Richard AOUN, Pierre MATHIEU, & Cagri SERT

Random walks on hyperbolic spaces: second order expansion of the rate function at the drift

RANDOM WALKS ON HYPERBOLIC SPACES: SECOND ORDER EXPANSION OF THE RATE FUNCTION AT THE DRIFT

by Richard Aoun, Pierre Mathieu & Cagri Sert

Abstract. — Let (X, d) be a separable geodesic Gromov-hyperbolic space, $o \in X$ a basepoint and μ a countably supported non-elementary probability measure on $\mathrm{Isom}(X)$. Denote by z_n the random walk on X driven by the probability measure μ. Supposing that μ has a finite exponential moment, we give a second-order Taylor expansion of the large deviation rate function of the sequence $\frac{1}{n} d(z_n, o)$ and show that the corresponding coefficient is expressed by the variance in the central limit theorem satisfied by the sequence $d(z_n, o)$. This provides a positive answer to a question raised in [6]. The proof relies on the study of the Laplace transform of $d(z_n, o)$ at the origin using a martingale decomposition first introduced by Benoist–Quint together with an exponential submartingale transform and large deviation estimates for the quadratic variation process of certain martingales.

Résumé (Marches aléatoires sur les espaces hyperboliques : dérivée seconde en la vitesse de fuite de la fonction de taux des grandes déviations)

Soit (X, d) un espace Gromov-hyperbolique, géodésique et séparable, $o \in X$ un point base et μ une mesure de probabilité non élémentaire et à support dénombrable sur le groupe $\mathrm{Isom}(X)$ des isométries de X. Notons par z_n la marche aléatoire sur X induite par μ. Sous l'hypothèse de moment exponentiel fini de μ, nous donnons un développement de Taylor d'ordre 2 de la fonction de taux des grandes déviations de la suite de variables aléatoires $\frac{1}{n} d(z_n, o)$ et exprimons la dérivée seconde en la vitesse de fuite en fonction de la variance dans le théorème central limite que vérifie la suite $d(z_n, o)$. Cela répond par l'affirmative à une question posée dans [6]. La preuve s'appuie sur l'étude de la transformée de Laplace de $d(z_n, o)$ en zéro en utilisant une approximation par une martingale introduite pour la première fois par Benoist-Quint, combinée avec une transformée exponentielle de martingales et des estimées de grandes déviations pour le crochet de certaines martingales.

Contents

1. Introduction

Let (X, d) be a separable geodesic Gromov-hyperbolic space, $G = \mathrm{Isom}(X)$, and $o \in X$ a base point of X. A probability measure μ on G defines a random walk on the group G and subsequently on the metric space X in the following way. Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables on G with distribution μ. We let $L_n = X_n \ldots X_1$ denote the successive positions of the random walk on G. The process $(z_n)_{n \in \mathbb{N}}$ on X defined by $z_n = L_n \cdot o$ constitutes a Markov chain on X that we shall refer to as a random walk on X. To avoid measurability issues, we will always suppose that the probability measure μ is countably supported.

Thanks to the subadditive ergodic theorem, under a finite first moment assumption, we have the following law of large numbers

$$(1.1) \qquad \frac{1}{n} d(z_n, o) \xrightarrow[n \to \infty]{\mathrm{a.s.}} \ell_\mu,$$

where $\ell_\mu \in [0, \infty)$ is a constant called the drift of the random walk. There has recently been substantial interest in the finer study of asymptotic properties of a random walks on Gromov-hyperbolic spaces. This recent progress shows that the resemblance between the asymptotic behaviour of random walk displacement and classical sums of i.i.d. real random variables is far more than the law of large numbers (1.1): a central limit theorem (CLT) with the optimal finite second moment assumption is proved by Benoist–Quint [3] (see also Horbez [22]) improving previous more restrictive versions by Ledrappier [23] and Björklund [5] — an alternative proof of the CLT was later given by Mathieu–Sisto [26] and in a more restrictive setting by Gouëzel [19]. These show that for a non-elementary probability measure μ with finite second moment (see below for the definitions), we have

$$(1.2) \qquad \frac{1}{\sqrt{n}} \left(d(z_n, o) - n\ell_\mu \right) \xrightarrow[n \to \infty]{\mathrm{law}} \mathcal{N}(0, \sigma_\mu^2).$$

The analogue of Cramér's theorem on large deviation principles was recently proved by Boulanger–Mathieu–Sert–Sisto [6] (see also Gouëzel [20]): they showed that for a non-elementary probability measure with a finite exponential moment, the sequence $\frac{1}{n} d(z_n, o)$ satisfies a large deviation principle (LDP) with a proper convex rate function $I : [0, \infty) \to [0, \infty]$ vanishing only at the drift ℓ_μ: for every (measurable) subset R of $[0, \infty)$, we have

$$
\begin{aligned}
(1.3) \qquad - \inf_{\alpha \in \mathrm{int}(R)} I(\alpha) &\leqslant \liminf_{n \to \infty} \frac{1}{n} \ln \mathbb{P}\Big(\frac{1}{n} d(z_n, o) \in R \Big) \\
&\leqslant \limsup_{n \to \infty} \frac{1}{n} \ln \mathbb{P}\Big(\frac{1}{n} d(z_n, o) \in R \Big) \leqslant - \inf_{\alpha \in \overline{R}} I(\alpha),
\end{aligned}
$$

where $\mathrm{int}(R)$ denotes the interior and $\overline{R}$ the closure of R.

Furthermore, concentration inequalities reminiscent of Hoeffding inequalities were recently shown by Aoun–Sert [1] and a local limit theorem for random walks on Gromov-hyperbolic groups was proved by Gouëzel [18].

However, establishing these results analogous to the classical setting of sums of i.i.d. real random variables involves overcoming serious issues by use of various approaches and techniques. Apart from mostly geometric approaches such as the ones used in [6, 20, 26], two classical methods are present — say in aforementioned different proofs of the CLT. These are Nagaev's analytic method [28] and Gordin–Lifšic's martingale method [16].

Nagaev's method can be seen as a version of the classical Fourier–Laplace transform and it relies on techniques of analytic perturbation theory, and in general, yields sharper estimates. However, implementing it requires proving a certain spectral gap result for a Markov operator acting on an appropriate boundary space. Although this is by-now standard, say, on classical hyperbolic spaces or on free groups, it is not well-developed in the generality of spaces, namely (not necessarily proper) geodesic Gromov-hyperbolic spaces that we shall be working with in this article. A thorough study of the analytical method in the case of Gromov-hyperbolic groups was done by Gouëzel [19, Prop. 3.6, §5].

We will extensively use the martingale approach — developed in this setting by Benoist–Quint [2, 3] and adapted to greater generality by Horbez [22] and Aoun–Sert [1] — to tackle the analytic problem of giving a second-order expansion of the limit Laplace transform of the sequence $d(z_n, o)$ (or by convex duality, of its large deviation rate function in (1.3)) and relating it to the variance in the central limit theorem (1.2). Similar results are known to hold in settings where spectral methods are available. We now expound on these notions and precisely state the main result of this note.

A geodesic metric space (X, d) is said to be Gromov-hyperbolic if there exists $\delta > 0$ such that for every $x, y, z, o \in X$, we have $(x|y)_o \geqslant (x|z)_o \wedge (z|y)_o - \delta$, where $(.|.)$ is the Gromov-product given by $(x|y)_o = \frac{1}{2}(d(x, o) + d(y, o) - d(x, y))$. A probability measure μ on $\mathrm{Isom}(X)$ is called *non-elementary* if its support S generates a semigroup that contains two independent loxodromic elements (see Section 3.2). For such a measure μ and $n \in \mathbb{N}$, we denote by μ^{*n} its nth convolution which is the law of the random variable L_n.

Given a probability measure μ, the limit Laplace transform of the sequence $\frac{1}{n}d(z_n, o)$ is the function $\Lambda : \mathbb{R} \to (-\infty, \infty]$ defined by

$$(1.4) \qquad \Lambda(\lambda) = \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda d(z_n, o)}].$$

Note that, since the increments are i.i.d and G acts by isometries on X, subadditivity implies that the limit in (1.4) exists. Under a finite super-exponential moment assumption, the Fenchel-Legendre transform of Λ is the rate function I of the large deviation principle satisfied by $\frac{1}{n}d(z_n, o)$ (see [6, Lem. C.2]). Using this and [6, Th. 1.1], one can deduce that the derivative of Λ at 0 is equal to the drift ℓ_μ.

The goal of this note is to prove the following result which answers part of [6, Quest. C.1] and which says that the convex function Λ has a second order Taylor expansion at 0 with second derivative equal to the variance in the central limit theorem:

Theorem 1.1. — *Let (X,d) be a separable geodesic Gromov-hyperbolic space and μ a non-elementary probability measure on $\mathrm{Isom}(X)$. Suppose that μ has a finite exponential moment, i.e., for some $\alpha > 0$, $\int e^{\alpha\, d(g\cdot o,o)}d\mu(g) < +\infty$. Then, we have*

$$\lim_{\lambda\to 0}\frac{\Lambda(\lambda) - \lambda\ell_\mu}{\lambda^2} = \frac{\sigma_\mu^2}{2}.$$

The proof uses extensively the martingale approach developed in this context by Benoist–Quint [2, 3]. The martingale decomposition proved in these works allows us to reduce the study of Λ near zero to the study of the limit Laplace transform of a martingale induced by an i.i.d. random walk on the group $\mathrm{Isom}(X)$. Once this reduction is done, the proof is divided into two parts: proving the lower bound, i.e., $\lim_{\lambda\to 0}(\Lambda(\lambda) - \lambda\ell_\mu)/\lambda^2 \geqslant \sigma_\mu^2/2$ and the upper bound, i.e., $\lim_{\lambda\to 0}(\Lambda(\lambda) - \lambda\ell_\mu)/\lambda^2 \leqslant \sigma_\mu^2/2$. The proof of the lower bound is based on a new exponential submartingale transform that we establish in Proposition 2.2. The latter extends a classical result of Freedman [15] to the case of martingales with unbounded differences. The proof of the upper bound uses ideas from martingale concentration inequalities. Another important tool is large deviation estimates for the quadratic variation of our martingales.

Remark 1.2

(1) (*Busemann cocycle*) A general version of Theorem 1.1 will be proved in Theorem 4.1 where the displacement $d(z_n, o)$ is replaced with the Busemann cocycle $\sigma(L_n, x)$ of L_n based at any point of x in the horofunction compactification of X. See also Question 4.8 for an ensuing problem.

(2) (*Translation distance*) Thanks to [6, Th. 1.3], when μ has bounded support, one can replace $d(z_n, o)$ by $\tau(L_n)$ in (1.4), where $\tau(.)$ denotes the translation distance given for $g \in \mathrm{Isom}(X)$ by $\tau(g) = \lim_{n\to\infty}\frac{1}{n}d(g^n \cdot o, o)$.

(3) (*Positivity of σ_μ*) By an argument of Benoist–Quint [3], it follows from the expression of σ_μ^2 (see (3.14)) that $\sigma_\mu > 0$ if any only if μ is non-arithmetic (see Remark 3.7).

Using the convexity of the rate function proved in [6] and standard results from convex analysis, we deduce

Corollary 1.3 (About the rate function). — *Keep the assumptions of Theorem 1.1 and let I be the rate function (1.3). Then, we have*

$$\lim_{\lambda\to 0}\frac{I(\ell_\mu + \lambda)}{\lambda^2} = \frac{1}{2\sigma_\mu^2}.$$

Finally, we note that our results are also valid for the right random walk $R_n = X_1\dots X_n$ since for every $n \in \mathbb{N}$, L_n and R_n have the same distribution.

