TU Delft

Delft University of Technology

Large deviations for geodesic random walks

Versendaal, Rik

DOI
10.1214/19-EJP351

Publication date
2019

Document Version
Final published version

Published in
Electronic Journal of Probability

Citation (APA)
Versendaal, R. (2019). Large deviations for geodesic random walks. *Electronic Journal of Probability*, *24*, 1-39. Article 93. https://doi.org/10.1214/19-EJP351

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Copyright for all articles in EJP is CC BY 4.0

Electron. J. Probab. **24** (2019), no. 93, 1–39.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP351

Large deviations for geodesic random walks

Rik Versendaal*

Abstract

We provide a direct proof of Cramér's theorem for geodesic random walks in a complete Riemannian manifold (M, g). We show how to exploit the vector space structure of the tangent spaces to study large deviation properties of geodesic random walks in M. Furthermore, we reveal the geometric obstructions one runs into.

To overcome these obstructions, we provide a Taylor expansion of the inverse Riemannian exponential map, together with appropriate bounds. Furthermore, we compare the differential of the Riemannian exponential map to parallel transport. Finally, we show how far geodesics, possibly starting in different points, may spread in a given amount of time.

With all geometric results in place, we obtain the analogue of Cramér's theorem for geodesic random walks by showing that the curvature terms arising in this geometric analysis can be controlled and are negligible on an exponential scale.

Keywords: large deviations; Cramér's theorem; geodesic random walks; Riemannian exponential map; Jacobi fields; spreading of geodesics.
AMS MSC 2010: 60F10; 60G50; 58C99.
Submitted to EJP on December 13, 2018, final version accepted on August 5, 2019.
Supersedes arXiv:1811.09507v2.

1 Introduction

Random walks are among the most extensively studied discrete stochastic processes. Given a sequence of random variables $\{X_n\}_{n\geq 1}$ in some vector space V, one defines the random walk with increments $\{X_n\}_{n\geq 1}$ as the random variable

$$S_n = \sum_{i=1}^{n} X_i.$$

When rescaled by a factor $\frac{1}{n}$, one can study large deviations for the so obtained sequence $\{\frac{1}{n}S_n\}_{n\geq 1}$. When the increments are independent and identically distributed, Cramér's

*Delft University of Technology, Netherlands. E-mail: R.Versendaal@tudelft.nl

theorem ([1, 5]) states that the sequence $\{\frac{1}{n}S_n\}_{n\geq 1}$ satisfies the large deviations principle. Intuitively, this means that there is some rate function $I : V \to [0, \infty]$ such that

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i \approx x\right) \approx e^{-nI(x)}.$$

More specifically, the rate function is given as the Legendre transform of the log moment generating function of the increments, i.e.,

$$I(x) = \sup_{\lambda} \left\{\langle \lambda, x\rangle - \Lambda(\lambda)\right\},$$

where $\Lambda(\lambda) = \log \mathbb{E}(e^{\langle \lambda, X_1\rangle})$. One may weaken the independence assumption to obtain for example the Gärtner-Ellis theorem, see e.g. [1, 5]. Also, Cramér's theorem can be generalized to the setting of topological vector spaces or Banach spaces. Furthermore, Cramér's theorem provides a basis for path space large deviations, such as Mogulskii's theorem (random walks) and Schilder's theorem (Brownian motion), see e.g. [1, 12, 2]. Recently, it was shown in [8] that the analogue of Cramér's theorem (as well as Mogulskii's theorem and Schilder's theorem) also holds in the Riemannian setting.

In [8], Cramér's theorem for geodesic random walks is obtained by first proving the Riemannian analogue of Moguslkii's theorem, the path space analogue of Cramér's theorem. As evaluation in the end point of trajectories is a continuous map, Cramér's theorem then follows by an application of the contraction principle (see e.g. [1, Chapter 4]). To obtain Mogulskii's theorem, the Feng-Kurtz formalism ([3]) is used. However, this is the reverse order in which the theorems are obtained naturally in the Euclidean case. In the Euclidean setting, one uses Cramér's theorem to prove Mogulskii's theorem by first proving large deviations for the finite dimensional distributions and then deducing from these the large deviations on path space. Furthermore, the Feng-Kurtz approach is only suitable for Markov processes and hence does not extend to the case where the increments are allowed to be dependent. This causes an obstruction in finding a Riemannian analogue of the Gärtner-Ellis theorem for example.

These observations raise the question whether it is possible to avoid the use of the Feng-Kurtz formalism and path space large deviations to obtain Cramér's theorem for geodesic random walks. It turns out that it is possible to only study the underlying geometry in order to prove Cramér's theorem. This gives us new insight in what geometrical aspects allow us to still obtain the large deviation principle for rescaled geodesic random walks, even though the geodesic random walk is in general no longer a simple function of its increments. Furthermore, this approach does not rely on the fact that the random walk is a Markov process, and thus seems suitable to be extended to random walks with dependent increments for example. This will be investigated further in future work.

The main difficulty in the Riemannian setting, is that we lack a vector space structure to define a random walk as sum of increments. The appropriate analogue is a geodesic random walk as introduced by Jørgensen in [6]. To define a geodesic random walk, we need to find a replacement for the additive structure, as well as a generalization of the increments. It turns out that as increments one uses tangent vectors, while the additive structure is replaced by an application of the Riemannian exponential map.

More precisely, we introduce a family of probability measures $\{\mu_x\}_{x\in M}$ such that for each $x \in M$, μ_x is a measure on T_xM, the tangent space at x. These measures $\{\mu_x\}_{x\in M}$ provide the space-dependent distribution of the increments. Now we start a random walk at some initial point $Z_0 = x_0 \in M$. Then recursively, we define for $k = 0, \ldots, n-1$ the random variable

$$Z_{k+1} = \mathrm{Exp}_{Z_k}\left(\frac{1}{n}X_{k+1}\right),$$

where X_{k+1} is distributed according to μ_{Z_k}. Hence, the random variable Z_n takes values in M and is the natural analogue of the empirical mean of the increments $X_1, \dots, X_n$. In Euclidean space, this definition reduces to the usual one, as the Riemannian exponential map is simply vector addition, i.e.,

$$\operatorname{Exp}_x v = x + v.$$

To obtain an analogue of Cramér's theorem, we also need to generalize the notion of the increments of the random walk being identically distributed, since the increments are no longer in the same space. To compare two distributions μ_x and μ_y, we need to identify the tangent spaces T_xM and T_yM. We do this by taking a curve γ connecting x and y and using parallel transport along γ. Because different curves lead to different identifications, we say that the distributions μ_x and μ_y are identical if for all curves γ from x to y we have

$$\mu_x = \mu_y \circ \tau_{yx;\gamma}^{-1},$$

where τ denotes parallel transport. Equivalently, one can characterize this property by assuming that the log moment generating functions are invariant under parallel transport, i.e.,

$$\Lambda_x(\lambda) = \Lambda_y(\tau_{xy;\gamma}\lambda),$$

where $\Lambda_x(\lambda) = \log \int_{T_xM} e^{\langle \lambda, v\rangle} \mu_x(\mathrm{d}v)$.

In Euclidean space, the end point of the random walk is a simple function of the increments. In the Riemannian setting, curvature ensures that this is in general no longer the case. For example, the endpoint in general depends on the order of the increments. Nonetheless, it is possible to utilize the vector space structure of the tangent spaces. By controlling the error induced by the curvature, the large deviations for the geodesic random walk Z_n can be obtained from the large deviations for $\frac{1}{n}\sum_{i=1}^n \tilde{X}_i$, the empirical mean of the appropriately pulled back increments in $T_{x_0}M$, were x_0 is the starting point of the random walk.

To support this claim, we can also define an alternative random walk in M. For this, we take a sequence of independent, identically distributed random variables $\{Y_n\}_{n\geq 1}$ in $T_{x_0}M$ with distribution μ_{x_0} and consider the process

$$\tilde{Z}_n = \operatorname{Exp}_{x_0}\left(\frac{1}{n}\sum_{i=1}^n Y_i\right).$$

In general, $\tilde{Z}_n$ is different from Z_n, even in distribution. Although our method of proving the large deviations for Z_n does not immediately allow us to conclude that Z_n and $\tilde{Z}_n$ are exponentially equivalent, the main idea of our proof does rely on the fact that we can (in some sense) relate and compare the geodesic random walk to a sum of independent, identically distributed random variables in the tangent space at x_0, following the distribution μ_{x_0}.

The paper is organised as follows. In Section 2 we introduce the main notions we use from large deviation theory to obtain our results, as well as some notation and results from differential geometry. Section 3 introduces the geodesic random walks. In Section 4 we give the precise statement of Cramér's theorem for geodesic random walks. Additionally, we provide an overview of the various steps that are needed for the proof. In Section 5 we obtain a Taylor expansions of the Riemannian exponential map with appropriate error bound. Furthermore, we compare the differential of the exponential map to parallel transport. Finally, we also provide bounds for how far geodesics, possibly starting at different points, can spread in a given amount of time. These geometric results are key ingredients in the proof of Cramér's theorem, which is given in Section 6.

2 Notation and important notions

In this section we collect some important notions and fix the notation we will be using. Firstly, we introduce large deviation principles, along with some general useful results from the theory. Following up, we introduce the necessary tools from differential geometry and fix the notation for the various objects.

2.1 Large deviation principle

Large deviation principles are concerned with the asymptotic behaviour on an exponential scale of a sequence of probability measures $\{\nu_n\}_{n\geq 1}$. This behaviour is governed by a rate function. We make this precise in the following definition.

Definition 2.1. *Let $\{\nu_n\}_{n\geq 1}$ be a sequence of probability measures with values in a metric space $\mathcal{X}$.*

1. *A* rate function *is a lower semicontinuous function $I : \mathcal{X} \to [0, \infty]$. A rate function is called* good *if the level sets $\{x \in \mathcal{X} | I(x) \leq c\}$ are compact for any $c \geq 0$.*

2. *The sequence $\{\nu_n\}_{n\geq 1}$ satisfies the* large deviation principle (LDP) *in $\mathcal{X}$ with rate function I if the following are satisfied:*

   (a) (Upper bound) For any closed $F \subset \mathcal{X}$

   $$\limsup_{n\to\infty} \frac{1}{n} \log \nu_n(F) \leq - \inf_{x\in F} I(x).$$

   (b) (Lower bound) For any open $G \subset \mathcal{X}$

   $$\liminf_{n\to\infty} \frac{1}{n} \log \nu_n(G) \geq - \inf_{x\in G} I(x).$$

3. *The sequence $\{\nu_n\}_{n\geq 1}$ is* exponentially tight *if for every $\alpha > 0$ there exists a compact set $K_\alpha \subset \mathcal{X}$ satisfying*

   $$\limsup_{n\to\infty} \frac{1}{n} \log \nu_n(K_\alpha^c) < -\alpha.$$

When a sequence of probability measures is exponentially tight, it is sufficient to know the upper bound of the large deviation principle only for compact sets. The upper bound then also immediately holds for all closed sets, see e.g. [1, Section 1.2].

2.2 Riemannian geometry

In this section we introduce the necessary notions from differential geometry, see for example [11] for a general introduction. We mainly focus on Riemannian geometry, for which we refer to [9] among others.

Let (M, g) be a Riemannian manifold of dimension N. As usual, we denote by TM the tangent bundle of M. For a point $x \in M$ we write T_xM for the tangent space at x. Tangent vectors are usually denoted by v. A smooth assignment of tangent vectors to all points at M is called a vector field, and the set of vector fields is denoted by $\Gamma(TM)$.

For $x \in M$ and $v, w \in T_xM$ we write the inner product as $\langle v, w\rangle_{g(x)}$, where the subscript is omitted when the tangent space is understood. Given the inner product, we define the length of $v \in T_xM$ by its usual formula

$$|v|_{g(x)} = \sqrt{\langle v, v\rangle_{g(x)}}.$$

Given a curve $\gamma : [a, b] \to M$, we define its length by

$$L(\gamma; [a, b]) = \int_a^b |\dot\gamma(t)| \, \mathrm{d}t.$$

Using this length function, we define the Riemannian distance d on M as

$$d(x, y) := \inf\{L(\gamma) | \gamma : [a, b] \to M, \gamma(a) = x, \gamma(b) = y, \gamma \text{ piecewise smooth}\}. \tag{2.1}$$

2.2.1 Connection and parallel transport

Associated to the Riemannian metric g is a unique connection ∇, the *Levi-Civita* connection, which is compatible with the metric and torsion free.

A vector field $v(t)$ along a curve $\gamma(t)$ is called parallel if $D_t v(t) := \nabla_{\dot\gamma(t)} v(t) = 0$. If the vector field $\dot\gamma(t)$ is parallel along $\gamma(t)$, then γ is called a geodesic. It turns out that optimal paths for the distance between points in M are geodesics for the Levi-Civita connection.

Equivalent to having a connection is having a notion of parallel transport. Given a curve $\gamma : [a, b] \to M$ and $v \in T_{\gamma(a)}M$, we can consider the unique solution $v(t)$ of the differential equation

$$\nabla_{\dot\gamma(t)} v(t) = 0, \qquad v(0) = v.$$

This allows us to define a linear map

$$\tau_{\gamma(a)\gamma(t);\gamma} : T_{\gamma(a)}M \to T_{\gamma(t)}M$$

by setting $\tau_{\gamma(a)\gamma(t);\gamma} v = v(t)$. The map $\tau_{\gamma(a)\gamma(t);\gamma}$ is called *parallel transport* along γ. We omit the reference to the curve γ when it is understood. Because ∇ is compatible with the Riemannian metric, parallel transport is in fact an isometry.

Conversely, we can use parallel transport to compute covariant derivatives. To this end, let $v, w \in \Gamma(TM)$ be vector fields and $x \in M$. Let γ be a curve with $\gamma(0) = x$ and $\dot\gamma(0) = v$. Then

$$\nabla_v w(x) = \lim_{h \to 0} \frac{\tau^{-1}_{x\gamma(h)} w(\gamma(h)) - w(x)}{h}.$$

2.2.2 Riemannian exponential map

Given $x \in M$, define for every $v \in T_x M$ the geodesic γ_v satisfying $\gamma_v(0) = x$ and $\dot\gamma_v(0) = v$. A priori, this geodesic does not exist for all time t. We say that the manifold M is *complete* if every such geodesic can be extended indefinitely. By the Hopf-Rinow theorem, this is equivalent to the completeness of M as a metric space with the Riemannian distance d defined in (2.1).

We now define the *Riemannian exponential map* $\mathrm{Exp}_x : \mathcal{E}(x) \to M$ by setting $\mathrm{Exp}_x v = \gamma_v(1)$, where $\mathcal{E}(x) \subset T_x M$ contains all $v \in T_x M$ for which γ_v as above exists at least on $[0, 1]$. If M is complete, we have $\mathcal{E}(x) = T_x M$. If additionally M is simply connected, it holds that Exp_x is surjective.

However, due to curvature, the exponential map is not necessarily injective. For $x \in M$ we define the *injectivity radius* $\iota(x) \in (0, \infty]$ as

$$\iota(x) = \sup\{\delta > 0 | \mathrm{Exp}_x \text{ is injective on } B(0, \delta)\}.$$

Given a set $A \subset M$, the injectivity radius of A is defined by

$$\iota(A) = \inf\{\iota(x) | x \in A\}. \tag{2.2}$$

It can be shown (see e.g. [7]) that the map $x \mapsto \iota(x)$ is continuous on M. Consequently, for a compact set K we have $\iota(K) > 0$.

The differential $\mathrm{d}(\mathrm{Exp}_x)$ of the exponential map at x is a linear map from $T(T_xM)$ into TM. Upon identifying $T_v(T_xM)$ with T_xM, we find that for any $v \in T_xM$ we have

$$\mathrm{d}(\mathrm{Exp}_x)_v : T_xM \to T_{\mathrm{Exp}_x v}M.$$

2.2.3 Jacobi fields

Let $\gamma : [0,1] \to M$ be a smooth curve. A *variation* of γ is a smooth map $\Gamma : (-\varepsilon, \varepsilon) \times [0,1] \to M$ such that $\Gamma(0,t) = \gamma(t)$ for all $t \in [0,1]$. Denoting by s the first variable, the *variational vector field* V of Γ is defined as

$$V(t) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \Gamma(s,t) =: \partial_s \Gamma(0,t).$$

Intuitively, V measures the speed at which the curve γ deforms.

We denote by D_t the covariant derivative along the curve $t \mapsto \Gamma(s,t)$, and similarly for D_s. Because the Levi-Civita connection is symmetric, we obtain the following symmetry lemma, see e.g. [9, Lemma 6.3].

Lemma 2.2 (Symmetry lemma)**.** *Let $\gamma : [0,1] \to M$ be a smooth curve and $\Gamma : (-\varepsilon, \varepsilon) \times [0,1] \to M$ a variation of γ. If M is equipped with the Levi-Civita connection, then*

$$D_s \partial_t \Gamma(s,t) = D_t \partial_s \Gamma(s,t).$$

Now suppose $\gamma : [0,1] \to M$ is a geodesic. Let $\Gamma : (-\varepsilon, \varepsilon) \times [0,1] \to M$ be a variation of γ such that for any $s \in (-\varepsilon, \varepsilon)$, the curve $t \mapsto \Gamma(s,t)$ is a geodesic. We call Γ a *variation of geodesics*, and the corresponding variational vector field is called a *Jacobi field* along γ.

It is possible to derive a second order differential equation satisfied by Jacobi fields. For this, we need to introduce the Riemann curvature endomorphism. The *Riemann curvature endomorphism* measures the commutativity of second order covariant derivatives of a vector field. More precisely, it is a map $R : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ defined by

$$R(v,w)u = \nabla_v \nabla_w u - \nabla_w \nabla_v u - \nabla_{[v,w]} u,$$

where $[v,w] = vw - wv$ is the commutator of the vector fields v and w.

One can show (see e.g. [9, Theorem 10.2] or [4, Section 10.1]) that a Jacobi field $J(t)$ along a geodesic γ satisfies

$$D_t^2 J(t) + R(J(t), \dot\gamma(t))\dot\gamma(t) = 0, \tag{2.3}$$

where R denotes the Riemann curvature endomorphism. Equation (2.3) is called the Jacobi equation.

