

# LARGE DEVIATION PRINCIPLE FOR RANDOM MATRIX PRODUCTS

BY CAGRI SERT

*ETH Zürich*

Under a Zariski density assumption, we extend the classical theorem of Cramér on large deviations of sums of iid real random variables to random matrix products.

**1. Introduction.** Let  $S$  be a set of  $d \times d$  real invertible matrices and  $\mu$  be a probability measure on  $S$ . Let  $X_1, X_2, \dots$  be independent  $S$ -valued random variables with distribution  $\mu$ . Consider the random product  $Y_n = X_n \dots X_1$ . One of the goals of the theory of random matrix products is to understand the limiting behaviour of this random product as  $n$  tends to infinity. A convenient way to do this is to study the *extensions* of classical limit theorems (law of large numbers, central limit theorem, Cramér's theorem and so on) for the norm of this random product. More precisely, choose a norm  $\|\cdot\|$  on  $\mathbb{R}^d$  and consider the associated operator norm  $\|\cdot\|$  on  $\text{Mat}_d(\mathbb{R})$  (the choice of norm is irrelevant to our discussion). One is interested in studying the probabilistic limiting behaviour of  $\log \|Y_n\|$ . Note that when  $d = 1$ , this is precisely a sum of independent identically distributed (iid) real random variables, i.e. the subject of study of classical limit theorems in probability theory. When  $d > 1$ , there are at least two new aspects: the operation is no longer commutative and the log-norm functional is only *subadditive*. In this article, we shall be working in a more general setting and we will consider a slightly more general multi-norm given by classical decompositions of Lie groups, which we now describe.

For the sake of exposition, let  $G$  be a connected semisimple linear Lie group, e.g.  $\text{SL}(d, \mathbb{R})$  (more generally, we prove our results in the setting of a group of  $k$ -points of a connected reductive algebraic group defined over a local field  $k$ ). The multi-norm that we shall consider comes from the classical Cartan decomposition: let  $\mathfrak{g}$  be the Lie algebra of  $G$ ,  $\mathfrak{a}$  be a Cartan subalgebra in  $\mathfrak{g}$  and  $\mathfrak{a}^+$  be a chosen Weyl chamber in  $\mathfrak{a}$ . Let  $K$  be a maximal compact subgroup of  $G$  for which we have the Cartan decomposition  $G = K \exp(\mathfrak{a}^+)K$ . This decomposition allows one to consider the mapping  $\kappa : G \rightarrow \mathfrak{a}^+$ , called the Cartan projection or multi-norm, satisfying for every

---

*MSC 2010 subject classifications:* Primary 60F10, 20P05; secondary 22E46

*Keywords and phrases:* Large deviation principle, random matrix products

$g \in G$ ,  $g \in k \exp(\kappa(g))u$  for some  $k, u \in K$ . In the case of  $G = \mathrm{SL}(d, \mathbb{R})$ , this is the usual polar decomposition and for an element  $g \in \mathrm{SL}(d, \mathbb{R})$ , the multi-norm  $\kappa(g)$  writes as  $\kappa(g) = (\log \|g\|, \log \frac{\|\wedge^2 g\|}{\|g\|^2}, \dots, \log \frac{\|\wedge^d g\|}{\|\wedge^{d-1} g\|})$ , where  $\wedge^k \mathbb{R}^d$ 's are endowed with their canonical Euclidean structures and  $\|\cdot\|$ 's denote the associated operator norms. The components of  $\kappa(g)$  are the logarithms of the singular values of  $g$ .

Now let  $\mu$  be a probability measure on  $G$  and  $X_1, X_2, \dots$  be  $G$ -valued iid random variables with distribution  $\mu$ . Consider the random product  $Y_n$  and its multi-norm  $\kappa(Y_n)$ . The first limit theorem that was proven for random matrix products is the analogue (extension) of the law of large numbers. Stating it in our setting, Furstenberg-Kesten's result [16] reads: if  $\mu$  is a probability measure on  $G$  with a finite first moment (i.e.  $\int \|\kappa(g)\| \mu(dg) < \infty$  for some norm  $\|\cdot\|$  on  $\mathfrak{a}$ ), then the  $\mu$ -random walk  $Y_n = X_n \dots X_1$  satisfies

$$\frac{1}{n} \kappa(Y_n) \xrightarrow[n \rightarrow \infty]{a.s.} \vec{\lambda}_\mu \in \mathfrak{a}$$

where  $\vec{\lambda}_\mu$  can be defined by this and is called the Lyapunov vector of  $\mu$ . Nowadays, this result is a corollary of Kingman's subadditive ergodic theorem.

A second important limit theorem that was established in increasing generality by Tutubalin [29], Le Page [21], Goldsheid-Guivarc'h [18], and Benoist-Quint [9], [8] is the central limit theorem (CLT). Benoist-Quint's CLT reads: if  $\mu$  is a probability measure on  $G$  with finite second order moment and such that the support of  $\mu$  generates a Zariski-dense semigroup in  $G$ , then  $\frac{1}{\sqrt{n}}(\kappa(Y_n) - n\vec{\lambda}_\mu)$  converges in distribution to a non-degenerate Gaussian law on  $\mathfrak{a}$ . A feature of this result is the Zariski density assumption which also appears in our result below. We note that the fact that the support  $S$  of the probability measure  $\mu$  generates a Zariski-dense semigroup can be read as: any polynomial that vanishes on  $\cup_{n \geq 1} S^n$  also vanishes on  $G$  (recall that when  $d = 1$ , a subset is Zariski dense if and only if it is infinite). Some other limit theorems whose analogues have been obtained are the law of iterated logarithm and local limit theorems, for which we refer the reader to the nice books of Bougerol-Lacroix [10] and more recently Benoist-Quint [8].

An essential and, until now, a rather incomplete aspect of these non-commutative limit theorems is concerned with large deviations. The main result in this direction is that of Le Page [21], (see also Bougerol [10]) and its extension by Benoist-Quint [8], stating the exponential decay of probabilities of large deviations off the Lyapunov vector. Before stating this result,

recall that a probability measure  $\mu$  on  $G$  is said to have a finite exponential moment, if there exists  $\alpha > 1$  such that  $\int \alpha^{\|\kappa(g)\|} \mu(dg) < \infty$ . We have

**THEOREM 1.1** (Le Page [21], Benoist-Quint [8]). *Let  $G$  be as before,  $\mu$  be a probability measure of finite exponential moment on  $G$  whose support generates a Zariski-dense semigroup in  $G$ . Then, for all  $\epsilon > 0$ , we have  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\|\frac{1}{n}\kappa(Y_n) - \vec{\lambda}_\mu\| > \epsilon) < 0$ .*

In our first main result, under the usual Zariski density assumption, we prove the matrix extension of Cramér's classical theorem about large deviations for iid real random variables. Let  $X$  be a topological space and  $\mathcal{F}$  be a  $\sigma$ -algebra on  $X$ .

**DEFINITION 1.2.** *A sequence  $Z_n$  of  $X$ -valued random variables is said to satisfy a large deviation principle (LDP) with rate function  $I : X \rightarrow [0, \infty]$ , if for every measurable subset  $R$  of  $X$ , we have*

$$-\inf_{x \in \text{int}(R)} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in R) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in R) \leq -\inf_{x \in \bar{R}} I(x)$$

where,  $\text{int}(R)$  denotes the interior and  $\bar{R}$  the closure of  $R$ .

With this definition, Cramér's theorem says that the sequence of averages  $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$  of real iid random variables of finite exponential moment satisfies an LDP with a proper convex rate function  $I$ , given by the convex conjugate (Legendre transform) of the Laplace transform of  $X_i$ 's. Our first main result reads

**THEOREM 1.3.** *Let  $G$  be a connected semisimple linear real algebraic group and  $\mu$  be a probability measure of finite exponential moment on  $G$ , whose support generates a Zariski dense semigroup of  $G$ . Then, the sequence of random variables  $\frac{1}{n}\kappa(Y_n)$  satisfies an LDP with a proper convex rate function  $I : \mathfrak{a} \rightarrow [0, \infty]$  having a unique zero at the Lyapunov vector  $\vec{\lambda}_\mu$  of  $\mu$ .*

**REMARK 1.4.** *1. In Theorem 3.4, without any moment assumptions on  $\mu$ , we also obtain a weaker result which is an extension of a result of Bahadur [2] for iid real random variables.*

*2. In Theorem 3.5, under a stronger exponential moment condition, by exploiting convexity of  $I$ , we are able to identify the rate function  $I$  with the convex conjugate of a limiting Laplace transform of the random variables  $\frac{1}{n}\kappa(Y_n)$ .*

*3. We note that the unique zero assertion for  $I$  in the previous theorem is a*

reformulation of the exponential decay result expressed in Theorem 1.1.

4. In Section 6, we conjecture that a similar LDP holds for the Jordan projection  $\lambda : G \rightarrow \mathfrak{a}^+$  in place of  $\kappa$  (see the definition of Jordan projection below).

REMARK 1.5. *Let us also mention that if the Zariski closure of the semigroup generated by the support of the measure  $\mu$  is compact or unipotent, the conclusion of this theorem is still valid. In this case the rate function  $I$  is degenerate, its effective support  $D_I := \{x \in \mathfrak{a} \mid I(x) < \infty\}$  equals  $\{0\} \subset \mathfrak{a}$ .*

Coming back to the initial setting of norms of matrices, let  $V$  be a finite dimensional real vector space and recall that a subgroup  $\Gamma$  of  $\mathrm{GL}(V)$  is said to be completely reducible if  $V$  is a direct sum of  $\Gamma$ -irreducible subspaces. By the so-called contraction principles for LDP's, Theorem 1.3 (see also Theorem 3.5) yields the following corollary:

COROLLARY 1.6. *Let  $\mu$  be a probability measure with finite exponential moment on  $\mathrm{GL}(V)$  and suppose that the group generated by the support of  $\mu$  is completely reducible. Then the sequence of random variables  $\frac{1}{n} \log \|Y_n\|$  satisfies an LDP with a proper convex rate function  $I : \mathbb{R} \rightarrow [0, \infty]$  having a unique zero at the first Lyapunov exponent of  $\mu$ .*

We note that Remark 1.4 also applies to this corollary.

In the second part of this article, we study the effective support of the rate function  $I$  given by the previous theorem. By convexity of  $I$ , the effective support  $D_I$  is clearly a convex subset of  $\mathfrak{a}$ . Our second main result gives more information on this set. One important feature is that when the support  $S$  of the probability measure  $\mu$  is a bounded subset of  $G$ , we show that the effective support of  $I$  is identified with a set of deterministic construction depending only on  $S$ , namely the joint spectrum  $J(S)$  of  $S$ , which we now describe: let  $G$  be a connected semisimple linear Lie group as before. Denote by  $\lambda : G \rightarrow \mathfrak{a}^+$  the Jordan projection of  $G$ : for an element  $g \in G$ , if  $g = g_e g_h g_u$  is the Jordan decomposition of  $g$  with  $g_e$  elliptic,  $g_h$  hyperbolic and  $g_u$  unipotent, then  $\lambda(g)$  is defined as  $\kappa(g_h)$ . Now let  $S$  be a bounded subset of  $G$  and suppose that  $S$  generates a Zariski dense semigroup in  $G$ . In [12], it is shown that both of the sequences  $\frac{1}{n} \kappa(S^n)$  and  $\frac{1}{n} \lambda(S^n)$  of subsets of  $\mathfrak{a}^+$  converge in the Hausdorff topology to a convex body (i.e. compact, convex subset with non-empty interior) in  $\mathfrak{a}^+$ . This limit set is called the joint spectrum of  $S$  (see [12]). In these terms our second result reads

**THEOREM 1.7.** *Let  $G$  be a connected semisimple linear Lie group and let  $\mu$  be a probability measure on  $G$ . Denote by  $S$  the support of  $\mu$  and suppose that the semigroup generated by  $S$  is Zariski dense in  $G$ . Let  $I$  be the rate function given by Theorem 3.4. Then,*

1. *The effective support  $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$  of  $I$  is a convex set with non-empty interior. Moreover, if  $\mu$  has a finite second order moment, we have  $\vec{\lambda}_\mu \in \text{int}(D_I)$ .*
2. *If  $S$  is a bounded subset of  $G$ , then  $\overline{D_I} = J(S)$  and  $\text{int}(D_I) = \text{int}(J(S))$ .*
3. *If  $S$  is a finite subset of  $G$ , then  $D_I = J(S)$ .*

**REMARK 1.8.** 1. *Since  $D_I$  has non-empty interior and  $I$  is convex, it follows that  $I$  is locally Lipschitz (in particular continuous) on the interior of  $D_I$ .*

2. *Convexity of  $I$  and the identification in 2. of the previous theorem allows us to show the existence of certain limits in large deviation probabilities (see Corollary 5.4) for sufficiently regular sets  $R \subseteq \mathfrak{a}$ .*

3. *In Section 5, we present an explicit example of a probability measure  $\mu$  of bounded support  $S$  such that  $D_I \neq J(S)$ .*

Let  $B$  be a bounded subset of the matrix algebra  $\text{Mat}(d, \mathbb{R})$  endowed with an operator norm  $\|\cdot\|$ . Recall from [23] that (the logarithm of) the joint spectral radius  $r(B)$  of  $B$  is the quantity  $\lim_{n \rightarrow \infty} \sup_{x \in B^n} \frac{1}{n} \log \|x\|$ . This limit exists by subadditivity and does not depend on the norm  $\|\cdot\|$ . This generalizes the usual notion of spectral radius. Recall furthermore that the joint spectral subradius  $r_{\text{sub}}(B)$  of  $B$  is the quantity similarly defined by replacing sup by inf in the definition of  $r(B)$ . From the previous theorem and Corollary 1.6, we deduce

**COROLLARY 1.9.** *Let  $\mu$  be a probability measure on  $\text{GL}(V)$  such that the group generated by its support is completely reducible and let  $I$  be the rate function given by Theorem 3.4 (as in Corollary 1.6). Then,*

1.  *$D_I \subseteq \mathbb{R}$  is an interval with non-empty interior. Moreover, if  $\mu$  has a finite second order moment, then  $\lambda_1 \in \text{int}(D_I)$ , where  $\lambda_1$  is the first Lyapunov exponent of  $\mu$ .*
2. *If the support  $S$  of  $\mu$  is a bounded subset of  $\text{GL}(V)$ , then  $\overline{D_I} = [r_{\text{sub}}(S), r(S)]$  and  $\text{int}(D_I) = (r_{\text{sub}}(S), r(S))$ .*
3. *If  $S$  is a finite subset of  $G$ , then  $D_I = [r_{\text{sub}}(S), r(S)]$ .*

Finally, the following question remains unsettled:

QUESTION 1.10. *Is the rate function  $I$  given by Theorem 1.3 strictly convex?*

Some partial results have recently been obtained by Guivarc'h-Le Page [20] using an analytic approach. We also note that a positive answer to this question would be considerably stronger than the exponential decay result of Le Page (Theorem 1.1) which itself may be considered to indicate that  $I$  is strictly convex at least around the Lyapunov vector  $\vec{\lambda}_\mu$ .

1.11. *Overview of the argument.* We now briefly sketch the proof of the existence of an LDP as claimed in Theorem 1.3. A key tool here will be the notion of an  $(r, \epsilon)$ -Schottky semigroup. For simplicity, we shall assume that the measure  $\mu$  is compactly supported. The general fact that we use to show the existence of LDP is Theorem 3.20: we have to show that the equality  $I_i = I_s$  in that theorem is satisfied.

To fix ideas, let us speculate that  $\kappa$  was an additive mapping (i.e.  $\kappa(gh) = \kappa(g) + \kappa(h)$ ). Then the equality  $I_i = I_s$  would follow rather easily from the independence of random walk increments and uniform continuity of  $\kappa$ . Of course,  $\kappa$  is not additive, but in fact a weaker form of additivity (i.e.  $|\kappa(gh) - \kappa(g) - \kappa(h)|$  is uniformly bounded for all  $g, h \in \text{supp}(\mu)$ ) is sufficient to insure the desired equality. A key result of Benoist (see Theorem 2.20 and Proposition 3.14) shows that this weak form of additivity is satisfied in any given  $(r, \epsilon)$ -Schottky semigroup ([4]). This already finishes the proof in the case when  $\mu$  is supported on such a semigroup. For the general case, we need an argument showing that we can restrict the random walk on Schottky semigroups with no loss in the exponential rate of probabilities involved. This is done by using, first a result of Abels-Margulis-Soifer [1] about the ubiquity of proximal elements in Zariski dense semigroups (which in turn uses a result of Benoist-Labourie [7] and Prasad [22]) together with the uniform continuity of the Cartan projection, and second, a simple partitioning and pigeonhole argument.

Abels-Margulis-Soifer show that for a Zariski dense semigroup  $\Gamma$  in  $G$ , there exists  $r > 0$  such that for every  $\epsilon > 0$ , one can find a *finite* subset  $F \subset \Gamma$  with the property that for all  $\gamma \in \Gamma$ , there exists  $f \in F$  such that  $\gamma.f$  is  $(r, \epsilon)$ -proximal (see Section 3). This allows one to see that (Lemma 3.9) if the Cartan projection of the random walk hits a region of  $\mathfrak{a}^+$  at some step with some probability, after a uniformly bounded number of steps, it will hit  $(r, \epsilon)$ -proximal elements, whose Cartan projection belong to a neighborhood of that region, and this with almost the same exponential rate of probability.

The next step in the proof consists in observing that one can further restrict the random walk to a  $(r, \epsilon)$ -Schottky semigroup, again keeping al-

most the same exponential rate of probability (Corollary 3.12). By doing so, we reduce the situation to a random walk on a semigroup on which the Cartan projection  $\kappa(\cdot)$  is almost additive and hence we can conclude as we mentioned in the beginning of the argument.

1.12. *Organization of the article.* In Section 2, we review some basic properties of reductive groups over local fields and we note some variants of classical results on  $(r, \epsilon)$ -Schottky semigroups. These results will be essential in our later arguments on large deviations. In Section 3 we give two precise versions of Theorem 1.3 and prove the existence of the LDP. Section 4 is devoted to the proof of the convexity of the rate function and other assertions of Theorem 3.5. In Section 5, we give the precise version of Theorem 1.7 and prove it. Finally, in Section 6 we collect some results on large deviations for Jordan projections, make a conjecture and present some examples.

*Acknowledgements.* These results are part of author's doctoral thesis realized under the supervision of Emmanuel Breuillard in Université Paris-Sud during 2013-2016. The author would like to take the opportunity to thank him for asking the original question and numerous discussions. The author also thanks to WWU Münster where part of this work was conducted and acknowledges the supports of DIM RDM-IdF, ERC Grant 617129 and SNF Grant 200021-152819.

**2. Preliminaries from  $(r, \epsilon)$ -Schottky semigroups.** We start by indicating related definitions and results for linear transformations, we then note some basic properties of linear reductive groups over local fields and finally give relevant definitions and some variants of results on  $(r, \epsilon)$ -Schottky semigroups. We also provide an example to illustrate some of the notions for the reader only interested in matrices for the case of  $G = \mathrm{SL}(d, \mathbb{R})$ .

Let  $k$  be a local field (locally compact topological field with respect to a non-discrete topology), i.e.  $k = \mathbb{R}$  or  $\mathbb{C}$  (Archimedean, characteristic zero case) or a finite extension of  $\mathbb{Q}_p$  (non-Archimedean, characteristic zero case) or a finite extension of  $\mathbb{F}_p((T))$  (non-Archimedean, positive characteristic case). When  $k$  is Archimedean, we denote by  $|\cdot|$  the usual absolute value on  $k$ . When  $k$  is non-Archimedean, we denote  $\mathcal{O}$  the ring of integers of  $k$ ,  $\mathfrak{m}$  the maximal ideal of  $\mathcal{O}$ ,  $q$  the cardinality of the residue field and  $\varpi$  a uniformizer of  $k$ , i.e. a generator of  $\mathfrak{m}$ . We denote by  $\nu(\cdot)$  the discrete valuation on  $k$  such that  $\nu(\varpi) = 1$  and we endow  $k$  with the ultrametric norm  $|\cdot| = q^{-\nu(\cdot)}$ .

Let  $V$  be a finite dimensional  $k$ -vector space,  $X = \mathbb{P}(V)$  its projective space. If  $k$  is Archimedean, we endow  $V$  with a Euclidean norm  $\|\cdot\|$ , and

if  $k$  is non-Archimedean, we endow  $V$  with an ultrametric sup-norm  $\|\cdot\|$  associated to a basis of  $V$ . We will work with the Fubini-Study metric on  $X$ : for  $x, y \in X$ , denoting by  $v_x$  and  $v_y$  any two vectors in  $V$  projecting respectively on  $x$  and  $y$ , we have  $d(x, y) := \frac{\|v_x \wedge v_y\|}{\|v_x\| \cdot \|v_y\|}$ , where  $\|\cdot\|$  also denotes the associated norm on  $\bigwedge^2 V$ . In the sequel, we will also denote by the same  $\|\cdot\|$ , the operator norm on the  $k$ -linear endomorphisms of  $V$ , associated to the norm  $\|\cdot\|$  on  $V$ . Finally, for a metric space  $(X, d)$ , we denote by  $d_H$  the corresponding Hausdorff distance on the set of subsets of  $X$ .

**2.1. Proximal transformations.** The notion of proximality of a linear transformation is related to an important contraction property of the dynamics of its projective action. It is, for example, of essential use in the Tits' original proof of the Tits alternative in [27] through the so called ping-pong lemma. It is also in close relation to Furstenberg's earlier (quasi-) projective transformations [15]. See Breuillard-Gelander's [11] for a more detailed account and Quint's [24] for a generalization.