Remark 1.4 (Role of hyperbolicity and possible extensions). — The key ingredients of the proof of our main result in which hyperbolicity plays a role are Lemma 3.1 (in combination with the solution of cohomological equation (3.4)) and Lemma 3.8 (a qualitative form of which is also sufficient for the purposes). It may therefore be

possible to establish this connection between large deviations and central limit theorem in similar geometric settings (see for example [22] and the recent monograph [8] and the articles [9, 10].)

The paper is organized as follows. In Section 2, we recall some preliminaries on submartingales and prove an exponential transform for submartingales. In Section 3, we recall basic definitions about Gromov-hyperbolic spaces and metric compactifications as well as results from the theory of random walks on hyperbolic spaces. In particular, we recall that $d(z_n, o) - n\ell_\mu$ is at bounded distance from a martingale and prove a large deviation estimate for the predictable quadratic variation of the latter. In Section 4 we prove Theorem 1.1 in its general form Theorem 4.1, by treating separately the lower bound (Section 4.2) and the upper bound (Section 4.3). In Section 4.4, we deduce Corollary 1.3, and finally, discuss some ensuing questions in Section 4.5.

Acknowledgements. — The authors would like to thank the anonymous referee for the careful reading of our paper and useful suggestions.

2. Preliminaries on martingales

In this section, we recall some preliminaries from the theory of martingales and prove a result about exponential martingale transforms that will play a crucial role in the proof of our main theorem.

Let us first fix our notation. We shall denote by $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$ an increasing sequence of σ-algebras (a filtration) on a fixed standard probability space Ω. Usually, we will consider the filtration to be fixed and omit it from the notation. The notation $M = (M_n)_{n \in \mathbb{N}}$ will be reserved for an adapted sequence of random variables that form either a martingale or submartingale. Denoting by ΔM the sequence of differences given by $\Delta_n M := M_n - M_{n-1}$, we recall that M being a submartingale means that for every $n \in \mathbb{N}$, M_n is $\mathcal{F}_n$-measurable, integrable, and it satisfies $\mathbb{E}[\Delta_n M | \mathcal{F}_{n-1}] \geqslant 0$. In the sequel, unless otherwise stated, we take $M_0 = 0$ a.s. The predictable quadratic variation (or conditional quadratic variation) of the submartingale M_n is denoted by $\langle M \rangle_n := \sum_{i=1}^n \mathbb{E}[(\Delta_i M)^2 | \mathcal{F}_{i-1}]$. Given a positive constant $a > 0$, we denote

$$(2.1) \qquad G_n^a = \sum_{i=1}^n \mathbb{E}[(\Delta_i M)^2 1_{|\Delta_i M| \leqslant a} | \mathcal{F}_{i-1}] - a \sum_{i=1}^n |\Delta_i M| 1_{|\Delta_i M| \geqslant a}.$$

Finally, the following special function defined on $\mathbb{R}$ will play a significant role: $\mathfrak{f}(\lambda) = e^{-\lambda} - 1 + \lambda$.

We start by recalling Freedman's submartingale transform whose statement and proof strategy will be used in our generalization below.

Proposition 2.1 ([15, Cor. 1.4(b), (3.7) & (3.9)])

(1) *Let X be an integrable random variable with $\mathbb{E}(X) = 0$ (resp. $\mathbb{E}[X] \geqslant 0$) and $X \geqslant -1$ (resp. $|X| \leqslant 1$) a.s. Then, for every $\lambda \geqslant 0$, we have*

$$\mathbb{E}[e^{\lambda X}] \geqslant e^{\mathfrak{f}(\lambda) \mathrm{Var}(X)}.$$

(2) *Let $(M_n)_{n\in\mathbb{N}}$ be a submartingale and such that for every $1 \leqslant n \in \mathbb{N}$, $|\Delta_n M| \leqslant 1$ almost surely. Then for every $\lambda \geqslant 0$, the sequence of random variables*

$$(\exp(\lambda M_n - \mathfrak{f}(\lambda)\langle M \rangle_n))_{n\in\mathbb{N}}$$

is a submartingale with respect to the same filtration.

We note that the second statement above is a consequence of the first one. The following result provides a generalization of Proposition 2.1 to submartingales with increments possessing a finite exponential moment.

Proposition 2.2 (Submartingale transform). — *Let $(M_n)_{n\in\mathbb{N}}$ be a submartingale. Suppose that there exists a constant $\alpha > 0$ such that for every $n \in \mathbb{N}$, we have $\mathbb{E}[e^{\alpha \sum_{k=1}^n |\Delta_k M|}] < \infty$. Then, given any $a > 0$, for every $\lambda > 0$ small enough, the sequence of random variables*

$$\left(\exp\left(\lambda M_n - \frac{\mathfrak{f}(\lambda a)}{a^2} G_n^a\right)\right)_{n\in\mathbb{N}}$$

is a submartingale with respect to the same filtration.

This is an extension (to unbounded differences) of Freedman's submartingale transform in his seminal work [15]. Indeed, if the difference sequence $\Delta_n M$ satisfies $|\Delta_n M| \leqslant 1$ a.s., the transform in the previous result boils down to Proposition 2.1. On the other hand, it applies, for instance, when there exists a constant $\alpha > 0$ such that for every $n \in \mathbb{N}$, $\mathbb{E}[e^{\alpha|\Delta_n M|}|\mathcal{F}_{n-1}] < +\infty$. This will be the case in our application. The counterparts of Proposition 2.2 for supermartingale transforms were obtained by Dzhaparidze–van Zanten [13] (see also Fan–Grama–Liu [14]).

Proof. — Let $a > 0$. By the finite exponential moment hypothesis on $(M_n)_{n\in\mathbb{N}}$, it is clear that for every $\lambda > 0$ small enough and for every $n \in \mathbb{N}$, $\exp(\lambda M_n - (\mathfrak{f}(\lambda a)/a^2)G_n^a)$ is $\mathcal{F}_n$-measurable and integrable. Therefore, by expanding the conditional expectation, one sees that it is enough to show the following: for any integrable random variable X with $\mathbb{E}[X] \geqslant 0$, for any $\lambda > 0$,

$$(2.2) \qquad \mathbb{E}\left[\exp\left(\lambda X + \frac{\mathfrak{f}(\lambda a)}{a}|X|\mathbb{1}_{|X|\geqslant a}\right)\right] \geqslant e^{(\mathfrak{f}(\lambda a)/a^2)\mathbb{E}[X^2\mathbb{1}_{|X|\leqslant a}]}.$$

Denote by ν the distribution of X.

– Case 1: $\mathbb{E}[X] = 0$ and ν is supported on two points $-c$ and d with $c, d > 0$ and both $c, d \leqslant a$. Let $r := \max\{c, d\}$. Since $\mathbb{E}[X] = 0$ and $X \geqslant -r$ almost surely, 1. of Proposition 2.1 (applied to X/r and to λr) yields

$$\mathbb{E}\left[\exp\left(\lambda X + \frac{\mathfrak{f}(\lambda a)}{a}|X|\mathbb{1}_{|X|\geqslant a}\right)\right] = \mathbb{E}[\exp(\lambda X)] \geqslant e^{(\mathfrak{f}(\lambda r)/r^2)\mathbb{E}[X^2]}.$$

Since the function $x \mapsto \mathfrak{f}(x)/x^2$ is decreasing on $\mathbb{R}$ and since $r \leqslant a$, we deduce that $\mathfrak{f}(\lambda r)/r^2 \geqslant \mathfrak{f}(\lambda a)/a^2$. Therefore

$$\mathbb{E}\left[\exp\left(\lambda X + \frac{\mathfrak{f}(\lambda a)}{a}|X|\mathbb{1}_{|X|\geqslant a}\right)\right] \geqslant e^{(\mathfrak{f}(\lambda a)/a^2)\mathbb{E}[X^2]} = e^{(\mathfrak{f}(\lambda a)/a^2)\mathbb{E}[X^2\mathbb{1}_{|X|\leqslant a}]}.$$

– Case 2: $\mathbb{E}[X] = 0$ and ν is supported exactly on two points $-c$ and d and we are not in Case 1. By Jensen's inequality, we obtain

$$(2.3) \qquad \mathbb{E}\Big[\exp\Big(\lambda X + \frac{\mathfrak{f}(\lambda a)}{a}\,|X|\mathbb{1}_{|X|\geqslant a}\Big)\Big] \geqslant \exp\Big(\lambda\mathbb{E}[X] + \frac{\mathfrak{f}(\lambda a)}{a}\,\mathbb{E}[|X|\mathbb{1}_{|X|\geqslant a}]\Big)$$
$$= \exp\Big(\frac{\mathfrak{f}(\lambda a)}{a}\,\mathbb{E}[|X|\mathbb{1}_{|X|\geqslant a}]\Big).$$

If both $c, d \geqslant a$, then the right hand side of (2.2) is equal to 1 and hence (2.2) holds in view of (2.3). So, since we are also not in Case 1, we can suppose that either $c > a$ and $d < a$, or $d > a$ and $c < a$. Let us treat the case $c > a$ and $d < a$. Notice also that since $\mathbb{E}[X] = 0$ we have $\mathbb{P}(X = -c) = d/(c+d)$ and $\mathbb{P}(X = d) = c/(c+d)$. By assumption on c, d and (2.3), these yield

$$(2.4) \qquad \mathbb{E}\Big[\exp\Big(\lambda X + \frac{\mathfrak{f}(\lambda a)}{a}\,|X|\mathbb{1}_{|X|\geqslant a}\Big)\Big] \geqslant \exp\Big(\frac{\mathfrak{f}(\lambda a)}{a}\,\frac{cd}{c+d}\Big).$$

On the other hand,

$$(2.5) \qquad \exp\Big(\frac{\mathfrak{f}(\lambda a)}{a^2}\,\mathbb{E}[X^2 1_{|X|\leqslant a}]\Big) = \exp\Big(\frac{\mathfrak{f}(\lambda a)}{a^2}\,\frac{d^2 c}{c+d}\Big).$$

Since $d \leqslant a$, (2.2) follows from combining (2.4) and (2.5). The case $d > a$ and $c \leqslant a$ can be treated similarly.

– Case 3: $\mathbb{E}[X] \geqslant 0$ and ν is supported on two points $\{-c, d\}$ with $c, d > 0$. We will study the behavior of the left-hand-side and the right-hand-side of (2.2) when we vary ν with the condition $\mathbb{E}[X] \geqslant 0$, while fixing a, c, d, λ. Since ν is supported on two points, it is enough to treat the behavior of these quantities when $\beta := \mathbb{P}(X = d)$ varies. Observe that since $\mathbb{E}[X] \geqslant 0$, we have $c/(c+d) \leqslant \beta \leqslant 1$. The function $\psi_1(\beta) := \mathbb{E}_\nu[\exp(\lambda X + (\mathfrak{f}(\lambda a)/a)|X|\mathbb{1}_{|X|\geqslant a})]$ is affine in β while the function $\psi_2(\beta) := e^{(\mathfrak{f}(\lambda a)/a^2)\mathbb{E}_\nu[X^2 \mathbb{1}_{|X|\leqslant a}]}$ is convex in β (being of the form $\psi_2(\beta) = e^{L(\beta)}$ with L an affine map). Thus the function $\psi := \psi_1 - \psi_2$ is concave on $[c/(c+d), 1]$. It suffices then to check that $\psi(c/(c+d)) \geqslant 0$ and that $\psi(1) \geqslant 0$. The case $\beta = c/(c+d)$ corresponds to the case $\mathbb{E}_\nu[X] = 0$ and hence, by cases 1 and 2, $\psi(c/(c+d)) \geqslant 0$. The case $\beta = 1$ corresponds to $\nu = \delta_d$. Clearly, $\psi(1) \geqslant 0$ when $d \geqslant a$. When $d < a$, the relation $\psi(1) \geqslant 0$ follows from the facts that the function $x \mapsto \mathfrak{f}(x)/x^2$ is decreasing and that $\mathfrak{f}(x) \leqslant x$ for every $x \geqslant 0$. This concludes the proof of (2.2) in this case.