If $J(0) = 0$ and $\dot J(0)$ is given, a Jacobi field along a geodesic γ satisfying these conditions is

$$J(t) = \mathrm{d}(\mathrm{Exp}_{\gamma(0)})_{t\dot\gamma(0)}(t\dot J(0)).$$

This can be seen by considering the variation $\Gamma(t,s) = \mathrm{Exp}_{\gamma(0)}(t(\dot\gamma(0) + s\dot J(0)))$. The condition that $J(0) = 0$ indicates that all geodesics in the variation start in the same point.

In Euclidean space, this Jacobi field reduces to $J(t) = t\dot J(0)$, which is indeed the variation field of the variation $\Gamma(t,s) = \gamma(0) + t(\dot\gamma(0) + s\dot J(0))$.

We conclude this section by collecting some properties of Jacobi fields that we need later on. We include the arguments for the reader's convenience.

Proposition 2.3. *Let $\gamma : [0,1] \to M$ be a geodesic and $J(t)$ a Jacobi field along γ. Then*

$$\langle J(t), \dot\gamma(t)\rangle = t\langle \dot J(0), \dot\gamma(0)\rangle + \langle J(0), \dot\gamma(0)\rangle$$

for all $t \in [0,1]$.

Proof. Define $f(t) = \langle J(t), \dot\gamma(t)\rangle$. Then

$$f'(t) = \langle D_t J(t), \dot\gamma(t)\rangle + \langle J(t), D_t\dot\gamma(t)\rangle = \langle D_t J(t), \dot\gamma(t)\rangle,$$

because γ is a geodesic. We are done once we show that $f''(t) = 0$. For this, notice that, using (2.3)

$$f''(t) = \langle D_t^2 J(t), \dot\gamma(t)\rangle = -\langle R(J(t), \dot\gamma(t))\dot\gamma(t), \dot\gamma(t)\rangle = 0.$$

Here, the last step follows from the symmetry properties of the Riemann curvature tensor. □

Proposition 2.4. *Let $\gamma : [0,1] \to M$ be a geodesic and $J(t)$ a Jacobi field along γ. For every $t \in [0,1]$ there exists $\xi_t \in (0,t)$ such that*

$$|\dot J(t)| = |\dot J(0)| - t\frac{1}{|\dot J(\xi_t)|}\langle R(J(\xi_t), \dot\gamma(\xi_t))\dot\gamma(\xi_t), \dot J(\xi_t)\rangle.$$

Proof. Define $f(t) = |\dot J(t)|$. We have

$$\begin{aligned} f'(t) &= \frac{1}{|\dot J(t)|}\langle \ddot J(t), \dot J(t)\rangle \\ &= -\frac{1}{|\dot J(t)|}\langle R(J(t), \dot\gamma(t))\dot\gamma(t), \dot J(t)\rangle. \end{aligned}$$

The statement now follows from the mean-value theorem. □

3 Geodesic random walks

In order to generalize Cramér's theorem to the setting of Riemannian manifolds, we first need to introduce the appropriate analogue of the sequence $\{\frac{1}{n}\sum_{i=1}^n X_i\}_{n\geq 0}$ for a sequence of increments $\{X_n\}_{n\geq 1}$. In order to do this, we introduce geodesic random walks, following the construction in [6]. Finally, we generalize the notion of identically distributed increments to geodesic random walks and characterize it using log moment generating functions.

3.1 Definition of geodesic random walks

We start by defining a *geodesic random walk* $\{\mathcal{S}_n\}_{n\geq 0}$ on M with increments $\{X_n\}_{n\geq 1}$. For this we need to generalize how to add increments together. This is achieved by using the Riemannian exponential map. Because the space variable determines in which tangent space the increment should be, we have to define the random walk recursively, which is the main difficulty in the definition below.

Definition 3.1. *Fix x_0 in M. A pair $(\{\mathcal{S}_n\}_{n\geq 0}, \{X_n\}_{n\geq 1})$ is called a* geodesic random walk *with increments $\{X_n\}_{n\geq 1}$ and started at x_0 if the following hold:*

1. *$\mathcal{S}_0 = x_0$,*
2. *$X_{n+1} \in T_{\mathcal{S}_n} M$ for all $n \geq 0$,*
3. *$\mathcal{S}_{n+1} = \mathrm{Exp}_{\mathcal{S}_n}(X_{n+1})$ for all $n \geq 0$.*

In what follows, the sequence $\{X_n\}_{n\geq 1}$ of increments will usually be omitted and we simply write that $\{\mathcal{S}_n\}_{n\geq 0}$ is a geodesic random walk with increments $\{X_n\}_{n\geq 1}$.

Note that in the above definition, we fix nothing about the distribution of the increments $\{X_n\}_{n\geq 1}$. The distribution is allowed to depend both on the space variable, as well as on time.

For $M = \mathbb{R}^N$, the Riemannian exponential map can be identified with addition, i.e., $\mathrm{Exp}_x(v) = x + v$. Hence, a geodesic random walk in $\mathbb{R}^N$ reduces to the usual random walk, i.e. $\mathcal{S}_n = \sum_{i=1}^n X_i$.

Next, we introduce the concept of time-homogeneous increments for geodesic random walks. For this, we need to fix the distribution of the increments independent of the time variable. Because the increments can take values in different tangent spaces, we need a collection of measures $\{\mu_x\}_{x\in M}$ such that μ_x is a probability measure on T_xM for every $x \in M$. We denote the set of probability measures on T_xM by $\mathcal{P}(T_xM)$. We have the following definition.

Definition 3.2. *Let $\{\mathcal{S}_n\}_{n\geq 0}$ be a geodesic random walk with increments $\{X_n\}_{n\geq 1}$ and started at x_0. Let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for every $x \in M$. We say the random walk $(\{\mathcal{S}_n\}_{n\geq 0}, \{X_n\}_{n\geq 1})$ is* compatible *with the collection $\{\mu_x\}_{x\in M}$ if $X_{n+1} \sim \mu_{\mathcal{S}_n}$ for every $n \geq 0$.*

Essentially, the collection of measures provides the distributions for the increments of the geodesic random walk. Because the collection of measures is independent of n, the increments are time-homogeneous.

Next, we want to define what it means for the increments of a geodesic random walk to be independent. Because the distribution of increment X_{n+1} depends on $\mathcal{S}_n$, we have that X_{n+1} is in general not independent of $\mathcal{A}_n = \sigma(\{X_1, \dots, X_n\})$ in the usual sense. However, this dependence is purely geometric, as $\mathcal{S}_n$ simply determines in which tangent space we have to choose X_{n+1}. If this is the only dependence of X_{n+1} on $\mathcal{A}_n$, we say the increments of $\{\mathcal{S}_n\}_{n\geq 0}$ are independently distributed. We make this precise in the following definition.

Definition 3.3. *Let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for every $x \in M$. Let $\{\mathcal{S}_n\}_{n\geq 0}$ be a geodesic random walk with increments $\{X_n\}_{n\geq 1}$, compatible with $\{\mu_x\}_{x\in M}$. For every $n \geq 1$, define the σ-algebra $\mathcal{F}_n$ by*

$$\mathcal{F}_n = \sigma(\{(\mathcal{S}_0, X_1), \dots, (\mathcal{S}_{n-1}, X_n)\}).$$

We say the increments of $\{\mathcal{S}_n\}_{n\geq 0}$ are independent*, if for every $n \geq 1$ and all bounded, continuous functions $f : M^n \to \mathbb{R}$ we have*

$$\mathbb{E}(f(X_1, \dots, X_n)|\mathcal{F}_{n-1}) = \int_{T_{\mathcal{S}_{n-1}}M} f(X_1, \dots, X_{n-1}, v)\mu_{\mathcal{S}_{n-1}}(\mathrm{d}v).$$

Remark 3.4. Because $\mathcal{S}_n = \mathrm{Exp}_{\mathcal{S}_{n-1}} X_n$, we have that $\mathcal{S}_n$ is $\mathcal{F}_n$-measurable. Consequently, we have $\sigma(\{\mathcal{S}_0, \dots, \mathcal{S}_n\}) \subset \mathcal{F}_n$. However, equality need not hold. Indeed, if the Riemannian exponential map Exp_x is not injective, one cannot retrieve the increments $X_1, \dots, X_n$ from $\mathcal{S}_0, \dots, \mathcal{S}_n$.

Remark 3.5. Let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for all $x \in M$. Let $\{\mathcal{S}_n\}_{n\geq 0}$ be a geodesic random walk with increments $\{X_n\}_{n\geq 1}$ compatible with $\{\mu_x\}_{x\in M}$. Suppose furthermore that the increments are independent. Then $\{\mathcal{S}_n\}_{n\geq 0}$ is a time-homogeneous, discrete time Markov process on M with transition operator

$$Pf(x) = \mathbb{E}(f(\mathcal{S}_1)|\mathcal{S}_0 = x) = \int_{T_xM} f(\mathrm{Exp}_x(v))\mu_x(\mathrm{d}v).$$

This is the point of view taken in [8].

3.1.1 Rescaled geodesic random walks

In Euclidean space, one commonly encounters rescaled versions of a random walk, for example for laws of large numbers and central limit theorems. On a general manifold, this rescaling cannot be achieved by multiplication.

Before we define the appropriate analogue of $\{\frac{1}{n}\sum_{i=1}^n X_i\}_{n\geq 0}$, we first need to define how to rescale a geodesic random walk by a factor $\alpha > 0$ independent of n. Note that in Euclidean space we can write $\alpha\sum_{i=1}^n X_i = \sum_{i=1}^n(\alpha X_i)$. This shows that we should rescale the increments of the random walk, which is possible in a manifold, because the increments are tangent vectors.

Definition 3.6. *Fix x_0 in M and $\alpha > 0$. A pair $(\{(\alpha * \mathcal{S})_n\}_{n\geq 0}, \{X_n\}_{n\geq 1})$ is called an* α-rescaled geodesic random walk *with increments $\{X_n\}_{n\geq 1}$ and started at x_0 if the following hold:*

1. $(\alpha * \mathcal{S})_0 = x_0$,
2. $X_{n+1} \in T_{(\alpha*\mathcal{S})_n}M$ *for all* $n \geq 0$,
3. $(\alpha * \mathcal{S})_{n+1} = \mathrm{Exp}_{(\alpha*\mathcal{S})_n}(\alpha X_{n+1})$ *for all* $n \geq 0$.

As with geodesic random walks, we will often omit the sequence of increments and simply write that $\{(\alpha * \mathcal{S})_n\}_{n\geq 0}$ is an α-rescaled geodesic random walk with increments $\{X_n\}_{n\geq 1}$.

Note that an α-rescaled geodesic random walk can itself be considered as a geodesic random walk. Indeed, if $(\alpha * \mathcal{S})_n$ is an α-rescaled geodesic random walk with increments $\{X_n\}_{n\geq 1}$, then it is a geodesic random walk with increments $\{\alpha X_n\}_{n\geq 1}$.

As for geodesic random walks, we say that an α-rescaled geodesic random walk $\{(\alpha * \mathcal{S})_n\}_{n\geq 0}$ with increments $\{X_n\}_{n\geq 1}$ is compatible with a collection of probability measures $\{\mu_x\}_{x\in M}$ if $X_{n+1} \sim \mu_{(\alpha*\mathcal{S})_n}$ for every $n \geq 0$. It follows that when considered as geodesic random walk, $\{(\alpha * \mathcal{S}_n)\}_{n\geq 0}$ is compatible with the collection of measures $\{\mu_x^\alpha\}_{x\in M}$ given by

$$\mu_x^\alpha = \mu_x \circ m_\alpha^{-1}$$

where $m_\alpha : T_xM \to T_xM$ denotes multiplication by α, i.e., $m_\alpha(v) = \alpha v$.

3.1.2 Empirical average process

We conclude this section by introducing the analogue of the sequence of empirical averages $\{\frac{1}{n}\sum_{i=1}^n X_i\}_{n\geq 0}$ for a sequence $\{X_n\}_{n\geq 1}$ of random variables.

Fix $x_0 \in M$ and let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for all $x \in M$. For every $n \geq 1$, let $\{(\frac{1}{n} * \mathcal{S})_j\}_{j\geq 0}$ be a $\frac{1}{n}$-rescaled geodesic random walk started at x_0 with increments $\{X_j^n\}_{j\geq 1}$, compatible with the measures $\{\mu_x\}_{x\in M}$. By considering the diagonal elements of $\{(\frac{1}{n} * \mathcal{S})_j\}_{n\geq 1, j\geq 0}$, we obtain for every $n \geq 1$ a random variable $(\frac{1}{n} * \mathcal{S})_n$ in M. If we now set the initial value of the sequence $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$ to be x_0, we obtain the Riemannian analogue of the sequence $\{\frac{1}{n}\sum_{i=1}^n X_i\}_{n\geq 0}$. We refer to this process as the *empirical average process* started at x_0 compatible with the collection of measures $\{\mu_x\}_{x\in M}$.

3.2 Identically distributed increments

For our purposes, we also need a notion of identically distributed increments. In general, the increments of a geodesic random walk do not live in the same tangent space. In order to overcome this problem, we use parallel transport to identify tangent spaces. Because the identification via parallel transport depends on the curve along which the vectors are transported, we need to make the following definition.

Definition 3.7. *Let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for all $x \in M$. Let $\{\mathcal{S}_n\}_{n\geq 0}$ be a geodesic random walk with increments $\{X_n\}_{n\geq 1}$, compatible with $\{\mu_x\}_{x\in M}$. We say the increments $\{X_n\}_{n\geq 1}$ are* identically distributed *if the measures satisfy the following* consistency property*: for any $y, z \in M$ and any smooth curve $\gamma : [a, b] \to M$ with $\gamma(a) = y$ and $\gamma(b) = z$ we have*

$$\mu_z = \mu_y \circ \tau^{-1}_{yz;\gamma}.$$

By the transitivity property of parallel transport, one can equivalently define the consistency property to hold for all piecewise smooth curves.

Note that in Euclidean space, our definition of independent increments implies that the measures are independent of the space variable, because parallel transport is the identity map. Hence, our definition reduces to the usual one, as we obtain that every increment has some fixed distribution μ.

Because parallel transport is an isometry, we can use distributions with spherical symmetry to construct a family of measures $\{\mu_x\}_{x\in M}$ satisfying Definition 3.7. We refer to [8, Section 4] for the details and more specific examples.

The consistency property in Definition 3.7 may also be characterised by a consistency assumption for the corresponding log-moment generating functions $\Lambda_x : T_xM \to \mathbb{R}$ of μ_x given by

$$\Lambda_x(\lambda) = \log \int_{T_xM} e^{\langle \lambda, v\rangle} \mu_x(\mathrm{d}v).$$

This is recorded in the following proposition, which can be found in [8, Section 4].

Proposition 3.8. *Let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for every $x \in M$. Assume that $\Lambda_x(\lambda) < \infty$ for all $x \in M$ and all $\lambda \in T_xM$. The following are equivalent:*

(a) The collection $\{\mu_x\}_{x\in M}$ satisfies the consistency property in Definition 3.7.

(b) For all $x, y \in M$ and all smooth curves $\gamma : [a, b] \to M$ with $\gamma(a) = x$ and $\gamma(b) = y$ and for all $\lambda \in T_xM$ we have

$$\Lambda_x(\lambda) = \Lambda_y(\tau_{xy;\gamma}\lambda).$$

The *Legendre transform* $\Lambda^*_x : T_xM \to \mathbb{R}$ of Λ_x is defined by

$$\Lambda^*_x(v) := \sup_{\lambda\in T_xM} \langle \lambda, v\rangle - \Lambda_x(\lambda).$$

If the collection of log-moment generating functions $\{\Lambda_x\}_{x\in M}$ satisfies the consistency property in (b) of the above proposition, then so does the collection $\{\Lambda^*_x\}_{x\in M}$ of their Legendre transforms.

4 Sketch of the proof of Cramér's theorem for Riemannian manifolds

In this section we provide a sketch of the proof of Cramér's theorem for geodesic random walks and stress what observations and properties are important to make the proof work. Before we get to this, let us first state the exact theorem we wish to prove.

4.1 Statement of Cramér's theorem

Cramér's theorem is concerned with the large deviations for the empirical average process $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 1}$ with independent, identically distributed increments.

Along with the large deviation principle, we need to identify the rate function. In Euclidean space, the rate function is given by

$$I(x) = \Lambda^*(x),$$

the Legendre transform of the log moment generating function of an increment. Note here that one can consider the vector x as the tangent vector of the straight line from the origin to the point x. Using this viewpoint, the analogue of the rate function in the Riemannian setting should be

$$I(x) = \inf\{\Lambda^*_{x_0}(v) | \mathrm{Exp}_{x_0} v = x\}.$$

Here, we have to take the infimum, because the Riemannian exponential map is not necessarily injective, i.e., there may be more than one geodesic connecting x_0 and x. We will show that this is indeed the correct rate function, as collected in the following theorem.

Theorem 4.1 (Cramér's theorem for Riemannian manifolds)**.** *Let (M, g) be a complete Riemannian manifold. Fix $x_0 \in M$ and let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for all $x \in M$. For every $n \geq 1$, let $\{(\frac{1}{n} * \mathcal{S})_j\}_{j\geq 0}$ be a $\frac{1}{n}$-rescaled geodesic random walk started at x_0 with independent increments $\{X^n_j\}_{j\geq 1}$, compatible with $\{\mu_x\}_{x\in M}$. Let $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$ be the associated empirical average process started at x_0. Assume the increments are bounded and have expectation 0. Assume furthermore that the collection $\{\mu_x\}_{x\in M}$ satisfies the consistency property in Definition 3.7. Then $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$ satisfies in M the LDP with good rate function*

$$I_M(x) = \inf\{\Lambda^*_{x_0}(v) | \mathrm{Exp}_{x_0} v = x\} \tag{4.1}$$

Due to geometrical influences, which become apparent when sketching the proof, we prove Cramér's theorem only in the case when the increments are bounded. This allows for a less technical proof of the theorem, but nevertheless introduces all geometrical obstructions that have to be dealt with. The details of the proof can be found in Section 6.