For  $g \in \text{End}(V)$ , denote by  $\lambda_1(g)$  the spectral radius of  $g$ . An element  $g \in \text{End}(V)$  is said to be proximal if it has a unique eigenvalue  $\alpha$  such that  $|\alpha| = \lambda_1(g)$ , and this eigenvalue is simple (in particular,  $\alpha \in k$ ). Denote by  $x_g^+$ , the element of  $X$  corresponding to the one dimensional eigenspace corresponding to  $\alpha$ . Let  $v_g^+$  be a vector of norm 1 on this line, and  $V_g^<$  the supplementary  $g$ -invariant hyperplane, and put  $X_g^< := \mathbb{P}(V_g^<) \subset X$ .

The following definition singles out special proximal elements: let  $0 < \epsilon \leq r$  and set  $b_g^\epsilon := \{x \in X \mid d(x, x_g^+) \leq \epsilon\}$  and  $B_g^\epsilon := \{x \in X \mid d(x, X_g^<) \geq \epsilon\}$ .

**DEFINITION 2.2** ([1],[3]). *Let  $0 < \epsilon \leq r$ . An element  $g \in \text{End}(V)$  is said to be  $(r, \epsilon)$ -proximal, if  $d(x_g^+, X_g^<) \geq 2r$ ,  $g(B_g^\epsilon) \subset b_g^\epsilon$ , and  $g|_{B_g^\epsilon}$  is an  $\epsilon$ -Lipschitz mapping.*

**REMARK 2.3.** *1. The notion of an  $(r, \epsilon)$ -proximal transformation, as well as the numbers  $0 < \epsilon \leq r$  depend on the choice of the norm on  $V$ .*

*2. Nevertheless, it is not hard to see that for every proximal transformation  $g$  and for any choice of norm on  $V$ , there exists  $r > 0$  such that for all  $k \in \mathbb{N}$  large enough,  $g^k$  is  $(r, \epsilon_k)$ -proximal with  $\epsilon_k \xrightarrow[k \rightarrow \infty]{} 0$ .*

**2.4. Two properties of  $(r, \epsilon)$ -proximal transformations.** The following lemma says that for  $\epsilon > 0$  small enough, the spectral radius of an  $(r, \epsilon)$ -proximal transformation can be controlled by the operator norm of this transformation:



LEMMA 2.5. *Let  $V$  be a finite dimensional  $k$ -vector space and  $0 < \epsilon \leq r$ . Then, there exist constants  $c_{r,\epsilon} \in ]0, 1[$  such that, for each  $r > 0$ , we have  $\lim_{\epsilon \rightarrow 0} c_{r,\epsilon} = 2r$ , and for every  $(r, \epsilon)$ -proximal endomorphism  $g$  of  $V$ , we have*

$$c_{r,\epsilon} \|g\| \leq \lambda_1(g) \leq \|g\|$$

PROOF. One notes that if  $(g_k)_{k \in \mathbb{N}}$  is a convergent sequence of  $(r, \epsilon_k)$ -proximal transformations such that for all  $k \in \mathbb{N}$ ,  $\|g_k\| = 1$  and  $\epsilon_k \xrightarrow[k \rightarrow \infty]{} 0$ , then  $\lim_{k \rightarrow \infty} g_k = \alpha p$ , where  $\alpha$  is a positive constant and  $p$  is a projection satisfying - denoting by  $v_p$  a non-zero vector in its image,  $x_p \in \mathbb{P}(V)$  its projective image, and by  $X_p \subset \mathbb{P}(V)$  the projective image of  $\ker p$  -  $d(x_p, X_p) \geq 2r$  (note also that the definition of an  $(r, \epsilon)$ -proximal transformation implies that  $r \leq \frac{1}{2}$ ). Since  $\|\alpha p\| = 1$ , it follows by elementary computations that we have  $\alpha \geq 2r$ , and the conclusion of lemma results from the compactness of the set of  $(r, \epsilon)$ -proximal transformations of norm 1 and continuity of the application  $\lambda_1(\cdot)$ .  $\square$

The following important proposition is a direct consequence of Lemma 1.4. in Benoist's [6] (see also Proposition 6.4. in [3]). It says that one can have a fairly good control over the spectral radii of the products of  $(r, \epsilon)$ -proximal elements in terms of the spectral radii of the factors, given that the successive factors satisfy a natural geometric condition.

PROPOSITION 2.6. *For all real numbers  $0 < \epsilon \leq r$ , there exist positive constants  $D_r$  and  $D_{r,\epsilon} > 0$  with the property that for each  $r > 0$ , we have  $\lim_{\epsilon \rightarrow 0} D_{r,\epsilon} = D_r$  and such that if  $g_1, \dots, g_l$  are  $(r, \epsilon)$ -proximal linear transformations of  $V$  satisfying (putting  $g_l = g_0$ )  $d(x_{g_{j-1}}^+, X_{g_j}^-) \geq 6r$ , for all  $j = 1, \dots, l$ , then for all  $n_1, \dots, n_l \geq 1$ , the linear transformation  $g = g_l^{n_l} \dots g_1^{n_1}$  is  $(2r, 2\epsilon)$ -proximal, and*

$$D_{r,\epsilon}^{-l} \leq \frac{\lambda_1(g_l^{n_l} \dots g_1^{n_1})}{\lambda_1(g_l)^{n_l} \dots \lambda_1(g_1)^{n_1}} \leq D_{r,\epsilon}^l$$

This proposition partly motivates the following definitions which will be of important use to us in the sequel (see also Definition 1.7 in [6]):

DEFINITION 2.7. 1. *A subset  $E$  of  $GL(V)$  is called an  $(r, \epsilon)$ -Schottky family if*

- a. *For all  $\gamma \in E$ ,  $\gamma$  is  $(r, \epsilon)$ -proximal, and*
- b.  *$d(x_\gamma^+, X_{\gamma'}^-) \geq 6r$ , for all  $\gamma, \gamma' \in E$ .*
2. *Let  $E \subset GL(V)$  be a subset consisting of proximal elements and  $a \geq 0$  be*

a real number. We say that the set  $E$  is  $a$ -narrow in  $\mathbb{P}(V)$ , if there exists a subset  $Y$  of  $\mathbb{P}(V)$  of diameter less than  $a$  such that for each  $\gamma \in E$ , we have  $x_\gamma^+ \in Y$ , and for every  $\gamma, \gamma' \in E$ , we have  $d_H(X_\gamma^<, X_{\gamma'}^<) < a$ .

REMARK 2.8. Note that, by definition, a Schottky family (i.e.  $(r, \epsilon)$ -Schottky family, for some  $r \geq \epsilon > 0$ ) cannot contain an element  $g \in GL(V)$  and its inverse  $g^{-1}$  at the same time.

The notion of proximality is related to only one special direction of the action of a linear transformation. We would like to have an equivalent property for the other/all eigenvalues and eigendirections. This property is reflected in the notion of a  $\theta$ -proximal element, which we shall shortly define.

2.9. *Connected reductive groups.* Let  $k$  be a local field,  $\mathbf{G}$  a connected reductive algebraic group defined over  $k$ . Set  $G = \mathbf{G}(k)$  and equip  $G$  with its natural locally compact topology.

Fix a maximal  $k$ -split torus  $\mathbf{A}$  of  $\mathbf{G}$ . Let  $\mathbf{Z}$  be the centralizer of  $\mathbf{A}$  in  $\mathbf{G}$  and  $\mathbf{S}$  be the derived  $k$ -subgroup of  $\mathbf{G}$ . Denote by  $d$  the  $k$ -rank of  $\mathbf{G}$  and by  $d_S$  that of  $\mathbf{S}$ . Let  $Z, A, S, G$  be the groups of  $k$ -points of  $\mathbf{Z}, \mathbf{A}, \mathbf{S}, \mathbf{G}$ , respectively.

Let  $X(\mathbf{A})$  denote the set of rational characters of  $\mathbf{A}$  (it is a free  $\mathbb{Z}$ -module of rank  $d$ ), set  $\mathfrak{a}^* = X(\mathbf{A}) \otimes_{\mathbb{Z}} \mathbb{R}$ , and let  $\mathfrak{a}$  denote the dual  $\mathbb{R}$ -vector space of  $\mathfrak{a}^*$ . There exists a unique morphism, that we denote by  $\log$ ,  $\log : Z \rightarrow \mathfrak{a}$  extending the natural morphism from  $A \rightarrow \mathfrak{a}$  (see [8] 7.1.). For any  $\chi \in X(A)$ , denote by  $\bar{\chi}$ , the unique element of  $\mathfrak{a}^*$  such that  $|\chi(\cdot)| = \exp(\bar{\chi}(\log(\cdot)))$ . In case  $k = \mathbb{R}$ ,  $\mathfrak{a}$  is the Lie algebra of  $A$ ,  $\log$  is the usual logarithm mapping (inverse of the exponential map on  $\mathfrak{a}$ ), and  $\bar{\chi}$  is the differential of  $\chi \in X(A)$ .

2.9.1. *Roots, Weyl chambers.* Let  $\Sigma$  be a root system of the pair  $(\mathbf{G}, \mathbf{A})$ , i.e. it is the set of non-trivial weights of the adjoint representation of  $A$  in the Lie algebra of  $G$ . Choose a set of positive roots  $\Sigma^+$  in  $\Sigma$ , and let  $\Pi = \{\alpha_1, \dots, \alpha_{d_S}\}$  be the simple roots in  $\Sigma^+$ . The set  $\bar{\Sigma} = \{\bar{\alpha} \in \mathfrak{a}^* \mid \alpha \in \Sigma\}$  is a root system in  $\mathfrak{a}^*$  and  $\bar{\Pi} = \{\bar{\alpha} \mid \alpha \in \Pi\}$  is a basis of this root system. Let  $W$  denote the Weyl group of this root system, put  $\mathfrak{a}^+ := \{x \in \mathfrak{a} \mid \forall \alpha \in \Sigma^+, \bar{\alpha}(x) \geq 0\}$  the closed Weyl chamber of  $\mathfrak{a}$  associated to the choice of  $\Sigma^+$ , and set  $Z^+ = \log^{-1}(\mathfrak{a}^+) \subset Z$ . Similarly, let  $\mathfrak{a}^{++} := \{x \in \mathfrak{a} \mid \forall \alpha \in \Sigma^+, \bar{\alpha}(x) > 0\}$  be the open Weyl chamber associated to  $\Sigma^+$ . The choice of  $\Sigma^+$  also induces a partial order on  $X(\mathbf{A})$ : for  $\chi_1, \chi_2$  in  $X(\mathbf{A})$ ,  $\chi_1 \geq \chi_2$  if and only if  $\bar{\chi}_1(x) \geq \bar{\chi}_2(x)$  for all  $x \in \mathfrak{a}^+$ .

We denote by  $\mathfrak{a}_C$  the subspace of  $\mathfrak{a}$  consisting of fixed points of the Weyl group  $W$ , and by  $\mathfrak{a}_S$ , the unique  $W$ -stable supplementary subspace of  $\mathfrak{a}_C$ .

We fix a  $W$ -invariant scalar product on  $\mathfrak{a}$ , and denote by  $(\bar{\omega}_1, \dots, \bar{\omega}_{d_S})$  fundamental weights of  $(\bar{\Sigma}, \bar{\Pi})$ , satisfying  $\bar{\omega}_i|_{\mathfrak{a}_C} \equiv 0$  for each  $i = 1, \dots, d_S$ . These are elements of  $\mathfrak{a}^*$  satisfying  $\frac{2\langle \bar{\omega}_i, \bar{\alpha}_j \rangle}{\langle \bar{\alpha}_j, \bar{\alpha}_j \rangle} = \delta_{ij}$  for all  $i, j = 1, \dots, d_S$ . Finally, fix a subset  $X_C$  of  $X(\mathbf{Z})$  (set of characters of  $\mathbf{Z}$ ), such that  $\bar{X}_C = \{\bar{\alpha} \mid \alpha \in X_C\}$  is a basis of  $\mathfrak{a}_C^*$  (subspace of  $W$ -fixed points of  $\mathfrak{a}^*$ ).

For a subset  $\theta$  of  $\bar{\Pi}$ , denote by  $\theta^c$ , the set  $\bar{\Pi} \setminus \theta$ . Put  $\mathfrak{a}_\theta = \bigcap_{\alpha \in \theta^c} \ker \alpha$ ,  $\mathfrak{a}_\theta^+ = \mathfrak{a}_\theta \cap \mathfrak{a}^+$ , and set  $\mathfrak{a}_\theta^{++} = \mathfrak{a}_\theta^+ \setminus (\bigcup_{\tau \in \theta} \mathfrak{a}_\tau^+)$ . The elements of the collection  $(\mathfrak{a}_\theta^+)_{\theta \subset \bar{\Pi}}$  are the faces of the convex polytope  $\mathfrak{a}^+$ . One notes that  $\mathfrak{a}_{\bar{\Pi}} = \mathfrak{a}$  and  $\mathfrak{a}_\theta$  is the subspace of  $\mathfrak{a}$  spanned by  $\bar{X}_C$ .

**2.9.2. Cartan and Jordan projections.** Let  $K$  be a maximal compact subgroup of  $G$  such that one has the Cartan decomposition  $G = KZ^+K$ . When  $\mathfrak{k}$  is Archimedean,  $K$  can be taken as the maximal compact subgroup whose Lie algebra is orthogonal to that of  $A$  for the Killing form. For the non-Archimedean case, see [13]. In the  $KZ^+K$  factorization of an element  $g \in G$ , the middle factor is uniquely defined. This allows us to define the Cartan projection  $\kappa : G \rightarrow \mathfrak{a}^+$  by requiring that for every  $g \in G$ ,  $g \in K \log^{-1}(\kappa(g))K$ . It is a proper continuous map on  $G$ .

In case  $\mathfrak{k} = \mathbb{R}$  or  $\mathbb{C}$ , every element  $g \in G$  admits a unique factorization into commuting elements as  $g = g_e g_h g_u$ , where  $g_e$  is an elliptic,  $g_h$  is an hyperbolic and  $g_u$  is a unipotent element. This is called the Jordan decomposition of  $g$ . The Jordan projection  $\lambda : G \rightarrow \mathfrak{a}^+$  is defined as  $\lambda(g) = \log(z_g)$ , where  $z_g$  is the unique element of  $Z^+$  such that  $g_h$  is conjugated to  $z_g$ . When  $\mathfrak{k}$  is non-Archimedean, such a decomposition still exists, but up to passing to a finite power of  $g$ , i.e. there exists  $n \geq 1$ , such that  $g^n = g_e g_h g_u$ , where  $g_h$  is semisimple with eigenvalues in  $\varpi^{\mathbb{Z}}$  ( $\varpi$  is the uniformizer of  $\mathfrak{k}$ ). The element  $g_h$  is conjugated to a unique element  $z_g$  of  $Z_g$ , and we set  $\lambda(g) = \frac{1}{n} \log(z_g)$ . This does not depend on  $n$ .

**2.9.3. Representations.** Let  $(V, \rho)$  be a  $\mathfrak{k}$ -rational representation of  $G$ . The weights of  $(V, \rho)$  are the characters  $\chi \in X(\mathbf{A})$  such that the associated weight space  $V_\chi = \{v \in V \mid \forall a \in A, \rho(a)v = \chi(a)v\}$  is non-trivial. If  $(V, \rho)$  is an irreducible  $\mathfrak{k}$ -rational representation, then the set of weights of  $(V, \rho)$  admits a maximal element  $\chi_\rho$  (for the partial order on  $X(\mathbf{A})$  induced by  $\mathfrak{a}^+$ ), called the highest weight of  $(V, \rho)$ . The irreducible representation  $(V, \rho)$  is said to be proximal, if  $\dim(V_{\chi_\rho}) = 1$ .

For the remaining part of this article, we **fix** the family of representations given by the next lemma. We shall refer to them as distinguished representations.

LEMMA 2.10. (*Tits [28]*) *Let  $G$  be as before. For each  $i = 1, \dots, d_S$ , there*

exists a proximal irreducible  $k$ -rational representation  $(V_i, \rho_i)$  with highest weight  $\chi_i$  such that  $\bar{\chi}_i$  is a multiple of the fundamental weight  $\bar{\omega}_i$ .

We note that for  $i = 1, \dots, d_S$ , all the other weights of  $(V_i, \rho_i)$  consist of  $(\bar{\chi}_i - \bar{\alpha}_i)$ 's and others of the form  $\bar{\chi}_i - \bar{\alpha}_i - \sum_{\beta \in \bar{\Pi}} n_\beta \bar{\beta}$  where  $n_\beta \in \mathbb{N}$ . As a consequence, for all  $g \in G$  and  $i = 1, \dots, d_S$ ,  $\rho_i(g)$  is a proximal linear transformation of  $V_i$  if and only if  $\bar{\alpha}_i(\lambda(g)) > 0$ . We also note that the mapping  $a \rightarrow (\bar{\chi}_1(a), \dots, \bar{\chi}_d(a))$ , where  $\{\bar{\chi}_{d_S+1}, \dots, \bar{\chi}_d\} = \bar{X}_C$  are the central weights, is an isomorphism of real vector spaces  $\mathfrak{a} \rightarrow \mathbb{R}^d$ .

For  $i = 1, \dots, d_S$ , we will also fix the norms  $\|\cdot\|_i$  on  $V_i$ 's, given by the next lemma.

LEMMA 2.11. ([8]) *Let  $G$  be as before and let  $(V, \rho)$  be an irreducible  $k$ -rational representation of  $G$ . Let  $\chi_\rho$  be the highest weight of  $(V, \rho)$ . Then, there exists a norm  $\|\cdot\|$  on  $V$  such that for all  $g \in G$ , we have*

1.  $\|\rho(g)\| = \exp(\bar{\chi}(\kappa(g)))$
2.  $\lambda_1(\rho(g)) = \exp(\bar{\chi}(\lambda(g)))$ .

We note that 2. does not depend on the norm and follows by definitions, and that the norm  $\|\cdot\|$  is Euclidean if  $k = \mathbb{R}$  or  $\mathbb{C}$ , and ultrametric if  $k$  is non-Archimedean (see 7.4.1. in [8]).

Lemma 2.10 and Lemma 2.11 allow us to control the Cartan and Jordan projection of an element  $g \in G$  by looking at the image of  $g$  by these projections with the central weights and  $g$ 's operator norm and spectral radius in the distinguished representations. We now see a first useful corollary of these two lemmata. We include its proof to illustrate their use.

COROLLARY 2.12 (Uniform continuity of Cartan projection). *Let  $G$  be as before and  $\kappa : G \rightarrow \mathfrak{a}^+$  be a Cartan projection of  $G$ . For every compact subset  $L$  of  $G$ , there exists a compact subset  $M$  of  $\mathfrak{a}$  such that for every  $g \in G$ , we have  $\kappa(LgL) \subseteq \kappa(g) + M$ .*

PROOF. By the paragraph following Lemma 2.10, it suffices to show that there exists a constant  $D \geq 0$  such that for every  $l_1, l_2 \in L$  and for every  $\bar{\chi} \in \{\bar{\chi}_1, \dots, \bar{\chi}_d\}$ , we have

$$(2.1) \quad |\bar{\chi}(\kappa(l_1 g l_2)) - \bar{\chi}(\kappa(g))| \leq D$$

Set  $L^{-1} = \{l^{-1} \mid l \in L\}$ ,  $C = \max_{l \in L \cup L^{-1}} \max_{i=1, \dots, d} |\bar{\chi}_i(\kappa(l))|$ ,  $D = 2C$  and let  $l_1, l_2$  be in  $L$ .

Then, for each central weight  $\bar{\chi}$  (i.e.  $\bar{\chi} = \bar{\chi}_i$  such that  $d_S + 1 \leq i \leq d$ ), we have  $\bar{\chi}(\kappa(l_1gl_2)) = \bar{\chi}(\kappa(l_1)) + \bar{\chi}(\kappa(g)) + \bar{\chi}(\kappa(l_2))$ , so that (2.1) is clearly satisfied.

Let now  $\bar{\chi}$  be the highest weight of a distinguished representation  $(V, \rho)$ . By Lemma 2.11, for all  $h \in G$ , we have  $\bar{\chi}(\kappa(h)) = \log \|\rho(h)\|$ . Then, since by submultiplicativity of the associated operator norms, for all  $x, y, u \in GL(V)$  for a normed vector space  $V$ , one has  $\|x^{-1}\|^{-1} \cdot \|y^{-1}\|^{-1} \cdot \|u\| \leq \|xuy\| \leq \|x\| \cdot \|u\| \cdot \|y\|$ , we get

$$\bar{\chi}(\kappa(g)) - 2C \leq \bar{\chi}(\kappa(l_1gl_2)) \leq \bar{\chi}(\kappa(g)) + 2C$$

and the result follows.  $\square$

*Example.* If one takes  $G = SL_d(\mathbb{R})$ , then we can write,  $\mathfrak{a} = \{(\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d \mid \sum \alpha_i = 0\}$ ,  $\mathfrak{a}^+ = \{(\alpha_1, \dots, \alpha_d) \mid \alpha_1 \geq \dots \geq \alpha_d\}$ ,  $\mathfrak{a}^{++} = \{(\alpha_1, \dots, \alpha_d) \mid \alpha_1 > \alpha_2 > \dots > \alpha_d\}$ , and  $K = SO_d(\mathbb{R})$ . The Cartan projection  $\kappa(\cdot)$  associates to an element  $g$  of  $SL_d(\mathbb{R})$ , the element of  $\mathfrak{a}$  consisting of the logarithms of the diagonal entries of the matrix  $A$  in  $KAK$  decomposition of  $g$ , i.e. it is the vector of logarithms of the singular values of  $g$  placed in decreasing order. Similarly, Jordan projection  $\lambda(\cdot)$  associates to  $g$ , the logarithms of the modules of eigenvalues of  $g$  in decreasing order.

