Large deviations and concentration for random walks on hyperbolic spaces

Cagri Sert

Universität Zürich

Geometrische Gruppentheorie Seminar, WWU Münster, 06.05.2021

joint works with Adrien Boulanger–Pierre Mathieu–Alessandro Sisto (1) and Richard Aoun (2)

What is in this talk?

1. Large deviations and concentrations: the classical case
2. Large deviations on hyperbolic spaces
   - Random walks on metric spaces
   - Limit theorems and large deviations on hyperbolic spaces
   - Connection with random matrix products and LDP for spectral radius
3. Concentrations of random walks on hyperbolic spaces and Applications
   - Concentrations around the drift
   - An application: Finite-time probabilistic Tits alternative

Law of large numbers and deviations

Let $X_1, X_2, \dots$ be a sequence of independent and identically distributed (iid) real random variables.

For the purposes of this section, one can take X_i's to be independent Bernoulli random variables with distribution $\mu = p\delta_1 + (1-p)\delta_0$ (heads and tails). Equivalently, one can think of X_i's as coordinate functions on the Borel space $\{0,1\}^{\mathbb{N}}$ endowed with the probability measure $\mu^{\otimes \mathbb{N}}$.

Law of large numbers (in some form going back to Bernoulli) say that $\mu^{\otimes \mathbb{N}}$-a.s.

$$\frac{1}{n}\sum_{k=1}^{n} X_i \to \mathbb{E}[X] \quad (= p).$$

This is a particular case of Birkhoff's ergodic theorem.

Deviations around the mean

Once a law of large numbers type result is available, a natural follow-up/refinement is to consider the deviations around the mean.

At a first approximation, one may think of two different *scales of deviations*:

Moderate deviations

A choice of magnification of the deviations so as to lead to an asymptotic law of deviations: Central Limit Theorem (with $\sqrt{n}$-scaling) yields a limiting Gaussian distribution.

Large deviations

This concerns the study of deviations at "the largest meaningful scale": for sums of random variables, this is the linear scale n: Cramér's Theorem giving a precise description of linear order deviations (and refining LLN).

Large deviations 1

We will primarily focus on large deviations on the one hand on their asymptotics (Cramér type result, LDP) on the other hand on their finite time bounds (concentrations). Let us now briefly describe Cramér's result and introduce the relevant terminology.

By law of large numbers

$$\mathbb{P}(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X_1]| > t) \overset{n\to\infty}{\longrightarrow} 0$$

for every $t > 0$.

Typically this decay rate is exponential (because it is deviations of the largest scale) and one is interested in a precise description of the exponential decay rates as a function of deviation $t > 0$.

Large deviations 2

Definition

The process X_i is said to satisfy the large deviation principle (LDP) with rate function $I : \mathbb{R} \to [0, \infty]$ if for every interval $J \subset \mathbb{R}$, we have

$$\mathbb{P}(\frac{1}{n}\sum_{i=1}^{n} X_i \in J) \sim e^{-n(\inf_{\alpha \in J} I(\alpha))}.$$

Cramér's theorem ensures the existence of LDP for iid sums. For example, for the Bernoulli process, we have $I(\alpha) = \alpha \log(\frac{\alpha}{p}) + (1-\alpha)\log(\frac{1-\alpha}{1-p})$.

Concentrations

Notice that none of the above limit theorems say anything at finite time, that is for a given $n \in \mathbb{N}$. A precise understanding of probabilistic behaviour at finite time (and not only asymptotically) is crucial in many application in pure and applied mathematics.

Concentrations are concerned with finite time bounds in large deviations. A basic and fundamental result in this direction is the Hoeffding–Azuma bounds:

Suppose X_i's are iid random variables with values in $[a, b] \subset \mathbb{R}$. Then *for every* $n \in \mathbb{N}$ and $t \geqslant 0$, we have

$$\mathbb{P}(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X_1]| \geqslant t) \leqslant 2\exp(\frac{-2nt^2}{(b-a)^2}).$$

Random walks on metric spaces

We will consider random walks on metric spaces via their isometry groups and study the asymptotic properties of these processes.