– Case 4: here we treat the general case (cf. proof of [15, Prop. 3.6]). Since $\mathbb{E}[X] \geqslant 0$, we can find a family $(\nu_\alpha)_{\alpha\in I}$ of probability measures, each supported on two points $-c_\alpha \leqslant 0$ and $d_\alpha > 0$ and of expectation $\geqslant 0$, and a probability measure θ on I such that $\nu = \int d\theta(\alpha)\nu_\alpha$. We have

$$\mathbb{E}\Big[\exp\Big(\lambda X + \frac{\mathfrak{f}(\lambda a)}{a}\,|X|\mathbb{1}_{|X|\geqslant a}\Big)\Big] = \int\Big(\int e^{\lambda x + (\mathfrak{f}(\lambda a)/a)|x|1_{|x|\geqslant a}}\,d\nu_\alpha(x)\Big)\,d\theta(\alpha)$$
$$\geqslant \int e^{(\mathfrak{f}(\lambda a)/a^2)\int x^2 1_{|x|\leqslant a}\,d\nu_\alpha(x)}\,d\theta(\alpha)$$
$$\geqslant e^{(\mathfrak{f}(\lambda a)/a^2)\int\!\int x^2 \mathbb{1}_{|x|\leqslant a}\,d\nu_\alpha(x)d\theta(\alpha)}$$
$$= e^{(\mathfrak{f}(\lambda a)/a^2)\mathbb{E}[X^2 1_{|X|\leqslant a}]},$$

where we applied (2.2) for each probability measure ν_α in the second inequality and Jensen in the third inequality. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3. Random walks on hyperbolic spaces

3.1. **Preliminaries on hyperbolic spaces.** — Let us first fix our notation. Let (X, d) be a geodesic metric space. Fix a base point $o \in X$. Recall that (X, d) is said to be δ-hyperbolic (where $\delta \geqslant 0$) if for every $x, y, z, o \in M$,

$$(3.1) \qquad\qquad (x|y)_o \geqslant (x|z)_o \wedge (z|y)_o - \delta,$$

where $(.|.)$ is the Gromov-product given by $(x|y)_o = \frac{1}{2}(d(x, o) + d(y, o) - d(x, y))$. For simplicity, we will often omit the basepoint o from the notation. We recall that this category of metric spaces comprises many usual spaces: trees, classical hyperbolic spaces, the fundamental group of compact surfaces of genus $\geqslant 2$. We recall that the definition of hyperbolicity is equivalent to geodesic triangles being thin. We refer to [11] for general properties of these spaces. Denote by $G := \mathrm{Isom}(X)$ the group of isometries of the metric space (X, d). The displacement of $g \in G$ is by definition

$$\kappa(g) := d(g \cdot o, o).$$

An element $\gamma \in G$ is said to be loxodromic if for any $x \in X$, the sequence $(\gamma^n x)_{n \in \mathbb{Z}}$ constitutes a quasi-geodesic (see [11, Ch. 3]). Equivalently, γ is loxodromic if and only if it fixes precisely two points x_γ^+, x_γ^- on the Gromov boundary ∂X of X [11, Ch. 9 & 10]. Two loxodromic elements γ_1, γ_2 are said to be independent if the sets of fixed points $\{x_{\gamma_i}^+, x_{\gamma_i}^-\}$ for $i = 1, 2$ are disjoint. Finally, a set S, or equivalently a probability measure with support S, is said to be non-elementary if the semigroup generated by S contains at least two independent loxodromic elements.

Now we recall the definition of the Busemann compactification of X (no need for hyperbolicity in this part). Denote by $\mathrm{Lip}^1(X)$ the set of real valued Lipschitz functions on X with Lipschitz constant 1, endowed with the topology of pointwise convergence. Fixing $o \in X$, for $x \in X$, let the function $h_x \in \mathrm{Lip}_o^1(X)$, defined by $h_x(m) = d(x, m) - d(x, o)$, where $\mathrm{Lip}_o^1(X)$ is the subspace of $\mathrm{Lip}^1(X)$ consisting of functions f satisfying $f(o) = 0$. If X is separable (as we suppose in the sequel), the closure of $\{h_x \mid x \in X\}$ is a compact metrizable subset of $\mathrm{Lip}_o^1(X)$, called the *horofunction compactification* of X (see e.g. [25, Prop. 3.1]). It will be denoted as $\overline{X}^h$. The map $x \mapsto h_x$ is injective on X (and an embedding when X is a proper metric space) and we usually identify X with its image in $\overline{X}^h$. The *horofunction boundary* of X is defined as $\partial_h X := \overline{X}^h \smallsetminus X$. The group of isometries $\mathrm{Isom}(X)$ acts on $\overline{X}^h$ by homeomorphisms given, for $g \in \mathrm{Isom}(X)$, $h \in \overline{X}^h$ and $m \in X$, by $(g \cdot h)(m) = h(g^{-1}m) - h(g^{-1}o)$. This extends equivariantly the isometric action of $\mathrm{Isom}(X)$ on X and the set $\partial_h X \subset \overline{X}^h$ is invariant under $\mathrm{Isom}(M)$. The *Busemann cocycle* $\sigma : \mathrm{Isom}(X) \times \overline{X}^h \to \mathbb{R}$ is defined by

$$\sigma(g, h) = h(g^{-1}o).$$

Observe that for every $g \in G$ and $x \in \overline{X}^h$,

$$(3.2) \qquad \sigma(g, o) = \kappa(g) \quad \text{and} \quad |\sigma(g, x)| \leqslant \kappa(g).$$

Finally, we recall that the Gromov-product can be extended to the whole Busemann compactification by setting $(x|y)_o := -\min_{z \in X} \frac{1}{2}(h_x(z) + h_y(z))$. In particular, one can infer that for $x \in \overline{X}^h$ and $y \in X$,

$$(3.3) \qquad (x|y)_o = \frac{1}{2}(d(y, o) - h_x(y)).$$

3.2. RANDOM WALKS. — There are two main goals in this section. The first one (discussed in Section 3.2.1) is to recall a martingale decomposition (Lemma 3.1) of the Busemann cocycle along non-elementary random walks on Gromov-hyperbolic spaces which is due to Benoist–Quint [2, 3] (see also an extension in [22]). We will use a slightly more general version of this worked out in [1]. The second goal (discussed in Section 3.2.2) is to prove Proposition 3.3 about large deviations of predictable quadratic variation and its consequence expressed in Corollary 3.9. The latter will be crucial in the proof of our main result.

3.2.1. *Benoist–Quint martingale decomposition*. — Let μ be a probability measure on the isometry group G of X with countable support. Recall that it is said to have a finite exponential moment (resp. finite second moment) if there exists $\alpha > 0$ such that $\int e^{\alpha d(g \cdot o, o)} d\mu(g) < \infty$ (resp. $\int \kappa(g)^2 d\mu(g) < \infty$). Let $L_n = X_n \dots X_1$ be the random walk on G and ℓ_μ the drift of the random walk on X defined in (1.1). Denote by $\mathcal{F}$ the natural filtration generated by the increments X_i's. Finally, we denote by P_μ the Markov operator on the horofunction compactification $\overline{X}^h$ induced by the random walk on G, i.e., $P_\mu f(x) = \int f(g \cdot x) d\mu(g)$ for every bounded measurable function f on $\overline{X}^h$. The starting point of the proof of Theorem 1.1 is the following.

LEMMA 3.1. — *Let μ be a non-elementary probability measure with finite second moment. Then, for every $x \in \overline{X}^h$, there exists a martingale $M_x = (M_{x,n})_{n \in \mathbb{N}}$ with respect to the filtration $\mathcal{F}$ starting at the origin and such that for every $n \in \mathbb{N}$,*

$$\sigma(L_n, x) - n\ell_\mu = M_{x,n} + O_{x,n}(1),$$

where $O_{x,n}(1)$ is a random variable whose absolute value is bounded uniformly in $n \in \mathbb{N}$ and $x \in \overline{X}^h$.

Proof. — When X is proper, Benoist–Quint [3, Prop. 4.6] showed that that there exists a bounded measurable function ψ on the Busemann boundary $\partial_h X$ such that

$$\psi - P_\mu \psi = \int \sigma(g, x) d\mu(g) - \ell_\mu.$$

It was then verified in [22] that this solution can be extended to the case when X is non-proper and also in [1] that ψ could be defined on the whole compactification $\overline{X}^h$ while preserving the boundedness of ψ. This is equivalent to finding a cocycle

$\sigma_0 : G \times \overline{X}^h \to \mathbb{R}$ with constant drift equal to ℓ_μ, i.e.,

$$\int \sigma_0(g, x) d\mu(g) = \ell_\mu$$

for every $x \in \overline{X}^h$, such that the following identity holds for every $(g, x) \in G \times \overline{X}^h$:

(3.4) $$\sigma(g, x) = \sigma_0(g, x) - \psi(g \cdot x) + \psi(x).$$

Let then

(3.5) $$M_{x,n} := \sigma_0(L_n, x) - n\ell_\mu.$$

The constant drift property of σ_0 implies that $M_x := (M_{x,n})_{n \in \mathbb{N}}$ is a martingale with respect to the filtration $\mathcal{F}$, which finishes the proof. $\qquad\square$

Remark 3.2. — Observe that since $\mathbb{E}(M_{x,n}) = \mathbb{E}(M_{x,0}) = 0$ for every $n \in \mathbb{N}$, we obtain the existence of some $C \geqslant 0$ such that for every $n \in \mathbb{N}$ and every $x \in M$,

$$n\ell_\mu - C \leqslant \mathbb{E}[\sigma(L_n, x)] \leqslant n\ell_\mu + C.$$

From now on, for every $x \in \overline{X}^h$ we denote by $M_x = (M_{x,n})_{n \in \mathbb{N}}$ the martingale defined in the proof of Lemma 3.1, i.e.,

$$M_{x,n} := \sigma_0(L_n, x) - n\ell_\mu.$$

Many properties of a martingale are encoded in its different notions of quadratic variation. For instance, a martingale whose predictable quadratic variation (see below for the definition) is almost surely bounded satisfies a Bennett–Bernstein concentration result (see [15] for the bounded difference case and [29, 13, 14] for the general case). Burkholder inequalities [7] are another instance of the relevance of the quadratic variation in studying martingales.