As examples of characters on  $A = \exp(\mathfrak{a})$  (elements of  $A$  are seen as diagonal matrices), we can exhibit  $L_i$ 's for  $i = 1, \dots, d$ , defined by  $L_i(\text{diag}(a_1, \dots, a_d)) = a_i$ . The set of roots are the weights of the Ad representation of  $SL(d, \mathbb{R})$ , i.e.  $R = \{\frac{L_i}{L_j} \mid i \neq j\}$ . For our choice of  $\mathfrak{a}^+$ , the positive roots are  $\Sigma^+ = \{\frac{L_i}{L_j} \mid i < j\}$  and the set of simple roots  $\Pi = \{\frac{L_i}{L_{i+1}} \mid i = 1, \dots, d-1\}$ . On  $\mathfrak{a}$ , we have, for example,  $\overline{\frac{L_i}{L_j}}(x_1, \dots, x_d) = x_i - x_j$ . The fundamental weights are  $\omega_i = \prod_{j=1}^i L_j$ .

Some examples of proximal irreducible representations are  $\sigma_1 = id$  or, more generally,  $\sigma_i : SL(\mathbb{R}^d) \rightarrow SL(\wedge^i \mathbb{R}^d)$  where  $\sigma_i(g) := \wedge^i g$  for  $i = 1, \dots, d-1$ . These are also the fundamental representations, meaning that their highest weights are the fundamental weights  $\omega_i$ 's. The partial ordering corresponding to the choice of  $\mathfrak{a}^+$  on the set of characters of  $A$  is simply described as: for  $\chi_1, \chi_2 : A_G \rightarrow ]0, \infty[$ , we have  $\chi_1 \geq \chi_2 \iff \chi_1(a) \geq \chi_2(a)$  for all  $a \in A^+ = \exp(\mathfrak{a}^+)$ .  $\diamond$

**2.13.  $\theta$ -proximal elements.** Let  $(V_i, \rho_i)$  be the distinguished representations of  $G$  for  $i = 1, \dots, d_S$ . For each  $g \in G$ , set  $\theta_g = \{\alpha_i \in \Pi \mid \rho_i(g) \text{ is a proximal linear transformation of } V_i\}$ . By the paragraph following Lemma 2.10 and by definition of  $\mathfrak{a}^{++}$  for a subset  $\theta \subseteq \bar{\Pi}$  (see 2.9.1),  $\theta_g$  is characterized by saying  $\lambda(g) \in \mathfrak{a}_{\theta_g}^{++}$ .

DEFINITION 2.14. [Benoist [4]] 1. Let  $\theta \subseteq \Pi$ . An element  $g \in G$  is said to be  $\theta$ -proximal if for each  $\alpha_i \in \theta$ ,  $\rho_i(g)$  is proximal.

2. Let  $0 < \epsilon \leq r$  and  $\theta \subseteq \Pi$ . An element  $g \in G$  is said to be  $(\theta, r, \epsilon)$ -proximal, if for each  $\alpha_i \in \theta$ ,  $\rho_i(g)$  is  $(r, \epsilon)$ -proximal as a linear transformation of  $V_i$ .

When  $\theta = \Pi$ , we say that  $g$  is  $k$ -regular or proximal. One notes from the definitions that  $\mathfrak{a}_\theta^+$  is increasing in  $\theta$  for inclusion partial orders. Again following Benoist [4], we also set

DEFINITION 2.15. Let  $\theta \subseteq \Pi$ . We say that a sub-semigroup  $\Gamma$  is of type  $\theta$ , if  $\theta$  is the smallest subset of  $\Pi$  such that  $\{\lambda(g) \mid g \in \Gamma\} \subseteq \mathfrak{a}_\theta^+$ .

If  $\Gamma$  is of type  $\theta$ , we will sometimes denote  $\theta = \theta_\Gamma$ . Note that  $\theta_\Gamma$  is also characterized by saying that for each  $\alpha_i \in \Gamma$ , there exists  $g \in \Gamma$  such that  $\rho_i(g)$  is proximal. In other words,  $\theta_\Gamma = \bigcup_{g \in \Gamma} \theta_g$ .

For a Zariski dense semigroup  $\Gamma$  in  $G$ , we have the following useful characterization of  $\theta_\Gamma$ :

LEMMA 2.16. [[4]]  $\alpha_i \in \theta_\Gamma$  if and only if  $\bar{\alpha}_i(\kappa(\Gamma))$  is unbounded.

REMARK 2.17. 1. In particular,  $\theta_\Gamma = \emptyset$  if and only if  $\Gamma$  is bounded modulo the centre of  $G$ .

2. In case  $k = \mathbb{R}$ , for a Zariski dense semigroup  $\Gamma$  in  $G$ , it follows by Goldsheid-Margulis [17] and Benoist-Labourie [7] (see also Prasad [22]) that  $\theta_\Gamma = \Pi$ . This is clearly not true for an arbitrary local field: indeed, for  $\Gamma = \mathrm{SL}(n, \mathbb{Z}_p)$  and  $G = \mathrm{SL}(n, \mathbb{Q}_p)$ , we have  $\theta_\Gamma = \emptyset$

2.18. Two properties of  $(\theta, r, \epsilon)$ -proximal elements. We now state the multidimensional counterparts of Lemma 2.5 and Proposition 2.6. We give a proof of the following lemma (see Lemma 4.5. in [4]) to illustrate the use of previous definitions.

PROPOSITION 2.19. Let  $G$  be as before and let  $\Gamma$  be a Zariski-dense semigroup in  $G$ . Let  $r > 0$  be a constant. Then, there exists a compact set  $M_r \subset \mathfrak{a}$  such that for every  $r \geq \epsilon > 0$ , there exists a compact set  $M_{(r, \epsilon)}$  in  $\mathfrak{a}$  satisfying  $\lim_{\epsilon \rightarrow 0} M_{(r, \epsilon)} \subseteq M_r$  (Hausdorff convergence), and such that for every  $(\theta_\Gamma, r, \epsilon)$ -proximal element  $g$  of  $\Gamma$ , we have  $\lambda(g) - \kappa(g) \in M_{(r, \epsilon)}$ .

PROOF. The statement is obvious if  $\theta_\Gamma = \emptyset$  by Lemma 2.16. If not, by the same lemma, choose  $C \geq 0$  such that for every  $\alpha_i \in \theta_\Gamma^c$ ,  $|\bar{\alpha}_i(\kappa(\Gamma))| \leq C$ . On the other hand, by Lemma 2.5, there exists a constant  $C_r$  such that for

every  $r \geq \epsilon \geq 0$ , there exist constants  $C_{(r,\epsilon)}$  satisfying  $\lim_{\epsilon \rightarrow 0} C_{(r,\epsilon)} = C_r$  and such that, by Lemma 2.11, for each  $\alpha_i \in \theta_\Gamma$  and all  $(\theta_\Gamma, r, \epsilon)$ -proximal element  $g$  of  $\Gamma$ ,  $|\bar{\chi}_i(\kappa(g)) - \bar{\chi}_i(\lambda(g))| \leq C_{(r,\epsilon)}$ . Finally, note that for every central weight  $\bar{\chi} \in \bar{X}_C$ , we have  $\bar{\chi}(\kappa(g)) = \bar{\chi}(\lambda(g))$ .

Now the result follows since  $\{\bar{\alpha}, \bar{\chi}_i, \bar{\chi} \mid \alpha \in \theta_\Gamma^c, \alpha_i \in \theta_\Gamma, \chi \in X_C\}$  is a basis of  $\mathfrak{a}^*$ .  $\square$

We also have the following important counterpart of Proposition 2.6. It is proved from this proposition using Lemma 2.16, Lemma 2.11 as in the proof of the previous proposition.

**THEOREM 2.20** (Benoist [3], [4]). *Let  $G$  be the group of  $k$ -points of a connected reductive algebraic group defined over  $k$  and let  $\Gamma$  be a Zariski dense semigroup in  $G$ . For every  $r \geq \epsilon > 0$ , there exist compact sets  $N_r$  and  $N_{(r,\epsilon)}$  in  $\mathfrak{a}$ , such that for each  $r > 0$ , we have a Hausdorff convergence  $\lim_{\epsilon \rightarrow 0} N_{(r,\epsilon)} \subseteq N_r$ , and such that if  $g_1, \dots, g_l$  are  $(\theta_\Gamma, r, \epsilon)$ -proximal elements of  $\Gamma$  having the property that (noting  $g_0 = g_l$ )  $d(x_{\rho_i(g_j)}^+, X_{\rho_i(g_{j+1})}^-) \geq 6r$  for all  $j = 0, \dots, l-1$  and for all  $i = 1, \dots, d$ , then we have that for all  $n_1, \dots, n_l \geq 1$ , the element  $g = g_l^{n_l} \dots g_1^{n_1}$  is  $(\theta_\Gamma, 2r, 2\epsilon)$ -proximal, and satisfies*

$$\lambda(g_l^{n_l} \dots g_1^{n_1}) - \sum_{i=1}^l n_i \lambda(g_i) \in l \cdot N_{(r,\epsilon)} \cap \mathfrak{a}_{\theta_\Gamma}$$

Motivated by this result, analogously to Definition 2.7, we single out the following

**DEFINITION 2.21.** 1. *Let  $G$  be as above,  $r \geq \epsilon > 0$  be given constants and let  $\theta \subseteq \Pi$ . A subset  $E$  of  $G$  is said to be an  $(\theta, r, \epsilon)$ -Schottky family, if for each  $\alpha_i \in \theta$ ,  $\rho_i(E)$  is an  $(r, \epsilon)$ -Schottky family.*

2. *A subset  $E$  of  $G$  consisting of  $\theta$ -proximal elements is said to be  $a$ -narrow, if for each  $\alpha_i \in \theta$ ,  $\rho_i(E)$  is  $a$ -narrow in  $\mathbb{P}(V_i)$ .*

### 2.22. Abels-Margulis-Soifer.

**LEMMA 2.23** (Simultaneous proximality, Lemma 5.15 [1]). *Let  $G$  be as before and  $\Gamma$  be a Zariski dense semigroup in  $G$ . Then,  $\Gamma$  contains a  $\theta_\Gamma$ -proximal element.*

The following important finiteness result of Abels-Margulis-Soifer [1] is a considerable refinement of the previous lemma. It says that in a Zariski dense semigroup  $\Gamma$  of  $G$ , for some  $r > 0$ , one can effectively generate many

$(\theta_\Gamma, r, \epsilon)$ -proximal elements. It will be of crucial use in our considerations. We also note that our Lemma 4.2 is inspired by the proof of this theorem, for which we refer the reader to the original [1] or for another treatment, to Benoist's [5], [4] or Quint's [25].

**THEOREM 2.24** (Abels-Margulis-Soifer [1]). *Let  $G$  and  $\Gamma$  be as before. Then, there exists  $0 < r = r(\Gamma)$  such that for all  $0 < \epsilon \leq r$ , there exists a finite subset  $F$  of  $\Gamma$  with the property that for every  $\gamma \in G$ , there exists  $f \in F$  such that  $\gamma f$  is  $(\theta_\Gamma, r, \epsilon)$ -proximal.*

**REMARK 2.25.** 1. *While dealing with the probability measures of uncountable support, we will use the following immediate extension of this result: there exists  $0 < r = r(\Gamma)$  such that for all  $0 < \epsilon \leq r$ , we can find a finite subset  $F$  of  $\Gamma$  and bounded neighbourhoods  $V_f$  in  $G$  of each  $f \in F$ , with the property that for each  $\gamma \in G$ , there exist a neighbourhood  $U_\gamma$  of  $\gamma$  in  $G$ , and  $f \in F$  such that for all  $f' \in V_f$  and  $\gamma' \in U_\gamma$ ,  $\gamma' f'$  is  $(\theta_\Gamma, r, \epsilon)$ -proximal. Indeed, this extension readily follows by: 1. The set of proximal elements in  $G$  is open in  $G$ . 2. The attracting direction  $x_g^+ \in \mathbb{P}(V)$  and the repulsive hyperplane  $X_g^- \subset \mathbb{P}(V)$  depend continuously on  $g \in GL(V)$ , where  $V$  is a finite dimensional vector space.*

2. *Up to enlarging  $r(\Gamma)$  given by the previous theorem, we will denote by the **same**  $r(\Gamma) > 0$ , the constant given by 1. This should not cause any confusion.*

**3. Existence of LDP.** This section is devoted to the proof of existence of LDP for the sequence  $\frac{1}{n}\kappa(Y_n)$  of random variables (i.e. existence of a rate function  $I : \mathfrak{a} \rightarrow [0, \infty]$  as in Definition 1.2). We first recall our setting and give more precise versions of Theorem 1.3 of the introduction.

3.1. *Statement of results.* Given a probability measure  $\mu$  on  $G$  (endowed with its Borel  $\sigma$ -algebra),  $Y_n$  denotes the  $n^{\text{th}}$ -step of the left  $\mu$ -random walk, i.e.  $Y_n = X_n \dots X_1$ , where the random walk increments  $X_i$ 's are  $G$ -valued independent random variables with distribution  $\mu$ , defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , henceforth fixed. Note that since the distributions of left and right random walks are the same, for the results of this article, the choice of left random walk is only a matter of convenience.

Our first Theorem 3.4 is a variant of Theorem 1.3: in this first result, we do not assume any moment condition on the probability measure  $\mu$ , in turn we have a slightly weaker conclusion. Namely, we obtain a weak LDP which we describe now (for more details see [14]).



In Definition 1.2, an LDP with a rate function  $I$  for a sequence of random variables  $Z_n$  (in our case, to be thought of as  $\frac{1}{n}\kappa(Y_n)$ ) with values in a topological space  $X$ , can be reformulated as saying

1. (Upper bound) For any closed set  $F \subset X$ ,  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in F) \leq -\inf_{x \in F} I(x)$ .
2. (Lower bound) For any open set  $O \subset X$ ,  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in O) \geq -\inf_{x \in O} I(x)$ .

The definition of a weak LDP is a slight weakening of the upper bound in the previous reformulation and it is the following:

**DEFINITION 3.2.** *A sequence of  $X$ -valued random variables  $Z_n$  is said to satisfy a weak LDP with a rate function  $I : X \rightarrow [0, \infty]$  if the upper bound 1. (above) holds for all compact sets and the lower bound 2. holds the same, for all open sets in  $X$ .*

In passing, we note the following

**REMARK 3.3.** *If  $X$  is locally compact or a polish space and a sequence of random variables  $Z_n$  on  $X$  satisfies a weak LDP with a rate function  $I$ , then  $I$  is unique.*

With this definition, our first result reads:

**THEOREM 3.4.** *Let  $k$  be a local field and let  $G$  be the group of  $k$ -points of a connected reductive algebraic group defined over  $k$ . Let  $\mu$  be a probability measure on  $G$  and suppose that its support generates a Zariski dense sub-semigroup in  $G$ . Then, the sequence of  $\mathfrak{a}^+$ -valued random variables  $\frac{1}{n}\kappa(Y_n)$  satisfies a weak LDP with a convex rate function  $I : \mathfrak{a}^+ \rightarrow [0, \infty]$ .*

The content of the next theorem is that under some moment hypotheses on  $\mu$ , one can strengthen the weak LDP of the previous theorem to a (full) LDP with a proper rate function, for which we can write an alternative expression.

Recall that a probability measure  $\mu$  on  $G$  is said to have a finite exponential moment if there exists  $c > 0$  such that  $\int e^{c\|\kappa(g)\|} \mu(dg) < \infty$ , where  $\|\cdot\|$  is an arbitrary norm on  $\mathfrak{a}$ . We shall say that  $\mu$  has a strong exponential moment, if  $\int e^{c\|\kappa(g)\|} \mu(dg) < \infty$  for all  $c > 0$ . This is clearly satisfied if  $\mu$  is of bounded support. Moreover, define the limit Laplace transform of the

sequence  $\frac{1}{n}\kappa(Y_n)$  as

$$\Lambda(\lambda) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda(\kappa(Y_n))}]$$

In these terms, we have

**THEOREM 3.5.** *Let  $G$  and  $\mu$  be as in Theorem 3.4. Suppose moreover that  $\mu$  has a finite exponential moment. Then, for the sequence  $\frac{1}{n}\kappa(Y_n)$  of random variables, a (full) LDP exists with a proper convex rate function  $I : \mathfrak{a}^+ \rightarrow [0, \infty]$ . Furthermore, if  $\mu$  has a strong exponential moment, then we can identify  $I$  with the Legendre transform of  $\Lambda$ , i.e. for all  $x \in \mathfrak{a}$ , we have  $I(x) = \sup_{\lambda \in \mathfrak{a}^*} (\lambda(x) - \Lambda(\lambda))$ .*

**REMARK 3.6.** *We observe in the previous theorems that if the support of the measure  $\mu$  instead generates a semigroup which is Zariski dense in a compact or unipotent subgroup of  $G$ , then it is still true that the LDP holds with the rate function  $I$  which takes the value 1 on  $0 \in \mathfrak{a}$  and  $\infty$  elsewhere.*

**REMARK 3.7.** *For  $g \in G$ , denote by  $\tau_g$  the automorphism conjugation by  $g$  and denote by  $\tau_{g*}\mu$  the push-forward of a probability measure  $\mu$  on  $G$  by  $\tau_g$ . Denote also by  $I_\mu$  the corresponding rate function of LDP given by Theorem 3.4. Then, for every  $g \in G$ , we have  $I_\mu = I_{\tau_{g*}\mu}$ . This also follows easily from Corollary 2.12 using the definition of  $I$  in Theorem 3.20.*

In the rest of this section, we prove the existence of weak LDP statement of Theorem 3.4. The convexity of the rate function and other assertions of Theorem 3.5 are proved in Section 4.

**3.8. Restricting the random walk to Schottky families.** The following first lemma relies on Theorem 2.24 and the uniform continuity of Cartan projections (Corollary 2.12). It says that if at some step, the Cartan projection of the walk hits a certain region of the Weyl chamber with a certain probability, then after some bounded number of steps, it will hit proximal elements whose Cartan projection is close to that region, and this will happen with a probability that is proportionally not arbitrarily small:

**LEMMA 3.9.** *Let  $0 < \epsilon < r = r(\Gamma)$ . There exist a compact set  $C = C(\Gamma, \epsilon) \subset \mathfrak{a}$ , a natural number  $i_0 = i_0(\epsilon, \Gamma, \mu)$ , and a constant  $d_1 = d_1(\epsilon, \Gamma, \mu) > 0$  such that for all  $n_0 \in \mathbb{N}$  and  $R \subset \mathfrak{a}^+$ , there exists a natural number  $n_1 \geq n_0$  with  $n_1 - n_0 \leq i_0$  such that we have*

$$\mathbb{P}(\kappa(Y_{n_1}) \in R + C \text{ and } Y_{n_1} \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal}) \geq d_1 \cdot \mathbb{P}(\kappa(Y_{n_0}) \in R)$$

PROOF. Let  $F = F(r, \epsilon)$  denote the finite subset of  $\Gamma$  given by Theorem 2.24 and  $V_f$  denote the neighbourhoods in  $G$  of elements  $f$  of  $F$  given by Remark 2.25. Fix  $i_0 \in \mathbb{N}$  such that  $F \subset \bigcup_{i=1}^{i_0} \text{supp}(\mu^{*i})$ , this is indeed possible since  $\text{supp}(\mu)$  generates  $\Gamma \supset F$ . Denote  $F = \{f_1, \dots, f_{|F|}\}$  and using Remark 2.25, define a covering of  $\Gamma$  by the subsets  $\Gamma_i := \{g \in \Gamma \mid gf'_i \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal for every } f'_i \in V_{f_i}\}$  for  $i = 1, \dots, |F|$ . Fix numbers  $k_1, \dots, k_{|F|} \leq i_0$  such that  $\mu^{*k_i}(V_{f_i}) =: \alpha_i > 0$ , where this latter inequality is strict by definition of support of a probability measure, here  $\mu^{*k_i}$ 's. Then, since,  $\Gamma_i$ 's cover  $\Gamma$ , we have

$$\mathbb{P}(\kappa(Y_{n_0}) \in R) \leq \sum_{j=1}^{|F|} \mathbb{P}(Y_{n_0} \in \Gamma_j \cap \kappa^{-1}(R))$$

so that there exists  $j_0 \in \{1, \dots, |F|\}$  such that

$$\mathbb{P}(Y_{n_0} \in \Gamma_{j_0} \cap \kappa^{-1}(R)) \geq \frac{\mathbb{P}(\kappa(Y_{n_0}) \in R)}{|F|}$$

Now, as  $|F|$  is finite and  $V_{f_i}$ 's are bounded, the set  $\bigcup_{i=1}^{|F|} \overline{V}_{f_i}$  is a compact set in  $G$ , and denote by  $C$  the compact subset  $M$  of  $\mathfrak{a}$  given by Corollary 2.12, in which we take  $L = \bigcup_{i=1}^{|F|} \overline{V}_{f_i}$ . Therefore, by this lemma, for every  $g \in \Gamma$  such that  $\kappa(g) \in R$  and for all  $f' \in \bigcup_{i=1}^{|F|} \overline{V}_{f_i}$ , we have  $\kappa(gf') \in R + C$ . Then, it follows by the independence of the random walk increments that

$$\begin{aligned} & \mathbb{P}(\kappa(Y_{n_0+k_{j_0}}) \in R + C \text{ and } Y_{n_0+k_{j_0}} \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal}) \\ & \geq \mathbb{P}(X_{n_0+k_{j_0}} \dots X_{k_{j_0}+1} \in \Gamma_{j_0} \cap \kappa^{-1}(R) \text{ and } X_{k_{j_0}} \dots X_1 \in V_{f_{j_0}}) \\ & = \mathbb{P}(Y_{n_0} \in \Gamma_{j_0} \cap \kappa^{-1}(R)) \cdot \mathbb{P}(Y_{k_{j_0}} \in V_{f_{j_0}}) \geq \frac{\mathbb{P}(\kappa(Y_{n_0}) \in R)}{|F|} \cdot \alpha_{j_0} \end{aligned}$$

Now, putting  $n_1 := n_0 + k_{j_0} \leq n_0 + i_0$  and  $\alpha_0 := \min_{k=1, \dots, |F|} \alpha_k > 0$ , we have

$$\mathbb{P}(\kappa(Y_{n_1}) \in R + C \text{ and } Y_{n_1} \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal}) \geq d_1 \mathbb{P}(\kappa(Y_{n_0}) \in R)$$

where we have put  $d_1 = \frac{\alpha_0}{|F|} = d_1(\epsilon, \mu, \Gamma)$ .  $\square$

The next lemma is an obvious observation on the relation between narrowness and  $(\theta, r, \epsilon)$ -Schottky properties of a set of proximal elements. It will prove to be useful in our considerations together with the lemma following it. In its proof and in what follows, recall that  $\Pi$  stands for the set of simple roots  $\alpha_1, \dots, \alpha_{d_S}$  of  $G$  and for each  $\alpha_i \in \Pi$ ,  $(\rho_i, V_i)$  is the corresponding distinguished representation of  $G$  (given by Lemma 2.10).