More precisely,

- Let (M, d) be a metric space and $G = \mathrm{Isom}(M)$ the group of isometries of M.
- Let μ be a probability measure on G and let X_i's denote iid G-valued random variables with distribution μ e.g. $\mu = \frac{1}{2}\delta_g + \frac{1}{2}\delta_h$ for your favorite metric space M and isometries $g, h \in G$.
- Fix a basepoint $o \in M$ and consider the M-valued (Markov) process $X_n \dots X_1 o =: R_n o$.

We will be interested in the probabilistic behaviour of $d(X_n \dots X_1 o, o)$

Examples 1

Let us illustrate this on **examples**:

• $M = \mathbb{R}$ with the canonical metric: any probability measure on $\mathbb{R}$ can be thought of as a probability measure on $\mathbb{R} < \mathrm{Isom}(\mathbb{R})$ and $d(X_n \dots X_1 o, o)$ corresponds to the absolute value of a sum. $\longrightarrow$ In fact this very particular case corresponds to setting of *classical probability theory with* $\mathbb{R}$*-valued random variables*.

As opposed to previous one, the examples to come are non-commutative.

• $M =$ symmetric spaces of simple Lie groups, e.g. let us consider the **hyperbolic plane** $M = \mathbb{H}^2$. In this case, the group of isometries $G = O(2,1) > G^o \simeq \mathrm{PSL}(2,\mathbb{R})$. One can consider a mass distribution on her favorite collection of matrices, e.g. $\mu = \frac{1}{2}\delta_{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}} + \frac{1}{2}\delta_{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}$.

Examples 2 and Gromov hyperbolic spaces

• Finally, an example of a graph with rich symmetries: a **regular tree** with the natural metric structure. One can visualize random walks to get an intuition.

$*\quad*\quad*$

In fact, the metric spaces in all above examples fall into a class called **Gromov hyperbolic spaces**.

In the rest of the talk we will be mainly interested in random walks on those metric spaces (non-elementary ones).

Familiar examples are classical hyperbolic spaces, any CAT(-1)-manifold, trees or tree-like graphs (namely, Cayley graphs of non-elementary Gromov hyperbolic groups $\circlearrowleft$).

Back to the limit theorems 1

So let us fix a Gromov hyperbolic metric space M and consider a probability measure μ on $G = \mathrm{Isom}(M)$. Recall that we are interested in the process $d(X_n \dots X_1 o, o)$.

Law of large numbers: Under a finite first moment condition (i.e. $\int d(go, o) d\mu(g) < \infty$),there exists a real $\ell_\mu \geqslant 0$ (called *the drift*) such that

$$\frac{1}{n} d(X_n \dots X_1 o, o) \xrightarrow{a.s.} \ell_\mu \quad \text{as } n \to \infty.$$

This result is general and follows from the general subadditive ergodic theorem of Kingman (It is very hard to find the explicit value of ℓ_μ except for very simple cases).

Back to the limit theorems 2

Central limit theorem: Suppose that μ has a finite second order moment and it is *non-elementary*. Then, we have

$$\frac{d(X_n \dots X_1 o, o) - n\ell_\mu}{\sqrt{n}} \to \mathcal{N}(0, \sigma^2).$$

This result is due to culmination of works of Ledrappier('01), Bjorklund('11), Benoist–Quint('16), Horbez ('17), Mathieu–Sisto('20).

Large deviations for random walks on Gromov hyperbolic spaces

In this setting, the analogue of the classical result of Cramér is provided by the following

Theorem (Boulanger–Mathieu–S'–Sisto, '20)

Let (M, d) be a hyperbolic metric space.
1. *Suppose that μ is non-elementary. Then, the sequence $\frac{1}{n}d(X_n \dots X_1 o, o)$ satisfies a Large Deviation Principle with a convex rate function $I : [0, \infty) \to [0, \infty]$.*

2. *If μ has a finite exponential moment, then I is proper and has a unique zero at the drift ℓ_μ.*

Remarks and questions on the LDP 1

Remark

Very recently, Gouëzel (Feb '21) has partially extended the second assertion of the above theorem by getting rid of the finite exponential moment assumption: he showed that for a non-elementary probability measure μ, we have $I(x) > 0$ for every $x < \ell_\mu$.