3.2.2. *Large deviation estimate for predictable quadratic variation of $M_{x,n}$.* — We now proceed with the second goal of Section 3.2, namely proving Proposition 3.3 below and deducing Corollary 3.9. We first give some observations and definitions regarding the martingale $(M_{x,n})_{n \in \mathbb{N}}$ introduced in Section 3.2.1. The martingale difference of $(M_{x,n})_{n \in \mathbb{N}}$ is

(3.6) $$\Delta_n M_x := M_{x,n} - M_{x,n-1} = \sigma_o(X_n, Z_{x,n-1}) - \ell_\mu,$$

where $(Z_{x,j} := L_j \cdot x)$ is the Markov chain on $\overline{X}^h$ induced by the random walk on G and starting at x. We recall that the (predictable) quadratic variation of $\langle M_x \rangle$ is the unique increasing predictable process such that $(M_{x,n}^2 - \langle M_x \rangle_n)_{n \in \mathbb{N}}$ is a martingale. We have

(3.7) $$\langle M_x \rangle_n = \sum_{j=1}^n \mathbb{E}(\Delta_j M_x^2 | \mathcal{F}_{j-1}) = \sum_{j=1}^n \int (\sigma_0(g, Z_{x,j-1}) - \ell_\mu)^2 d\mu(g).$$

We now come to the main result of this section. Its statement contains the expression

(3.8) $$\sigma_\mu^2 := \iint (\sigma_0(g, x) - \ell_\mu)^2 d\mu(g) d\nu(x),$$

where ν is any μ-stationary probability measure on $\overline{X}^h$ — we will see that the integral does not depend on ν. This constant σ_μ^2 is also the variance appearing in the central limit theorem (1.2) (see proof of [3, Th. 4.7.b] or [22, Th. 1.3]).

Proposition 3.3 (Large deviation estimates for the quadratic variation)

Let μ be a non-elementary probability measure with finite second moment. Then, for every $\varepsilon > 0$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \sup_{x \in X} \mathbb{P}(|\langle M_x \rangle_n - n\sigma_\mu^2| \geqslant n\varepsilon) < 0.$$

To proceed to prove this result, we first observe that we can reformulate the statement as a statement about large deviations for an additive functional of a Markov chain. Indeed, for $x \in \overline{X}^h$, defining

$$(3.9) \qquad \phi(x) := \int (\sigma_0(g,x) - \ell_\mu)^2 d\mu(g),$$

the expression (3.7) shows that

$$\langle M_x \rangle_n = \sum_{j=1}^{n} \phi(Z_{x,j}).$$

Benoist–Quint showed a large deviation estimate for functionals along Markov chains [2, Prop. 3.1], which is a quantitative refinement of Breiman's law of large numbers. In the aforementioned paper, the authors work with continuous functions in the framework of Markov–Feller operators on compact metric spaces. However, in the generality that we work with, we were not able to prove the continuity of ϕ. Note that by the expression (3.4) of the cocycle σ_0, the continuity of ϕ would follow from the continuity of the Gromov-product on the Busemann compactification $\overline{X}^h$. Up to our knowledge, the latter is known in familiar cases including trees and classical hyperbolic spaces but not in our generality (note that by [27, §10] the Gromov-product on the Busemann compactification of a general metric space X may fail to be continuous even if X is proper and geodesic). To overcome this issue, we will adapt the statement of Benoist–Quint by relaxing the continuity assumption.

Proposition 3.4 ([2, Prop. 3.1] modified). — *Let $(Z_n)_{n \in \mathbb{N}}$ be a Markov chain on a state space E, P its Markov operator and $\phi : E \to \mathbb{R}$ a measurable bounded function. Suppose that*

$$(3.10) \qquad \frac{1}{n} \sum_{j=1}^{n} P^j \phi(x) \longrightarrow \ell_\phi \in \mathbb{R},$$

uniformly in $x \in E$. Then the following large deviation estimate holds: for every $\varepsilon > 0$

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{P}_x\left(\sum_{i=1}^{n} \phi(Z_i) \notin [n\ell_\phi - n\varepsilon, n\ell_\phi + n\varepsilon]\right) < 0,$$

uniformly in $x \in E$.

The proof is an adaptation of Benoist–Quint's proof of [2, Prop. 3.1]. We include it for the convenience of the reader.

Proof. — We begin with a general result, which can be seen as a quantitative version of the law of large numbers stated in [4, Th. 1.6]. If $\xi = (\xi_n)_{n \in \mathbb{N}}$ is a sequence of bounded real random variables adapted to a filtration $\mathcal{F} = \{\mathcal{F}_n \mid n \in \mathbb{N}\}$, then for every $\varepsilon' > 0$ and $n \in \mathbb{N}$

$$(3.11) \qquad \mathbb{P}\big(\big|\textstyle\sum_{j=1}^n (\xi_j - \mathbb{E}[\xi_j | \mathcal{F}_{j-1}])\big| \geqslant n\varepsilon'\big) \leqslant 2\exp(-n\varepsilon'^2/8\|\xi\|_\infty^2).$$

Indeed the sequence $\{\xi_n - \mathbb{E}[\xi_n | \mathcal{F}_{n-1}] \mid n \in \mathbb{N}\}$ is a bounded martingale difference sequence with respect to the filtration $\mathcal{F}$ and hence (3.11) follows from Azuma–Hoeffding's concentration inequality for martingales with bounded differences.

In the second step, we show that for every $m \in \mathbb{N}$, $\frac{1}{n}\sum_{j=1}^n \phi(Z_j)$ concentrates around the Cesàro average $\frac{1}{m}\sum_{\ell=1}^m P^\ell \phi(Z_j)$; more precisely for every $\varepsilon' > 0$, $n, m \in \mathbb{N}$, and $x \in X$,

$$(3.12) \quad \mathbb{P}_x\Big(\big|\textstyle\sum_{j=1}^n \big[\phi(Z_j) - \frac{1}{m}\sum_{\ell=1}^m P^\ell \phi(Z_j)\big]\big| \geqslant mn\varepsilon' + 2m\|\phi\|_\infty\Big)$$
$$\leqslant 2m^2 \exp(-n\varepsilon'^2/8\|\phi\|_\infty^2).$$

Indeed, let $1 \leqslant \ell \leqslant m$. We write

$$\sum_{j=\ell+1}^{\ell+n} (\phi(Z_j) - \mathbb{E}_x[\phi(Z_j)|\mathcal{F}_{j-\ell}]) = \sum_{k=0}^{\ell-1}\sum_{j=\ell+1}^{\ell+n} (\mathbb{E}_x[\phi(Z_j)|\mathcal{F}_{j-k}] - \mathbb{E}_x[\phi(Z_j)|\mathcal{F}_{j-k-1}]),$$

where $\mathcal{F}$ is the filtration induced by the Markov chain. For each $k \in \{0, \cdots, \ell-1\}$, we apply (3.11) with the sequence of random variables $\{\xi_{j,k} = \mathbb{E}_x[\phi(Z_j)|\mathcal{F}_{j-k}] \mid j \in \mathbb{N}\}$ which are adapted to the filtration $\{\mathcal{F}_{j-k} \mid j \in \mathbb{N}\}$ and bounded by $\|\phi\|_\infty$. Combining the resulting ℓ estimates, we obtain that

$$\mathbb{P}_x\big(\textstyle\sum_{j=\ell+1}^{\ell+n} (\phi(Z_j) - \mathbb{E}_x[\phi(Z_j)|\mathcal{F}_{j-\ell}]) \geqslant \ell n\varepsilon'\big) \leqslant 2\ell\exp(-n\varepsilon'^2/8\|\phi\|_\infty^2).$$

Noticing that $\mathbb{E}_x[\phi(Z_j)|\mathcal{F}_{j-\ell}] = P^\ell\phi(Z_{j-\ell})$, the previous estimate gives (after killing the boundary terms using the boundedness of ϕ) that

$$\mathbb{P}_x\big(\big|\textstyle\sum_{j=1}^n \big[\phi(Z_j) - P^\ell\phi(Z_j)\big]\big| \geqslant \ell n\varepsilon' + 2\ell\|\phi\|_\infty\big) \leqslant 2\ell\exp(-n\varepsilon'^2/8\|\phi\|_\infty^2).$$

Estimate (3.12) immediately follows.

Finally, let $\varepsilon > 0$ be given. The uniform convergence (3.10) yields an integer m_0 such that $\mathbb{P}_x$-almost surely for every $j \in \mathbb{N}$,

$$(3.13) \qquad\qquad \frac{1}{m_0}\sum_{\ell=1}^{m_0} P^\ell\phi(Z_j) \in [\ell_\phi - \varepsilon/3, \ell_\phi + \varepsilon/3].$$

Plugging (3.13) into (3.12) with $m = m_0$ and $\varepsilon' = \varepsilon/3m_0$ gives some constant $C(\varepsilon) > 0$ such that for every $n \geqslant 6m_0\|\phi\|_\infty/\varepsilon$,

$$\mathbb{P}_x\big(\textstyle\sum_{i=1}^n \phi(Z_i) \in [n\ell_\phi - n\varepsilon, n\ell_\phi + n\varepsilon]\big) \leqslant \exp(-nC(\varepsilon)),$$

concluding the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 3.5. — If E is a compact metric space, P a Markov Feller operator and ϕ is a continuous function which has a unique average with respect to stationary measures on E, then (3.10) is fulfilled. As mentioned earlier, this is the case, for instance, for random walks on trees, classical hyperbolic spaces and also for strongly irreducible and proximal random walks on projective spaces (see for instance [4]).

We now check that (3.10) is satisfied for our function ϕ defined in (3.9) and the Markov operator $P = P_\mu$ of the Markov chain on $\overline{X}^h$ induced by the random walk on G (see Section 3.2).

Lemma 3.6. — Let $\phi : \overline{X}^h \to \mathbb{R}$ as defined in (3.9). Then the sequence of functions $f_n(x) := \frac{1}{n} \sum_{j=1}^n P_\mu^j \phi(x)$ converges uniformly on $\overline{X}^h$ to $\sigma_\mu^2 \geqslant 0$. The limit σ_μ^2 can be expressed as

$$(3.14) \qquad \sigma_\mu^2 := \iint (\sigma_0(g,x) - \ell_\mu)^2 d\mu(g) d\nu(x),$$

where ν is any μ-stationary measure on $\overline{X}^h$.

Remark 3.7. — It follows from (3.14) and the argument in the proof of [3, Th. 4.7.b] that $\sigma_\mu = 0$ if any only if there exists a constant $C > 0$ such that for every $n \in \mathbb{N}$ and $g \in \operatorname{supp}(\mu^{*n})$, we have $|\kappa(g) - n\ell_\mu| \leqslant C$. It follows that $\sigma_\mu > 0$ if any only if μ is non-arithmetic. Here, a probability measure μ on $\operatorname{Isom}(X)$ is said to be non-arithmetic if there exists $n \in \mathbb{N}$ and $g, g' \in \operatorname{supp}(\mu^{*n})$ such that $\tau(g) \neq \tau(g')$ where τ is the translation distance, $\tau(g) = \lim_{n \to \infty} \kappa(g^n)/n$.

The proof of the previous lemma is based on showing that $f_n(x) - f_n(y)$ converges uniformly to zero (see (3.17)), which imposes the limit to be the average σ_μ^2 as defined in (3.14). To prove this, we express $f_n(x)$ as the variance of $M_{x,n}/\sqrt{n}$ (see (3.18)). Using Burkholder's inequalities, the proof boils down to showing deviation inequalities for $\sigma(L_n, x) - \sigma(L_n, y)$ uniformly in $x, y \in \overline{X}^h$ (see (3.22)). For the latter fact, we will use the following lemma which is a direct consequence of uniform punctual deviation estimates given in [6, Prop. 2.12].

