the proposition $C(j)$: "for all even integer $k < 2m_j$, we have $\gamma_k^{w^j} \mathbb{A}^{\varepsilon} \gamma_{k+1}^{w^j}$". Let $j \geq 0$ and assume $C(j)$, we consider two cases, if we are in case (2) the for all even $k < 2m_{j+1}$, we also have $k < 2m_j$ and $\gamma_k^{w^{j+1}} = \gamma_k^{w^j}$ and $\gamma_{k+1}^{w^{j+1}} = \gamma_{k+1}^{w^j}$ (indeed only $\gamma_{2m_{j+1}}^{w^j}$ changes) so we have $\gamma_k^{w^{j+1}} \mathbb{A}^{\varepsilon} \gamma_{k+1}^{w^{j+1}}$ by $C(j)$, which proves $C(j+1)$, otherwise, we are in case (1), then for all even $k < 2m_{j+1}$, we either have $k < 2m_j$, in which case the above reasoning still works and we have $\gamma_k^{w^{j+1}} \mathbb{A}^{\varepsilon} \gamma_{k+1}^{w^{j+1}}$ or we have $k = 2m_j$, then since the condition of (1) applies, we have $\gamma_{2m_j}^{w^j} \mathbb{A}^{\varepsilon} \gamma_{2j+1}$ and $\gamma_{2m_j}^{w^{j+1}} = \gamma_{2m_j}^{w^j}$ and $\gamma_{2m_j+1}^{w^{j+1}} = \gamma_{2j+1}^{w^j}$, which proves $C(j+1)$.

Now we get to the tricky part. We will show by induction on j the property $S(j)$: for all odd integer $k \leq 2m_j$, we have $\gamma_k^{w^j} \widetilde{\mathbb{A}} \widetilde{\gamma}_{k+1}^{w^j}$. Where the notation $\gamma \widetilde{\mathbb{A}} \tilde{g}$ is defined inductively and means that either $\tilde{g}$ has only one letter and $\gamma \mathbb{A}^{\varepsilon} g$ and $\mathrm{sqz}(g) \geq \lambda$; or there is a decomposition of $\tilde{g}$ into three words $\tilde{g} = \tilde{g}_1 \odot (s) \odot \tilde{g}_2$ such that $g_1 \mathbb{A}^{\varepsilon} s \mathbb{A}^{\varepsilon} g_2$ and $\mathrm{sqz}(s) \geq \lambda$ and $\gamma \widetilde{\mathbb{A}} \tilde{g}_1$. First we show $S(j)$ by induction. Let $j \geq 0$ and assume $S(j)$. Assume that we are in case (1) in Definition 4.17, then for all $k < m_j$ we have nothing to prove and for $k = 2m_j + 1$, we have $\gamma_{2m_j+1}^{w^{j+1}} \mathbb{A}^{\varepsilon} \gamma_{2j+2}^{w^j}$ and $\widetilde{\gamma}_{2j+2}^{w^j}$ has only one letter which proves $S(j+1)$. Now assume that we are in case (2), then for some odd $k < 2m_{j+1}$, we always have $\gamma_k^{w^{j+1}} = \gamma_k^{w^j}$ and if $k \neq 2m_{j+1} - 1$, then we also have $\widetilde{\gamma}_{k+1}^{w^{j+1}} = \widetilde{\gamma}_{k+1}^{w^j}$ so there is nothing to prove. If $k = 2m_{j+1} - 1$, then we have the decomposition:

$$\widetilde{\gamma}_{2m_{j+1}}^{w^{j+1}} = \widetilde{\gamma}_{2m_{j+1}}^{w^j} \odot \left(\gamma_{\overline{w}_{2m_{j+1}+1}^j} \right) \odot \left(\gamma_{\overline{w}_{2m_{j+1}+2}^j}, \dots, \gamma_{2j+2} \right). \tag{52}$$

Now we have $\gamma_k^{w^{j+1}} \widetilde{\mathbb{A}} \widetilde{\gamma}_{2m_{j+1}}^{w^j}$ by $S(j)$ and:

$$\gamma_{2m_{j+1}}^{w^j} \mathbb{A}^{\varepsilon} \gamma_{\overline{w}_{2m_{j+1}+1}^j} \mathbb{A}^{\varepsilon} \left(\gamma_{\overline{w}_{2m_{j+1}+2}^j} \cdots \gamma_{2j+2} \right)$$

because we are in case (2) with $m_{j+1} > 0$. This proves $S(j)$ for all j. Now we need to show that for all $\gamma \in \Gamma$ and all $\tilde{g} \in \mathrm{Paths}(\Gamma)$, such that $\gamma\widetilde{\mathbb{A}}\tilde{g}$, we have $\mathrm{sqz}(g) \geq \lambda$ and $\gamma\mathbb{A}^{\frac{\varepsilon}{2}}g$. We write $\lambda' := 2|\log(\varepsilon)| + 4\log(2)$ and we show by induction we show $D(\tilde{g})$: "$\tilde{g}$ can be decomposed into a product:

$$\tilde{g} = \tilde{g}^1 \odot \cdots \odot \tilde{g}^l$$

such that $\gamma\mathbb{A}^{\varepsilon}g^1\mathbb{A}^{\frac{\varepsilon}{2}} \cdots \mathbb{A}^{\frac{\varepsilon}{2}}g^l$ and $\mathrm{sqz}(g^i) \geq \lambda'$". Note that when $\tilde{g}$ has length one, we have such a decomposition with $l = 1$ because $\lambda \geq \lambda'$ Note also that by lemma 2.18 (14), $D(\tilde{g})$ implies that we have $\mathrm{sqz}(g) \geq \lambda$ and $\gamma\mathbb{A}^{\frac{\varepsilon}{2}}g$. For $\tilde{g}$ of length more than one, we decompose $\tilde{g} := \tilde{g}_1 \odot (s) \odot \tilde{g}_2$, then by $D(\tilde{g}_1)$, we have a decomposition:

$$\tilde{g}_1 = \tilde{g}^1 \odot \cdots \odot \tilde{g}^{l-1} \tag{53}$$

we write $\tilde{g}^l := (s) \odot \tilde{g}_2$. Then we have $g_1\mathbb{A}^{\varepsilon}s$ so by Lemma 2.18 (19), we have $g^{l-1}\mathbb{A}^{\frac{\varepsilon}{2}}g^l$ and by Lemma 2.13, we have $\mathrm{sqz}(g^l) \geq \mathrm{sqz}(s) + \mathrm{sqz}(g_2) - 2|\log(\varepsilon)| \geq \lambda'$. In conclusion, $S(j)$ implies (4) for all even $k \geq 2$ and (3) for all odd k and $C(j)$ implies (3) for all even k and (2) implies (4) for all odd k. ☐

Now we need to prove that the pivotal algorithm does converge almost surely for some specific sequences.

Lemma 4.19 (Pivotal concatenation)**.** *Let Γ be a measurable semi-group endowed with a binary relation $\mathbb{A}$ and $\rho < 1$. Let $g \in \Gamma$ be fixed. Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of random variables in Γ that respects a filtration $(\mathcal{F}_k)$ in the sense of Definition B.7 and such that for all odd n, the distribution of γ_n knowing $(\mathcal{F}_n)$ is ρ-Schottky in the sense of Definition 4.17. Then if one writes l for the smallest integer such that $(g\overline{\gamma}_l)\mathbb{A}\gamma_l$, one has:*

$$\forall k \in \mathbb{N},\ \mathbb{P}(l > 2k+1) \leq \rho^k.$$

As a consequence l has a finite exponential moment

Proof. First note that the family of events $(l > 2k+1)_{k\geq 0}$ is non-increasing for the inclusion. By Schottky property, for all k, we have $\mathbb{P}((g\overline{\gamma}_{2k+1})\mathbb{A}\gamma_{2k+1} \,|\, \mathcal{F}_{2k+1}) \geq (1-\rho)$ because $(g\overline{\gamma}_{2k+1})$ is $\mathcal{F}_{2k+1}$-measurable. This means that $\mathbb{P}(l > 2k+1 \,|\, \mathcal{F}_{2k+1}) < \rho$. Moreover, if $k \neq 0$ the event $(l > 2k-1)$ is $\mathcal{F}_{2k+1}$-measurable so $\mathbb{P}(l > 2k+1 \,|\, l > 2k-1) = \frac{\mathbb{P}(l>2k+1)}{\mathbb{P}(l>2k-1)} \leq \rho$. When $k = 0$, we have $\mathbb{P}(l > 1) \leq 1$ because $\mathbb{P}$ is a probability measure so we can conclude by induction. ☐

Now we want to show that the pivotal algorithm converges. To have some nice probabilistic estimates, we will change the algorithm the following way:

Definition 4.20 (Weighted pivot algorithm)**.** *Let (γ_n) be an acceptable sequence in a category Γ, and $\mathbb{A}$ a binary relation on Γ. Let $(t_k) \in [0,1]^{\mathbb{N}}$, let $F : \Gamma^3 \to [0,1]$ and let $R : \Gamma^4 \to [0,1]$ be two measurable functions. We define a sequence of integers $(w_k) \in \mathbb{N}^{\mathbb{N}} \sqcup \mathbb{N}^{(\mathbb{N})} \times \{+\infty\}$ as follows:*
We create by induction over $j \in \mathbb{N}$, a family of sequences $(w_k^j)_{0\leq k} \in \mathbb{N}^{\mathbb{N}}$ and a family of integers (m_j) that we initialise with the value $w_k^0 := 1$ for all k and $m_0 := 0$. the for all $j \geq 0$, we do as follows

1. *if $t_{2j+1} < F(\gamma_{2m_j}^{w^j}, \gamma_{2j+1}, \gamma_{2j+2})$, change $m_{j+1} := m_j + 1$ and keep $w_k^{j+1} = w_k^j$ for all k,*
2. *otherwise, and if there is an integer $l \in \{1, \dots, m_j - 1\}$ such that:*

$$t_{\overline{w}^j_{2l+1}} < R\left(\gamma_{2l}^{w^j}, \gamma_{2l+1}^{w^j}, \gamma_{2l+2}^{w^j}, \gamma_{\overline{w}^j 2l+2} \cdots \gamma_{2j+2}\right)$$

then take l to be the largest such integer and take $l = 0$ otherwise and change $w_l^{j+1} := w_l^j + \cdots + w_{2m_j}^j + 2$, take $m_{j+1} := l$ and for all $k > 2m_{j+1}$, set $w_k^{j+1} := 1$.

If there is an index k such that $(w_k^j)_j$ does not converge then write m for the smallest such integer and set $w_m := +\infty$, otherwise, we say that the pivot algorithm converges and write $m = +\infty$. For all $k < m$, the sequence $(w_k^j)_j$ converges (is stationary) and we set w_k to be its limit. This creates a sequence of integers $(w_k)_{0\leq k<m}$ with $m = \liminf m_j$ and $w_k = \lim w_k^j$.

Lemma 4.21. *Let $0 < \rho < \frac{1}{5}$. Let η be a ρ-Schottky probability distribution on a measurable semi-group Γ for a binary relation $\mathbb{A}$. Let $\gamma \sim \eta$. Write:*

$$\forall f, g, h \in \Gamma,\ F(f, g, h) := \frac{\mathbb{1}(f\mathbb{A}g\mathbb{A}h)}{\mathbb{P}\left(g\mathbb{A}\gamma\mathbb{A}h\right)}(1 - 2\rho)$$
$$\forall f, g, h, h' \in \Gamma,\ R(f, g, h, h') := \frac{\mathbb{1}(f\mathbb{A}^\varepsilon g\mathbb{A}^\varepsilon h)\mathbb{1}(g\mathbb{A}h')}{\mathbb{P}\left((g\mathbb{A}\gamma\mathbb{A}h) \cap (\gamma\mathbb{A}h')\right)}(1 - 3\rho)$$

Let $(\gamma_{2k+1})_{k\in\mathbb{N}} \sim \eta^{\otimes\mathbb{N}}$, let $(\gamma_{2k})_{k\in\mathbb{N}}$ be any random sequence and $(t_k)_{k\in\mathbb{N}} \in [0, 1]^{\mathbb{N}}$ be a uniformly distributed sequence independent of $(\gamma_n)_{n\in\mathbb{N}}$. Then the weighted pivotal algorithm associated to the sequences $(\gamma_n)_{n\in\mathbb{N}}$ and $(t_k)_{k\in\mathbb{N}}$ and to the functions F and R converges almost surely. Moreover, if we write $\mathcal{P}$ for the σ-algebra generated by the construction of $(w^j_k)_{j\in\mathbb{N},k\in\mathbb{N}}$, $(j)_{j\in\mathbb{N}}$ and $(m_j)_{j\in\mathbb{N}}$, then $(\gamma_{2k})_{k\in\mathbb{N}}$ is independent of $\mathcal{P}$ and for all $n \in \mathbb{N}$, the distribution of γ^w_{2n+1} knowing $\mathcal{P}$ and $(\gamma_{2k})_{k\in\mathbb{N}}$ and $(\gamma^w_{2n'+1})_{n'\neq n}$ are absolutely continuous with respect to η and $(\frac{\rho}{1-2\rho})$-Schottky.

Proof. Let j be a step of the algorithm, Write $\mathcal{P}_j$ the σ-algebra generated by the construction of the pivotal extraction up to step j. We have:

$$\mathbb{E}\left(F(\gamma^{w^j}_{2m_j}, \gamma_{2j+1}, \gamma_{2j+2}) \,\middle|\, (\gamma_{2k})_{k\in\mathbb{N}}, (\gamma_{2k+1})_{k\neq j}, \mathcal{P}_j\right) = 1 - 2\rho$$
$$\mathbb{P}\left(t_{2j+1} < F(\gamma^{w^j}_{2m_j}, \gamma_{2j+1}, \gamma_{2j+2}) \,\middle|\, (\gamma_{2k})_{k\in\mathbb{N}}, (\gamma_{2k+1})_{k\neq j}, \mathcal{P}_j\right) = 1 - 2\rho \tag{54}$$

It means that we are in (1) of Definition 4.20 with probability $1 - 2\rho$. Moreover, for all $l \leq m_j$, if we write $2k + 1 = \overline{w}^j_{2l+2}$, and $\gamma'_l := \gamma^{w^j}_{2l+2} \cdots \gamma^{w^j}_{2m_j}\gamma_{2j+1}\gamma_{j+2}$ we have:

$$\mathbb{P}\left(t_{2k+1} < R\left(\gamma^{w^j}_{2l}, \gamma_{2k+1}, \gamma_{2k+2}, \gamma'_l\right) \,\middle|\, (\gamma_n)_{n\neq 2k}, \mathcal{P}_j\right) = \frac{1 - 3\rho}{1 - 2\rho}. \tag{55}$$

Indeed, knowing $\mathcal{P}_j$ and knowing that $2k+1$ is a pivotal time at time j, we have $t_k < F\left(\gamma^{w^j}_{2l}, \gamma_{2k+1}, \gamma_{2k+2}\right)$ and $\mathbb{E}\left(\frac{R}{F} \,\middle|\, F > 0\right) = \frac{1-3\rho}{1-2\rho}$. Those three estimate determine the law of $(w^j_k)_{k\in\mathbb{N}}$ and j and m_j for all step j and they do not depend on the sequence $(\gamma_{2k})_{k\in\mathbb{N}}$. For all $k \in \mathbb{N}$, we take j_k to be the last time such that $m_{j_k} = k$. Note that $w_0 = 2r_0 - 1$ and in fact for all k, we have $\overline{w}_{2k} = 2r_k - 1$ and $w_{2k} = 2(r_{k+1} - r_k) - 1$. Now we need to prove that all of this is well defined, for that, we need to show that m_j satisfies large deviations inequalities around a positive escape speed in the sense of Definition B.27. Note that conditionally to m_j, the possible values for m_{j+1} are $\{0, \dots, m_j - 1, m_j + 1\}$. By (54), we have $\mathbb{P}(m_{j+1} = m_j + 1|m_j) = 1 - 2\rho$. By (55), we have for all value of m_j:

$$\forall 0 < k < m_j, \qquad \mathbb{P}(m_{j+1} < k \,|\, m_{j+1} < k + 1, m_j, \dots, m_0) = \frac{\rho}{1 - 2\rho}$$
$$\forall 0 \leq k < m_j, \qquad \mathbb{P}(m_{j+1} \leq k \,|\, m_j, \dots, m_0) = 2\rho\left(\frac{\rho}{1 - 2\rho}\right)^{m_j - 1 - k}$$

Now we compute the conditional expectation for some $j \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{E}(m_{j+1}\,|\,m_j,\dots,m_0) &= m_j + \mathbb{P}(m_{j+1} = m_j+1) - \sum_{k=0}^{m_j-1}(m_j-k)\mathbb{P}(m_{j+1}=k)\\
&= m_j + \mathbb{P}(m_{j+1} = m_j+1) + \sum_{k=0}^{m_j-1}\mathbb{P}(m_{j+1}\leq k)\\
&= m_j + 1 - 2\rho - 2\rho\sum_{k=0}^{m_j-1}\left(\frac{\rho}{1-2\rho}\right)^{m_j-1-k}\\
&= m_j + 1 - 2\rho - 2\rho\frac{1-2\rho}{1-3\rho}\left(1-\left(\frac{\rho}{1-2\rho}\right)^{m_j}\right)\\
&= m_j + \frac{1-2\rho}{1-3\rho}(1-5\rho) + \left(\frac{\rho}{1-2\rho}\right)^{m_j} 2\rho\frac{1-2\rho}{1-3\rho}\\
&\geq m_j + \frac{1-2\rho}{1-3\rho}(1-5\rho)
\end{aligned}$$

By assumption $\rho < \frac{1}{5}$ so $\sigma_\rho := \frac{1-2\rho}{1-3\rho}(1-5\rho) > 0$. then by Lemma B.26, the sequence $(m_j)_j$ satisfies large deviations inequalities below the speed σ_ρ which is positive. As a consequence, w_0 has finite exponential moment. Now with the same reasoning as above, we get for all $k \in \mathbb{N}$, and for all $j \in \mathbb{N}$:

$$\mathbb{E}(m_{j+1+j_{k-1}}\,|\,m_{j+j_{k-1}},\dots,m_0) = m_j + \frac{1-2\rho}{1-3\rho}(1-5\rho) + \left(\frac{\rho}{1-2\rho}\right)^{m_j-k} 2\rho\frac{1-2\rho}{1-3\rho}.$$

This means that w_{2k} has the same distribution law as w_0 for all k so the $(w_{2k})_{k\geq 0}$ are i.i.d and have finite exponential moment by the large deviations inequality for $\alpha = 0 < \sigma_\rho$. Then note that for all j, the distribution of $\gamma_{2n+1}^{w^j}$ knowing $\mathcal{P}_j$ and $(\gamma_{2k})_{k\in\mathbb{N}}$ and $(\gamma_{2k+1}^{w^j})_{k\neq n}$ is simply the normalised restriction of η to the set $\left\{\gamma \in \Gamma \,\middle|\, \gamma_{2n}^{w^j}\mathbb{A}\gamma\mathbb{A}\gamma_{\overline{w}^j_{2n+2}}\right\}$ so it is bounded by $\frac{\eta}{1-2\rho}$. Therefore it is $\frac{\rho}{1-2\rho}$-Schottky. $\square$

Lemma 4.22. *Let ν be a strongly irreducible and proximal distribution over $\Gamma := \mathrm{End}(E)$. Let $N : \Gamma \to [0,+\infty]$ be a measurable map such that $N(\nu)$ is almost surely finite. Let $0 < \rho < \frac{1}{3}$. There are constants $\varepsilon' > 0$ such that for all $\lambda > 0$, there is an exponentially integrable and eventually irreducible extraction $(X,\tilde{\kappa})$ of $(\{*\},\nu)$ and three constants $l_a \in \mathbb{N}$ and $B \in \mathbb{R}_{\geq 0}$ and $C \in \mathbb{R}_{\geq 0}$ and a non-empty family of edges $\mathcal{S} \subset \mathcal{E}(X,\tilde{\kappa})$ such that, we have:*

1. *The Markov bundle $(X,\tilde{\kappa})$ is λ-proximal and $\mathbb{A}^{\frac{\varepsilon'}{2}}$ aligned.*
2. *For all $(a:b) \in \mathcal{S}$, the product distribution $\kappa_{a,b}$ is ρ Schottky for $\mathbb{A}^{\frac{\varepsilon}{2}}$ and almost all Markov chain in $(X,p_{\tilde{\kappa}})$ reaches $(a:b)$.*
3. *Given $(a:b) \in \mathcal{S}$ and a random word $\tilde{h} \sim \tilde{\kappa}_{a,b}$, the length $L(\tilde{h}) = l_a$ and for all $0 \leq k < l_a$, we have $N\left(\chi_k\left(\tilde{h}\right)\right) \leq B$.*
4. *For all $(c:d) \in \mathcal{E}(X,\tilde{\kappa})$ and for $\tilde{h} \sim \tilde{\eta}_{c,d}$ a random word, we have:*

$$\forall t \geq B,\ \forall l > k,\ \mathbb{P}\left(N\left(\chi_k(\tilde{g})\right) > t \,|\, L(\tilde{g}) = l\right) \leq CN(\nu)(t,+\infty).$$

Proof. Let $\rho' := \frac{\rho}{1+2\rho} < \frac{1}{5}$. Let $\varepsilon > 0$ and $\lambda \geq 4|\log(\varepsilon)| + 4\log(2)$ and B be as in Lemma 4.14. Let $(\{a,b\},\tilde{\eta})$ be the extraction of ν that we constructed in Lemma 4.14. Consider $(\tilde{g}_k)$ an ornamented Markov chain in $(\{a,b\},\tilde{\eta})$ that starts at b. For all k, we write $g_k := \Pi(\tilde{g}_k)$ and $l_k := L(\tilde{g}_k)$ and we write $(\gamma_n) := \bigodot_{k\in\mathbb{N}}\tilde{g}_k \sim \nu^{\otimes\mathbb{N}}$. Then construct $\mathbb{A}$ to be a discrete alignment relation such that $\mathbb{A}^\varepsilon \subset \mathbb{A} \subset \mathbb{A}^{\frac{\varepsilon}{2}}$ as described in 2.24. Write $\mathbb{A} = \bigsqcup_{(i,j)\in A} L_i \times R_j$ for $(L_i)_{1\leq i\leq M}$ and $(R_j)_{1\leq j\leq M}$ two measurable partitions of $\Gamma := \mathrm{End}(E)$. Then write (w_k) for the extraction constructed in Definition 4.20 for $\eta := \Pi\eta_a$ and for the sequence $(g_k)_{k\in\mathbb{N}}$ and for $(t_n) \in [0,1]^{\mathbb{N}}$ taken uniformly and independently of $(\tilde{g}_k)$. Then by Lemma 4.21, the sequence (w_k) is well defined and independent of $(\tilde{g}_{2k})_{k\in\mathbb{N}}$. So by (45), we have:

$$\forall t \geq B,\ \mathbb{P}\left(N(\gamma_k) > t \,|\, (\gamma_{k'})_{k'\neq k}, (w_n)_{n\in\mathbb{N}}\right) \leq 2N(\nu)(t,+\infty). \tag{56}$$

Now we want the sequence (g_k^w) to be λ squeezing. For that, we shift w to the left We define $w'_{2k} := w_{2k+1}$ and $w'_{2k-1} := 1$ for all $k \geq 1$ and $w'_0 := w_0 + 1 + w_2$. Then $(g_k^{w'})$ is λ-squeezing and $\mathbb{A}^{\frac{\varepsilon}{2}}$-aligned. Then write $X_s := \{s\}$ and $X'_a := \{a_1, \dots, a_M\}$ and $Y'_a := \{a'_1, \dots, a'_M\}$ and $X'_b := \{b_1, \dots, b_M\}$ three abstract sets. Then for all time k, we write (i_k, j_k) for the pair of indices such that $g_{2k}^{w'} \in L_{i_k}$ and $g_{\overline{w}'_{2k}} \in R_{j_k}$. Then we write $x_0 = s$ and $x_1 := a'_{i_0}$ and for all $k \geq 1$, we write $x_{2k} = b_{j_k}$ and $x_{2k+1} := l_{i_k}$. Now note that the $(\tilde{g}_{2k}^{w'})_{k\geq 1}$ are i.i.d by construction. Write $\tilde{\kappa}'_b$ their distribution. Draw $\tilde{g}' \sim \tilde{\kappa}'_b$, write $i(\tilde{g})$ for the index such that $\Pi(\tilde{g}') \in L_{i(\tilde{g}')}$. Then for all $1 \leq i, j \leq M$, write:

$$\begin{aligned} q(i) &:= \mathbb{P}\left(\Pi(\tilde{g}') \in L_i\right) \\ p(j) &:= \mathbb{P}\left(\chi_0(\tilde{g}') \in R_j\right) \end{aligned}$$

Let j be such that $p(j) > 0$, write $\tilde{\kappa}_{b_j}$ for the distribution of $(b_j : \tilde{g} : a_{i(\tilde{g})})$ knowing that $(\chi_0(\tilde{g}) \in R_j)$. Then write $\tilde{\kappa}_s$ for the distribution of $(x_0 : \tilde{g}_0^w : x_1)$ where $x_1 = a'_i$ for $1 \leq i \leq M$ the index such that $g_0^{w'} \in L_i$ and write $s(i) := \mathbb{P}(g_0^{w'} \in L_i)$.

Now note that for all n, the condition $g_{2n}^{w'} \mathbb{A} g_{2n+1}^{w'} \mathbb{A} g_{\overline{w}'_{2n+2}}$ only depends on (i_n, j_{n+1}) so the distribution of $\tilde{g}_{2k+1}^{w'}$ knowing $(w'_k)_{k\in\mathbb{N}}$ and $(\tilde{g}_k)_{k\neq\overline{w}'_n}$ only depends on $(x_{2n+1} : x_{2n+2}) = \left(a_{i_n} : b_{j_{n+1}}\right)$. For all pair $1 \leq i, j \leq M$, we write $\tilde{\kappa}_{a_i, b_j}$ for the normalised restriction of $\tilde{\eta}_a$ to the set:

$$S_{i,j} := \Pi^{-1}\left(\bigsqcup_{(i',j)\in A} \bigsqcup_{(i,j')\in A} L_{i'} \cap R_{j'}\right).$$

Note that $\tilde{\eta}_a(S_{i,j}) \geq 1 - 2\rho'$ for all pair i, j which proves 3 for $\mathcal{S}$. Then write:

$$\tilde{\kappa}_{a_i} := \sum_{1=j}^{M} p(j)(a_i : \tilde{\kappa}_{a_i,b_j} : b_j)$$

Now write $X_b := \{b_j \,|\, p(j) > 0\}$ and $X_a := \{a_i \,|\, q(i) > 0\}$ and $Y_a := \{a_i \,|\, s(i) > 0\}$ and $X := X_s \sqcup Y_a \sqcup X_a \sqcup X_b$. Then the sequence $\left(x_k : \tilde{g}_k^{w'} : x_{k+1}\right)_{k\in\mathbb{N}}$ is an ornamented Markov chain in $(X, \tilde{\kappa})$. Note that it is eventually irreducible. Indeed, for all $a \in X_a$ there is $b \in X_b$ such that $p_{\tilde{\kappa}}(b, a) > 0$ and for all $b \in X_b$ we have $p_{\tilde{\kappa}}(b, a) > 0$ for all $a \in X_a$. Moreover $X_a \cup X_b$ is $p_{\tilde{\kappa}}$-stable so it is irreducible. Then $p_{\tilde{\kappa}}(s, Y_a) = 1$ and $p_{\tilde{\kappa}}(a', X_b) = 1$ for all $a' \in Y_a$ so all orbit ends in $X_b \cup X_a$ after time 3.Write $\mathcal{S} := \{(a : b) \,|\, a \in X_a, b \in X_b\}$ note that we have (3) because for all $a \in X_a$ and all $b \in X_b$, $\tilde{\kappa}_{a,b}$ is absolutely continuous with respect to η_a. Then for all other edge $(c : d)$ note that $\tilde{\kappa}_{c,d}$ is a fraction of the distribution of $\tilde{g}_0^{w'}$ (when $c = s$) or of the distribution of $\tilde{g}_2^{w'}$ (when $c \in X_b$) or the norm of each letter is bounded when $c \in Y_a$. Then by (56), we have:

$$\begin{aligned} &\forall t \geq B,\ \forall k \in \mathbb{N},\ \mathbb{P}\left(N(\gamma_k) > t \,|\, (w'_n), (l_n), (\gamma_{k'})_{k'\neq k}\right) \leq 2N(\nu)(t,+\infty) \\ &\forall t \geq B,\ \forall k \in \mathbb{N},\ \mathbb{P}\left(N(\gamma_{k+\bar{l}_2^{w'}}) > t \,\middle|\, (w_n), (l_n), (\gamma_{k'})_{k'\neq k}\right) \leq 2N(\nu)(t,+\infty). \end{aligned}$$

Note that there are only finitely many possibilities for (x_1, x_2, x_3). We write:

$$\delta := \min\{\mathbb{P}(x_1 = y_1, x_2 = y_2, x_3 = y_3) \,|\, y_1, y_2, y_3 \in X, \mathbb{P}(x_1 = y_1, x_2 = y_2, x_3 = y_3) > 0\}$$

and we have $\delta > 0$ because it is the minimum of finitely many positive numbers. Then we have for all $a' \in Y_a$ and all $b \in X_b$ and all $a \in X_a$ such that $p_{\tilde{\kappa}}(b, a) > 0$:

$$\forall k \in \mathbb{N},\ \mathbb{P}\left(N(\gamma_k) > t \,|\, (w'_n), (l_n), x_1 = a'\right) \leq \frac{2}{\delta} N(\nu)(t,+\infty) \tag{57}$$

$$\forall k \in \mathbb{N},\ \mathbb{P}\left(N(\gamma_{k+\bar{l}_2^{w'}}) > t \,\middle|\, (w'_n), (l_n), x_2 = b, x_3 = a\right) \leq \frac{2}{\delta} N(\nu)(t,+\infty). \tag{58}$$

This proves (4) for $C := \frac{2}{\delta}$. Indeed, (57) implies (4) when $\tilde{g} \sim \kappa_{s,a'}$ and (58) implies (4) when $\tilde{g} \sim \kappa_{b,a}$. Then for all $a \in X_a \cup Y_a$ and $b \in X_b$ the distribution of $\tilde{\kappa}_{a,b}$ is bounded by $\frac{\tilde{\eta}_a}{1-2\rho}$ so it is $\frac{\rho}{1+2\rho}$-Schottky. $\square$

Now we can conclude the proof.

proof of Theorem 4.9. Consider $\varepsilon' > 0$, as in Lemma 4.22 and write $\varepsilon := \frac{\varepsilon'}{4}$. Then for all $\lambda \geq 4|\log(\varepsilon)| + 4\log(2)$, write $(X, \tilde{\kappa}')$ for the extraction constructed in Lemma 4.22 with the constants B', l_a. By Lemma 2.18, the bundle $(X, \tilde{\kappa}')$ is strongly $\mathbb{A}^{\varepsilon}$-aligned. Write $B = B'$. Then draw $(x_k : \tilde{g}_k : x_{k+1})$ an ornamented Markov chain in $(X, \tilde{\kappa}')$ and write $(\gamma_k) := \bigodot \tilde{g}_k$ and $l_k := L(\tilde{g}_k)$ for all k. Then take an edge $(a : b) \in \mathcal{S}$. Write $\tilde{\kappa}_a := \tilde{\kappa}'_{a,b}$ which is ρ-Schottky by (3). Then write k_0 for the first time such that $(x_{k_0} : x_{k_0+1}) = (a : b)$ and write $\tilde{\kappa}_s$ for the distribution of $\tilde{g}_0 \odot \cdots \odot \tilde{g}_{k_0}$ Then write $k_1 > k_0$ for the second time such that $(x_{k_1} : x_{k_1+1}) = (a : b)$ and write $\tilde{\kappa}_b$ for the distribution of $\tilde{g}_{k_0+2} \odot \cdots \odot \tilde{g}_{k_1}$. Note that k_0 and k_1 both have finite exponential moment by B.22. Then the distribution of $(x_{k_0} : \tilde{g}_{k+k_0} : x_{k+k_0+1})$ is the same as $(x_{k_1} : \tilde{g}_{k+k_0} : x_{k+k_1+1})$ so we have:

$$\tilde{\kappa}_s \odot (\tilde{\kappa}_a \odot \tilde{\kappa}_b)^{\odot \mathbb{N}} = \nu^{\otimes \mathbb{N}}.$$

This proves that $(X, \tilde{\kappa})$ is an extraction of $(\{*\}, \nu)$. By Lemma 4.22, for all $m \in \mathbb{N}$, we have:

$$\forall t \geq B,\ \mathbb{P}\left(N(\gamma_m) > t \,|\, (x_k)_{k\in\mathbb{N}}, (l_k)_{k\in\mathbb{N}}, (\gamma_{m'})_{m'\neq m}\right) \leq CN(\nu)(t, +\infty).$$

Moreover the sequence $(k_n)_{n\in\mathbb{N}}$ only depends on $(x_k)_{k\in\mathbb{N}}$ so we have:

$$\forall t \geq B,\ \mathbb{P}\left(N(\gamma_n) > t \,|\, (k_n)_{k\in\mathbb{N}}, (l_k)_{k\in\mathbb{N}}, (\gamma_{m'})_{m'\neq m}\right) \leq CN(\nu)(t, +\infty)$$
$$\forall t \geq B,\ \mathbb{P}\left(N(\gamma_{n+\bar{l}_{k_0+1}}) > t \,\middle|\, (k_n)_{n\in\mathbb{N}}, (l_k)_{k\in\mathbb{N}}, (\gamma_{n'})_{n'\neq n}\right) \leq CN(\nu)(t, +\infty).$$

This proves (38). $\square$

5 Proof of the results

In this section E is a standard vector space of finite dimension d over a local field $\mathbb{K}$ and Γ is the semi-group of endomorphisms over E.

