An invariance principle for random walks on graphs

Fabian Gerle

Universität Duisburg-Essen

Research training group 2131
January 11, 2016
Introduction

**Invariance principle**

- $S = (E, d, \nu)$ metric measure space
- $X$: stochastic process on $S$ determined by the geometric structure of $S$, e.g.
  - simple nearest neighbor random walk on a graph
  - one dimensional drift-free diffusion
  - Brownian motion on manifolds
- $\forall n \in \mathbb{N}: S^n = (E^n, d^n, \nu^n)$ and associated processes $X^n$.
- Invariance principle: $S^n \xrightarrow{n \to \infty} S \Rightarrow X^n \xrightarrow{n \to \infty} X$.
  - For random walks on trees: Athreya, Löhr, and Winter (to appear)
  - For random walks on (sub-)Riemannian manifolds: Gordina and Laetsch (2014)
  - For Brownian motion on a class of metric measure spaces: Suzuki (to appear)
Motivation

Donsker’s invariance principle

- Let \((X_i)_{i \geq 1}\) be a sequence of iid random variables with
  \[\Pr(X_1 = 1) = \Pr(X_1 = -1) = \frac{1}{2}p, \quad p \in (0, 1)\]
  and
  \[\Pr(X_1 = 2) = \Pr(X_1 = -2) = \frac{1}{2}(1 - p).\]

- Then \(\mathbb{E}[X_1] = 0\) and \(\text{Var}(X_1) = 4 - 3p =: \sigma^2\).

- For \(n \in \mathbb{N}\) let \(S^n = (S^n_t)_{t \geq 0}\) with
  \[S^n_t := \frac{1}{\sqrt{\sigma^2 n}} \sum_{i=1}^{\lfloor nt \rfloor} X_i\]

- If \(B = (B_t)_{t \geq 0}\) is a standard BM, then by Donsker’s invariance principle
  \[\mathcal{L}(S^n) \xrightarrow{n \to \infty} \mathcal{L}(B)\]
  weakly in Skorohod space \(D([0, \infty))\).
Motivation

Random walks on graphs

- For $n \in \mathbb{N}$ let $\Gamma^n = (G^n, \mu^n, \nu^n)$ be a weighted measure graph with
  - vertices $G^n := \mathbb{Z}$,
  - symmetric map $\mu^n : G^n \times G^n \to \mathbb{R}^+$ with
    $$\mu^n_{xy} = \begin{cases} \sqrt{np} =: \mu^n_1, & \text{if } |x - y| = 1, \\ \frac{\mu^n_1(1-2p)}{2p} = \sqrt{n\frac{1}{2}}(1 - 2p) =: \mu^n_2, & \text{if } |x - y| = 2. \end{cases}$$
  - and an atomic measure $\nu^n : G \to \mathbb{R}^+$ with
    $$\nu^n_x = \frac{1}{\sqrt{n}}$$
- $Y^n = (Y^n_t)_{t \geq 0}$: (continuous-time) Markov chain on $\Gamma^n$, which jumps from $x \in \mathbb{Z}$ to $y \in \mathbb{Z}$ at rate
  $$\gamma^n_{xy} = \frac{\mu^n_{xy}}{\nu^n_x} \quad (x \sim_n y :\Leftrightarrow \mu^n_{xy} > 0)$$
Motivation
Random walk with jumps as nearest neighbor random walk on graphs

\( Y^n: \text{speed-} \nu^n \text{ random walk on } (G^n, \mu^n). \)

Observations

- \( Y^n \) jumps at a constant rate of \( \gamma^n = \gamma^n_x = \sum_{y:y \sim n x} \gamma^n_{xy} = n \)
- \( \Gamma^n \) "converges" to \( \mathbb{R} \) equipped with the Lebesgue measure.
- Donsker’s theorem implies \( Y^n \xrightarrow{n \to \infty} B \) pathwise.
- Restriction to jumps of length 2 is unnecessary as long as \( \mathbb{E}[X] = 0 \) and \( \text{Var}(X) < \infty \).
Motivation
Invariance principle for random walks on trees

Athreya, Löhr, and Winter (to appear) have verified an invariance principle for random walks on trees:

- If metric measure trees $(T^n, r^n, \nu^n)$ "converge" to some compact $(T, r, \nu)$
- then the speed-$\nu^n$ random walks converge to a limit process.
- Limit is called speed-$\nu$ motion.