Lemma 3.8 (Uniform punctual deviations). — Keep the hypotheses of Proposition 3.3. Then there are constants $C, \beta > 0$ such that for any $k \in \mathbb{N}$ and any $x \in \overline{X}^h$, $R > 0$, we have

$$\mathbb{P}(\kappa(L_k) - \sigma(L_k, x) > R) \leqslant Ce^{-\beta R}.$$

Proof. — Notice that for $g \in \operatorname{Isom}(M)$ and $x \in \overline{X}^h$, by (3.3) we have $\kappa(g) - \sigma(g,x) = 2(g^{-1}o|x)$. In particular, when $x \in X$, the statement precisely corresponds to [6, Prop. 2.12] applied with the image $\check{\mu}$ of μ by the map $g \mapsto g^{-1}$ on $\operatorname{Isom}(X)$. To extend it to $\overline{X}^h$, given $x \in \overline{X}^h$, let x_n be a sequence in X such that $x_n \to x$ in $\overline{X}^h$. By continuity of $\sigma(g, \cdot)$, we have $\kappa(g) - \sigma(g,x) = \lim_{n \to \infty} 2(g^{-1}o|x_n)$. Therefore, given $R > 0$,

$$(3.15) \qquad \mu^{*k}\{g \mid \kappa(g) - \sigma(g,x) > R\} = \check{\mu}^{*k}\{g \mid \lim_{n \to \infty}(go|x_n) > R/2\},$$

for every $k \in \mathbb{N}$. Denoting by $h_n(\cdot)$, the map $g \mapsto \mathbb{1}_{(go|x_n)>R/2}$, by (3.15) we have

$$\mathbb{P}(\kappa(L_k) - \sigma(L_k, x) > R) = \int \lim_{n\to\infty} h_n(g) d\breve{\mu}^{*k}(g) = \lim_{n\to\infty} \breve{\mu}^{*k}\{g \mid (go|x_n) > R/2\},$$

where we used dominated convergence in the last equality. Hence the statement follows from [6, Prop. 2.12]. $\hfill\square$

Proof of Lemma 3.6. — First, we reduce the problem to showing that

$$(3.16) \qquad\qquad f_n(x) - f_n(y) \longrightarrow 0$$

uniformly in x and y in $\overline{X}^h$. Indeed, let us assume for a while that (3.16) holds. Fix any μ-stationary measure ν on $\overline{X}^h$ (the latter exists by compactness of $\overline{X}^h$). We have for every $n \in \mathbb{N}$,

$$(3.17) \qquad f_n(x) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[\phi(Z_j)|Z_0 = x] = \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[\int (\sigma_0(g, L_j \cdot x) - \ell_\mu)^2 d\mu(g)].$$

Since ν is μ^{*n}-stationary for every $n \in \mathbb{N}$, we deduce that for every $n \in \mathbb{N}$,

$$\int_{\overline{X}^h} f_n(x) d\nu(x) = \iint_{G \times \overline{X}^h} (\sigma_0(g, x) - \ell_\mu)^2 d\mu(g) d\nu(x) := \sigma_{\mu,\nu}^2.$$

Let $\varepsilon > 0$ and $y \in \overline{X}$. We can find n_0 depending only on ε such that for every $n \geqslant n_0$ and for every $x \in \overline{X}^h$, $f_n(x) - \varepsilon \leqslant f_n(y) \leqslant f_n(x) + \varepsilon$. Integrating on both sides with respect $d\nu(x)$, we obtain that $|f_n(y) - \sigma_{\mu,\nu}^2| \leqslant \varepsilon$ for every $n \geqslant n_0$, concluding the proof of the uniform convergence of the sequence of functions $(f_n)_{n\in\mathbb{N}}$ towards $\sigma_{\mu,\nu}^2$. It also shows that $\sigma_{\mu,\nu}^2$ is independent of the choice of the stationary measure ν.

From now on, we focus on showing the convergence (3.16) uniformly in $x, y \in \overline{X}^h$. Since $(M_{x,n}^2 - \langle M_x \rangle_n)_{n\in\mathbb{N}}$ is a martingale starting at zero, we have that $\mathbb{E}[\langle M_x \rangle_n] = \mathbb{E}[M_{x,n}^2]$ for every $n \in \mathbb{N}$ so that by (3.17):

$$(3.18) \qquad\qquad f_n(x) = \frac{1}{n} \mathbb{E}[M_{x,n}^2] = \mathbb{E}\left[\left(M_{x,n}/\sqrt{n}\right)^2\right].$$

Let us check that the sequence $\{(M_{x,n}/\sqrt{n})^2 \mid n \in \mathbb{N}, x \in \overline{X}^h\}$ is uniformly bounded in L^p for every $p > 1$; and hence in particular uniformly integrable. Indeed, by Burkholder's inequality ([7, Th. 9]), we have for every $k > 2$,

$$(3.19) \qquad\qquad \mathbb{E}[|M_{x,n}|^k] \leqslant C_k \mathbb{E}[[M_x]_n^{k/2}],$$

where $C_k > 0$ is a constant depending only on k, $[M_x]_n = \sum_{j=1}^{n} (\Delta_j M_x)^2$ is the quadratic variation of M_x. By Jensen's inequality, we have

$$[M_x]_n^{k/2} \leqslant n^{k/2-1} \sum_{j=1}^{n} |\Delta_j M_x|^k,$$

so that

$$(3.20) \qquad \mathbb{E}[[M_x]_n^{k/2}] \leqslant n^{k/2-1} \sum_{j=1}^{n} \mathbb{E}[|\Delta_j M_x|^k] \leqslant n^{k/2} \sup_{j\in\mathbb{N}} \mathbb{E}[|\Delta_j M_x|^k].$$

Remembering that $\sigma(g, x) = \sigma_0(g, x) - \psi(g \cdot x) + \psi(x)$, $|\sigma(g, x)| \leqslant \kappa(g)$, and that ψ is bounded on $\overline{X}^h$, we get that for every $n \in \mathbb{N}$,

$$(3.21) \qquad |\Delta_n M_x| = |\sigma_0(X_n, L_{n-1} \cdot x) - \ell_\mu| \leqslant \kappa(X_n) + \ell_\mu + 2\|\psi\|_\infty.$$

Since the X_i's have the same distribution, by plugging (3.21) and (3.20) in (3.19) we get

$$\mathbb{E}[M_{x,n}^k] \leqslant C_k n^{k/2} \mathbb{E}[|\kappa(X_1) + \ell_\mu + 2\|\psi\|_\infty|^k].$$

The right-hand-side is finite (since μ has a finite moment of any order $k > 2$) and does not depend neither on n nor on x, showing the boundedness in $L^{k/2}$ of $(M_{x,n}/\sqrt{n})^2$ uniformly in n and x.

Let now $\varepsilon > 0$. It follows from the uniform integrability of the family $M_{x,n}/\sqrt{n}$ that there exists $L(\varepsilon) > 0$ such that for every $n \in \mathbb{N}$ and $x, y \in \overline{X}^h$ we have

$$\mathbb{P}\big(\max\{|M_{x,n}|/\sqrt{n}, |M_{y,n}|/\sqrt{n}\} > L(\varepsilon)\big) < \varepsilon.$$

Using Lemma 3.8 together with the fact that σ differs from σ_0 by a bounded function on $\overline{X}^h$, we obtain some $T(\varepsilon) > 0$ such that for every $n \in \mathbb{N}$, $x, y \in \overline{X}^h$,

$$(3.22) \qquad \mathbb{P}(|\sigma_0(L_n, x) - \sigma_0(L_n, y)| > T(\varepsilon)) \leqslant \varepsilon.$$

Hence, if $A_{x,y,n}$ denotes the event

$$A_{x,y,n} := \{|\sigma_0(L_n, x) - \sigma_0(L_n, y)| > T(\varepsilon)\} \cup \{\max\{|M_{x,n}|/\sqrt{n}, |M_{y,n}|/\sqrt{n}\} > L(\varepsilon)\},$$

we have for every $n \in \mathbb{N}$, $x, y \in \overline{X}^h$ that $\mathbb{P}(A_{x,y,n}) < 2\varepsilon$. Now we write

$$|f_n(x) - f_n(y)| \leqslant \underbrace{\mathbb{E}\Big[\Big|\big(M_{x,n}/\sqrt{n}\big)^2 - \big(M_{y,n}/\sqrt{n}\big)^2\Big|\mathbb{1}_{A_{x,y,n}}\Big]}_{a_{x,y,n}}$$

$$+ \underbrace{\mathbb{E}\Big[\Big|\big(M_{x,n}/\sqrt{n}\big)^2 - \big(M_{y,n}/\sqrt{n}\big)^2\Big|\mathbb{1}_{A_{x,y,n}^C}\Big]}_{b_{x,y,n}}.$$

Let us estimate $a_{x,y,n}$. By Cauchy–Schwarz inequality, we have for every $n \in \mathbb{N}$,

$$a_{x,y,n}^2 \leqslant 2 \max_{x,y} \mathbb{E}\big[\big(M_{x,n}/\sqrt{n}\big)^4\big] \mathbb{P}(A_{x,y,n}) \leqslant C_4 \varepsilon,$$

where $C_4 > 0$ is a constant independent of n, x and ε; guaranteed by the uniform boundedness in L^4 shown at the beginning of the proof. Finally, we estimate $b_{x,y,n}$. Since the function $x \mapsto x^2$ is uniformly continuous on $[-L(\varepsilon), L(\varepsilon)]$, we can find $\delta(\varepsilon) > 0$ such that $|t^2 - t'^2| < \varepsilon$ whenever $|t - t'| < \delta(\varepsilon)$ and $\max\{|t|, |t'|\} \leqslant L(\varepsilon)$. Let $n_0(\varepsilon) \in \mathbb{N}$ be such that $T(\varepsilon)/\sqrt{n_0(\varepsilon)} < \delta(\varepsilon)$. From the definition of the event $A_{x,y,n}^C$, we deduce that for every $n \geqslant n_0(\varepsilon)$, $x, y \in \overline{X}^h$, $b_{x,y,n} < \varepsilon$. Hence for $n \geqslant n_0(\varepsilon)$, $x, y \in \overline{X}^h$

$$|f_n(x) - f_n(y)| < \sqrt{C_4 \varepsilon} + \varepsilon,$$

which finishes the proof of the uniform convergence (3.16). $\qquad\square$

We end this section with the following consequence of Proposition 3.3. In the statement below, for every $x \in \overline{X}^h$, we use the transform G_n^a introduced in (2.1) associated to the martingale M_x. To ease the notation, we omit the dependence on x in G_n^a.