5.1 Asymptotic estimates for the singular gap

In this section, we define the escape speed of a random product of matrices using the formalism of pivotal extractions. First we show that we have large deviations from below as long as there is a pivotal, squeezing and strongly aligned extraction (which is always the case for strongly irreducible and proximal distributions according to Theorem 4.9). Then we show that the upper bound of all the numbers $\alpha > 0$ such that $\mathbb{P}(\mathrm{sqz}(\overline{\gamma}_n) \leq \alpha_n)$ is exponentially small is in fact the almost sure limit of $\frac{\mathrm{sqz}(\overline{\gamma}_n)}{n}$.

Lemma 5.1 (Large deviations on a Markov bundle over $\mathbb{R}$)**.** *Let (X, ν) be a Markov bundle over $\mathbb{R}$ that is eventually irreducible. Assume that there are constants $C, \beta > 0$ such that $\mathbb{E}(\exp(-\beta\nu_x)) \leq C$ for all $x \in X$. Write ξ for the p_ν-invariant measure on X and write $\sigma := \sum_{x\in X} \xi(x)\mathbb{E}(\nu_x)$. Then for (γ_n) an ornamented Markov chain in (X, ν), the sequence $(\overline{\gamma}_n)$ satisfies large deviations inequalities below the speed σ in the sense of Definition B.27.*

Proof. For all e be an edge in the support of ξ, write $\xi(e) := \xi(\theta_0(e))p_\nu(e)$ the invariant measure of e. Then we have $\sigma := \sum_{e\in\mathcal{E}(X)} \xi(e)\mathbb{E}(\nu_e)$. We write $(\overline{w}^e_k)_{k\geq 1}$ the sequence of times such that $\theta\left(\gamma_{\overline{w}^e_k}\right) = e$. Then write $\overline{w}^e_k = w^e_0 + \cdots + w^e_{k-1}$ with $w^e_0 := \overline{w}^e_1$ the time of first visit in x a random variable that has finite exponential moment and $(w^e_k)_{k\geq 1}$ the length of the consecutive loops of the chain $(\theta(\gamma_n))$ around e, which is an i.i.d. sequence of integers that have finite exponential moment and expectation $\frac{1}{\xi(e)}$. For all time n, write $\overline{r}^e_n := \#\{k < n \,|\, \theta(\gamma_k) = x\}$. Then by Lemma B.26 applied to $(w^e_k)_{k\geq 1}$, the sequence $(\overline{w}^e_k - w^e_0)_{k\in\mathbb{N}}$ satisfies large deviations inequalities above and below the speed $\frac{1}{\xi(e)}$. Moreover w^e_0 is no,n-negative and has finite exponential moment so $(\overline{w}^e_k)_{k\in\mathbb{N}}$ satisfies large deviations inequalities around the speed $\frac{1}{\xi(e)}$. Then by Lemma B.29 applied to $(\overline{w}_k)$, the sequence (r^e_n) satisfies large deviations inequalities around the speed $\xi(e)$.

Let e be an edge in the support of ξ, we define the sequence (S^e_n) by $S^e(n) := \sum_{k=1}^n \gamma_{\overline{w}^e_k}$ for all n. Then $(S^e(n))_{n\in\mathbb{N}}$ satisfies large deviations inequalities around the speed $\mathbb{E}(\nu_e)$. Let t_0 be the first time such that $\xi(\theta(\gamma_{t_0})) > 0$. The support of ξ is eventually reached by definition so t_0 has finite exponential moment and by Lemma B.24 the negative part of $\overline{\gamma_{t_0}}$ has finite exponential moment. Moreover, for all $t \geq t_0$, we have $\xi(\theta(\gamma_t)) > 0$ so for all $n \in \mathbb{N}$, we have the decomposition:

$$\overline{\gamma_n} = \overline{\gamma}_{t_0} + \sum_{e\in\mathbf{supp}(\xi)} S^e(\overline{r}^e_n). \tag{59}$$

So by Lemma B.29, the sequence $(\overline{\gamma_n})$ satisfies large deviations inequalities below the speed σ. □

Definition 5.2. *Let $(Y,\tilde{\kappa})$ be an eventually irreducible Markov bundle over a category of words. Let ξ be the invariant probability distribution over Y, we write:*

$$L(Y,\tilde{\kappa}) := \sum_{y\in Y} \xi\{y\}\mathbb{E}(L(\tilde{\kappa}_y)). \tag{60}$$

Lemma 5.3 (Escape speed of a strongly aligned bundle)**.** *Let $\varepsilon > 0$ and let (X,ν) be a strongly $\mathbb{A}^\varepsilon$-aligned, and irreducible Markov bundle. There is a constant $\sigma(X,\nu) \in [0,+\infty]$ such that for (γ_n) an ornamented Markov chain in (X,ν), we have $\frac{\mathrm{sqz}(\overline{\gamma}_n)}{n} \to \sigma(X,\nu)$ and:*

$$\forall \alpha < \sigma(X,\nu), \exists C,\beta > 0, \forall n \in \mathbb{N}, \mathbb{P}(\mathrm{sqz}(\overline{\gamma}_n) \leq \alpha n) \leq C\exp(-\beta n). \tag{61}$$

If moreover (X,ν) is λ-proximal for some $\lambda > 2|\log(\varepsilon)|$, then $\sigma(X,\nu) \geq \lambda - 2|\log(\varepsilon)| > 0$.

Proof. The fact that (X,ν) is strongly $\mathbb{A}^\varepsilon$-aligned means that for almost every ornamented Markov chain (γ_n) in (X,ν), and every $N \in \mathbb{N}\setminus\{0\}$, by Lemma 2.18 (17), we have:

$$\mathrm{sqz}(\overline{\gamma}_n) \geq \sum_{k=0}^{\lfloor \frac{n}{N}\rfloor} \mathrm{sqz}\left(\gamma_{kN}\cdots\gamma_{(k+1)N-1}\right) + \mathrm{sqz}\left(\gamma_{\lfloor\frac{n}{N}\rfloor N}\cdots\gamma_{n-1}\right) - \frac{2n|\log(\varepsilon)|}{N}. \tag{62}$$

For every integer $N \geq 0$ that is co-prime with the period of (X,ν), (*i.e,* such that $\left(X,\nu^{*N}\right)$ is eventually irreducible), we define:

$$\sigma_N := \frac{1}{N}\sum_{x_0\in X} \mathbb{E}\left(\mathrm{sqz}\left(\nu^{*N}_{x_0}\right)\right)\xi(x_0).$$

Then by Lemma 5.1 and (62) for any ornamented Markov chain (γ_n) in (X,ν), one has:

$$\forall \alpha < \sigma_N - \frac{2|\log(\varepsilon)|}{N}, \exists C,\beta > 0,\ \mathbb{P}(\mathrm{sqz}(\overline{\gamma}_{nN}) \leq \alpha n) \leq C\exp(-\beta n).$$

Then since (γ_n) is an aligned sequence, $\mathrm{sqz}(\overline{\gamma}_n)$ is increasing with n, so we have:

$$\forall \alpha < \sigma_N - \frac{2|\log(\varepsilon)|}{N}, \exists C,\beta > 0,\ \mathbb{P}(\mathrm{sqz}(\overline{\gamma}_n) \leq \alpha n) \leq C\exp(-\beta n).$$

Then by taking the expectation of the above inequality, we get:

$$\begin{aligned}
\sigma_{NM} &\geq \sigma_N - \frac{2|\log(\varepsilon)|}{N} && \forall M,N\geq 1\\
\sigma_M &\geq \sigma_N(1-N/M) - \frac{2|\log(\varepsilon)|}{N} && \forall M\geq N\geq 1\\
\liminf_{M\to\infty}\sigma_M &\geq \sigma_N - \frac{2|\log(\varepsilon)|}{N} && \forall N\geq 1\\
\liminf_{M\to\infty}\sigma_M &\geq \limsup_{N\to\infty}\sigma_N.
\end{aligned}$$

As a consequence the limit $\sigma(X,\nu) := \lim_{N\to\infty} \sigma_N$ exists in $[0,+\infty]$ and we have (61). Then to show that we have almost sure convergence, we use Kingmann's sub-additive ergodic Theorem [Kin68]. We consider the space Ω of ornamented Markov chains in (X,ν) whose starting point has distribution ξ and $\theta : \Omega \to \Omega$ the shift map. For all $\omega = (x_n : \gamma_n : x_n+1)_{n\in\mathbb{N}}$, we write $F_n(\omega) := -\mathrm{sqz}(\overline{\gamma}_n) - 2\log(\varepsilon)$. Then by Lemma 2.18, the sequence (F_n) is θ sub-additive in the sense that for all m,n we have $F_{n+m} \leq F_n + F_m \circ \theta^n$. Moreover, θ is ergodic so $\frac{F_n}{n}$ almost surely converges to a constant limit by the sub-additive ergodic theorem. Moreover by definition, we have $\frac{\mathbb{E}(F_n)}{n} \to \sigma(X,\nu)$ so the limit in question is $\sigma(X,\nu)$. □

Lemma 5.4 (alignment on both sides)**.** *Let $\varepsilon > 0$, let $\rho < 1$ and let $\lambda \geq 2|\log(\varepsilon)| + 2\log(4)$. Let (X,κ) be a ρ-ping-pong $\mathbb{A}^\varepsilon$-aligned and λ-proximal Markov bundle over Γ. Let $(x_n : \gamma_n : x_{n+1})_{n\in\mathbb{N}}$ be an ornamented Markov chain in (X,κ) and $g_n, h_n \in \mathrm{End}(E)$ be non-random sequences (or independent of (γ_n)). There are two sequences of integers (l_n) such that for all n, one of the following holds:*

- *we have $(g_n\gamma_0\cdots\gamma_{l_n-1})\mathbb{A}^\varepsilon\gamma_{l_n}$ and $\gamma_{n-r_n-1}\mathbb{A}^\varepsilon(\gamma_{n-r_n}\cdots\gamma_{n-1}h_n)$.*
- *$r_n + l_n \geq n$.*

Moreover l_n is a stopping time i.e, , conditionally to l_n the sequence $(\gamma_{l_n+1+k})_{k\geq 0}$ is an ornamented Markov chain in (X,ν). Moreover, l_n and r_n have bounded exponential moment conditionally to the even times i.e, there are constants $C,\beta > 0$ such that :

$$\forall n\in\mathbb{N}, \forall t\in\mathbb{N}, \mathbb{P}\left(l_n + r_n \geq t | (\gamma_{2k})_{k\in\mathbb{N}}, g_n, h_n\right) \leq C\exp(-\beta k). \tag{63}$$

Proof. Consider an integer $n\in\mathbb{N}$. We define the forward filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ as the family of σ-algebras initialized with $\mathcal{F}_0 := \langle(\gamma_{2k})_{k\in\mathbb{N}}, g_n, h_n\rangle$ and with the induction relation $\mathcal{F}_{k+1} := \langle\gamma_k, \mathcal{F}_k\rangle$ for all $k\in\mathbb{N}$. If there is an odd integer $l \leq n$ such that $(g_n\gamma_0\cdots\gamma_{l-1})\mathbb{A}^\varepsilon\gamma_l$, then we take l_n to be the smallest such l, otherwise, we take $l_n = n$. Then l_n is a stopping time for $(\mathcal{F}_k)$ because the event $(l_n < l)$ only depends on $(g_n,\dots,\gamma_{l-1})$ so it is $\mathcal{F}_l$-measurable. In particular, knowing $(l_n > l)$ for some odd $l \leq n-2$, the distribution of γ_{l+2} is still ρ-Schottky so we have:

$$\forall l \leq n-2, \mathbb{P}(l_n > l+2 | l_n > l) \leq \rho. \tag{64}$$

The above implies that the distribution of l_n is dominated by a geometric distribution so it has exponential moment. Then to define r_n, we use the same technique, note that drawing an ornamented Markov chain in a ping-pong bundle amounts to drawing $\gamma_0 \sim \kappa_s$ and then $(\gamma_{2k-1},\gamma_{2k}) \sim \kappa_a\otimes\kappa_b$ independently for all $k\geq 1$. We draw the sequence backwards. We define a backwards filtration $(\mathcal{B}_k)$ by $\mathcal{B}_0 := \langle(\gamma_{2k})_{k\in\mathbb{N}}, g_n, h_n\rangle$ and $\mathcal{B}_{k+1} := \langle\gamma_{n-k}, \mathcal{B}_k\rangle$ for all $k\in\mathbb{N}$. Then we define r_n as the smallest integer that has the parity of n and such that $\gamma_{n-r_n-1}\mathbb{A}^\varepsilon(\gamma_{n-r_n}\cdots\gamma_{n-1}h_n)$ or $r_n = n$ if there is no such r. The same reasoning as above holds and we have:

$$\forall l \leq n-2, \mathbb{P}(r_n > l+2 | r_n > l) \leq \rho. \tag{65}$$

Then by (64) and (65), and by lemma B.36, we have for all $t\in\mathbb{N}$:

$$\mathbb{P}\left(l_n + r_n \geq t | (\gamma_{2k})_{k\in\mathbb{N}}, g_n, h_n\right) \leq 2\rho^{\frac{t-2}{4}}.$$

Moreover, the bound on right does not depend on n, which proves (63). □

Lemma 5.5. *Let ν be a probability distribution on Γ. Let $\varepsilon > 0$, let $\rho < 1$ and let $\lambda \geq 2|\log(\varepsilon)| + 2\log(4)$. Let $(X,\tilde{\kappa})$ be an exponentially integrable extraction of $(\{*\},\nu)$. Assume that (X,κ) be a ρ-ping-pong $\mathbb{A}^\varepsilon$-aligned and λ-proximal Markov bundle over Γ. Assume also that $L(\tilde{\kappa}_a)$ is a constant. Let $(\gamma_n)_{n\in\mathbb{N}} \sim \nu^{\otimes\mathbb{N}}$ and let (w_n) be a sequence of random integers such that $(\widetilde{\gamma}^w_n)_{n\in\mathbb{N}}$ be an ornamented Markov chain in (X,κ). Then we have:*

$$\forall\alpha < \frac{\sigma(X,\kappa)}{L(X,\tilde{\kappa})}, \exists C,\beta > 0, \forall n\in\mathbb{N}, \mathbb{P}(\mathrm{sqz}(\overline{\gamma}_n \leq \alpha n) \leq C\exp(-\beta n). \tag{66}$$

Proof. We use the language of large deviations as in Definition B.27. Consider an integer time $n\in\mathbb{N}$, we write m_n for the largest integer such that $\overline{w}_{m_n} \leq n$. The sequence of random integers (m_n) satisfies large deviations inequalities around the speed $\frac{1}{L(X,\tilde{\kappa})}$ by Lemma B.26 applied to $(w_{2n})_{n\geq 1}$ and Lemma B.29 with

(1) applied to w_0 and (6) to switch to m_n. We also define two integers l_n and r_n as in Lemma 5.4 so that $\gamma^w_{l_n}\mathbb{A}^\varepsilon\gamma^w_{l_n}$ and $\gamma^w_{m_n-r_n-1}\mathbb{A}^\varepsilon(\gamma_{\overline{w}_{m_n-r_n}}\cdots\gamma_{n-1})$. Then by Lemma 5.4 l_n and r_n both have bounded exponential moment and by Lemma B.29 (1), the sequence $(m_n - l_n - r_n - 2)_n$ satisfies large deviations inequalities around the speed $\frac{1}{L(X,\tilde{\kappa})}$. Since l_n is a stopping time, by Lemma 5.3, the sequence $\left(\mathrm{sqz}(\widetilde{\gamma}^w_{l_n+1}\cdots\widetilde{\gamma}^w_{l_n+2+k})\right)_k$ satisfies large deviations inequalities below the speed $\sigma(X,\kappa)$. Moreover, by Lemma 2.18, we have for all n:

$$\mathrm{sqz}(\overline{\gamma}_n) \geq \mathrm{sqz}(\widetilde{\gamma}^w_{l_n+1}\cdots\widetilde{\gamma}^w_{m_n-r_n-2}). \tag{67}$$

Then by Lemma B.29 (5), the term on the right of (67) satisfies large deviations inequalities below the speed $\frac{\sigma(X,\kappa)}{L(X,\tilde{\kappa})}$ and so does $(\mathrm{sqz}(\overline{\gamma}_n))_n$. □

Lemma 5.6. *Let ν be a strongly irreducible and proximal probability distribution over $\Gamma := \mathrm{End}(E)$. Let $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$. There are constants $\varepsilon > 0$, $l \in \mathbb{N}$ and $C > 0$ that depend on ν and such that for all $g \in \Gamma \setminus \{0\}$, we have:*

$$\mathbb{P}(\forall n \geq l, g\mathbb{A}^\varepsilon\overline{\gamma}_n) \geq C. \tag{68}$$

Proof. Consider $\rho := \frac{1}{3}$. Apply Theorem 4.9 to get some $\varepsilon > 0$, some $\lambda \geq 4|\log(\varepsilon)| + 4\log(2)$ and an extraction $(\{s,a,b\},\tilde{\kappa})$ of ν whose product is ρ-ping-pong $\mathbb{A}^\varepsilon$-aligned and λ-squeezing and write l_0 the constant length of $L(\widetilde{\kappa}_{a,b})$. Then $\tilde{\kappa}_{a,b}$ is absolutely continuous with respect to $\widetilde{\nu}^{\odot l_0}$ so there is a constant $\delta > 0$ such that $\delta\tilde{\kappa}_{a,b} \leq \widetilde{\nu}^{\odot l_0} = \nu^{\otimes l_0}$. Then draw the sequence $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$ as follows. With probability δ, draw $(\gamma_n)_{n\geq 0} \sim \tilde{\kappa}_{a,b} \otimes \nu^{\otimes\mathbb{N}}$ and with probability $(1-\delta)$ draw $(\gamma_n)_{n\geq 0} \sim \frac{1}{1-\delta}(\nu^{\otimes l_0} - \delta\tilde{\kappa}_{a,b}) \otimes \nu^{\otimes\mathbb{N}}$. Write $g_n := \gamma_{n+l_0}$ for all $n \in \mathbb{N}$. Then write (w_k) a sequence of integers independent of $(\gamma_0,\dots,\gamma_{l_0-1})$ such that $(\widetilde{g}^w_n)$ is an ornamented Markov chain in $(\{s,a,b\},\tilde{\kappa})$. Then for all $n \geq l_0$, write m_n the largest integer such that $\overline{w}_{m_n} \geq n - l_0$ and write r_n for the smallest integer such that $g^w_{m_n-r_n-1}\mathbb{A}^\varepsilon(\gamma_{\overline{w}_{m_n-r_n}+l_0}\cdots\gamma_{n-1})$. Then $(m_n - r_n)$ satisfies large deviations inequalities below a positive speed so we have constants $C,\beta > 0$ such that $\mathbb{P}(m_n - r_n \leq 0) \leq C\exp(-\beta n)$ for all $n \in \mathbb{N}$. In particular, there is an integer $l_1 \geq l_0$ such that:

$$\mathbb{P}\left(\exists n \geq l_1,\ m_n - r_n \leq 0\right) \leq \frac{1}{2}. \tag{69}$$

Moreover, $\kappa_{a,b}$ is ρ-Schottky, so we have

$$\mathbb{P}\left(g\mathbb{A}^\varepsilon\overline{\gamma}_{l_0}\mathbb{A}^\varepsilon g^w_0 \middle| (g_n),(w_n)\right) \geq \delta(1-2\rho). \tag{70}$$

Moreover, when $m_n - r_n \geq 0$ and $g\mathbb{A}^\varepsilon\overline{\gamma}_{l_0}\mathbb{A}^\varepsilon g^w_0$, we have $g\mathbb{A}^{\frac{\varepsilon}{2}}\overline{\gamma}_n$ so we have:

$$\mathbb{P}(\forall n \geq l_1, g\mathbb{A}^{\frac{\varepsilon}{2}}\overline{\gamma}_n) \geq \frac{\delta}{2}(1-2\rho).$$ □

Remark 5.7. *Up to taking the transpose of ν, the conclusion of Lemma 5.6 also holds for the right to left product and we also have:*

$$\mathbb{P}(\forall n \geq l, (\gamma_{n-1}\cdots\gamma_0)\mathbb{A}^\varepsilon g) \geq C. \tag{71}$$

Theorem 5.8 (Escape speed of the singular gap)**.** *Let ν be a strongly irreducible and proximal probability distribution over $\Gamma := \mathrm{End}(E)$. Let $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$. Define $\sigma(\nu)$ to be the essential supremum of $\limsup \frac{\mathrm{sqz}(\overline{\gamma}_n)}{n}$. Then the random sequence $(\mathrm{sqz}(\overline{\gamma}_n))_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed $\sigma(\nu)$. As a consequence $\frac{\mathrm{sqz}(\overline{\gamma}_n)}{n} \to \sigma(\nu)$ almost surely.*

Proof. We start with the $\mathbb{A}^\varepsilon$ aligned and λ-squeezing ρ-ping-pong extraction $(X,\tilde{\kappa})$ constructed in Theorem 4.9 for some $\varepsilon > 0$, some $\lambda > 2\log(\varepsilon) + 4\log(2)$ and some $\rho < \frac{1}{2}$. Assume also that ε is small enough to apply Lemma 5.6. Draw two random sequences $(\overline{\gamma}_n) \sim \nu^{\otimes\mathbb{N}}$ and (w_n) such that $(\widetilde{\gamma}^w_n)$ is an ornamented Markov chain in $(X,\tilde{\kappa})$. By lemma 5.3, we have $\frac{\mathrm{sqz}(\gamma^w_1\cdots\gamma^w_k)}{k} \to \sigma(X,\kappa)$ with $\sigma(X,\kappa) > 0$ a constant. By Lemma 5.5, the sequence $(\mathrm{sqz}(\overline{\gamma}_n))_n$ satisfies large deviations inequalities below the speed $\frac{\sigma(X,\kappa)}{L(X,\tilde{\kappa})}$ so $\sigma(\nu) \geq \frac{\sigma(X,\kappa)}{L(X,\tilde{\kappa})}$.

We now need to show that $\sigma(\nu) \leq \frac{\sigma(Y,\tilde{\eta})}{L(Y,\tilde{\eta})}$, consider some $\alpha < \sigma(\nu)$ and $\delta > 0$. Note that by definition of the essential supremum, we have $\mathbb{P}\left(\limsup \frac{\mathrm{sqz}(\overline{\gamma}_n)}{n} \geq \alpha\right) > 0$. Write $C_\alpha := \mathbb{P}\left(\limsup \frac{\mathrm{sqz}(\overline{\gamma}_n)}{n} \geq \alpha\right)$. Then by Lemma 5.6, up to taking ε small enough, there is an integer l and a constant C_l such that:

$$\forall m \in \mathbb{N},\ \mathbb{P}\left(\forall m' \geq m + l,\ \overline{\gamma}_m\mathbb{A}^\varepsilon(\gamma_m\cdots\gamma_{m'-1})\,|\,(\gamma_n)_{n<m}\right) \geq C_l. \tag{72}$$

Then for all $m_1 \in \mathbb{N}$, we have:

$$\mathbb{P}(\exists m \geq m_1,\ \mathrm{sqz}(\overline{\gamma}_m) \geq m\alpha) \geq C_\alpha. \tag{73}$$

Consider $m_0 \leq m_1$ two integers. With probability at least C_α, there is an integer $m_2 \geq m_1$ such that $\mathrm{sqz}(\gamma_{m_0} \cdots \gamma_{m_0+m_2-1}) \geq m_2\alpha$. We define m_2 to be the smallest such integer and $+\infty$ otherwise then m_2 is a stopping time for the cylinder filtration associated to the sequence $(\gamma_{m+m_0})_{m\in\mathbb{N}}$. In particular, for $g = (\gamma_{m_0} \cdots \gamma_{m_2-1})$, we have for all finite $m \geq m_1$:

$$\mathbb{P}\left(\forall m \leq m_0 - l, \forall m' \geq m_2 + l, (\gamma_m \cdots \gamma_{m_0-1})\mathbb{A}^\varepsilon g \mathbb{A}^\varepsilon (\gamma_{m_2} \cdots \gamma_{m'-1}) \,|\, m_2 = m\right) \geq C_l^2. \tag{74}$$

Now we write n_0, for the largest integer such that $\overline{w}_{n_0} \leq m_0 - l$ and n_1 for the smallest integer such that $m_2 + l \leq \overline{w}_{n_1}$. Then by Lemma 2.18 and (74), we have for all finite m:

$$\mathbb{P}\left(\mathrm{sqz}(\gamma^w_{n_0} \cdots \gamma^w_{n_1-1}) \geq m_2\alpha \,\middle|\, m_2 = m\right) \geq C_l^2 \tag{75}$$

Now take m_0 large enough so that:

$$\mathbb{P}(n_0 = 0) \leq \frac{C_\alpha C_l^2}{4} \tag{76}$$

and take m_1 large enough so that for $n(m)$ the smallest integer such that $\overline{w}_n \geq m + l$, we have:

$$\mathbb{P}\left(\exists m \geq m_1, n(m)(1-\delta)L(X,\tilde{\kappa}) > m\right) \leq \frac{C_\alpha C_l^2}{4}. \tag{77}$$

If we combine (75) with (77) and (76), we get that:

$$\mathbb{P}\left(\mathrm{sqz}(\gamma^w_1 \cdots \gamma^w_{n_1-1}) \geq n_1\alpha(1-\delta)L(X,\tilde{\kappa})\right) \geq \frac{C_\alpha C_l^2}{2} \tag{78}$$

Now this is true for all m_1 and we may assume that $n_1 \geq \frac{m_1}{2L(X,\tilde{\kappa})}$ with probability at least $1 - \frac{C_\alpha C_l^2}{4}$, so we get:

$$\mathbb{P}\left(\exists n \geq \frac{m_1}{2L(X,\tilde{\kappa})}, \frac{\mathrm{sqz}(\gamma^w_1 \cdots \gamma^w_n)}{n} \geq \alpha(1-\delta)L(X,\tilde{\kappa})\right) \geq \frac{C_\alpha C_l^2}{2}. \tag{79}$$

The above is true for all m_1, which means that the essential supremum of $\limsup \frac{\mathrm{sqz}(\gamma^w_1 \cdots \gamma^w_n)}{n}$ is at least $\alpha(1-\delta)L(X,\tilde{\kappa})$ so by Lemma 5.3, $\sigma(X,\kappa) \geq \alpha(1-\delta)L(X,\tilde{\kappa})$. Moreover this is true for all $\alpha < \sigma(\nu)$ and all $\delta > 0$ so $\sigma(X,\kappa) \geq L(X,\tilde{\kappa})\sigma(\nu)$. Now the convergence is a consequence of the large deviations inequalities from below by Remark B.28. ☐

Theorem 5.9 (Quantitative convergence to the boundary)**.** *Let ν be a strongly irreducible probability distribution over $\Gamma := \mathrm{End}(E)$. Consider a random sequence $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$. Let $u \in \mathbf{P}(E) \setminus \ker(\nu)$ be a generic line (in the sense that u is not included in $\ker(\nu)$). Then we have a random limit $u_\infty \in \mathbf{P}(E)$ such that for all $\alpha < \sigma(\nu)$, we have constants $C, \beta > 0$ such that for all $n \in \mathbb{N}$:*

$$\forall u \in \mathbf{P}(E),\ \mathbb{P}(\mathrm{d}(\overline{\gamma}_n u, u_\infty) \geq \exp(-\alpha n)) \leq C\exp(-\beta n), \tag{80}$$

$$\mathbb{P}\left(\exists u \in U^+(\overline{\gamma}_n),\ \mathrm{d}(u, u_\infty) \geq \exp(-\alpha n)\right) \leq C\exp(-\beta n). \tag{81}$$

