Non-Markovian coupling of sub-Riemannian diffusions

Robert Neel

Department of Mathematics
Lehigh University

September 9, 2024
Workshop: Stochastics and Geometry
BIRS, Banff, Canada

Acknowledgments

I would like to thank Masha, Todd, Jing, and Tai for organizing the workshop and for the invitation.

This is based on joint work with Liangbing Luo (Queen's University, Ontario).

Couplings

Consider a diffusion, which we'll call Brownian motion, B_t on a manifold (say, Riemannian or sub-Riemannian), starting from any $x \in M$. A coupling of such BMs is a process $(B_t, \tilde{B}_t)$ on $M \times M$, from $(x, \tilde{x})$ such that each marginal is a BM.

Couplings

Consider a diffusion, which we'll call Brownian motion, B_t on a manifold (say, Riemannian or sub-Riemannian), starting from any $x \in M$. A coupling of such BMs is a process $(B_t, \tilde{B}_t)$ on $M \times M$, from $(x, \tilde{x})$ such that each marginal is a BM.

Goal:

- get the processes to meet quickly;
- that is, to find a joint distribution so that the coupling time $\tau = \inf\{t > 0 : B_t = \tilde{B}_t\}$ is as small as possible. This leads to "reflection-style" couplings.
- (Another possible goal is to keep B_t and $\tilde{B}_t$ a.s. as close as possible. This leads to "parallel-style" couplings.)

Consequences of "reflection-type" couplings

The Aldous inequality says the total variation distance between the laws of the processes satisfies

$$\operatorname{dist}_{\mathrm{TV}}\left(\mathcal{L}\left(B_t\right), \mathcal{L}\left(\tilde{B}_t\right)\right) \leq \mathbb{P}\left(\tau > t\right).$$

With appropriate dependence of τ on $\operatorname{dist}\left(x, \tilde{x}\right)$, this gives gradient bounds for the heat semigroup.

Also connections to Liouville properties, first eigenvalue on a compact manifold, etc.

Quality of reflection couplings

- A coupling is *successful* if $\tau < \infty$ a.s.
- A coupling is *efficient* if

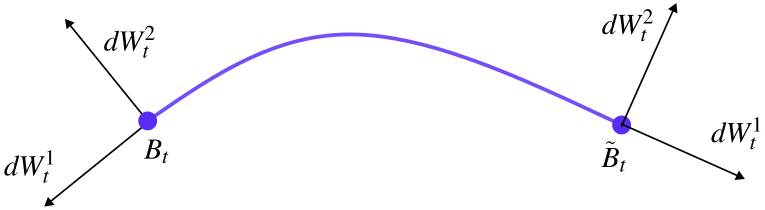
$$\frac{\mathbb{P}\left(\tau > t\right)}{\mathrm{dist}_{\mathrm{TV}}\left(\mathcal{L}\left(B_t\right), \mathcal{L}(\tilde{B}_t)\right)}$$

  stays bounded as $t \to \infty$.
- A coupling is *maximal* if

$$\mathrm{dist}_{\mathrm{TV}}\left(\mathcal{L}\left(B_t\right), \mathcal{L}\left(\tilde{B}_t\right)\right) = \mathbb{P}\left(\tau > t\right)$$

  for all $t > 0$.

Kendall-Cranston mirror coupling



On a Riemannian manifold, we have a Markovian reflection scheme, where we infinitesimally reflect along the minimal geodesic from B_t to $\tilde{B}_t$.

This gives an SDE on $M \times M$ for the joint process. (See Elton Hsu's book.)

Properties of this mirror coupling

- Technically, the cut locus is an issue. Resolved using random walk approxiamtion (von Renesse '04).

Properties of this mirror coupling

- Technically, the cut locus is an issue. Resolved using random walk approxiamtion (von Renesse '04).
- Because it's Markov, Itô's lemma gives an SDE for the distance between the particles, which is compatible with standard comparison geometry.
- It yields the sharp lower bound on $\lambda_1(M)$ for both
  - compact manifolds with positive lower Ricci bound, in terms of the bound,
  - and compact manifolds with non-negative Ricci curvature, in terms of diameter (Zhong and Yang '84).

Maximal couplings on model spaces

The sharpness above is related to the fact that this mirror coupling is maximal on the model spaces $\mathbb{S}^n$ and $\mathbb{R}^n$.

Maximal couplings on model spaces

The sharpness above is related to the fact that this mirror coupling is maximal on the model spaces $\mathbb{S}^n$ and $\mathbb{R}^n$.

In general, there is an abstract existence result for maximal couplings (Sverchkov-Smirnov '90). They are not unique, even for $\mathbb{R}$ (Hsu-Sturm '13, credited to Fitzpatrick).