Like in the Euclidean setting, we prove Cramér's theorem for geodesic random walks by separately proving the upper and lower bound for the large deviation principle of $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$. In Section 4.2 we give an overview of the steps one needs to take to prove the upper bound, while in Section 4.3 we sketch how to prove the lower bound.

4.2 Sketch of the proof of the upper bound

In the Euclidean case, one proves the upper bound in Cramér's theorem by using Chebyshev's inequality. More precisely, the key step is to show that for $\Gamma \subset \mathbb{R}^d$ compact one has (see e.g. [5, 1])

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} S_n \in \Gamma\right) \leq - \inf_{x\in\Gamma} \sup_{\lambda\in\mathbb{R}^d} \left\{ \langle \lambda, x\rangle - \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(e^{n\langle \lambda, \frac{1}{n} S_n\rangle}\right)\right\}.$$

The upper bound is then extended to all closed sets by proving exponential tightness. The idea is to follow a similar procedure in the Riemannian case. However, because $(\frac{1}{n} * \mathcal{S})_n$ is M-valued, its moment generating function is not defined.

4.2.1 Step 1: Analogue of the moment generating function $\mathbb{E}(e^{n\langle \lambda, \frac{1}{n} S_n\rangle})$

To overcome the problem of not having a moment generating function of $(\frac{1}{n} * \mathcal{S})_n$, we want to identify points in M with tangent vectors in $T_{x_0}M$. For this we use the Riemannian

exponential map. However, this map is not necessarily injective. Hence, we first assume that for each $n \geq 1$, the $\frac{1}{n}$-rescaled geodesic random walk stays within the injectivity radius $\iota(x_0)$ of its initial point x_0 up to time n. Consequently, because Exp_{x_0} is injective on $B(0, \iota(x_0)) \subset T_{x_0}M$, we can uniquely define $v_k^n \in T_{x_0}M$ satisfying $|v_k^n| < \iota(x_0)$ and

$$\mathrm{Exp}_{x_0}^{-1}(v_k^n) = \left(\frac{1}{n} * \mathcal{S}\right)_k.$$

Ideally, we would like to prove the large deviation principle for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$ by proving the large deviation principle for $\{v_n^n\}_{n\geq 0}$ in $T_{x_0}M$ and then apply the contraction principle (see e.g. [1, Chapter 4]) with the continuous function Exp_{x_0}. For this to work, we would need to show that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{n\langle \lambda, v_n^n\rangle}) = \Lambda_{x_0}(\lambda). \tag{4.2}$$

Unfortunately, using the estimate for $\mathbb{E}(e^{n\langle \lambda, v_n^n\rangle})$ found in Step 2 as explained below, we are only able to show that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{n\langle \lambda, v_n^n\rangle}) \leq \Lambda_{x_0}(\lambda) + C|\lambda| \tag{4.3}$$

and likewise

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{n\langle \lambda, v_n^n\rangle}) \geq \Lambda_{x_0}(\lambda) - C|\lambda|, \tag{4.4}$$

where the constant only depends on the curvature and the uniform bound of the increments.

4.2.2 Step 2: Upper bound for the moment generating function of v_n^n

In $\mathbb{R}^d$ we simply have $v_n^n = \frac{1}{n}\sum_{i=1}^n X_i$ and hence its moment generating function is given by

$$\mathbb{E}(e^{n\langle \lambda, v_n^n\rangle}) = \prod_{i=1}^n \mathbb{E}(e^{\langle \lambda, X_i\rangle}) = \mathbb{E}(e^{\langle \lambda, X_1\rangle})^n.$$

Here we use the fact that we can write $v_k^n = v_{k-1}^n + \frac{1}{n}X_k$. This fails in the Riemannian setting, which results in the fact that we can only estimate $\mathbb{E}(e^{n\langle \lambda, v_n^n\rangle})$ as mentioned above in (4.3) and (4.4).

In a Riemannian manifold we replace the identity $v_k^n = v_{k-1}^n + \frac{1}{n}X_k$ by the Taylor expansion of $\mathrm{Exp}_{x_0}^{-1}$ (see Section 5.1, Proposition 5.4). This results in

$$v_k^n = v_{k-1}^n + \frac{1}{n}\mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n + \mathcal{O}\left(\frac{1}{n^2}\right). \tag{4.5}$$

Here one needs to be careful that the constant in the error term may depend on curvature properties of the manifold around $(\frac{1}{n} * \mathcal{S})_{k-1}$. Because we assume the increments are uniformly bounded, there exists a compact set $K \subset M$ such that for all $n \geq 1$ and all $0 \leq j \leq n$ we have $(\frac{1}{n} * \mathcal{S})_j \in K$. This allows us to control the constant in the error term.

However, the problem arises that this expression does not yet allow us to use the assumption that the increments of the geodesic random walk are identically distributed, which essentially means that the distribution of the increments is invariant under parallel transport.

Consequently, we need to argue that $\mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n}$ can be approximated well enough by parallel transport. It turns out there exists a constant $C > 0$ such that

$$|\mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n - \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X_k^n| \leq C|v_{k-1}^n|^2 |X_k^n|, \tag{4.6}$$

see Section 5.2 for details, in particular Corollary 5.8. By the same reasoning as before, the constant C may be controlled independent of k.

Combining (4.5) and (4.6) and using that $v_n^n = \sum_{k=1}^n v_k^n - v_{k-1}^n$, we have

$$\left| v_n^n - \frac{1}{n}\sum_{k=1}^n \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X_k^n \right| \lesssim \frac{1}{n} + 1 \tag{4.7}$$

Consequently, using the Cauchy-Schwarz inequality, we find

$$\begin{aligned} \mathbb{E}(e^{n\langle \lambda, v_n^n\rangle}) &\leq e^{C|\lambda|} e^{nC|\lambda|} \mathbb{E}\left(e^{\sum_{i=1}^n \langle \lambda, \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{n-1}} X_k^n\rangle} \right) \\ &= e^{C|\lambda|} e^{nC|\lambda|} \mathbb{E}\left(e^{\langle \lambda, X_1\rangle} \right)^n. \end{aligned} \tag{4.8}$$

Here, the last line uses that the increments are independent and identically distributed. From this it follows that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{n\langle \lambda, v_n^n\rangle}) \leq C|\lambda| + \Lambda_{x_0}(\lambda),$$

so that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(v_n^n \in F) \leq - \inf_{v\in F} \sup_{\lambda \in T_{x_0}M} \{\langle \lambda, v\rangle - \Lambda_{x_0}(\lambda) - C|\lambda|\}.$$

It remains to get rid of the $C|\lambda|$ term. In the next step we show how to reduce the order n term in the upper bound in (4.8), so that we can still use the above estimating procedure to obtain the upper bound of the large deviation principle for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$.

4.2.3 Step 3: Reducing the upper bound in Step 2 by splitting the random walk in pieces

The problematic factor in estimate (4.8) arises from the replacement of the differential of the exponential map with parallel transport as done in Step 2. This error depends on $|v_k^n|$, i.e., the distance from x_0 to $(\frac{1}{n} * \mathcal{S})_k$. Note that in Step 2, we simply estimated $|v_k^n|$ uniformly in k. However, if we write r for the uniform bound on the increments, we actually have $|v_k^n| \leq \frac{k}{n} r$. Consquently, we can reduce the upper bound if the amount of steps for which we need to compare parallel transport and the differential of the exponential map becomes smaller.

To do this, the idea is to cut the random walk in finitely many pieces, say m, each consisting of (roughly) $m^{-1}n$ steps. We can then consider each of these pieces as separate random walks which we need to identify with a vector in some tangent space. In the end, we can then let the amount of pieces tend to infinity by considering the limit $m \to \infty$, so that the part of the upper bound which we want to reduce vanishes entirely.

More precisely, fix $m \in \mathbb{N}$, and define for $l = 0, \dots, m-1$ the indices $n_l = l\lfloor m^{-1}n \rfloor$ and set $n_m = n$. This divides the random walk in m pieces, where a piece starts in $(\frac{1}{n} * \mathcal{S})_{n_l}$ and consists of $\lfloor m^{-1}n \rfloor$ increments. Now recall there is a compact set $K \subset M$ such that for all n and all $0 \leq j \leq n$ we have $(\frac{1}{n} * \mathcal{S})_j \in K$. Because $\iota(K) > 0$, we can choose m sufficiently large, such that for all n, all $l = 1, \dots, m$ and all $k = 1, \dots, \lfloor m^{-1}n \rfloor$ we have

$$\left(\frac{1}{n} * \mathcal{S}\right)_{n_{l-1}+k} \in B\left(\left(\frac{1}{n} * \mathcal{S}\right)_{n_{l-1}}, \iota(K) \right).$$

Consequently, we may follow the same procedure as in Step 1, so that for every $l = 1, \dots, m$ and every $k = 1, \dots, \lfloor m^{-1}n \rfloor$ we can uniquely define $\tilde{v}_k^{n,m,l} \in T_{(\frac{1}{n} * \mathcal{S})_{n_l - 1}} M$

such that

$$\tilde{v}_k^{n,m,l} \in \operatorname{Exp}^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_{l-1}}} \left(\left(\frac{1}{n} * \mathcal{S} \right)_{n_{l-1}+k} \right)$$

and $|\tilde{v}_k^{n,m,l}| < \iota((\frac{1}{n} * \mathcal{S})_{n_{l-1}})$. Finally, we define $v_k^{n,m,l} \in T_{x_0}M$ by

$$v_k^{n,m,l} = \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{n_{l-1}}} \tilde{v}_k^{n,m,l},$$

where the parallel transport can be taken along any path connecting x_0 and $\left(\frac{1}{n} * \mathcal{S}\right)_{n_{l-1}}$, as long as it is measurable with respect to $\mathcal{F}_{n_{l-1}} = \sigma(X_1, \dots, X_{n_{l-1}})$.

This associates to $(\frac{1}{n} * \mathcal{S})_n \in M$ a tuple

$$\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor} \right) \in (T_{x_0}M)^m.$$

Following the procedure in Step 2, apart from some technical details, we find

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(e^{n\langle \lambda, v^{n,m,l}_{\lfloor m^{-1}n \rfloor}\rangle} \right) \leq C|\lambda| \frac{1}{m^3} + \frac{1}{m} \Lambda_{x_0}(\lambda),$$

for all $\lambda \in T_{x_0}M$. From here it is possible to show that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(e^{n\sum_{l=1}^m \langle \lambda_l, v^{n,m,l}_{\lfloor m^{-1}n \rfloor}\rangle} \right) \leq C \frac{1}{m^3} \sum_{l=1}^m |\lambda_l| + \frac{1}{m} \sum_{l=1}^m \Lambda_{x_0}(\lambda_l)$$

for all $(\lambda_1, \dots, \lambda_m) \in (T_{x_0}M)^m$. Consequently, we find that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor} \right) \in F \right) \\ &\leq - \inf_{(v_1,\dots,v_m)\in F} \frac{1}{m} \sum_{l=1}^m \sup_{\lambda \in T_{x_0}M} \{\langle \lambda, m v_l \rangle - \Lambda_{x_0}(\lambda) - \frac{1}{m^2} C|\lambda|\}. \end{aligned}$$

4.2.4 Step 4: Upper bound for the large deviation principle of $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$

To prove the large deviation upper bound for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$, we notice that the map sending $(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor})$ to $(\frac{1}{n} * \mathcal{S})_n$ is continuous. Hence, if $F \subset M$ is closed, there exists a closed set $\tilde{F} \subset (T_{x_0}M)^m$ such that

$$\mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S} \right)_n \in F \right) = \mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor} \right) \in \tilde{F} \right).$$

From this it follows that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S} \right)_n \in F \right) \\ &\leq - \inf_{(v_1,\dots,v_m)\in \tilde{F}} \frac{1}{m} \sum_{l=1}^m \sup_{\lambda \in T_{x_0}M} \{\langle \lambda, v_l \rangle - \Lambda_{x_0}(\lambda) - \frac{1}{m^2} C|\lambda|\}. \end{aligned}$$

Now note that for every $v \in \operatorname{Exp}_{x_0}^{-1} F$ we have that $(\frac{1}{m}v, \dots, \frac{1}{m}v) \in \tilde{F}$. Furthermore, by convexity, the infimum in the upper bound is attained when all v_i are equal. Consequently, the upper bound reduces to

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S} \right)_n \in F \right) \\ &\leq - \inf_{v\in \operatorname{Exp}_{x_0}^{-1} F} \sup_{\lambda \in T_{x_0}M} \{\langle \lambda, v \rangle - \Lambda_{x_0}(\lambda) - \frac{1}{m^2} C|\lambda|\}. \end{aligned}$$

The desired upper bound now follows by considering the limit $m \to \infty$.

4.3 Sketch of the proof of the lower bound

To prove the lower bound of the large deviation principle for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$, it suffices to show that if $G \subset M$ is open, then

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in G\right) \geq -I_M(x),$$

for all $x \in G$. Because $I_M(x) = \inf_{v\in \operatorname{Exp}_{x_0}^{-1} x} \Lambda^*_{x_0}(v)$, it is in fact sufficient to show that

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in G\right) \geq -\Lambda^*_{x_0}(v)$$

for any $v \in \operatorname{Exp}_{x_0}^{-1} G$. Consequently, we again need to transfer the problem to the tangent space $T_{x_0}M$.

4.3.1 Transfer to the tangent space $T_{x_0}M$

Similar to how estimate (4.7) is derived, we find that

$$\left| v^n_{\lfloor m^{-1}n\rfloor} - \frac{1}{n} \sum_{k=1}^{\lfloor m^{-1}n\rfloor} \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X^n_k \right| \lesssim \frac{1}{nm} + \frac{1}{m^3}. \tag{4.9}$$

Consequently, by choosing m sufficiently large, we can get $v^n_{\lfloor m^{-1}n\rfloor}$ arbitrarily close to $\frac{1}{n}\sum_{k=1}^{\lfloor m^{-1}n\rfloor} \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X^n_k$. Using the fact that the increments of the geodesic random walk are independent and identically distributed, we prove that $\sum_{k=1}^{\lfloor m^{-1}n\rfloor} \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X^n_k$ is a sum of independent random variables, each distributed according to μ_{x_0}. Consequently, by Cramér's theorem for vector spaces, for every $m \in \mathbb{N}$ the sequence $\{\frac{1}{n}\sum_{k=1}^{\lfloor m^{-1}n\rfloor} \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X^n_k\}_{n\geq 0}$ satisfies the large deviation principle in $T_{x_0}M$ with good rate function $I(v) = \frac{1}{m}\Lambda^*_{x_0}(mv)$.

Putting everything together, after some technicalities, we find that if $\varepsilon > 0$ is small enough, there exists a constant $c \in (0,1)$ such that for m large enough

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(v^n_{\lfloor m^{-1}n\rfloor} \in B(v,\varepsilon)) \\
&\geq \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{\lfloor m^{-1}n\rfloor} \tau^{-1}_{x_0 \frac{1}{n}\mathcal{S}_{k-1}} X^n_k \in B(v, c\varepsilon^2)\right) \\
&\geq \frac{1}{m}\Lambda^*_{x_0}(mv).
\end{aligned} \tag{4.10}$$

In order to make use of this fact, we again need to divide the random walk in pieces, like in Step 3 in Section 4.2. Consequently, we again first identify $(\frac{1}{n} * \mathcal{S})_n \in M$ with a tuple

$$\left(\tilde{v}^{n,m,1}_{\lfloor m^{-1}n\rfloor}, \dots, \tilde{v}^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right) \in T_{(\frac{1}{n}*\mathcal{S})_{n_0}} M \times \cdots \times T_{(\frac{1}{n}*\mathcal{S})_{n_m}} M.$$

However, this time we need to be careful how we transport these vectors to $T_{x_0}M$. Indeed, we wish to do this in such a way that

$$\left(v^{n,m,1}_{\lfloor m^{-1}n\rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right) \in B(v, c\varepsilon)^m \Rightarrow \left(\frac{1}{n} * \mathcal{S}\right)_n \in B(\operatorname{Exp}_{x_0} v, \varepsilon). \tag{4.11}$$

The key to making the correct choice is given by Proposition 5.10, which gives us control over how far geodesics can spread in a short time when starting in different

points of the manifold. This result shows us how to choose the parallel transport based on the vector v, so that the curvature has only little effect. Essentially, one first transports a vector to an associated point on the geodesic with speed v which connects x_0 and x. After that, one transports the vector along this geodesic to x_0. More precisely, we do the following:

1. Consider the geodesic $\gamma(t) = \operatorname{Exp}_{x_0}(tv)$ and for $i = 0, \dots, m$ define the points $y_i = \gamma(\frac{i}{m})$. Note that $y_0 = x_0$.

2. For every $i = 0, \dots, m$ and every $x \in M$, choose a geodesic of minimal length connecting y_i and x and define $\tau_{y_i x}$ to be parallel transport along this geodesic.

3. Now define for $i = 1, \dots, m$ the vector $v^{n,m,1}_{\lfloor m^{-1}n \rfloor} \in T_{x_0}M$ by

$$v^{n,m,i}_{\lfloor m^{-1}n \rfloor} = \tau^{-1}_{y_0 y_i} \tau^{-1}_{y_i (\frac{1}{n} * \mathcal{S})_{n_i - 1}} \tilde{v}^{n,m,i}_{\lfloor m^{-1}n \rfloor}$$

Now, given $G \subset M$ open, $x \in G$ and $v \in \operatorname{Exp}_{x_0}^{-1} x$, by (4.11) we have

$$\mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in G\right) \geq \mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor}\right) \in B(v, c\varepsilon^2)^m\right).$$

Using this, an approach similar to the one used to obtain (4.10), also using that the increments are independent and identically distributed, gives us that

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in G\right) \\
&\geq \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor}\right) \in B(v, c\varepsilon^2)^m\right) \\
&\geq -\Lambda^*_{x_0}(v),
\end{aligned}$$

which is as desired.

5 Geometric results for the proof

This section focuses on geometric results needed for the proof of Cramér's theorem for geodesic random walks as sketched in Section 4. We obtain a Taylor expansion for the inverse Riemannian exponential map and estimate the residual term. Furthermore, we bound the difference between the differential of the Riemannian exponential map and parallel transport. This heavily relies on the theory of Jacobi fields, which have been introduced in Section 2.2.3. We conclude this section by proving how far geodesics can spread in a short time interval when starting in different points on the manifold.