Proof. Consider some $\rho < 1$, $\varepsilon > 0$ and $\lambda \geq 2|\log(\varepsilon)| + \log(4)$ such that there is $(X,\tilde{\kappa})$ a λ-squeezing, strongly $\mathbb{A}^\varepsilon$-aligned and ρ-ping-pong extraction of ν as in Theorem 4.9. Then take a random sequence $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$ and a random sequence (w_k) of integers such that $(\tilde{\gamma}^w_k)$ is a decorated Markov chain in $(X,\tilde{\kappa})$. By Lemma 2.18, for all integers $m \geq n$, we have $\overline{\gamma}^w_n \mathbb{A}^\varepsilon \left(\gamma^w_n \cdots \gamma^w_{m-1}\right)$ so by Lemma 2.13, we have:

$$\forall u \in U^+(\overline{\gamma}^w_n),\ , \forall u' \in U^+\left(\overline{\gamma}^w_m\right),\ \mathrm{d}(u,u') \leq \frac{2}{\varepsilon}\exp\left(-\mathrm{sqz}\left(\overline{\gamma}^w_n\right)\right). \tag{82}$$

Then (82) is true for all $m \in \mathbb{N}$ so we have:

$$\forall u \in U^+(\overline{\gamma}^w_n),\ \forall u' \in \bigcap_{k\geq 0} \overset{\mathbf{cl}}{\bigcup_{m\geq k}} U^+(\overline{\gamma}^w_m),\ \mathrm{d}(u,u') \leq \frac{2}{\varepsilon}\exp\left(-\mathrm{sqz}\left(\overline{\gamma}^w_n\right)\right). \tag{83}$$

Now (83) is true for all $n \in \mathbb{N}$ and $\exp(-\mathrm{sqz}(\overline{\gamma}_n)) \to 0$ so $\bigcap_{k \geq 0} \bigcup_{m \geq k}^{\mathbf{cl}} U^+(\overline{\gamma}_m^w)$ has only one random point that we call u_∞. For all $n \in \mathbb{N}$, write m_n for the largest integer such that $\overline{w}_{m_n} \leq n$ and r_n for the smallest integer such that $\gamma^w_{m_n - r_n - 1} \mathbb{A}^\varepsilon \left(\gamma^w_{m_n - r_n} \cdots \gamma_{n-1} h\right)$. By Lemma 4.19, The sequence (r_n) is uniformly exponentially integrable and by Lemma B.29, the sequence $(m_n - r_n)$ satisfies large deviations inequalities below the speed $\frac{1}{L(X, \tilde{\kappa})}$. Moreover, by Lemma 2.13, we have:

$$\forall u \in U^+(\overline{\gamma}_n h),\ \forall u' \in U^+(\overline{\gamma}^w_{m_n - r_n}),\ \mathrm{d}(u, u') \leq \frac{2}{\varepsilon} \exp\left(-\mathrm{sqz}\left(\overline{\gamma}^w_{m_n - r_n}\right)\right) \tag{84}$$

So by triangular inequality with (83) and (84), we have:

$$\forall u \in U^+(\overline{\gamma}_n h),\ \mathrm{d}(u, u_\infty) \leq \frac{2}{\varepsilon} \exp\left(-\mathrm{sqz}\left(\overline{\gamma}^w_{m_n - r_n}\right)\right) \tag{85}$$

Moreover, the random sequence $(\overline{w}_n)$ satisfies large deviations inequalities around the speed $L(X, \tilde{\kappa})$ by Lemma 5.1. By Lemma B.29 6 random sequence (m_n) satisfies large deviations inequalities around the speed $\frac{1}{L(X, \tilde{\kappa})}$ and by Lemma B.29 1 $(m_n - r_n)$ also does. Then by Lemma B.29 5 and Theorem 5.8, the random sequence $\left(\mathrm{sqz}\left(\overline{\gamma}^w_{m_n - r_n}\right)\right)_{n \in \mathbb{N}}$ satisfies large deviations inequalities below the speed $\sigma(\nu)$. If we take $h = \mathrm{Id}$ in (85) then the large deviations inequalities for $\left(\mathrm{sqz}\left(\overline{\gamma}^w_{m_n - r_n}\right)\right)_{n \in \mathbb{N}}$ give (81). Now, if we take h to be a rank one endomorphism of image $u \in \mathbf{P}(E)$, then $U^+(\overline{\gamma}_n h) = \overline{\gamma}_n u$ so (85) and the large deviations inequalities for $\left(\mathrm{sqz}\left(\overline{\gamma}^w_{m_n - r_n}\right)\right)_{n \in \mathbb{N}}$ give (80). □

Corollary 5.10. *Let ν be any strongly irreducible and proximal distribution of positive rank on* $\mathrm{End}(E)$*. There is a unique ν-invariant probability distribution ξ_ν on $\mathbf{P}(E)$. Moreover, There is a constant C, β such that for all distribution ξ on $\mathbf{P}(E) \setminus \ker(\nu)$ and for all Lipschitz function $f : \mathbf{P}(E) \to \mathbb{R}$ with Lipschitz constant $\lambda(f)$, we have:*

$$\forall n \in \mathbb{N},\ \left| \int_{\mathbf{P}(E)} f \mathrm{d}\xi_\nu - \int_{\mathbf{P}(E)} f \mathrm{d}\nu^n * \xi \right| \leq \lambda(f) C \exp(-\beta n). \tag{86}$$

Proof. Let $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$. Consider some line $u \in \mathbf{P}(E) \setminus \ker(\nu)$, with $\ker(\nu)$ as in Definition 3.30. Write $u_n := \overline{\gamma}_n u$ and $u'_n := \gamma_1 \cdots \gamma_n u$ for all $n \in \mathbb{N}$. Then by Theorem 5.9 and by Borel Cantelli's Lemma, (u'_n) and (u_n) both converge to a random limit that does not depend on u. We write $u_\infty := \lim u_n$ and $u'_\infty :=: lim u'_n$. Then u_∞ and u'_∞ only depend on $(\gamma_n)_{n \geq 0}$ and $(\gamma_{n+1})_{n \geq 0}$ so both have the same distribution that we write ξ_∞. Then we have $\gamma_0 u'_n = u_{n+1}$ for all $n \in \mathbb{N}$ so $\gamma_0 u'_\infty = u_\infty$ so ξ_∞ is ν-invariant. Now consider ξ a ν-invariant probability distribution. Draw $u_0, (\gamma_n) \sim \xi \otimes \nu^{\otimes \mathbb{N}}$. Assume that $\xi(\ker(\nu)) > 0$, then by definition of $\ker(\nu)$, we have $\mathbb{P}(\gamma_0 u = 0) > 0$ so $\xi\{\{0\}\} > 0$ which contradicts the fact that ξ is supported on $\mathbf{P}(E)$. As a consequence, we have $\overline{\gamma}_n u_0 \to u_\infty$ so in law, we have $\nu^{*n} * \xi \to \xi_\infty$ but $\xi = \nu^{*n} * \xi$ for all $n \in \mathbb{N}$ so $\xi = \xi_\infty$. Now to prove (86) draw $u_0, (\gamma_n) \sim \xi \otimes \nu^{\otimes \mathbb{N}}$ for ξ any probability distribution on $\mathbf{P}(E) \setminus \ker(\nu)$. Write $u_n := \overline{\gamma}_n u_0$ for all $n \in \mathbb{N}$ and $u_\infty := \lim u_n$. Then we have $\mathbb{E}(f(u_n)) = \int f(u) \mathrm{d}\nu^n * \xi(u)$ for all $n \in \mathbb{N}$ and $\mathbb{E}(f(u_\infty)) = \int f(u) \mathrm{d}\xi(u)$. Now consider some $0 < \alpha < \sigma(\nu)$, we have some constants $C, \beta > 0$ that satisfy (80) for all $n \in \mathbb{N}$. Then since f is λ-Lipschitz we have:

$$\begin{aligned} \mathbb{E}(f(u_n) - f(u_\infty)) &\leq \lambda(f) \mathbb{E}(\mathrm{d}(u_n, u_\infty)) \\ &\leq \lambda(f) \left(\mathbb{P}(\mathrm{d}(u_n, u_\infty) > \exp(-\alpha n)) + \exp(-\alpha n)\right) \\ &\leq \lambda(f) (C \exp(-\beta n) + \exp(-\alpha n)). \end{aligned}$$

□

We have shown that the left dominant space $U^+(\overline{\gamma}_n)$ converges exponentially fast to a limit l_∞ whose distribution is the unique ν-invariant distribution on $\mathbf{P}(E)$. In the meantime, the right dominant space $W^+(\overline{\gamma}_n)$ does not converge, in fact it has a mixing behaviour.

Corollary 5.11 (Mixing property for the right dominant space)**.** *Let ν be any strongly irreducible and proximal distribution of positive rank on* $\mathrm{End}(E)$*. Let ξ' be the only ν-invariant probability distribution on $\mathbf{P}(E^*)$ i.e, such that $\xi' * \nu = \xi'$. Let $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$ and $w \in \mathbf{P}(E^*) \setminus \ker(\nu^*)$. Then we have:*

$$\frac{1}{n} \sum_{k=0}^{n} \delta_{w \overline{\gamma}_n} \longrightarrow \xi'.$$

5.2 Asymptotic estimates for the spectral gap

Now that we understand the behaviour of the singular gap, we can study the spectral gap using the following trick.

Lemma 5.12 (Pivotal concatenation)**.** *Let Γ be a measurable semi-group endowed with a binary relation $\mathbb{A}$ and $\rho < 1$. Let n be an odd integer and let $(\gamma_0, \cdots, \gamma_{n-1})$ be a sequence of random variables such that for all odd $1 \leq k \leq n-2$, the distribution of γ_k knowing $(\gamma_j)_{j \neq k}$ is ρ-Schottky. Write $c \leq \frac{n-1}{2}$ for the smallest odd integer such that:*

$$(\gamma_{n-c} \cdots \gamma_{n-1} \overline{\gamma}_c) \mathbb{A} \gamma_c \tag{87}$$

$$\gamma_{n-c_n-1} \mathbb{A} (\gamma_{n-c} \cdots \gamma_{n-1} \overline{\gamma}_{c+1}) \tag{88}$$

and take $c := \frac{n+1}{2}$ if there is no such c. Then we have:

$$\forall k \in \mathbb{N},\ \mathbb{P}(c > 2k+1) \leq (2\rho - \rho^2)^k. \tag{89}$$

Proof. Consider some odd integer $n \in \mathbb{N}$ and a random sequence $(\gamma_k)_{0 \leq k < n}$ as in Lemma 5.12. For all $k \leq \frac{n-1}{4}$, write $\mathcal{C}_k$ for the σ-algebra generated by the families $(\gamma_{2j})_{0 \leq j \leq \frac{n-1}{2}}$ and $(\gamma_{2j-1})_{1 \leq j \leq k}$ and $(\gamma_{n-2j})_{1 \leq j \leq k}$. Then for all $0 \leq k \leq \frac{n-3}{4}$, the distribution of γ_{2k+1} knowing $\mathcal{C}_k$, is ρ-Schottky, and $(\gamma_{n-2k+1} \cdots \gamma_{n-1} \overline{\gamma}_{2k-1})$ is $\mathcal{C}_k$-measurable so we have:

$$\mathbb{P}\left((\gamma_{n-2k-1} \cdots \gamma_{n-1} \overline{\gamma}_{2k+1}) \mathbb{A} \gamma_{2k+1} \,\middle|\, \mathcal{C}_k\right) \geq (1-\rho). \tag{90}$$

Then the distribution of γ_{n-2k-2} knowing $\mathcal{C}_k$ and γ_{2k+1} is also ρ-Schottky, so we have:

$$\mathbb{P}\left(\gamma_{n-2k-2} \mathbb{A} (\gamma_{n-2k-1} \cdots \gamma_{n-1} \overline{\gamma}_{2k+2}) \,\middle|\, \mathcal{C}_k, \right) \geq (1-\rho). \tag{91}$$

Then note that for all $0 \leq k' < k$, the event $c = 2k'+1$ in $\mathcal{C}_k$-measurable so if we combine (90) and (91), we get that:

$$\mathbb{P}(c \leq 2k+1 \,|\, \mathcal{C}_k) \geq (1-\rho)^2$$

$$\mathbb{P}(c > 2k+1 \,|\, c > 2k-1) \leq 2\rho - \rho^2. \tag{92}$$

Moreover, for $k = \langle \frac{n+1}{4} \rangle$, we have $\mathbb{P}(c > 2k+1) = 0$. Then if we apply (92), by induction initializing with $\mathbb{P}(c \geq 1) = 1$, we have (89). ☐

Theorem 5.13 (large deviations inequalities for the spectral gap and convergence of the dominant eigenspace)**.** *Let ν be a strongly irreducible and proximal probability distribution over* $\mathrm{End}(E)$*. Let $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$ and $\sigma(\nu)$ be as in Theorem 5.8. We have $\sigma(\nu) = \liminf \frac{\mathrm{prox}(\overline{\gamma}_n)}{n}$ and:*

$$\forall \alpha < \sigma(\nu), \exists C, \beta > 0, \forall n \in \mathbb{N}, \mathbb{P}\left(\mathrm{prox}(\overline{\gamma}_n) \leq \alpha n\right) \leq C \exp(-\beta n). \tag{93}$$

Moreover, if we write $E^+(\overline{\gamma}_n)$ for the dominant eigenspace of $\overline{\gamma}_n$. Then for u_∞ as defined in Theorem 5.9, we have:

$$\forall \alpha < \sigma(\nu),\ \exists C, \beta > 0,\ \forall n \in \mathbb{N},\ \mathbb{P}(\mathrm{d}(E^+(\overline{\gamma}_n), u_\infty) \geq \exp(-\alpha n)) \leq C \exp(-\beta n) \tag{94}$$

Proof. Consider some $\rho < 1$, $\varepsilon > 0$ and $\lambda \geq 4|\log(\varepsilon)| + 4\log(2)$ such that there is $(X, \tilde{\kappa})$ a λ-squeezing, strongly $\mathbb{A}^\varepsilon$-aligned and ρ-ping-pong extraction of ν as in Theorem 4.9. Then take a random sequence $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$ and a random sequence (w_k) of integers such that $(\widetilde{\gamma}_k^w)$ is a decorated Markov chain in $(X, \tilde{\kappa})$. For all $n \in \mathbb{N}$, write m_n for the largest even integer such that $\overline{w}_{m_n} \leq n$. then write $c_n \leq \frac{m_n}{2}$ for the smallest odd integer such that:

$$\left(\gamma^w_{m_n+1-c_n} \cdots \gamma^w_{m_n-1} (\gamma_{\overline{w}_{m_n}} \cdots \gamma_{n-1}) \overline{\gamma}^w_{c_n}\right) \mathbb{A}^\varepsilon \gamma_{c_n} \tag{95}$$

and

$$\gamma^w_{m_n-c_n} \mathbb{A}^\varepsilon \left(\gamma^w_{m_n+1-c_n} \cdots \gamma^w_{m_n-1} (\gamma_{\overline{w}_{m_n}} \cdots \gamma_{n-1}) \overline{\gamma}^w_{c_n+1}\right) \tag{96}$$

and take $c_n = \frac{m_n}{2}$ otherwise. Then by lemma 5.12 applied to the sequence:

$$\left(\gamma_0^w, \dots, \gamma_{m_n-1}^w, (\gamma_{\overline{w}_{m_n}} \cdots \gamma_{n-1})\right),$$

we know that c_n has bounded exponential moment. So $(m_n - 2c_n)$ satisfies large deviations inequalities below the speed $\frac{1}{L(X,\tilde{\kappa})}$. Now assume that $m_n - 2c_n > 0$. Then by (95) and by Lemma 2.18, we have:

$$\operatorname{sqz}\left(\gamma_{m_n+1-c_n}^w \cdots \gamma_{m_n-1}^w (\gamma_{\overline{w}_{m_n}} \cdots \gamma_{n-1}) \overline{\gamma}_{c_n+1}^w\right) \geq \lambda - 2|\log(\varepsilon)| - 2\log(2), \tag{97}$$

$$\left(\gamma_{m_n+1-c_n}^w \cdots \gamma_{m_n-1}^w (\gamma_{\overline{w}_{m_n}} \cdots \gamma_{n-1}) \overline{\gamma}_{c_n+1}^w\right) \mathbb{A}^{\frac{\varepsilon}{2}} \gamma_{c_n+1}^w. \tag{98}$$

Define $k_n := \overline{w}_{m_n - c_n}$ for $h_n := \gamma_{k_n} \cdots \gamma_{n-1} \overline{\gamma}_{k_n - 1}$. Note that k_n satisfies large deviations inequalities around the speed 1 by lemma B.29. Moreover, by Lemma 2.18, we have $h_n \mathbb{A}^{\frac{\varepsilon}{2}} h_n$ and:

$$\operatorname{sqz}(h_n) \geq \operatorname{sqz}\left(\gamma_{c_n+1}^w \cdots \gamma_{k_n}\right). \tag{99}$$

Then by Lemma 5.3, the random sequence $(\operatorname{sqz}(h_n))_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed $\sigma(\nu)$. Moreover, by Corollary 2.27, we have:

$$\operatorname{prox}(h_n) \geq \operatorname{sqz}(h_n) - 2|\log(\varepsilon)| - 4\log(2) \tag{100}$$

$$\forall u \in U^+(h_n),\ \operatorname{d}\left(u, E^+(h_n)\right) \leq \exp\left(-\operatorname{sqz}(h_n)\right) \frac{4}{\varepsilon}. \tag{101}$$

Moreover, h_n is cyclically conjugated to $\overline{\gamma}_n$ for all n such that $m_n - 2c_n > 0$ so $\operatorname{prox}(h_n) = \operatorname{prox}(\gamma_n)$ and $E^+(\overline{\gamma}_n) = \overline{\gamma}_{k_n} E^+(h_n))$. Moreover, by Lemma 2.18, we also have $\overline{\gamma}_{k_n} \mathbb{A}^{\frac{\varepsilon}{2}} h_n$. Now consider n large enough so that $\operatorname{sqz}(h_n) \geq 2|\log(\varepsilon)| - 4\log(2)$ and by Lemma 2.13 and (101), we have:

$$\forall u \in U^+\left(\overline{\gamma}_{k_n}\right),\ \operatorname{d}\left(u, E^+\left(\overline{\gamma}_n\right)\right) \leq \exp\left(-\operatorname{sqz}\left(\overline{\gamma}_{k_n}\right)\right) \frac{4}{\varepsilon} \tag{102}$$

Then by Theorem 5.9, we have a sequence $U^+\left(\overline{\gamma}_{k_n}\right)$ is in a ball of radius at most $\exp(-\lambda_n)$ around a random limit $u_\infty \in \mathbf{P}(E)$, for a random sequence (λ_n) that satisfies large deviations inequalities below the speed $\sigma(\nu)$. Then by lemma 5.3 $(\operatorname{sqz}(k_n))$ also satisfies large deviations inequalities below the speed $\sigma(\nu)$. Then by (102), by triangular inequality and by lemma B.29, the random sequence $(-\log(\operatorname{d}(E^+(\overline{\gamma}_n), u_\infty))$ satisfies large deviations inequalities below the speed $\sigma(\nu)$. □

Proof of Theorems 1.2 and 1.3. Consider ν a strongly irreducible probability distribution on $\Gamma := \operatorname{End}(E)$. Let $(\gamma_n)_{n\in\mathbb{N}} \sim \nu^{\otimes\mathbb{N}}$. If ν has rank 0, then by lemma 3.29, we have two constants $C, \beta > 0$ such that $\mathbb{P}(\overline{\gamma}_n \neq 0) \leq C\exp(-\beta n)$. By convention, we have $\operatorname{sqz}(0) = +\infty$ so we have Theorem 1.2 for $\sigma(\nu) = +\infty$ and Theorem 1.3 doesn't apply. If ν is not proximal then we write $\sigma(\nu) := 0$ and then by Lemma 3.38, the sequences $\left(\operatorname{sqz}\left(\overline{\gamma}_n\right)\right)_{n\in\mathbb{N}}$ is bounded and $\operatorname{prox}\left(\overline{\gamma}_n\right) = 0$ for all $n \in \mathbb{N}$ so (3) and (4) are trivial. When ν is proximal, (3) is a reformulation of Theorem 5.8 and (4) is a reformulation of Theorem 5.13. To show Theorem 1.3, we simply need to show that the limit $u_\infty \in \mathbf{P}(E)$ defined in Theorem 5.9 can be expressed as the image of the sequence $(\gamma_n)_{n\in\mathbb{N}}$ by a measurable, shift-invariant map $\Gamma^{\mathbb{N}} \to \mathbf{P}(E)$. We simply define $U_\infty((g_n)_{n\in\mathbb{N}})$ to be the limit point of $U^+(\overline{g}_n)$ when for all sequences $x_n \in U^+(\overline{g}_n)$, the sequence $(\mathbb{K}x_n)$ converges to the same limit and when the same is true for all the shifted sequences $(g_{n+m})_{n\in\mathbb{N}}$ for $m \in \mathbb{N}$. □

5.3 Limit flag for absolutely strongly irreducible distributions

Now we can give the following corollary which is a reformulation of the former results made to look like the stable and unstable spaces decomposition in Oseledets' Theorem with large deviations from below.

Definition 5.14. *Let ν be a distribution on* $\operatorname{End}(E)$ *that is absolutely strongly irreducible in the sense of Definition 3.34. For all $1 \leq k < \dim(E)$, we write $\sigma_k(\nu) := \sigma\left(\bigwedge^k \nu\right)$ as in Theorem 5.8. We write:*

$$\Theta(\nu) := \{1 \leq k < \dim(E) \,|\, \sigma_k(\nu) \neq 0\} \tag{103}$$

We call Θ the set of spectral gaps of ν.

Theorem 5.15 (Convergence of the Cartan projection with large deviations)**.** *Let E be a standard vector spaces and ν be an absolutely strongly irreducible probability distribution on* $\mathrm{End}(E)$ *of rank at least* $\dim(E)-1$*. Let $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$. For $\widetilde{\sigma}(\nu) := (\sigma_j(\nu))_{1 \le j \le d-1} \in [0,+\infty]^{\dim(E)-1}$, we have almost surely $(\frac{\mathrm{sqz}_j(\overline{\gamma}_n)}{n})_j \to \tilde{\sigma}(\nu)$. Moreover, for all $\widetilde{\alpha} < \widetilde{\sigma}(\nu)$ i.e, $\alpha_k < \sigma_k(nu)$ for all k, we have constants $C, \beta > 0$ such that:*

$$\forall n \in \mathbb{N}, \forall 1 \le j < d,\ \mathbb{P}\left(\mathrm{sqz}_j(\overline{\gamma}_n) \le \alpha_j n\right) \le C\exp(-\beta n). \tag{104}$$

$$\forall n \in \mathbb{N},\ \mathbb{P}\left(\widetilde{\mathrm{prox}}(\overline{\gamma}_n) \le \alpha n\right) \le C\exp(-\beta n). \tag{105}$$

If we moreover assume that $N(\nu)$ has finite first moment, then we have almost surely $\frac{\mathrm{prox}_j(\overline{\gamma}_n)}{n}_j \to \sigma_j(\nu)$ for all j.

Proof. Line (104) is a reformulation of Theorem 5.8 for $\bigwedge^j(\nu)$. Line 105 is a reformulation of Theorem 5.13. Now assume that $N(\nu)$ has finite first moment. It means that $\mathbb{E}\frac{N(\overline{\gamma}_n)}{n}$ is bounded. Moreover, we have $N = \sum \mathrm{sqz}_j$ so $\frac{N(\overline{\gamma}_n)}{n} \to L := \sum \sigma_j(\nu) < +\infty$ so all the $\sigma_j(\nu)$ are finite. Then note that $\sum \mathrm{prox}_j \le N$ so $\limsup \sum \frac{\mathrm{prox}_j(\overline{\gamma}_n)}{n} \le L$. Moreover, by (105), we have $\liminf \frac{\mathrm{prox}_j(\overline{\gamma}_n)}{n} = \sigma_j(\nu)$ for all j. Now take $k \in \{1, \dots, \dim(E)-1\}$. We have:

$$\begin{aligned}
\mathrm{prox}_k(\overline{\gamma}_n) &\le N(\overline{\gamma_n}) - \sum_{j \ne k} \mathrm{prox}_j(\overline{\gamma}_n) \\
\limsup \frac{\mathrm{prox}_k(\overline{\gamma}_n)}{n} &\le \limsup \frac{N(\overline{\gamma}_n)}{n} - \sum \liminf \frac{\mathrm{prox}_j(\overline{\gamma}_n)}{n} \\
&\le L - \sum_{j \ne k} \sigma_j(\nu) \\
&\le \sigma_k(\nu).
\end{aligned}$$

It implies that $\frac{\mathrm{prox}_k(\overline{\gamma}_n)}{n} \to \sigma_k(\nu)$ for all j. □

Theorem 5.16 (Convergence to the limit flag)**.** *Let E be a standard vector space and $\Theta \subset \{1, \dots, \dim(E)-1\}$. There is a partially defined measurable map:*

$$F^\infty : \quad \mathrm{End}(E)^{\mathbb{N}} \quad \longrightarrow \quad \mathrm{Fl}_\Theta(E) \sqcup \{\text{undefined}\}$$

Such that for all absolutely strongly irreducible probability distribution ν on $\mathrm{End}(E)$ *such that $\Theta(\nu) \supset \Theta$, the map F is $\nu^{\otimes\mathbb{N}}$ almost surely defined. Moreover F^∞ is shift-equivariant, in the sense that for all sequence $(\gamma_n)_{n\in\mathbb{N}}$ such that $F^\infty\left((\gamma_n)_{n\in\mathbb{N}}\right)$ is defined, we have:*

$$F^\infty\left((\gamma_n)_{n\in\mathbb{N}}\right) = \gamma_0 F^\infty\left((\gamma_{n+1})_{n\in\mathbb{N}}\right) \tag{106}$$

Moreover, for an absolutely strongly irreducible probability distribution ν and for a generic flag $F \in \mathrm{Fl}_{\Theta(\nu)}$, for all $k \in \Theta(\nu)$, the random sequence:

$$-\log\left(\mathrm{d}\left(F_k^\infty((\gamma_n)_{n\in\mathbb{N}}), \overline{\gamma}_n F_k\right)\right)$$

satisfies large deviations inequalities below the speed $\sigma_k(\nu)$.

Proof. Let $(\gamma_n) \in \mathrm{End}(E)^{\mathbb{N}}$ be any sequence and $k < \dim(E)$. For all $g \in \mathrm{End}(E)$, we write:

$$U_k^+(g) := \left\{V \in \mathrm{Gr}_k(E) \,\middle|\, \bigwedge^k V \subset U^+\left(\bigwedge^k g\right)\right\}.$$

If $U_k^+(\overline{\gamma}_n)$ converges to a point and $\mathrm{sqz}_k(\overline{\gamma}_n) \to \infty$, we write $F_k^\infty((\gamma_n)_{n\in\mathbb{N}})$ for the value of this point, otherwise, $F_k^\infty((\gamma_n)_{n\in\mathbb{N}})$ is undefined. Then for all $\Theta \subset \{1, \dots, \dim(E)-1\}$, if there is an integer $k \in \Theta$ and $m \in \mathbb{N}$ such that $F_k^\infty((\gamma_{n+m})_{n\in\mathbb{N}})$ is undefined or $F_k^\infty((\gamma_{n+m})_{n\in\mathbb{N}}) \cap \ker(\overline{\gamma}_m) \ne \{0\}$, then $F^\infty((\gamma_n)_{n\in\mathbb{N}})$ is undefined, otherwise, we write $F^\infty((\gamma_n)_{n\in\mathbb{N}}) = \left(F_k^\infty((\gamma_n)_{n\in\mathbb{N}})\right)_{k\in\Theta}$. Now assume that $F^\infty((\gamma_n)_{n\in\mathbb{N}})$ is defined, then $F^\infty((\gamma_{n+1})_{n\in\mathbb{N}})$ also is. Moreover $\ker(\gamma_0)$ is closed so after a time n_0 and for some $\varepsilon > 0$, for all $n \geq n_0$ and all $x \in \bigcup U_k^+(\gamma_1 \cdots \gamma_{n-1})$, we have $\|\gamma_0(x)\| \geq \varepsilon \|\gamma_0\| \|x\|$ so for all unitary $v \in V^+\left(\bigwedge^k \overline{\gamma}_n\right)$, we have $\|\gamma_1 \cdots \gamma_{n-1} v\| \geq \varepsilon \mu_k(\gamma_1 \cdots \gamma_{n-1})$ so by Lemma 2.13, we have:

$$\mathrm{d}\left(\gamma_1 \cdots \gamma_{n-1} v, U_k^+(\gamma_1 \cdots \gamma_{n-1})\right) \leq \exp(-\mathrm{sqz}(\gamma_1 \cdots \gamma_{n-1}))/\varepsilon.$$

Then for all $\varepsilon' > 0$ there is an integer $n_1 \geq n_0$ such that for all $n \geq n_1$, we have:

$$\mathrm{d}\left(\gamma_1 \cdots \gamma_{n-1} v, U_k^+(\gamma_1 \cdots \gamma_{n-1})\right) \leq \varepsilon \varepsilon'.$$

Then by triangular inequality, we have

$$\mathrm{d}\left(\gamma_0 \cdots \gamma_{n-1} v, \gamma_0 U_k^+(\gamma_1 \cdots \gamma_{n-1})\right) \leq 2\varepsilon'/\varepsilon.$$

If we go to the limit $\varepsilon' \to 0$, this proves that F_k^∞ is shift-equivariant.

Now assume that $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$ for some absolutely strongly irreducible ν. then for all k such that $\sigma_k(\nu) > 0$, the space $F_k^\infty((\gamma_n)_{n\in\mathbb{N}})$ is almost surely well defined and we have the large deviations inequalities by Theorem 5.9 applied to $\bigwedge^k \nu$. Moreover, the distribution of $F_k^\infty((\gamma_{n+m})_{n\in\mathbb{N}})$ gives measure 0 to $\ker\left(\bigwedge^k \nu\right)$ by Corollary 5.10. It means that $\mathbb{P}\left(F_k^\infty((\gamma_{n+m})_{n\in\mathbb{N}}) \cap \ker(\overline{\gamma_m}) \neq 0\right) = 0$ so $F^\infty((\gamma_n)_{n\in\mathbb{N}})$ is almost surely defined. ☐

Note that in the case of ν an absolutely strongly irreducible probability distribution on $\mathrm{SL}(E)$, one can actually take the pivotal extraction to be aligned in all Cartan projections. With a correct adaptation of the works of [CFFT22], one should be able to prove the following.