**Goal:** Show a similar result for graphs converging to trees.
Invariance principle for random walks on trees

Trees

Let \((T, r)\) be a compact metric space such that

- \((T, r)\) is 0-hyperbolic:
  - For all \(x, y, z, p \in T\) the following four point condition holds:
    \[ r(x, y) + r(z, p) \leq \max\{r(x, z) + r(y, p), r(x, p) + r(y, z)\}. \]
  - "No cycles"

- \((T, r)\) is fine:
  - For all \(x_1, x_2, x_3 \in T\) there exists a point \(c = c(x_1, x_2, x_3) \in T\) such that for \(i, j \in \{1, 2, 3\}, i \neq j\)
    \[ r(x_i, c) + r(x_j, c) = r(x_i, x_j) \]
  - "Branchpoints exist"

Then \((T, r)\) is a metric tree.
A rooted metric measure tree (RMMT) is a quadruple \((T, r, \nu, \varrho)\) such that

- \((T, r)\) is a compact metric tree,
- \(\nu\) is a finite measure on \(T\) with \(\text{supp}(\nu) = T\),
- \(\varrho \in T\).

Two RMMT \((T, r, \nu, \varrho)\) and \((T', r', \nu', \varrho')\) are equivalent, iff there exists an isometry \(\varphi\) between \((T, r)\) and \((T', r')\) such that

- \(\varphi(\varrho) = \varrho'\) and
- \(\nu \circ \varphi^{-1} = \nu'\).

\(\mathcal{T}\) : equivalence class of (compact) rooted metric measure trees.
Invariance principle for random walks on trees

The speed-$\nu$ motion

$(T, r, \nu)$ a compact metric measure tree. Athreya, Eckhoff, and Winter (2013) and Athreya, Löhr, and Winter (2014) have shown

**Proposition Occupation time formula**

There exists a unique strong Markov process $X$ (up to $\nu$-equivalence) on $(T, r)$ satisfying

$$
\mathbb{E}_X \left[ \int_0^{\tau_y} f(X_s) \, ds \right] = 2 \int_T r(y, c(x, y, z))f(z) \, d\nu(z)
$$

for all $x, y \in T$ and positive, bounded measurable $f$. $X$ is called *speed-$\nu$ motion* on $(T, r)$.

$\tau_y := \inf \{ t \geq 0 \mid X_t = y \}$: First hitting time of $y$. 
Invariance principle for random walks on trees
Gromov-Hausdorff-weak topology

For $n \in \mathbb{N} \cup \{\infty\}$ let $\mathcal{X}^n = (T^n, r^n, \nu^n, \varrho^n) \in \mathcal{T}$. Then $\mathcal{X}^n \xrightarrow{n \to \infty} \mathcal{X}^\infty$ pointed Gromov-Hausdorff-weakly iff

$\exists (E, d_E, \varrho_E)$ compact rooted metric space and $\varphi_n : T_n \to E$ isometries with $\varphi_n(\varrho_n) = \varrho_E$ such that

1. the supports of $\nu^n$ converge in Gromov-Hausdorff topology, i.e.
   $$\text{supp}((\varphi_n)_* \nu_n) \xrightarrow{n \to \infty} \text{supp}((\varphi_\infty)_* \nu_\infty)$$
   in the Hausdorff sense

2. and
   $$(\varphi_n)_* \nu_n \xrightarrow{n \to \infty} (\varphi_\infty)_* \nu_\infty$$ weakly
Invariance principle for random walks on trees

The invariance principle

**Theorem** Athreya, Löhr, and Winter (2014)

Let \((\mathcal{X}^n)_{n \in \mathbb{N}}\) and \(\mathcal{X}\) be in \(T\) and \(X^n, X\) the speed-\(\nu^n\) and speed-\(\nu\) motions started in \(\varrho^n\) and \(\varrho\) respectively. If \(\mathcal{X}_n \xrightarrow{n \to \infty} \mathcal{X}\) Gromov-Hausdorff-weakly, then

\[
\mathcal{X}^n \xrightarrow{n \to \infty} \mathcal{X}
\]

weakly in path space.
Invariance principle for random walks on trees

Example

**Reflected Brownian motion**

Let

- \( T^n = \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\} \),
- \( r^n(x, y) := |x - y| \),
- \( \nu^n = \frac{1}{n} \).