Corollary 3.9. — *Suppose μ has finite exponential moment. Then for every $x \in \overline{X}^h$,*

$$\lim_{a \to +\infty} \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}[\exp(-\lambda(G_n^a - n\sigma_\mu^2))] = 0.$$

Proof. — Let $x \in \overline{X}^h$. The result will follow from Cauchy–Schwarz inequality and the following two estimates

$$(3.23) \qquad \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}[\exp(-\lambda(\langle M_x \rangle_n - n\sigma_\mu^2))] \leqslant 0$$

and

$$(3.24) \qquad \limsup_{a \to +\infty} \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}[\exp(\lambda(\langle M_x \rangle_n - G_n^a))] \leqslant 0.$$

(i) We start by proving (3.23). To ease the notation, let $Y_n = \langle M_x \rangle_n - n\sigma_\mu^2$. Let $\varepsilon > 0$. By Proposition 3.3, there exists $\alpha(\varepsilon) > 0$ and $n_0(\varepsilon) \in \mathbb{N}$ such that for every $n \geqslant n_0(\varepsilon)$,

$$\mathbb{P}(Y_n \leqslant -n\varepsilon) \leqslant \exp(-\alpha(\varepsilon)n).$$

Noticing that $Y_n \geqslant -n\sigma_\mu^2$, we write

$$Y_n = Y_n \mathbb{1}_{Y_n \geqslant 0} + Y_n \mathbb{1}_{-n\varepsilon \leqslant Y_n \leqslant 0} + Y_n \mathbb{1}_{-n\sigma_\mu^2 \leqslant Y_n \leqslant -n\varepsilon}.$$

Since $\mathbb{E}[\exp(-\lambda Y_n)\mathbb{1}_{-n\varepsilon \leqslant Y_n \leqslant 0}] \leqslant \exp(n\lambda\varepsilon)$ and, for $n \geqslant n_0(\varepsilon)$,

$$\mathbb{E}[\exp(-\lambda Y_n)\mathbb{1}_{-n\sigma_\mu^2 \leqslant Y_n \leqslant -n\varepsilon}] \leqslant \exp(n\lambda\sigma^2)\mathbb{P}(Y_n \leqslant -n\varepsilon) \leqslant \exp(n(\sigma_\mu^2\lambda - \alpha(\varepsilon))),$$

we get that for every $n \geqslant n_0(\varepsilon)$,

$$\mathbb{E}[\exp(-\lambda(\langle M_x \rangle_n - n\sigma_\mu^2))] \leqslant 1 + \exp(n\lambda\varepsilon) + \exp(n\sigma_\mu^2\lambda - \alpha(\varepsilon)n).$$

Keeping ε and $\lambda > 0$ (small enough) fixed, we let $n \to +\infty$ and deduce that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}[\exp(-\lambda(\langle M_x \rangle_n - n\sigma_\mu^2))] \leqslant \max\{\lambda\varepsilon, \lambda\sigma_\mu^2 - \alpha(\varepsilon)\}.$$

Since $\alpha(\varepsilon) > 0$, we get by letting $\lambda \to 0^+$ (while keeping ε fixed) that

$$\limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}[\exp(-\lambda(\langle M_x \rangle_n - n\sigma_\mu^2))] \leqslant \varepsilon.$$

Letting $\varepsilon \to 0$, we conclude that

$$\limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}[\exp(-\lambda(\langle M_x \rangle_n - n\sigma_\mu^2))] \leqslant 0.$$

This shows (3.23).

(ii) Finally, we show (3.24). Using the expression (2.1) for G_n^a, we see that

$$\langle M_x \rangle_n - G_n^a = \sum_{i=1}^n \mathbb{E}[(\Delta_i M_x)^2 \mathbb{1}_{|\Delta_i M_x| > a} | \mathcal{F}_{i-1}] + a \sum_{i=1}^n |\Delta_i M_x| \mathbb{1}_{|\Delta_i M_x| \geqslant a}.$$

Observe that by the expression of our martingale difference (3.6) and by the decomposition (3.4), we have a.s. for every $i \in \mathbb{N}$,

$$(3.25) \qquad |\Delta_i M_x| = |\sigma_0(X_i, L_{i-1} \cdot x) - \ell_\mu| \leqslant \kappa(X_i) + \ell_\mu + \|\psi\|_\infty := \zeta_i.$$

Since the ζ_i's have the same distribution,

$$\mathbb{E}[(\Delta_i M_x)^2 \mathbb{1}_{|\Delta_i M_x| > a} | \mathcal{F}_{i-1}] \leqslant \mathbb{E}(\zeta_1^2 \mathbb{1}_{\zeta_1 > a}) := h_\mu(a).$$

Observe that the constant $h_\mu(a)$ is independent of i. Since μ has finite second moment, we deduce that

$$(3.26) \quad \limsup_{a \to \infty} \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}\big[e^{\lambda \sum_{i=1}^n \mathbb{E}[(\Delta_i M_x)^2 \mathbb{1}_{|\Delta_i M_x| > a} | \mathcal{F}_{i-1}]} \big]$$
$$\leqslant \limsup_{a \to \infty} h_\mu(a) = 0.$$

On the other hand, for every $a > 0$, the random variables $a \zeta_i \mathbb{1}_{|\zeta_i| > a}$ are i.i.d random variables. Denote by ζ_a their common distribution and by Λ_{ζ_a} the Laplace transform of ζ_a. The latter is differentiable at 0 for every $a > 0$ as ζ_1 has finite exponential moment (because μ has finite exponential moment). It follows that

$$(3.27) \quad \limsup_{a \to +\infty} \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}\big[e^{\lambda a \sum_{i=1}^n |\Delta_i M_x| \mathbb{1}_{|\Delta_i M_x| \geqslant a}} \big]$$
$$\leqslant \limsup_{a \to +\infty} \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}\big[e^{\lambda \sum_{i=1}^n a |\zeta_i| \mathbb{1}_{|\zeta_i| \geqslant a}} \big] = \limsup_{a \to +\infty} \Lambda'_{\zeta_a}(0).$$

But $\Lambda'_{\zeta_a}(0) = a \mathbb{E}[|\zeta_1| \mathbb{1}_{|\zeta_1| > a}] \leqslant \mathbb{E}[\zeta_1^2 \mathbb{1}_{|\zeta_1| > a}] \xrightarrow[a \to \infty]{} 0$. This concludes the proof of (3.24) and hence the corollary. □

4. Proof of the main result

Having established the submartingale transform from Section 2 and the exponential decay of large deviation probabilities of the predictable quadratic variation from Section 3.2, we are now ready to give the proofs of Theorem 1.1 and Corollary 1.3. In fact, we will prove a slightly more general version given by Theorem 4.1 below.

4.1. STATEMENT OF THE MAIN RESULT. — To state the more general version of Theorem 1.1, we recall and introduce some notation. We are given a separable geodesic Gromov-hyperbolic space X with a fixed based point $o \in X$. The Busemann cocycle $\sigma : \mathrm{Isom}(X) \times \overline{X}^h \to \mathbb{R}$ with respect to the base point o is as defined in Section 3.1. Given a countably supported probability measure μ on $\mathrm{Isom}(X)$ and $x \in \overline{X}^h$,

we define the upper Λ_x^+ and lower Λ_x^- limit Laplace transforms as

$$\Lambda_x^+(\lambda) := \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda \sigma(L_n, x)}]$$

and

$$\Lambda_x^-(\lambda) := \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda \sigma(L_n, x)}].$$

Whenever μ has finite exponential moment both functions have values in $\mathbb{R}$ in a neighborhood of $0 \in \mathbb{R}$.

We will omit sub/super-scripts when $x \in X$, indeed, for every $x \in X$, we have $\Lambda_x^+ \equiv \Lambda_x^- \equiv \Lambda_o$ (since $|\sigma(g, x) - \sigma(g, y)| \leqslant 2d(x, y)$ for $g \in G$ and $x, y \in X$). This common function $\Lambda = \Lambda_o$ is the notation used in Theorem 1.1 where we work with the basepoint $x = o \in X$.

THEOREM 4.1. — *Let (X, d) be a separable geodesic Gromov-hyperbolic space and μ a non-elementary probability measure on $\mathrm{Isom}(X)$. Suppose that μ has a finite exponential moment. Then for every $x \in \overline{X}^h$,*

$$\lim_{\lambda \to 0} \frac{\Lambda_x^-(\lambda) - \ell\lambda}{\lambda^2} = \lim_{\lambda \to 0} \frac{\Lambda_x^+(\lambda) - \ell\lambda}{\lambda^2} = \frac{\sigma_\mu^2}{2}.$$

Sections 4.2 and 4.3 are devoted to the proof of Theorem 4.1.

4.2. PROOF OF THE LOWER BOUND. — Here we prove the following.

PROPOSITION 4.2. — *Keep the setting of Theorem 4.1. Then, for every $x \in \overline{X}^h$*

$$\lim_{\lambda \to 0} \frac{\Lambda_x^-(\lambda) - \ell_\mu\lambda}{\lambda^2} \geqslant \frac{\sigma_\mu^2}{2}.$$

Proof of Proposition 4.2. — Given a probability measure μ as in the statement and $x \in \overline{X}^h$, let M_x be the martingale given by Lemma 3.1. It satisfies

$$\sigma(L_n, x) - n\ell_\mu = M_{x,n} + O_{x,n}(1),$$

for every $n \in \mathbb{N}$, where $O_{x,n}(1)$ is a random variable that is bounded (in absolute value) uniformly in $x \in \overline{X}^h$ and $n \in \mathbb{N}$. Let $\sigma_\mu^2 > 0$ be as defined in (3.14). Let $x \in \overline{X}^h$ be fixed for the rest of the proof. For every $\lambda \in \mathbb{R}$, we have

$$\begin{aligned}
\Lambda_x^-(\lambda) - \lambda\ell_\mu &= \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda(\sigma(L_n, x) - n\ell_\mu)}] \\
&= \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda(M_{x,n} + O_{x,n}(1))}] \\
&= \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\lambda M_{x,n}}],
\end{aligned}$$

where we used the fact that the random variables $O_{x,n}(1)$ are bounded below and above uniformly in $n \in \mathbb{N}$. Notice that since μ has finite exponential moment, for every λ in a neighborhood of $0 \in \mathbb{R}$ (independent of $x \in \overline{X}^h$), the last quantity in the above displayed equation is finite.

We first prove that

$$(4.1) \qquad \frac{\sigma_\mu^2}{2} \leqslant \liminf_{\lambda \to 0^+} \frac{\Lambda_x^-(\lambda) - \lambda \ell_\mu}{\lambda^2}.$$

Let $n \in \mathbb{N}$, $a > 0$, and $\lambda > 0$ small enough. By Proposition 2.2, we have

$$1 \leqslant \mathbb{E}\Big[\exp\Big(\lambda M_{x,n} - \frac{\mathfrak{f}(a\lambda)}{a^2} G_n^a\Big)\Big],$$

where, we recall,

$$G_n^a = \sum_{i=1}^n \mathbb{E}[(\Delta_i M_x)^2 1_{|\Delta_i M_x| \leqslant a} | \mathcal{F}_{i-1}] - a \sum_{i=1}^n |\Delta_i M_x| 1_{|\Delta_i M_x| \geqslant a}.$$

Let $p > 1$. By Hölder inequality, we get

$$1 \leqslant \mathbb{E}[\exp(p\lambda M_{x,n})]^{1/p} \mathbb{E}\Big[\exp\Big(-q\,\frac{\mathfrak{f}(a\lambda)}{a^2} G_n^a\Big)\Big]^{1/q},$$

where $q \geqslant 1$ satisfies $1/p + 1/q = 1$. Taking logarithm and dividing by $n\lambda^2$, adding and subtracting the term $(\mathfrak{f}(a\lambda)/(a\lambda)^2)\sigma_\mu^2$ gives

$$(4.2) \quad 0 \leqslant \frac{1}{np\lambda^2} \log \mathbb{E}[\exp(p\lambda M_{x,n})]$$

$$+ \frac{1}{nq\lambda^2} \log \mathbb{E}\Big[\exp\Big(-q\,\frac{\mathfrak{f}(a\lambda)}{a^2}\,(G_n^a - n\sigma_\mu^2)\Big)\Big] - \frac{\mathfrak{f}(a\lambda)}{(a\lambda)^2}\sigma_\mu^2.$$

Using the elementary fact $\liminf_{n \to +\infty}(a_n + b_n) \leqslant \liminf_n a_n + \limsup_n b_n$ for real sequences a_n and b_n, letting $n \to \infty$ in (4.2), we get

$$0 \leqslant p\frac{\Lambda_x^-(\lambda p) - \lambda p \ell_\mu}{(\lambda p)^2} + \frac{1}{q\lambda^2} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}\Big[\exp\Big(-q\,\frac{\mathfrak{f}(a\lambda)}{a^2}\,(G_n^a - n\sigma_\mu^2)\Big)\Big] - \frac{\mathfrak{f}(a\lambda)}{(a\lambda)^2}\sigma_\mu^2.$$

Letting $\lambda \to 0^+$ while noting that $\mathfrak{f}(a\lambda) \underset{\lambda \to 0}{\sim} (a\lambda)^2/2$ and in particular $\eta = q\mathfrak{f}(a\lambda)/a^2 \xrightarrow[\lambda \to 0]{} 0^+$, we obtain:

$$\frac{\sigma_\mu^2}{2} \leqslant p \liminf_{\lambda \to 0^+} \frac{\Lambda_x^-(\lambda) - \lambda \ell_\mu}{\lambda^2} + \frac{1}{2} \limsup_{\eta \to 0^+} \frac{1}{\eta} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{E}[\exp(-\eta(G_n^a - n\sigma_\mu^2))].$$

Letting $a \to \infty$, we deduce from Corollary 3.9 that we have

$$\frac{\sigma_\mu^2}{2} \leqslant p \liminf_{\lambda \to 0^+} \frac{\Lambda_x^-(\lambda) - \lambda \ell_\mu}{\lambda^2}.$$

The desired inequality (4.1) is now proved by taking $p \to 1$.