Conjecture 5.17 (Poisson boundary)**.** *Let ν be an absolutely strongly irreducible probability distribution on $\mathrm{SL}(E)$. Assume that ν has finite entropy, then the Poisson boundary of ν is isomorphic to $\mathrm{Fl}_{\Theta(\nu)}(E)$ endowed with the ν-invariant probability distribution.*

5.4 Law of large numbers for the coefficients

In this section we give a proof of Theorem 5.18. We consider the map $N(\gamma) := \log\|\gamma\| + \log\|\gamma^{-1}\|$ defined on the group $\Gamma = \mathrm{GL}(E)$.

Theorem 5.18 (Strong law of large numbers for the coefficients)**.** *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible and proximal. There are constants C, β such that for all unitary $w \in E^*$ and $u \in E$, for all n and for $\overline{\gamma}_n$ a random matrix of distribution ν^{*n}, we have for all $t \geq 0$:*

$$\mathbb{P}\left(\log\frac{\|\overline{\gamma}_n\|}{|w\overline{\gamma}_n u|} > t\right) \leq C\exp(-\beta n) + \sum_{k=1}^{\infty} C\exp(-\beta k) N(\nu)\left(\frac{t}{k}, +\infty\right). \tag{107}$$

$$\mathbb{P}\left(\log\frac{\|\overline{\gamma}_n\|\|u\|}{\|\overline{\gamma}_n u\|} > t\right) \leq \sum_{k=1}^{\infty} C\exp(-\beta k) N(\nu)\left(\frac{t}{k}, +\infty\right) \tag{108}$$

Proof. Let $0 < \rho < \frac{1}{3}$, consider $0 < \varepsilon \leq 1$ and $(X, \tilde{\kappa})$ the ping-pong extraction as constructed in Theorem 4.9. Write $\mathbb{A} := \mathbb{A}^\varepsilon$. Consider $(\gamma_n) \sim \nu^{\otimes\mathbb{N}}$ a random sequence and (w_n) a random sequence of positive integers such that (γ_n^w) is an ornamented Markov chain in $(X, \tilde{\kappa})$ and (τ_n) an independent sequence taken uniformly in $[0,1]^{\mathbb{N}}$. Consider $C \in \mathbb{R}_{\geq 0}$ large enough so that :

$$\forall t \geq C,\ \forall k \in \mathbb{N},\ \mathbb{P}\left(N(\gamma_k) > t \,|\, (w_n)\right) \leq C N(\nu)(t, +\infty). \tag{109}$$

Such a C exists by Theorem 4.9, C exists (take C the maximum of C and B as defined in the Theorem). Now consider $\gamma' \sim \Pi\tilde{\kappa}_a$, we define the function:

$$F(f, g, h) := \frac{\mathbb{1}(f\mathbb{A}g\mathbb{A}h)}{\mathbb{P}(f\mathbb{A}\gamma'\mathbb{A}h))}(1 - 2\rho). \tag{110}$$

Now consider $g, h \in \mathrm{End}(E) \setminus \{0\}$ and n any integer and write m_n for the only integer such that $\overline{w}_{2m_n} \leq n < \overline{w}_{2m_n+2}$. Then by Lemma B.23, the distribution of the random integer $n - \overline{w}_{m_n}$ is bounded in law by an exponentially integrable distribution. By Lemma 5.1, the random sequence m_n satisfies large deviations inequalities below the speed $L(X,\tilde\kappa)^{-1}/2$. Write $\alpha := \frac{1}{6}L(X,\tilde\kappa)^{-1}$. Then for C large enough and for $\beta > 0$ small enough, we have:

$$\mathbb{P}(m_n < 2\lfloor\alpha n\rfloor + 2) \leq C\exp(-\beta n) \tag{111}$$

When $m_n > 2\lfloor\alpha n\rfloor + 2$ write l_n for the smallest integer such that $l_n = \lfloor\alpha n\rfloor + 1$ or:

$$\tau_{l_n} < F\left(g\overline{\gamma}^w_{2l_n}, \gamma^w_{2l_n+1}, \gamma^w_{2l_n+2}\right).$$

Write r_n for the largest integer such that either $m_n - r_n = \lfloor\alpha n\rfloor + 1$ or :

$$\tau_{r_n} < F\left(\overline{\gamma}^w_{2r_n-2}, \gamma^w_{2r_n-1}, \gamma_{\overline{w}2r_n}\cdots\gamma_{n-1}h\right).$$

Note that γ^w_{2k+1} has the same distribution as γ' for all k so knowing $(\widetilde{\gamma}^w_{2k})_{k\in\mathbb{N}}$, so the integers l_n and $m_n - r_n$ follow geometric distributions of scale factor 2ρ stopped at $\lfloor\alpha n\rfloor + 1$. Moreover, l_n and $m_n - r_n$ are independent because the intervals $\{2k+1 \,|\, k \leq \lfloor\alpha n\rfloor\}$ and $\{2k-1 \,|\, m_n - k \leq \lfloor\alpha n\rfloor\}$ are disjoint when $m_n > 2\lfloor\alpha n\rfloor + 2$. So we have:

$$\mathbb{P}\left(\text{either}\begin{cases} l_n = \lfloor\alpha n\rfloor + 1 & \text{or} \\ m_n - r_n = \lfloor\alpha n\rfloor + 1 & \text{or} \\ m_n < 2\lfloor\alpha n\rfloor + 2 \end{cases}\right) \leq 2(2\rho)^{\lfloor\alpha n\rfloor+1} + C\exp(-\beta n)$$

$$\mathbb{P}((g\overline{\gamma}^w_{2l_n+1})\mathbb{A}\gamma^w_{2l_n+1}\mathbb{A}\dots\mathbb{A}\gamma^w_{2r_n-1}\mathbb{A}(\gamma_{\overline{w}_{2r_n}}\cdots\gamma_{n-1}h)) \geq 1 - 2(2\rho)^{\alpha n} - C\exp(-\beta n)$$

Up to replacing β by $\min\{\beta, \alpha\log(2\rho)\}$ and replacing C by $C+2$, we have for $\beta > 0$ small enough and for C large enough:

$$\mathbb{P}((g\overline{\gamma}^w_{2l_n})\mathbb{A}\gamma^w_{2l_n+1}\mathbb{A}\dots\mathbb{A}\gamma^w_{2r_n-1}\mathbb{A}(\gamma_{\overline{w}2r_n}\cdots\gamma_{n-1}h)) \geq 1 - C\exp(-\beta n) \tag{112}$$

Then by Lemma B.24, and because $(X,\tilde\kappa)$ is exponentially integrable, for C large enough and $\beta > 0$ small enough, we have:

$$\forall k\in\mathbb{N},\ \mathbb{P}\left(\overline{w}_{2l_n+1} + n - \overline{w}_{2r_n} = k \,|\, (\widetilde{\gamma}^w_{2k})_{k\in\mathbb{N}}\right) \leq C\exp(-\beta k). \tag{113}$$

Moreover, by independence, by (109) and because all the letters of $\widetilde{\gamma}^w_{2k+1}$ have norm at most C, we have:

$$\forall t \geq C,\ \forall k\in\mathbb{N},\ \mathbb{P}\left(N(\gamma_k) > t \,|\, (w_n), l_n, r_n\right) \leq CN(\nu)(t,+\infty). \tag{114}$$

Then by (113), (112) and Lemma B.40 we have for C large enough, for all $t \geq 0$ and for all $n\in\mathbb{N}$:

$$\mathbb{P}\left(\sum_{k=0}^{\overline{w}_{2l_n}-1} N(\gamma_k) + \sum_{k=\overline{w}_{2r_n}}^{n-1} N(\gamma_k) > t\right) \leq C\exp(-\beta n) + \sum_{k=1}^{\infty} C\exp(-\beta k)N(\nu)\left(\frac{t}{k} - C, +\infty\right). \tag{115}$$

Now write $D := 2|\log(\varepsilon)| + 2\log(2)$ and:

$$\Delta_n := \sum_{k=0}^{\overline{w}_{2l_n}-1} N(\gamma_k) + \sum_{k=\overline{w}_{2r_n}}^{n-1} N(\gamma_k) + D. \tag{116}$$

Then write $C := C'C''$. By Lemma B.40, we have for all $t \geq 0$:

$$\begin{aligned}\mathbb{P}\left(\Delta_n > t\right) &= \sum_{k=0}^{\infty}\mathbb{P}(\overline{w}_{2l_n} + n - \overline{w}_{2r_n} = k)\mathbb{P}\left(\Delta_n > t \,|\, \overline{w}_{2l_n} + n - \overline{w}_{2r_n} = k\right) \\ &\leq \sum_{k=1}^{+\infty} C\exp(-\beta k)(B\wedge N(\nu))^{\star k}(t - 2|\log(\varepsilon)| - 2\log(2), +\infty).\end{aligned}$$

Moreover, we have:

$$g\overline{\gamma}^w_{l_n}\mathbb{A}\gamma^w_{2l_n+1}\mathbb{A}\cdots\mathbb{A}\gamma^w_{2r_n-1}\mathbb{A}\gamma_{\overline{w}2r_n}\cdots\gamma_{n-1}h, \tag{117}$$

So by Lemma 2.18, and (13) in Lemma 2.13 we have:

$$\|g\overline{\gamma}_n h\| \geq \|g\overline{\gamma}^w_{l_n}\|\frac{\varepsilon}{2}\|\gamma^w_{2l_n+1}\cdots\gamma^w_{2r_n-1}\|\frac{\varepsilon}{2}\|\gamma_{\overline{w}2r_n}\cdots\gamma_{n-1}h\|$$

Moreover, by sub-multiplicativity, we have:

$$\|\overline{\gamma}_n\| \leq \prod_{k=0}^{\overline{w}_{2l_n}-1}\|\gamma_k\|\cdot\|\gamma^w_{2l_n+1}\cdots\gamma^w_{2r_n-1}\|\prod_{k=\overline{w}_{2r_n}}^{n-1}\|\gamma_k\| \tag{118}$$

and

$$\|g\| = \|g\overline{\gamma}^w_{l_n}(\overline{\gamma}^w_{l_n})^{-1}\|$$

$$\|g\| \leq \|g\overline{\gamma}^w_{l_n}\|\prod_{k=0}^{\overline{w}_{2l_n}-1}\|\gamma_k^{-1}\| \tag{119}$$

$$\|h\| \leq \|\gamma_{\overline{w}_{2r_n}}\cdots\gamma_{n-1}h\|\prod_{k=\overline{w}_{2r_n}}^{n-1}\|\gamma_k^{-1}\| \tag{120}$$

Now write for $g = ew$ and $h := ue'$ with $e \in E$ and $e' \in E^*$ unitary. Note that $\|g\| = \|w\|$ and $\|h\| = \|u\|$ and $\|g\overline{\gamma}_n h\| = |w\overline{\gamma}_n|$. We combine $\log(118) + \log(119) + \log(120) - \log(117)$, for $g = ew$ and $h := ue'$ with $e \in E$ and $e' \in E^*$ unitary, we get:

$$\log\|\overline{\gamma}_n\| + \log\|u\| + \log\|w\| - \log|w\overline{\gamma}_n u| \leq \Delta_n. \tag{121}$$

This is true as long as l_n and r_n are well defined, then we apply Remark B.41 and Lemma B.40 to Δ_n to get (107). To get (108), we remove the condition $m_n - r_n \leq \lfloor \alpha n\rfloor + 1$ and simply impose that $r_n \geq 0$. That way, we have:

$$\overline{\gamma}^w_{2r_n}\mathbb{A}^{\frac{\varepsilon}{2}}\gamma_{\overline{w}2r_n}\cdots\gamma_{n-1}h.$$

Because the identity is aligned with everyone in the case $r_n = 0$ and because of Lemma 2.18 in the case $r_n \geq 0$, then we can apply the above reasoning without the error term $C\exp(-\beta n)$ from (112) in (115). ☐

Definition 5.19. *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible and proximal. Let C, β be as in Theorem 5.18. We define the distribution ζ_ν as:*

$$\forall t \geq 0, \zeta_\nu(t, +\infty) = \min\left\{1, \sum_{k=1}^{\infty} C\exp(-\beta k)N(\nu)\left(\frac{t}{k} - C, +\infty\right)\right\}. \tag{122}$$

Remark 5.20 (What about exponential moment)**.** *Note that when $N(\nu)$ has a finite exponential moment, one can not conclude that ζ_ν also has using* (107) *alone. It is however a known result, written in [BQ16a, p 231]. It is also possible to prove it using the pivotal method (ant the proof in actually less technical than the one of Theorem 5.18). For that, note that* (121) *still holds for Δ_n as defined in* (116)*. Moreover, by Lemma B.24, if we assume that the $N(\gamma_k)$ are i.i.d and have finite exponential moment then Δ_n also has because the size of the sum has finite exponential moment.*

Corollary 5.21 (Qualitative convergence)**.** *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible, proximal. Assume moreover that for $\gamma \sim \nu$, the random variables $\log\|\gamma\|$ and $\log\|\gamma^{-1}\|$ both have finite expectation. Then for every $w \in E^*\setminus\{0\}$, and every $u \in E\setminus\{0\}$ we have almost surely $\frac{\log|w\overline{\gamma}_n u|}{n} \to \rho(\nu)$ for $\rho(\nu) := \lim\frac{\log\|\overline{\gamma}_n\|}{n}$.*

Proof. By Theorem 2 in [Fur63], the quantity $\rho(\nu) := \lim \frac{\log \|\overline{\gamma}_n\|}{n}$ is well defined and it is a finite constant. Moreover, we have $\rho(\nu) \geq \mathbb{E}(\log|\det(\nu)|)$. Moreover, we have $\limsup \frac{\log|w\overline{\gamma}_n u|}{n} \leq \rho(\nu)$ by definition of the operator norm. Then note that we have:

$$\left(\liminf \frac{\log|w\overline{\gamma}_n u|}{n} < \overline{\sigma}(\nu)\right) = \bigcup_{t>0} \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} \left(\log\left(\frac{|u\overline{\gamma}_n v|}{\|\overline{\gamma}_n\|}\right) \leq -t\right).$$

Then by homogeneity, one may assume that x and y are unitary. Indeed, the term $\frac{\log\|x\|+\log\|y\|}{n}$ goes to 0. Then, we have:

$$\mathbb{P}\left(\liminf \frac{\log|u\overline{\gamma}_n v|}{n} < \overline{\sigma}(\nu)\right) \leq \sup_{t>0} \lim_{m\to\infty} \sum_{n\geq m} \zeta_\nu\left(tn, +\infty\right) + C\exp(-\beta n). \tag{123}$$

Then by B.39, the distribution ζ_ν has finite expectation. So we have:

$$\sum_{n\geq m} \zeta_\nu\left(tn, +\infty\right) \leq \frac{\mathbb{E}(\zeta_\nu \mathbb{1}[tm, +\infty])}{t} \underset{m\to\infty}{\longrightarrow} 0.$$ ☐

Corollary 5.22 (Central limit theorem for the coefficients)**.** *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible, proximal. Assume moreover that for $\gamma \sim \nu$, the random variables $\log\|\gamma\|$ and $\log\|\gamma^{-1}\|$ both have finite L^2 moment. Then for every $w \in E^* \setminus \{0\}$, and every $u \in E \setminus \{0\}$, we have:*

$$\frac{\log|w\overline{\gamma}_n u| - n\rho(\nu)}{\sqrt{n}} \longrightarrow \mathcal{N}(0, V). \tag{124}$$

Proof. We know from [BQ16b] that $\frac{\log\|\overline{\gamma}_n\| - n\rho(\nu)}{\sqrt{n}}$ converges in law to a centred Gaussian distribution of variance V. Then by Theorem 1.5, ζ_ν has finite L^2 moment so $\frac{\log\|\overline{\gamma}_n\| - \log|w\overline{\gamma}_n u|}{\sqrt{n}}$ converges to 0 almost surely and as a consequence it converges in law to the Dirac mass at 0. ☐

Corollary 5.23 (Regularity of the measure)**.** *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible and proximal. Let p be such that $N(\nu)$ is weakly L^p. Then the ν-invariant distribution ξ_∞ on $\mathbf{P}(E)$ satisfies for some constant $C \geq 0$:*

$$\forall x \in \mathbf{P}(E), \forall r > 0, \xi_\infty(\mathcal{B}(x, r)) \leq \frac{C}{|\log(r)|^p}. \tag{125}$$

Proof. Write (γ_n) a random sequence of distribution $\nu^{\otimes\mathbb{N}}$ and $y \in \mathbf{E}$. By (80), in Theorem 5.9, there is a random line $u_\infty \in \mathbf{P}(E)$ such that almost surely and for all y, we have $u_\infty = \lim \mathbb{K}\overline{\gamma}_n y$ and some constants $\alpha > 0, \beta > 0, C'$, such that $\mathbb{P}(\mathrm{d}(\overline{\gamma}_n y, l_\infty) \geq \exp(-\alpha n)) \leq C'\exp(-\beta n)$. As a consequence, we have for all $r > 0$, $\mathrm{d}(\overline{\gamma}_n y, l_\infty) \leq r$ after a random time $n_0(r)$ who is bounded in law by a geometric law whose expectation is proportional to $|\log(r)|$. Then take some unitary vector $v \in y$ and a non-trivial linear form w such $w(x) = 0$ and a number $r > 0$, we have for all $n \geq n_0(r)$:

$$(\mathrm{d}(l_\infty, x) \leq r) \subset (\mathrm{d}(\overline{\gamma}_n y, x) \leq 2r) \subset \left(\frac{|w\overline{\gamma}_n v|}{\|\overline{\gamma}_n\|} \leq 2r\right). \tag{126}$$

Then by Theorem 5.18, we have $\mathbb{P}(\mathrm{d}(l_\infty, x) \leq r) \leq \zeta_\nu(|\log(2r)|, +\infty)$, and by Lemma B.39, ζ_ν is weakly L^p so there is a constant $C = 2W_p(\zeta_\nu) + \log(2)$ such that $\zeta_\nu(t - \log(2), +\infty) \leq Ct^{-p}$ for all $t \geq 0$. ☐

We said in the introduction that 5.23 is an amelioration of a result by Benoist and Quint in [BQ16b]. We show that the inequality in Corollary 5.23 is actually optimal by considering $\nu := \nu_A * \nu_K$ where ν_K is the Haar measure on the (compact) group of isometries $\mathrm{O}(E)$ and A is the distribution of the matrix

$$M := \begin{pmatrix} \exp(T) & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

where T is a random variable that is not weakly L^p. Then ν is strongly irreducible and proximal and it actually has full support in $\mathrm{PGL}(E)$. Then write ξ_∞ for the invariant distribution on $\mathbf{P}(E)$ and for x the first base vector. Then we have $\xi_\infty := \nu_A * \xi_K$ for ξ_K the Lebesgue measure on $\mathbf{P}(E)$. Indeed $\nu_K * \xi = \xi_K$ for all ξ, by property of the Haar measure. As a consequence, we have for all $r \geq 0$:

$$\xi_\infty(\mathcal{B}(x,r)) \geq \frac{1}{d}\mathbb{P}(T \geq \log(d) - \log(r)).$$

Where $\frac{1}{d}$ is the probability that a random variable of distribution law ξ_K has dominant first coordinate.

Another interesting question that is asked in [BQ16a, p 231] is to ask whether Theorem 5.18 still works if we drop the proximality assumption and replace it by an absolute string irreducibility assumption. Indeed theorem 5.16 tells us that if we take a distribution ν and write $p(\nu)$ its proximality rank *i.e,* $p(\nu) := \min \Theta(\nu)$, then we have a limit space of dimension $p(\nu)$. Then with the same trick as in the proof of 5.18, we show that the coefficient $w\overline{\gamma}_n u$ is up to an exponentially small error the product of a linear form w' and a vector u' whose norms are controlled in law by the same ζ_ν. However, the fact that the kernel of w' is orthogonal to a $p(\nu)$-dimensional space that contains u' does not give a lower bound on the product $|w'(u')|$. For example in dimension 2, we can take ν to be the law of a random rotation of angle $2^{-n}\pi$ with probability $\exp(-\exp(\exp(n)))$. Then the random walk $(\overline{\gamma}_n)$ is recurrent so if we take w, u such that $w(u) = 0$, then we almost surely have $|w\overline{\gamma}_n u| = 0$ for infinitely many times $n \in \mathbb{N}$.

Remark 5.24. *Using the same trick as in Theorem 5.18 to create the cyclically aligned decomposition in 5.13 and with the setting of Theorem 5.18, we can show that:*

$$\mathbb{P}\left(\log\left(\frac{\mu(\overline{\gamma}_n)}{|\lambda_1(\overline{\gamma}_n)|}\right) > t\right) \leq \sum_{k=1}^{\infty} C\exp(-\beta k)N(\nu)\left(\frac{t}{k} - C, +\infty\right).$$

We only need to introduce an auxiliary function so that the distribution of c_n as in Lemma 5.12 is explicit conditionally to the even $\widetilde{\gamma}_n^w$ and scroll down the proof of Theorem 5.18 using Corollary 2.27 to get a lower bound on $|\lambda_1(\overline{\gamma}_n)|$. This is stronger than 5.15 because we do not need to assume absolute strong irreducibility. This result can also be proven using Theorem 1.5 and Theorem 5.8 as a black box to conclude with corollary 2.27.

Theorem 5.25. *Let ν be a probability measure on $\mathrm{GL}(E)$ that is strongly irreducible and proximal. There are constants C, β such that for all $n \in \mathbb{N}$ and all $t \geq 0$:*

$$\mathbb{P}\left(\log\left(\frac{\|\overline{\gamma}_n\|}{|\lambda_1(\overline{\gamma}_n)|}\right) > t\right) \leq \sum_{k=1}^{\infty} C\exp(-\beta k)N(\nu)\left(\frac{t}{k} - C, +\infty\right). \tag{127}$$

Proof. Take $(X, \tilde{\kappa})$ the ping-pong extraction as constructed in Theorem 4.9. Write $\mathbb{A}$ for $\mathbb{A}^\varepsilon$ the associated alignment relation and $\mathbb{A}'$ for $\mathbb{A}^{\frac{\varepsilon}{2}}$. Consider $(\gamma, \gamma') \sim \kappa_a^{\otimes 2}$. We write:

$$F(g,h,g') := \frac{\mathbb{1}(g\mathbb{A}h \wedge (gh)\mathbb{A}g')}{\mathbb{P}(\gamma\mathbb{A}h \wedge (\gamma h)\mathbb{A}\gamma')}(1-2\rho). \tag{128}$$

Now we consider $(\gamma_n) \sim \nu^{\otimes \mathbb{N}}$ and (w_n) a random sequence such that $(\widetilde{\gamma}_n^w)$ is an ornamented Markov chain in $(X, \tilde{\kappa})$. Fix an integer $n \in \mathbb{N}$ and write m_n for the largest odd integer such that $\overline{w}_{m_n} \leq n$. Now for all $0 \leq c \leq \frac{m_n - 1}{2}$, we write :

$$\begin{aligned}
h_c &:= \gamma^w_{m_n - 2c+1} \cdots \gamma^w_{m_n} \gamma_{\overline{w}_{m_n+1}} \cdots \gamma_{n-1} \gamma^w_0 \cdots \gamma^w_{2c} \\
g_c &:= \gamma^w_{m_n - 2c} \\
g'_c &:= \gamma^w_{2c+1}.
\end{aligned}$$

Then consider $(\tau_c)_{c\in\mathbb{N}}$ independent, uniform in $[0,1]$ and independent of $(\widetilde{\gamma}_k^w)_{k\in\mathbb{N}}$. Now take c_n to be the smallest integer such that $c_n = \frac{m_n - 1}{2}$ or:

$$\tau_c \leq F(g_c, h_c, g'_c). \tag{129}$$

The integer c_n is a stopping time for the filtration $(\mathcal{F}_c)$ generated by $(\widetilde{\gamma}^w_{2k})_{k\in\mathbb{N}}$ and by $(\tau_{c'}, g_{c'}, g'_{c'})_{c'<c}$. Moreover, we have $\mathbb{P}(c_n = c \,|\, \mathcal{F}_c, c_n \geq c) = 1 - 2\rho$ because the distribution of (g_c, g'_c) knowing $\mathcal{F}_c$ is precisely $\kappa_a^{\otimes 2}$. This means that c_n is independent of $(\widetilde{\gamma}^w_{2k})_{k\in\mathbb{N}}$ and has finite exponential moment. Then by Lemma B.40 and Remark B.41, we have for some C and some $\beta > 0$:

$$\mathbb{P}\left(\sum_{k=\overline{w}_{m_n-2c+1},0}^{n-1,\overline{w}_{2c_n}-1} N(\gamma_k) > t\right) \leq \sum_{k=1}^{\infty} C\exp(-\beta k) N(\nu)\left(\frac{t}{k} - C, +\infty\right). \tag{130}$$

Moreover, we have $h_{c_n}\mathbb{A}'(\gamma^w_{2c+1}\cdots\gamma^w_{m_n-2c})\mathbb{A}' h_{c_n}$. Then by Corollary 2.27, we have:

$$\begin{aligned}|\lambda_1(\overline{\gamma}_n)| &\geq \mu_1(\gamma^w_{2c+1}\cdots\gamma^w_{m_n-2c})\mu_1\left(h_{c_n}\right)\frac{\varepsilon^2}{4} \\ &\geq \mu_1(\gamma^w_{2c+1}\cdots\gamma^w_{m_n-2c})\frac{\varepsilon^2}{4}\prod_{k=\overline{w}_{m_n-2c+1},0}^{n-1,\overline{w}_{2c_n}-1}\mu_d(\gamma_k).\end{aligned}$$

And by sum-multiplicativity, we have:

$$\mu_1(\overline{\gamma}_n) \leq \mu_1(\gamma^w_{2c+1}\cdots\gamma^w_{m_n-2c})\prod_{k=\overline{w}_{m_n-2c+1},0}^{n-1,\overline{w}_{2c_n}-1}\mu_1(\gamma_k).$$

Then we conclude using (130). □

5.5 About modelization

Definition 5.26 (Wasserstein distance)**.** *Let η, ν be probability distributions on a metric space (X, d). We define the L^1 Wasserstein distance $\mathcal{W}_{\mathrm{d},1}(\eta, \nu)$ as the minimum of $\mathbb{E}(\mathrm{d}(g, h))$ where $g \sim \nu$ and $h \sim \eta$ are defined on the same probability space but not necessarily independent.*

In the proof of the main results, all constants are constructed explicitly. In fact they can be measured with an algorithm. Moreover, in the invertible case, we can construct the constants C and β of each Theorems can be taken continuous with respect to the Wasserstein distance associated to d the distance in $\mathbf{P}\mathrm{End}(E)$ for Theorems 1.2 and 1.3 and to the metric induced by $N + \mathrm{d}$ for Theorem 1.5. It is possible to give explicit bounds on the constants C and β that depend on Wasserstein continuous quantities. Now the principal issue is that a generic probability distribution ν on $\mathrm{End}(E)$ tends to have rank 0. Indeed, if we take any nontrivial weighted barycenter of ν and any rank 0 distribution like δ_0 for example, we get a rank 0 distribution. In fact a satisfying result would be to prove that for any strongly irreducible and proximal probability distribution, we have Theorems 5.8 and 5.9 with the conditional probability $\mathbb{P}(\cdot|\overline{\gamma}_n \neq 0)$ of $\mathbb{P}(\cdot|\overline{\gamma}_n u \neq 0)$. These results seem to be true and a similar proof strategy should work. The main ingredient is to note that the extractions constructed in Section 4 do not depend on whether the partial products $\gamma_a \cdots \gamma_b$ are 0 or not for all pair $a < b$.

A Basics in Finite dimensional geometry

In this section we will consider $\mathbb{K}$ a local field *i.e,* a field endowed with an absolute value that makes it locally compact. We will also consider normed $\mathbb{K}$-vector spaces *i.e,* vector spaces endowed with a sub-additive norm $\|\cdot\|$ that takes non negative real values and such that for all $\lambda \in \mathbb{K}, x \in E$, we have $\|\lambda x\| = |\lambda|\|x\|$. We call unitary a vector or a scalar that has norm one, we will also make sure that the ultra-metric norms we introduce are good norms in the sense that all vector line contains at least one unitary vector.

In this section we will introduce a vocabulary between local fields that allows us to prove the results of the article without having to distinguish cases for every lemma. This appendix is interesting for two reasons, the first is that we construct the Cartan decomposition on locally compact fields, the second reason is that the intuition that one needs to have to understand proximal and strongly irreducible random walks is a mix of the Euclidean and ultra-metric views.

A.1 Real and complex Euclidean spaces

In this section we assume $\mathbb{K}$ to be an Archimedean locally compact field *i.e,* $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Given $z \in \mathbb{C}$, we write $\mathrm{re}(z)$ and $\mathrm{im}(z)$ for the real and imaginary parts of z, we write $\overline{z}$ for the complex conjugate of z and $|z|$ for the modulus of z.

Definition A.1 (Scalar product)**.** *Let E be a $\mathbb{K}$-vector space. We call scalar product on E a left-anti-linear and right-linear map $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{K}$ that satisfies:*

- *symmetry: for all $x, y \in E$, we have $\langle x, y \rangle = \overline{\langle y, x \rangle}$,*
- *separability: for all $x \in E \setminus \{0\}$, $\langle x, x \rangle \in \mathbb{R}_{>0}$.*

We call Euclidean (or standard Archimedean) vector space a finite dimensional vector space endowed with a scalar product.