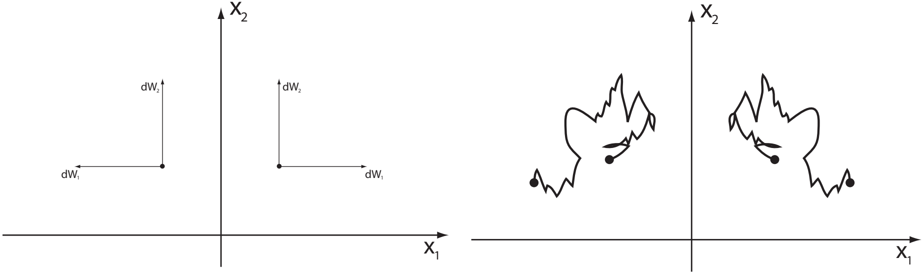
Maximal couplings on model spaces

The sharpness above is related to the fact that this mirror coupling is maximal on the model spaces $\mathbb{S}^n$ and $\mathbb{R}^n$.

In general, there is an abstract existence result for maximal couplings (Sverchkov-Smirnov '90). They are not unique, even for $\mathbb{R}$ (Hsu-Sturm '13, credited to Fitzpatrick).

Markovian maximal couplings are closely related to global reflections on M (Kuwada '07, '09).

Reflection on $\mathbb{R}^2$



On $\mathbb{R}^2$, unlike a more generic Riemannian manifold, the infinitesimal reflection along the geodesic from B_t to $\tilde{B}_t$ (on the left) extends to a global reflection of the space (on the right).

The reflection is through the x_2-axis, chosen to be the bisector between the starting points, which is known as soon as the starting points are chosen.

The Heisenberg group, $\mathbb{H}$

The most fundamental example of a sub-Riemannian manifold is the (3D) Heisenberg group:

The Heisenberg group, $\mathbb{H}$

The most fundamental example of a sub-Riemannian manifold is the (3D) Heisenberg group:

- topologically $\mathbb{R}^3$;
- $$X = \partial_x - \frac{y}{2}\partial_z \quad \text{and} \quad Y = \partial_y + \frac{x}{2}\partial_z,$$
  are an orthonormal set determines the horizontal distribution $\mathcal{H} = \text{span}\{X, Y\}$ and the inner product on it;
- $\mathbb{H}$ is a Lie group, with the group law
  $$(x, y, z) \cdot (x', y', z') = \left(x + x', y + y', z + z' + \frac{1}{2}\left(xy' - yx'\right)\right);$$
- if $Z = \partial_z$, then X, Y, Z are left-invariant and
  $$[X, Y] = Z \quad \text{and} \quad [X, Z] = [Y, Z] = 0.$$

Heisenberg Brownian motion

- $\pi : \mathbb{H} \to \mathbb{R}^2$ given by $\pi(x, y, z) = (x, y)$ is a submersion;
- horizontal curves are curves $(x(t), y(t))$ in xy-plane lifted by letting $z(t)$ be the swept area from the origin; their length is their $\mathbb{R}^2$-length; geodesics are (lifts of) circles;
- Heisenberg BM has generator $\frac{1}{2}\left(X^2 + Y^2\right) = \frac{1}{2}\Delta\mathcal{H}$;
- Heisenberg BM is planar BM (x_t, y_t) lifted by the associated Lévy area

$$z_t = z_0 + \frac{1}{2}\int_0^t x_s \, dy_s - \frac{1}{2}\int_0^t y_s \, dx_s.$$

Coupled Heisenberg BMs

Successful Markovian couplings of $\mathbb{H}$-BMs (from any starting points) is possible (Ben Arous-Cranston-Kendall '95, Kendall '07-), but…

Coupled Heisenberg BMs

Successful Markovian couplings of $\mathbb{H}$-BMs (from any starting points) is possible (Ben Arous-Cranston-Kendall ’95, Kendall ’07-), but… no Markovian coupling can be efficient, in particular for BMs started from the same vertical fiber (Banerjee-Gordina-Mariano ’18).

Banerjee-Gordina-Mariano ’18 do construct a non-Markovian (and non-co-adapted) efficient coupling from two points on the same fiber (and apply it). The construction is somewhat intricate, based on the Karhunen-Loève expansion and properties of the Brownian bridge, iterated on intervals of time of different lengths.

Coupled Heisenberg BMs

Successful Markovian couplings of $\mathbb{H}$-BMs (from any starting points) is possible (Ben Arous-Cranston-Kendall '95, Kendall '07-), but… no Markovian coupling can be efficient, in particular for BMs started from the same vertical fiber (Banerjee-Gordina-Mariano '18).

Banerjee-Gordina-Mariano '18 do construct a non-Markovian (and non-co-adapted) efficient coupling from two points on the same fiber (and apply it). The construction is somewhat intricate, based on the Karhunen-Loève expansion and properties of the Brownian bridge, iterated on intervals of time of different lengths.