The inequality

$$(4.3) \qquad \frac{\sigma_\mu^2}{2} \leqslant \liminf_{\lambda \to 0^-} \frac{\Lambda_x^-(\lambda) - \lambda \ell_\mu}{\lambda^2}$$

is proved in precisely the same way replacing the martingale M_x by the martingale $-M_x$ using the fact that both martingales have same transforms G^a. This completes the proof of Proposition 4.2. □

4.3. **Proof of the upper bound.** — Here we prove the following.

Proposition 4.3. — *Keep the setting of Theorem* 4.1. *Then, for every* $x \in \overline{X}^h$

$$(4.4) \qquad \lim_{\lambda \to 0} \frac{\Lambda_x^+(\lambda) - \ell_\mu \lambda}{\lambda^2} \leqslant \frac{\sigma_\mu^2}{2}.$$

The proof is based on showing that for large $n \in \mathbb{N}$ the random variable $\frac{1}{n}\sigma(L_n, x) - \ell_\mu$ has a subgaussian behaviour in a neighborhood of 0. This is shown in the following proposition which controls the limit Laplace transform of the sequence of random variables $\frac{1}{n}\sigma(L_n, x) - \ell_\mu$. The proof is based on the martingale decomposition given in Lemma 3.1 and standard techniques for concentration results for martingales (in particular [31, Th. 2.19]). With the notation of Section 3, the main tool for the proof of Proposition 4.3 is the following.

Proposition 4.4. — *Let*

$$(4.5) \qquad v(\mu) := \sup_{\xi \in \overline{X}^h} \mathbb{E}\Big[(\sigma_0(X_1, \xi) - \ell_\mu)^2\Big] = \sup_{\xi \in \overline{X}^h} \mathbb{E}[M_{\xi,1}^2].$$

Then there exists $C > 0$ *such that for every* $\varepsilon > 0$, *there exists* $b > 0$ *such that for every* $|\lambda| < v(\mu)/b$, *every* $n \in \mathbb{N}$ *and every* $x \in \overline{X}^h$,

$$(4.6) \qquad \mathbb{E}\Big[e^{\lambda(\sigma(L_n, x) - n\ell_\mu)}\Big] \leqslant \exp\Big(\frac{\lambda^2(v(\mu) + \varepsilon)n}{2} + C|\lambda|\Big).$$

This proposition will yield (4.4) with σ_μ^2 replaced by the larger quantity $v(\mu)$. To obtain (4.4), we will use an acceleration technique speeding up the random walk, see the proof of Proposition 4.3.

We now proceed with proving Proposition 4.4. The proof is based on the following control of the conditional expectation of the martingale difference ΔM_x:

Lemma 4.5. — *For every* $\varepsilon > 0$, *there exists a constant* $b > 0$ *such that for every* $|\lambda| < b$, $n \in \mathbb{N}$ *and* $x \in \overline{X}^h$, *the following inequality holds almost surely:*

$$\mathbb{E}\left[\exp(\lambda \Delta_n M_x) \,|\, \mathscr{F}_{n-1}\right] \leqslant \exp\Big(\frac{\lambda^2(v(\mu) + \varepsilon)}{2}\Big).$$

Proof. — By expanding the expression (3.5) of the martingale M_x and taking conditional expectation, it suffices to show that for every $\varepsilon > 0$, there exists a constant $b > 0$ such that for every $\xi \in \overline{X}^h$ and $|\lambda| < b$

$$(4.7) \qquad \int e^{\lambda(\sigma_0(g, \xi) - \ell_\mu)} d\mu(g) \leqslant \exp\Big(\frac{\lambda^2(v(\mu) + \varepsilon)}{2}\Big).$$

Using the exponential moment assumption on μ, let $\alpha > 0$ be such that

$$\int e^{\alpha\kappa(g)} d\mu(g) < \infty.$$

Thanks to (3.4), we have that for every $g \in \text{Isom}(X)$ and $\xi \in \overline{X}^h$, $|\sigma_0(g, x) - \ell_\mu| \leqslant \kappa(g) + 2(\|\psi\| + \ell_\mu)$. Therefore, for every $|\lambda| < \alpha$, using dominated convergence, we

have

$$\int e^{\lambda(\sigma_0(g,\xi)-\ell_\mu)}d\mu(g) = 1 + \frac{\lambda^2}{2}\mathbb{E}[(\sigma_0(X_1,\xi)-\ell_\mu)^2] + \sum_{k=3}^{\infty}\frac{\lambda^k}{k!}\mathbb{E}[(\sigma_0(X_1,\xi)-\ell_\mu)^k],$$

where we have used the fact that $\sigma_0(X_1,\xi)-\ell_\mu = M_{\xi,1}$ has mean zero (as the cocycle σ_0 has constant drift). Now using again the fact that $|\sigma_0(g,x)-\ell_\mu| \leqslant \kappa(g) + 2(\|\psi\| + \ell_\mu)$ and that μ has finite exponential moment, we get that there exists $C > 0$ (independent of ξ) such that for every $|\lambda| < \alpha$

$$\int e^{\lambda(\sigma_0(g,\xi)-\ell_\mu)}d\mu(g) \leqslant 1 + \frac{\lambda^2 v(\mu)}{2} + C\lambda^3.$$

This readily implies (4.7) and hence finishes the proof. $\qquad\square$

Proof of Proposition 4.4. — By the tower property of the conditional expectation, we deduce from Lemma 4.5 that for every $\varepsilon > 0$, $|\lambda| < b$ (where $b = b(\mu,\varepsilon)$ is given by the aforementioned lemma), every $n \in \mathbb{N}$ and $x \in \overline{X}^h$,

$$\mathbb{E}\big[e^{\lambda M_{x,n}}\big] = \mathbb{E}\big[e^{\lambda M_{x,n-1}}\mathbb{E}[e^{\lambda\Delta M_{x,n}}|\mathcal{F}_{n-1}]\big] \leqslant \exp\Big(\frac{\lambda^2(v(\mu)+\varepsilon)}{2}\Big)\mathbb{E}\big[e^{\lambda M_{x,n-1}}\big].$$

Iterating the same process, we deduce that

$$\mathbb{E}[\exp(\lambda M_{x,n})] \leqslant \exp\Big(\frac{\lambda^2(v(\mu)+\varepsilon)n}{2}\Big).$$

Finally, recall that $\sigma(L_n,x) - n\ell_\mu = M_{x,n} + R_{x,n}$ where $|R_{x,n}| \leqslant 2\|\psi\|_\infty := C$. This finishes the proof of the proposition. $\qquad\square$

A remark on the proof Proposition 4.4 is in order.

REMARK 4.6. — Given a martingale M with unbounded differences, controlling various quantities involving the *conditional expectation* of the martingale difference sequence ΔM is generally an important step to prove concentration results for the martingale M; see the works of de La Peña [30], Dzhaparidze–van Zanten [13], Fan–Grama–Liu [14] and Liu–Watbled [24] who prove Bennett–Bernstein type concentration inequalities generalizing results of Freedman [15] to the case of unbounded differences. Proposition 4.4 avoids using these more sophisticated results thanks to Lemma 4.5 which, exploiting the special form of our martingales (namely, coming from an i.i.d. random walk on a group), gives a deterministic bound for the exponential of the conditional expectation.

We are now ready to give:

Proof of Proposition 4.3. — Using Proposition 4.4 and taking logarithm and dividing by n on both sides of (4.6), letting first $n \to +\infty$, then $\lambda \to 0$, and finally $\varepsilon \to 0$, we get that

$$(4.8) \qquad\qquad \limsup_{\lambda \to 0} \frac{\Lambda_x^+(\lambda) - \lambda\ell_\mu}{\lambda^2} \leqslant \frac{v(\mu)}{2}.$$

This yields (4.4) with σ_μ^2 replaced with the larger quantity $v(\mu)$. We now employ an acceleration trick. More precisely, consider, for every $k \in \mathbb{N}$, the probability measure μ^{*k} (distribution of L_k), which is a non-elementary probability measure with finite exponential moment. Denote by $\Lambda(\mu^{*k}, .)$ the Laplace transform based at $x = o$ for the μ^{*k}-random walk $(L_{nk})_{n \in \mathbb{N}}$. In particular, $\Lambda(\mu, .) = \Lambda(.)$. Applying (4.8) for the μ^{*k}-random walk, we deduce that for every $k \geqslant 1$,

$$\limsup_{\lambda \to 0} \frac{\Lambda(\mu^{*k}, \lambda) - \lambda \ell_{\mu^{*k}}}{\lambda^2} \leqslant \frac{v(\mu^{*k})}{2}$$

and $\Lambda(\mu^{*k}, .) = k\Lambda(.)$. Hence, for every $k \geqslant 1$,

$$\limsup_{\lambda \to 0} \frac{\Lambda(\lambda) - \ell_\mu \lambda}{\lambda^2} \leqslant \frac{v(\mu^{*k})}{2k}.$$

It remains to check that

(4.9)
$$\lim_{k \to \infty} \frac{v(\mu^{*k})}{k} = \sigma_\mu^2.$$

By definition of $v(\mu^{*k})$ given in (4.5) and using (3.18) (with the notation of Lemma 3.6), we get that for every $k \geqslant 1$,

$$\frac{v(\mu^{*k})}{k} = \sup_{\xi \in \overline{X}^h} f_k(\xi).$$

Finally, the uniform convergence given in Lemma 3.6 for the sequence $(f_k)_{k \in \mathbb{N}}$ implies (4.9) and finishes the proof of the proposition. $\qquad \square$

4.4. Proof of Corollary 1.3. — If $\sigma_\mu = 0$, it is easy to deduce from Remark 3.7 that the rate function I satisfies $I(\ell_\mu) = 0$ and $I(x) = \infty$ for every $x \in \mathbb{R} \smallsetminus \{\ell_\mu\}$ and hence Corollary 1.3 is true in that case. We therefore suppose $\sigma_\mu^2 > 0$. To treat this case, we will use some standard terminology from convex analysis, for which we refer the reader to [21]. Let, as usual, Λ denote the limit Laplace transform of the sequence $\frac{1}{n} d(L_n \cdot o, o)$. Note that Λ is convex (as it follows by a direct application of Hölder inequality), and, thanks to the finite exponential moment assumption, it takes finite values on an interval of type $(-\infty, \alpha)$ with $\alpha > 0$ and hence it is continuous on this interval. Let Λ^* be its Fenchel–Legendre transform. By Theorem 1.1 and [21, Prop. 6.1.2], we have $\partial \Lambda^*(\ell_\mu) = \{0\}$ where $\partial \Lambda^*$ is the multi-valued subdifferential function of Λ^*. Moreover, by Theorem 1.1, Λ has a second-order development at 0 and therefore, by [21, Th. 5.1.2] its subdifferential $\partial \Lambda$ is differentiable in the sense of [21, Def. 5.1.1]. Since also $\sigma_\mu^2 > 0$, we can apply [21, Prop. 6.2.5] (see also [17, Prop. 4.5]) and deduce that Λ^* satisfies