Remark A.2. *One can also see a scalar product as an anti-linear[8] bijection $E \to E^*; x \mapsto x^\top$ where $x^\top : y \mapsto \langle x, y \rangle$ and the separability condition becomes $x^\top(x) \in \mathbb{R}_{>0}$ for all $x \neq 0$. In physics, the map $\cdot^\top$ is sometimes written $\cdot^\dagger$ in the complex case to distinguish both cases but this is not our aim to distinguish local fields.*

Definition A.3 (Euclidean norm)**.** *Let E be a Euclidean vector space and $x \in E$. We call norm of x and write $\|x\|$ the quantity:*

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

Remark A.4. *Let E be a Euclidean space and $x, y \in E$. We have $|\langle x, y \rangle| \leq \|x\|\|y\|$.*

Proof. Take two unitary vectors $x, y \in E$, take $\lambda \in \mathbb{C}$ a unitary complex number such that $\langle x, y \rangle$ is in the same half line as λ. Then we have $\overline{\lambda}\langle x, y \rangle = |\langle x, y \rangle| \in \mathbb{R}$ and

$$\begin{aligned}
0 &\leq \langle \lambda x - y, \lambda x - y \rangle \\
0 &\leq \langle \lambda x, \lambda x \rangle + \langle y, y \rangle - \langle \lambda x, y \rangle - \langle y, \lambda x \rangle \\
0 &\leq |\lambda|^2\|x\|^2 + \|y\|^2 - \overline{\lambda}\langle x, y \rangle - \lambda\overline{\langle x, y \rangle} \\
0 &\leq \|x\|^2 + \|y\|^2 - 2|\langle x, y \rangle| \\
|\langle x, y \rangle| &\leq 1
\end{aligned}$$

Then when x and y are not unitary, we have by linearity, $x = \|x\|a$ and $y = \|y\|b$ for some $a, b \in E$ unitary and then $\langle x, y \rangle = \|x\|\|y\|\langle a, b \rangle$. Moreover by the above reasoning we have $|\langle a, b \rangle| \leq 1$ so $\langle x, y \rangle \leq \|x\|\|y\|$. ☐

Proposition A.5. *Let E be a Euclidean vector space. There is a metric* d *on E defined by* $\mathrm{d}(x, y) = \|x - y\|$ *and this metric makes E locally compact i.e, the unit ball is compact.*

[8] a map f is said to be anti-linear if it is an additive group morphism and if for all $\lambda \in \mathbb{K}$, we have $f(\lambda x) = \overline{\lambda} f(x)$. Note that real anti-linear maps are linear.

Proof. The distance is symmetric by bi-linearity of the scalar product. The distance between two distinct points is never zero by separability of the scalar product. The distance satisfies the triangular inequality because for all $x, y \in E$, we have:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\mathrm{re}\langle x, y\rangle.$$

Moreover, by Remark A.4, we have $2\mathrm{re}\langle x, y\rangle \leq 2|\langle x, y\rangle| \leq 2\|x\|\|y\|$ so $\|x+y\|^2 \leq (\|x\|+\|y\|)^2$ and the square root of that is the triangular inequality for d. Then note that if we take an orthonormal basis of E, we have an identification between E and $\mathbb{R}^d$ for $d = \dim_{\mathbb{R}}(E) = \dim_{\mathbb{K}}(E)[\mathbb{K} : \mathbb{R}]$, note that the unit ball is included in $[-1,1]^d$ and contains $\left[-d^{-\frac{1}{2}}, +d^{-\frac{1}{2}}\right]^d$ so the topology of the norm is the product topology. Moreover, the unit ball is closed in $[-1,1]^d$ because the norm is continuous for its own topology and $[-1,1]^d$ is compact for the product topology by Tychonoff's theorem and Weierstrass' lemma. □

Definition A.6 (Orthogonality)**.** *Let E be a $\mathbb{K}$-vector space endowed with a scalar product. We say that two vectors $x, y \in E$ are orthogonal and write $x \perp y$ when $\langle x, y\rangle = 0$. Given an arbitrary subset $A \subset E$, we write $A^\perp$ the set $\{x \in E; \forall y \in A, x \perp y\}$. Given another arbitrary subset $B \subset E$, we write $B \perp A$ if $x \perp y$ for all $x \in A$ and all $y \in B$, note that it means that $B \subset A^\perp$.*

Note that the binary relation $\perp$ determines the scalar product $\langle \cdot, \cdot \rangle$ up to multiplication y a positive real number.

Definition A.7 (Orthogonal sum)**.** *Let E be a Euclidean space. We say that a given family of subspaces $(V_i)_{i \in I}$ is in orthogonal sum if for all $J \subset I$, we have:*

$$\left(\bigoplus_{j \in J} V_j\right) \perp \left(\bigoplus_{i \notin J} V_i\right).$$

Proposition A.8. *Let E be a Euclidean vector space and $(V_i)_{i \in I}$ a family of vector subspaces. Then $(V_i)_{i \in I}$ is in orthogonal sum if and only if we have $V_i \perp V_j$ for all $i \neq j$.*

Proof. One sense is trivial, just take $J = \{j\}$ and $V_j \perp \left(\bigoplus_{i \neq j} V_i\right)$ directly implies that $V_i \perp V_j$ for all $i \neq j$. Then if we take an arbitrary $J \subset I$ and two vector $x \in \bigoplus_{j \in J} V_j$ and $y \in \bigoplus_{i \notin J} V_i$, we may write $x = \sum_{j \in J} x_j$ with $x_j \in V_j$ for all j and $y = \sum_{i \notin J} y_i$ with $y_i \in V_i$ for all i. Then for all i, j, we have $\langle x_j, y_i\rangle = 0$ and $\langle x, y\rangle = \sum_{j \in J} \sum_{i \notin J} \langle x_j, y_i\rangle = 0$. □

Definition A.9 (Orthonormal basis)**.** *Let E be a Euclidean vector space. We call orthonormal basis of E a basis $(e_1, \dots, e_d)$ such that for all i, j, we have $\langle e_i, e_j\rangle = \delta_{i,j}$.*

Lemma A.10. *Let E be a Euclidean vector space. There exists an orthonormal basis $(e_1, \dots, e_d)$ and for all $x \in E$, if we write $x_i = \langle e_i, x\rangle$, we have $x = \sum x_i e_i$ and for all $x, y \in E$, we have $\langle x, y\rangle = \sum_i \overline{x_i} y_i$.*

Proof. We construct the orthonormal basis inductively. For all $k \in \{1, \dots, d\}$, we take e_k to be any unitary vector in $\{e_1, \dots, e_k\}^\perp$. Then note that $(e_i^\top)$ form a basis of E^* so a vector x is characterized by the family $(x_1, \dots, x_d)$ where $x_i = e_i^\top x$ then since $\cdot^\top$ is anti-linear, we have $x^\top = \sum \overline{x_i} e_i^\top$ and $\langle x, y\rangle = \sum_{i,j} \overline{x_i} y_j \langle e_i, e_j\rangle = \sum_i \overline{x_i} y_i$. □

Lemma A.11 (Contraction property)**.** *Let $V \subset E$ be a vector subspace. We have $E = V \oplus V^\perp$ and write π_V the projection onto V along $V^\perp$. Then π_V is contracting. Moreover, given a decomposition $E = V \oplus W$, we have $W \perp V$ if and only if the projection onto V along W is contracting.*

Proof. Write $k = \dim(V)$ and $d = \dim(E)$. Take $(e_1, \dots e_k)$ an orthonormal basis of V and $(e_{k+1}, \dots, e_d)$ an orthonormal basis of $V^\perp$. Then $(e_1, \dots, e_d)$ is an orthonormal basis of E, indeed for all $0 \leq i, j \leq k$, such that $i \leq k < j$ or $j \leq k < i$ we have $\langle e_i, e_j\rangle = 0 = \delta_{i,j}$ because $V \perp V^\perp$ and if $i, j \leq k$ of $k < i, j$ then $\langle e_i, e_j\rangle = \delta_{i,j}$ because we have taken orthonormal bases. As a consequence, for all $x \in E$, we have $\|x\|^2 = \sum_{i=1}^d |x_i|^2 \geq \sum_{i=1}^k |x_i|^2 = \|\pi_V(x)\|^2$ so π_v is contracting. Conversely, take V, W two vector spaces that are not orthogonal, then there are two vectors $x \in V, y \in W$ such that $\langle x, y\rangle \neq 0$ and up to a scalar multiplication, we may assume that $\langle x, y\rangle > 0$, then for all $t > 0$, we have $\|x - ty\|^2 = \|x\| + t^2\|y\|^2 - 2t\langle x, y\rangle$ so for $t = \frac{\langle x,y\rangle}{\|y\|^2}$, we have $\|x - ty\|^2 = \|x\| - \frac{\langle x,y\rangle^2}{\|y\|^2} < \|x\|$ but x is the projection of $x - ty$ along W so this projection is not contracting. □

Corollary A.12. *Let E be a Euclidean vector space and $(e_1, \dots, e_d)$ a basis of E composed of unitary vectors. Then (e_i) is an orthonormal basis if and only if for all $x = \sum x_i e_i$ we have $\|x\| \geq \max |x_i|$.*

A.2 Standard ultra-metric vector spaces and orthogonality

In this section, we assume $\mathbb{K}$ to be a non-Archimedean local field. That means that $\mathbb{K}$ is a field endowed with an absolute value $|\cdot| = \mathbb{K} \to \mathbb{R}_{\geq 0}$ that satisfies:

(S) for all $a \in \mathbb{K}$, $|a| = 0_{\mathbb{R}}$ if and only if $a = 0_{\mathbb{K}}$,

(M) for all $a, b \in \mathbb{K}$, $|a \cdot b| = |a| \cdot |b|$,

(UM) for all $a, b \in \mathbb{K}$, $|a + b| \leq \max\{|a|, |b|\}$,

(LC) the distance $\mathrm{d} : (a, b) \mapsto |a - b|$ makes the unit ring $\mathcal{R}(\mathbb{K}) := \mathcal{B}(0_{\mathbb{K}}, 1_{\mathbb{R}}) = \{a \in \mathbb{K}; |a| \leq 1\}$ compact.

Remark A.13. *Let d be an ultra-metric distance on a space X. That means that d is a distance map such that for all $x, y, z \in X$, we have $\mathrm{d}(x, z) \leq \max\{\mathrm{d}(x, y), \mathrm{d}(y, z)\}$. Then for all $x, y, z \in X$ such that $\mathrm{d}(x, z) < \max\{\mathrm{d}(x, y), \mathrm{d}(y, z)\}$, we have $\mathrm{d}(x, y) = \mathrm{d}(y, z)$. In other words all triangle has two sides of equal length.*

Proof. Consider three points $x, y, z \in X$ and assume that $\mathrm{d}(x, z) < \max\{\mathrm{d}(x, y), \mathrm{d}(y, z)\}$. Then by symmetry and up to exchanging x with z, one may assume that $\mathrm{d}(y, x) \geq \mathrm{d}(y, z)$. Then by triangular inequality applied to the triple y, x, z, we have $\mathrm{d}(y, x) \leq \max\{\mathrm{d}(y, z), \mathrm{d}(z, x)\}$ but by assumption $\mathrm{d}(z, x) < \mathrm{d}(y, z)$ so we necessarily have $\mathrm{d}(y, x) \leq \mathrm{d}(y, z)$ and by double inequality $\mathrm{d}(y, x) = \mathrm{d}(y, z)$. □

Remark A.14. *Let $(a_i)_{i\in I}$ and $(b_i)_{i\in I}$ be two finite families of non-negative real numbers. Then we have:*

$$\max_{i\in I} a_i b_i \leq \max_{i\in I} a_i \max_{i\in I} b_i$$

It is known that an ultra-metric field $\mathbb{K}$ is either finite or a finite extension of one of the following: $\mathbb{Q}_p$ for p a prime number; or $\mathbb{F}((t))$ for $\mathbb{F}$ a finite field. For the study of random walks one does not need to have the full classification in mind, the following proposition is enough. We write $\mathbb{K}^*$ for the multiplicative group $\mathbb{K} \setminus \{0\}$.

Proposition A.15. *Let $(\mathbb{K}, |\cdot|)$ be an ultra-metric locally compact field. The image of $\mathbb{K}^*$ by $|\cdot|$ is a discrete sub-group of $\mathbb{R}_{>0}$, i.e, there is a real number $r > 0$ and a group morphism $v : \mathbb{K}^* \to \mathbb{Z}$ such that $|x| = r^{v(x)}$ for all $x \in \mathbb{K}^*$.*

Proof. First note that by (LC), the unit group $\mathcal{U}(\mathbb{K}) := \mathcal{R}(\mathbb{K})^{\times} := \{a \in \mathbb{K}; |a| = 1\}$ is compact, indeed is is closed because $|\cdot|$ is trivially continuous and it is included in the ring of integers R which is compact. Then note that for all $\phi \neq 0$, the map $a \mapsto \phi a$ is continuous and bijective. Then take a real $r \in |\mathbb{K}^*|$ and a scalar $\rho \in \mathbb{K}$ such that $|\rho| = r$. Then the r-sphere $\{a \in \mathbb{K}; |a| = r\}$ is the image of the unit group by the multiplication by ρ so it is compact. Moreover, if we assume that $r \geq \frac{1}{2}$, then the ball $\mathcal{B}\left(\rho, \frac{1}{2}\right) := \left\{a \in \mathbb{K}; |a - \rho| < \frac{1}{2}\right\}$ is included in the r-sphere. Indeed, if we take $a \in B\left(\rho, \frac{1}{2}\right)$, we have $|a| \leq \min\{|a - \rho|, |\rho|\} = |\rho| \leq \min\{|a|, |\rho - a|\}$ so $|a| = r$. Now note that for all $x \in \mathbb{K}$, $x \in \mathcal{B}\left(x, \frac{1}{2}\right)$ so we have:

$$R = \bigcup_{x\in R} \mathcal{B}\left(x, \frac{1}{2}\right) = \mathcal{B}\left(0_{\mathbb{K}}, \frac{1}{2}\right) \bigcup_{x\in R; |x|\geq\frac{1}{2}} \mathcal{B}\left(x, \frac{1}{2}\right).$$

Since R is compact, there is a finite family $0_{\mathbb{K}} =: x_0, x_1, \dots, x_N$ such that $\frac{1}{2} \leq |x_i| \leq 1$ for all $i \in \{1, \dots, N\}$ and:

$$R = \bigcup_{i=0}^{N} \mathcal{B}\left(x_i, \frac{1}{2}\right).$$

So for all $\rho \in \mathbb{K}$ such that $\frac{1}{2} \leq |\rho| \leq 1$ we have $|\rho| \in \mathcal{B}\left(x_i, \frac{1}{2}\right)$ for some i, as a consequence, we have $|\rho| = |x_i|$ so the multiplicative group $|\mathbb{K}^*|$ has only finitely many elements in $\left[\frac{1}{2}, 1\right]$, so it is discrete. □

Then we want to study the monoid $\Gamma := \mathrm{End}_{\mathbb{K}}(E)$ for $E = \mathbb{K}^d$ for $d \geq 2$ an integer. Note that when $\mathbb{K}$ is finite, Γ also is so random walks on Γ are simply random walks on a finite set. From now on, we will assume that $\mathbb{K}$ is not finite. We endow E with a norm $\|\cdot\| : E \to \mathbb{R}$ defined by:

$$\|(x_1, \dots, x_d)\| := \max_{i \in \{1, \dots, d\}} |x_i|. \tag{131}$$

Note that the values that the norm $\|\cdot\|$ may take are the same as the absolute value $|\cdot|$ may take, so by Proposition A.15, the norm $\|\cdot\|$ takes values in a discrete group.

Proposition A.16. *The norm $\|\cdot\|$ defined above satisfies the following properties.*

(S) For all $x \in E$, we have $\|x\| = 0$ if and only if $x = 0$.

(H) For all $x \in E$ and $a \in \mathbb{K}$, we have $\|ax\| = |a|\|x\|$.

(UM) For all $x, y \in E$, we have $\|x + y\| \leq \max\{\|x\| + \|y\|\}$.

(LC) The distance $\mathrm{d} : (x, y) \mapsto \|x - y\|$ makes $\mathbb{K}$ second-countable and the unit ball $B(1)\{x \in E; \|x\| \leq 1\}$ compact.

Proof. Take $a \in \mathbb{K}$, $x = (x_1, \dots, x_d) \in E$ and $y = (y_1, \dots, y_d) \in E$. If $\|x\| = 0$ then $|x_i| = 0$ for all $i \in \{1, \dots, d\}$ so $x = (0, \dots, 0) = 0_E$. Then $\|ax\| = \max |a||x_i| = |a| \max |x_i|$. Then $\|x + y\| = \max_i \|x_i + y_i\| \leq \max_i \max\{|x_i|, |y_i|\} = \max\{\max_i x_i, \max_i y_i\}$. Then we have $B(1) = R^d$ and R is compact by (LC) so $B(1)$ also is. □

Definition A.17 (Ultra-metric norm)**.** *Let E be a $\mathbb{K}$-vector space. We say that a map $\|\cdot\| : E \to \mathbb{R}_{\geq 0}$ is an ultra-metric norm if all the following conditions hold:*

1. *For all $x \in E$, we have $\|x\| = 0$ if and only if $x = 0$.*
2. *For all $x \in E$ and all $a \in \mathbb{K}$, we have $\|ax\| = |a|\|x\|$.*
3. *For all $x, y \in E$, we have $\|x + y\| \leq \max\{\|x\| + \|y\|\}$.*

Definition A.18 (Standard ultra-metric space)**.** *Let E be a $\mathbb{K}$ vector space endowed with an ultra-metric norm $\|\cdot\|$. We say that a given basis $(e_1, \dots, e_d)$ of E is standard for $\|\cdots\|$ if:*

$$\forall (x_1, \dots, x_d) \in \mathbb{K}^d, \left\|\sum_{k=1}^{d} x_k e_k\right\| = \max_{1 \leq k \leq d} |x_k|. \tag{132}$$

We call standard ultra-metric vector space the data $(E, \|\cdot\|)$ of a finite dimensional $\mathbb{K}$-vector space E and of an ultra-metric norm $\|\cdot\|$ that admits a standard basis.

Remark A.19. *Note that just like Euclidean and Hermitian spaces, standard ultra-metric spaces are characterized, up to an isometry, by their dimension. Indeed* (132) *means that a standard basis of a vector space E induces an isometry between $(E, \|\cdot\|)$ and $(\mathbb{K}^d, \|\cdot\|)$. Moreover, we can characterize the property of being an orthonormal basis with Corollary A.12, indeed, by ultra-metric inequality, we always have $\left\|\sum_{k=1}^{d} x_k e_k\right\| \leq \max_{1 \leq k \leq d} |x_k|$ when the (e_i)'s are unitary so assuming only $\left\|\sum_{k=1}^{d} x_k e_k\right\| \geq \max_{1 \leq k \leq d} |x_k|$ for all $x \in \mathbb{K}^d$ implies* (132)*.*

Now we do a little warm up to get familiar with standard ultra-metric vector spaces.

Proposition A.20 (Sub-spaces)**.** *Let $(E, \|\cdot\|)$ be a standard vector space and F a subspace of E. Then F endowed with the restriction of $\|\cdot\|$ to F is also a standard ultra-metric vector space.*

Proof. We prove it only for F of co-dimension 1 and conclude by induction. Take d an integer and F an hyperplane in $E = \mathbb{K}^d$. Then there is a linear form $f \in E^* \setminus \{0\}$ such that $F = \ker(f)$, we write $f_i = f(e_i)$ for all $i \in \{1, \ldots, d\}$. Then take i_0 such that $|f_{i_0}| = \max |f_i| =: \|f\|$. Then if we replace f by $\frac{f}{f_{i_0}}$ we may assume that $f_{i_0} = 1$ without changing the kernel of f. Then for all $i \in \{1, \ldots, d\}$, we write e'_i the projection of e_i to F along e_{i_0}. Now take $x \in F$ any vector, one has $x = \sum_{i=1}^d x_i e_i$ in the basis $(e_i)_{1 \leq i \leq d}$ but $x \in F$ so $x = \sum_{i=1}^d x_i e'_i = \sum_{i \neq i_0} x_i e'_i$ because $e'_{i_0} = 0$ by definition. Then one has $f(x) = 0$ so $\sum_{i=1}^d x_i f_i = 0$ and $f_{i_0} = 1$ so $\sum_{i \neq i_0} f_i x_i = x_{i_0}$, then by *(UM)*, one has $\max_{i<d} |x_i f_i| \geq |x_{i_0}|$ and $|f_i| \leq 1$ for all i so $\max_{i \neq i_0} |x_i| \geq |x_d|$ so we have $\|x\| = \max_{i \neq i_0} |x_i|$ so $(e'_i)_{i \neq i_0}$ is a standard basis of F. ☐

Proposition A.21 (Dual space)**.** *Let $(E, \|\cdot\|)$ be a standard ultra-metric vector space. The dual space $E^* := \mathrm{Hom}_{\mathbb{K}}(E, \mathbb{K})$ is a standard ultra-metric vector space for the norm $\|\cdot\| : f \mapsto \max_{x \in E \setminus \{0\}} \frac{|f(x)|}{\|x\|}$.*

Proof. Take $(e_1, \ldots, e_d)$ a standard basis of E. Note that since (e_i) is a basis, a linear form f in E^* is characterized by coordinates $(f_1, \ldots, f_d) \in \mathbb{K}^d$ with $f_i = f(e_i)$ for all i. With this notation we have $\|f\| \geq \max |f_i|$ because the e_i's are unitary and then if we take x any vector, we have $|f(x)| = \max |f_i x_i| \leq \max |f_i| \max |x_i|$ so $\|f\| = \max |f_i|$. Note that we also have $f = \sum f_i e'_i$ with e'_i characterized by $e'_i(e_j) = \delta_{i,j}$ for all $1 \leq i, j \leq d$. So the family (e'_i) is a standard basis of E^*. ☐

Proposition A.22 (Homomorphisms)**.** *Let E, F be two standard ultra-metric vector spaces over the same field $\mathbb{K}$. The space $\mathrm{Hom}(E, F)$ endowed with the operator norm:*

$$\|h\| := \max_{x \in E \setminus \{0\}} \frac{\|h(x)\|}{\|x\|}$$

is a standard ultra-metric vector space.

Proof. Consider (e_i) a standard basis of E and (f_i) a standard basis of F. Then for every pair (i, j) write $f_j e'_i \in \mathrm{Hom}(E, F)$ the map that sends e_i to f_j and e_k to 0 for all $k \neq i$. Note that $f_j e'_i$ is unitary for all (i, j), we want to show that the family $(f_j e'_i)_{i,j}$ is actually a standard basis of $\mathrm{Hom}(E, F)$. Consider an arbitrary $h \in \mathrm{End(E)}$. For all i, one has $h(e_i) \in F$ so one may write $h(e_i) = \sum_j h_{i,j} f_j$. Then we have $h = \sum_{i,j} h_{i,j} f_j e'_i$ so $(f_j e'_i)$ is indeed a basis of $\mathrm{Hom}(E, F)$. Then if we take $x = \sum_i x_i e_i \in E$ any vector, we have by ultra-metric inequality and by Remark A.14:

$$\|h(x)\| = \max_j \left| \sum_i h_{i,j} x_i \right| \leq \max_{i,j} |h_{i,j} x_i| \leq \max_{i,j} |h_{i,j}| \max_i |x_i|.$$

The above is true for all x so $\|h\| \leq \max_{i,j} |h_{i,j}|$. Now if we take two indices i_0, j_0 such that $|h_{i_0,j_0}|$ is maximal we have $\|h(e_{i_0})\| = |h_{i_0,j_0}|$ and e_{i_0} is unitary so $\|h\| \geq \max_{i,j} |h_{i,j}|$. By double inequality, we have $\|h\| = \max_{i,j} |h_{i,j}|$ so $(f_j e'_i)$ is indeed a standard basis of $\mathrm{Hom}(E, F)$ and as a consequence $\|\cdot\|$ is a standard ultra-metric norm on $\mathrm{Hom}(E, F)$. ☐

Note that this is false for standard Archimedean vector spaces. For example if we take $a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ then a and b are not co-linear, yet $\|a\| + \|b\| = \|a + b\| = 3$ so the operator norm is not strictly convex and therefore not Euclidean.

Definition A.23 (Orthogonal spaces and projections)**.** *Let $(E, \|\cdot\|)$ be an ultra-metric vector space and V, W two subspaces of E. We say that V, W are orthogonal and write $V \perp W$ if $V \cap W = \{0\}$ and if the projections π_V and π_W of $V + W$ to V along W and to W along V respectively are contracting. In general, we way that a projection π on E (i.e, a linear map such that $\pi \circ \pi = \pi$) is orthogonal if it is contracting for the norm $\|\cdot\|$.*

Remark A.24. *Let V, W be two supplementary spaces. Then the projection π_V is contracting if and only π_W is.*

Proof. Take x a vector such that $\|\pi_V(x)\| > \|x\|$. Then note that $x = \pi_V(x) + \pi_W(x)$, and $\|x\| < \max\{\|\pi_V(x)\|, \|\pi_W(x)\|\}$ so by Remark A.13, we have $\|\pi_V(x)\| = \|\pi_W(x)\|$. ☐

Definition A.25. *We say that a family of vector space $(V_i)_{i\in I}$ is orthogonal if for all $i \in I$, V_i is orthogonal to $\sum_{j\neq i} V_j$. We say that a family of vectors $(x_1, \ldots, x_k)$ is orthogonal if the family $(\mathbb{K}x_1, \ldots, \mathbb{K}x_k)$ is orthogonal. We say that $(x_1, \ldots, x_k)$ is orthonormal if moreover the x_k's are unitary.*

Proposition A.26. *Let E be a standard vector space of dimension d and $e_1, \ldots, e_d \in E$. The family $(e_1, \ldots, e_d)$ is a standard basis of E if and only if it is orthonormal.*

Proposition A.27 (Orthogonal decomposition)**.** *Let $(E, \|\cdot\|)$ be a standard ultra-metric vector space and V, W two orthogonal subspaces of E. For all $x \in V + W$, we have $\|x\| = \max\{\|\pi_V(x)\|, \|\pi_W(x)\|\}$.*

Proof. By *(UM)*, we have $\|x\| \leq \max\{\|\pi_V(x)\|, \|\pi_W(x)\|\}$, but by orthogonality, we also have $\|\pi_V(x)\| \leq \|x\|$ and $\|\pi_W(x)\| \leq \|x\|$. ☐

Corollary A.28 (Incomplete orthonormal basis)**.** *Let E be a standard ultra-metric vector space and V, W be two orthogonal supplementary spaces in E. Let $(e_1, \ldots, e_k)$ be an orthonormal basis of V and $(e_{k+1}, \ldots, e_d)$ an orthonormal basis of W. Then $(e_1, \ldots, e_k)$ is an orthonormal basis of E.*

Proof. Then if we take a vector $x \in E$, we may write $x = \sum x_i e_i$ because V and W are supplementary spaces. Then we have $\|\pi_V(x)\| = \max\{|x_i|\ 1 \leq i \leq k\}$ and $\|\pi_W(x)\| = \max\{|x_i|\ k < i \leq d\}|$ so by Proposition A.27 $\|x\| = \max\{|x_i|\ 1 \leq i \leq d\}$ and $(e_1, \ldots, e_k)$ is indeed an orthonormal basis of E. ☐

Proposition A.29 (Orthogonal supplementary)**.** *Let E be a standard ultra-metric vector space and V a subspace of E. Then there exists subspace $W \subset E$ such that $V \perp W$ and $V + W = E$. However V is not unique in general.*

Proof. Write $c := \dim(E) - \dim(V)$ for the co-dimension of V and take $(e_1, \ldots, e_d)$ a standard basis of E. Then take $f : E \to \mathbb{K}^c$ a linear map such that $V = \ker(f)$. Then write $(f_{i,j})_{1\leq i\leq c, 1\leq j\leq d}$, the coordinates of f, and take $j_1 \in \{1, \ldots, d\}$ the index such that $|f_{1,j_1}|$ is maximal, it is not 0 because f is invertible, then for all j, replace $f_{1,j}$ with $\frac{f_{1,j}}{f_{1,j_1}}$ for all $i > 1$, replace $f_{i,j}$ with $f_{i,j} - \frac{f_{i,j_1}}{f_{1,j_1}} f_{1,j}$, this does not change the kernel of f. Then we construct a sequence $j_1, \ldots, j_c$ all distinct and such that $f_{i,j_i} = 1$ for all i and $|f_{i,j}| < 1$ for all i, j. Then up to rearranging the line of f (this does not change the kernel), we may assume that $j_1 < \cdots < j_c$ and write $W = \langle e_{j_1}, \ldots, e_{j_c}\rangle$, then for all x, we have $\pi_W(x) = \sum f_{i,j} x_j e_{j_i}$ which is contracting because $\max f_{i,j} = 1$ and $\pi_V(x) = x - \sum f_{i,j} x_j e_{j_i}$ which is also contracting by the ultra-metric inequality. We have $\pi_W(x) \in W$ and $x = \pi_V(x) + \pi_W(x)$ directly and to show that $\pi_V(x) \in V$, we simply do the computation, $f(\pi_V(x)) = (\sum f_{i,j} x_j (1 - f_{i,j_i}))_{1\leq i\leq c} = 0$ because $f_{i,j_i} = 1$ by definition. ☐

Lemma A.30. *Let E be a standard $\mathbb{K}$-vector space. Let $w \in E^*$ and $v \in E$ be unitary vectors. Then $|w(v)| = 1$ if and only if $v \perp \ker(w)$.*

Proof. First assume that $|w(v)| = 1$. Take $x \in E$, one has $x = \frac{w(x)}{w(v)} v + \left(x - \frac{w(x)}{w(v)} v\right)$ and $w\left(x - \frac{w(x)}{w(v)} v\right) = 0$ so $\frac{w(x)}{w(v)} v$ is the projection of x onto $\mathbb{K}v$ along $\ker(w)$. Moreover w is unitary so we have $\left|\frac{w(x)}{w(v)}\right| = |w(x)| \leq \|x\|$ and v is unitary so we have $\left\|\frac{w(x)}{w(v)} v\right\| = \left|\frac{w(x)}{w(v)}\right| \leq \|x\|$. This means that the projection onto $\mathbb{K}v$ is contracting and by Remark A.24, $v \perp \ker(w)$.

Now we assume that $v \perp \ker(w)$. Then take x a unitary vector such that $|w(x)| = 1$. By orthogonality the projection of x onto $\mathbb{K}v$ along $\ker(w)$ is contracting so $\left\|\frac{w(x)}{w(v)} v\right\| \leq 1$ and therefore $|w(v)| \geq 1$. Moreover, we always have $|w(v)| \leq 1$ because w and v are unitary so $|w(v)| = 1$. ☐

Lemma A.31. *Let E be a standard ultra-metric vector space, there is a contracting, norm-preserving map $E \to E^*, v \mapsto v^\top$ such that:*

$$\forall v \in E, \|v^\top(v)\| = \|v\|^2.$$

Proof. Take $(e_1,\dots,e_d)$ a standard basis of E and $(e_1^*,\dots,e_d^*)$ the associated basis of E^* (*i.e,* such that $e_i^*(e_j)=\delta_{i,j}$). Then for all $v=\sum v_ie_i\in E$, write j the smallest index such that $|v_j|=\|v\|$ and define $v^\top:=v_je^*j$. Note that with this definition, $\|v^\top\|=\|v\|$. Then take $x,y\in E$, if $\|x-y\|=\max\{\|x\|,\|y\|\}$ then this is direct, otherwise $\|x\|=\|y\|$ and for all index j such that $|x_j|=\|x\|$, we also have $|y_j|=\|x\|=\|y\|$ (otherwise $|x_j-y_j|=\|x\|$ which contradicts the hypothesis $\|x-y\|<\max\{\|x\|,\|y\|\}$). So we have $x^\top=x_je_j^*$ and $y^\top=y_je^j$ for the same j and $\|x^\top-y^\top\|=|x_j-y_j|\le\|x-y\|$. □

A.3 Exterior algebra and Cartan decomposition

In this section $(\mathbb{K},|\cdot|)$ is an Archimedean or Ultra-metric local field and $(E,\|\cdot\|)$ is a standard $\mathbb{K}$- vector space of dimension d.