5.1 Taylor expansion of the inverse Riemannian exponential map

The Riemannian exponential map $\operatorname{Exp}_x : T_xM \to M$ is a local diffeomorphism around 0. More precisely, it is a diffeomorphism between $B(0, \iota(x)) \subset T_xM$ and $\operatorname{Exp}_x(B(0, \iota(x)))$. Now suppose $\gamma(t)$ is a curve in $\operatorname{Exp}_x(B(0, \iota(x)))$. There exists a unique curve $w(t)$ in $B(0, \iota(x)) \subset T_xM$ such that $\operatorname{Exp}_x w(t) = \gamma(t)$. Our aim is to find a Taylor expansion for $w(t)$ around $t = 0$. Although this seems to be folklore, we also find a precise estimate of the residual term of the Taylor approximation.

Before we can do this, we first need two lemmas that will help us control the error term in the first order Taylor polynomial for the inverse of the Riemannian exponential map.

Lemma 5.1. *Let $K \subset M$ be compact and for any $x \in K$, let $K_x \subset T_xM$ be compact. Assume there exists a $C > 0$ such that $K_x \subset \overline{B(0,C)}$ for any $x \in K$. Then*

$$\sup_{x\in K}\sup_{v\in K_x} |\mathrm{d}(\mathrm{Exp}_x)_v| < \infty$$

Proof. Because the sets K_x are uniformly bounded and K is compact, it follows that

$$\{(x,v) \in TM | x \in K, v \in K_x\}$$

is compact.

Now fix $x \in M$ and $v \in T_xM$. Because the Riemannian exponential map $\mathrm{Exp} : TM \to M \times M$ is continuous, there exists a neighbourhood $U \subset TM$ of (x,v) such that

$$\mathrm{Exp}(U) \subset B(x, \iota(x)) \times B(\mathrm{Exp}_x v, \iota(\mathrm{Exp}_x v)).$$

Now for $(y,w) \in U$, and any $u \in T_{\mathrm{Exp}_x v}M$ we define

$$F_{u,x,v} : (y,w) \mapsto |\tau_{\mathrm{Exp}_y w \mathrm{Exp}_x v}\mathrm{d}(\mathrm{Exp}_y)_w \tau_{xy} u|$$

where parallel transport is taken along the unique minimizing geodesic connecting the two points, which exists by the choice of U. We argue that $F_{u,x,v}$ is continuous for any $u \in T_{\mathrm{Exp}_x v}M$. By the choice of parallel transport, $\tau_{\mathrm{Exp}_y w \mathrm{Exp}_x v}$ and τ_{xy} are continuous. Furthermore, note that we can write

$$\mathrm{d}(\mathrm{Exp}_y)_w \tilde{u} = \mathrm{dExp}((y,w), \tilde{u}).$$

Because Exp is smooth, it follows that dExp is continuous. Consequently, $F_{u,x,v}$ is a composition of continuous maps, and hence continuous on U.

Since TM is locally Euclidean, we can find a relatively compact set $U_{(x,v)}$ containing (x,v), such that $\overline{U_{(x,v)}} \subset U$.

Because the set $\{(x,v) \in TM | x \in K, v \in K_x\}$ is compact, we can find $(x_1,v_1), \ldots, (x_k,v_k)$ such that

$$\{(x,v) \in TM | x \in K, v \in K_x\} \subset \bigcup_{i=1}^{k} U_{(x_i,v_i)}.$$

Consequently, we have that

$$\sup_{x\in K}\sup_{v\in K_x} |\mathrm{d}(\mathrm{Exp}_x)_v| \leq \max_{i=1}^{k} \sup_{(x,v)\in\overline{U_{(x_i,v_i)}}} |\mathrm{d}(\mathrm{Exp}_x)_v|.$$

It follows that we are done once we show that

$$\sup_{(x,v)\in\overline{U_{(x_i,v_i)}}} |\mathrm{d}(\mathrm{Exp}_x)_v| < \infty$$

for all $i = 1, \ldots, k$.

For this, remember that F_{u,x_i,v_i} is continuous on $\overline{U_{(x_i,v_i)}}$, and hence bounded for any u, since $\overline{U_{(x_i,v_i)}}$ is compact. Consequently, it follows from the uniform boundedness principle that

$$\sup_{(x,v)\in\overline{U_{(x_i,v_i)}}} |\tau_{\mathrm{Exp}_x v \mathrm{Exp}_{x_i} v_i}\mathrm{d}(\mathrm{Exp}_x)_v \tau_{x_i x}| < \infty$$

However, because parallel transport is an isometry, we have

$$|\mathrm{d}(\mathrm{Exp}_x)_v| = |\tau_{\mathrm{Exp}_x v \mathrm{Exp}_{x_i} v_i}\mathrm{d}(\mathrm{Exp}_x)_v \tau_{x_i x}|,$$

which concludes the proof. □

As long as one restricts to a set where the inverse of the Riemannian exponential map is well-defined, one obtains in a similar way a bound for the differential of the inverse Riemannian exponential map.

Lemma 5.2. *Let $K \subset M$ be compact and for any $x \in K$, let $K_x \subset B(0, \iota(x)) \subset T_xM$ be compact. Assume that there exists a constant $C > 0$ such that $K_x \subset \overline{B(0,C)}$ for any $x \in K$. Then*

$$\sup_{x\in K} \sup_{v\in K_x} |\mathrm{d}(\mathrm{Exp}_x)_v^{-1}| < \infty.$$

Remark 5.3. When we take $K = \{x_0\}$ in Lemma 5.2, the statement simplifies as follows: If $\tilde{K} \subset B(0, \iota(x_0))$ is compact, then

$$\sup_{v\in \tilde{K}} |\mathrm{d}(\mathrm{Exp}_{x_0})_v^{-1}| < \infty.$$

We are now in a position to find a first order Taylor expansion of the inverse Riemannian exponential map and control the error term appropriately.

Proposition 5.4. *Fix $x_0 \in M$ and let $K \subset B(0, \iota(x_0))$ be compact. Define $\tilde{K} = \mathrm{Exp}_{x_0} K$ and let $x \in \tilde{K}$ and $v \in T_xM$. Consider the geodesic $\gamma_v : [0,T] \to M$ defined by $\gamma_v(t) = \mathrm{Exp}_x(tv)$, where T is such that the image of γ_v is contained in $\tilde{K}$. Restrict Exp_{x_0} to K and set $w(t) = \mathrm{Exp}_{x_0}^{-1}(\gamma_v(t)) \in K \subset T_{x_0}M$. Then there exists a constant $C > 0$ such that*

$$|w(t) - w(0) - t\mathrm{d}(\mathrm{Exp}_{x_0})_{w(0)}^{-1}(v)|_{g(x_0)} \leq Ct^2$$

for all $t \in [0,T]$. Here, the constant C only depends on the compact set $\tilde{K}$.

Proof. First observe that $w(t)$ is well-defined, because $K \subset B(0, \iota(x_0))$ so that the restriction of Exp_{x_0} to K is injective. Moreover, it is actually a diffeomorphism onto $\tilde{K}$, and thus $\mathrm{d}(\mathrm{Exp}_{x_0})_w$ is also injective. By the inverse function theorem, Exp_{x_0} has a differentiable inverse, whose derivative at w is given by

$$\mathrm{d}(\mathrm{Exp}_{x_0}^{-1})(w) = \mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1}$$

Consequently, by Taylor's theorem, we find for any $t \in [0,T]$ that

$$w(t) = w + t\mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1}(v) + t^2 w''(\xi_t)$$

for some $\xi_t \in (0,t)$.

To control the error term, we estimate $|w''(t)|$. We have

$$w''(t) = \lim_{h\to 0} \frac{\mathrm{d}(\mathrm{Exp}_{x_0})_{w(t+h)}^{-1}(\dot{\gamma}_v(t+h)) - \mathrm{d}(\mathrm{Exp}_{x_0})_{w(t)}^{-1}(\dot{\gamma}_v(t))}{h}.$$

We estimate the numerator to find a desired bound on $w''(t)$. Set

$$u = \mathrm{d}(\mathrm{Exp}_{x_0})_{w(t)}^{-1}(\dot{\gamma}_v(t)) \in T_{x_0}M \tag{5.1}$$

and

$$\tilde{u} = \mathrm{d}(\mathrm{Exp}_{x_0})_{w(t+h)}^{-1}(\dot{\gamma}_v(t+h)) \in T_{x_0}M. \tag{5.2}$$

Then

$$\dot{\gamma}_v(t) = \mathrm{d}(\mathrm{Exp}_{x_0})_{w(t)}(u)$$

and

$$\dot{\gamma}_v(t+h) = \mathrm{d}(\mathrm{Exp}_{x_0})_{w(t+h)}(\tilde{u}).$$

As γ_v is a geodesic, we have $\dot\gamma_v(t+h) = \tau_{\gamma_v(t)\gamma_v(t+h)}\dot\gamma_v(t)$. Consequently, we obtain

$$\mathrm{d}(\mathrm{Exp}_{x_0})_{w(t+h)}(\tilde u) = \tau_{\gamma_v(t)\gamma_v(t+h)}\mathrm{d}(\mathrm{Exp}_{x_0})_{w(t)}(u). \tag{5.3}$$

Define the curves ψ_1, ψ_2 in $T_{x_0}M$ by

$$\psi_1(s) = w(t) + su, \qquad \psi_2(s) = w(t+h) + s\tilde u$$

and the corresponding curves ϕ_1, ϕ_2 in M by

$$\phi_1(s) = \mathrm{Exp}_{x_0}(w(t) + su), \qquad \phi_2(s) = \mathrm{Exp}_{x_0}(w(t+h) + s\tilde u).$$

The aim is to control $|u - \tilde u|_{g(x_0)}$. For this, take normal coordinates around x_0 (which can be taken to cover all of $\tilde K$, because $\tilde K \subset \mathrm{Exp}_{x_0}[B(0, \iota(x_0))]$). In these coordinates, let us write $u = u^i\partial_i(x_0)$ and $\tilde u = \tilde u^j\partial_j(x_0)$. Note that in coordinates

$$\phi_1(s) = (w^1(t) + su^1, \dots, w^d(t) + su^d)$$

and

$$\phi_2(s) = (w^1(t+h) + s\tilde u^1, \dots, w^d(t+h) + s\tilde u^d).$$

Consequently,

$$\dot\phi_1(s) = u^i\partial_i(\phi_1(s))$$

and

$$\dot\phi_2(s) = \tilde u^j\partial_j(\phi_2(s)).$$

By equation (5.3) we have $\dot\phi_2(0) = \tau_{\gamma_v(t)\gamma_v(t+h)}\dot\phi_1(0)$. But then we find that the coefficients of $\dot\phi_2(0)$ satisfy the equations

$$\dot V^k(s) + \Gamma^k_{ij}(\gamma_v(t+s))\dot\gamma^i_v(t+s)V^j(s) = 0$$

with $V^k(0) = \dot\phi^k_1(0)$. Consequently, using a Taylor expansion, we find

$$\dot\phi^k_2(0) = \dot\phi^k_1(0) - h\Gamma^k_{ij}(\gamma_v(t))\dot\gamma^i_v(t)\dot\phi^j_1(0) + \mathcal{O}(h^2)$$

Using that $\dot\phi^k_1(0) = u^k$ and $\dot\phi^k_2(0) = \tilde u^k$, we obtain

$$u^k - \tilde u^k = h\Gamma^k_{ij}(\gamma_v(t))\dot\gamma^i_v(t)u^j + \mathcal{O}(h^2). \tag{5.4}$$

Because we are using normal coordinates around x_0, we have

$$|u - \tilde u|^2_{g(x_0)} = \sum_{k=1}^d (u^k - \tilde u^k)^2.$$

If we plug in expression (5.4), we get

$$|u - \tilde u|^2_{g(x_0)} = h^2\sum_{k=1}^d (\Gamma^k_{ij}(\gamma_v(t))\dot\gamma^i_v(t)u^j)^2 + \mathcal{O}(h^3).$$

As the Christoffel symbols are continuous, they are bounded on our compact set $\tilde K$ by some constant C_1. Furthermore, the coefficients g_{ij} of the metric are also continuous, and in particular, by the positive definiteness of the metric, there exists a uniform constant $\delta > 0$ such that $g_{ii}(x) \geq \delta$ for all $x \in \tilde K$ and all $i = 1, \dots, d$. In particular, this implies

$$(\dot\gamma^i_v(t))^2 \leq \delta^{-1}|\dot\gamma_v(t)|^2_{g(\gamma_v(t))} = \delta^{-1}|v|^2_{g(\gamma_v(0))}.$$

Similarly, we have

$$(u^j)^2 \leq |\mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{w(t)}(\dot{\gamma}_v(t))|^2_{g(x_0)} \leq C_2^2 |\dot{\gamma}_v(t)|^2_{g(\gamma_v(t))} = C_2^2 |v|^2_{g(\gamma_v(0))},$$

where we used Lemma 5.2 to find the constant C_2, which again only depends on the compact set $\tilde{K}$.

Collecting everything, we find

$$|u - \tilde{u}|^2_{g(x_0)} \leq C_1^2 C_2^2 |v|^2_{g(\gamma_v(t))} h^2 + \mathcal{O}(h^3)$$

Recalling the definition of u and $\tilde{u}$ in (5.1) and (5.2) respectively, we find after taking the limit $h \to 0$ that

$$|w''(t)|_{g(x_0)} \leq C_1 C_2 |v|_{g(\gamma_v(0))},$$

which provides the desired constant, because C_1, C_2 only depend on $\tilde{K}$. □

5.2 Differential of the Riemannian exponential map and parallel transport

Next, we wish to understand the relation between the differential of the Riemannian exponential map and parallel transport. Before we can make the appropriate comparison, we first need a version of Taylor's theorem suitable for vector fields along a curve on a manifold.

Proposition 5.5 (Taylor's theorem)**.** *Let γ be a curve in M and v a vector field along γ. Define $D_t v(t) := \nabla_{\dot{\gamma}(t)} v(t)$ and D_t^k as the k-th covariant derivative in this way. Fix $n \in \mathbb{N}$. For every $t > 0$ there exists $\xi_t \in (0, t)$ such that*

$$v(t) = \sum_{k=0}^{n} \frac{t^k}{k!} \tau_{\gamma(0)\gamma(t)} D_t^k v(0) + \frac{t^{k+1}}{(k+1)!} \tau_{\gamma(\xi_t)\gamma(t)} D_t^{k+1} v(\xi_t).$$

Proof. Consider the map $f(t) = \tau^{-1}_{\gamma(0)\gamma(t)} v(t)$, mapping into $T_{\gamma(0)}M$. Because f is smooth, by Taylor's theorem, given $n \in \mathbb{N}$ and $t > 0$, there exists $\xi_t \in (0, t)$ such that

$$f(t) = \sum_{k=0}^{n} \frac{t^k}{k!} f^{(k)}(0) + \frac{t^{k+1}}{(k+1)!} f^{(k+1)}(\xi_t).$$

Let us compute the derivatives of f. Note that

$$\begin{aligned} f'(t) &= \lim_{h \to 0} \frac{f(t+h) - f(t)}{h} \\ &= \lim_{h \to 0} \frac{\tau^{-1}_{\gamma(0)\gamma(t+h)} v(t+h) - \tau^{-1}_{\gamma(0)\gamma(t)} v(t)}{h} \\ &= \tau^{-1}_{\gamma(0)\gamma(t)} \lim_{h \to 0} \frac{\tau^{-1}_{\gamma(t)\gamma(t+h)} v(t+h) - v(t)}{h} \\ &= \tau^{-1}_{\gamma(0)\gamma(t)} D_t v(t). \end{aligned}$$

Using induction, one can show that

$$f^{(k)}(t) = \tau^{-1}_{\gamma(0)\gamma(t)} D_t^k v(t)$$

for all $k \in \mathbb{N}$. But then we find that

$$\tau^{-1}_{\gamma(0)\gamma(t)} v(t) = \sum_{k=0}^{n} \frac{t^k}{k!} D_t^k v(0) + \frac{t^{k+1}}{(k+1)!} \tau^{-1}_{\gamma(0)\gamma(\xi_t)} D_t^{k+1} v(\xi_t).$$

Applying $\tau_{\gamma(0)\gamma(t)}$ to both sides and observing that $t > \xi_t$ gives the desired result. □

We are now able to compare the differential of the Riemannian exponential map and parallel transport. The Taylor series of the differential of the exponential map may be found in e.g. [13, Appendix A]. The error term for finite Taylor polynomials seems to belong to folklore, but we insert a proof here for the reader's convenience.

Proposition 5.6. *Let $x_0 \in M$ and take $w, u \in T_{x_0}M$. Consider the geodesic $\gamma_w : [0,1] \to M$ given by $\gamma_w(t) = \mathrm{Exp}_{x_0}(tw)$. For every $t \in [0,1]$ there exists $\xi_t \in (0,t)$ such that*

$$\mathrm{d}(\mathrm{Exp}_{x_0})_{tw}(u) = \tau_{\gamma_w(0)\gamma_w(t)}u + \frac{1}{2}t\tau_{\gamma_w(\xi_t)\gamma_w(t)}R_{\gamma_w(\xi_t)}(\mathrm{d}(\mathrm{Exp}_{x_0})_{\xi_t w}(\xi_t u), \dot\gamma_w(\xi_t))\dot\gamma_w(\xi_t).$$

Proof. Consider the vector field $J(t) = \mathrm{d}(\mathrm{Exp}_{x_0})_{tw}(tu)$ along $\gamma_w(t)$. As argued in Section 2.2.3, $J(t)$ is a Jacobi field along $\gamma(t)$ with $J(0) = 0$ and $\dot J(0) = u$. By the Jacobi equation (2.3), the second derivative is given by

$$D_t^2 J(t) = -R_{\gamma_w(t)}(J(t), \dot\gamma_w(t))\dot\gamma_w(t).$$

Consequently, by Proposition 5.5 we find there exists some $\xi_t \in (0,t)$ such that

$$J(t) = t\tau\gamma_w(0)\gamma_w(t)u - \frac{1}{2}t^2\tau_{\gamma_w(\xi_t)\gamma_w(t)}R_{\gamma_w(\xi_t)}(\mathrm{d}(\mathrm{Exp}_{x_0})_{\xi_t w}(\xi_t u), \dot\gamma_w(\xi_t))\dot\gamma_w(\xi_t).$$

The result now follows after dividing by t. □

This proposition allows us to obtain the following estimate.