LEMMA 3.10. *Let  $\epsilon$  and  $r$  be two real numbers such that  $0 < 6\epsilon \leq r$  and let  $\theta$  be a non-empty subset of  $\Pi$ . Then, an  $r$ -narrow set  $E$  of  $(\theta, r, \epsilon)$ -proximal elements in  $G$  is a  $(\theta, r_1, \epsilon)$ -Schottky family, where we can take  $r_1 = \frac{r}{6}$ .*

PROOF. Observe first that, by definition, if  $\gamma$  is  $(\theta, r, \epsilon)$ -proximal, then  $\gamma$  is also  $(\theta, r_1, \epsilon_1)$ -proximal for all  $r_1 \leq r$  and  $\epsilon_1 \geq \epsilon$  such that  $r_1 \geq \epsilon_1$ . Therefore, to prove the lemma, one just notes that for all  $\gamma, \gamma' \in E$  and  $\alpha_i \in \theta$ , since  $d(x_{\rho_i(\gamma)}^+, X_{\rho_i(\gamma)}^{\leq}) \geq 2r$  and  $d(x_{\rho_i(\gamma)}^+, x_{\rho_i(\gamma')}^+) < r$ , we have  $d(x_{\rho_i(\gamma)}^+, X_{\rho_i(\gamma')}^{\leq}) \geq 2r - r = r$ . Hence putting  $r_1 = \frac{r}{6}$  we have by hypothesis,  $r_1 \geq \epsilon$  and  $d(x_{\rho_i(\gamma)}^+, X_{\rho_i(\gamma')}^{\leq}) \geq 6r_1$  as in the definition of a  $(\theta, r_1, \epsilon)$ -Schotky family.  $\square$

We shall now proceed with the following lemma, which is a consequence of the compactness of projective spaces of  $V_i$ 's. We will put it to good use on two occasions; once, together with Lemma 3.10 to obtain a useful corollary, and once in the proof of convexity.

LEMMA 3.11. *Let  $r \geq \epsilon > 0$  and a positive constant  $a$  be given. Let  $\theta$  be a non-empty subset of  $\Pi$ . Then, there exists a strictly positive constant  $d_2 = d_2(a)$  such that for every subset  $E$  of  $G$  consisting of  $(\theta, r, \epsilon)$ -proximal elements, and for all  $n \in \mathbb{N}$ , there exists an  $a$ -narrow subset  $E_n$  of  $E$  such that, we have  $\mathbb{P}(Y_n \in E_n) \geq d_2 \mathbb{P}(Y_n \in E)$ .*

PROOF. Indeed, for each  $\alpha_i \in \theta$ , by compactness of  $\mathbb{P}(V_i)$ , we can choose a partition  $Y_1^i, \dots, Y_{s_i}^i$  of  $\mathbb{P}(V_i)$  with  $\text{diam}(Y_j^i) < a$  and where  $s_i = s_i(a)$ . Similarly, we can find hyperplanes  $H_1^i, \dots, H_{t_i}^i$  in  $V_i$  with  $t_i = t_i(a)$ , and with the property that - denoting by  $Z_j^i$  the  $a$ -neighbourhood of  $H_j^i$  in  $\mathbb{P}(V_i)$  - the projection  $\mathbb{P}(H)$  of any given hyperplane  $H$  of  $V_i$  is contained in one of  $Z_j^i$ 's. Up to re-indexing  $\alpha_i$ 's, write  $\theta = \{\alpha_1, \dots, \alpha_c\}$  for some integer  $1 \leq c \leq d_S$ . Let  $\underline{i}, \underline{j}$  denote multi-indices of the form  $\underline{i} = (i_1, \dots, i_c)$  and  $\underline{j} = (j_1, \dots, j_c)$  where, for each  $k = 1, \dots, c$ ,  $i_k \in \{1, \dots, s_k\}$  and  $j_k \in \{1, \dots, t_k\}$ . Now, let  $E \subset \Gamma$  be given as in the statement and for multi-indices  $\underline{i}, \underline{j}$ , denote by  $E_{\underline{i}}^{\underline{j}}$  the following subset of  $E$ :

$$E_{\underline{i}}^{\underline{j}} := \{ \gamma \in E \mid x_{\rho_k(\gamma)}^+ \in Y_{i_k}^{k} \text{ and } X_{\rho_k(\gamma)}^{\leq} \subset Z_{j_k}^k \}$$

By the choice of  $Y_j^i$ 's and  $Z_j^i$ 's, the family  $E_{\underline{i}}^{\underline{j}}$  covers  $E$  and we thus have for every  $n \in \mathbb{N}$

$$\mathbb{P}(Y_n \in E) \leq \sum_{\underline{i}, \underline{j}} \mathbb{P}(Y_n \in E_{\underline{i}}^{\underline{j}})$$

It follows that for every  $n \in \mathbb{N}$ , there exist at least two multi-indices  $\underline{j}_0$  and  $\underline{j}'_0$  such that  $\mathbb{P}(Y_n \in E_{\underline{j}_0}^{\underline{j}'_0}) \geq \frac{\mathbb{P}(Y_n \in E)}{s_1 \dots s_{c t_1} \dots t_c}$ . Hence, putting  $d_2 = d_2(a) = \frac{1}{s_1 \dots s_{c t_1} \dots t_c}$  and  $E_n = E_{\underline{j}_0}^{\underline{j}'_0}$ , we have the result of the lemma.  $\square$

**COROLLARY 3.12.** *Let  $r$  and  $\epsilon$  be two real numbers with  $r \geq 6\epsilon > 0$  and let  $\theta \subseteq \Pi$ . Then, there exists a constant  $d_3 = d_3(r) > 0$  such that for every subset  $E$  of  $G$  consisting of  $(\theta, r, \epsilon)$ -proximal elements and for all  $n \in \mathbb{N}$ , there exists an  $(\theta, r_1, \epsilon)$ -Schottky family  $E_n \subset E$  with  $r_1 \geq \frac{r}{6} \geq \epsilon$  and such that  $\mathbb{P}(Y_n \in E_n) \geq d_3 \cdot \mathbb{P}(Y_n \in E)$ .*

**PROOF.** If  $\theta = \emptyset$  the statement is trivial; if not, choose  $a = r$  in Lemma 3.11 and apply Lemma 3.10.  $\square$

**3.13. Cartan projections of powers of Schottky families.** The next proposition says that the images in  $\mathfrak{a}^+$  of the Cartan projections of the  $n^{\text{th}}$ -power of an  $(\theta, r, \epsilon)$ -Schottky family in  $\Gamma$  is contained, up to compact perturbation, in the  $n$ -dilation of the images in  $\mathfrak{a}^+$  of the Cartan projections of that family. It follows from Benoist's Lemma 2.16, Proposition 2.19 and Theorem 2.20.

**PROPOSITION 3.14.** *There exists a compact subset  $K$  of  $\mathfrak{a}$ , depending on  $r, \epsilon$  and  $\Gamma$ , with the property that for every  $(\theta_\Gamma, r, \epsilon)$ -Schottky family  $E$  in  $\Gamma$  and  $n \in \mathbb{N}$ , we have  $\kappa(E^n) \subset n \cdot (\text{co}(\kappa(E)) + K)$ , where  $E^n := \{\gamma_1 \dots \gamma_n \mid \gamma_i \in E\}$ ,  $\kappa(E) := \{\kappa(\gamma) \mid \gamma \in E\}$ ,  $\kappa(E) + K := \{x + k \mid x \in \kappa(E), k \in K\}$  and  $\text{co}(\cdot)$  stands for the convex hull.*

**PROOF.** We first note that the statement is clear if  $\theta_\Gamma = \emptyset$ . Indeed, in this case, by Lemma 2.16, for each  $i = 1, \dots, d_S$ ,  $\bar{\alpha}_i(\kappa(\Gamma))$  is bounded. On the other hand, for all central weight  $\chi \in X_C$  and  $g, h \in G$ , we have  $\bar{\chi}(\kappa(gh)) = \bar{\chi}(\kappa(g)) + \bar{\chi}(\kappa(h))$  and the statement follows since  $\bar{\Pi} \cup \bar{X}_C$  is a basis of  $\mathfrak{a}^*$ .

Now suppose that  $\theta_\Gamma \neq \emptyset$  and let  $g_1, \dots, g_n \in E$ . It follows by definition of a  $(\theta_\Gamma, r, \epsilon)$ -Schottky family and Theorem 2.20 that for every  $n \geq 1$ , the product  $g_1, \dots, g_n$  is  $(\theta_\Gamma, 2r, 2\epsilon)$ -proximal. Now, let  $N = N_{(r, \epsilon)}$  be the compact subset of  $\mathfrak{a}$  given by Theorem 2.20 and let  $M = M_{(2r, 2\epsilon)}$  be the compact subset of  $\mathfrak{a}$  given by Proposition 2.19. Rewrite the difference  $\kappa(g_1, \dots, g_n) - \sum_{i=1}^n \kappa(g_i)$  as

$$(\kappa(g_1, \dots, g_n) - \lambda(g_1, \dots, g_n)) + (\lambda(g_1, \dots, g_n) - \sum_{i=1}^n \lambda(g_i)) + (\sum_{i=1}^n (\lambda(g_i) - \kappa(g_i)))$$

In this expression, observe that the first term belongs to  $M$  by Proposition 2.19 and the above remark, the second term belongs to  $n \cdot N$  by Theorem

2.20, and the third term belongs to  $n.co(M)$  by Proposition 2.19. Now the statement of our proposition easily follows: denote by  $M^{-1}$  the set  $\{-x \mid x \in M\}$  and put  $\tilde{M} = co(M^{-1} \cup M)$ . Finally set  $K = 2.\tilde{M} + N$  and observe that by above, we have  $\kappa(g_1, \dots, g_n) \in \sum_{i=1}^n \kappa(g_i) + n.K$  proving the statement.  $\square$

3.15. *Controlling deviations in bounded steps.* For later convenient use, we single out the following topological notion and note two obvious facts about it in the following lemma.

DEFINITION 3.16. *Let  $X$  be a topological space and  $O_1 \subset O_2$  two open subsets of  $X$ . We say that  $O_1$  is super-strictly contained in  $O_2$  if  $\overline{O_1} \subseteq O_2$ .*

LEMMA 3.17. 1. *Let  $V$  be a finite dimensional real normed vector space and  $O_1$  and  $O_2$  two open bounded subsets of  $V$ ,  $O_1$  super-strictly contained in  $O_2$ . Then, for all bounded set  $K \subset V$ , there exists a constant  $R(O_1, O_2, K) \in \mathbb{R}^+$  such that for all  $Q \geq Q(O_1, O_2, K)$ , we have  $Q.O_1 + K \subset Q.O_2$*

2. *Let  $O_1$  and  $O_2$  be as above. Then, there exists a real number  $q(O_1, O_2) < 1$  such that for all  $n_1, n_2 \in \mathbb{N}$  with  $1 \geq \frac{n_1}{n_2} > q(O_1, O_2)$ , we have  $n_1 O_1 \subset n_2 O_2$ .*

PROOF. Both statements are obvious. Remark that the hypothesis implies that  $d(O_1, O_2^c) > 0$  and one can take  $Q(O_1, O_2, K)$  and  $1 > q(O_1, O_2)$  any real numbers larger than respectively  $\frac{\text{diam}(K)}{d(O_1, O_2^c)}$  and  $1 - \frac{d(O_1, O_2^c)}{\sup_{x \in O_1} \|x\|}$ .  $\square$

We shall need one last lemma before proceeding to prove the theorem. It relies on the uniform continuity of the Cartan projections (Corollary 2.12) and says that if the averages of the Cartan projections of the random product hits a certain region of the Cartan subalgebra at periodic times, then it will hit any open neighbourhood of this region at any time with at least the same asymptotic exponential rate of probability:

LEMMA 3.18. *Let  $O_1$  and  $O_2$  be two open bounded convex subsets of  $\mathfrak{a}^+$ ,  $O_1$  super-strictly contained in  $O_2$ . Suppose that there exist  $n_0 \in \mathbb{N}$  and  $\alpha \geq 0$  such that for all  $k \geq 1$ , we have  $\mathbb{P}(\kappa(Y_{n_0 k}) \in k n_0 O_1) \geq e^{-n_0 k \alpha}$ . Then we have  $\liminf_n \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \kappa(Y_n) \in O_2) \geq -\alpha$ .*

PROOF. For all  $n \in \mathbb{N}$ , let  $k_n \in \mathbb{N}$  be defined by  $n_0(k_n + 1) > n \geq n_0 k_n$ . By  $\sigma$ -compactness, we can choose a compact subset  $L_{n_0}$  of  $G$  containing  $e \in G$  and such that  $\mu^{*i}(L_{n_0}) \geq \frac{1}{2}$  for each  $i = 1, \dots, n_0$ . Let  $M_{n_0}$  be the compact subset  $M$  of  $\mathfrak{a}$  given by Corollary 2.12, by taking in it  $L = L_{n_0}$ .

By definition of super-strict inclusion and the fact that the ambient space is a normed real vector space, we can pick  $O_{12}$  such that each of the inclusions  $O_1 \subset O_{12} \subset O_2$  is super-strict. Now, let  $Q_{n_0} := Q(O_{12}, O_2, M_{n_0}) \in \mathbb{R}$  and  $q := q(O_1, O_{12}) < 1$  where these last quantities are as defined in Lemma 3.17. Then, for all  $n \in \mathbb{N}$  such that  $n \geq Q_{n_0}$  and  $1 - \frac{n_0}{n} > q$ , we have the following sequence of inclusions of events:

$$\{\kappa(Y_n) \in k_n n_0 O_1 + M_{n_0}\} \subset \{\kappa(Y_n) \in n O_{12} + M_{n_0}\} \subset \{\kappa(Y_n) \in n O_2\}$$

where the first inclusion is by 2. and the second by 1. of Lemma 3.17.

As a result, by independence of random walk increments, for all  $n \in \mathbb{N}$ , we have

$$(3.1) \quad \begin{aligned} \mathbb{P}\left(\frac{1}{n}\kappa(Y_n) \in O_2\right) &\geq \mathbb{P}(\kappa(Y_{k_n n_0 + (n - k_n n_0)}) \in k_n n_0 O_1 + M_{n_0}) \geq \\ &\mathbb{P}(\kappa(Y_{k_n n_0}) \in k_n n_0 O_1) \cdot \mathbb{P}(Y_{n - k_n n_0} \in L_{n_0}) \geq e^{-n_0 k_n \alpha} \frac{1}{2} \end{aligned}$$

where the last inequality follows by hypothesis and the construction of  $L_{n_0}$ . Now, in (3.1), taking logarithm, dividing by  $n$ , and taking  $n$  to infinity, we obtain the result of the lemma.  $\square$

3.19. *Proof of existence of weak LDP.* We are now ready to prove the existence of weak LDP statement in Theorem 3.4 by using the following general fact:

**THEOREM 3.20** (see Theorem 4.1.11 in [14]). *Let  $X$  be a topological space endowed with its Borel  $\sigma$ -algebra  $\beta_X$ , and  $Z_n$  be a sequence of  $X$ -valued random variables. Denote by  $\mu_n$  the distribution of  $Z_n$ . Let  $\mathcal{A}$  be a base of open sets for the topology of  $X$ . For each  $x \in X$ , define:*

$$I_{li}(x) := \sup_{\substack{A \in \mathcal{A} \\ x \in A}} -\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A) \quad \text{and} \quad I_{ls}(x) := \sup_{\substack{A \in \mathcal{A} \\ x \in A}} -\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A)$$

*Suppose that for all  $x \in X$ , we have  $I_{li}(x) = I_{ls}(x)$ . Then, the sequence  $Z_n$  satisfies an LDP with rate function  $I$  given by  $I(x) := I_{li}(x) = I_{ls}(x)$ .*

**REMARK 3.21.** *In a polish space  $X$ , the hypothesis of the previous theorem is actually equivalent to the existence of a weak LDP (see [14]).*

We note that below if  $\theta_\Gamma = \emptyset$ , the proof simplifies to a great extent and the main relevant part is at the end where we make use of Proposition 3.14.

PROOF OF THEOREM 3.4, (EXISTENCE OF LDP). For all  $n \geq 1$ , denote by  $\mu_n$  the distribution of the random variable  $\frac{1}{n}\kappa(Y_n)$ . It is a probability measure supported on the closed subset  $\mathfrak{a}^+$  of the vector space  $\mathfrak{a}$ . To establish the weak LDP for this sequence of probability measures, we use Theorem 3.20 and argue by contradiction.

Let  $I_{li}$  and  $I_{ls}$  denote the functions on  $\mathfrak{a}$ , associated to the sequence  $\mu_n$  as in Theorem 3.20, where we take the norm-open balls in  $\mathfrak{a}$  as a base of topology. Suppose now for a contradiction that there exists  $x \in \mathfrak{a}$  such that  $I_{li}(x) > I_{ls}(x) \geq 0$ . We can suppose that  $x$  is in the closed Weyl chamber  $\mathfrak{a}^+$  since for all  $n \in \mathbb{N}$ ,  $\text{supp}(\mu_n) \subset \mathfrak{a}^+$ .

By definitions of the functions  $I_{li}$  and  $I_{ls}$ , this implies that there exists an open ball  $O_5 \subset \mathfrak{a}$  with  $x \in O_5$  and such that

$$(3.2) \quad -\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_5) > \sup_{\substack{O \subset \mathfrak{a} \\ x \in O}} -\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O) + 4\eta$$

for some  $\eta > 0$  small enough.

We then choose  $x \in O_1 \subset O_2 \subset O_3 \subset O_4 \subset O_5$  open balls around  $x$ , where each inclusion is super-strict, such that (3.2) yields

$$-\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_5) > -\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_1) + 3\eta$$

Now, let  $r = r(\Gamma)$  be given by Theorem 2.24 and choose  $\epsilon \leq \frac{r}{6}$ . Let  $d_1 = d_1(r, \epsilon, \Gamma)$  and  $i_0 = i_0(\epsilon, \Gamma, \mu)$  be the constants given by Lemma 3.9,  $C = C(\Gamma, \epsilon)$  be the compact subset of  $\mathfrak{a}$  also given by Lemma 3.9,  $d_3 = d_3(r)$  be the constant given by Corollary 3.12,  $K = K(r, \epsilon)$  be the compact subset of  $\mathfrak{a}$  given by Proposition 3.14. Let us also fix a real number  $Q \geq \max_{i < j} (Q(O_i, O_j, C) \vee Q(O_i, O_j, K))$  where these latter quantities are as defined in Lemma 3.17 and let  $q := q(O_1, O_5)$  where again this is defined as in Lemma 3.17. Choose  $n_0 \in \mathbb{N}$  such that

- (i)  $-\frac{1}{n_0} \log \mu_{n_0}(O_1) + 2\eta < -\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_5)$
- (ii)  $e^{-n_0\eta} \leq d_1 d_3$
- (iii)  $n_0 \geq Q$
- (iv)  $\frac{n_0}{n_0 + i_0} > q$

Put  $\alpha := -\frac{1}{n_0} \log \mu_{n_0}(O_1)$  and  $\beta := -\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_5)$  so that by item (i) in the choice of  $n_0$ ,

$$(3.3) \quad \alpha + 2\eta < \beta$$



Setting  $R = n_0 O_1$  in Lemma 3.9, we obtain that for some  $n_1$  such that  $n_1 - n_0 \leq i_0$

$$(3.4) \quad \mathbb{P}(\kappa(Y_{n_1}) \in n_0 O_1 + C \text{ and } Y_{n_1} \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal}) \geq e^{-n_1 \alpha} \cdot d_1$$

The choice of  $n_0$  (respectively items (iii) and (iv) above) implies by Lemma 3.17 that  $n_0 O_1 + C \subset n_0 O_2$  and  $n_0 O_2 \subset n_1 O_3$  so that (3.4) becomes

$$(3.5) \quad \mathbb{P}(\kappa(Y_{n_1}) \in n_1 O_3 \text{ and } Y_{n_1} \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal}) \geq e^{-n_1 \alpha} \cdot d_1$$

Applying Corollary 3.12 by taking  $L = \kappa^{-1}(n_1 O_3) \cap \Gamma_{(r, \epsilon)}$ , which is non-empty by (3.5), and where  $\Gamma_{(r, \epsilon)}$  is the set of  $(\theta_\Gamma, r, \epsilon)$ -proximal elements in  $\Gamma$ , using also (3.5), we obtain that there exists an  $(\theta_\Gamma, r_1, \epsilon)$ -Schottky family  $E \subset L \subset \Gamma$  such that we have

$$\mathbb{P}(\kappa(Y_{n_1}) \in n_1 O_3 \text{ and } Y_{n_1} \in E) \geq e^{-n_1 \alpha} d_1 d_3 \geq e^{-n_1(\alpha + \eta)}$$

where the last inequality follows by item (ii) of the choice of  $n_0$  and since  $n_1 \geq n_0$ .