Then \( U(T^n, \frac{1}{n}) = \{ x \in \mathbb{R} | \exists y \in T^n : |x - y| \leq \frac{1}{n} \} \subset [0, 1] \)

\( \Rightarrow d_{GH}(T^n, [0, 1]) \leq \frac{1}{n} \)

\( \nu^n \) converges to the Lebesgue measure \( \lambda \) weakly.

Invariance principle \( \Rightarrow \) speed-\( \nu^n \) motion converges to speed-\( \lambda \) motion on \([0, 1]\), i.e. reflected Brownian motion on \([0, 1]\) pathwise.
Example

**Moran Model**

- from mathematical biology
- models the frequency of an allele in a population
- \( n \in \mathbb{N} \) individuals; fixed
- state space: gene frequencies \( \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\} \)
- stochastic process \((X_t^n)_{t \in \mathbb{N}}\): At each timepoint
  - one randomly selected individual dies
  - and another (randomly selected) individual gives birth (to an identical copy)
- transition probabilities:
  \[
p(x, x \pm 1/n) = x(1-x), \quad p(x, x) = 1 - 2x(1-x)
\]
- scaling limit: Wright-Fisher diffusion \( X \) solves the SDE
  \[
dX_t = 1_{[0,1]}(X_t) \sqrt{2X_t(1-X_t)} dB_t
\]
Invariance principle for random walks on trees

Example

The Moran model as speed-$\nu^n$ random walk on a tree

- $T^n = \left\{ \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n} \right\}$, $r^n(x, y) = |x - y|
- $X^n$ speed-$\nu^n$ motion on $(T^n, r^n)$
- $X^n$ jumps in $x \in T^n$ at rate

$$
\gamma^n_x = \frac{2n}{\nu^n_x} = \frac{2n^2}{x(1 - x)}
$$

- $\Rightarrow \nu^n(x) = x(1 - x)/n$
- then $(T^n, r^n, \nu^n) \underset{n \to \infty}{\longrightarrow} ((0, 1), d_\mathbb{R}, \nu)$ Gromov-Hausdorff weakly with

$$
\nu(dx) = \frac{dx}{x(1 - x)}
$$
Invariance principle for random walks on trees
Sketch of the proof

Usual procedure for proving limit theorems for stochastic processes:

1. **Existence**
   - show tightness of the family \((X^n)_{n \in \mathbb{N}}\), i.e. show that cluster points exist.

2. **Uniqueness**
   - show that all limit points agree almost surely.
Invariance principle for random walks on trees

Sketch of the proof

1. **Tightness**

   Corollary to **Aldous’ tightness criterion**: If for all $\varepsilon > 0$

   $$\lim_{t \to 0} \lim_{n \to \infty} \sup_{x \in T^n} \mathbb{P}^x(r^n(x, X^n_t) \geq \varepsilon) = 0$$

   then the family $(X^n)_{n \in \mathbb{N}}$ is tight.

Proceed in two steps:

- Hitting time bound: Upper bound on $\mathbb{P}_x(\tau_y < t)$ as $t \to 0$.

- Topological bound: Number of directions to get "far" away is uniformly bounded in $n$.

- Both steps rely heavily on the tree structure of $X^n$
Invariance principle for random walks on trees
Sketch of the proof

2 Uniqueness

Proof consists of two parts: Show

- all limit points possess the strong Markov property
- all limit points satisfy the occupation time formula

⇒ Uniqueness of the limit.

- This proof can easily be adapted to the graph-case.
- Showing tightness as main task!
A compact weighted measure graph is a triple \((G, \mu, \nu)\), where

- **\(G\)** is a finite set (vertices),
- **\(\mu : G \times G \rightarrow \mathbb{R}^+\)** a symmetric map (edge conductances),
- **\(\nu : G \rightarrow \mathbb{R}^+\)** an atomic measure.
- Two points \(x, y \in G\) are connected by an edge (\(x \sim y\)), if \(\mu_{xy} > 0\).
- Assume that any two points \(x, y \in G\) are connected, i.e. \(\exists (x_0, \ldots, x_n) = (x_i)_0^n \in G\) with \(x_0 = x, x_n = y\) and \(x_i \sim x_{i-1}, 1 \leq i \leq n\).
- A path of length \(n\) connecting \(x, y \in G\) is a tuple of points \(x_0, \ldots, x_n \in G\) as above.
- The weighted vertex degree of \(x \in G\): \(\mu_x := \sum_{y : x \sim y} \mu_{xy}\)
Graphs