Definition A.32. *Let E be a $\mathbb{K}$ vector space. For all $k\in\mathbb{N}$, we define $\bigotimes^k E'$ to be the space of k-linear forms i.e, the set of maps $\phi:E^k\to\mathbb{K}$ such that for all index $i\in\{1,\dots,k\}$ and for all family $(x_j)_{j\in\{1,\dots,k\}\setminus\{i\}}\in E^{k-1}$, the map $x_i\mapsto\phi(x_1,\dots,x_k)$ is in E^*. Then $\bigotimes^k E'$ is a vector space and we define $\bigotimes^k E$ to be the dual space of $\bigotimes^k E'$. Given a family $(x_1,\dots,x_k)\in E^k$, we write $x_1\otimes\cdots\otimes x_k:\phi\mapsto\phi(x_1,\dots,x_k)$.*

Proposition A.33. *Let E be a standard vector space with standard basis $(e_1,\dots,e_d)$. Then $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$ is a basis of $\bigotimes^k E$ and the norm induced on $\bigotimes^k E$ by the basis $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$ does not depend on the choice of the basis $(e_1,\dots,e_d)$.*

Proof. First we prove that $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$ generates $\bigotimes^k E$ as a vector space. Consider ϕ a k-linear form such that $\phi=0$ on the vector space generated by $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$. Then we have $\phi(e_{i_1},\cdots,e_{i_k})=0$ for all $1\le j_1,\dots,j_k\le d$. Consider a family $(x_1,\dots,x_k)\in E^k$ and for all indices $i\in\{1,\dots,d\},j\in\{1,\dots,k\}$, write $x_{i,j}$ for the i-th coordinate of x_j in the basis $(e_1,\dots,e_d)$. Then we have:

$$\begin{aligned}\phi(x_1,\dots,x_k)&=\sum_{i_1=1}^{d}x_{i_1,1}\phi(e_{i_1},x_2,\dots,x_k)\\&\ \ \vdots\\\phi(x_1,\dots,x_k)&=\sum_{i_1=1}^{d}\cdots\sum_{i_k=1}^{d}x_{i_1,1}\cdots x_{i_k,k}\phi(e_{i_1},\dots,e_{i_k})\\\phi(x_1,\dots,x_k)&=0.\end{aligned}$$

This means that $\phi=0$ so $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$ generates $\bigotimes^k E$ as a vector space.

Now consider a family of indices $1\le i_1,\dots,i_k\le d$ and write $\phi_{i_1,\dots,i_k}:(x_1,\dots,x_k)\mapsto x_{i_1,1}\cdots x_{i_k,k}$. Then $\phi=0$ on the space generated by all the $(e_{i'_1}\otimes\cdots\otimes e_{i'_k})$ but $e_{i_1}\otimes\cdots\otimes e_{i_k}$ so $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$ is a minimal generating family and therefore a basis.

Now we want to show that the norm induced by $(e_{i_1}\otimes\cdots\otimes e_{i_k})_{1\le i_1,\dots,i_k\le d}$ on $\bigotimes^k E$ does not depend on $(e_1,\dots,e_d)$. Consider $\alpha=x_1\otimes\cdots\otimes x_k\in\bigotimes^k E$ a pure k-vector. For all pair i,j of indices, write $x_{i,j}$ for the i-th coordinate of x_j in the basis $(e_1,\dots,e_d)$. For all family $1\le i_1,\dots,i_k\le d$ write $\alpha_{i_1,\dots,i_k}:=x_{i_1,1}\cdots x_{i_k,k}$. First assume that $\mathbb{K}$ is ultra-metric. Then by by Remark A.14, we have:

$$\begin{aligned}\alpha&=\sum\alpha_{i_1,\dots,i_k}e_{i_1}\otimes\cdots\otimes e_{i_k}\\\|\alpha\|&=\max|\alpha_{i_1,\dots,i_k}|\\\|\alpha\|&=\|x_1\|\cdots\|x_k\|.\end{aligned}$$

As a consequence, the norm of a pure k-vector does not depend on the choice of the basis. Then for all $a\in\bigotimes^k E$, the norm $\|a\|$ is the minimum for all path $a=\alpha_1+\cdots+\alpha_N$ (where the α_n's are pure k-vectors) of the quantity $\max\|\alpha_n\|$.

In the euclidean case, take $\beta : y_1 \otimes \cdots \otimes y_k$ another pure k-vector. Then for $\langle\cdot,\cdot\rangle$ the scalar product induced by the basis $(e_{i_1} \otimes \cdots \otimes e_{i_k})_{1 \le i_1,\dots,i_k \le d}$, we have

$$\langle\alpha,\beta\rangle = \sum_{i_1=1}^{d} \cdots \sum_{i_k=1}^{d} \alpha_{i_1,\dots,j_k} \overline{\beta}_{i_1,\dots,j_k}$$

$$\langle\alpha,\beta\rangle = \sum_{i_1=1}^{d} \cdots \sum_{i_k=1}^{d} \prod_{j=1}^{k} x_{i_j,j} \overline{y}_{i_j,j}$$

$$\langle\alpha,\beta\rangle = \prod_{j=1}^{k} \left(\sum_{i=1}^{d} x_{i,j} \overline{y}_{i,j} \right)$$

$$\langle\alpha,\beta\rangle = \prod_{j=1}^{k} \langle x_j, y_j \rangle.$$

It means that the scalar product of pure k-vectors does not depend on the choice of the basis of E. If we take $(e'_1,\dots,e'_d)$ another orthonormal basis, then the basis $(e'_{i_1} \otimes \cdots \otimes e'_{i_k})_{1 \le i_1,\dots,i_k \le d}$ is orthonormal for the scalar product induced by the basis $(e_{i_1} \otimes \cdots \otimes e_{i_k})_{1 \le i_1,\dots,i_k \le d}$, which means that they both induce the same norm on $\bigotimes^k E$. □

Definition A.34 (Exterior algebra)**.** *Let k be an integer and E be a vector space. We write $\bigwedge^k E$ the quotient of $\bigotimes^k E$ by the relation $\sim_\wedge$ which is the minimal linear relation such that for all map $\sigma : \{1,\dots,k\} \to \{1,\dots,k\}$, and for all $x_1,\dots,x_k \in E$, we have:*

$$x_1 \otimes \cdots \otimes x_k \sim_\wedge \varepsilon_{\mathbb{K}}(\sigma) x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(k)}.$$

Where $\varepsilon_{\mathbb{K}}$ is the image of the signature morphism on $\mathfrak{S}_k$ and 0 on the set of non invertible maps. Given $x_1,\dots,x_k \in E^k$, we write $x_1 \wedge \cdots \wedge x_k \in \bigwedge^k E$ for the $\sim_\wedge$-equivalence class of $x_1 \otimes \cdots \otimes x_k$.

Proposition A.35. *Let E be a vector space, let j,k be two integers. Let $x,x' \in \bigotimes^j E$ and $y,y' \in \bigotimes^k E$. If $x \sim_\wedge x'$ and $y \sim_\wedge y'$ then $x \otimes y \sim_\wedge x' \otimes y'$. Moreover, $x \otimes y \sim_\wedge y \otimes x$. In other words, if we write $v \in \bigwedge^j E$ for the $\sim_\wedge$ class of x and $w \in \bigwedge^k E$ for the $\sim_\wedge$ class of y, then the product $v \wedge w$ defined as the $\sim_\wedge$ class of $x \otimes y$ is well defined and the operation $\wedge$ defined that way is commutative.*

Remark A.36. *Definition A.34 gives a natural metric to $\bigwedge^k E$ which is the distance between sheets, it is well defined because the relation $\sim_\wedge$ is linear and closed. However this metric is not the one we want to use in the Archimedean case because we do not have $\|e_1 \wedge \cdots \wedge e_k\| = 1$ for an orthonormal family $(e_1,\dots,e_k)$.*

Definition A.37 (Norm on the exterior product)**.** *Let E be a standard vector space and let $1 \le k \le d$. We define the norm $\|\cdot\|$ on $\bigwedge^k E$ to be equal to a constant C times the distance between sheets in $\bigotimes^k E$. Where $C = 1$ in the ultra-metric case and $C = \sqrt{k!}$ in the Archimedean case.*

Proposition A.38. *Let E be a standard vector space and $1 \le k \le d$ and $(e_1,\dots,e_k)$ be an orthonormal family. We have $\|e_1 \wedge \cdots \wedge e_k\| = 1$.*

Proof. The norm on $\bigotimes^k E$ is standard so the distance between the sheet of 0 and the sheet of $x := e_1 \otimes \cdots \otimes e_k$ is the distance between 0 and its orthogonal projection on the sheet of x. In the ultra-metric case, we claim that x is an orthogonal projection of 0. Consider $(e_1,\dots,e_d)$ a completion of $(e_1,\dots,e_k)$ into an orthonormal basis. Take $y \sim_\wedge x$, we want to show that $\|y\| \ge \|x\|$. We write y in the basis $(e_{i_1} \otimes \cdots \otimes e_{i_k})$:

$$y = \sum_{i_1=1}^{d} \cdots \sum_{i_k=1}^{d} y_{i_1,\dots,i_k} e_{i_1} \dots e_{i_k} \tag{133}$$

The fact that $y \sim \wedge x$ implies that:

$$\sum_{(i_1,\dots,i_k) \in \mathfrak{S}_k} \sigma(i_1,\dots,i_k) y_{i_1,\dots,i_k} = 1 \tag{134}$$

By the ultra-metric inequality, at least one of the $y_{i_1,\dots,i_k}$ has absolute value 1 so $\|y\| \geq 1$. Now in the Euclidean case, (134) still holds and implies that:

$$\sum_{(i_1,\dots,i_k)\in\mathfrak{S}_k} y_{i_1,\dots,i_k}^2 \geq \frac{1}{k!} \tag{135}$$

It means that $\|y\| \geq \sqrt{\frac{1}{k!}}$. Moreover, for $y = \frac{1}{k!}\sum_{(i_1,\dots,i_k)\in\mathfrak{S}_k} \sigma(i_1,\dots,i_k)e_{i_1}\dots e_{i_k}$, we have $y \sim_\wedge x$ and $\|y\| = \sqrt{\frac{1}{k!}}$ so the distance between the sheets of x and 0 is $\sqrt{\frac{1}{k!}}$. In both cases, the sheets of x and 0 are at distance $1/C$ with C as in Definition A.37 so $\|e_1 \wedge \cdots \wedge e_k\| = 1$. □

Definition A.39. *Let E be a vector space, let j, k be two integers. Let $V \subset \bigwedge^j E$ and $W \subset \bigwedge^k E$ be two vector subspaces. We define $V \wedge W \subset \bigwedge^{j+k} E$ to be the vector space spanned by $\{v \wedge w \,|\, v \in V, w \in W\}$.*

Lemma A.40. *Let E be a standard vector space of dimension d. Let $1 \leq k \leq d$ and $x_1, \dots, x_k \in E \setminus \{0\}$. We have:*

$$\|x_1 \wedge \cdots \wedge x_k\| \leq \|x_1\| \cdots \|x_k\| \tag{136}$$

with equality if and only if $(x_1, \dots, x_k)$ is an orthogonal family.

Proof. Consider a family $x_1, \dots, x_k \in E$ such that $x_1 \wedge \cdots \wedge x_k \neq 0$, otherwise (136) is trivial. First note that by a Graham-Schmidt algorithm, there is an orthonormal basis $(e_1, \dots, e_d)$ such that for all $1 \leq i \leq k$, we have $x_i \in \bigoplus_{j=1}^i \mathbb{K}e_j$. For all $1 \leq i \leq k$, write a_i for the i-th coefficient of x_i in the basis $(e_1, \dots, e_d)$. Then we have $\|x_1 \wedge \cdots \wedge x_k\| = |a_1| \cdots |a_k|$ by Proposition A.38. Moreover, we have $|a_i| \leq \|x_i\|$ for all i because $(e_1, \dots, e_d)$ is an orthonormal family and $|a_i| = \|x_i\|$ if and only if $x_i \perp \bigoplus_{j=0}^i e_i$. Then note that if $x_1 \wedge \cdots \wedge x_k \neq 0$, then for all i, we have $\bigoplus_{j=0}^i e_i = \bigoplus_{j=0}^i x_i$ for all $i \leq k$ so $|a_i| = \|x_i\|$ for all i if and only if $(x_1, \dots, x_k)$ is orthogonal. □

Lemma A.41. *Let E be a standard vector space and $h \in \mathrm{End}(E)$. There is a unitary vector $v \in E$ such that $\|h(v)\| = \|h\|$ and for all such v,there is a unitary linear form $w \in E^*$ such that $|w(v)| = 1$ and $h(\ker(w)) \perp h(v)$.*

Proof. Note that this is trivial when $h = 0$. Otherwise, $\|h\|$ is in the image of the valuation $|\cdot|$, which is a multiplicative group so there is a scalar λ such that $|\lambda| = \|h\|^{-1}$ and λh is unitary. Take v a vector such that $\|h(v)\| = \|h\|\|v\|$, by the same argument as above, one may assume that v is unitary. Then take $w := (h(v))^\top \circ h$, (with $\top$ as in Lemma A.31 in the ultra-metric case and as in Remark A.2 in the Archimedean case). Then one has $|w(v)| = 1$. Moreover, w is unitary because h is and $h(\ker(w)) = \ker(h(v)^\top)$ is orthogonal to v by Lemma A.30. □

Definition A.42 (Singular values)**.** *Let $h \in \mathrm{End}(E)$. We call singular values of h the non-increasing (by Proposition A.40) sequence of non-negative real numbers $(\mu_k(h))_{k\geq 0}$ such that:*

$$\left\|\bigwedge^k h\right\| = \mu_1(h) \cdots \mu_k(h).$$

Proposition A.43 (Cartan decomposition)**.** *Let E be a standard vector space and $h \in \mathrm{End}(E)$. There an orthonormal family $(v_1, \dots, v_d)$ of E such that $(h(v_1), \dots, h(v_d))$ is orthogonal and $\|h(v_1)\| \geq \cdots \geq \|h(v_d)\|$.*

Proof. We define $(v_1, \dots, v_d)$ by induction using an auxiliary sequence $(w_1, \dots, w_d)$. Assume that for some $k \in \{1, \dots, d\}$ we have constructed an orthonormal family $(v_1, \dots, v_{k-1})$ in E and an orthonormal family $(w_1, \dots, w_{k-1})$ in E^* such that $|w^j(v_j)| = 1$ for all $j \in \{1, \dots, k-1\}$ and $w^j(v_i) = 0$ for all $j \neq i$. Write for all j, V_j the vector space spanned by $(v_1, \dots, v_j)$ and $w^j := \bigcap_{i\leq j} \ker(w^j)$. Then one has $V_j \perp w^j$ for all j, assume also that $h(V_j) \perp h(w^j)$ and that $\|h_{|w^j}\| \leq \|h(v_j)\|$. Then take $(v_k, w'_k) \in W_{k-1} \times W_{k-1}^*$ as in Lemma A.41, and set w_k to be w'_k on W_{k-1} and 0 on V_{k-1}. Then one has: $\|h_{|W_k}\| \leq \|h_{|W_{k-1}}\| = \|h(v_k)\|$; $h(W_k) \perp h(v_k)$ by Lemma A.41 and $h(W_k) \perp V_{k-1}$ because $h(W_k) \subset h(W_{k-1})$. □

Lemma A.44. *Let $h \in \mathrm{End}(E)$ and $(v_1, \dots, v_d)$ be as in Proposition A.43. Then one has for all $i \in \{1, \dots, d\}$ $\|h(v_i)\| = \mu_i(h)$.*

Proof. We prove this by induction on $k = 1, \dots, d$. Since $(v_1, \dots, v_d)$ is an orthonormal basis, we have a family of indices $1 \leq i_1 < \dots < i_k \leq d$ such that $h(v_{i_1} \wedge \dots \wedge v_{i_k}) = \mu_1(h) \dots \mu_k(h)$ and by proposition A.40, we have $h(v_{i_1} \wedge \dots \wedge v_{i_k}) \leq \|h(v_{i_1})\| \cdots \|h(v_{i_k})\|$ but the $h(v_i)$ are decreasing so $\mu_1(h) \dots \mu_k(h) \leq \|h(v_1)\| \cdots \|h(v_k)\|$. Then to have the other inequality, note that by orthogonality $\|h(v_1)\| \cdots \|h(v_k)\| = h(v_1 \wedge \dots \wedge v_k) \leq \mu_1(h) \dots \mu_k(h)$ and by induction hypothesis $\mu_1(h) \dots \mu_{k-1}(h) \leq \|h(v_1)\| \cdots \|h(v_{k-1})\|$ and if this product is 0 then we have $\mu_k(h) = \|h(v_k)\| = 0$ because the sequences are non-increasing and otherwise, we can simplify and we have $\mu_k(h) = \|h(v_k)\|$. ☐

A.4 Projective space and Grassmanian variety

Lemma A.45. *Let E be a Euclidean or Hermitian vector space. Let x, y, z be unitary vectors in E and w be a unitary linear form, we have:*

$$\|x \wedge z\| \leq \|x \wedge y\| + \|y \wedge z\| \tag{137}$$

$$||wy| - |wz|| \leq \|y \wedge z\|. \tag{138}$$

Proof. Take $(e_1, \dots, e_d)$ and orthonormal basis of E such that $e_1 = y$. Then we write $(x_1, \dots, x_d)$ and $(z_1, \dots, z_d)$ the coordinates of x and z respectively and write $x' = x - x_1 e_1$ and $y' := y - y_1 e_1$. Then we have $\|x \wedge y\| = \|x'\|$ and $\|y \wedge z\| = \|z'\|$ but $x \wedge z = x_1 e_1 \wedge z' - z_1 e_1 \wedge x' + x' \wedge z'$ and all three terms are orthogonal so:

$$\begin{aligned} \|x \wedge z\|^2 &= |x_1|^2 \|z'\|^2 + |z_1|^2 \|x'\|^2 + \|x' \wedge z'\|^2 \\ &\leq \|z'\|^2 \|x\|^2 + |z_1|^2 \|x'\|^2 \\ &\leq \|x'\|^2 + \|z'\|^2 \\ &\leq \|x'\|^2 + \|z'\|^2 + 2\|x'\|\|z'\| \end{aligned} \qquad \begin{aligned} \|x \wedge z\| &= \max\{x_1 \|z'\|, z_1 \|x'\|, \|x' \wedge z'\|\} \\ \|x \wedge z\| &\leq \max\{x_1 \|z'\|, z_1 \|x'\|\} \\ \|x \wedge z\| &\leq \max\{\|z'\|, \|x'\|\} \\ \|x \wedge z\| &\leq \max\{\|z \wedge y\|, \|x \wedge y\|\}. \end{aligned}$$

Then taking the square root, we have $\|x \wedge z\| \leq \|x \wedge y\| + \|y \wedge z\|$. Now to prove (138), consider a basis such that $w(e_1) = 1$ and $w(e_j) = 0$ for all $j > 1$, then note that

$$\begin{aligned} y \wedge z &= (y - z) \wedge z \\ &= \sum_{j=1}^{d} z_j (y - z) \wedge e_j \\ &= \sum_{j=1}^{d} \sum_{i=1}^{d} z_j (y_i - z_i) e_i \wedge e_j \end{aligned}$$

Moreover, this sum is orthogonal so we have:

$$\begin{aligned} \|y \wedge z\| &\geq \left\| \sum_{j=1}^{d} |z_j| (y_1 - z_1) e_1 \wedge e_j \right\| \\ \|y \wedge z\| &\geq |w(y) - w(z)| \|z\| \\ \|y \wedge z\| &\geq ||w(y)| - |w(z)||. \end{aligned}$$

☐

Lemma A.46. *If E is a standard ultra-metric vector space then* d *on* $\mathbf{P}(E)$ *defined by* $\mathrm{d}(\mathbb{K}x, \mathbb{K}y) := \frac{\|x \wedge y\|}{\|x\| \cdot \|y\|}$ *is ultra-metric. Moreover, the metric space* $(\mathbf{P}(E), \mathrm{d})$ *is compact.*

Proof. Take three unitary vectors x, y, z, take π_y an orthogonal projection onto $\mathbb{K}y$, then we have $z \wedge y = (z - \pi_y(z)) \wedge y$ and by Lemma A.40 $\|(z - \pi_y(z)) \wedge y\| = \|z - \pi_y(z)\|$, then by linearity, we have $x \wedge z = x \wedge (z - \pi_y(z)) + x \wedge \pi_y(z)$ and $\pi_y(z) = ay$ for some $|a| \leq 1$ so $\|x \wedge \pi_y(z)\| \leq \|x \wedge y\|$ and by Lemma A.40 we have $\|x \wedge (z - \pi_y(z))\| \leq \|z - \pi_y(z)\| = \|y \wedge z\|$, then by ultra-metric inequality on $\bigwedge^k E$, we have $\|x \wedge z\| \leq \max\{\|x \wedge y\|, \|y \wedge z\|\}$. Then we claim that $\mathbf{P}(E)$ is the quotient of the unit sphere of E (which is compact by *(LC)*) by the closed equivalence relation $x \sim y$ if there is $a \in \mathbb{K}$ unitary such that

$x = ay$. The equivalence relation is closed because for all unitary x, $\mathbb{K}x$ is closed in E and isometric to $\mathbb{K}$ and the unit group $\mathcal{U}(\mathbb{K})$ is closed in $\mathbb{K}$. Moreover, by the above reasoning, for all $x, y \in \mathbf{P}(E)$, we have $\|x \wedge y\| = \|x - \pi_y(x)\|$ and by ultra-metric inequality, we either have $\|\pi_y(x)\| = 1$, in which case or $\pi_y(x) \sim y$ or $\|x - \pi_y(x)\| = 1$ in which case $\|x - y\| \geq 1$ because the projection is contracting but x and y are unitary so $\|x - y\| = 1 = \|x \wedge y\|$. So $\|x \wedge y\|$ is indeed the distance between the image of x by the unit group and the image of y by the unit group. Then it is straightforward that the quotient of a compact by a closed equivalence relation is compact. Indeed a covering of $\mathbf{P}(E)$ by open sets lifts to a covering of the unit sphere which is invariant by the action of the unit group $\mathcal{U}(\mathbb{K})$, such a covering has a finite sub-covering by compactness and the latter projects back to a finite sub-covering of $\mathbf{P}(E)$. □

Note that this implies that the distance map on $\mathbf{P}(E)$ gives an ordered family of partitions by the following Lemma. Note that we already know by Proposition A.15 that the distance takes values in a set of type $\{\exp(-\beta n) | n \in \mathbb{N}\}$.

Lemma A.47 (Ball coverings of ultra-metric spaces)**.** *Let (X, d) be a compact ultra-metric space i.e, such that for all $x, y, z \in X$, we have $\mathrm{d}(x, z) \leq \max\{\mathrm{d}(x, y), \mathrm{d}(y, z)\}$. Then for all $\varepsilon > 0$, the relation $\sim_\varepsilon := (\mathrm{d} < \varepsilon)$ is an equivalence relation that has finitely many equivalence classes and as a consequence d takes its non-zero values in a discrete subset of $\mathbb{R}_{>0}$.*

Proof. Let $x, y, z \in X$ be such that $\mathrm{d}(x, y) < \varepsilon$ and $\mathrm{d}(y, z) < \varepsilon$, then by ultra-metric inequality we have $\mathrm{d}(x, z) < \varepsilon$, the reflexivity comes from the fact that the distance is symmetric. Then the covering of X by balls of radius ε has no non-trivial sub-covering since the balls are disjoint so there are only finitely many balls by compactness. Moreover, for all $\alpha \geq \varepsilon$ the relation $\sim_\alpha$ is coarser that $\mathrm{d} > \varepsilon$ so if we write $X = \bigsqcup_{i=1}^N C_i$ where the C_i's are the balls or radius ε then we have a non decreasing map $[\varepsilon, +\infty] \to \mathfrak{p}_N$ where $\mathfrak{p}_N$ is the set of equivalence relations on (a-k-a partitions of) $\{1, \dots, N\}$ endowed with the order relation $\sim \leq \sim' := \forall i, j,\ i \sim j \Rightarrow i \sim' j$. Since $\mathfrak{p}_N$ is a finite partially ordered set, $\sim_\alpha$ takes only finitely many values for $\alpha > \varepsilon$ and for all $\alpha < \alpha'$, we have $\sim_\alpha < \sim_{\alpha'}$ if and only if there is a pair $x, y \in X$ such that $\alpha \leq \mathrm{d}(x, y) < \alpha'$ therefore, there are only finitely many possible values taken by the distance map d in $[\varepsilon, +\infty]$ so they are all isolated except maybe ε, but this is true for all $\varepsilon > 0$ so all the non-zero values taken by d are isolated. □

Definition A.48 (Grassmannian variety)**.** *Let E be a standard vector space over a local field $\mathbb{K}$. Let $1 \leq k \leq \dim(E)$. We define the Grassmannian variety $\mathrm{Gr}_k(E)$ to be the set of subspaces of E that have dimension k. We endow V with the metric induced by the embedding:*

$$\begin{array}{ccc} \mathrm{Gr}_k(E) & \longrightarrow & \mathbf{P}\left(\bigwedge^i E\right) \\ \bigoplus_{i=1}^k \mathbb{K}x_i & \longmapsto & \mathbb{K} \bigwedge_{i=1}^k x_i. \end{array} \tag{139}$$

We write $\mathrm{Gr}_0(E) := \{\{0\}\}$ and $\mathrm{Gr}(E) := \bigsqcup_{k=0}^{\dim(E)} \mathrm{Gr}_k(E)$ for the general Grassmannian variety.

Definition A.49 (Flag variety)**.** *Let E be a standard vector space. Let $\Theta \subset \{1, \dots, \dim(E)\}$. We define the flag variety:*

$$\mathrm{Fl}_\Theta(E) := \left\{ (F_i)_{i \in \theta} \in \prod_{i \in \Theta} \mathrm{Gr}_i(E) \,\middle|\, \forall i \leq j, F_i \subset F_j \right\}. \tag{140}$$

We write $\mathrm{Fl}(E) := \bigsqcup_\Theta \mathrm{Fl}_\Theta(E)$ for the general flag variety and $\mathrm{Fl}(E) := \mathrm{Fl}_{\{1,\dots,d\}}(E)$ for the total flag variety.

Note that the data of a general flag is simply the data of a completely ordered subset of the general Grassmannian variety.

Definition A.50. *Let E be a standard vector space. Let $\gamma \in \mathrm{End}(E)$ and $F \in \mathrm{Fl}(E)$. We define the product $\gamma \cdot F$ as the collection $\{\gamma f \,|\, f \in F\}$.*

Proposition A.51. *Let E be a standard vector space. Let $\Theta \subset \{1, \dots, \dim(E)\}$. The flag variety $\mathrm{Fl}_\Theta(E)$ is compact.*

Proof. Let $d := \dim(E)$, let $k \leq d$ and let $\phi \in \bigwedge^k E$. Write:

$$\phi = \sum_{1 \leq i_1 < \cdots < i_k \leq d} \phi_{i_1,\dots,i_d} e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

Then ϕ is a pure k-vector if and only if for all $1 \leq i_1 < \cdots < i_k \leq d$, we have $\phi_{i_1,\dots,i_k} = x_{i_1,1} \cdots x_{i_k,k}$, for some $(x_{i,j})_{i \leq d, j \leq k} \in \mathbb{K}^{dk}$. This means that The set of pure k-vectors is closed. $\square$

A.5 Spectral radius and real spectrum

More details about this construction can be found in [Die54]

Lemma A.52. *Let E be a standard vector space and let $h \in \mathrm{End}(E)$. The sequence $\left(\|h^n\|^{\frac{1}{n}}\right)$ converges. Moreover, for all $k \in \{1, \dots, \dim(E)\}$, the sequence $\left(\mu_k(h^n)^{\frac{1}{n}}\right)$ also does.*

Proof. Let $m, n \in \mathbb{N}$. We have $\|h^{n+m}\| \leq \|h^n\| \cdot \|h^m\|$ so by sub-multiplicativity the sequence $\left(\|h^n\|^{\frac{1}{n}}\right)$ converges. Write ρ its limit, we claim that $\rho = 0$ if and only if h is nilpotent. Let $d := \dim(E)$, then h^d is a bijection on its image, write r the norm of its inverse, then we have $\|h^{nd}\| \geq r^{-n}$ so $\rho \geq r^{-d} > 0$. Then by the same argument, the sequence $\left(\| \bigwedge^k h^n\|^{\frac{1}{n}}\right)$ converges for all $k \in \{1, \dots, \dim(E)\}$. Moreover, for all $k \in \{2, \dots, \dim(E)\}$, if $\bigwedge^k h$ is nilpotent, then $\bigwedge^{k-1} h$ also is. As a consequence $\left(\mu_k(h^n)^{\frac{1}{n}}\right)$ converges because it is either the quotient of two convergent sequences that have positive limit or stationary to 0. $\square$

Definition A.53 (Real spectrum)**.** *Let E be a standard vector space and let $h \in \mathrm{End}(E)$ and let $k \in \{1, \dots, \dim(E)\}$. We define the k-th absolute eigenvalue of h as:*

$$\rho_k(h) := \lim_{n \to \infty} \mu_k(h^n)^{\frac{1}{n}}.$$

Theorem A.54 (Jordan-Dunford decomposition)**.** *Let $\mathbb{K}$ be a field, $d \geq 2$ an integer and E be a d-dimensional $\mathbb{K}$-vector space. Let $h \in \mathrm{End}(E)$. Let:*

$$\chi_h(X) := \det(h - X\mathrm{id}).$$

Let $\mathbb{K}'$ be the smallest extension of $\mathbb{K}$ on which $\chi_h(X)$ is a product of linear factors and let $E' := \mathbb{K}' \otimes E$. For all $\lambda \in \mathbb{K}'$, we define the eigenspace:

$$E_\lambda(h) := \ker\left((h - \lambda\mathrm{id})^d\right).$$

Then for all $\lambda \in \mathbb{K}'$, the vector space $E_\lambda(h)$ is stable by h and the dimension of $E_\lambda(h)$ is the multiplicity of λ as a root of $\chi_h(X)$. Moreover, we have:

$$E' = \bigoplus_{\lambda \in \mathbb{K}'} E_\lambda(h).$$

Moreover, for all factor $R(X) | \chi_h(X)$ in $\mathbb{K}[X]$, we have $E_R(h) \subset E$ a h-stable subspace such that:

$$\bigoplus_{\lambda \in \mathbb{K}', R(\lambda) = 0} E_\lambda(h) = \mathbb{K}' \otimes E_R(h).$$

Lemma A.55. *Let $\mathbb{K}$ be a local field and $P(X)$ be irreducible in $\mathbb{K}[X]$. let $\mathbb{K}'$ be an algebraic extension of $\mathbb{K}$. Then there is only one absolute value on $\mathbb{K}'$ that matches with the absolute value on $\mathbb{K}$ and all the roots of $P(X)$ in $\mathbb{K}'$ have the same absolute value.*

Definition A.56 (Ordered spectrum)**.** *Let $\mathbb{K}$ be a field, $d \geq 2$ an integer and E be a d-dimensional $\mathbb{K}$-vector space. Let $h \in \mathrm{End}(E)$. We call ordered spectrum of h a family $(\lambda_1, \dots, \lambda_d) \in \overline{\mathbb{K}}^d$ such that:*

$$\chi_h(X) = \prod_{i=1}^{d} (\lambda_i - X).$$

Lemma A.57 (Jordan-Dunford decomposition of the exterior product). *Let $\mathbb{K}$ be a field, $d \geq 2$ an integer and E be a d-dimensional $\mathbb{K}$-vector space. Let $h \in \mathrm{End}(E)$. Let $\mathbb{K}'$ be the smallest extension of $\mathbb{K}$ on which $\chi_h(X)$ is a product of linear factors. Let $2 \leq k \leq d$. Then the characteristic polynomial $\chi_{\bigwedge^k h}(X)$ is a product of linear factors on $\mathbb{K}'$ moreover, we have for all $\lambda \in \mathbb{K}'$:*

$$E_\lambda\left(\bigwedge^k h\right) = \bigoplus_{\lambda_1 \cdots \lambda_k = \lambda} \bigwedge_{i=1}^{k} E_{\lambda_i}(h).$$

As a consequence, if $(\lambda_i)_{1 \leq i \leq d}$ is an ordered spectrum of h, then $(\lambda_{i_1} \cdots \lambda_{i_k})_{1 \leq i_1 < \cdots < i_k \leq d}$ is an ordered spectrum of $\bigwedge^k h$.