Next, observe that by the construction of  $L$  and since  $E \subset L$ , we have  $\kappa(E) \subset n_1 O_3$  and therefore, as  $O_3$  is convex,  $\text{co}(\kappa(E)) \subset n_1 O_3$ . Then, by Proposition 3.14, we obtain that for each  $k \geq 1$ ,  $\kappa(E^k) \subset k \cdot (\text{co}(\kappa(E)) + K) \subset k \cdot (n_1 O_3 + K) \subset k n_1 O_4$  where the last inclusion follows also from item (iii) of the choice of  $n_0$  and since  $n_1 \geq n_0$ .

Finally, for all  $k \geq 1$ , by the independence of the random walk increments, we have that  $\mathbb{P}(Y_{n_1 k} \in E^k) \geq \mathbb{P}(Y_{n_1} \in E)^k$  and thus we obtain

$$\mathbb{P}(\kappa(Y_{n_1 k}) \in k n_1 O_4) \geq \mathbb{P}(Y_{n_1 k} \in E^k) \geq \mathbb{P}(Y_{n_1} \in E)^k \geq e^{-n_1 k(\alpha + \eta)}$$

Therefore, Lemma 3.18 establishes that  $\beta = -\liminf \frac{1}{n} \log \mathbb{P}(\kappa(Y_n) \in O_5) \leq \alpha + \eta$  which together with (3.3) yields  $\alpha + 2\eta < \beta \leq \alpha + \eta$ , a contradiction.  $\square$

#### 4. Convexity of the rate function.

4.1. *A dispersion lemma.* Our first lemma in this section is a key dispersion result which is in fact a corollary of the proof of Theorem 2.24 in Abels-Margulis-Soifer's [1]. Namely, it says that, by the Zariski density of  $\Gamma$  in  $G$  and connectedness of  $G$ , one can find finite sets in  $\Gamma$  such that for each point of the projective spaces of the distinguished representation spaces  $V_i$ 's, some elements of these finite sets of  $\Gamma$  will, by their action, disperse that point in the projective spaces. It will be useful on several occasions, particularly by its relation to the 1. (b) of Definition 2.7.

LEMMA 4.2 (Dispersion lemma). *For all  $t \in \mathbb{N}$ , there exist a strictly positive constant  $\eta_t = \eta(t, \Gamma)$ , depending only on  $t$  and  $\Gamma$ , and a finite set  $M_t \subset \Gamma$  with the following properties: for every  $\bar{x} = (x_1, \dots, x_{d_S}) \in \prod_{i=1}^{d_S} \mathbb{P}(V_i)$ , where  $V_i$ 's are the distinguished representation spaces of  $G$ , there exist  $\gamma_1, \dots, \gamma_t \in M_t$  such that*

(i) *For each  $i = 1, \dots, d_S$  and for all  $j \neq k \in \{1, \dots, t\}$ ,*

$$d_i(\rho_i(\gamma_j).B_i(x_i, \eta_t), \rho_i(\gamma_k).B_i(x_i, \eta_t)) > \eta_t$$

(ii) *For all  $i = 1, \dots, d_S$  and for every subset  $\{\gamma_{i_1}, \dots, \gamma_{i_k}\}$  of  $\{\gamma_1, \dots, \gamma_t\}$  of cardinality less than  $k \leq \dim V_i$ , for all  $y_i^1, \dots, y_i^k, z_i \in B_i(x_i, \eta_t)$ , denoting by  $\langle \rho_i(\gamma_{i_1})y_i^1, \dots, \rho_i(\gamma_{i_k})y_i^k \rangle$  the projective image of the subspace generated by these lines, and for all  $j \notin \{i_1, \dots, i_k\}$ , we have,*

$$d_i(\langle \rho_i(\gamma_{i_1})y_i^1, \dots, \rho_i(\gamma_{i_k})y_i^k \rangle, \rho_i(\gamma_j)z_i) > \eta_t$$

PROOF. We start by inductively finding elements  $\gamma_1^{\bar{x}}, \dots, \gamma_t^{\bar{x}} \in \Gamma$  for each element  $\bar{x} = (x_1, \dots, x_{d_S})$  of  $\prod_{i=1}^{d_S} \mathbb{P}(V_i)$ : choose  $\gamma_1^{\bar{x}} \in \Gamma$  arbitrarily. Having constructed  $\gamma_1^{\bar{x}}, \dots, \gamma_k^{\bar{x}}$  for some  $k < t$ , put

$$G_{i,k+1} := \{\gamma \in G \mid \rho_i(\gamma).x_i \text{ does not belong to the proper subspaces of } V_i \text{ generated by the lines } \rho_i(\gamma_j).x_i \text{ for } j \in \{1, \dots, k\}\}$$

Since there are finitely many such proper spaces of  $V_i$ , and the condition of not belonging to a proper subspace is a Zariski open condition in  $G$ ,  $G_{i,k+1}$  is a finite intersection of Zariski open sets which are also non-empty since the distinguished representations,  $\rho_i$ 's are irreducible. Consequently,  $G_{i,k+1}$  is a non-empty Zariski open set in  $G$ . Similarly, the set  $G_{k+1}$  defined by  $G_{k+1} := \bigcap_{i=1}^{d_S} G_{i,k+1}$  is Zariski open.  $\Gamma$  being, by assumption, Zariski dense in  $G$ , the intersection  $G_{k+1} \cap \Gamma$  is non-empty; choose one element  $\gamma_{k+1}^{\bar{x}} \in G_{k+1} \cap \Gamma$ .

By induction, we then have constructed  $\gamma_1^{\bar{x}}, \dots, \gamma_t^{\bar{x}} \in \Gamma$  for each  $\bar{x} \in \prod \mathbb{P}(V_i)$  such that for each  $i = 1, \dots, d_S$ , the elements of  $\{\rho_i(\gamma_1^{\bar{x}}).x_i, \dots, \rho_i(\gamma_t^{\bar{x}}).x_i\}$  are in general position. Now choose  $\eta_t^{\bar{x}} > 0$ , such that

$$d_i(\langle \rho_i(\gamma_{i_1}^{\bar{x}}).x_i, \dots, \rho_i(\gamma_{i_k}^{\bar{x}}).x_i \rangle, \rho_i(\gamma_j^{\bar{x}}).x_i) > 2\eta_t^{\bar{x}}$$

for all  $i = 1, \dots, d_S$ ,  $k \leq \dim V_i - 1$ ,  $i_1, \dots, i_k \in \{1, \dots, t\}$  and  $j \notin \{i_1, \dots, i_k\}$ . Such an  $\eta_t^{\bar{x}} > 0$  indeed exists by our construction of the  $\gamma_i^{\bar{x}}$ 's.

Now, by continuity of the action of  $G$  on  $\mathbb{P}(V_i)$ 's, for all  $\bar{x} = (x_1, \dots, x_{d_S}) \in \prod \mathbb{P}(V_i)$ , there exists a neighbourhood  $W^{\bar{x}} = W_{x_1}^{\bar{x}} \times \dots \times W_{x_{d_S}}^{\bar{x}} \subset \prod \mathbb{P}(V_i)$

such that for all  $i = 1, \dots, d_S$ , for all  $k \leq \dim V_i - 1$ , and for all  $(y_1^i, \dots, y_k^i) \in W_i^{\bar{x}}$ ,  $z^i \in W_i^{\bar{x}}$  and  $\gamma_i$ 's as above; we have

$$(4.1) \quad d_i(\langle \rho_i(\gamma_{i_1}^{\bar{x}}) \cdot y_1^i, \dots, \rho_i(\gamma_{i_k}^{\bar{x}}) \cdot y_k^i \rangle, \rho_i(\gamma_j^{\bar{x}}) \cdot z^i) > \eta_t^{\bar{x}}$$

Up to reducing  $\eta_t^{\bar{x}}$ , we can suppose that for each  $i = 1, \dots, d_S$ ;  $B_i(x_i, 2\eta_t^{\bar{x}}) \subset W_i$ . Now, cover the compact set  $\prod \mathbb{P}(V_i)$  by the open sets  $\bigcup_{\bar{x} \in \prod \mathbb{P}(V_i)} \prod_{i=1}^{d_S} B_i(x_i, \eta_t^{\bar{x}})$

and extract a finite subcover. Let us call the elements  $\bar{x}^1, \dots, \bar{x}^n \in \prod \mathbb{P}(V_i)$  such that  $(\prod_{i=1}^{d_S} B_i(x_i^j, \eta_t^{\bar{x}^j}))_{j=1, \dots, n}$  is the extracted finite subcover, and put  $\eta_t := \min_{j=1, \dots, n} \eta_t^{\bar{x}^j}$  and  $M_t := \bigcup_{j=1}^n \{\gamma_1^{\bar{x}^j}, \dots, \gamma_t^{\bar{x}^j}\}$ .

Then, the result of the lemma readily follows: as in the assertion of the lemma, let  $\bar{x} = (x_1, \dots, x_{d_S}) \in \prod \mathbb{P}(V_i)$ . Let also, up to re-indexing,  $\bar{x}^1$  be such that for each  $i = 1, \dots, d_S$ ;  $d_i(x_i, x_i^1) < \eta_t^{\bar{x}^1}$  and take  $\gamma_1^{\bar{x}^1}, \dots, \gamma_t^{\bar{x}^1} \in M_t$ . Then,

- (i) To see the first statement, fix  $i \in \{1, \dots, d_S\}$  and  $j \neq k \in \{1, \dots, t\}$ , and consider  $y_i, z_i \in B_i(x_i, \eta_t)$ . Since  $d_i(x_i, x_i^1) < \eta_t^{\bar{x}^1}$ ,  $\eta_t \leq \eta_t^{\bar{x}^1}$  and  $B_i(x_i^1, 2\eta_t^{\bar{x}^1}) \subset W_i^{\bar{x}^1}$ , we have  $B_i(x_i, \eta_t) \subset B_i(x_i^1, 2\eta_t^{\bar{x}^1}) \subset W_i^{\bar{x}^1}$ , so that by (4.1)  $d_i(\rho_i(\gamma_j) \cdot y_i, \rho_i(\gamma_k) \cdot z_i) > \eta_t^{\bar{x}^1} \geq \eta_t$ , establishing the claim.
- (ii) The proof of the second statement is similar. Fix  $i \in \{1, \dots, d_S\}$  and  $i_1, \dots, i_k, j \in \{1, \dots, t\}$  with  $j \notin \{i_1, \dots, i_k\}$  and set  $k = \dim V_i - 1$ . For all  $y_{i_1}, \dots, y_{i_k}, z_i \in B_i(x_i, \eta_t)$ , exactly as above, we have  $y_{i_1}, \dots, y_{i_k}, z_i \in B_i(x_i, \eta_t) \subset W_i^{\bar{x}^1}$  so that (4.1) again proves the claim.

□

REMARK 4.3. *A similar observation as Remark 2.25 of the Abels-Margulis-Soifer finiteness result, clearly applies to this finiteness result as well. Namely, for all  $t \in \mathbb{N}$ , there exists a constant  $\eta_t \in \Gamma$ , a finite subset  $M_t$  of  $\Gamma$  and for each  $\gamma \in M_t$ , bounded neighbourhoods  $V_\gamma$  of  $\gamma$  in  $G$  such that we have the conclusions of the lemma for every  $\gamma'_i \in V_{\gamma_i}$ , instead of only  $\gamma_i$ 's for  $i = 1, \dots, d_S$ . We shall use the **same constants**  $\eta_t$  for this extended result and Lemma 4.2.*

#### 4.4. Dealing with two Schottky families.

LEMMA 4.5. *Let  $V$  be a finite dimensional  $k$ -vector space and  $g \in GL(V)$ . For the action of  $GL(V)$  on  $\mathbb{P}(V)$  (endowed with the Fubini-Study metric),  $g$  is a  $\|\Lambda^2 g\| \cdot \|g^{-1}\|^2$ -Lipschitz transformation.*

PROOF. Indeed, for  $x, y \in \mathbb{P}(V)$ , we have

$$d(gx, gy) = \frac{\|gx \wedge gy\|}{\|gx\| \cdot \|gy\|} \leq \frac{\|\Lambda^2 g\| \cdot \|x \wedge y\|}{\|g^{-1}\|^{-2} \cdot \|x\| \cdot \|y\|} = \|\Lambda^2 g\| \cdot \|g^{-1}\|^2 d(x, y)$$

□

Accordingly, for an element  $\gamma \in G$ , put

$$(4.2) \quad L(\gamma) := \max_{i=1, \dots, d} \|\Lambda^2 \rho_i(\gamma)\| \cdot \|\rho_i(\gamma)^{-1}\|^2 \in [1, \infty[.$$

The next technical lemma is based on the observation that if a proximal element  $g$ , when multiplied on the left by an arbitrary element  $\gamma$ , gives a proximal element  $\gamma g$ , then the projective hyperplane  $X_{\gamma g}^<$  is close to that of  $g$ , while the attracting directions  $x_{\gamma g}^+$  and  $x_g^+$  may differ arbitrarily. The rest of the proof is along the same lines as the so called Tits proximality criterion (See [27] 3.8, [1] 2.1, [4] Lemme 6.2).

LEMMA 4.6. *Let  $g$  be a  $(\theta, r, \epsilon)$ -proximal element of  $G$  and  $\gamma \in G$  such that  $L(\gamma) \cdot \epsilon < 1$ . Put  $1 > \epsilon_1 := L(\gamma)\epsilon \geq \epsilon$  and suppose there exists a  $\delta$  with  $\delta > 6\epsilon_1$  such that for each  $\alpha_i \in \theta$ , we have  $d_i(\rho_i(\gamma)x_{\rho_i(g)}^+, X_{\rho_i(g)}^<) > \delta$ . Then,  $\gamma g$  is  $(\theta, \frac{\delta}{3}, 2\epsilon_1)$ -proximal. Moreover, for each  $\alpha_i \in \theta$ , we have  $d(x_{\rho_i(\gamma g)}^+, \gamma x_{\rho_i(g)}^+) < \epsilon_1$  and  $d_H(X_{\gamma g}^<, X_g^<) < \epsilon$ .*

PROOF. To ease the notation, we will dismiss the representations  $\rho_i$ . By our definition of  $L(\cdot)$  in (4.2), our reasonings apply simultaneously to each representation  $\rho_i$  such that  $\alpha_i \in \theta$ .

We first establish that  $\gamma g$  is proximal. One first observes that we have

$$(4.3) \quad \gamma g B_g^\epsilon \subseteq \gamma b_g^\epsilon \subseteq B(\gamma x_g^+, \epsilon L(\gamma)) \subseteq B_g^{4\epsilon_1}$$

where the first inclusions is by  $(r, \epsilon)$ -proximality of  $g$  and the last by our hypothesis that  $d(\gamma x_g^+, X_g^<) > \delta \geq 6\epsilon_1$ .

Moreover, the restriction of the action of  $\gamma g$  on  $B_g^\epsilon$  is  $L(\gamma)\epsilon = \epsilon_1$  Lipschitz with, by hypothesis,  $\epsilon_1 < 1$ . Therefore,  $\gamma g$  is a continuous contraction of the compact  $B_g^\epsilon$  into  $B_g^{4\epsilon_1} \subseteq \text{int}(B_g^\epsilon)$  and thus, by Banach fixed point theorem, has a unique attracting fixed point, of basin of attraction containing  $B_g^\epsilon$ . This indeed implies that  $\gamma g$  is proximal. One also sees from (4.3) that we must have  $x_{\gamma g}^+ \in B(\gamma x_g^+, \epsilon_1)$  and  $d_H(X_{\gamma g}^<, X_g^<) < \epsilon$ .

To get the complete statement of the lemma, in view of the definition of a  $(\theta, \frac{\delta}{3}, 2\epsilon_1)$ -proximal element, one checks that

- (i) Since by above  $x_{\gamma g}^+ \in B(\gamma x_g^+, \epsilon_1)$  and  $d_H(X_{\gamma g}^<, X_g^<) < \epsilon$ , and by hypothesis  $d(\gamma x_g^+, X_g^<) > \delta \geq 6\epsilon_1$ , we have  $d(x_{\gamma g}^+, X_{\gamma g}^<) \geq \delta - \epsilon - \epsilon_1 \geq \delta - 2\epsilon_1 > 2\frac{\delta}{3}$ .
- (ii) Similarly, we have  $\gamma g B_{\gamma g}^{2\epsilon_1} \subseteq \gamma g B_g^\epsilon \subseteq B(\gamma x_g^+, \epsilon_1) \subseteq b_{\gamma g}^{2\epsilon_1}$ .

- (iii) Finally, the restriction of the action of  $\gamma g$  on  $B_{\gamma g}^{2\epsilon_1} \subseteq B(g)^\epsilon$  is  $\epsilon_1 = \epsilon L(\gamma)$  Lipschitz, as observed above.

These establish our claim. □

In the next proposition, we exploit more deeply the observation mentioned before the last lemma, in its relation with the result of Lemma 4.2 and the notion of narrowness of a set of proximal elements. It says that the union of left translates by suitable elements of two sufficiently narrow and contracting Schottky families is a Schottky family. By its probabilistic Corollary 4.9, it will be of crucial use in proving the convexity of the rate function.

Let us fix some notation before stating it: let  $t$  be a fixed natural number with  $t > 2 \sum_{i=1}^{d_S} (\dim V_i - 1)$ . Let  $\eta_t > 0$  and the finite subset  $M_t$  of  $\Gamma$  be as given by Lemma 4.2. For a subset  $M$  of  $G$ , denote by  $L(M) = \max_{\gamma \in M} (L(\gamma) \vee L(\gamma^{-1})) \in [1, \infty]$  where  $L(\gamma)$  is defined as in (4.2). Observe that by Lemma 4.5, for any  $M \subset G$  contained in a compact of  $G$ , we have  $L(M) < \infty$ . With these notations, we have:

**PROPOSITION 4.7.** *Let  $E_1$  and  $E_2$  be two  $(\theta_\Gamma, r, \epsilon)$ -Schottky families in  $\Gamma$  with  $\epsilon < \frac{\eta_t}{96L(M_t)^2}$ . Suppose also that  $E_1$  and  $E_2$  are  $\frac{\eta_t}{4L(M_t)^2}$ -narrow. Then, there exist  $\gamma_1$  and  $\gamma_2$  in  $M_t$  such that  $\gamma_1 E_1 \cup \gamma_2 E_2$  is  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky family and we can take  $r_1 = \frac{\eta_t}{48L(M_t)}$  and  $\epsilon_1 = 2\epsilon L(M_t)$ .*

**PROOF.** To simplify the notation, we will only work in one fixed representation  $(\rho, V)$  among  $(\rho_i, V_i)_i$ 's such that  $\alpha_i \in \theta_\Gamma$  and dismiss this from the notation as in the proof of the previous lemma. Our reasonings are such that they simultaneously apply to all representations  $(\rho_i, V_i)_i$  with  $\alpha_i \in \theta_\Gamma$ ; except at one point at the very end of the proof, where of course we will take into account all representations (we explicitly indicate that point).

By hypothesis, there exist  $Y^1$  and  $Y^2$ , subsets of  $\mathbb{P}(V)$  of diameter less than  $\frac{\eta_t}{4L(M_t)^2}$  and such that for  $i = 1, 2$ , for all  $g \in E_i$ , we have  $x_g^+ \in Y^i$ . Let  $y_1$  and  $y_2$  be respectively in  $Y^1$  and  $Y^2$  such that for  $i = 1, 2$ ;  $E_i^+ := \{x_g^+ \mid g \in E_i\} \subseteq B(y_i, \frac{\eta_t}{4L(M_t)^2})$ . Take elements  $\gamma_{1,1}, \dots, \gamma_{1,t}$  and  $\gamma_{2,1}, \dots, \gamma_{2,t}$  from  $M_t$  satisfying the conclusions of Lemma 4.2 respectively for the points  $y_1$  and  $y_2$ .