Bénéfice ('23) gave similar constructions for $\mathrm{SL}(2, \mathbb{R})$ and $\mathrm{SU}(2)$.

(Baudoin-Gordina-Mariano '20 give a comparison of coupling techniques with Γ-calculus for some Kolmogorov-type diffusions.)

An elementary maximal coupling

W.l.o.g, two points on the same fiber are $(0,0,0)$ and $(0,0,2a)$ for $a > 0$.

Let $B_t = (x_t, y_t, z_t)$ be $\mathbb{H}$-BM from $(0,0,0)$, and σ_a the first time z_t hits a. Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection across the line of at angle θ through the origin.

An elementary maximal coupling

W.l.o.g, two points on the same fiber are $(0,0,0)$ and $(0,0,2a)$ for $a>0$.

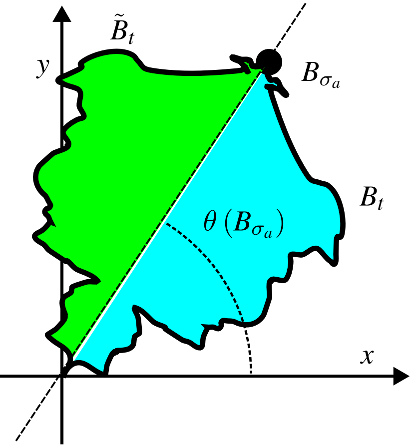
Let $B_t=(x_t,y_t,z_t)$ be $\mathbb{H}$-BM from $(0,0,0)$, and σ_a the first time z_t hits a. Let $R_\theta:\mathbb{R}^2\to\mathbb{R}^2$ be reflection across the line of at angle θ through the origin.

Then

$$\tilde{B}_t=(\tilde{x}_t,\tilde{y}_t,\tilde{z}_t)=\begin{cases}\left(R_{\theta(x_{\sigma_a},y_{\sigma_a})}\left(x_t,y_t\right),2a-z_t\right) & \text{for } t\le\sigma_a\\ (x_t,y_t,z_t) & \text{for } t>\sigma_a\end{cases}\qquad (*)$$

is a $\mathbb{H}$-BM from $(0,0,2a)$, which couples with B_t at time σ_a.

The picture



Blue area = Lévy area of $B_{\sigma_a} = a$
Green area = Lévy area of $\tilde{B}_{\sigma_a} = -a$

In the total space $\mathbb{H}$, we're inducing a (twisted) reflection across the plane $\{z = a\}$.

Properties

- The construction of the coupling is elementary, using no fine properties of paths.
- The coupling time τ is reduced (or "reduced") to a hitting time σ_a for a canonical 1D marginal process.

Properties

- The construction of the coupling is elementary, using no fine properties of paths.
- The coupling time τ is reduced (or "reduced") to a hitting time σ_a for a canonical 1D marginal process.
- Relative to the global isometry implied by $(*)$, z_t satisfies a reflection principle, just like 1D BM

$$\mathbb{P}\left(\sigma_a > t\right) = 1 - 2\mathbb{P}\left(z_t > a\right).$$

Properties

- The construction of the coupling is elementary, using no fine properties of paths.
- The coupling time τ is reduced (or "reduced") to a hitting time σ_a for a canonical 1D marginal process.
- Relative to the global isometry implied by $(*)$, z_t satisfies a reflection principle, just like 1D BM

$$\mathbb{P}\left(\sigma_a > t\right) = 1 - 2\mathbb{P}\left(z_t > a\right).$$

- From this, the coupling is immediately seen to be maximal (for such initial points).

Properties

- The construction of the coupling is elementary, using no fine properties of paths.
- The coupling time τ is reduced (or "reduced") to a hitting time σ_a for a canonical 1D marginal process.
- Relative to the global isometry implied by $(*)$, z_t satisfies a reflection principle, just like 1D BM

$$\mathbb{P}\left(\sigma_a > t\right) = 1 - 2\mathbb{P}\left(z_t > a\right).$$

- From this, the coupling is immediately seen to be maximal (for such initial points).
- The construction directly generalizes to $\mathrm{SL}(2, \mathbb{R})$, its universal cover $\widetilde{\mathrm{SL}(2, \mathbb{R})}$, and $\mathrm{SU}(2)$, and beyond to anisotropic Heisenberg groups of any dimension,...

Computing τ

It's more or less classical that z_t has density $\frac{1}{t} \operatorname{sech}\left(\pi \frac{z}{t}\right) dz$ (e.g. Lévy, Yor, Baudoin).

Computing τ

It's more or less classical that z_t has density $\frac{1}{t}\,\mathrm{sech}\left(\pi \frac{z}{t}\right) dz$ (e.g. Lévy, Yor, Baudoin).