Proposition A.58. *Let E be a standard vector space and let $h \in \mathrm{End}(E)$. For all $r \geq 0$, the number of indices $k \in \{1, \ldots, \dim(E)\}$ such that $\rho_k = r$ is equal to the dimension of:*

$$E_r(h) := \bigoplus_{|\lambda| = r} E_\lambda(h).$$

Proof. Let $(\lambda_1, \ldots, \lambda_d)$ be an ordered spectrum of h such that $|\lambda_1| \geq \cdots \geq |\lambda_d|$. Then for all vector x in, we write $x = \sum_{\lambda \in \mathbb{K}'} x_\lambda$. Then we have $\|h^n x_\lambda\|^{\frac{1}{n}} \to |\lambda|$ for all λ so $\|h^n\|^{\frac{1}{n}} \to |\lambda_1|$. Then by taking the exterior product for all $k \in \{1, \ldots, d\}$, we get that:

$$\rho_1(h) \cdots \rho_k(h) = |\lambda_1 \cdots \lambda_k|.$$

As a consequence $\rho_k(h) = |\lambda_k|$ for all k. □

Definition A.59. *Let E be a standard vector space and let $h \in \mathrm{End}(E)$. We define:*

$$E^+(h) := E_{\rho_1(h)}(h).$$

B About probabilities

In this appendix, we prove basic results in the study of stochastic processes. More details about the following results can be found in [Sen06, p. 11-24] page 11.

B.1 Definitions and preliminaries

Definition B.1 (σ-Algebra). *Let Ω be a set. We call algebra in Ω a collection $\mathcal{A}$ of parts of Ω that is stable by union, by intersection and by complementary i.e, a sub-algebra of $(\mathcal{P}(E), \cup, \cap)$. We say that $\mathcal{A}$ is a σ-algebra if it moreover satisfies the following equivalent conditions:*

1. *$\mathcal{A}$ is stable by countable union,*
2. *$\mathcal{A}$ is stable by disjoint countable union,*
3. *$\mathcal{A}$ is stable by countable intersection,*
4. *$\mathcal{A}$ is stable by taking the limit of monotonous (resp. increasing; resp. decreasing) sequences.*

We call measurable space the data $(\Omega, \mathcal{A})$ where Ω is a set and $\mathcal{A}$ a is σ-algebra on Ω. We sometimes omit the σ-algebra and simply call elements of $\mathcal{A}$ measurable subsets of Ω.

Definition B.2 (Probability space). *We call probability space a triple $(\Omega, \mathcal{A}, \mathbb{P})$ where Ω is a set, $\mathcal{A}$ is a σ-algebra and $\mathbb{P} : \mathcal{A} \to [0, 1]$ a σ-additive map i.e, a map $\mathbb{P}$ such that $\mathbb{P}(\Omega) = 1$ and that satisfies one of the following equivalent conditions:*

1. *for all sequence $(A_n) \in \mathcal{A}^{\mathbb{N}}$, that are pairwise disjoint, we have $\mathbb{P}(\bigcup A_n) = \sum \mathbb{P}(A_n)$*

2. *for all monotonous sequence* $(A_n) \in \mathcal{A}^{\mathbb{N}}$*, we have* $\lim \mathbb{P}(A_n) = \mathbb{P}(\lim A_n)$ *and for all* A, B *disjoint,* $\mathbb{P}(A \sqcup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Just like for measurable spaces, we will simply write $(\Omega, \mathbb{P})$ *to specify a probability space.*

Definition B.3 (Basis)**.** *We say that a* σ*-algebra* $\mathcal{A}$ *has basis* $\mathcal{B}$ *(or is generated by* $\mathcal{B}$*) if* $\mathcal{B} \subset \mathcal{A}$ *and no proper sub-algebra of* $\mathcal{A}$ *contains* $\mathcal{B}$ *or equivalently, if* $\mathcal{A}$ *is the intersection of all* σ*-algebras that contain* $\mathcal{B}$*.*

Proposition B.4 (Carathéodory's theorem)**.** *Let* A *be a* σ*-algebra and* $\mathcal{B}$ *a countable basis of* A *that is stable by finite union and by complementary i.e, a* $\cap, \cup$*-sub-algebra. Assume that there is an additive map* $\mathbb{P} : \mathcal{B} \to [0, 1]$ *such that* $\mathbb{P}(\Omega) = 1$*. Then* $\mathbb{P}$ *extends uniquely to a probability distribution on* $\mathcal{A}$*.*

Definition B.5 (Random variable)**.** *Let* $(\Omega, \mathcal{A}_\Omega)$ *and* $(\Gamma, \mathcal{A}_\Gamma)$ *be two Measurable spaces. We say that a function* $\gamma : \Omega \to \Gamma$ *is measurable if* $\gamma^{-1}(\mathcal{A}_\Gamma) \subset \mathcal{A}_\Omega$*. We call random variable the joint data of a measurable map* $\gamma : \Omega \to \Gamma$ *and a probability measure* $\mathbb{P}$ *on* Ω*. We then say that* γ *is defined on* $(\Omega, \mathbb{P})$ *and valued in* $(\Gamma, \mathcal{A}_\Gamma)$*.*

Definition B.6 (Filtered probability spaces)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space. We call filtration on* Ω *a non-decreasing sequence* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *of sub-algebras of* $\mathcal{A}$*. We call the data of the triple* $(\Omega, (\mathcal{F}_n), \mathbb{P})$ *a filtered probability space. Note that the data of the* σ*-algebra* $\mathcal{A}$ *is encoded in* $\mathbb{P}$*.*

Definition B.7. *Let* $(\Omega, (\mathcal{F}_n), \mathbb{P})$ *be a filtered probability space and* X *a measurable space. We say that a random sequence* $(x_n) : \Omega \to X^{\mathbb{N}}$ *respects the filtration* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *if for all* $n \in \mathbb{N}$*, the data of* $(x_0, \dots, x_{n-1})$ *is* $\mathcal{F}_n$*-measurable.*

Definition B.8 (Stopping time)**.** *Let* $(\Omega, (\mathcal{F}_n), \mathbb{P})$ *be a filtered probability space. Let* $t \in \mathbb{N}$ *be a random integer, we say that* t *is a stopping time for* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *if for all* $n \in \mathbb{N}$*, the event* $(t \geq n)$ *is* $\mathcal{F}_n$*-measurable.*

Definition B.9 (Extracted filtration)**.** *Let* $(t_m)_{m\in\mathbb{N}}$ *be a sequence of stopping times such that* $t_m \leq t_{m+1}$ *for all* $m \in \mathbb{N}$*. We write* $(\mathcal{F}_{t_m})_{m\in\mathbb{N}}$ *for the extracted filtration generated by:*

$$\mathcal{F}_{t_m} := \langle F \cap (t_m \geq n)\,;\, n \in \mathbb{N}, F \in \mathcal{F}_n \rangle\,.$$

It is characterized by the property that for all sequence $(x_n)_{n\in\mathbb{N}}$ *that respects* $\mathcal{F}_n$ *and for all* m*, the data of* $(x_0, \dots, x_{t_m-1})$ *is* $\mathcal{F}_{t_m}$*-measurable.*

Definition B.10 (Conditional probability)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space. For two measurable events* $A, B \subset \Omega$*, we define the conditional probability* $\mathbb{P}(A|B) := \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$*. And for* $\mathcal{B}$ *a sub-algebra of* $\mathcal{A}$*, we define* $\mathbb{P}(A|\mathcal{B})$ *as the only (up to equivalence)* $\mathcal{B}$*-measurable function valued in* $[0, 1]$ *such that* $\mathbb{P}(A|B) = \int_B \mathbb{P}(A|\mathcal{B})\mathrm{d}\mathbb{P}$ *for every* $B \in \mathcal{B}$*.*

Definition B.11 (Distribution)**.** *Let* $(\Omega, \mathbb{P})$ *be a probability space and* X *a measurable space. Let* $x\Omega \to X$ *be a random variable. We call distribution of* x *the probability distribution on* X *defined by* $A \mapsto \mathbb{P}(x \in A)$ *for all measurable* $A \subset X$*.*

Let X, Y *be two measurable spaces and* $f : X \to Y$ *be a measurable map. Let* x *be a random variable valued in* X*. Let* ν *be a probability measure on* X*. We write* $x \sim \nu$ *to say that* x *has distribution* ν *and in this case, we write* $f(\nu)$ *for the distribution of* $f(x)$*.*

Definition B.12 (Expectation)**.** *Let* $(\Omega, \mathbb{P})$ *be a probability space. Let* $x : \Omega \to \mathbb{R}_{\geq 0}$ *be a real random variable of distribution* ν*. We define:*

$$\mathbb{E}(\nu) := \int_{t=0}^{+\infty} \nu(t, +\infty)dt = \mathbb{E}(x) := \int_{t=0}^{+\infty} \mathbb{P}(x > t)dt.$$

Definition B.13 (Conditional expectation)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space and* $\mathcal{B}$ *a sub-algebra of* $\mathcal{A}$*. Let* x *be a non negative real* $\mathcal{A}$*-measurable random variable. We define the non negative (but possibly infinite)* $\mathcal{B}$*-measurable random variable:*

$$\mathbb{E}(X \,|\, \mathcal{B}) := \int_{t=0}^{\infty} \mathbb{P}(x \geq t \,|\, \mathcal{B})dt.$$

The random variable $\mathbb{E}(x \,|\, \mathcal{B})$ *is (up to equivalence) the only* $\mathcal{B}$*-measurable random variable that satisfies* $\mathbb{E}(y\mathbb{E}(x|\mathcal{B})) = \mathbb{E}(XY)$ *for all* $\mathcal{B}$*-measurable random variable* y*.*

B.2 Markov chains

Definition B.14 (Markov chain)**.** *We call finite Markov space the data of a finite set X and a transition kernel p i.e, map $p : X \times X \to [0,1]$ such that $\sum_{y \in X} p(x,y) = 1$ for all $x \in X$. We say that a random sequence $(x_n) \in X^{\mathbb{N}}$ defined over a probability space $(\Omega, \mathbb{P})$ is a Markov chain in (X,p) if for every index $n \geq 1$, we have $\mathbb{P}(x_n = y | x_0, \cdots, x_{n-1}) = p(x_{n-1}, y)$. If X is a singleton, we say that the Markov space (X,p) is trivial and if there is a point $y_0 \in X$ such that $p(x, y_0) = 1$ for all $x \in X$, we say that (X,p) is semi-trivial.*

Remark B.15. *The distribution law of a Markov chain (x_n) is encoded in the data of a Markov space (X,p) and a starting law ξ_0, that is the distribution law of x_0.*

Definition B.16 (Admissible paths)**.** *One can see a finite or countable Markov space (X,p) as an oriented graph with weighted vertices. We call edge of X a pair $(x,y) \in X \times X$ such that $p(x,y) > 0$ and write $\mathcal{E}(X,p) := \{(x,y) \in X | p(x,y) > 0\}$. We say that a sequence $(x_i)_{i \in I}$ for an interval $I \subset \mathbb{Z}$ is p-admissible if for every consecutive pair $i, i+1 \in I$, we have $(x_i, x_{i+1})\mathcal{E}(X,p)$.*

Definition B.17 (Image of a distribution and iterated Markov chain)**.** *Let (X,p) be a Markov space. Given a measure ξ on X (i.e, a function $\xi : X \to \mathbb{R}$) and a subset $A \subset X$, we write $p(\xi, A) := \sum_{x \in X, y \in A} \xi(x) p(x,y)$, that way, one has $p(x,y) = p(\delta_x, \{y\})$. We say that a measure ξ is p-stationary if $p(\xi) = \xi$. For every integer n, we write p^n the iterated Markov kernel associated to p i.e, such that $p^n(x,y) = \mathbb{P}(x_n = y)$ for (x_n) a Markov chain in (X,p) with $x_0 = x$. This is equivalent to saying that $p^{n+m}(x,z) = \sum_{y \in X} p^n(x,y) + p^m(y,z)$ for all $x, y \in X$ and $p^1 = p$.*

Now we define the notions of irreducible, aperiodic irreducible, eventually irreducible and eventually aperiodic irreducible Markov chains. Note that even though the definitions and demonstrations rely on probability theory, these notions are purely properties of the oriented graph $(X, \mathcal{E})$. The only purely probabilistic quantity is the spectral gap.

Definition B.18 (Aperiodic and irreducible Markov spaces)**.** *We say that a Markov space is irreducible if for every $x, y \in X$, there exists an integer $n > 0$ such that $p^n(x,y) > 0$ i.e, if the oriented graph $(X, \mathcal{E}(X,p))$ is strongly connected. We say that (X,p) is aperiodic irreducible if for every $x, y \in X$ the set $\{n \in \mathbb{N} | p^n(x,y) = 0\}$ is finite. We call irreducible components of a Markov space a p-invariant subset that is irreducible.*

Remark B.19. *Note that if we assume a given Markov space (X,p) to be aperiodic irreducible, then for any given integer $n \geq 1$, the Markov space (X, p^n) is also aperiodic irreducible.*

Lemma B.20. *Let (X,p) be an irreducible Markov space. There is a unique probability distribution ξ_p on X that is p-stationary and has full support. If moreover (X,p) is aperiodic then for every probability distribution ξ on X, $p^n * \xi \to \xi_p$ exponentially fast.*

Proof. For every point $x \in X$, write $g(x,y) := \lim_{m \to \infty} \sum_{n=1}^{m} \frac{p^n(x,y)}{m}$, the limit exists and is in $[0,1]$ because $0 \leq p^n(x,y) \leq 1$ for all n, x, y. Then write $g_m(x,y) := \sum_{n=1}^{m} \frac{p^n(x,y)}{m}$, then we decompose each path of length $n \geq 1$ into the concatenation of a prefix of length $n-1$ and a last step and we have:

$$\begin{aligned} \sum_{n=0}^{m} \frac{p^n(x,y)}{m} &= \frac{1}{m} + \sum_{z \in X} \sum_{n=1}^{m} \frac{p^{n-1}(x,z)}{m} p(z,y) \\ g_m(x,y) &= \frac{1}{m} + \frac{m-1}{m} \sum_{z \in X} g_{m-1}(x,z) p(z,y) \\ g(x,y) &= \sum_{z \in X} g(x,z) p(z,y) \end{aligned}$$

So $y \mapsto g(x,y)$ is a p-stationary distribution on X. Now assume that we have two p-stationary distributions ξ_1 and ξ_2 on X. Then one has for all $y \in X$, and all $n \in \mathbb{N}$, $\sum_{x \in X} (\xi_1(x) - \xi_2(x)) p(x,y) = \xi_1(y) - \xi_2(y)$. Now take y such that $\xi_1(y) - \xi_2(y)$ is maximal. Then for every $x \in X$ we have either $(\xi_1 - \xi_2)(x) = (\xi_1 - \xi_2)(y)$ or $p^n(x,y) = 0$ for every $n \in \mathbb{N}$. If we assume that (X,p) is irreducible, then $(\xi_1 - \xi_2)$ is constant on X. So there is a unique p-stationary distribution ξ_p. Moreover if there is a point x such that $\xi_p(x) = 0$, then for every $n \in \mathbb{N}$, one would have $\sum_y p^n(y,x)\xi_p(y) = 0$, and then by irreducibility $\xi_p = 0$ which is absurd. Now assume that (X,p) is aperiodic. Take $m, \varepsilon > 0$ such that $p^m(x,y) \geq \varepsilon$ for all x, y. Then take a function $f \in \mathbb{R}^X$ such that $\sum f(x) = 0$ and take y such that $f(y)$ is maximal. Then one has for all z, $\sum f(x)p^m(x,z) \leq f(y)(1-\varepsilon)$ and so if we write $f_n = p^{nm+k} * \xi - \xi_p$ then $f_n \to 0$ so $p^n(\xi) \to \xi_p$ exponentially fast. □

Lemma B.21. *Let (X,p) be an eventually irreducible Markov space. Then there exists a unique p-stationary distribution on X. If it is moreover aperiodic, then for every probability distribution ξ on X, the image $p^n(\xi)$ converges to the p-stationary distribution.*

Proof. We simply need to show that for every $x \in X^{\mathrm{w}}$ and every probability distribution ξ, we have $p^n(\xi, x) \to 0$. Therefore take an integer m such that $p^m(x, X^{\mathrm{s}}) > 0$ for every $x \in X^{\mathrm{w}}$, such an m exists because (X,p) is eventually irreducible. Then we have $p^m(x, X^{\mathrm{w}}) \leq \rho 1$ for every $x \in X^{\mathrm{w}}$ and since there is no admissible path from X^{s} to X^{w}, we have:

$$p^{m+k}(x, X^{\mathrm{w}}) = \sum_{y \in X^{\mathrm{w}}} p^k(x,y)p^m(y, X^{\mathrm{w}})$$

So for every $n \in \mathbb{N}$ and every probability distribution ξ, we have $p^n(\xi, X^{\mathrm{w}}) \leq \rho^{\left\lfloor \frac{n}{m} \right\rfloor} \to 0$. □

Lemma B.22. *Let (X,p) be an eventually irreducible Markov space. Let ξ be the p-invariant probability distribution. There are constants $C, \beta > 0$ such that for all $b \in X$ such that $\xi(b) > 0$ and for all Markov chain (x_n) in (X,p), if we write n_0 for the first time such that $x_{n_0} = b$ then:*

$$\mathbb{P}(n_0 \geq k) \leq C \exp(-\beta_k). \tag{141}$$

Proof. Let l_0 be the Diameter of X *i.e,* the smallest integer such that for all $x, y \in X$ such that $\xi(y) > 0$, we have an admissible path of length at most l_0 from x to y. Then Write $\varepsilon := \min_{p(x,y>0)} p(x,y)$. Then for all $n \in \mathbb{N}$, we have:

$$\mathbb{P}(n_0 \geq n + l_0 \,|\, n_0 \geq l) \leq 1 - \varepsilon^{l_0}.$$

So if we write $\beta := -\frac{\log(1-\varepsilon^{l_0})}{l_0}$ and $C := \exp(\beta l_0)$, then we have (141). □

B.3 About exponential large deviations inequalities

Lemma B.23 (Distribution of the current step)**.** *Let $(w_k)_{k \in \mathbb{N}}$ be a random sequence of positive integers that respects a filtration $(\mathcal{F}_k)$. For all $n \geq 0$, we write r_n for the largest integer such that $w_0 + \cdots + w_{r_n - 1} < n$. Assume that for some $\eta : \mathbb{N} \to \mathbb{R}_{\geq 0}$, we have:*

$$\forall t \geq 0, \forall k \geq 0, \mathbb{P}(w_k > t \,|\, \mathcal{F}_k) \leq \eta(t).$$

Then we have:

$$\forall t \in \mathbb{N},\ \forall n \in \mathbb{N},\ \mathbb{P}(w_{r_n} = t) \leq t\eta(t) \tag{142}$$

In particular if there are constants $C, \beta > 0$ such that $\eta(t) \leq C \exp(-\beta t)$ for all integer t, then we have:

$$\forall t \geq 0, \forall n \geq 0, \mathbb{P}(w_{r_n} > t \,|\, \mathcal{F}_k) \leq \frac{C(1+t\beta)}{\beta^2} \exp(-\beta t). \tag{143}$$

Proof. First have a look at the green function $\mathcal{G} : y \mapsto \mathbb{E}(\#\{r \in \mathbb{N}; \overline{w}_r = y\})$. One has $\mathcal{G}(y) \leq 1$ for all $y \in \mathbb{N}$ because the (w_k)'s are positive so there is at most one index r such that $\overline{w}_r = y$. It means that we have

$\sum_{r=0}^{\infty} \mathbb{P}(\overline{w}_r = y) \leq 1$ for all $y \in \mathbb{N}$. Now consider $0 \leq t, n$ two integers.

$$\begin{aligned}
\mathbb{P}(w_{r_n} = t) &= \sum_{r=0}^{+\infty} \mathbb{P}((r = r_n) \wedge (w_r = t)) \\
&= \sum_{r=0}^{+\infty} \mathbb{P}((n - t \leq \overline{w}_r < n) \wedge (w_r = t)) \\
&= \sum_{r=0}^{+\infty} \sum_{u=t}^{+\infty} \mathbb{P}(\overline{w}_r = n - u) \mathbb{P}(w_r = t \,|\, \overline{w}_r = n - u) \\
&\leq \sum_{u=1}^{t} \sum_{r=0}^{\infty} \mathbb{P}(\overline{w}_r = n - u) \eta(t) \\
&\leq t\eta(t).
\end{aligned}$$

This proves (142). Now assume that $\eta(t, +\infty) \leq C \exp(-\beta t)$ for some $C, \beta > 0$ and for all $t \in \mathbb{N}$. Then for all $t, n \in \mathbb{N}$, we have:

$$\begin{aligned}
\mathbb{P}\left(w_{r_n} > t\right) &= \sum_{u>t} \mathbb{P}\left(w_{r_n} = u\right) \\
&\leq \sum_{u>t} u\eta(u) \\
&\leq \sum_{u>t} uC \exp(-\beta u) \\
&\leq \frac{C(1 + t\beta)}{\beta^2} \exp(-\beta t)
\end{aligned}$$

□

We now give some standard large deviations inequalities for sums of random variables guided by a finite Markov chain.

Lemma B.24 (Sum of random variables that have finite exponential moment)**.** *Let $(\Omega, \mathbb{P})$ be a probability space endowed with a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Let w be a random integer that is exponentially integrable (i.e, such that $\mathbb{E}(\lambda^w) < +\infty$ for some constant $\lambda > 1$) and let (X_n) be a random sequence of $(\mathcal{F}_{n-1})$-measurable real non-negative random variables that are uniformly relatively exponentially integrable i.e, there exists constants $C, \beta > 0$ such that:*

$$\forall n \in \mathbb{N}, \mathbb{E}(\exp(\beta X_n) | \mathcal{F}_n) \leq C. \tag{144}$$

Then $\sum_{k=0}^{w} X_n$ is exponentially integrable. Moreover there are some constants $C', \beta' > 0$ that can be expressed as functions of $C, \beta, \mathbb{E}(\lambda^w), \lambda$ and such that $\mathbb{E}\left(\exp\left(\beta' \sum_{k=0}^{w} X_n\right) \mid \mathcal{F}_{n-1}\right) \leq C'$.

Proof. Take some $0 < \varepsilon$. Write $Y := \sum_{k=0}^{w} X_n$ and for every j, write $Z_j := \sum_{k=0}^{j-1} X_n$. Then for all constant t, one has:

$$\mathbb{P}(Y \geq t) \leq \mathbb{P}(w \geq t\varepsilon) + \mathbb{P}\left(Z_{\lfloor t\varepsilon \rfloor} \geq t\right). \tag{145}$$

Then write $\zeta_j := \exp(\beta Z_j)$, by equation (144), and by induction on j, one has $\mathbb{E}(\zeta_j) \leq C^j$. Indeed, $\zeta_0 = 1$ and for all $j \geq 0$, ζ_j is $\mathcal{F}_j$ measurable, so by Definition B.13, one has:

$$\begin{aligned}
\mathbb{E}(\zeta_{j+1}) &= \mathbb{E}(\mathbb{E}(\zeta_{j+1} | \mathcal{F}_j)) \\
&= \mathbb{E}(\zeta_j \mathbb{E}(\exp(\beta X_j) | \mathcal{F}_j)) \\
&\leq C\mathbb{E}(\zeta_{j-1}).
\end{aligned}$$

So for all $j \in \mathbb{N}, t \in \mathbb{R}_{\geq 0}$, one has by Markov's inequality: $\mathbb{P}(\zeta_j \geq \exp(\beta t)) \leq \frac{C^j}{\exp(\beta t)}$, and taking $j = \lfloor t\varepsilon \rfloor$, we get:

$$\mathbb{P}\left(Z_{\lfloor t\varepsilon \rfloor} \geq t\right) \leq \exp(-(\beta - \varepsilon \log(C))t).$$

Note that $C \geq 1$ because the X_n's are non-negative. Moreover for ε small enough, one has $\beta - \varepsilon \log(C) > 0$. Since w is exponentially integrable and by Markov's inequality, there are constants $C', \beta' > 0$ such that $\mathbb{P}(w \geq t\varepsilon) \leq C' \exp(-\beta' t)$ (namely $C' := \mathbb{E}(\lambda^w)$ and $\beta' := \varepsilon \log(\lambda)$). So if we write $\beta'' := \min\{\beta', \beta - \varepsilon \log(C)\} > 0$ and $C'' := C' + 1$ then we have $\mathbb{P}(Y \geq t) \leq C'' \exp(-\beta'' t)$. □

Lemma B.25 (Exponential moments approximate the expectation)**.** *Let $M, \sigma, C, \beta > 0$. For all $\alpha < \sigma$, there is a constant $\beta_\alpha > 0$ that depends on $(M, \sigma, C, \beta, \alpha)$ such that for all random variable x that satisfies $\mathbb{E}(\min\{x, M\}) \geq \sigma$ and $\mathbb{P}(x \leq t) \leq C \exp(\beta t)$ for all $t \in \mathbb{R}$, we have:*

$$\mathbb{E}(\exp(-\beta_\alpha x)) \leq \exp(-\beta_\alpha \alpha).$$

Proof. Let x be a random variable such that $\mathbb{P}(x \leq t) \leq C \exp(\beta t)$ for all $t \in \mathbb{R}$. For all $\beta' < \beta$ and for all $m \in \mathbb{R}$ we have:

$$\begin{aligned}
\mathbb{E}(\exp(-\beta' x)\mathbb{1}(x \leq m)) &= \int_{m=\exp(-\beta' m)}^{+\infty} \mathbb{P}(\exp(-\beta' x) \geq t)\mathrm{d}t \\
&\leq \int_{t=\exp(\beta' m)}^{+\infty} t^{-\frac{\beta}{\beta'}} \mathrm{d}t \\
&\leq \frac{\beta}{\beta - \beta'} \exp((\beta - \beta')m) =: F(m, \beta')
\end{aligned}$$

Now assume moreover that $\mathbb{E}(\min\{x, M\}) \geq \sigma > 0$. Write x' for the random variable such that $x' = x$ when $m \leq x \leq M$, $x' = m$ when $x \leq m$ and $x' = M$ when $x \geq M$. Note that $\mathbb{E}(x') \geq \sigma$, note also that exp is a convex function. Therefore we have:

$$\begin{aligned}
\exp(-\beta' x') &\leq \frac{x' - m}{M - m} \exp(-\beta' M) + \frac{M - x'}{M - m} \exp(-\beta' m) \\
\mathbb{E}(\exp(-\beta x')) &\leq \frac{M \exp(-\beta' m) - m \exp(-\beta' M)}{M - m} \\
&\quad - \mathbb{E}(x') \frac{(\exp(-\beta' m) - \exp(-\beta' M))}{M - m} \\
\mathbb{E}(\exp(-\beta x')) &\leq \frac{M \exp(-\beta' m) - m \exp(-\beta' M)}{M - m} \\
&\quad - \frac{(\exp(-\beta' m) - \exp(-\beta' M))}{M - m} \sigma =: L(m, \beta')
\end{aligned}$$

Now if we take $m' := \frac{2\log(\beta') - 2K}{\beta}$ for some K, then for $\beta' < \frac{\beta}{2}$, we get that $L(m', \beta') \leq 2\exp(-K)\beta'$. Moreover, $\beta' m \to 0$ as $\beta' \to 0$ and exp is $\mathcal{C}^1$ at the neighbourhood of 0 so $\frac{1 - L(m', \beta')}{\beta} \to \sigma$ as $\beta' \to 0$. Therefore for all $\alpha < \alpha' < \sigma$, we can take K such that $2\exp(-K) \leq \alpha' - \alpha$. Then β small enough so that $L(m', \beta') \leq 1 - \beta\alpha'$ and then for all random variable x that satisfies the hypothesis of Lemma B.25, we get:

$$\begin{aligned}
\mathbb{E}(\exp(-\beta' x)) &\leq L(m', \beta') + F(m', \beta') \\
&\leq 1 - \beta\alpha' + (\alpha' - \alpha)\beta' \\
&\leq \exp(-\beta' \alpha).
\end{aligned}$$

□

Lemma B.26 (Classical large deviations from below)**.** *let (x_n) be a sequence or real random variables that respect a filtration $\mathcal{F}_n$. Assume that there exist constants C, β such that for all $n \in \mathbb{N}$, we have $\mathbb{E}(\exp(-\beta x_n)|\mathcal{F}_n) \leq C$ and a non-decreasing family $(\sigma_t)_{t\in\mathbb{R}}$ such that $\mathbb{E}(\min\{x_n, t\}|\mathcal{F}_n) \geq \sigma_t$ for all t, n with. Write $\sigma := \lim_{t\to+\infty} \sigma_t$. Then we have:*

$$\forall \alpha < \sigma,\ \exists \beta_\alpha > 0,\ \forall n \in \mathbb{N},\ \mathbb{P}(\overline{x}_n \leq \alpha n) \leq \exp(-\beta_\alpha n).$$

Proof. Consider a constant $\alpha_0 < \alpha < \alpha' < \alpha'' < \sigma$. Let $t \in \mathbb{R}$ be such that $\mathbb{E}(\min\{x_n, t\}|\mathcal{F}_n) \geq \alpha''$ for all $n \in \mathbb{N}$. Then by Lemma B.25, there is a constant $\beta' > 0$ such that for all $n \in \mathbb{N}$, we have:

$$\mathbb{E}(\exp(-\beta' x_n)|\mathcal{F}_n) \leq \exp(-\beta'\alpha').$$

Then by characterization of the conditional expectation, we get $\mathbb{E}(\exp(-\beta'\overline{x}_n)) \leq \exp(-n\beta'\alpha')$ for all $n \in \mathbb{N}$. Then by Markov's inequality, we get $\mathbb{P}(\overline{x}_n \leq \alpha n) \leq \exp(-n\beta'(\alpha' - \alpha))$. □

Definition B.27 (Large deviations inequalities)**.** *Let (x_n) be a random sequence of real numbers. We say that (x_n) satisfies some large deviations inequalities below a speed $\sigma \in \mathbb{R} \cup \{+\infty\}$ if we have:*

$$\forall \alpha < \sigma,\ \exists C, \beta > 0,\ \forall n \in \mathbb{N},\ \mathbb{P}(x_n \leq \alpha n) \leq \exp(-\beta n).$$

We then say that $(-x_n)$ satisfies large deviations inequalities above $-\sigma$.