Reformulating the conclusion 2) of Lemma 4.2; we have that for each hyperplane  $H \subset V$ ; there exist at most  $k$  distinct indices  $i_1, \dots, i_k \subset \{1, \dots, t\}$  with  $k \leq \dim V - 1$ , such that for each  $l = 1, \dots, k$ ,  $\mathbb{P}(H) \cap \gamma_{1,i_l} \cdot B(y_1, \eta_t) \neq \emptyset$ . Indeed, otherwise there exist  $u_1, \dots, u_{\dim V} \in B(y_1, \eta_t)$  and  $\gamma_{1,i_1}, \dots, \gamma_{1,i_{\dim V}} \in M_t$  such that  $\mathbb{P}(H)$  contains the projective image

of the span of the lines  $\{\gamma_{1,i_1} \cdot u_1, \dots, \gamma_{1,i_{\dim V}} \cdot u_{\dim V}\}$  contradicting the conclusion of Lemma 4.2. (Of course, the same conclusion holds true for  $\gamma_{1,i_j}$ 's replaced by  $\gamma_{2,i_j}$ 's and  $y_1$  by  $y_2$ )

Meanwhile, note that for each  $\gamma \in M_t$ ,  $x \in \mathbb{P}(V)$  and  $\delta \geq 0$ , by definition of  $L(M_t)$ , we have

$$(4.4) \quad \gamma B(x, \delta) \subseteq B(\gamma \cdot x, L(M_t)\delta) \subseteq \gamma B(x, L(M_t)^2\delta)$$

Now, we claim that there are at most  $\dim V - 1$  distinct elements  $\gamma_{1,i_1}, \dots, \gamma_{1,i_k}$  among  $\{\gamma_{1,1}, \dots, \gamma_{1,t}\}$  such that

$$(4.5) \quad B(\gamma_{1,i_j} y_1, \frac{\eta_t}{2L(M_t)}) \cap E_1^< \neq \emptyset$$

where we have put  $E_1^< = \bigcup_{g \in E_1} X_g^<$ .

Indeed, if  $i \in \{1, \dots, t\}$  is such that  $B(\gamma_{1,i} y_1, \frac{\eta_t}{2L(M_t)}) \cap E_1^< \neq \emptyset$ , then since by hypothesis for all  $g, h \in E_1$ , one has  $d_H(X_g^<, X_h^<) < \frac{\eta_t}{4L(M_t)^2}$ , we have that for each  $g \in E_1$ ;  $B(\gamma_{1,i} y_1, \frac{1+2L(M_t)}{4L(M_t)^2} \eta_t) \cap X_g^< \neq \emptyset$ . But by (4.4), since  $L(M_t) \geq 1$ , this implies that  $\gamma_{1,i} B(y_1, \frac{1+2L(M_t)}{4L(M_t)} \eta_t) \cap X_g^< \neq \emptyset$  for each  $g \in E_1$ . Therefore, as  $E_1 \neq \emptyset$ , we have found an hyperplane  $\mathbb{P}(H)$  in  $\mathbb{P}(V)$  (take  $H = X_g^<$  for an element  $g \in E_1$ ) such that for each  $i \in \{1, \dots, t\}$  satisfying (4.5), we have  $\gamma_{1,i} B(y_1, \frac{1+2L(M_t)}{4L(M_t)} \eta_t) \cap \mathbb{P}(H) \neq \emptyset$ . Since  $\frac{1+2L(M_t)}{4L(M_t)} < 1$ , the above reformulation of the conclusion of Lemma 4.2 tells us that there are at most  $\dim V - 1$  such indices  $i \in \{1, \dots, t\}$ . Put

$$D_1 := \{i \in \{1, \dots, t\} \mid B(\gamma_{1,i} y_1, \frac{\eta_t}{2L(M_t)}) \cap E_1^< \neq \emptyset\}$$

so that  $|D_1| \leq \dim V - 1$ .

Observe then that for each  $i \in \{1, \dots, t\} \setminus D_1$ ,  $g \in E_1$  and  $x \in X_g^<$ , we have

$$(4.6) \quad d(B(\gamma_{1,i} y_1, \frac{\eta_t}{4L(M_t)}), x) \geq \frac{\eta_t}{4L(M_t)}$$

Therefore, since  $E_1^+ \subseteq B(y_1, \frac{\eta_t}{4L(M_t)^2})$ , by (4.4) we have that for each  $\gamma \in M_t$ ;  $\gamma E_1^+ \subset B(\gamma \cdot y_1, \frac{\eta_t}{4L(M_t)})$  so that (4.6) implies

$$(4.7) \quad d(\gamma_{1,i} x_g^+, X_h^<) \geq \frac{\eta_t}{4L(M_t)}$$

for all  $g, h \in E_1$  and for each  $i \in \{1, \dots, t\} \setminus D_1$ .

As a consequence, since by hypothesis  $\epsilon < \frac{1}{L(M_t)}$  and  $6\epsilon L(M_t) < \frac{\eta t}{4L(M_t)}$ , Lemma 4.6 is in force and gives that for each  $i \in \{1, \dots, t\} \setminus D_1$  and  $g \in E_1$ ;  $\gamma_{1,i}g$  is  $(\frac{\eta t}{12L(M_t)}, 2\epsilon L(M_t))$ -proximal. Moreover,  $d(x_{\gamma_{1,i}g}^+, \gamma_{1,i}x_g^+) < 2\epsilon L(M_t)$  and  $d_H(X_{\gamma_{1,i}g}^<, X_g^<) < \epsilon$ .

Combining these last two inequalities with (4.7), one sees that for all  $g, h \in E_1$ , and for each  $i \in \{1, \dots, t\} \setminus D_1$ , we have

$$(4.8) \quad d(x_{\gamma_{1,i}g}^+, X_{\gamma_{1,i}h}^<) \geq \frac{\eta t}{2L(M_t)} - 2\epsilon L(M_t) - \epsilon \geq \frac{\eta t}{8L(M_t)}$$

Hence, it follows that for each  $i \in \{1, \dots, t\} \setminus D_1$ ,  $\gamma_{1,i}E_1$  is a  $(\frac{\eta t}{48L(M_t)}, 2\epsilon L(M_t))$ -Schottky family.

Repeating exactly the same argument for  $E_2$ , one finds a subset  $D_2$  of  $\{1, \dots, t\}$  such that  $|D_2| \leq \dim V - 1$  and for each  $i \in \{1, \dots, t\} \setminus D_2$ , one has that  $\gamma_{2,i}E_2$  is a  $(\frac{\eta t}{48L(M_t)}, 2\epsilon L(M_t))$ -Schottky family.

Again, the same reasoning, replacing in (4.5)  $E_1^<$  by  $E_2^<$ , allows us to see that there exist at most  $\dim V - 1$  indices  $i \in \{1, \dots, t\}$ , denoting the set of these by  $D_{12}$ , such that for each  $g \in E_1$ ,  $h \in E_2$  and  $i \in \{1, \dots, t\} \setminus D_{12}$ ; we have  $d(\gamma_{1,i}x_g^+, X_h^<) \geq \frac{\eta t}{4L(M_t)}$ . By the same token, we get  $D_{21} \subset \{1, \dots, t\}$  with the corresponding properties.

By consequent, it follows that for each  $i_1 \in \{1, \dots, t\} \setminus D_1 \cup D_{12}$  and  $i_2 \in \{1, \dots, t\} \setminus D_2 \cup D_{21}$ ,  $\gamma_{1,i_1}E_1 \cup \gamma_{2,i_2}E_2$  is a  $(\frac{\eta t}{48L(M_t)}, 2\epsilon L(M_t))$ -Schottky family in  $\mathbb{P}(V)$ .

At this point, as indicated at the beginning of the proof, regarding the construction of the index sets  $D_1, D_2, D_{12}, D_{21}$ , we must take into account each of the representations  $\rho_i$  such that  $\alpha_i \in \theta_\Gamma$ . Hence, repeating the same procedure for each such  $\rho_i$ , we get index subsets  $D_1^j, D_2^j, D_{12}^j, D_{21}^j$  of  $\{1, \dots, t\}$  for each  $j$  such that  $\alpha_j \in \theta_\Gamma \subseteq \{\alpha_1, \dots, \alpha_{d_S}\}$  and with cardinality at most  $\dim V_j - 1$ . Up to re-indexing, set  $\theta_\Gamma = \{\alpha_1, \dots, \alpha_{d_\Gamma}\}$ , where  $d_S \geq d_\Gamma := |\theta_\Gamma|$ .

Finally, denoting  $\tilde{D}_1 := \bigcup_{j=1}^{d_\Gamma} (D_1^j \cup D_{12}^j)$  and  $\tilde{D}_2 := \bigcup_{j=1}^{d_\Gamma} (D_2^j \cup D_{21}^j)$ , since for  $i = 1, 2$ ,  $t > 2 \sum_{j=1}^{d_S} (\dim V_j - 1) \geq |\tilde{D}_i|$ , we have  $\{1, \dots, t\} \setminus \tilde{D}_i \neq \emptyset$ . As a result, choosing  $\gamma_i \in \{1, \dots, t\} \setminus \tilde{D}_i$  for  $i = 1, 2$ , we get that  $\gamma_1 E_1 \cup \gamma_2 E_2$  is a  $(\theta_\Gamma, \frac{\eta t}{48L(M_t)}, 2\epsilon L(M_t))$ -Schottky family, proving the proposition.  $\square$

REMARK 4.8. *One notes from the proof that this proposition is also true with  $\gamma_i$  replaced by any  $\gamma'_i$  in the neighbourhood  $V_{\gamma'_i}$  of  $\gamma_i$  given by Remark 4.3 for  $i = 1, 2$ , and  $L(M_t)$  by  $L(\cup_{\gamma \in M_t} V_\gamma)$ .*

Combining the previous proposition with Lemma 3.11 and Corollary 3.12, we obtain the following technical probabilistic corollary which will be an essential step in our proof of convexity of the rate function. In the corollary, we denote by  $L$ , the Lipschitz constant  $L(\cup_{\gamma \in M_t} V_\gamma)$  of the union of neighbourhoods of elements of  $M_t$  given by Remark 4.3. Since  $M_t$  is a finite set and  $V_\gamma$ 's are bounded, we have  $L \in [1, \infty)$ .

**COROLLARY 4.9.** *Let  $\epsilon$  and  $r$  be given with  $0 < \epsilon < \frac{r}{6} \wedge \frac{\eta t}{96L^2}$ . Then, there exist a natural number  $i_1 = i_1(\mu, M_t)$ , a constant  $d_4 > 0$  depending on the probability measure and a compact subset  $\tilde{K}$  of  $\mathfrak{a}$  with the property that for all subsets  $E_1$  and  $E_2$  of  $\Gamma$  consisting of  $(\theta_\Gamma, r, \epsilon)$ -proximal elements, for all  $n_1, n_2 \in \mathbb{N}$  there exist two natural numbers  $n_1 + i_1 \geq n_{1,1} \geq n_1$  and  $n_2 + i_1 \geq n_{2,2} \geq n_2$ , two  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky families  $\tilde{E}_1$  and  $\tilde{E}_2$  such that  $\tilde{E}_1 \cup \tilde{E}_2$  is an  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky family and for  $i = 1, 2$ ,  $\mathbb{P}(S_{n_i} \in \tilde{E}_i) \geq \mathbb{P}(S_{n_i} \in E_i).d_4$ . Moreover, we have  $\kappa(\tilde{E}_i) \subset \kappa(E_i) + \tilde{K}$ , and one can choose  $r_1 = \frac{\eta t}{48L}$  and  $\epsilon_1 = 2\epsilon L$ .*

**PROOF.** Write  $M_t = \{\gamma_1, \dots, \gamma_m\}$  and put  $i_1 = i_1(\mu, M_t)$  a natural number such that  $M_t \subset \bigcup_{i=1}^{i_1} (\text{supp}(\mu^{*i}))$ . For each  $i = 1, \dots, m$ , take neighbourhoods  $V_{\gamma_i}$  of  $\gamma_i$ 's as in Remark 4.3, set  $k_i \leq i_1$  such that  $\mu^{*k_i}(V_{\gamma_i}) =: \beta_i > 0$  and finally put  $\beta := \min_{1 \leq i \leq m} \beta_i > 0$ . Furthermore, taking the compact subset  $\cup_{i=1}^m \bar{V}_{\gamma_i}$  of  $G$  as  $L$  in Corollary 2.12, get a compact subset  $\tilde{K}$  of  $\mathfrak{a}$  satisfying the conclusion of Corollary 2.12. Let also  $d_2 = d_2(t, \Gamma) > 0$  be the constant given by Lemma 3.11, in which we take  $a = \frac{\eta t}{4L^2}$ ,  $d_3 = d_3(r) > 0$  be the constant given by Corollary 3.12 and finally set  $d_4 = d_2 d_3 \beta > 0$ .

Let now  $E_1$  and  $E_2$  be two given subsets of  $\Gamma$  consisting of  $(\theta_\Gamma, r, \epsilon)$ -proximal elements and  $n_1, n_2 \in \mathbb{N}$ . Applying Corollary 3.12 for  $E_1$  and  $E_2$ , there exist two  $(\theta_\Gamma, \frac{r}{6}, \epsilon)$ -Schottky families,  $E'_1 \subset E_1$  and  $E'_2 \subset E_2$  such that for  $i = 1, 2$

$$(4.9) \quad \mathbb{P}(S_{n_i} \in E'_i) \geq \mathbb{P}(S_{n_i} \in E_i).d_3$$

Noting that subsets of  $(\theta_\Gamma, r, \epsilon)$ -Schottky families are themselves  $(\theta_\Gamma, r, \epsilon)$ -Schottky families, using (4.9) and applying Lemma 3.11 twice with  $a = \frac{\eta t}{4L^2}$  for respectively  $E'_1, E'_2$  and  $n_1, n_2$ , we get two  $\frac{\eta t}{4L}$ -narrow  $(\theta_\Gamma, \frac{r}{6}, \epsilon)$ -Schottky families  $\hat{E}_1 \subset E'_1$  and  $\hat{E}_2 \subset E'_2$  such that for  $i = 1, 2$

$$(4.10) \quad \mathbb{P}(S_{n_i} \in \hat{E}_i) \geq \mathbb{P}(S_{n_i} \in E_i)d_3d_2$$

Now applying Proposition 4.7 (and Remark 4.8) to the  $(\theta_\Gamma, \frac{r}{6}, \epsilon)$ -Schottky families  $\hat{E}_1$  and  $\hat{E}_2$ , remarking that the hypotheses of that proposition is



satisfied by the constructions of  $\hat{E}_1$  and  $\hat{E}_2$ , we get that, up to reindexing, there exist  $\gamma_1, \gamma_2$  in  $M_t$  such that, setting for  $i = 1, 2$ ,  $\tilde{E}_i := V_{\gamma_i} \hat{E}_i$ ,  $\tilde{E}_1 \cup \tilde{E}_2$  is an  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky family, where we can take  $r_1 = \frac{\eta_t}{48L}$  and  $\epsilon_1 = 2\epsilon L$ .

Then, setting  $n_{1,1} := n_1 + k_1 \leq n_1 + i_1$  and  $n_{2,2} = n_2 + k_2 \leq n_2 + i_1$ ; by independence of random walk increments, for  $i = 1, 2$ , we have

$$\begin{aligned} \mathbb{P}(S_{n_{i,i}} \in \tilde{E}_i) &\geq \mathbb{P}(X_{n_i+k_i} \cdots X_{n_i} \in V_{\gamma_i} \text{ and } S_{n_i} \in \hat{E}_i) \\ &= \mathbb{P}(S_{n_i} \in \hat{E}_i) \mathbb{P}(S_{k_i} \in V_{\gamma_i}) \geq \mathbb{P}(S_{n_i} \in E_i) \beta d_3 d_2 = \mathbb{P}(S_{n_i} \in E_i) d_4 \end{aligned}$$

Finally, one remarks that for  $i = 1, 2$ , we have  $\tilde{E}_i \subset M_t \hat{E}_i \subset M_t E_i$  so that by choice of  $\tilde{K}$ , Corollary 2.12 implies that  $\kappa(\tilde{E}_i) \subset \kappa(E_i) + \tilde{K}$ , establishing the last claim.  $\square$

4.10. *Proof of convexity.* We are now in a position to prove the convexity result:

PROOF OF THEOREM 3.4 (CONVEXITY OF THE RATE FUNCTION). Denoting the rate function by  $I$ , start by observing that, by lower semi-continuity, it is sufficient to show that for all  $x_1, x_2 \in \mathfrak{a}$ , we have  $I(\frac{x_1+x_2}{2}) \leq \frac{I(x_1)}{2} + \frac{I(x_2)}{2}$ . For this, we can indeed suppose that  $x_1, x_2$  belongs to the effective domain  $D_I$  of  $I$ , where  $D_I := \{x \in \mathfrak{a} \mid I(x) < \infty\}$ . We shall argue by contradiction.

Suppose there exists  $x_1, x_2 \in D_I$  with  $I(\frac{x_1+x_2}{2}) > \frac{I(x_1)}{2} + \frac{I(x_2)}{2} + 5\xi$  for some  $\xi > 0$ . By the weak LDP and Remark 3.21,  $I$  satisfies

$$(4.11) \quad I(x) = \sup_{\substack{O \text{ open} \\ x \in O}} - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O) = \sup_{\substack{O \text{ open} \\ x \in O}} - \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O)$$

Hence, we can find neighbourhoods  $O_1^{12} \subset O_2^{12}$  of  $\frac{x_1+x_2}{2}$ ; where the inclusions are super-strict and such that

$$(4.12) \quad - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_2^{12}) \geq \frac{I(x_1)}{2} + \frac{I(x_2)}{2} + 4\xi.$$

By (4.11) and (4.12), for  $i = 1, 2$ , one can also find neighborhoods  $x_i \in O_1^i \subset O_2^i \subset O_3^i$  where the inclusions are super-strict and  $O_i^j$ 's are such that  $O_3^1 \cap O_3^2 = \emptyset$ ,  $\frac{O_3^1 + O_3^2}{2} \subset O_1^{12}$  and

$$(4.13) \quad - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_2^{12}) \geq \frac{1}{2} \sum_{i=1}^2 - \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_1^i) + 3\xi$$

It follows from (4.13) that, there exists  $N_0 \in \mathbb{N}$  such that for all  $m \geq N_0$ , we have

$$(4.14) \quad -\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_2^{12}) \geq \frac{1}{2} \sum_{i=1}^2 -\frac{1}{m} \log \mu_m(O_1^i) + 2\xi$$

Now, let  $r = r(\Gamma) > 0$  be as given by Theorem 2.24,  $t = 1 + 2 \sum_i^{d_S} (\dim V_i - 1)$ ,  $\eta_t > 0$ , the finite set  $M_t \subset \Gamma$  as given by Lemma 4.2, for each  $\gamma \in M_t$ , its neighbourhood  $V_\gamma$  as in Remark 4.3 and set  $L \geq 1$  to be the Lipschitz constant  $L(\cup_{\gamma \in M_t} V_\gamma)$ . Choose  $\epsilon < \frac{r}{6} \wedge \frac{\eta_t}{96L^2}$ . Put  $r_1 = \frac{\eta_t}{48L}$  and  $\epsilon_1 = 2\epsilon L$ . Let also the constants  $d_1 = d_1(\epsilon, \Gamma, \mu)$ ,  $i_0 = i_0(\epsilon, \Gamma, \mu)$  and the compact subset  $C = C(\epsilon, \Gamma)$  of  $\mathfrak{a}$  be as given by Lemma 3.9. Denote by  $K$  the compact set  $K(r_1, \epsilon_1) \subset \mathfrak{a}$  given by Proposition 3.14. Let also the compact set  $\tilde{K}$  and the constants  $d_4 > 0$ ,  $i_1 = i_1(\mu, M_t)$  be as in Corollary 4.9. Finally, fix  $Q \in \mathbb{N}$  with for  $i = 1, 2$ ,  $Q \geq Q(O_1^i, O_2^i, C + \tilde{K}) \vee Q(O_2^i, O_3^i, K)$  and  $q = q(O_1^{12}, O_2^{12}) < 1$ , where  $Q(\cdot, \cdot, \cdot)$  and  $q(\cdot, \cdot)$  are as defined in Lemma 3.17.

Now, choose  $n_0 \in \mathbb{N}$  with

- (i)  $n_0 \geq N_0$
- (ii)  $e^{-n_0 \xi} \leq d_1 d_4$
- (iii)  $n_0 \geq Q$
- (iv)  $\frac{n_0}{n_0 + i_0 + i_1} > q$

and put for  $i = 1, 2$ ,  $\alpha_i = -\frac{1}{n_0} \log \mu_{n_0}(O_1^i)$  and  $\beta = -\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O_2^{12})$  so as to have by item (i) of the choice of  $n_0$  and (4.14) that

$$(4.15) \quad \beta \geq \frac{\alpha_1 + \alpha_2}{2} + 2\xi$$

Applying Lemma 3.9 twice, once with taking  $A = n_0 O_1^1$  and the other  $A = n_0 O_1^2$  in that lemma, one gets  $n_1, n_2 \in \mathbb{N}$  with for  $i = 1, 2$   $n_0 + i_0 \geq n_i \geq n_0$  and

$$(4.16) \quad \mathbb{P}(\kappa(Y_{n_i}) \in n_0 O_1^i + C \text{ and } Y_{n_i} \text{ is } (\theta_\Gamma, r, \epsilon)\text{-proximal}) \geq e^{-n_0 \alpha_i} d_1$$

Setting for  $i = 1, 2$ ;  $E_i := \kappa^{-1}(n_0 O_1^i + C) \cap \Gamma_{(r, \epsilon)}$ , where  $\Gamma_{(r, \epsilon)}$  denotes  $(\theta_\Gamma, r, \epsilon)$ -proximal elements of  $\Gamma$ , by (4.16)  $E_i$ 's are non-empty and by our choices of  $r$  and  $\epsilon$ , they satisfy the hypotheses of Corollary 4.9. This corollary therefore gives that for some  $n_{11}, n_{22} \in \mathbb{N}$  with for  $i = 1, 2$ ;  $n_0 + i_0 + i_1 \geq n_{ii} \geq n_0$ , there exist two  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky families  $\tilde{E}_i$  such that  $\tilde{E}_1 \cup \tilde{E}_2$  is also an  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky family with

$$(4.17) \quad \mathbb{P}(Y_{n_{ii}} \in \tilde{E}_i \text{ and } \kappa(Y_{n_{ii}}) \in n_0 O_1^i + C + \tilde{K}) \geq e^{-n_0 \alpha_i} d_1 d_4 \geq e^{-n_0(\alpha_i + \xi)}$$

by the definitions of  $E_i$  above and the last statement of Corollary 4.9 and where the last equality follows from the choice of  $n_0$ , namely item (ii). Furthermore, by item (iii) in the choice of  $n_0$ , (4.17), implies

$$(4.18) \quad \mathbb{P}(Y_{n_{ii}} \in \tilde{E}_i \text{ and } \kappa(Y_{n_{ii}}) \in n_0 O_2^i) \geq e^{-n_0 \alpha_i} d_1 d_4 \geq e^{-n_0(\alpha_i + \xi)}$$

for  $i = 1, 2$ .