Corollary 5.7. *Fix $x_0 \in M$ and let $w \in B(0, \iota(x_0)) \subset T_{x_0}M$. Define the geodesic $\gamma_w : [0,1] \to M$ by $\gamma_w(t) = \mathrm{Exp}_{x_0}(tw)$. There exists a constant $C > 0$ only depending on some compact set containing γ_w such that*

$$|\mathrm{d}(\mathrm{Exp}_{x_0})_w(u) - \tau_{\gamma_w(0)\gamma_w(1)}u|_{g(\gamma_w(1))} \leq C|u|_{g(x_0)}|w|^2_{g(x_0)}$$

for all $u \in T_{x_0}M$.

Proof. By Proposition 5.6 there exists $\xi \in (0,1)$ such that

$$\mathrm{d}(\mathrm{Exp}_{x_0})_w(u) - \tau_{\gamma_w(0)\gamma_w(1)}u = -\frac{1}{2}\tau_{\gamma_w(\xi),\gamma_w(1)}R_{\gamma_w(\xi)}(\mathrm{d}(\mathrm{Exp}_{x_0})_{\xi w}(\xi u), \dot\gamma_w(\xi))\dot\gamma_w(\xi).$$

Now taking norms on both sides, we first observe that the norm of the Riemann curvature endomorphism is bounded on compact sets, because it is continuous (in coordinates the norm can be expressed as a continuous functions of the coefficients). Furthermore, by Lemma 5.1 we have that $w \mapsto |\mathrm{d}(\mathrm{Exp}_{x_0})_w|$ is bounded on compact sets.

We thus obtain constants $C_1, C_2 > 0$, only depending on some compact set containing the curve γ_w such that

$$\begin{aligned}
&|\mathrm{d}(\mathrm{Exp}_{x_0})_w(u) - \tau_{\gamma_w(0)\gamma_w(1)}u|_{g(\gamma_w(1))} \\
&\leq \frac{1}{2}|R_{\gamma_w(\xi)}(\mathrm{d}(\mathrm{Exp}_{x_0})_{\xi w}(\xi u), \dot\gamma_w(\xi))\dot\gamma_w(\xi)|_{g(\gamma_w(\xi))} \\
&\leq C_1|\mathrm{d}(\mathrm{Exp}_{x_0})_{\xi w}(\xi u)|_{g(\gamma_w(\xi))}|\dot\gamma_w(\xi)|^2_{g(\gamma_w(\xi))} \\
&\leq C_1C_2|u|_{g(x_0)}|w|^2_{g(x_0)}.
\end{aligned}$$

Here, in the last line we used that $\xi < 1$ and the fact that γ_w is a geodesic. □

The result in the latter corollary can also be used to compare the inverse of the differential of the exponential map to the inverse of parallel transport, which itself is parallel transport, but in the opposite direction.

Corollary 5.8. *Let $x_0 \in M$ and fix $w \in B(0, \iota(x_0)) \subset T_{x_0}M$. Define the geodesic $\gamma_w : [0,1] \to M$ by $\gamma_w(t) = \mathrm{Exp}_{x_0}(tw)$. Then there exists a constant $C > 0$ only depending on some compact set containing γ_w, such that*

$$|\mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1}(u) - \tau_{\gamma_w(0)\gamma_w(1)}^{-1} u|_{g(\gamma_w(1))} \leq C|u|_{g(\gamma_w(1))}|w|_{g(x_0)}^2$$

for all $u \in T_{\gamma_w(1)}M$.

Proof. Fix $u \in T_{\gamma_w(1)}M$ and consider $\mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1} u \in T_{x_0}M$. By Corollary 5.7, there exists a constant $C > 0$ only depending on a compact set containing γ_w such that

$$|u - \tau_{\gamma_w(0)\gamma_w(1)}\mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1} u|_{g(\gamma_w(1))} \leq C|\mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1} u|_{g(x_0)}|w|_{g(x_0)}^2.$$

Because parallel transport is an isometry, the left hand side is equal to

$$|\tau_{\gamma_w(1)\gamma_w(0)} u - \mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1} u|_{g(\gamma_w(1))}.$$

For the right hand side, we observe that by Lemma 5.2 there exists a constant $\tilde{C} > 0$, only depending on some compact set containing γ_w such that

$$|\mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1} u|_{g(x_0)} \leq \tilde{C}|u|_{g(\gamma_w(1))}.$$

Putting everything together, we find

$$|\tau_{\gamma_w(1)\gamma_w(0)} u - \mathrm{d}(\mathrm{Exp}_{x_0})_w^{-1} u|_{g(\gamma_w(1))} \leq C\tilde{C}|u|_{g(\gamma_w(1))}|w|_{g(x_0)}^2$$

as desired. □

5.3 Spreading of geodesics

We conclude this section with a result on how far geodesics, possibly starting in different points, can spread in a given amount of time. To shed some light on the upcoming result, we first consider the Euclidean case. For this, let $\gamma(t) = \gamma(0) + t\dot{\gamma}(0)$ and $\phi(t) = \phi(0) + t\dot{\phi}(t)$ be two straight lines. Then

$$|\gamma(t) - \phi(t)|^2 = |\gamma(0) - \phi(0)|^2 + 2t\langle \dot{\gamma}(0) - \dot{\phi}(0), \gamma(0) - \phi(0)\rangle + t^2|\dot{\gamma}(t) - \dot{\phi}(t)|^2.$$

It turns out that in a Riemannian manifold, this formula is analogous up to first order. The curvature terms show up in the second order term. Before we prove this, we first need a lemma.

Lemma 5.9. *Let $K \subset M$ be compact and fix $L > 0$. Let $0 < r < \iota(K)$. Let $\phi : [0,T] \to M$ and $\gamma : [0,T] \to M$ be two geodesics contained in K. Assume that $d(\phi(0), \gamma(0)) \leq \frac{r}{2}$ and $|\dot{\phi}(0)|, |\dot{\gamma}(0)| \leq L$. Then there exists a $t_0 > 0$, only depending on K, L and r, such that for all $0 \leq t \leq t_0$ we have*

$$d(\phi(t), \gamma(t)) < r.$$

Proof. Because $d : M \times M \to \mathbb{R}$ is continuous, and $K \times K$ is compact, $d(\cdot,\cdot)$ is uniformly continuous on $K \times K$. Consequently, pick $\varepsilon > 0$ such that $|d(x,y) - d(x',y')| < \frac{r}{2}$, whenever $d(x,x') < \varepsilon$ and $d(y,y') < \varepsilon$.

Now observe that $d(\phi(t), \phi(0)) \leq t|\dot{\phi}(0)| \leq tL$ and likewise $d(\gamma(t), \gamma(0)) \leq tL$. Hence, if we take $t_0 < \varepsilon L^{-1}$, then for all $0 \leq t \leq t_0$ we have $d(\phi(t), \phi(0)) < \varepsilon$ and $d(\gamma(t), \gamma(0)) < \varepsilon$. By the choice of ε, it follows that

$$|d(\phi(0), \gamma(0)) - d(\phi(t), \gamma(t))| < \frac{r}{2}.$$

Since $d(\phi(0), \gamma(0)) \leq \frac{1}{2}r$, the above then implies that $d(\phi(t), \gamma(t)) < r$ as desired. □

Proposition 5.10. *Let $K \subset M$ be compact and fix $L > 0$. Let $0 < r < \iota(K)$ and fix $t_0 > 0$ as in Lemma 5.9. Let $\phi : [0, t_0] \to M$ and $\gamma : [0, t_0] \to M$ be two geodesics in K such that $d(\gamma(0), \phi(0)) \leq \frac{r}{2}$ and $|\dot\phi(0)|, |\dot\gamma(0)| \leq L$. Finally, let $\tilde K$ be a compact set containing all geodesics of minimal length between points in K. Then for all $0 \leq t \leq t_0$ we have*

$$\begin{aligned}&d(\gamma(t), \phi(t))^2\\ &\leq d(\gamma(0), \phi(0))^2 + 2t\langle \tau^{-1}_{\phi(0)\gamma(0)}\dot\gamma(0) - \dot\phi(0), \mathrm{Exp}^{-1}_{\phi(0)}\gamma(0)\rangle + t^2 C(|\dot\gamma(0)| + |\dot\phi(0)|),\end{aligned}$$

where the constant $C > 0$ only depends on $\tilde K, L$ and r.

Proof. Define $f(t) = d(\gamma(t), \phi(t))^2$. By the choice of t_0, Lemma 5.9 gives us that

$$d(\phi(t), \gamma(t)) < r < \iota(K)$$

for every $0 \leq t \leq t_0$. Consequently, $\phi(t)$ and $\gamma(t)$ may be joined by a unique geodesic of minimal length. Moreover, by restricting Exp, we have $f(t) = |\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)|^2$. Consequently, we can compute

$$\begin{aligned}f'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}|\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)|^2\\ &= 2\langle \nabla_{\dot\phi(t)}\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t), \mathrm{Exp}^{-1}_{x_0}\gamma(t)\rangle.\end{aligned}$$

Now define the variation of curves $\Gamma : [0, t_0] \times [0, 1] \to M$ by

$$\Gamma(t, s) = \mathrm{Exp}_{\phi(t)}(s\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)).$$

Then for each t, the curve $s \mapsto \Gamma(t, s)$ is the geodesic of minimal length between $\phi(t)$ and $\gamma(t)$. Hence, $\Gamma([0, t_0] \times [0, 1]) \subset \tilde K$. Furthermore, because Γ is a variation of geodesics, the vector field

$$J_t(s) = \partial_t\Gamma(t, s)$$

is a Jacobi field along the curve $\Gamma_t(s) := \Gamma(t, s)$ for all $0 \leq t \leq t_0$.

Now note that by the Symmetry Lemma (Lemma 2.2), we have

$$\nabla_{\dot\phi(t)}\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t) = D_t\partial_s\Gamma(t, 0) = D_s\partial_t\Gamma(t, 0) = \dot J_t(0).$$

Consequently, we obtain

$$f'(t) = 2\langle \dot J_t(0), \mathrm{Exp}^{-1}_{x_0}\gamma(t)\rangle = 2\langle \dot J_t(0), \partial_s\Gamma(t, 0)\rangle.$$

By Proposition 2.3 we find

$$\begin{aligned}f'(t) &= 2\langle \dot J_t(0), \partial_s\Gamma(t, 0)\rangle\\ &= 2\langle J_t(1), \partial_s\Gamma(t, 1)\rangle - 2\langle J_t(0), \partial_s\Gamma(t, 0)\rangle\\ &= 2\langle \dot\gamma(t), -\mathrm{Exp}_{\gamma(t)}\phi(t)\rangle - 2\langle \dot\phi(t), \mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)\rangle\\ &= 2\langle \tau^{-1}_{\phi(t)\gamma(t)}\dot\gamma(t) - \dot\phi(t), \mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)\rangle.\end{aligned}$$

Consequently, we have

$$f'(0) = 2\langle \tau^{-1}_{\phi(0)\gamma(0)}\dot\gamma(0) - \dot\phi(0), \mathrm{Exp}^{-1}_{\phi(0)}\gamma(0)\rangle.$$

By Taylor's theorem, we find that

$$d(\gamma(t), \phi(t))^2 \leq d(\gamma(0), \phi(0))^2 + 2t\langle \tau^{-1}_{\phi(0)\gamma(0)}\dot\gamma(0) - \dot\phi(0), \mathrm{Exp}^{-1}_{\phi(0)}\gamma(0)\rangle + \frac{1}{2}t^2 \sup_{\xi\in[0,t]} |f''(\xi)|.$$

We now turn to estimating the residual term. For this, we compute $f''(t)$ as follows:

$$\begin{aligned}\frac{1}{2}f''(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\langle\dot\gamma(t), -\mathrm{Exp}_{\gamma(t)}\phi(t)\rangle - \frac{\mathrm{d}}{\mathrm{d}t}\langle\dot\phi(t), \mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)\rangle\\ &= -\langle\dot\gamma(t), \nabla_{\dot\gamma(t)}\mathrm{Exp}^{-1}_{\gamma(t)}\phi(t)\rangle - \langle\dot\phi(t), \nabla_{\dot\phi(t)}\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)\rangle\\ &= \langle\dot\gamma(t), \partial_t\Gamma(t,1)\rangle - \langle\dot\phi(t), \partial_t\Gamma(t,0)\rangle\\ &= \langle\dot\gamma(t), \dot J_t(1)\rangle - \langle\dot\phi(t), \dot J_t(0)\rangle.\end{aligned}$$

Here we used that $\nabla_{\dot\phi(t)}\dot\phi(t) = \nabla_{\dot\gamma(t)}\dot\gamma(t) = 0$, since ϕ and γ are geodesics. Consequently, we have

$$\frac{1}{2}|f''(t)| \le |\dot\gamma(t)||\dot J_t(1)| + |\dot\phi(t)||\dot J_t(0)| = |\dot\gamma(0)||\dot J_t(1)| + |\dot\phi(0)||\dot J_t(0)|,$$

where we again used that γ and ϕ are geodesics. It follows that we are done once we can bound $|\dot J_t(0)|$ and $|\dot J_t(1)|$. For this, we first obtain a more specific expression for the Jacobi field J_t. To this end, we define for every $0 \le t \le t_0$ the vector fields

$$J^1_t(s) = \mathrm{d}(\mathrm{Exp}_{\phi(t)})_{s\partial_s\Gamma(t,0)}(s\dot J^1_t(0))$$

and

$$J^2_t(s) = \mathrm{d}(\mathrm{Exp}_{\gamma(t)})_{-s\partial_s\Gamma(t,1)}(s\dot J^2_t(0)),$$

where

$$\dot J^1_t(0) = \mathrm{d}(\mathrm{Exp}_{\phi(t)})^{-1}_{\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)}\dot\gamma(t) \in T_{\phi(t)}M$$

and likewise

$$\dot J^2_t(0) = \mathrm{d}(\mathrm{Exp}_{\gamma(t)})^{-1}_{\mathrm{Exp}^{-1}_{\gamma(t)}\phi(t)}\dot\phi(t) \in T_{\gamma(t)}M.$$

As explained in Section 2.2.3, J^1_t and J^2_t are Jacobi fields along Γ_t. Note that $J^1_t(0) = J^2_t(0) = 0$ and $J^1_t(1) = \dot\gamma(t)$ and $J^2_t(1) = \dot\phi(t)$. Because J_t is the unique Jacobi field along Γ_t with $J_t(0) = \dot\phi(t)$ and $J_t(1) = \dot\gamma(t)$, it follows that

$$J_t(s) = J^1_t(s) + J^2_t(1-s).$$

Using the above decomposition, we show how to bound $|\dot J_t(0)|$. The bound for $|\dot J_t(1)|$ may be obtained similarly. By the triangle inequality, we have

$$|\dot J_t(0)| \le |\dot J^1_t(0)| + |\dot J^2_t(1)|.$$

Note that

$$|\dot J^1_t(0)| = |\mathrm{d}(\mathrm{Exp}_{\phi(t)})^{-1}_{\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)}\dot\gamma(t)| \le |\mathrm{d}(\mathrm{Exp}_{\phi(t)})^{-1}_{\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)}||\dot\gamma(t)|$$

Consequently, by Lemma 5.2 there exists a constant $C > 0$ only depending on K and r (since $|\mathrm{Exp}^{-1}_{\phi(t)}\gamma(t)| = d(\phi(t),\gamma(t)) \le r$) such that

$$|\dot J^1_t(0)| \le C|\dot\gamma(t)| = C|\dot\gamma(0)|.$$

For the other term, it follows from Proposition 2.4 that

$$\begin{aligned}|\dot J^2_t(1)| &\le |\dot J^2_t(0)| + \sup_{s\in[0,1]}|R_{\Gamma(t,s)}(J^2_t(s), \partial_s\Gamma(t,s))\partial_s\Gamma(t,s)|\\ &\le C|\dot\phi(0)| + |\partial_s\Gamma(t,0)|^2\sup_{s\in[0,1]}|R_{\psi_t(s)}||J^2_t(s)|\\ &\le C|\dot\phi(0)| + \tilde C d(\gamma(t),\phi(t))^2\sup_{s\in[0,1]}|J^2_t(s)|\\ &\le C|\dot\phi(0)| + \tilde C r^2\sup_{s\in[0,1]}|J^2_t(s)|.\end{aligned}$$

Here we used in the second line again Lemma 5.2 as above, together with the fact that the curves $\Gamma_t(s)$ are geodesics. Furthemore, we used that the curvature is continuous, and hence bounded on compact sets, so that $\tilde{C}$ only depends on $\tilde{K}$, since the variation Γ is contained in $\tilde{K}$. In the last line, we used that $d(\gamma(t), \phi(t)) \leq r$ for all $0 \leq t \leq t_0$ by choice of t_0.

Finally, we have for any $s \in [0,1]$

$$\begin{aligned} |J_t^2(s)| &= |\mathrm{d}(\mathrm{Exp}_{\gamma(t)})_{-s\partial_s\Gamma(t,1)}(s\dot{J}_t^2(0))| \\ &\leq s|\mathrm{d}(\mathrm{Exp}_{\gamma(t)})_{-s\partial_s\Gamma(t,1)}||\dot{J}_t^2(0))| \\ &\leq C'|\dot{\phi}(0)|, \end{aligned}$$

where in the last line we used Lemma 5.1. Collecting everything, there exists a constant $C > 0$, only depending on $\tilde{K}$ and r, such that

$$|\dot{J}_t^2(1)| \leq C|\dot{\phi}(0)|.$$

Putting everything together, we find that

$$|\dot{J}_t(0)| \leq |\dot{J}_t^1(0)| + |\dot{J}_t^2(1)| \leq C(|\dot{\gamma}(0)| + |\dot{\phi}(0)|)$$

for some $C > 0$ only depending on $\tilde{K}$ and r. Obtaining a similar bound for $|\dot{J}_t(1)|$ now proves the claim. □

6 Proof of Cramér's theorem for geodesic random walks

In this section we provide a proof of Cramér's theorem for geodesic random walks with independent and identically distributed increments, which are bounded and have expectation 0. The proof relies on an analysis of the geometric properties of a geodesic random walk. To prove the theorem, we follow the steps as discussed in Section 4. We provide the details and show how we use the geometric results from Section 5. For completeness, let us recall the statement of the theorem.