Remark (Bounded support case)

Suppose that the support of μ is bounded (e.g. μ is finitely supported), then for some $M > 0$, we have $\frac{1}{n}d(X_n \dots X_1 o, o) \leqslant M$ (deterministically), so that the effective support $D_I := \{x \mid I(x) < \infty\}$ is compact (interval by convexity). In this case, we can describe D_I as a Hausdorff limit of deterministic objects. In particular it depends only on the support of μ.

Remarks and questions on the LDP 2

Question

Here is the expression of the LDP rate function of the nearest neighborhood random walk on the $2q$-regular tree:

$$I(x) = \frac{1+x}{2}\log(1+x) + \frac{1-x}{2}\log(1-x) + \log(q) - \frac{1+x}{2}\log(2q-1),$$

for $x \in [0,1]$ and ∞ elsewhere. One observes

1) *Unique zero on the drift $1 - 1/q$, the value at zero is given by the Kesten spectral radius.*

2) *It is strictly convex, analytic, and the second derivative at the drift can be expressed in terms of the variance of the associated CLT.*

LDP for random matrix products 1

Let us consider the analogous problem for random matrix products: namely, let μ be a probability measure on $\mathrm{SL}_d(\mathbb{R})$ and consider the process $\frac{1}{n}\log||X_n \dots X_1||$, where X_i's are iid with distribution μ.

In fact the above process fits in the general context of random walks on metric spaces (not necessarily Gromov hyperbolic) presented before:

Namely one can express the $d(X_n \dots X_1 o, o)$ in the associated Riemannian symmetric space by $\log||X_n \dots X_1||$'s. E.g. for $\mathrm{SL}(2,\mathbb{R}) \curvearrowright \mathbb{H}^2$, we have $2\log||X_n \dots X_1|| = d(X_n \dots X_1 o, o)$.

LDP for random matrix products 2

The existence of LDP (with convex rate function I) for the process $\frac{1}{n}\log||X_n \dots X_1||$ was settled earlier (S. '16, thesis).

Remark: In that context, the analogous quantity to drift ℓ_μ is the top Lyapunov exponent $\lambda_1(\mu)$. The analogous result to one in BMSS that $\lambda_1(\mu)$ is the unique zero of the rate function I corresponds to a result of Le Page, proved much earlier ('82) using spectral methods.

* * *

A conjecture in (S.) says that the <u>spectral radius</u> $\frac{1}{n}\log\rho(X_n \dots X_1)$ of random matrix products

1) satisfies a LDP

2) and this, with the same rate function as for the norm $||.||$.

LDP for translation distance

The analogous quantity to spectral radius in metric geometry is the translation distance of an isometry. Namely, for $g \in \mathrm{Isom}(M)$, we set $\tau(g) = \lim_{n\to\infty} \frac{1}{n} d(g^n o, o)$. For a Gromov hyperbolic metric space this is also equal to (up to an additive constant) $\inf_{x \in M} d(gx, x)$.

Theorem (Boulanger–Mathieu–S'–Sisto, '20)

Given a Gromov hyperbolic space and a non-elementary probability measure μ on $\mathrm{Isom}(M)$ with bounded support, the process $\frac{1}{n}\tau(X_n \ldots X_1)$ satisfies the same LDP as the average displacement process $\frac{1}{n} d(X_n \ldots X_1 o, o)$.

Whereas it is known that the spectral radius (or τ) satisfies several of the other limit laws in the same way as norm (or displacement) the LDP conjecture for (non-hyperbolic) random matrix products remain open.

1 Large deviations and concentrations: the classical case

2 Large deviations on hyperbolic spaces
- Random walks on metric spaces
- Limit theorems and large deviations on hyperbolic spaces
- Connection with random matrix products and LDP for spectral radius

3 Concentrations of random walks on hyperbolic spaces and Applications
- Concentrations around the drift
- An application: Finite-time probabilistic Tits alternative

Section summary

I remind that the previous LDP theorems are by nature asymptotic and do not yield information on fixed finite time behaviour. In recent ('21) collaboration with Aoun, we refined the large deviation result (for displacement) in the direction of concentration estimates and obtained several applications.

Subgaussian concentrations around the drift 1

Let (M, d) be a proper Gromov hyperbolic metric space. Proper means that closed balls are compact; this covers all examples discussed above.