Observe now that by our initial choice of open sets, we have  $O_3^1 \cap O_3^2 = \emptyset$ , so that up to taking their intersections, respectively with  $\kappa^{-1}(n_0 O_2^1)$  and  $\kappa^{-1}(n_0 O_2^2)$ , we can suppose that  $\tilde{E}_1$  and  $\tilde{E}_2$  are disjoint and are such that for  $i = 1, 2$ ,  $\kappa(\tilde{E}_i) \subseteq n_0 O_2^i$ . Now, for all  $k_1, k_2 \geq 0$  define the collection of subsets  $E^{k_1, k_2}$  of  $\Gamma$  by

$$E^{k_1, k_2} = \{\gamma_1 \dots \gamma_{k_1+k_2} \mid |\{i \mid \gamma_i \in \tilde{E}_j\}| = k_j \text{ for } j = 1, 2\}$$

Making key use of the fact that  $\tilde{E}_1 \cup \tilde{E}_2$  is an  $(\theta_\Gamma, r_1, \epsilon_1)$ -Schottky family, 3.14 implies that for all  $k_1, k_2 \geq 0$ ,

$$(4.19) \quad \kappa(E^{k_1, k_2}) \subset k_1(n_0 O_2^1 + K) + k_2(n_0 O_2^2 + K) \subset k_1 n_0 O_3^1 + k_2 n_0 O_3^2$$

where the last inclusion is due to item (iii) of the choice of  $n_0$ . Hence, for all  $k \geq 0$ , choosing  $k = k_1 = k_2$ , since  $\frac{O_3^1 + O_3^2}{2} \subseteq O_1^{12}$ , it follows from (4.19) that  $\kappa(E^{k, k}) \subseteq 2kn_0 O_1^{12}$ . Moreover, item (iv) of the choice of  $n_0$  implies by Lemma 3.17 that for all  $k \geq 0$ , we have  $2kn_0 O_1^{12} \subseteq k(n_{11} + n_{22}) O_2^{12}$ .

Consequently, we have the following inclusion of events for each  $k \geq 1$ :

$$(4.20) \quad \{Y_{kn_{11} + kn_{22}} \in E^{k, k}\} \subset \left\{ \frac{1}{kn_{11} + kn_{22}} \kappa(Y_{kn_{11} + kn_{22}}) \in O_2^{12} \right\}$$

Now, using, respectively, (4.20), independence of random walk increments and (4.18), for all  $k \geq 1$ , we have

$$\begin{aligned} \mathbb{P}\left(\frac{\kappa(Y_{kn_{11} + kn_{22}})}{kn_{11} + kn_{22}} \in O_2^{12}\right) &\geq \mathbb{P}(Y_{kn_{11} + kn_{22}} \in E^{k, k}) \\ &\geq \mathbb{P}(Y_{n_{11}} \in \tilde{E}_1)^k \mathbb{P}(Y_{n_{22}} \in \tilde{E}_2)^k \\ &\geq e^{-kn_0(\alpha_1 + \xi)} e^{-kn_0(\alpha_2 + \xi)} \end{aligned}$$

As a result, in the above inequality, taking logarithm, dividing by  $k$ , it follows that

$$-\beta(n_{11} + n_{22}) \geq \limsup_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}\left(\frac{\kappa(Y_{k(n_{11} + n_{22})})}{k(n_{11} + n_{22})} \in O_2^{12}\right) \geq -2n_0\left(\frac{\alpha_1 + \alpha_2}{2} + \xi\right)$$

where the first inequality is immediate by definition of  $\beta$  above.

Finally, dividing this last inequality by  $-(n_{11} + n_{22})$ , using (4.15), we get  $\frac{\alpha_1 + \alpha_2}{2} + 2\xi \leq \beta \leq \frac{\alpha_1 + \alpha_2}{2} + \xi$ , a contradiction.  $\square$

The rest of this section is devoted to completing the proof of Theorem 3.5. It remains to show that the (full) LDP holds under a finite exponential moment condition and that we can give an alternative expression for the rate function under a strong exponential moment condition.

4.11. *Existence of (full) LDP under exponential moment condition.* The following classical notion of large deviations theory enables one to formulate a sufficient condition (see Lemma 4.13) to strengthen a weak LDP to an LDP with proper rate function:

DEFINITION 4.12. *A sequence of random variables  $Z_n$  on a topological space  $X$  is said to be exponentially tight, if for all  $\alpha \in \mathbb{R}$ , there exists a compact set  $K_\alpha \subset X$  such that  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in K_\alpha^c) < -\alpha$ .*

The following lemma (see [14]) explains the interest of this notion:

LEMMA 4.13. *If an exponentially tight sequence of random variables on  $X$  satisfies a weak LDP with a rate function  $I$ , then it satisfies a (full) LDP with a proper rate function  $I$ .*

In view of this lemma, to prove the existence of an LDP with a proper rate function in Theorem 3.5, we only need to show that a finite exponential moment condition on  $\mu$  implies that the sequence  $\frac{1}{n}\kappa(Y_n)$  of random variables is exponentially tight. This is done in the following proposition.

Recall that a probability measure  $\mu$  on  $G$  is said to have a finite exponential moment if there exists  $c > 0$  such that  $\int \exp(c\|\kappa(g)\|) < \infty$ . For convenience, we endow  $\mathfrak{a}$  with the  $l^\infty$ -norm for the dual basis of the characters  $\bar{\chi}_i$  for  $i = 1, \dots, d$ , where these latter are as in the paragraph following Lemma 2.10 (namely, for  $i = 1, \dots, d_S$ ,  $\chi_i$ 's defined by this lemma and for  $i = d_S + 1, \dots, d$ , the central characters  $\chi_i \in X_C$  are defined in paragraph 2.9.1.). Note that by Lemma 2.11 and submultiplicativity of an associated operator norm, this norm satisfies the subadditive property  $\|\kappa(gh)\| \leq \|\kappa(g)\| + \|\kappa(h)\|$  for all  $g, h \in G$ . We have:

PROPOSITION 4.14. *If  $\mu$  has a finite exponential moment, then the sequence random variables  $\frac{1}{n}\kappa(Y_n)$  is exponentially tight.*

PROOF. In view of the above discussion, we only need to show that

$$\lim_{t \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n}\|\kappa(Y_n)\| \geq t\right) = -\infty$$

By Chebyshev inequality, for every  $s \geq 0$ , we have

$$\mathbb{P}(\|\kappa(Y_n)\| \geq tn) \leq \mathbb{E}[e^{s\|\kappa(Y_n)\|}]e^{-stn}$$

In this inequality, taking log, dividing by  $n$  and specializing to some  $s_0 \in \mathbb{R}$  such that  $c \geq s_0 > 0$ , we get

$$\frac{1}{n} \log \mathbb{P}(\|\kappa(Y_n)\| \geq tn) \leq -(s_0 t - \frac{1}{n} \log \mathbb{E}[e^{s_0 \|\kappa(Y_n)\|}])$$

On the other hand, it follows by the independence of random walk increments and the subadditivity of  $\|\cdot\|$  that for all  $n \geq 1$ , we have  $\frac{1}{n} \log \mathbb{E}[e^{s_0 \|\kappa(Y_n)\|}] \leq \log \mathbb{E}[e^{s_0 \|\kappa(X_1)\|}]$ . Therefore, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \|\kappa(Y_n)\| \geq t) \leq -(s_0 t - \mathbb{E}[e^{s_0 \|\kappa(X_1)\|}])$$

Since  $\mathbb{E}[e^{s_0 \|\kappa(X_1)\|}]$  is finite by the exponential moment condition and the choice of  $s_0 > 0$ , the result follows by taking limit in both sides as  $t$  goes to  $+\infty$ .  $\square$

*4.15. Identification of the rate function.* In this last part of this section, under a strong exponential moment condition (see below), we give an alternative expression for the rate function  $I$  as the Legendre transform of a limit Laplace transform of the distributions of  $\frac{1}{n}\kappa(Y_n)$ . For this, we follow a standard path in large deviations theory using the Fenchel-Moreau duality and Varadhan's integral lemma.

Define the limit Laplace transform of  $\frac{1}{n}\kappa(Y_n)$  as  $\Lambda : \mathfrak{a}^* \rightarrow \overline{\mathbb{R}}$  as

$$\Lambda(\lambda) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda(\kappa(Y_n))}]$$

We note in passing that nice properties (e.g. differentiability, steepness) of this function have implications for LDP (e.g. Gärtner-Ellis theorem). For a recent, analytic approach to the study of this function, see Guivarc'h-Le Page [20]. In the next lemma, we write a straightforward observation on the locus of finiteness of  $\Lambda$ . Below, for a  $\lambda \in \mathfrak{a}^*$ ,  $\|\lambda\|_1$  denotes its  $l^1$ -norm in the basis  $(\bar{\chi}_i)_{i=1, \dots, d}$  of  $\mathfrak{a}^*$ , and for convenience, we use the same norm  $\|\cdot\|$  on  $\mathfrak{a}$  as in the proof of Proposition 4.14.

**LEMMA 4.16.** *Let  $\mu$  be a probability measure of finite exponential moment on  $G$ . Accordingly, let  $c > 0$  be such that  $\int e^{c\|\kappa(g)\|} \mu(dg) < \infty$ . Then,*

$$D_\Lambda := \{\lambda \in \mathfrak{a}^* \mid \Lambda(\lambda) \in \mathbb{R}\} \supset \{\lambda \in \mathfrak{a}^* \mid \|\lambda\|_1 \leq c\}$$

PROOF. By definition of the norm  $\|\cdot\|$  on  $\mathfrak{a}$ , for all  $t \in \mathbb{R}$  and  $i = 1, \dots, d$ , we have

$$\log \mathbb{E}[e^{-|t|\|\kappa(Y_n)\|}] \leq \log \mathbb{E}[e^{t \cdot \bar{\chi}_i(\kappa(Y_n))}] \leq \log \mathbb{E}[e^{|t|\|\kappa(Y_n)\|}]$$

where  $Y_n = X_n \dots X_1$  is as usual the  $\mu$ -random walk.

Using this and the fact that the sequence on the right hand side is subadditive and the one on the left hand side is superadditive, one deduces that for  $\lambda = \sum_{i=1}^d \lambda_i \bar{\chi}_i$ , we have

$$\log \mathbb{E}[e^{-\|\lambda\|_1 \|\kappa(X_1)\|}] \leq \frac{1}{n} \log \mathbb{E}[e^{\lambda(\kappa(Y_n))}] \leq \log \mathbb{E}[e^{\|\lambda\|_1 \|\kappa(X_1)\|}]$$

The result follows by the exponential moment hypothesis in the statement of the lemma.  $\square$

We now complete the

PROOF OF THEOREM 3.5 (IDENTIFICATION OF THE RATE FUNCTION). It follows from Lemma 4.16 that if  $\mu$  has a strong exponential moment, then for all  $\lambda \in \mathfrak{a}^*$ ,  $\Lambda(\lambda) < \infty$ . Then, it follows from Varadhan's integral lemma (see [14] section 4.3) that in fact for all  $\lambda \in \mathfrak{a}^*$ , one has

$$\Lambda(\lambda) = \lim_n \frac{1}{n} \log \mathbb{E}[e^{\lambda(\kappa(Y_n))}] = \sup_{x \in \mathfrak{a}} (\langle \lambda, x \rangle - I(x))$$

where  $I$  is the proper rate function of the LDP.

Now, for a function  $f$  on  $\mathfrak{a}$ , denote its convex conjugate (Legendre transform) on  $\mathfrak{a}^*$  by  $f^*(\cdot)$ , where  $f^*(\lambda) := \sup_{x \in \mathfrak{a}} (\langle \lambda, x \rangle - f(x))$ . The above conclusion of Varadhan's integral lemma hence reads as  $\Lambda(\lambda) = I^*(\lambda)$ . Now, since  $I$  is a convex rate function, Fenchel-Moreau duality tells us that  $I(x) = I^{**}(x) = \Lambda^*(x)$ , identifying  $I(x)$  with  $\Lambda^*(x)$  and completing the proof.  $\square$

By the expression of  $I$  given by this identification, one gets an information on the shape of the rate function (which is non-trivial if the support of  $\mu$  is unbounded):

COROLLARY 4.17. *We have  $\lim_{x \rightarrow \infty} \frac{I(x)}{\|x\|} = +\infty$ .*  $\square$

**5. Support of the rate function.** The aim of this section is to prove a more precise version of Theorem 1.7.

Recall that if  $G$  is the group of  $k$ -points of a connected reductive algebraic group  $\mathbf{G}$  defined over a local field  $k$ , and  $S$  is a bounded subset of  $G$  generating a Zariski dense semigroup in  $G$ , then *the joint spectrum of  $S$* , denoted  $J(S)$ , is the Hausdorff limit of both of the sequences  $\frac{1}{n}\kappa(S^n)$  and  $\frac{1}{n}\lambda(S^n)$  ([12]). This is a compact, convex subset of  $\mathfrak{a}^+$ . If  $k = \mathbb{R}$ , then the minimal affine subspace of  $\mathfrak{a}$  containing  $J(S)$  also contains an affine copy of  $\mathfrak{a}_S$ . In particular, when  $k = \mathbb{R}$ , if  $\mathbf{G}$  is semisimple,  $J(S)$  is a convex body in  $\mathfrak{a}$ , and if  $\mathbf{G}$  reductive and  $S$  is symmetric (i.e.  $S = S^{-1} := \{g^{-1} \mid g \in S\}$ ), then  $J(S) \cap \mathfrak{a}_S$  is of non-empty interior in  $\mathfrak{a}_S$ . In the below statement,  $\text{int}(\cdot)$  denotes the interior, and  $\text{ri}(\cdot)$  denotes the relative interior of a set, i.e. its interior in the affine hull of this set. With these definitions, our result reads

**THEOREM 5.1.** *Let  $k$  be a local field and  $G$  be the group of  $k$ -points of a connected reductive algebraic group  $\mathbf{G}$  defined over  $k$ . Let  $\mu$  be a probability measure on  $G$ , whose support  $S$  generates a Zariski-dense semigroup in  $G$ . Then,*

1. *The effective support  $D_I = \{x \in \mathfrak{a} \mid I(x) < \infty\}$  is a convex subset of  $\mathfrak{a}$ . If  $\mathbf{G}$  is semisimple and  $k = \mathbb{R}$ , it is of non-empty interior, and  $\bar{\lambda}_\mu \in \text{int}(D_I)$  if moreover  $\mu$  has a finite second moment.*
2. *If  $S$  is a bounded subset of  $G$ , then  $\bar{D}_I = J(S)$  and  $\text{ri}(D_I) = \text{ri}(J(S))$ .*
3. *If  $S$  is a finite subset of  $G$ , then  $D_I = J(S)$  and  $I$  is bounded above by  $-\min_{g \in S} \log \mu(g)$  on  $\bar{D}_I$ .*

*In any case,  $I$  is locally Lipschitz (in particular continuous) on the relative interior of  $D_I$ .*

**REMARK 5.2.** *Let  $\mu$  be as in the previous theorem and  $S$  denote its support (possibly unbounded). Let  $\Gamma$  be the semi-group generated by  $S$  and  $B_\Gamma$  be the Benoist limit cone of  $\Gamma$  in  $\mathfrak{a}^+$  ([4]). Then, our proof of 2. of the previous theorem (see Proposition 5.7) in fact shows that the projective images  $\mathbb{P}(D_I)$  and  $\mathbb{P}(B_\Gamma)$  of  $D_I$  and  $B_\Gamma$  have the same interior and the same closure in  $\mathbb{P}(\mathfrak{a})$ . In particular, the Benoist limit cone is characterized by the support of the rate function  $I$  (for any probability distribution  $\mu$  of support  $S$ ).*

**REMARK 5.3.** *Note that 2. of the previous theorem says that the rate function  $I$  is finite on the joint spectrum, except possibly on its relative boundary (i.e.  $J(S) \setminus \text{ri}(J(S))$ ). One can easily construct examples of random walks where the corresponding rate function explodes on the boundary (see*

the example below). Moreover, remark that if  $D_I \neq J(S)$ , we have  $\text{Im}(I) = [0, \infty]$ . Indeed,  $0 \in \text{Im}(I)$  since  $I(\vec{\lambda}_\mu) = 0$  and the fact that  $\text{Im}(I)$  fills the whole set  $[0, \infty]$  then follows by convexity and lower semi-continuity of  $I$  using Theorem 5.1.

Using the definition of LDP, we obtain the following result as an immediate corollary of the last (continuity) statement of the previous theorem:

**COROLLARY 5.4.** *Let  $R$  be a subset of  $\mathfrak{a}$  such that  $\text{int}(R) \cap J(S) \neq \emptyset$  and  $\overline{\text{int}(R)} = \bar{R}$  (e.g. a convex body). Then, we have  $\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n} \kappa(Y_n) \in R) = -\inf_{x \in R} I(x)$ .  $\square$*

5.5. *An example of a rate function exploding on the boundary.* In the following, we exhibit an example of a random matrix products whose large deviation rate function explodes on the boundary of the joint spectrum.

*Example.* Let  $G = \text{SL}(2, \mathbb{R})$ ,  $U = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$ ,  $L = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$  and for  $k \in \mathbb{N}$ ,

set  $A_k = \begin{pmatrix} a_k & \\ & a_k^{-1} \end{pmatrix}$ , where  $a_k = e^{4 - \frac{1}{k}}$  and let  $\alpha_k$  be positive real numbers such that  $\sum \alpha_k = 1$ . Consider the probability measure  $\mu = \frac{1}{4}(\delta_U + \delta_L) + \frac{1}{2} \sum_{k \geq 1} \alpha_k \delta_{A_k}$  on  $G$ . Its support  $S$  is bounded and indeed generates a Zariski dense semigroup in  $G$  and hence Theorem 3.5 applies. Let  $I$  be the corresponding proper convex rate function for large deviations of the

random variables  $\frac{1}{n} \log \|Y_n\|$  where  $Y_n = X_n \dots X_1$  is the  $\mu$ -random walk and  $\|\cdot\|$  some associated operator norm. The joint spectrum  $J(S)$  is indeed  $[0, 4]$ , and in particular, by Theorem 5.1,  $\text{int}(D_I) \supseteq (0, 4)$ . Moreover it is obvious that  $0 \in D_I$ . We show  $I(4) = \infty$ : for  $k, n \in \mathbb{N}$ , define the random variables  $P_{k,n}, Q_{k,n}$  and  $R_{k,n}$  as

$P_{k,n} :=$  the number of occurrences of elements of  $\{U, L, A_i \mid i < k\}$  in  $\{X_1, \dots, X_n\}$ ,

$Q_{k,n} :=$  the number of occurrences of elements of  $\{A_i \mid k \leq i < 3k\}$  in  $\{X_1, \dots, X_n\}$ ,

$R_{k,n} := n - P_{k,n} - Q_{k,n}$ .

Then, for all  $k \geq 1$ , one has

$$\mathbb{P}\left(\frac{1}{n} \log \|Y_n\| \geq 4 - \frac{1}{3k}\right) = \sum_{T \in \{P, Q, R\}} \mathbb{P}\left(\frac{1}{n} \log \|Y_n\| \geq 4 - \frac{1}{3k} \text{ and } T_{k,n} \geq \frac{n}{3}\right)$$

Observe that in this last sum, the term corresponding to  $T = P$  is zero by submultiplicativity of the operator norm and the other two terms are asymptotically bounded above by  $\frac{n}{3}$ -powers of respectively  $c \cdot \sum_{i=k}^{3k} \alpha_i$  and



$c \cdot \sum_{i \geq 3k} \alpha_i$ , where  $c > 0$  is a fixed constant. Since these sums converge to zero, by definition of LDP, this shows that  $I(4) = \infty$ .  $\diamond$

5.6. *Proof of Theorem 5.1.* The following proposition shows the key first statement of 2. of Theorem 5.1.

PROPOSITION 5.7. *Let  $G$ ,  $\mu$  and  $S$  be as in 2. of Theorem 5.1. Then  $\overline{D}_I = J(S)$ .*

PROOF. We first show  $\overline{D}_I \subseteq J(S)$ . Since  $J(S)$  is closed by definition, we show  $D_I \subseteq J(S)$ . Let  $x \in D_I$  and  $O_x$  be a neighbourhood of  $x$  in  $\mathfrak{a}$ . Then, by Theorem 3.4, the LDP inequality implies that

$$-\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} \kappa(Y_n) \in O_x\right) \leq \inf_{y \in O_x} I(y) \leq I(x) < \infty$$

In particular, for all  $n \in \mathbb{N}$  large enough,  $\mathbb{P}(\frac{1}{n} \kappa(Y_n) \in O_x) > 0$ , implying that for all  $n$  large enough,  $\frac{1}{n} \kappa(S^n) \cap O_x \neq \emptyset$ . By definition of  $J(S)$ , since  $O_x$  is arbitrary, it follows that  $x \in J(S)$ .