(4.10)
$$\Lambda^*(\ell_\mu + x) = \frac{1}{2\sigma_\mu^2} x + o(x^2)$$

as $x \to 0$. Now, let $\alpha > 0$ be the constant appearing in the finite exponential moment condition, i.e., $\int e^{\alpha d(g \cdot o, o)} d\mu(g) < \infty$. Then, it follows by Varadhan's integral lemma [12, Th. 4.3.1] for every $\lambda < \alpha$, we have $I^*(\lambda) = \Lambda(\lambda)$, where I^* is the Fenchel–Legendre transform of I. But since the second-order term $\sigma_\mu^2/2$ in the second-order

expansion of Λ at 0 (given by Theorem 1.1) is positive, it follows that the Fenchel–Legendre transforms of I^* and Λ coincides in a neighborhood of ℓ_μ, i.e.,

$$(4.11) \qquad\qquad I^{**}(x) = \Lambda^*(x)$$

for every $x \in (\ell_\mu - \beta, \ell_\mu + \beta)$ for some $\beta > 0$. But since by [6, Th. 1.1], the function I is convex (and lower semi-continuous), thanks to Fenchel–Legendre duality, we have $I \equiv I^{**}$ and hence the corollary follows from (4.10) and (4.11). $\qquad\square$

REMARK 4.7 (On finite time large deviation estimates). — Corollary 1.3 is an asymptotic statement obviously in its expression (as $\lambda \to 0$) but also concerning the rate function I itself (which controls, from below and above, the exponential rate of decay of probabilities of large deviations of $\frac{1}{n}d(L_n \cdot o, o) - \ell_\mu$ as $n \to \infty$). In regard to giving upper bounds for the large deviation probabilities, Corollary 1.3 parallels Proposition 4.3. However, in the spirit of concentration estimates, as in the proof of Proposition 4.3, we could have directly used Proposition 4.4 together with the Chernoff bound, to obtain *finite time estimates* for the large deviations of $\frac{1}{n}d(L_n \cdot o, o) - \ell_\mu{}^{(1)}$. This is in line with the recent work [1] where, under additional assumptions, the appearing constants are made explicit (e.g. relating with the spectral radius of the probability measure μ in the regular representation $L^2(G)$ of the isometry group G).

4.5. CONCLUDING REMARKS AND QUESTIONS. — In this final part, we include two questions motivated by our results and and make some brief comments on them.

4.5.1. *Limit Laplace transform of the Busemann cocycle*. — As a direct consequence of Theorem 4.1, we have that the functions Λ_x^+ and Λ_x^- have the same derivatives at 0 for every $x \in \overline{X}^h$. Moreover, it is not hard to see that $\Lambda_x^+ = \Lambda_y^-$ on $[0, +\infty)$ for every $x, y \in \overline{X}^h$. These suggest the following questions:

QUESTION 4.8. — *Is it true that $\Lambda_x^+ = \Lambda_x^-$ for every $x \in \overline{X}^h$? More importantly, does there exist a neighborhood of 0 such that $\Lambda_x^+ = \Lambda_y^+$ for every $x, y \in \overline{X}^h$ (and similarly $\Lambda_x^- = \Lambda_y^-$)?*

The answer to Question 4.8 is positive for $x, y \in \partial_h X$ in standard cases when an analytic approach can be implemented. These include random walks on free groups or on classical hyperbolic spaces $\mathbb{H}^n$. Regarding the last part of the question, we note that there are simple examples which show that one cannot ask that the functions Λ_x and Λ_y coincide throughout the region where they are finite/well-defined — take for example the random walk on the group $F_2 = \langle a, b \rangle$ driven by the measure $\mu = \frac{1}{2}(\delta_a + \delta_b)$ and consider $x = a^{+\infty} \in \partial F_2$ and $y = a^{-\infty} \in \partial F_2$.

---

(1) These finite time estimates then can be used, with the acceleration trick, to prove that $\lim_{\lambda \to 0} I(\ell_\mu + \lambda)/\lambda^2 \geqslant 1/2\sigma_\mu^2$.

4.5.2. *Second-order expansion below the drift without exponential moment.* — The rate function I appearing in (1.3) for $\frac{1}{n}\kappa(L_n)$ exists without any moment assumption [6, Th. 2.8]. Moreover, if μ fails to have finite exponential moment, then the rate function I vanishes on $[\ell_\mu, +\infty)$ (see [6, Rem. 3.2]). On the other hand, it follows from Gouëzel's [20, Th. 1.2] that I is positive on $[0, \ell_\mu)$ when μ has finite first moment. This suggests the following question

Question 4.9. — *Suppose μ is a non-elementary probability measure with finite second order moment. Is it true that*

$$\lim_{\lambda \to 0^-} \frac{I(\lambda + \ell_\mu)}{\lambda^2} = \frac{1}{2\sigma_\mu^2}?$$

Moreover, we note that thanks to Benoist–Quint [3, §5], the definition of the variance σ_μ^2 given in (3.14) even makes sense under the finite first moment hypothesis supposing that the isometry group $\mathrm{Isom}(X)$ acts cocompactly on X. Therefore, this suggests the subsequent question as to whether the second-order term in the second-order expansion of I below the drift vanishes when $\sigma_\mu^2 = \infty$. Similar questions can be asked about the second-order expansion of the limit Laplace transform Λ below zero.

References

[1] R. Aoun & C. Sert – "Random walks on hyperbolic spaces: concentration inequalities and probabilistic Tits alternative", *Probab. Theory Related Fields* **184** (2022), no. 1-2, p. 323–365.

[2] Y. Benoist & J.-F. Quint – "Central limit theorem for linear groups", *Ann. Probability* **44** (2016), no. 2, p. 1308–1340.

[3] _____, "Central limit theorem on hyperbolic groups", *Izv. Ross. Akad. Nauk Ser. Mat.* **80** (2016), no. 1, p. 5–26.

[4] _____, *Random walks on reductive groups*, Ergeb. Math. Grenzgeb. (3), vol. 62, Springer, Cham, 2016.

[5] M. Björklund – "Central limit theorems for gromov hyperbolic groups", *J. Theoret. Probab.* **23** (2010), no. 3, p. 871–887.

[6] A. Boulanger, P. Mathieu, C. Sert & A. Sisto – "Large deviations for random walks on hyperbolic spaces", 2020, to appear in *Ann. Sci. École Norm. Sup. (4)*, `arXiv:2008.02709`.

[7] D. L. Burkholder – "Martingale transforms", *Ann. Math. Statist.* **37** (1966), p. 1494–1504.

[8] I. Choi – "Limit laws on outer space, Teichmüller space, and CAT(0) spaces", 2022, `arXiv:2207.06597v1`.

[9] _____, "Random walks and contracting elements I: deviation inequality and limit laws", 2022, `arXiv:2207.06597v2`.

[10] _____, "Random walks and contracting elements III: outer space and outer automorphism group", 2022, `arXiv:2212.12122`.

[11] M. Coornaert, T. Delzant & A. Papadopoulos – *Géométrie et théorie des groupes. Les groupes hyperboliques de Gromov*, Lect. Notes in Math., vol. 1441, Springer-Verlag, Berlin, 1990.

[12] A. Dembo & O. Zeitouni – *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010.

[13] K. Dzhaparidze & J. H. van Zanten – "On Bernstein-type inequalities for martingales", *Stochastic Process. Appl.* **93** (2001), no. 1, p. 109–117.

[14] X. Fan, I. Grama & Q. Liu – "Martingale inequalities of type Dzhaparidze and van Zanten", *Statistics* **51** (2017), no. 6, p. 1200–1213.

[15] D. A. Freedman – "On tail probabilities for martingales", *Ann. Probability* **3** (1975), p. 100–118.

[16] M. I. Gordin & B. A. Lifšic – "Central limit theorem for stationary Markov processes", *Dokl. Akad. Nauk SSSR* **239** (1978), no. 4, p. 766–767.

[17] G. GORNI – "Conjugation and second-order properties of convex functions", *J. Math. Anal. Appl.* **158** (1991), no. 2, p. 293–315.

[18] S. GOUËZEL – "Local limit theorem for symmetric random walks in gromov-hyperbolic groups", *J. Amer. Math. Soc.* **27** (2014), no. 3, p. 893–928.

[19] ———, "Analyticity of the entropy and the escape rate of random walks in hyperbolic groups", *Discrete Anal.* (2017), article no. 7 (37 pages).

[20] ———, "Exponential bounds for random walks on hyperbolic spaces without moment conditions", *Tunis. J. Math.* **4** (2022), no. 4, p. 635–671.

[21] J.-B. HIRIART-URRUTY & C. LEMARÉCHAL – *Convex analysis and minimization algorithms. I*, Grundlehren Math. Wiss., vol. 305, Springer-Verlag, Berlin, 1993.

[22] C. HORBEZ – "Central limit theorems for mapping class groups and Out(F_N)", *Geom. Topol.* **22** (2018), no. 1, p. 105–156.

[23] F. LEDRAPPIER – "Some asymptotic properties of random walks on free groups", in *Topics in probability and Lie groups: boundary theory*, CRM Proc. Lecture Notes, vol. 28, American Mathematical Society, Providence, RI, 2001, p. 117–152.

[24] Q. LIU & F. WATBLED – "Exponential inequalities for martingales and asymptotic properties of the free energy of directed polymers in a random environment", *Stochastic Process. Appl.* **119** (2009), no. 10, p. 3101–3132.

[25] J. MAHER & G. TIOZZO – "Random walks on weakly hyperbolic groups", *J. reine angew. Math.* **742** (2018), p. 187–239.

[26] P. MATHIEU & A. SISTO – "Deviation inequalities for random walks", *Duke Math. J.* **169** (2020), no. 5, p. 961–1036.

[27] H. MIYACHI – "Unification of extremal length geometry on Teichmüller space via intersection number", *Math. Z.* **278** (2014), no. 3-4, p. 1065–1095.

[28] S. V. NAGAEV – "Some limit theorems for stationary Markov chains", *Teor. Veroyatnost. i Primenen.* **2** (1957), p. 389–416.

[29] V. H. DE LA PEÑA – "A general class of exponential inequalities for martingales and ratios", *Ann. Probability* **27** (1999), no. 1, p. 537–564.

[30] H. VICTOR – "A general class of exponential inequalities for martingales and ratios", *Ann. Probability* **27** (1999), no. 1, p. 537–564.

[31] M. J. WAINWRIGHT – *High-dimensional statistics. a non-asymptotic viewpoint*, Cambridge Series in Statistical and Probabilistic Math., vol. 48, Cambridge University Press, Cambridge, 2019.

RICHARD AOUN, University Gustave Eiffel, Champs-sur-Marne
5 boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France
E-mail : richard.aoun@univ-eiffel.fr

PIERRE MATHIEU, Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373
13453 Marseille, France
E-mail : pierre.mathieu@univ-amu.fr

CAGRI SERT, Institut für Mathematik, Universität Zürich
190, Winterthurerstrasse, 8057 Zürich, Switzerland
E-mail : cagri.sert@math.uzh.ch