Theorem 6.1 (Cramér's theorem for Riemannian manifolds)**.** *Let (M, g) be a complete Riemannian manifold. Fix $x_0 \in M$ and let $\{\mu_x\}_{x\in M}$ be a collection of measures such that $\mu_x \in \mathcal{P}(T_xM)$ for all $x \in M$. For every $n \geq 1$, let $\{(\frac{1}{n} * \mathcal{S})_j\}_{j\geq 0}$ be a $\frac{1}{n}$-rescaled geodesic random walk started at x_0 with independent increments $\{X_j^n\}_{j\geq 1}$, compatible with $\{\mu_x\}_{x\in M}$. Let $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$ be the associated empirical average process started at x_0. Assume the increments are bounded and have expectation 0. Assume furthermore that the collection $\{\mu_x\}_{x\in M}$ satisfies the consistency property in Definition 3.7. Then $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$ satisfies in M the LDP with good rate function*

$$I_M(x) = \inf\{\Lambda_{x_0}^*(v) | v \in \mathrm{Exp}_{x_0}^{-1} x\} \tag{6.1}$$

In Section 6.1 we prove the upper bound of the large deviation principle for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 1}$ in M, while in Section 6.2 we prove the lower bound. More specifically, Theorem 6.1 follows immediately from Proposition 6.9 together with Proposition 6.11.

However, before we can prove the upper and lower bound of the large deviation principle for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 1}$, we first need some general results and estimates. From here on, we fix $r > 0$ to be the uniform bound on the increments of the random walk. By the triangle inequality, we find

$$d\left(\left(\frac{1}{n} * \mathcal{S}\right)_k, x_0\right) \leq \frac{1}{n}\sum_{l=1}^{k} |X_k^n| \leq \frac{k}{n} r \leq r$$

for all $0 \leq k \leq n$. Consequently, for every $n \geq 0$ and $1 \leq k \leq n$ we have

$$\left(\frac{1}{n} * \mathcal{S}\right)_k \in \overline{B(x_0, r)} =: K.$$

By completeness of M, K is compact since it is closed and bounded.

Now consider the process Z_n in $T_{x_0}M$ given by

$$Z_n = \frac{1}{n} \sum_{k=1}^{n} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n.$$

Here, the parallel transport $\tau_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}$ is considered along the piecewise geodesic path traced out by the geodesic random walk. From Cramér's theorem for vector spaces it follows that $\{Z_n\}_{n\geq 0}$ satisfies the large deviation principle in $T_{x_0}M$, which we will show in the following proposition.

Proposition 6.2. *Let the assumptions of Theorem 6.1 be satisfied. For every $n \geq 0$, define $Z_n = \frac{1}{n}\sum_{k=1}^{n} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n \in T_{x_0}M$. Let $\Lambda_{x_0}(\lambda) = \log \mathbb{E}(e^{\lambda X_1})$ be the log moment generating function of the increments. Then $\{Z_n\}_{n\geq 0}$ satisfies the large deviation principle in $T_{x_0}M$ with good rate function*

$$I(v) = \Lambda^*_{x_0}(v) := \sup_{\lambda \in T_{x_0}M} \{\langle \lambda, v\rangle - \Lambda_{x_0}(\lambda)\}.$$

Proof. Define $Y_k^n = \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n \in T_{x_0}M$. We compute for any $\lambda \in T_{x_0}M$

$$\begin{aligned}
\mathbb{E}(e^{\langle \lambda, Y_k^n\rangle}) &= \mathbb{E}\left(\mathbb{E}\left(e^{\langle \lambda, \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n\rangle} | \mathcal{F}_{k-1}\right)\right) \\
&= \mathbb{E}\left(\int_{T_{(\frac{1}{n}*\mathcal{S})_{k-1}}M} e^{\langle \lambda, \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} v\rangle} \mu_{(\frac{1}{n}*\mathcal{S})_{k-1}}(\mathrm{d}v)\right) \\
&= \mathbb{E}\left(\int_{T_{x_0}M} e^{\langle \lambda, v\rangle} \mu_{x_0}(\mathrm{d}v)\right) \\
&= \int_{T_{x_0}M} e^{\langle \lambda, v\rangle} \mu_{x_0}(\mathrm{d}v).
\end{aligned}$$

Here we used in the second line that $\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}$ is measurable with respect to $\mathcal{F}_{k-1}$, together with the fact that the increments are independent (see Definition 3.3). In the third line we applied Proposition 3.8, using that the increments are identically distributed. It follows that Y_k^n is distributed according to μ_{x_0}.

Consequently, the result follows from Cramér's theorem once we show that Y_k^n and Y_l^n are independent whenever $k \neq l$. To this end, assume without loss of generality that $l < k$. Then for measurable sets $A, B \subset T_{x_0}M$ we find in a similar way as above that

$$\begin{aligned}
&\mathbb{P}(Y_l^n \in A, Y_k^n \in B) \\
&= \mathbb{E}(I(Y_l^n \in A)\mathbb{E}(I(Y_k^n \in B)|\mathcal{F}_{k-1})) \\
&= \mathbb{E}\left(I(Y_l^n \in A)\int_{T_{(\frac{1}{n}*\mathcal{S})_{k-1}}M} I\left(\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} v \in B\right) \mu_{(\frac{1}{n}*\mathcal{S})_{k-1}}(\mathrm{d}v)\right) \\
&= \mathbb{E}\left(I(Y_l^n \in A)\int_{T_{x_0}M} I\left(v \in B\right) \mu_{x_0}(\mathrm{d}v)\right) \\
&= \mathbb{E}(I(Y_l^n \in A))\mathbb{E}(I(Y_k^n \in B)) \\
&= \mathbb{P}(Y_l^n \in A)\mathbb{P}(Y_k^n \in B),
\end{aligned}$$

where I denotes the indicator function. Above, we used in the one but last line that Y_k^n is distributed according to μ_{x_0}. We conclude that the Y_l^n and Y_k^n are independent. □

Remark 6.3. Note that in the proof of Proposition 6.2 we did not use along which path we performed the parallel transport $\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}$, only that it was measurable with respect to $\mathcal{F}_{k-1}$. Consequently, the result holds for any choice of parallel transport, as long as it is measurable with respect to $\mathcal{F}_{k-1}$.

Proposition 6.2 suggests we should try to map the sequence $\{(\frac{1}{n}*\mathcal{S})_n\}_{n\geq 0}$ from M to $T_{x_0}M$ in such a way that it will be close to the sequence $\{Z_n\}_{n\geq 0}$.

To this end, recall that if we assume that $r < \iota(x_0)$, then for all n and all $0 \leq k \leq n$ we can uniquely define

$$v_k^n \in \operatorname{Exp}_{x_0}^{-1}\left(\left(\frac{1}{n}*\mathcal{S}\right)_k\right) \subset T_{x_0}M$$

with $|v_k^n| < \iota(x_0)$, because $d((\frac{1}{n}*\mathcal{S})_k, x_0) \leq r < \iota(x_0)$.

As explained in Step 2 of Section 4.2, we have the following estimate. The first term of the upper bound in (6.2) follows from replacing v_n^l with a sum of differentials of the Riemannian exponential map, while the second term follows from replacing these differentials with parallel transport.

Proposition 6.4. *Let the assumptions of Theorem 6.1 be satisfied. Additionally, let r be the uniform bound of the increments and assume that $r < \iota(x_0)$. Then there exists a constant $C > 0$ such that*

$$\left| v_l^n - \frac{1}{n}\sum_{k=1}^{l} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n \right| \leq C\frac{l}{n^2} + Cr^2\frac{l^3}{n^3} \tag{6.2}$$

for all n and all $1 \leq l \leq n$.

Proof. Recall that for all n and all $0 \leq k \leq n$ we have that $(\frac{1}{n}*\mathcal{S})_k$ is in the compact set $K = \overline{B(x_0, r)}$. Consequently,

$$v_k^n \in \overline{B(0,r)} \subset T_{x_0}M$$

for all n and all $0 \leq k \leq n$. But then it follows from Proposition 5.4 that for every $0 \leq k \leq n$ there exists a constant $C_k > 0$ only depending on the norms of v_k^n, v_{k+1}^n and X_k^n such that

$$\left| v_{k+1}^n - \left(v_k^n + \frac{1}{n}\mathrm{d}(\operatorname{Exp}_{x_0})^{-1}_{v_k^n} X_{k+1}^n \right) \right| \leq C_k \frac{1}{n^2}. \tag{6.3}$$

Because each of the norms $|v_k^n|, |v_{k+1}^n|$ and $|X_k^n|$ are bounded by r, we conclude that we can take $C_k = C$ independent of k.

Turning to the proof of the statement, by the triangle inequality we have

$$\begin{aligned}
&\left| v_l^n - \frac{1}{n}\sum_{k=1}^{l} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n \right| \\
&\leq \left| v_l^n - \frac{1}{n}\sum_{k=1}^{l} \mathrm{d}(\operatorname{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n \right| + \frac{1}{n}\sum_{k=1}^{l} \left| \mathrm{d}(\operatorname{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n - \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n \right|.
\end{aligned}$$

We estimate both terms separately.

For the first term, we write v_l^n as the telescoping sum

$$v_l^n = \sum_{k=1}^{l} (v_k^n - v_{k-1}^n).$$

Consequently, we obtain

$$\begin{aligned}\left| v_l^n - \frac{1}{n}\sum_{k=1}^{l} \mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n \right| &\leq \sum_{k=1}^{l} |v_k^n - v_{k-1}^n - \mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n| \\ &\leq C\frac{l}{n^2},\end{aligned} \tag{6.4}$$

where the last line follows from the estimate in (6.3).

For the other term, observe that by Corollary 5.8, there exists a constant $C > 0$ only depending on the compact set $\overline{B(0,r)}$ and r, such that

$$|\mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n - \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n| \leq C|v_{k-1}^n|^2 \tag{6.5}$$

But then we find

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{l} |\mathrm{d}(\mathrm{Exp}_{x_0})^{-1}_{v_{k-1}^n} X_k^n - \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n| &\leq C\frac{1}{n}\sum_{k=1}^{l} |v_{k-1}^n|^2 \\ &\leq Cr^2\frac{l^3}{n^3},\end{aligned}$$

where in the last line we used that $|v_{k-1}^n| \leq r\frac{k-1}{n} \leq r\frac{l}{n}$ for any $1 \leq k \leq l$. □

Remark 6.5. The estimate in Proposition 6.4 is one of the most important ingredients of the proof of Theorem 6.1. Indeed, it allows us in some sense to connect large deviations for $\{(\frac{1}{n}*\mathcal{S})_n\}_{n\geq 0}$ in M to large deviations for the sums $\{\frac{1}{n}\sum_{k=1}^{n} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n\}_{n\geq 0}$ in the tangent space $T_{x_0}M$. Consequently, by making appropriate assumptions on the sequence $\{\frac{1}{n}\sum_{k=1}^{n} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n\}_{n\geq 0}$, for example in the spirit of the Gärtner-Ellis theorem (see e.g. [5, 1]), we can obtain more general results than Cramér's theorem for geodesic random walks in a similar way.

One might hope to combine Propositions 6.2 and 6.4 to prove that $\{v_n^n\}_{n\geq 0}$ satisfies in $T_{x_0}M$ the large deviation principle. Unfortunately, the upper bound found in Proposition 6.4 gives an unwanted contribution on the exponential scale. Indeed, taking $l = n$, we find that the upper bound in (6.2) is $\mathcal{O}(1)$, which results in the fact that we get stuck with a constant as explained in Step 1 of Section 4.2. In an attempt to reduce this term in the upper bound, we cut up the random walk in finitely many pieces and analyse the pieces separately.

To this end, recall that

$$d\left(\left(\frac{1}{n}*\mathcal{S}\right)_k, x_0\right) \leq \frac{1}{n}\sum_{l=1}^{k} |X_k^n| \leq \frac{k}{n}r.$$

Now observe that $\iota(\overline{B(x_0,r)}) > 0$, because $\overline{B(x_0,r)}$ is compact (see (2.2) for the definition of the injectivity radius of a set). Consequently, if $k \leq \frac{n\iota(\overline{B(x_0,r)})}{2r}$, then

$$d\left(\left(\frac{1}{n}*\mathcal{S}\right)_k, x_0\right) \leq \frac{\iota(\overline{B(x_0,r)})}{2} < \iota(\overline{B(x_0,r)}). \tag{6.6}$$

Now let $m \in \mathbb{N}$ such that $m \geq \frac{2r}{\iota(\overline{B(x_0,r)})}$. For $0 \leq l \leq m-1$ we define $n_l = l\lfloor m^{-1}n\rfloor$ and $n_m = n$. By (6.6), for every $0 \leq l \leq m-1$ and $1 \leq k \leq n_{l+1} - n_l$ we can uniquely define

$$\tilde{v}_k^{n,m,l} \in \mathrm{Exp}^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_l}}\left(\left(\frac{1}{n}*\mathcal{S}\right)_{n_l+k}\right) \subset T_{(\frac{1}{n}*\mathcal{S})_{n_l}}M \tag{6.7}$$

with $|\tilde{v}_k^{n,m,l}| < \iota((\frac{1}{n} * \mathcal{S})_{n_l})$, because $n_{l+1} - n_l \leq nm^{-1} \leq \frac{n\iota(\overline{B(x_0,r)})}{2r}$. Finally, we set

$$v_k^{n,m,l} = \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{n_l}} \tilde{v}_k^{n,m,l} \in T_{x_0}M,$$

where parallel transport $\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{n_l}}$ is taken along the piecewise geodesic path through the points $(\frac{1}{n} * \mathcal{S})_{n_1}, \ldots, (\frac{1}{n} * \mathcal{S})_{n_{l-1}}$.

Alongside this division of the random walk into pieces, we define a map $\Psi_m : (T_{x_0}M)^m \to M$ to identify the tuple $(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \ldots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor})$ with $(\frac{1}{n} * \mathcal{S})_n$, just like we used the Riemannian exponential map to identify v_n^n and $(\frac{1}{n} * \mathcal{S})_n$ before. Essentially, Ψ_m is an m time recursive application of the Riemannian exponential map.

More precisely, let $(v_1, \ldots, v_m) \in (T_{x_0}M)^m$ be given and define $x_1 = \mathrm{Exp}_{x_0}(v_1)$. Now, suppose $x_1, \ldots, x_i$ are given. Denote by $\tau_{x_0x_i}$ parallel transport along the constructed piecewise geodesic path via $x_1, \ldots, x_{i-1}$. Then we define $\tilde{v}_{i+1} = \tau_{x_0x_i} v_{i+1}$ and set $x_{i+1} = \mathrm{Exp}_{x_i}(\tilde{v}_{i+1})$. Finally, we define $\Psi_m(v_1, \ldots, v_m) = x_m$. In particular, we have for every $x \in M$ and $v \in \mathrm{Exp}_{x_0}^{-1}x$ that $(\frac{1}{m}v, \ldots, \frac{1}{m}v) \in \Psi_m^{-1}x$. To see this, observe that the path that Ψ_m constructs is precisely the geodesic $\gamma_v(t) = \mathrm{Exp}_{x_0}(tv)$, because the speed of a geodesic is parallel along the geodesic. Furthermore, the map Ψ_m is continuous as a composition of continuous maps.

Remark 6.6. Strictly speaking, if we divide the random walk into m pieces as above, for the last piece we can only guarantee that it has at most $\lfloor m^{-1}n \rfloor + m$ increments, since n need not be divisible by m. Additionally, this implies that $\Psi_m(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \ldots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor})$ is only equal to $\left(\frac{1}{n} * \mathcal{S}\right)_n$ when n is divisible by m. However, for every $m \in \mathbb{N}$ it holds that

$$d\left(\Psi_m(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \ldots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor}), \left(\frac{1}{n} * \mathcal{S}\right)_n\right) = \mathcal{O}\left(\frac{1}{n}\right).$$

Since in the proofs to follow we always first let n tend to infinity before m, this has no influence on the results and arguments. Therefore, to avoid unnecessarily complicated notation and reasoning, we proceed with the above.

6.1 Upper bound of the large deviation principle for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$

In this section we prove the large deviation upper bound for $\{(\frac{1}{n} * \mathcal{S})_n\}_{n\geq 0}$. Before we can do this, we first need some preliminary results.

Proposition 6.7 (Upper bound for $\mathbb{E}(e^{n\langle \lambda, v_n^n\rangle})$)**.** *Let the assumptions of Theorem 6.1 be satisfied. Additionally, let r be the uniform bound of the increments and assume that $r < \iota(x_0)$. Then there exists a constanct $C > 0$ such that for all n and all $1 \leq l \leq n$*

$$\mathbb{E}(e^{n\langle \lambda, v_l^n\rangle}) \leq e^{ln^{-1}|\lambda|C} e^{|\lambda|Cr^2l^3n^{-2}} M_{x_0}(\lambda)^l$$

for all $\lambda \in T_{x_0}M$. Here, $M_{x_0}(\lambda) = \int_{T_{x_0}M} e^{\langle \lambda, v\rangle} \mu_{x_0}(\mathrm{d}v)$.