Remark B.28. *Let (x_n) be a random sequence of real numbers that satisfies large deviations inequalities below a speed $\sigma \in \mathbb{R} \cup \{+\infty\}$. Then we have almost surely $\liminf \frac{x_n}{n} \geq \sigma$. If moreover (x_n) satisfies large deviations inequalities above $\sigma \in \mathbb{R}$, then $\lim \frac{x_n}{n} = \sigma$ almost surely.*

Proof. First assume that (x_n) is a random sequence of real numbers that satisfies large deviations inequalities below a speed $\sigma \in \mathbb{R} \cup \{+\infty\}$. We want to show that for all α the set $A_\alpha := \left\{n \in \mathbb{N} \,\middle|\, \frac{x_n}{n} \leq \alpha\right\}$ is almost surely finite. Consider some constants $C, \beta > 0$ such that $\mathbb{P}(x_n \leq \alpha n) \leq \exp(-\beta n)$ for all $n \in \mathbb{N}$, then we get $\mathbb{E}(\#A_\alpha) \leq \frac{C}{\beta}$. In particular A_α is almost surely finite. Then consider (α_k) an increasing sequence such that $\alpha_k \to \sigma$. For all $k \in \mathbb{N}$, we have $\liminf \frac{x_n}{n} \geq \alpha_k$ with probability one so the countable intersection $\liminf \frac{x_n}{n} \geq \sigma$ also has probability one.

If we now assume that (x_n) satisfies large deviations inequalities above $\sigma \in \mathbb{R}$. Then by the above reasoning, we have almost surely $\liminf \frac{x_n}{n} \geq \sigma$ and $\liminf \frac{-x_n}{n} \geq -\sigma$ so $\sigma = \liminf \frac{x_n}{n} = \limsup \frac{x_n}{n} = \lim \frac{x_n}{n}$. □

Now we see that sequences that satisfy large deviations inequalities behave well under some compositions.

Lemma B.29. *Let $(x_n)_{n\in\mathbb{N}}$ and $(x'_n)_{n\in\mathbb{N}}$ be two random sequences of real numbers that satisfy large deviations inequalities below speeds σ and σ' respectively. Let (y_n) be a sequence of random variables whose negative parts have bounded exponential moment. Let $(k_n)_{n\in\mathbb{N}}$ be a random non-decreasing sequence of non-negative integers. Then:*

1. *The shifted sequence $(x_n - y_n)_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed σ.*
2. *The minimum $(\min\{x_n, x'_n\})_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed $\min\{\sigma, \sigma'\}$.*
3. *The maximum $(\max\{x_n, x'_n\})_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed $\max\{\sigma, \sigma'\}$.*
4. *For all $\lambda > 0, \lambda' \geq 0$, the sum $(\lambda x_n + \lambda' x'_n)_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed $\lambda\sigma + \lambda'\sigma'$.*
5. *If $(k_n)_{n\in\mathbb{N}}$ satisfies large deviations inequalities below a speed $\kappa \in (0, +\infty)$ then the composition $(x_{k_n})_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed $\kappa\sigma$.*
6. *Let $(r_m)_{m\in\mathbb{N}}$ be the reciprocal function of κ defined by $r_m := \max\{n \in \mathbb{N}; k_n \leq m\}$ for all $m \in \mathbb{N}$. Then for $\kappa > 0$, the sequence $(r_m)_{m\in\mathbb{N}}$ satisfies large deviations inequalities above the speed κ^{-1} if and only if $(k_n)_{n\in\mathbb{N}}$ satisfies large deviations inequalities below the speed κ.*

Proof. We first prove 1 Let $\alpha < \alpha' < \sigma$. By assumption, there are two constants C_x, β_x such that $\mathbb{P}(x_n \leq \alpha'_n) \leq C\exp(-\beta n)$. Then saying that negative part of y has finite exponential moment means that there is a constant $\beta > 0$ such that $\mathbb{E}(\exp(-\beta y_n)) =: C_y < +\infty$. By Markov's inequality, we get that $\mathbb{P}(y_n \leq (\alpha - \alpha')t) \leq C_y \exp(-\beta(\alpha' - \alpha)t)$ for all $t \in \mathbb{R}$. Write $\beta_y := \beta(\alpha' - \alpha)$. Then we have:

$$\mathbb{P}(y_n + x_n \leq \alpha n) \leq \mathbb{P}(x_n \leq \alpha' n) + \mathbb{P}(y \leq (\alpha - \alpha')n)$$
$$\mathbb{P}(y_n + x_n \leq \alpha n) \leq C_x \exp(-n\beta_x) + C_y \exp(-\beta_y n)$$
$$\mathbb{P}(y_n + x_n \leq \alpha n) \leq (C_y + C_x)\exp(-n\min\{\beta_x, \beta_y\})$$

Now to prove 2 assume that $\sigma \leq \sigma'$. Then we get for all $\alpha < \sigma$, the inequality $\mathbb{P}(\min\{x_n, x'_n\} \leq \alpha n) \leq \mathbb{P}(x_n \leq \alpha n) + \mathbb{P}(x'_n \leq \alpha n)$ and both quantities are exponentially small. To prove 3, we still assume that $\sigma \leq \sigma'$. Then for all $\alpha' < \sigma'$, we get that $\mathbb{P}(\min\{x_n, x'_n\} \leq \alpha n) \leq \mathbb{P}(x'_n \leq \alpha' n)$ which is exponentially small.

Now to prove 4 consider some $\alpha_+ < \lambda\sigma + \lambda'\sigma'$, then we have some $\alpha < \sigma$ and $\alpha' < \sigma'$ such that $\alpha_+ = \lambda\alpha + \lambda'\alpha'$. Then consider $C, \beta > 0$ such that $\mathbb{P}(x_n \leq \alpha n) \leq C\exp(-\beta n)$ for all $n \in \mathbb{N}$ and $C', \beta' > 0$ such that $\mathbb{P}(x'_n \leq \alpha' n) \leq C'\exp(-\beta' n)$. Then we have for all $n \in \mathbb{N}$:

$$\begin{aligned}\mathbb{P}(\lambda x_n + \lambda' x'_n \leq \alpha_+ n) &\leq \mathbb{P}(x_n \leq \alpha n) + \mathbb{P}(x'_n \leq \alpha' n)\\ &\leq C\exp(-\beta n) + C'\exp(-\beta' n)\\ &\leq (C + C')\exp(-\min\{\beta, \beta'\} n)\end{aligned}$$

Now we prove 5. Take $\alpha < \kappa\sigma$ and write $\alpha = \alpha'\alpha''$ with $\alpha' < \sigma$ and $\alpha'' < \kappa$. Then take some adapted constants $C_x, \beta_x > 0$ such that $\mathbb{P}(x_n \leq \alpha' n) \leq \exp(-\beta n)$ for all $n \in \mathbb{N}$ and C_k, β_k such that $\mathbb{P}(k_n \leq \alpha'' n) \leq C_k \exp(-\beta_k n)$ for all $n \in \mathbb{N}$. Then we have for all $n \in \mathbb{N}$:

$$\begin{aligned}\mathbb{P}(x_{k_n} \leq \alpha n) &\leq \mathbb{P}(x_{k_n} \leq \alpha' k_n | k_n \geq \alpha'' n) + \mathbb{P}(k_n \leq \alpha'' n)\\ \mathbb{P}(x_{k_n} \leq \alpha n) &\leq \sum_{k \geq \alpha'' n} \mathbb{P}(x_k \leq \alpha' k) + \mathbb{P}(k_n \leq \alpha'' n)\\ \mathbb{P}(x_{k_n} \leq \alpha n) &\leq \sum_{k \geq \alpha'' n} C_x \exp(-\beta_x k) + C_k \exp(-\beta_k n)\\ \mathbb{P}(x_{k_n} \leq \alpha n) &\leq \frac{C_x}{\beta_x}\exp(-\beta_x \alpha'' n) + C_k \exp(-\beta_k n)\\ \mathbb{P}(x_{k_n} \leq \alpha n) &\leq \left(\frac{C_x}{\beta_x} + C_k\right)\exp(-\min\{\beta_x\alpha'', \beta_k\} m_0).\end{aligned}$$

To prove 6, we use a similar method. Take some $\alpha < \kappa$ and $C, \beta > 0$. First assume that $\mathbb{P}(k_n \leq \alpha n) \leq C\exp(-\beta n)$ for all $n \in \mathbb{N}$. Then we have for all $m_0 \in \mathbb{N}$:

$$\begin{aligned}\mathbb{P}(r_{m_0} \geq \alpha^{-1} m_0) &\leq \mathbb{P}(\exists m \geq m_0,\ r_m \geq \alpha^{-1} m)\\ &\leq \mathbb{P}(\exists m \geq m_0,\ \exists n \in \mathbb{N},\ (n \geq \alpha^{-1} m) \wedge (k_n \leq m))\\ &\leq \mathbb{P}(\exists n \geq \alpha^{-1} m_0,\ k_n \leq \alpha n)\\ &\leq \frac{C}{\beta}\exp(-\beta\alpha^{-1} m_0).\end{aligned}$$

Now note that for all $\alpha' > \kappa^{-1}$, we have $\alpha^{-1} < \kappa$. So the above tells us that (r_m) satisfies large deviations inequalities above the speed κ^{-1}. Reciprocally, assume that $\mathbb{P}(r_m \geq \alpha^{-1} m) \leq C\exp(-\beta m)$ for all $m \in \mathbb{N}$. Then we have for all $n_0 \in \mathbb{N}$:

$$\begin{aligned}\mathbb{P}(k_{n_0} \leq \alpha n) &\leq \mathbb{P}(\exists n \geq n_0,\ k_n \leq \alpha n)\\ &\leq \mathbb{P}(\exists n \geq n_0,\ k_n \leq \lfloor \alpha n \rfloor)\\ &\leq \mathbb{P}(\exists m \geq \lfloor \alpha n_0 \rfloor,\ \ r_m \geq \alpha^{-1} m)\\ &\leq \frac{C\exp(\beta)}{\beta}\exp(-\beta\alpha n_0).\end{aligned}$$

Therefore (k_n) satisfies large deviations inequalities below the speed κ. This proves 6 by double implication. □

B.4 About moments

Definition B.30 (L^p-integrability)**.** *Let $p \in (0, +\infty)$. Let η be a probability distribution on $\mathbb{R}_{\geq 0}$. We say that η is strongly L^p integrable or has moment of order p if:*

$$M_p(\eta) := \int_{t=0}^{+\infty} t^p \mathrm{d}\eta(t) < +\infty. \tag{sL^p}$$

We say that η is weakly L^p integrable if:

$$W_p(\eta) := \sup_{t\geq 0} t^p \eta\,(t,+\infty) < +\infty. \tag{wL^p}$$

Remark B.31. *Note that a probability distribution η on $\mathbb{R}_{\geq 0}$ is characterized by the map $t \mapsto \eta(t,+\infty)$. Moreover for all non-increasing $f : \mathbb{R}_{\geq 0} \to [0,1]$ such that $\lim_{t\to\infty} f(t) = 0$. There is a probability distribution η such that $f(t) = \eta(t,+\infty)$ for all $t \geq 0$. Indeed, we simply have to write $\eta(\{0\}) = 1 - f(0)$ and for all $0 \leq a < b$ we write $\eta(a,b] := f(a) - f(b) \geq 0$. This is clearly an additive function on half-open intervals. Then we only need to see that intervals of type $(a,b]$ generate the Borel σ-algebra or $\mathbb{R}_{\geq 0}$.*

Definition B.32 (Trunking)**.** *Let η' be a distribution on $\mathbb{R}_{\geq 0}$ that has finite total mass i.e, a multiple of a probability distribution. We call trunking of η' the distribution η characterized by:*

$$\forall t \geq 0,\ \eta(t,+\infty) = \min\{1, \eta'(t,+\infty)\}.$$

Given a semi-norm $\|\cdot\|$ on the set of probability distributions on $\mathbb{R}_{\geq 0}$, we write $\|\eta'\| := \|\eta\|$. Note that if η' has total mass less that one, then $\eta = \eta' + 1 - \eta'(\mathbb{R}_{\geq 0})\delta_0$.

Given η a probability distribution on $\mathbb{R}_{\geq 0}$ and $C \in \mathbb{R}_{\geq 0}$, Definition B.32 allows us to define a probability distribution $C\eta$ on $\mathbb{R}\geq 0$.

Definition B.33 (Push up Distribution)**.** *Let κ and η be probability distributions on $\mathbb{R}\geq 0$. We define the B-push-up $\kappa \vee \eta$ by:*

$$\forall t \geq 0,\ \kappa \wedge \eta(t,+\infty) := \max\{\kappa(t,+\infty) + \eta(t,+\infty), 1\}.$$

In other words, $\kappa \vee \eta$ is the distribution of $\max\{x(t), y(t)\}$ for t taken uniformly in $[0,1]$ and $x(t) \sim \kappa$ and $y(t) \sim \eta$ with x non-increasing and y non-decreasing. For C a constant, we write $C \vee \eta$ for the distribution $\delta_C \vee \eta$

Lemma B.34. *Let η be a probability distribution and $B \in \mathbb{R}\geq 0$ and $p \in \mathbb{R}_{>0}$. We have:*

$$M_p(\kappa \vee \eta) \leq M_p(\kappa) + M_p(\eta) \tag{146}$$
$$W_p(\kappa \vee \eta) \leq W_p(\kappa) + W_p(\eta) \tag{147}$$

Proof. We do the calculations:

$$\begin{aligned} M_p(\kappa \vee \eta) & \int_{t=0}^{\infty} t^p \mathrm{d}(\kappa \vee \eta)\,(t) \\ &= \int_{t=0}^{\infty} t^p \mathrm{d}(\kappa)\,(t) + \int_{t=0}^{\infty} t^p \mathrm{d}(\eta)\,(t) \\ &\leq M_p(\kappa) + M_p(\eta). \end{aligned}$$

This proves (146). For (147), the same reasoning holds:

$$\begin{aligned} M_p(\kappa \vee \eta) & \sup_{t\geq 0} t^p (\kappa \vee \eta)\,(t,+\infty) \\ &\leq \sup_{t\geq 0} t^p \kappa\,(t,+\infty) + t^p \eta\,(t,+\infty) \\ &\leq W_p(\kappa) + W_p(\eta). \end{aligned}$$

□

Definition B.35 (Coarse convolution)**.** *Let η be a probability distribution on $\mathbb{R}_{\geq 0}$. We define the coarse convolution $\eta^{\uparrow k}$ as:*

$$\forall t \geq 0,\ \eta^{\uparrow k}(t,+\infty) := \min\left\{1, k\eta\left(\frac{t}{k}, +\infty\right)\right\}.$$

Lemma B.36. *Let $x_1, \dots, x_k$ be random variables such that:*

$$\forall t \geq 0,\ \forall i \in \{1,\dots,k\},\ \mathbb{P}(x_i > t) \leq \eta(t,+\infty). \tag{148}$$

Then we have:

$$\forall t \geq 0,\ \mathbb{P}(x_1 + \cdots + x_k > t) \leq \eta^{\uparrow k}(t,+\infty). \tag{149}$$

Proof. Consider some $t \geq 0$, we have:

$$\begin{aligned}\mathbb{P}(x_1 + \cdots + x_k > t) &\leq \mathbb{P}\left(\exists i \in \{1, \ldots, k\},\ x_i > \frac{t}{k}\right) \\ &\leq k\eta\left(\frac{t}{k}, +\infty\right). \qquad \square\end{aligned}$$

Lemma B.37. *Let η, be a probability distributions on $\mathbb{R}_{\geq 0}$. Let $k \in \mathbb{N}_{\geq 1}$. Let $\|\cdot\|_{\sim p}$ be $\|\cdot\|_{\mathrm{L}^p}$ or $\|\cdot\|_{\mathrm{L}^p_*}$. We have:*

$$W_p(\eta^{\uparrow k}) \leq k^{p+1} W_p(\eta) \tag{150}$$

$$M_p(\eta^{\uparrow k}) \leq k^{p+1} M_p(\eta) \tag{151}$$

Proof. For the weak moment, we have:

$$\begin{aligned}W_p(\eta^{\uparrow k}) &= \max t^p \eta^{\uparrow k}(t, +\infty) \\ &\leq \max t^p k\eta\left(\frac{t}{k}, +\infty\right) \\ &\leq k^{p+1} W_p(\eta).\end{aligned}$$

This proves (151). For the strong moment, by integration by parts and by linear change of variables, we have:

$$\begin{aligned}M_p(\eta^{\uparrow k}) &= \int_{t=0}^{+\infty} t^{p-1}\eta^{\uparrow k}(t, +\infty)\mathrm{d}t \\ &\leq \int t^{p-1} k\eta\left(\frac{t}{k}, +\infty\right)\mathrm{d}t \\ &\leq k^{p+1} M_p(\eta).\end{aligned}$$

This proves (150). $\square$

Lemma B.38. *Let η, η' be probability distribution over $\mathbb{R}_{\geq 0}$ and $C \in \mathbb{R}_{\geq 0}$. We have:*

$$M_p(\eta + C\eta') \leq M_p(\eta) + CM_p(\eta') \tag{152}$$

$$W_p(\eta + C\eta') \leq W_p(\eta) + CW_p(\eta'). \tag{153}$$

Proof. This is a direct consequence of the linearity of the integral and sub-additivity of the supremum just like for Lemma B.34. $\square$

Lemma B.39. *Let η be a probability distribution on $\mathbb{R}_{\geq 0}$. Let $B, C, \beta > 0$ be constants. Consider the distribution:*

$$\kappa := \sum_{k=0}^{\infty} C\exp(-\beta k)\eta^{\uparrow k}.$$

Then κ is supported on $\mathbb{R}_{\geq 0}$ and has finite total mass. Moreover, given $p \in \mathbb{R}_{>0}$, if η is strongly (resp. weakly) L^p, then κ also is.

Proof. We first show that κ has finite total mass and is supported on $\mathbb{R}_{\geq 0}$. Consider some $\varepsilon > 0$. let k_0 be such that $\sum_{k=k_0}^{\infty} C\exp(-\beta k) \leq \frac{\varepsilon}{2}$ and t be such that for all $k \in \{0, \ldots, k_0-1\}$, we have $C\exp(-\beta k)(B \vee \eta)^{\uparrow k}(t, +\infty) \leq \frac{\varepsilon}{2k_0}$. Then $\kappa(t, +\infty) \leq \varepsilon$ so $\kappa(t, +\infty) \to 0$, which means that κ has finite mass and is supported on $\mathbb{R}_{\geq 0}$.

Write $L_p \in \{M_p, W_p\}$. By Lemma B.38, we have $L_p(\kappa) \leq \frac{C}{\beta^{p+2}} L_p(\eta)$. $\square$

Lemma B.40. *Let n be a random integer and $x_1, \ldots, x_n$ be non-negative real random variables. Let B, C_1 be such that:*

$$\forall t \geq B,\ \forall k \in \mathbb{N},\ \forall m \leq k,\ \mathbb{P}(x_m \geq t \,|\, k = n) \leq C_1\eta(t) \tag{154}$$

Let $C_2, \beta > 0$ be such that:

$$\forall k \in \mathbb{N},\ \mathbb{P}(n = k) \leq C_2 \exp(-\beta k). \tag{155}$$

Then for $C := C_1 C_2$, we have:

$$\forall t \geq 0,\ \mathbb{P}(x_1 + \cdots + x_n > t) \leq \sum_{k=0}^{\infty} C \exp(-\beta k)(B \vee \eta)^{\uparrow k}$$

Proof. First note that (154) with Definition B.33 and Lemma B.36 implies that for all $k \in \mathbb{N}$, we have :

$$\forall t \geq 0, \mathbb{P}(x_1 + \cdots + x_k > t \,|\, n = k) \leq C_1 (B \vee \eta)^{\uparrow k}(t, +\infty)$$

We do the computation, for all $t \geq 0$:

$$\begin{aligned} \mathbb{P}(x_1 + \cdots + x_n > t) &= \sum_{k=0}^{\infty} \mathbb{P}(n = k)\mathbb{P}(x_1 + \cdots + x_k > t \,|\, n = k) \\ &\leq \sum_{k=0}^{\infty} C_2 \exp(-\beta k) C_1 (B \vee \eta)^{\uparrow k}(t, +\infty). \end{aligned}$$

□

Remark B.41. *Let η be a non-trivial probability distribution on $\mathbb{R}_{\geq 0}$ i.e, $\eta \neq \delta_0$. Then, for all $B, C, \beta > 0$, there are constants C_0, β_0 such that for all $t \in \mathbb{R}$, we have:*

$$\sum_{k=0}^{\infty} C \exp(-\beta k)(B \vee \eta)^{\uparrow k}(t, +\infty) \leq \sum_{k=0}^{\infty} C_0 \exp(-\beta_0 k)\eta(t/k, +\infty).$$

Proof. First note that $(B \vee \eta)^{\uparrow k}(t, +\infty) \leq k\eta(t/k - B, +\infty)$. Then for all $\beta'' < \beta$, we have $k << \exp((\beta - \beta'')k)$ so for C'' large enough, we have $k \exp(-\beta k) \leq C'' \exp(-\beta'' k)$ for all k. Take such a $\beta'' > 0$ and such a C''. Now we re-index the sum by taking $k' = 2k$ and write $\beta' := \beta/2$ and $C' = CC''$, then we have:

$$\begin{aligned} \sum_{k=0}^{\infty} C \exp(-\beta k)(B \vee \eta)^{\uparrow k}(t, +\infty) &\leq \sum_{k=0}^{\infty} CC'' \exp(-\beta'' k)\eta(t/k - B, +\infty) \\ &\leq \sum_{k'=0}^{\infty} C' \exp(-\beta' k')\eta(2t/k' - B, +\infty) \\ &\leq \sum_{k'=0}^{\lceil t/B \rceil} C' \exp(-\beta' k')\eta(t/k', +\infty) + \sum_{k'=\lceil t/B \rceil}^{\infty} C' \exp(-\beta' k') \\ &\leq \sum_{k'=0}^{+\infty} C' \exp(-\beta' k')\eta(t/k', +\infty) + \frac{C'}{\beta'} \exp(-\beta' t/B). \end{aligned}$$

Now we use the fact that $\eta \neq \delta_0$ and take $a > 0$ such that $\eta(a, +\infty) > 0$. Then we have for all $t \in \mathbb{R}$ and all $C_0 \geq 0$ and all $\beta_0 > 0$:

$$\sum_{k=0}^{+\infty} C_0 \exp(-\beta_0 k)\eta(t/k, +\infty) \geq C_0 \exp(-\beta_0 \lceil t/a \rceil)\eta(a, +\infty).$$

Now for β_0 small enough, we have $-\beta_0 \lceil t/a \rceil \leq \beta' t/B$ for all $t \geq 0$ and $\beta_0 \leq \beta'$ and for C_0 large enough, we have $C_0 \geq 2C'/\beta'$ and $C_0 \geq 2C'$. Then we have:

$$\sum_{k'=0}^{+\infty} C' \exp(-\beta' k')\eta(t/k', +\infty) + \frac{C'}{\beta'} \exp(-\beta' t/B) \leq \sum_{k=0}^{+\infty} C_0 \exp(-\beta_0 k)\eta(t/k, +\infty)$$

$$\sum_{k=0}^{\infty} C \exp(-\beta k)(B \vee \eta)^{\uparrow k}(t, +\infty) \leq \sum_{k=0}^{+\infty} C_0 \exp(-\beta_0 k)\eta(t/k, +\infty).$$

□

References

[AG20] Richard Aoun and Yves Guivarc'h. Random matrix products when the top lyapunov exponent is simple, 2020.

[AMS95] H. Abels, G. A. Margulis, and G. A. Soifer. Semigroups containing proximal linear maps. *Israel Journal of Mathematics*, 91(1):1–30, 1995.

[Aou11] Richard Aoun. Random subgroups of linear groups are free. *Duke Mathematical Journal*, 160(1):117 – 173, 2011.

[Aou20] Richard Aoun. The central limit theorem for eigenvalues, 2020.

[AS21] Richard Aoun and Cagri Sert. Law of large numbers for the spectral radius of random matrix products, 2021.

[Bel54] Richard Bellman. Limit theorems for non-commutative operations. I. *Duke Mathematical Journal*, 21(3):491 – 500, 1954.

[BL85] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Prog. Probab. Stat.* Birkhäuser, Boston, MA, 1985.

[BMSS20] Adrien Boulanger, Pierre Mathieu, Cagri Sert, and Alessandro Sisto. large deviations for random walks on gromov-hyperbolic spaces, 2020.

[BQ16a] Yves Benoist and Jean-François Quint. *Random walks on reductive groups*, volume 62 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Cham, 2016.

[BQ16b] Yves Benoist and Jean-François Quint. Central limit theorem for linear groups. *The Annals of Probability*, 44(2):1308 – 1340, 2016.

[CDJ16] Christophe Cuny, Jerome Dedecker, and Christophe Jan. Limit theorems for the left random walk on gld (r), 2016.

[CDM17] Christophe Cuny, Jérôme Dedecker, and Florence Merlevède. On the komlós, major and tusnády strong approximation for some classes of random iterates, 2017.

[CDM23] C Cuny, J Dedecker, and F Merlevède. Limit theorems for iid products of positive matrices, 2023.

[CFFT22] Kunal Chawla, Behrang Forghani, Joshua Frisch, and Giulio Tiozzo. The poisson boundary of hyperbolic groups without moment conditions, 2022.

[Cho22] Inhyeok. Choi. Limit laws on outer space, teichmüller space, and $cat(0)$ spaces. *Preprint*, 2022.

[Die54] Jean Dieudonné. Groupes de Lie algébriques (travaux de Chevalley). In *Séminaire Bourbaki : années 1951/52 - 1952/53 - 1953/54, exposés 50-100*, number 2 in Séminaire Bourbaki. Société mathématique de France, 1954. talk:57.

[FK60] H. Furstenberg and H. Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, 1960.

[Fur63] Harry Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108(3):377–428, 1963.

[Fur73] Harry Furstenberg. Boundary theory and stochastic processes on homogeneous spaces. *Harmonic analysis on homogeneous spaces*, 26:193–229, 1973.

[GM89] Ilya Ya. Goldsheid and G. A. Margulis. Lyapunov indices of a product of random matrices. *Russian Mathematical Surveys*, 44:11–71, 1989.

[Gou22] Sébastien Gouëzel. Exponential bounds for random walks on hyperbolic spaces without moment conditions. *Tunisian Journal of Mathematics*, 4(4):635–671, December 2022.

[GP16] Y. Guivarc'h and É. Le Page. Spectral gap properties for linear random walks and Pareto's asymptotics for affine stochastic recursions. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 52(2):503 – 574, 2016.

[GQX20] Ion Grama, Jean-François Quint, and Hui Xiao. A zero-one law for invariant measures and a local limit theorem for coefficients of random walks on the general linear group, 2020.

[GR86] Y. Guivarc'h and A. Raugi. Products of random matrices: convergence theorems. In *Random matrices and their applications (Brunswick, Maine, 1984)*, volume 50 of *Contemp. Math.*, pages 31–54. Amer. Math. Soc., Providence, RI, 1986.

[GR89] Yves Guivarc'h and Albert Raugi. Propriétés de contraction d'un semi-groupe de matrices inversibles. coefficients de liapunoff d'un produit de matrices aléatoires indépendantes. *Israel Journal of Mathematics*, 65(2):165–196, 1989.

[Hen97] H. Hennion. Limit theorems for products of positive random matrices. *The Annals of Probability*, 25(4):1545 – 1587, 1997.

[Kin68] J. F. C. Kingman. The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B*, 30:499–510, 1968.

[KS84] Harry Kesten and Frank Spitzer. Convergence in distribution of products of random matrices. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(4):363–386, 1984.

[MS20] P. Mathieu and A. Sisto. Deviation inequalities for random walks. *Duke Mathematical Journal*, 169(5), Apr 2020.

[Muk87] Arunava Mukherjea. Convergence in distribution of products of random matrices: A semigroup approach. *Transactions of the American Mathematical Society*, 303(1):395–411, 1987.

[Sen06] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer Series in Statistics. Springer, 2006.

[Ser18] Cagri Sert. Large deviation principle for random matrix products, 2018.

[XGL21] Hui Xiao, Ion Grama, and Quansheng Liu. Limit theorems for the coefficients of random walks on the general linear group, 2021.

[XGL22] Hui Xiao, Ion Grama, and Quansheng Liu. Edgeworth expansion and large deviations for the coefficients of products of positive random matrices, 2022.