Graph distance

- Need a metric on \((G, \mu)\) to apply metric measure space theory.
- Natural metric: weighted graph distance, i.e. for \(x, y \in G\)

\[
    d_\mu(x, y) := \inf \left\{ \sum_{i=1}^{n} \mu_{x_{i-1}x_{i}}^{-1} \left| (x_i)_0^n \text{ is a path from } x \text{ to } y \right. \right\}.
\]

- "Right" metric on trees but not suitable for graphs.
- Ignores graph-structure, e.g.

\[
\begin{align*}
\text{mutual distances w.r.t. } d_\mu \text{ are the same in both graphs.}
\end{align*}
\]
Graphs
Speed-\(\nu\) random walks on graphs

\((G, \mu, \nu)\) a weighted graph.

**Speed-\(\nu\) random walk on** \((G, \mu)\):

- continuous time Markov chain \(X\) on \(G\)
- with generator

\[
Af(x) := \frac{\mu_x}{\nu_x} \left( \sum_{y : x \sim y} \frac{\mu_{xy}}{\mu_x} f(y) - f(x) \right)
\]

and \(\mathcal{D}(A) = \{f : G \to \mathbb{R}\}\)

- and semigroup

\[
P_t f(x) := \mathbb{E}^x [f(X_t)] = e^{tA} f(x)
\]
For $f \in \mathcal{D}(A)$ define the **energy** of $f$

$$\mathcal{E}(f, f) := \frac{1}{2} \sum_{x, y \in G} \mu_{xy} (f(y) - f(x))^2.$$ 

$\mathcal{E}$ can be made into a regular *Dirichlet form* by polarization.

For $x, y \in G$ define the **effective resistance** $\mathcal{R}(x, y)$ between $x$ and $y$:

$$\mathcal{R}(x, y) := (\inf \{ \mathcal{E}(f, f) | f \in \mathcal{D}(A), f(x) = 1, f(y) = 0 \})^{-1},$$

where $\inf \emptyset := \infty$.

Can be extended to the effective resistance between sets.
Graphs
Effective resistance metric

Facts
- $\mathcal{R}$ is a metric on $G$.
- $\mathcal{R} \leq d_\mu$.
- $\mathcal{R}$ and $d_\mu$ coincide if and only if $(G, d_\mu)$ is a tree.
- There exists a unique minimizer $f^y_x$ for the variational problem and $f^y_x$ is harmonic on $G \setminus \{x, y\}$, i.e.

$$Af^y_x(z) = 0 \quad \forall z \in G \setminus \{x, y\}.$$
Graphs
Effective resistance metric and Speed-$\nu$ motion

Why should the effective resistance metric be the right metric to formulate the invariance principle for random walks on graphs?

- Let $\tau_x := \inf \{ t \geq 0 \mid X_t = x \}$ and $\tau^+_x := \inf \{ t \geq 0 \mid X_t = x, \exists s < t : X_s \neq x \}$, then

\[ \mathbb{P}_x(\tau_y < \tau^+_x) = \frac{1}{\mathcal{R}(x, y)\mu_x}. \]

- If $\xi$ separates $x$ from $y$ (i.e. every path from $x$ to $y$ contains $\xi$) then

\[ \mathbb{P}_\xi(\tau_x < \tau_y) = \frac{\mathcal{R}(\xi, y)}{\mathcal{R}(x, y)}. \]
Graphs
Example

For \( n \in \mathbb{N} \) consider the graphs on the right. Then, using Kirchhoff’s law,

\[
R(x_1, y_1) = \frac{1}{1 + \frac{1}{1+2/n}} = \frac{2 + n}{2 + 2n}
\]

and hence

\[
P_{x_1}(\tau_{z_1} < \tau_{y_1}) = \frac{2 + 2n}{4 + 3n}
\]

\[
\xrightarrow{n \to \infty} \frac{2}{3}.
\]

\[\Rightarrow \quad \mu_{xy} = 2\]
Graphs
Balls in the effective resistance metric

Consider the connected graph

\[
\begin{align*}
\mu_{xy} &= 2 \\
\mu_{yi} &= 1 \\
\mu_{iz} &= 1 \\
\end{align*}
\]