To prove  $\overline{D}_I \supseteq J(S)$ , we shall show that for all  $x \in K(S)$  and  $\delta > 0$ , we have  $B(x, \delta) \cap D_I \neq \emptyset$ . Let such  $x$  and  $\delta$  be given. By definition of  $J(S)$ , there exists  $N_\delta$  such that for each  $n \geq N_\delta$ ,  $\frac{1}{n} \kappa(S^n) \cap B(x, \frac{\delta}{4}) \neq \emptyset$ . Let  $n_0 \in \mathbb{N}$  be large enough (to be specified later), such that  $x_{n_0} \in \frac{1}{n_0} \kappa(S^{n_0})$  and  $x_{n_0} \in B(x, \frac{\delta}{4})$ . Denote by  $g_{n_0}$  an element of  $S^{n_0}$  such that  $x_{n_0} = \frac{1}{n_0} \kappa(g_{n_0})$ , and let  $U_{n_0}$  be a neighbourhood of  $g_{n_0}$  in  $G$  such that  $\frac{1}{n_0} \kappa(U_{n_0}) \subseteq B(x, \frac{\delta}{4})$ . Take a compact  $C$  of  $\mathfrak{a}$  such that  $\frac{1}{n} \kappa(S^n) \subseteq C$  for each  $n \geq 1$ . This is indeed possible since  $S$  is bounded. Finally, put  $\tilde{C} = \max_{x \in C} \|x\|$ .

Denote by  $\Gamma$  the Zariski dense sub-semigroup of  $G$  generated by  $S$  and let  $r = r(\Gamma)$  be as given by Theorem 2.24. Fix  $0 < \epsilon \leq r$  such that  $6\epsilon \leq r$  and let  $F = F_{(r, \epsilon)}$  be the finite subset of  $\Gamma$  given by Theorem 2.24. For each  $f \in F$ , fix a neighbourhood  $V_f$  of  $f$  in  $G$  as in Remark 2.25. Let  $f_0$  be an element of  $F$  such that  $g_{n_0} f_0$  is  $(\theta_\Gamma, r, \epsilon)$ -proximal. Up to reducing  $U_{n_0}$ , we can suppose by Remark 2.25 that for every  $g \in U_{n_0}$  and  $f' \in V_{f_0}$ ,  $gf'$  is  $(\theta_\Gamma, r, \epsilon)$ -proximal.

Furthermore, let  $M$  be the compact subset of  $\mathfrak{a}$  obtained by Corollary 2.12, applying it with  $L = \overline{V}_f$ . Put  $K = K_{(\frac{r}{6}, \epsilon)}$  the compact subset of  $\mathfrak{a}$  given by Proposition 3.14. Fix  $i_0 \in \mathbb{N}$  such that  $f_0 \in S^{i_0}$ , let  $d_3 = d_3(r) > 0$  be as given by Corollary 3.12 and denote  $d_7 = d_3 \mathbb{P}(Y_{i_0} \in V_{f_0}) > 0$ . Finally, set  $\beta_0 = \mathbb{P}(Y_{n_0} \in U_{n_0}) > 0$ .

In Corollary 3.12, taking  $E = U_{n_0}V_{f_0}$  and using it with  $n_1 = n_0 + i_0$ , we get an  $(\theta_\Gamma, \frac{r}{6}, \epsilon)$ -Schottky family  $E_{n_1} \subseteq E$  such that

$$(5.1) \quad \mathbb{P}(Y_{n_1} \in E_{n_1}) \geq d_3 \mathbb{P}(Y_{n_1} \in E)$$

Now, using Proposition 3.14, one sees that if  $n_0 \in \mathbb{N}$  satisfies  $n_0 \geq 16 \frac{i_0 \tilde{C} + \text{diam}(M) + \text{diam}(K)}{\delta} \vee N_\delta$ , then for all  $k \geq 1$ , and  $h_1, \dots, h_k \in E_{n_1}$ , we have  $d(x_{n_0}, \frac{\kappa(h_1 \dots h_k)}{n_1 k}) < \frac{\delta}{2}$ . Therefore, using this, the independence of random walk increments and (5.1), we have

$$\begin{aligned} \mathbb{P}\left(\frac{1}{n_1 k} \kappa(Y_{n_1 k}) \in B(x_{n_0}, \frac{\delta}{2})\right) &\geq \mathbb{P}(Y_{kn_1} \in E_{n_1}^k) \geq \mathbb{P}(Y_{n_1} \in E_{n_1})^k \geq \\ &d_3^k \mathbb{P}(Y_{n_1} \in E)^k \geq d_3^k \mathbb{P}(X_{n_1} \dots X_{i_0+1} \in U_{n_0} \text{ and } Y_{i_0} \in V_{f_0})^k = \\ &d_3^k \mathbb{P}(Y_{n_0} \in U_{n_0})^k \mathbb{P}(Y_{i_0} \in V_{f_0})^k \geq (\beta_0 d_7)^k > 0 \end{aligned}$$

This readily implies that

$$\limsup_{m \rightarrow \infty} \frac{1}{m} \log \mathbb{P}\left(\frac{1}{m} \kappa(Y_m) \in B(x_n, \frac{\delta}{2})\right) \geq \frac{\log(\beta_0 d_7)}{n_1} > -\infty$$

Now, using the definition of LDP, by Theorem 3.4, we get

$$\inf_{y \in B(x, \frac{\delta}{2})} I(y) \leq -\frac{\log(\beta_0 d_7)}{n_1} < \infty$$

In particular,  $D_I \cap B(x, \delta) \neq \emptyset$ , what we wanted to show.  $\square$

We are now ready to complete the proof of Theorem 5.1. We note that in 1., the proof of  $\vec{\lambda}_\mu \in \text{int}(D_I)$  is the same as the proof of  $\vec{\lambda}_\mu \in \text{int}(J(S))$  in [12], when the measure  $\mu$  is supported on a bounded set  $S$  generating a Zariski dense semigroup. This uses the non-degeneracy of the limit Gaussian distribution in central limit theorem (of Goldsheid-Guivarc'h [18], Guivarc'h [19] and Benoist-Quint [9]) together with Abels-Margulis-Soifer's Theorem 2.24 and Benoist estimates (in the form of Proposition 3.14). We omit its proof to avoid lengthy repetitions.

**PROOF OF THEOREM 5.1.** 1. Convexity of  $D_I$  follows immediately from convexity of the rate function  $I$ . Thus,  $D_I$  is convex by Theorem 3.4. If  $\mathbf{G}$  is semisimple,  $k = \mathbb{R}$  and  $S$  is bounded, that  $\text{int}(D_I) \neq \emptyset$  follows by 2. and the fact that in this case  $\text{int}(J(S)) \neq \emptyset$  (see [12] or [26]). If  $S$  is unbounded, then we can find a bounded subset  $S_0$  of  $S$  generating a Zariski dense sub-semigroup in  $G$  and such that  $\mu(S_0) > 0$ . Let  $\mu_0$  be the the

probability measure obtained by restricting  $\mu$  to  $S_0$  and let  $I_0$  be the LDP rate function given by Theorem 3.4 applied to  $\mu_0$ -random walk on  $G$ . Then, by the expression of a rate function in Theorem 3.20, one sees that  $D_{I_0} \subseteq D_I$  and hence we conclude as before.

2.  $\overline{D}_I = J(S)$  is proved in Proposition 5.7. The second assertion  $\text{ri}(J(S)) = \text{ri}(D_I)$  follows from this, since both sets  $J(S)$  and  $D_I$  are convex.

3. We show that  $I$  is bounded above by  $-\min_{g \in S} \log \mu(g)$  on  $D_I$ , the rest follows from lower semi-continuity of  $I$ . Let  $x \in D_I$  so that  $I(x) < \infty$ . It follows by the expression of  $I(x)$  in Theorem 3.20 that there exists a neighbourhood  $O$  of  $x$  in  $\mathfrak{a}$  such that  $\mathbb{P}(\frac{1}{n}\kappa(Y_n) \in O) \neq 0$ , for all  $n$  large enough. Therefore for all such  $n$ , there exist  $g_1, \dots, g_n \in S$ , such that  $\frac{1}{n}\kappa(g_n \dots g_1) \in O$ . Using the independence of random walk increments  $X_i$ 's, we get  $\mathbb{P}(\frac{1}{n}\kappa(Y_n) \in O) \geq \mathbb{P}(X_i = g_i \text{ for each } i = 1, \dots, n) = \prod_{i=1}^n \mathbb{P}(X_i = g_i) \geq (\min_{g \in S} \mu(g))^n$ . Now using again the expression of  $I(x)$  in Theorem 3.20, we conclude that  $I(x) \leq -\min_{g \in S} \log \mu(g)$ .

Finally, the last assertion is a classical fact on convex functions. □

REMARK 5.8. *An interesting observation on the proof of 3. of the previous theorem is the following: (at least) when the support  $S$  of  $\mu$  is a finite set, the Hausdorff convergence of the sequence  $\frac{1}{n}\kappa(S^n)$  is a necessary condition (which is conjectured to hold without any assumptions on  $S$ ) for an LDP to hold for the sequence  $\frac{1}{n}\kappa(Y_n)$  of random variables. This is relevant when one tries to generalize Theorem 3.4 to a random walk governed by a probability measure supported on arbitrary set.*

**6. LDP for Jordan projections.** In this section, we gather some results and examples on large deviations of Jordan projections and make a conjecture.

Although we know that the probabilistic behaviours of averages of Cartan and Jordan projections along a random walk  $Y_n$  are very close (see below), in this article we are not able to prove an LDP for the sequence  $\frac{1}{n}\lambda(Y_n)$  of random variables. Indeed, the following observation of Benoist-Quint ([8]) expresses this close behaviour of  $\frac{1}{n}\kappa(Y_n)$  and  $\frac{1}{n}\lambda(Y_n)$ :

PROPOSITION 6.1 (Lemma 13.13. [8]). *In the setting of Theorem 3.5, for all  $\epsilon > 0$  there exists  $c > 0$  and  $l_0$  such that for every  $n \geq l \geq l_0$ , we have*

$$\mathbb{P}(\|\kappa(Y_n) - \lambda(Y_n)\| > \epsilon l) \leq e^{-c l}$$

From this proposition, one deduces that the averages  $\frac{1}{n}\kappa(Y_n)$  and  $\frac{1}{n}\lambda(Y_n)$  satisfy the *same* limit laws of law of large numbers (with the same limit),

central limit theorem (with the same limit Gaussian distribution), law of iterated logarithm (with the same constant) and exponential decay of probabilities off the Lyapunov vector (i.e. if the sequence  $\frac{1}{n}\lambda(Y_n)$  also satisfies an LDP, its rate function has the same unique zero as that of  $\frac{1}{n}\kappa(Y_n)$ ). On the other hand, it does not seem possible to deduce the same LDP from this proposition. Nevertheless, we believe that the following holds

**CONJECTURE 6.2.** *Let  $G$  be a connected reductive real linear algebraic group and  $\mu$  be a probability measure on  $G$  whose support generates a discrete Zariski-dense semigroup in  $G$ . Then, the sequence  $\frac{1}{n}\lambda(Y_n)$  of random variables satisfies an LDP with the same rate function  $I : \mathfrak{a}^+ \rightarrow [0, \infty]$  given by Theorem 3.5.*

**6.3. Domination of Jordan rate function and some examples.** Regarding this conjecture, the following proposition says that under the hypotheses of Theorem 3.4, if an LDP holds for  $\frac{1}{n}\lambda(Y_n)$ , then one side of the equality of rate functions in the above conjecture is satisfied:

**PROPOSITION 6.4.** *Under the same hypotheses as Theorem 3.4, for  $x \in \mathfrak{a}^+$ , setting*

$$\tilde{J}(x) = \sup_{\substack{O \subset \mathfrak{a} \text{ open} \\ x \in O}} - \limsup_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{n}\lambda(Y_n) \in O\right)$$

*we have  $\tilde{J}(x) \leq I(x)$ , where  $I$  is the rate function given by Theorem 3.4. In particular, if the sequence  $\frac{1}{n}\lambda(Y_n)$  satisfies an LDP with rate function  $J : \mathfrak{a}^+ \rightarrow [0, \infty]$ , then we have  $J(x) \leq I(x)$  for all  $x \in \mathfrak{a}^+$ .*

This proposition is proved along the same lines as the existence of LDP in Theorem 3.4. We provide a brief proof.

**PROOF.** By Theorem 3.20, it suffices to show that for any open set  $O'$  super-strictly containing  $O$ , we have

$$(6.1) \quad \alpha := - \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n}\kappa(Y_n) \in O\right) \geq - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n}\lambda(Y_n) \in O'\right)$$

Let  $0 < \delta < \alpha$  be small enough and  $n_k$  be a sequence such that for all  $k \geq 1$ ,  $\mathbb{P}\left(\frac{1}{n_k}\kappa(Y_{n_k}) \in O\right) \geq e^{-(\alpha+\delta)n_k}$ . Apply, Lemma 3.9 for some  $\epsilon > 0$  small enough, and Corollary 3.12 to get for every  $k \geq 1$  a sequence  $n'_k$  (such that for some  $i_0$  depending only on  $\mu$  and for every  $k \geq 1$ , we have  $|n_k - n'_k| \leq i_0$ ) and an  $(\theta_\Gamma, r, \epsilon)$ -Schottky family  $E_k$  (where  $\Gamma$  denotes the

semigroup generated by the support of  $\mu$  as usual) such that, for all  $k$  large enough we have

$$\mathbb{P}\left(\frac{1}{n'_k} \kappa(Y_{n'_k}) \in O_1 \text{ and } Y_{n'_k} \in E_k\right) \geq d \cdot e^{-(\alpha+\delta)n'_k}$$

where  $O_1$  is an open subset of  $\mathfrak{a}^+$  containing  $O$  and super-strictly contained in  $O'$ , and  $d > 0$  is a positive constant which only depends on  $\mu$  and  $\epsilon$ . Now, by Theorem 2.20 and independence of random walk increments, for each  $t \in \mathbb{N}$  and  $k \geq 1$  large enough, we have

$$\mathbb{P}\left(\frac{1}{t \cdot n'_k} \lambda(Y_{t \cdot n'_k}) \in O'\right) \geq \mathbb{P}(Y_{t \cdot n'_k} \in E_k^t) \geq \mathbb{P}(Y_{n'_k} \in E_k)^t \geq d^t e^{-(\alpha+\delta)n'_k \cdot t}$$

Since  $\delta > 0$  can be taken arbitrarily small, this indeed proves (6.1) and finishes the proof.  $\square$

REMARK 6.5. *If  $e$  denotes the identity in  $G$  and  $\mu(e) > 0$ , one can strengthen this proposition by changing  $\limsup$  to  $\liminf$  in the definition of  $\tilde{J}$ .*

In the following, we give some examples where the above conjecture holds true. As usual,  $G$  denotes a  $k$ -points of a connected reductive linear algebraic group defined over a local field  $k$ .

*Example.* 1. The first example is in a sense trivial, but we mention it to contrast it with the second example: let  $\theta \subseteq \Pi$  and  $r > \epsilon > 0$  be given and let  $E$  be an  $(\theta, r, \epsilon)$ -Schottky family in  $G$ . Let  $\Gamma$  be the semigroup generated by  $E$  and let  $\mu$  be a finitely supported probability measure supported on  $\Gamma$ . Then the conclusion of Conjecture 6.2 holds for the  $\mu$ -random walk. Indeed, for every  $\gamma \in \Gamma$ , by Proposition 2.19, one has  $\|\kappa(\gamma) - \lambda(\gamma)\| \leq M$ , where  $M$  depends only on  $\Gamma$  and the assertion follows easily from this. Note that this example is a purely semigroup case i.e. such a  $\Gamma$  never contains an element and its inverse. Note also that we do not suppose that  $\Gamma$  is Zariski dense, indeed for a  $\mu$  supported on such a semigroup, one does not need the Zariski density hypothesis for the conclusion of Theorem 3.4 to hold.

2. The following situation is more interesting since one does not have the uniform closeness of Cartan and Jordan projections as above: let  $E$  be a *free*  $(\theta, r, \epsilon)$ -Schottky family in  $G$  and let  $E'$  be a subset of  $E \cup E^{-1}$ . Let  $\Gamma$  be the semigroup generated by  $E'$  and  $\mu$  be a finitely supported probability measure on  $\Gamma$ . Then, the conclusion of Conjecture 6.2 holds for the  $\mu$ -random walk. This follows from an elementary calculation using essentially the fact that on a cyclically reduced element (seen as a word in the letters of  $E^{-1} \cup E$ )

the Cartan and Jordan projections are uniformly close (i.e. Proposition 2.19) together with Corollary 2.12.  $\diamond$

### References.

- [1] H. Abels, G.A. Margulis and G.A. Soifer, Semigroups containing proximal linear maps, *Israel journal of mathematics* 91(1-3) (1995), 1-30.
- [2] R.R. Bahadur, Some limit theorems in statistics, *Society for industrial and applied mathematics*, (1971).
- [3] Y. Benoist, Actions propres sur les espaces homogènes réductifs, *Annals of mathematics*, (1996), 315-347.
- [4] Y. Benoist, Propriétés asymptotiques des groupes linéaires, *Geometric & Functional Analysis GAFA* 7(1) (1997), 1-47.
- [5] Y. Benoist, Lectures notes of École Européenne de théorie des groupes, (1997), 1-70.
- [6] Y. Benoist, Propriétés asymptotiques des groupes linéaires, II. *Advanced Studies Pure Math* 26 (2000), 33-48.
- [7] Y. Benoist and F. Labourie, Sur les difféomorphismes d'Anosov affines à feuilletages stable et instable différentiables, *Inventiones mathematicae* 111(1) (1993), 285-308.
- [8] Y. Benoist and J.F. Quint, *Random walks on reductive groups*, Springer International Publishing, (2016).
- [9] Y. Benoist and J.F. Quint, Central limit theorem for linear groups., *Ann. Probab.* 44(2) (2016), 1308-1340.
- [10] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, *Progress in Probability and Statistics*, 8. Birkhäuser Boston Inc., Boston, MA, (1985).
- [11] E. Breuillard and T. Gelander, A topological Tits alternative, *Annals of mathematics*, (2007), 427-474.
- [12] E. Breuillard and C. Sert, Joint spectrum on reductive linear groups, (in preparation).
- [13] F. Bruhat and J. Tits, Groupes réductifs sur un corps local, *Publ. Math. IHES* 41 (1972), 5-251, and 60 (1984), 5-184.
- [14] A. Dembo and O. Zeitouni, *Large deviations techniques and applications (Vol. 38)*, Springer Science & Business Media, (2009).
- [15] H. Furstenberg, Boundary theory and stochastic processes on homogeneous spaces. *Harmonic analysis on homogeneous spaces* 26 (1973), 193-229.
- [16] H. Furstenberg and H. Kesten, Products of random matrices, the *Annals of Mathematical Statistics* 31(2) (1960), 457-469.
- [17] I. Goldsheid and G. Margulis, Lyapunov Indices of a Product of Random Matrices, *Russian Math. Surveys* 44 (1989), 11-81.
- [18] I.Y. Goldsheid and Y. Guivarc'h, Zariski closure and the dimension of the Gaussian law of the product of random matrices, I. *Probability theory and related fields* 105(1) (1996), 109-142.
- [19] Y. Guivarc'h, On the spectrum of a large subgroup of a semisimple group, *Journal of modern dynamics* 2(1) (2008), 15-42.
- [20] Y. Guivarc'h and E. 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) (2016), 503-574.
- [21] E. Le Page, Théorèmes limites pour les produits de matrices aléatoires, in *Probability measures on groups*, Springer Berlin Heidelberg, (1982), 258-303.
- [22] G. Prasad,  $\mathbb{R}$ -regular elements in Zariski dense subgroups, *The Quarterly Journal of Mathematics* 45(4) (1994), 542-545.

- [23] G.C. Rota and W. Strang, A note on the joint spectral radius, (1960).
- [24] J.F. Quint, Cônes limites des sous-groupes discrets des groupes réductifs sur un corps local, *Transformation groups* 7(3) (2002), 247-266.
- [25] J.F. Quint, Divergence exponentielle des sous-groupes discrets en rang supérieur. *Commentarii Mathematici Helvetici* 77(3) (2002), 563-608.
- [26] C. Sert, Joint spectrum and large deviation principle for random matrix products. *Comptes Rendus Mathématique*, 355(6) (2017), 718-722.
- [27] J. Tits, Free subgroups in linear groups. *Journal of Algebra* 20(2) (1972), 250-270.
- [28] J. Tits, Représentations linéaires irréductibles d'un groupe réductif sur un corps quelconque, *Journal für die reine und angewandte Mathematik* 247 (1971), 196-220.
- [29] V.N. Tutubalin, A central limit theorem for products of random matrices and some of its applications, in *Symposia Mathematica* Vol. 21 (1977), 101-116.

DEPARTEMENT MATHEMATIK, ETH ZÜRICH, RÄMISTRASSE 101, ZÜRICH, SWITZERLAND,  
E-MAIL: [cagri.sert@math.ethz.ch](mailto:cagri.sert@math.ethz.ch)