Proof. By Proposition 6.4 and the Cauchy-Schwarz inequality, there exists a constant $C > 0$ such that

$$\begin{aligned}\langle \lambda, v_l^n\rangle - \frac{1}{n}\sum_{k=1}^{l} \langle \lambda, \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n\rangle &\leq |\lambda| \left| v_l^n - \frac{1}{n}\sum_{k=1}^{l} \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}} X_k^n \right| \\ &\leq C|\lambda|\frac{l}{n^2} + C|\lambda|r^2\frac{l^3}{n^3}.\end{aligned}$$

But then we can estimate

$$\begin{aligned}\mathbb{E}\left(e^{n\langle\lambda,v_l^n\rangle}\right) &= \mathbb{E}\left(e^{\sum_{k=1}^{l}\langle\lambda,\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n\rangle}e^{n\langle\lambda,v_l^n\rangle-\sum_{k=1}^{l}\langle\lambda,\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n\rangle}\right)\\ &\leq e^{C|\lambda|ln^{-1}}e^{C|\lambda|r^2l^3n^{-2}}\mathbb{E}\left(e^{\sum_{k=1}^{l}\langle\lambda,\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n\rangle}\right).\end{aligned}$$

As shown in the proof of Proposition 6.2, for every $1 \leq k \leq n$ we have that $\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n$ is distributed accodring to μ_{x_0} and is independent of $\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{l-1}}X_l^n$ for any $l \neq k$. Consequentely, we find that

$$\begin{aligned}\mathbb{E}\left(e^{\sum_{k=1}^{l}\langle\lambda,\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n\rangle}\right) &= \prod_{k=1}^{l}\mathbb{E}\left(e^{\langle\lambda,\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n\rangle}\right)\\ &= M_{x_0}(\lambda)^l,\end{aligned}$$

where the last step follows from Proposition 3.8. □

Using Proposition 6.7, we obtain the following inequality, which is key in deriving the large deviations upper bound for $\{(\frac{1}{n}*\mathcal{S})_n\}_{n\geq 0}$.

Proposition 6.8. *Let the assumptions of Theorem 6.1 be satisfied. Denote by r the uniform bound on the increments of the geodesic random walk. Then for any $m \in \mathbb{N}$ such that $m \geq \frac{2r}{\iota(\overline{B(x_0,r)})}$ and any closed $F \subset (T_{x_0}M)^m$ we have*

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n\rfloor},\dots,v^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right)\in F\right)\\ &\leq -\inf_{(v_1,\dots,v_m)\in F}\sup_{(\lambda_1,\dots,\lambda_m)\in(T_{x_0}M)^m}\frac{1}{m}\sum_{i=1}^{m}\left\{\langle\lambda_i,mv_i\rangle-\Lambda_{x_0}(\lambda_i)-m^{-2}C|\lambda_i|r^2\right\}.\end{aligned}$$

Here, C is a constant depending on the curvature of the compact set $\overline{B(0,r)}$ and the bound r.

Proof. We first prove the upper bound for compact sets, so let $\Gamma \subset (T_{x_0}M)^m$ be compact. Following the proof of Cramér's theorem (see e.g. [1, 5]) for the vector space $(T_{x_0}M)^m$, we have

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n\rfloor},\dots,v^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right)\in\Gamma\right)\\ &\leq -\inf_{(v_1,\dots,v_m)\in\Gamma}\sup_{(\lambda_1,\dots,\lambda_m)\in(T_{x_0}M)^m}\left\{\sum_{i=1}^{m}\langle\lambda_i,v_i\rangle-\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left(e^{n\sum_{i=1}^{m}\langle\lambda_i,v^{n,m,i}_{\lfloor m^{-1}n\rfloor}\rangle}\right)\right\}.\end{aligned}$$

Recall that for $0 \leq i \leq m-1$ we write $n_i = i\lfloor m^{-1}n\rfloor$ and $n_m = n$. By Proposition 6.4 (which we may apply, because m is chosen large enough) there exists a $C > 0$ such that for any $1 \leq i \leq m$ we have

$$\begin{aligned}\left|\tilde{v}^{n,m,i}_{\lfloor m^{-1}n\rfloor}-\frac{1}{n}\sum_{k=n_{i-1}+1}^{n_i}\tau^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_{i-1}}(\frac{1}{n}*\mathcal{S})_k}X_k^n\right| &\leq C\frac{\lfloor m^{-1}n\rfloor}{n^2}+Cr^2\frac{\lfloor m^{-1}n\rfloor^3}{n^3}\\ &\leq C\frac{1}{nm}+Cr^2\frac{1}{m^3}.\end{aligned}$$

But then we also have that

$$\left|v^{n,m,i}_{\lfloor m^{-1}n\rfloor}-\frac{1}{n}\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{n_{i-1}}}\sum_{k=n_{i-1}+1}^{n_i}\tau^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_{i-1}}(\frac{1}{n}*\mathcal{S})_k}X_k^n\right| \leq C\frac{1}{nm}+Cr^2\frac{1}{m^3}, \tag{6.8}$$

because parallel transport is an isometry.

Now define

$$Y_i^n = \tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{n_{i-1}}} \sum_{k=n_{i-1}+1}^{n_i} \tau^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_{i-1}}(\frac{1}{n}*\mathcal{S})_k} X_k^n \in T_{x_0}M.$$

Using (6.8), it follows from the Cauchy-Schwarz inequality and the triangle inequality that

$$\left| \sum_{i=1}^m \langle \lambda_i, v^{n,m,i}_{\lfloor m^{-1}n \rfloor} \rangle - \frac{1}{n} \sum_{i=1}^m \langle \lambda_i, Y_i^n \rangle \right| \leq C \left(\frac{1}{nm} + r^2 \frac{1}{m^3} \right) \sum_{i=1}^m |\lambda_i|.$$

Consequently, we find that

$$\mathbb{E}\left(e^{n \sum_{i=1}^m \langle \lambda_i, v^{n,m,i}_{\lfloor m^{-1}n \rfloor} \rangle} \right) \leq e^{Cm^{-1} \sum_{i=1}^m |\lambda_i|} e^{Cr^2 m^{-3} n \sum_{i=1}^m |\lambda_i|} \mathbb{E}\left(e^{\sum_{i=1}^m \langle \lambda_i, Y_i^n \rangle} \right).$$

Now note that, like in the proof of Proposition 6.2, we can show that for $i \neq j$ the random variables Y_i^n and Y_j^n are independent. Consequently, we have that

$$\mathbb{E}\left(e^{\sum_{i=1}^m \langle \lambda_i, Y_i^n \rangle} \right) = \prod_{i=1}^m \mathbb{E}\left(e^{\langle \lambda_i, Y_i^n \rangle} \right).$$

Moreover, again following the proof of Proposition 6.2, one can show that

$$\mathbb{E}\left(e^{\langle \lambda_i, Y_i^n \rangle} \right) = M_{x_0}(\lambda_i)^{\lfloor m^{-1}n \rfloor}.$$

Combining everything, we find that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(e^{n \sum_{i=1}^m \langle \lambda_i, v^{n,m,i}_{\lfloor m^{-1}n \rfloor} \rangle} \right) \\
&\leq \limsup_{n\to\infty} \left\{ \frac{C}{mn} \sum_{i=1}^m |\lambda_i| + \frac{Cr^2}{m^3} \sum_{i=1}^m |\lambda_i| + \frac{\lfloor m^{-1}n \rfloor}{n} \sum_{i=1}^m \Lambda_{x_0}(\lambda_i) \right\} \\
&= \frac{Cr^2}{m^3} \sum_{i=1}^m |\lambda_i| + \frac{1}{m} \sum_{i=1}^m \Lambda_{x_0}(\lambda_i).
\end{aligned}$$

Putting everything together, we obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \ldots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor} \right) \in \Gamma \right) \\
&\leq - \inf_{(v_1,\ldots,v_m)\in\Gamma} \sup_{(\lambda_1,\ldots,\lambda_m)\in(T_{x_0}M)^m} \sum_{i=1}^m \left\{ \langle \lambda_i, v_i \rangle - m^{-1} \Lambda_{x_0}(\lambda_i) - m^{-3} Cr^2 |\lambda_i| \right\}.
\end{aligned}$$

This concludes the proof of the upper bound for compact sets.

To extend the upper bound to all closed sets, one should simply notice that

$$\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \ldots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor} \right) \in \overline{B(0,r)}^m$$

almost surely, where r is the uniform bound of the increments. Since M is complete, $\overline{B(0,r)}^m$ is compact, so that the sequence is exponentially tight. □

It now remains to transfer the upper bound in Proposition 6.8 for the process in $(T_{x_0}M)^m$ related to $\left\{ \left(\frac{1}{n} * \mathcal{S} \right)_n \right\}_{n\geq 0}$ to the upper bound of the large deviation principle for $\left\{ \left(\frac{1}{n} * \mathcal{S} \right)_n \right\}_{n\geq 0}$. With all preperations done, the only thing that remains to be shown, is that the upper bound has the desired form.

Proposition 6.9. *Let the assumptions of Theorem 6.1 be satisfied. Then for any $F \subset M$ closed we have*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in F\right) \leq -\inf_{x\in F} I_M(x),$$

where

$$I_M(x) = \inf\{\Lambda^*_{x_0}(v) | v \in \operatorname{Exp}^{-1}_{x_0} x\}.$$

Proof. Let $F \subset M$ be closed and pick $m \in \mathbb{N}$ such that $m \geq \frac{2r}{\iota(\overline{B(x_0,r)})}$, where r denotes the uniform bound of the increments. Let $\Psi_m : (T_{x_0}M)^m \to M$ be the recursive application of the Riemannian exponential map defined just above Section 6.1. Because Ψ_m is continuous, we have that $\Psi_m^{-1}F \subset (T_{x_0}M)^m$ is closed. Hence, by Proposition 6.8 we find that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in F\right) \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n\rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right) \in \Psi_m^{-1}F\right) \\
&\leq -\inf_{(v_1,\dots,v_m)\in\Psi_m^{-1}F} \sup_{(\lambda_1,\dots,\lambda_m)\in(T_{x_0}M)^m} \frac{1}{m}\sum_{i=1}^m \left\{\langle\lambda_i, mv_i\rangle - \Lambda_{x_0}(\lambda_i) - m^{-2}Cr^2|\lambda_i|\right\}.
\end{aligned}$$

Now observe that for every $\lambda \in T_{x_0}M$ we have $|\lambda| \leq |\lambda|^2 + 1$. Plugging this into the above estimate, keeping in mind the minus sign in front, we find that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in F\right) \\
&\leq \frac{Cr^2}{m^2} - \inf_{(v_1,\dots,v_m)\in\Psi_m^{-1}F} \sup_{(\lambda_1,\dots,\lambda_m)\in(T_{x_0}M)^m} \frac{1}{m}\sum_{i=1}^m \left\{\langle\lambda_i, mv_i\rangle - \Lambda_{x_0}(\lambda_i) - m^{-2}Cr^2|\lambda_i|^2\right\}.
\end{aligned}$$

We now focus on the infimum in the above expression. The necessity of replacing $|\lambda|$ with $|\lambda|^2$, and making the upper bound slightly worse, will become clear when we try to calculate this infimum further.

First, consider the map $\Lambda_m : T_{x_0}M \to \mathbb{R}$ defined by

$$\Lambda_m(\lambda) = \Lambda_{x_0}(\lambda) + \frac{1}{m^2}Cr^2|\lambda|^2,$$

and denote by Λ^*_m its Legendre transform. Then

$$\sup_{(\lambda_1,\dots,\lambda_m)\in(T_{x_0}M)^m} \frac{1}{m}\sum_{i=1}^m \left\{\langle\lambda_i, mv_i\rangle - \Lambda_{x_0}(\lambda_i) - m^{-2}Cr^2|\lambda_i|\right\} = \frac{1}{m}\sum_{i=1}^m \Lambda^*_m(mv_i).$$

The latter may be interpreted as

$$\int_0^1 \Lambda^*_m(\dot\gamma_m(t))\,\mathrm{d}t,$$

where γ_m is piecewise geodesic on intervals of the form $[\frac{(i-1)}{m}, \frac{i}{m}]$ with speed $m\tilde v_i$, where $\tilde v_i = \tau_{x_0\gamma_m(\frac{(i-1)}{m})}v_i$.

Now note that since Λ_{x_0} is differentiable and convex, we find that Λ_m is differentiable and strictly convex. Furthermore, we have for every $u \in T_{x_0}M$ that

$$\Lambda^*_m(u) = \sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,u\rangle - \Lambda_{x_0}(\lambda) - \frac{1}{m^2}Cr^2|\lambda|^2\right\} \leq \sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,u\rangle - \frac{1}{m^2}Cr^2|\lambda|^2\right\} < \infty.$$

Here we used that Λ_{x_0} is non-negative, because the expectation of μ_{x_0} is 0. Consequently, we find that Λ_m^* is everywhere finite. Note that this does not contradict the fact that the rate function might be infinite, since Λ_m^* merely provides a lower bound of the rate function. Because Λ_m^* is everywhere finite, it follows from Lemma A.1 that Λ_m^* is strictly convex and differentiable.

The above shows that we can apply [8, Proposition 8.3], giving us that minimizing trajectories for the functional

$$\int_0^1 \Lambda_m^*(\dot\gamma(t))\,\mathrm{d}t$$

are geodesics. Because for every $x \in F$ and every $v \in \operatorname{Exp}_{x_0}^{-1} x$ we have that $(\frac{1}{m}v, \dots, \frac{1}{m}v) \in \Psi_m^{-1}F$, we find that

$$\begin{aligned}
&- \inf_{(v_1,\dots,v_m)\in\Psi_m^{-1}F} \sup_{(\lambda_1,\dots,\lambda_m)\in(T_{x_0}M)^m} \frac{1}{m}\sum_{i=1}^m \left\{\langle\lambda_i, mv_i\rangle - \Lambda_{x_0}(\lambda_i)\right\} - m^{-2}Cr^2|\lambda_i|^2\Big\} \\
&= - \inf_{v\in\operatorname{Exp}_{x_0}^{-1}F} \sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda) - m^{-2}Cr^2|\lambda|^2\right\}.
\end{aligned}$$

Now note that

$$\begin{aligned}
&\lim_{m\to\infty} \sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda) - m^{-2}Cr^2|\lambda|^2\right\} \\
&= \sup_{\lambda\in T_{x_0}M} \lim_{m\to\infty} \left\{\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda) - m^{-2}Cr^2|\lambda|^2\right\} \qquad (6.9)\\
&= \sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda)\right\},
\end{aligned}$$

because $\Lambda_m(\lambda) = \langle\lambda, v\rangle - \Lambda_{x_0}(\lambda) - m^{-2}Cr^2|\lambda|^2$ is increasing in m for every $\lambda \in T_{x_0}M$. Furthermore, we have

$$\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda) - m^{-2}Cr^2|\lambda| \geq \langle\lambda, v\rangle - r|\lambda| - m^{-2}Cr^2|\lambda|^2,$$

because the support of μ_{x_0} is contained in $\overline{B(0,r)}$. Furthermore, one may compute that if $|v| > r$, then

$$\sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - r|\lambda| - m^{-2}Cr^2|\lambda|^2\right\} = \frac{m^2}{4Cr^2}(|v| - r)^2. \qquad (6.10)$$

Now write

$$\operatorname{Exp}_{x_0}^{-1}F = \left(\operatorname{Exp}_{x_0}^{-1}F \cap \overline{B(0,2r)}\right) \cup \left(\operatorname{Exp}_{x_0}^{-1}F \cap \overline{B(0,2r)}^C\right).$$

Note that by (6.10), we find that

$$\begin{aligned}
&\lim_{m\to\infty} \inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}^C} \sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda) - m^{-2}Cr^2|\lambda|^2\right\} \\
&\geq \lim_{m\to\infty} \inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}^C} \sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - r|\lambda| - m^{-2}Cr^2|\lambda|^2\right\} \\
&\geq \lim_{m\to\infty} \frac{m^2}{4Cr^2}r^2 \\
&= \infty,
\end{aligned}$$

where we used in the one but last line that $|v| \geq 2r$. Also, because $|v| \geq 2r \geq r$, we have

$$\sup_{\lambda\in T_{x_0}M} \left\{\langle\lambda, v\rangle - \Lambda_{x_0}(\lambda)\right\} = \infty,$$

so that

$$\begin{aligned}&\lim_{m\to\infty}\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}^C}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)-m^{-2}Cr^2|\lambda|^2\right\}\\&=\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}^C}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)\right\}.\end{aligned}$$

For the other part, because $\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}$ is compact, it follows from (6.9) that

$$\begin{aligned}&\lim_{m\to\infty}\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)-m^{-2}Cr^2|\lambda|^2\right\}\\&=\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)\right\}.\end{aligned}$$

Collecting everything, we find that

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(\frac{1}{n}*\mathcal{S}\right)_n\in F\right)\\&\le\lim_{m\to\infty}\left(\frac{Cr^2}{m^2}-\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)-m^{-2}Cr^2|\lambda|^2\right\}\right)\\&=-\lim_{m\to\infty}\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F\cap\overline{B(0,2r)}}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)-m^{-2}Cr^2|\lambda|^2\right\}\\&=-\inf_{v\in\operatorname{Exp}_{x_0}^{-1}F}\sup_{\lambda\in T_{x_0}M}\left\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)\right\}\\&=-\inf_{x\in F}I_M(x),\end{aligned}$$

which concludes the proof. □

6.2 Lower bound of the large deviation principle for $\{(\frac{1}{n}*\mathcal{S})_n\}_{n\ge0}$

In this section we prove the large deviation lower bound for $\{(\frac{1}{n}*\mathcal{S})_n\}_{n\ge0}$. In order to do this, we need a refinement of Proposition 6.2, which may be proven in a similar way.

Proposition 6.10. *Let the assumptions of Theorem 6.1 be satisfied. Let $m\in\mathbb{N}$ and set $Z_n^m=\frac{1}{n}\sum_{k=1}^{\lfloor m^{-1}n\rfloor}\tau^{-1}_{x_0(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n$. Finally, define $\Lambda_{x_0}(\lambda)=\log\mathbb{E}(e^{\langle\lambda,X_1\rangle})$. Then $\{Z_n^m\}_{n\ge1}$ satisfies in $T_{x_0}M$ the large deviation principle with good rate function*

$$I_m(v)=\frac{1}{m}\Lambda^*_{x_0}(mv),$$

where $\Lambda^*_{x_0}(v)=\sup_{\lambda\in T_{x_0}M}\{\langle\lambda,v\rangle-\Lambda_{x_0}(\lambda)\}$.

We are now able to prove the large deviations lower bound for $\{(\frac{1}{n}*\mathcal{S})_n\}_{n\ge0}$.