Simple calculations show

\[
\mathcal{R}(x, z) = \frac{9}{10} \quad \text{and} \quad \mathcal{R}(x, v_i) = \frac{11}{10}, \quad i \in \{1, \ldots, 5\}. 
\]

\[\Rightarrow z \in B(x, 1) \text{ but } v_i \notin B(x, 1). \text{ Hence } B(x, 1) \text{ is not connected!}\]
Graphs
Gromov hyperbolic spaces

Hyperbolicity as a measure for "tree-likeness"

Let $\delta \geq 0$. A metric space $(E, d)$ is $\delta$-hyperbolic, iff for all $x, y, z, p \in G$ the four point condition holds:

$$d(x, y) + d(z, p) \leq \max\{d(x, z) + d(y, p), d(y, z) + d(x, p)\} + 2\delta.$$

If $(E, d)$ is a length space then Gromov hyperbolicity is equivalent to

- for each $x_1, x_2, x_3 \in E \exists \delta \geq 0$:

\[ [x_1, x_2] \subset U_\delta([x_1, x_3]) \cup U_\delta([x_2, x_3]), \]

where $[x_i, x_j]$ is the geodesic segment joining $x_i$ and $x_j$ and $U_\delta(A)$ is the $\delta$-neighborhood of $A \subset E$.

- Geodesic triangles are $\delta$-thin.
Graphs
Examples

- Trees are length spaces and 0-hyperbolic.
- In general, graphs are not length spaces.
- The graphs

\[
\begin{align*}
\mu_{x_1x_2}^n &= n \\
\mu_{x_1y_1}^n &= 1 \\
\mu_{y_1y_2}^n &= n \\
\mu_{x_1z_1}^n &= 1 \\
\mu_{y_1z_1'}^n &= 1
\end{align*}
\]

are \( \frac{1}{n} \)-hyperbolic w.r.t. \( d_\mu \) and \( \frac{1}{2n^2} \)-hyperbolic w.r.t. the effective resistance metric.

- The complete symmetric graph is 0-hyperbolic.
  (But no tree!)
Graphs

Graphs converging to trees

For $n \in \mathbb{N}$ let

- $\Gamma^n = (G^n, \mu^n, \nu^n)$ weighted measure graph
- $\mathcal{R}^n$ effective resistance metric on $(G^n, \mu^n)$

Assume $(T, d)$ is a metric tree and

$$(G^n, \mathcal{R}^n) \xrightarrow{n \to \infty} (T, r)$$ in the Gromov-Hausdorff sense.

$\Rightarrow (G^n, \mathcal{R}^n)$ is $\delta_n$-hyperbolic and $\delta_n \xrightarrow{n \to \infty} 0$.

Conjecture

Let $(G, \mathcal{R})$ be $\delta$-hyperbolic then $\exists$ metric tree $(T, r)$ and $A \subset T$ s.t.

$$d_{\text{GH}}(G, A) \leq C\delta,$$

where $C > 0$ is a universal constant.
An invariance principle for random walks on graphs

Idea

**Tightness** (Idea&Plan)

- \((G^n, \mu^n, \nu^n, \varrho^n)\) rooted weighted measure graphs, \(X^n\) speed-\(\nu^n\) random walks, and \((T, r, \nu, \varrho)\) RMMT, 

- \((G^n, \mathcal{R}^n, \nu^n, \varrho^n) \xrightarrow{n \to \infty} (T, r, \nu, \varrho)\) pointed Gromov-Hausdorff-weakly.

- For some \(C > 0\) find metric trees \((T^n, r^n)\) and \(\varphi^n : G^n \to T^n\) s.t. \(d_{GH}(G^n, \varphi^n(G^n)) < C\delta_n\) and

- choose a measure \(\hat{\nu}^n\) on \((T^n, r^n)\) s.t. 
  \((T^n, r^n, \hat{\nu}^n) \xrightarrow{n \to \infty} (T, r, \nu)\) Gromov-Hausdorff-weakly.

- From the speed-\(\hat{\nu}^n\) random walks on \((T^n, r^n)\) construct stochastic processes \(Z^n\) on \((G^n, \mathcal{R}^n)\) via coupling s.t. 
  \(d_{Sk}(X^n, Z^n) \leq C\delta_n\).

- Finally use tightness of \(Y^n\) to show tightness of \(X^n\)
Thank you for your attention!