It is known that if M is proper, than the isometry group $\mathrm{Isom}(M)$ with its natural topology is locally compact and hence carries a Haar measure μ_G. Using this, given a probability measure μ on $\mathrm{Isom}(M)$, one associates a bounded linear operator $\rho(\mu)$ on $L^2(G, \mu_G)$: it is simply the μ-averaging operator, $\rho(\mu)f(x) = \int f(g^{-1}x)d\mu(g)$.

For simplicity, we assume μ to be symmetric $(\mu(g) = \mu(g^{-1}))$ and denote by λ_μ the norm $||\rho(\mu)||_2$.

Subgaussian concentrations around the drift 2

Finally, given a bounded set $S \subset G$, we denote $\kappa_S = \sup\{d(go, o) \mid g \in S\}$. Our result reads

Theorem (Aoun–S', '21)

Let M be a proper Gromov hyperbolic metric space such that $\mathrm{Isom}(M) \curvearrowright M$ cocompactly. Then, there is an explicit positive function $D()$ such that for every non-elementary probability measure μ with bounded support S, for every $t \geqslant 0$ and $n \in \mathbb{N}$, we have

$$\mathbb{P}(|\frac{1}{n} d(X_n \dots X_1 o, o) - \ell_\mu| \geqslant t) \leqslant 2 \exp(\frac{-nt^2}{\kappa_S^2 D(\kappa_S, \lambda_\mu)}).$$

Remarks on the statement

Remark (On the statement)

1. The only additional term with respect to the classical Azuma–Hoeffding inequalities is the red term $D(\kappa_S, \lambda_\mu) = 2^9(\log^+(\kappa_S) + A_M)\frac{1}{(1-\sqrt{\lambda_\mu})^4}$*.*

2. The constant A_M *is a doubling constant depending only on* M *and in concrete situations its explicit value can readily be calculated. Similarly, in concrete cases, the operator norm* λ_μ *can be effectively bounded from above which yields explicit numerical estimates in the above result.*

Remarks on the consequences

Remark (On consequences)

0. It allows one to probabilistically locate the drift ℓ_μ.

1. One can deduce a result on the continuity of the drift $\mu \mapsto \ell_\mu$.

2. With some further work, one can prove a finite-time probabilistic Tits alternative.

Classical and probabilistic Tits alternatives 1

Recall that a (say countable) group Γ is called amenable if it has a Folner sequence: there exists an increasing sequence of finite subset $F_n \subset \Gamma$ such that for every $\gamma \in \Gamma$, we have $\frac{\#(\gamma F_n \Delta F_n)}{\# F_n} \to 0$ as $n \to \infty$.

This notion was introduced by von Neumann ('29) and was partly motivated by the Banach–Tarski paradox. It has a number of equivalent formulations.

Von Neumann conjectured that a non-amenable group must contain non-abelian free subgroups; this was refuted by Olshanskii in '80s.

However, von Neumann's prediction is true for some familiar classes of groups, this is the content of Tits alternatives.

Classical and probabilistic Tits alternatives 2

Jacques Tits ('72) proved that Von Neumann's conjecture holds for linear groups. Namely, any linear group is either amenable or contains a free subgroup.

Since then, results of Tits' type have been proven for a variety of classes of groups (in particular Gromov hyperbolic groups).

By a probabilistic Tits alternative for a group Γ, we understand a statement of the following type: for any probability measure μ on Γ either the support of μ generates an amenable subgroup or

$$\mathbb{P}((R_n, R_n') \text{ generates a free subgroup}) > 0 \text{ or better } \to 1.$$

In the setting of the original result of Tits (linear groups), his result was improved to a probabilistic Tits alternative with exponential speed by Aoun ('11).

In our joint work, in the setting of hyperbolic groups, using our concentration results, we proved the following finite-time probabilistic Tits alternative

Theorem (Aoun-S', March '21)

Let M be a proper Gromov hyperbolic geodesic metric space. Let $\Gamma < \mathrm{Isom}(M)$ be a countable subgroup acting cocompactly and properly on M. Then, there exists an explicit function $T > 0$ such that for any probability measure μ on Γ, either the support of μ is amenable, or for every n larger than an explicit constant $C(M, \mu)$, we have

$$\mathbb{P}((R_n, R_n') \textit{ generates a free subgroup}) > 1 - 84\exp(-nT(\kappa_S, \lambda_\mu)).$$

Thank you

Thanks for your attention!