Proposition 6.11. *Let the assumptions of Theorem 6.1 be satisfied. Then for any $G\subset M$ open,*

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(\frac{1}{n}*\mathcal{S}\right)_n\in G\right)\ge-\inf_{x\in G}I_M(x),$$

where I_M is as in (6.1).

Proof. It suffices to show that

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left(\frac{1}{n}*\mathcal{S}\right)_n\in G\right)\ge-I_M(x)$$

for every $x\in G$.

So fix $x \in G$ and pick $v \in \mathrm{Exp}_{x_0}^{-1} x$. Because G is open, there exists an $\varepsilon > 0$ such that $B(x, \varepsilon) \subset G$. Let $m \in \mathbb{N}$ such that $m \geq \frac{2r}{\iota(\overline{B(x_0, r)})}$, where r is the uniform bound on the increments of the geodesic random walk.

We again need to identify the geodesic random walk with a tuple in $(T_{x_0}M)^m$. However, this time the parallel transport back to $T_{x_0}M$ is carried out by first transporting to a well-chosen point on the geodesic $\gamma_v(t) = \mathrm{Exp}_{x_0}(tv)$ and then to x_0 along this geodesic.

More precisely, we define a map $\Psi_{m,x,v} : (T_{x_0}M)^m \to M$ that allows us to identify the random variable $(\frac{1}{n} * \mathcal{S})_n \in M$ with a vector of random variables in $(T_{x_0}M)^m$. To this end, define for $0 \leq i \leq m$ the points $y_i = \mathrm{Exp}_{x_0}(\frac{i}{m} v)$. For $1 \leq i \leq m$ we define $\tau_{x_0 y_i}$ as parallel transport along the geodesic $\mathrm{Exp}_{x_0}(tv)$. Furthermore, for every $z \in M$ and every $0 \leq i \leq m-1$, we choose a geodesic $\gamma_{y_i x}$ of minimum length and denote by $\tau_{y_i x}$ parallel transport along this geodesic. We now define $\Psi_{m,x,v}(v_1, \ldots, v_m)$ as follows. Define $x_1 = \mathrm{Exp}_{x_0}(\frac{1}{m} v_1)$ and if x_i is defined, we set $\tilde{v}_{i+1} = \tau_{y_i x_i} \tau_{x_0 y_i} v_i$ and define $x_{i+1} = \mathrm{Exp}_{x_i}(\frac{1}{m} \tilde{v}_{i+1})$. Finally, we set $\Psi_{m,x,v}(v_1, \ldots, v_m) = x_m$.

Now note that by the triangle inequality, we have

$$d(x_i, x_0) \leq \frac{1}{m} \sum_{j=1}^{i} |v_j| \leq \frac{1}{m} \sum_{j=1}^{m} |v_j|$$

for any $1 \leq i \leq m$. Consequently, if $(v_1, \ldots, v_m) \in B(v, 1)^m$ we have $|v_j| \leq |v| + 1$, so that

$$d(x_i, x_0) \leq |v| + 1$$

for any $1 \leq i \leq m$. Because also $d(x_0, y_i) \leq \frac{i}{m}|v| \leq |v|$, we find that $x_i, y_i \in \overline{B(x_0, |v|+1)}$ for all $1 \leq i \leq m$.

Writing $\eta = |v| + 1$, we will show that there exists a constant $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$ we have

$$(v_1, \ldots, v_m) \in B(v, \varepsilon^2/(8\eta))^m \Rightarrow \Psi_{m,x,v}(v_1, \ldots, v_m) \in B(x, \varepsilon), \tag{6.11}$$

whenever $\varepsilon > 0$ is small enough.

To this end, let $K \subset M$ be a compact set, such that all geodesics of minimal length between points $x, y \in \overline{B(x_0, \eta)}$ are contained in K. Because K is compact, its injectivity radius $\iota(K)$ is strictly positive.

Fix $0 < \delta < \iota(K)$. We first show that for ε small enough and m large enough we have

$$d(x_i, y_i)^2 \leq \frac{i-1}{2m} \varepsilon^2 + \frac{i}{m^2} C \tag{6.12}$$

for $1 \leq i \leq m$. Here, $C > 0$ is some constant only depending on K and δ. We proceed by induction.

First consider the case $i = 1$. By taking m large enough, we can apply Proposition 5.10 to obtain a constant $C > 0$ (depending only on K and δ) such that

$$d(x_1, y_1)^2 = d\left(\mathrm{Exp}_{x_0}\left(\frac{1}{m} v_1\right), \mathrm{Exp}_{x_0}\left(\frac{1}{m} v\right)\right) \leq \frac{1}{m^2} C.$$

Now suppose that $d(x_i, y_i)^2 \leq \frac{i-1}{2m}\varepsilon^2 + \frac{i}{m^2} C$. Then in particular we have

$$d(x_i, y_i)^2 \leq \frac{\varepsilon^2}{2} + \frac{1}{m} C,$$

which can be made smaller than $\frac{\delta}{2}$ by taking ε sufficiently small and m sufficiently large. In that case, we may again apply Proposition 5.10, so that for the same constant $C > 0$

as above, we have

$$\begin{aligned} d(x_{i+1},y_{i+1})^2 &= d\left(\mathrm{Exp}_{x_i}\left(\frac{1}{m}\tau_{y_ix_i}\tau_{x_0y_i}v_{i+1}\right),\mathrm{Exp}_{y_i}\left(\frac{1}{m}\tau_{x_0y_i}v\right)\right)\\ &\leq d(x_i,y_i)^2+2\frac{1}{m}\langle\tau_{x_0y_i}v_{i+1}-\tau_{x_0y_i}v,\mathrm{Exp}_{y_i}^{-1}x_i\rangle+\frac{1}{m^2}C\\ &\leq \frac{i-1}{2m}\varepsilon^2+\frac{i}{m^2}C+2\frac{1}{m}|\tau_{x_0y_i}v_{i+1}-\tau_{x_0y_i}v||\mathrm{Exp}_{y_i}^{-1}x_i|+\frac{1}{m^2}C\\ &= \frac{i-1}{2m}\varepsilon^2+\frac{i+1}{m^2}C+\frac{2}{m}|v_i-v|d(x_i,y_i). \end{aligned}$$

Now, observe that $d(x_i,y_i)\leq 2\eta$ since $x_i,y_i\in\overline{B(x_0,\eta)}$. Using this, together with the induction hypothesis and the fact that $|v_i-v|\leq\frac{\varepsilon^2}{8\eta}$, we find

$$d(x_i,y_i)^2\leq\frac{i-1}{2m}\varepsilon^2+\frac{i+1}{m^2}C+\frac{1}{2m}\varepsilon^2=\frac{2i}{2m}\varepsilon^2+\frac{i+1}{m^2}C,$$

as desired.

Now taking $i=m$ in (6.12), we obtain

$$d(x_m,y_m)^2\leq\frac{\varepsilon^2}{2}+C\frac{1}{m},$$

whenever $(v_1,\dots,v_m)\in B(v,\varepsilon^2/(8\eta))^m$. Consequently, if we take $m_0>\frac{2C}{\varepsilon^2}$, we obtain for $m>m_0$ that

$$d\left(\Psi_{m,x,v}(v_1,\dots,v_m),x\right)^2=d(x_m,y_m)^2<\frac{\varepsilon^2}{2}+\frac{\varepsilon^2}{2}=\varepsilon^2$$

as desired.

Fixing m_0 and C as above, let $m\geq m_0$ be large enough so that we can define

$$\tilde{v}_k^{n,m,l}\in\mathrm{Exp}^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_l}}\left(\left(\frac{1}{n}*\mathcal{S}\right)_{n_l+k}\right)\subset T_{(\frac{1}{n}*\mathcal{S})_{n_l}}M$$

like in (6.7). Different from before, we now define the vectors

$$v_k^{n,m,l}=\tau^{-1}_{x_0y_{n_l}}\tau^{-1}_{y_{n_l}(\frac{1}{n}*\mathcal{S})_{n_l}}\tilde{v}_k^{n,m,l}\in T_{x_0}M, \tag{6.13}$$

using the parallel transport procedure used in the definition of the map $\Psi_{m,x,v}$.

Consequently, by construction we obtain

$$\Psi_{m,x,v}\left(v^{n,m,1}_{\lfloor m^{-1}n\rfloor},\dots,v^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right)=\left(\frac{1}{n}*\mathcal{S}\right)_n.$$

Using this, together with the implication in (6.11), we find

$$\begin{aligned} \mathbb{P}\left(\left(\frac{1}{n}*\mathcal{S}\right)_n\in G\right)&\geq\mathbb{P}\left(\left(\frac{1}{n}*\mathcal{S}\right)_n\in B(x,\varepsilon)\right)\\ &\geq\mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n\rfloor},\dots,v^{n,m,m}_{\lfloor m^{-1}n\rfloor}\right)\in B(v,\varepsilon^2/(8\eta))^m\right) \end{aligned}$$

Now define for $1\leq i\leq m$ the random variables

$$Y_i^n=\tau^{-1}_{x_0y_{n_{i-1}}}\tau^{-1}_{y_{n_{i-1}}(\frac{1}{n}*\mathcal{S})_{n_{i-1}}}\sum_{k=n_{i-1}+1}^{n_i}\tau^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_{i-1}}(\frac{1}{n}*\mathcal{S})_{k-1}}X_k^n\in T_{x_0}M,$$

where the parallel transport $\tau^{-1}_{(\frac{1}{n}*\mathcal{S})_{n_{i-1}}(\frac{1}{n}*\mathcal{S})_{k-1}}$ is carried out along the trajectory of the geodesic random walk. The sum is then transported from $T_{(\frac{1}{n}*\mathcal{S})_{n_{i-1}}}M$ to $T_{x_0}M$ as in the definition of $v_k^{n,m,l}$ as in (6.13).

In the same way as we obtained (6.8) in the proof of Proposition 6.8, we find that there exists a constant $\tilde{C} > 0$ such that

$$\left| v^{n,m,1}_{\lfloor m^{-1}n \rfloor} - Y_i^n \right| \leq \tilde{C}\frac{1}{nm} + \tilde{C}r^2\frac{1}{m^3}.$$

Consequently, we may take m large enough such that almost surely we have

$$\left| v^{n,m,1}_{\lfloor m^{-1}n \rfloor} - Y_i^n \right| < \frac{\varepsilon^2}{16\eta}.$$

But then we find that if $Y_i^n \in B(v, \varepsilon^2/(16\eta))$, then $v^{n,m,1}_{\lfloor m^{-1}n \rfloor} \in B(v, \varepsilon^2/(8\eta))$. This implies that

$$\begin{aligned}&\mathbb{P}\left(\left(v^{n,m,1}_{\lfloor m^{-1}n \rfloor}, \dots, v^{n,m,m}_{\lfloor m^{-1}n \rfloor}\right) \in B(v, \varepsilon^2/(8\eta))^m\right)\\ &\geq \mathbb{P}\left((Y_1^n, \dots, Y_m^n) \in B(v, \varepsilon^2/(16\eta))^m\right).\end{aligned}$$

Now note that, like in the proof of Proposition 6.2, we can show that the random variables Y_i^n and Y_j^n are independent and identically distributed for $i \neq j$, so that

$$\begin{aligned}\mathbb{P}\left((Y_1^n, \dots, Y_m^n) \in B(v, \varepsilon^2/(16\eta))^m\right) &= \prod_{i=1}^m \mathbb{P}\left(Y_i^n \in B(v, \varepsilon^2/(16\eta))\right)\\ &= \mathbb{P}\left(Y_1^n \in B(v, \varepsilon^2/(16\eta))\right)^m.\end{aligned}$$

Furthermore, by Proposition 6.10 we have

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(Y_1^n \in B(v, \varepsilon^2/(16\eta))\right) \geq -\frac{1}{m}\Lambda^*_{x_0}(v).$$

Combining everything, we find that

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in G\right) &\geq m \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(Y_1^n \in B(v, \varepsilon^2/(16\eta))\right)\\ &\geq -\Lambda^*_{x_0}(v).\end{aligned}$$

Since this holds for all $v \in \operatorname{Exp}_{x_0}^{-1} x$, we find that

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left(\frac{1}{n} * \mathcal{S}\right)_n \in G\right) \geq -\inf_{v \in \operatorname{Exp}_{x_0}^{-1} x} \Lambda^*_{x_0}(v) = -I_M(x),$$

which concludes the proof. □

A Some convex analysis

In this appendix we collect a result from convex analysis. Although this is probably well-known, we provide a proof for the reader's convenience.

Lemma A.1. *Let V be a vector space, and let $F : V \to \mathbb{R}$ be strictly convex and differentiable. Then its Legendre transform F^* is strictly convex and differentiable on the interior of its domain $\mathcal{D}^\circ_{F^*}$.*

Proof. The differentiability of F^* follows from [10, Theorem 26.3].

For the strict convexity, we first prove that for each $v \in \mathcal{D}^\circ_{F^*}$, there exists a $\lambda^*_v \in V$ such that

$$F^*(v) = \langle \lambda^*_v, v\rangle - F(\lambda^*_v).$$

Indeed, suppose this is not the case. Because $F^*(v) < \infty$, we can find a sequence λ_n such that

$$F^*(v) = \lim_{n\to\infty} \langle \lambda_n, v\rangle - F(\lambda_n).$$

Because the map $\lambda \mapsto \langle \lambda, v\rangle - F_x(\lambda)$ is continuous, the sequence λ_n cannot contain a convergent subsequence, else the limit of this subsequence would serve as λ^*_v. Consequently, we must have that $\lim_{n\to\infty} |\lambda_n| = \infty$.

But then there exists a $w \in V$ such that $\lim_{n\to\infty}\langle \lambda_n, w\rangle = \infty$. To see this, suppose such a w does not exist. Denoting by $e_1, \dots, e_d$ a basis of V, we must have that $\langle \lambda_n, e_i\rangle$ is a bounded sequence for all $i = 1, \dots, d$. But then, by taking subsequences, we find $\langle \lambda_n, e_i\rangle$ converges for all $i = 1, \dots, d$, which contradicts the fact that $\lim_{n\to\infty} |\lambda_n| = \infty$.

Now consider $v + \varepsilon w \in V$ and let λ_n be the sequence found above. We have that

$$F^*(v + \varepsilon w) \geq \lim_{n\to\infty} \langle \lambda_n, v + \varepsilon w\rangle - F(\lambda_n) = F^*(v) + \varepsilon \lim_{n\to\infty} \langle \lambda_n, w\rangle = \infty.$$

We conclude that $v + \varepsilon w \notin \mathcal{D}_{F^*}$ for any $\varepsilon > 0$, which contradicts the assumption that $v \in \mathcal{D}^\circ_{F^*}$.

We are now ready to prove that F^* is strictly convex on $\mathcal{D}^\circ_{F^*}$. To this end, fix $v, w \in \mathcal{D}^\circ_{F^*}$, $v \neq w$ and $t \in (0, 1)$ and assume that

$$F^*(tv + (1-t)w) = tF^*(v) + (1-t)F^*(w). \tag{A.1}$$

Now let λ^*_t be such that

$$F^*(tv + (1-t)w) = \langle tv + (1-t)w, \lambda^*_t\rangle - F(\lambda^*_t).$$

We find that

$$tF^*(v) + (1-t)F^*(w) = t(\langle \lambda^*_t, v\rangle - F(\lambda^*_t)) + (1-t)(\langle \lambda^*_t, w\rangle - F(\lambda^*_t)).$$

But then we find that

$$F^*(v) = \langle v, \lambda^*_t\rangle - F(\lambda^*_t)$$

and

$$F^*(w) = \langle w, \lambda^*_t\rangle - F(\lambda^*_t).$$

Now, because F is everywhere differentiable, it must be that $\nabla F(\lambda^*_t) = v$ and $\nabla F(\lambda^*_t) = w$, which contradicts the assumption that $v \neq w$. We conclude that F^* is strictly convex on $\mathcal{D}^\circ_{F^*}$. □

References

[1] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, second ed., Applications of Mathematics (New York), vol. 38, Springer-Verlag, New York, 1998. MR-1619036

[2] Jean-Dominique Deuschel and Daniel W. Stroock, *Large deviations*, Pure and Applied Mathematics, vol. 137, Academic Press, Inc., Boston, MA, 1989. MR-0997938

[3] Jin Feng and Thomas G. Kurtz, *Large deviations for stochastic processes*, Mathematical Surveys and Monographs, vol. 131, American Mathematical Society, 2006. MR-2260560

[4] Theodore Frankel, *The geometry of physics*, second ed., Cambridge University Press, Cambridge, 2004, An introduction. MR-2060795

[5] Frank den Hollander, *Large deviations*, Fields Institute Monographs, vol. 14, American Mathematical Society, Providence, RI, 2000. MR-1739680

[6] Erik Jørgensen, *The central limit problem for geodesic random walks*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **32** (1975), 1–64. MR-0400422

[7] Wilhelm Klingenberg, *Riemannian geometry*, de Gruyter Studies in Mathematics, vol. 1, Walter de Gruyter & Co., Berlin-New York, 1982. MR-0666697

[8] Richard Kraaij, Frank Redig, and Rik Versendaal, *Classical large deviations theorems on complete riemannian manifolds*, to appear in Stochastic Processes and their Applications, 2018.

[9] John M. Lee, *Riemannian manifolds: An introduction to curvature*, Graduate Texts in Mathematics, vol. 176, Springer-Verlag, New York, 1997. MR-1468735

[10] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR-0274683

[11] Michael Spivak, *A comprehensive introduction to differential geometry. Vol. I*, second ed., Publish or Perish, Inc., Wilmington, Del., 1979. MR-0532833

[12] D. W. Stroock, *An introduction to the theory of large deviations*, Universitext, Springer-Verlag, New York, 1984. MR-0755154

[13] Stefan Waldmann, *Geometric wave equations*, Lecture notes; ArXiv:1208.4706v1 (2012), no. 1208.4706v1.

Acknowledgments. The author thanks Frank Redig for helpful discussions, suggestions for certain arguments and the verification of various proofs. Thanks goes also to Jan van Neerven for suggestions on improving the clarity of certain definitions. The author acknowledges financial support from the Peter Paul Peterich Foundation via TU Delft University Fund.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System `http://www.vtex.lt/en/ejms.html`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`