We can integrate this to find

$$\mathbb{P}\left(\sigma_a > t\right) = \frac{4}{\pi} \arctan\left(\tanh\left(\frac{\pi}{2} \cdot \frac{a}{t}\right)\right)$$

which is then an exact expression for $\mathrm{dist}_{\mathrm{TV}}\left(\mathcal{L}\left(B_t\right), \mathcal{L}\left(\tilde{B}_t\right)\right)$. Note that

$$2\frac{a}{t} - \frac{\pi^2}{3}\left(\frac{a}{t}\right)^3 < \mathbb{P}\left(\sigma_a > t\right) < 2\frac{a}{t}.$$

Spatial dependence and gradients

Theorem (Luo-N. '24, also Banerjee-Gordina-Mariano '18, also it's immediate from the density of z_t...)

Let $P_t = e^{\frac{t}{2}\Delta_{\mathcal{H}}}$ be the heat semigroup on $\mathbb{H}$ and consider $f \in L^\infty(\mathbb{H})$. Then at any point $(x, y, z) \in \mathbb{H}$ and for any time $t > 0$, we have

$$|ZP_t f(x, y, z)| = |\nabla_V P_t f(x, y, z)| \leq \frac{1}{t} \|f\|_\infty.$$

Spatial dependence and gradients

Theorem (Luo-N. '24, also Banerjee-Gordina-Mariano '18, also it's immediate from the density of z_t...)

Let $P_t = e^{\frac{t}{2}\Delta_{\mathcal{H}}}$ be the heat semigroup on $\mathbb{H}$ and consider $f \in L^\infty(\mathbb{H})$. Then at any point $(x, y, z) \in \mathbb{H}$ and for any time $t > 0$, we have

$$|ZP_t f(x, y, z)| = |\nabla_V P_t f(x, y, z)| \leq \frac{1}{t} \|f\|_\infty.$$

This is stronger than a horizontal gradient bound or upper gradient bound (in the sense of metric measure spaces), for points on the same vertical fiber, because $\mathrm{dist}_{\mathrm{sR}}((0,0,0),(0,0,z)) = 2\sqrt{\pi}\sqrt{z}$ on $\mathbb{H}$.

Two-stage coupling

Two general points of $\mathbb{H}$ can be taken to be $(0,0,0)$ and $(h,0,v)$. Banerjee-Gordina-Mariano '18 construct a two-stage coupling:

- First, couple the xy-marginals by reflection coupling on $\mathbb{R}^2$; let the z-processes come along for the ride;
- the xy-marginals will couple at some random time τ_1, and the z-marginals will then differ by some random a;
- control the joint distribution of τ_1 and a;

Two-stage coupling

Two general points of $\mathbb{H}$ can be taken to be $(0,0,0)$ and $(h,0,v)$. Banerjee-Gordina-Mariano '18 construct a two-stage coupling:

- First, couple the *xy*-marginals by reflection coupling on $\mathbb{R}^2$; let the *z*-processes come along for the ride;
- the *xy*-marginals will couple at some random time τ_1, and the *z*-marginals will then differ by some random a;
- control the joint distribution of τ_1 and a;
- now run a vertical coupling for points on the same fiber (I claim, as above) from the random a;
- when this second stage succeeds, the particles have coupled in the total space;
- the total coupling time is the combined time of both stages, which we can estimate (BGM '18):

Two-stage coupling

Two general points of $\mathbb{H}$ can be taken to be $(0,0,0)$ and $(h,0,v)$. Banerjee-Gordina-Mariano '18 construct a two-stage coupling:

- First, couple the xy-marginals by reflection coupling on $\mathbb{R}^2$; let the z-processes come along for the ride;
- the xy-marginals will couple at some random time τ_1, and the z-marginals will then differ by some random a;
- control the joint distribution of τ_1 and a;
- now run a vertical coupling for points on the same fiber (I claim, as above) from the random a;
- when this second stage succeeds, the particles have coupled in the total space;
- the total coupling time is the combined time of both stages, which we can estimate (BGM '18):

$$\mathbb{P}(\tau > t) \leq C\left(\frac{h}{\sqrt{t}} + \frac{|v|}{t}\right) \quad \text{for } t > \max\{h^2, 2|v|\}.$$

$\widetilde{\mathrm{SL}(2,\mathbb{R})}$

From our perspective, $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ is $\mathbb{H}$ with $\mathbb{R}^2$ replaced by the hyperbolic plane and z by the hyperbolic swept area. Using (normal) polar coordinates on the hyperbolic plane, the BM is given by

$\widetilde{\mathrm{SL}(2,\mathbb{R})}$

From our perspective, $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ is $\mathbb{H}$ with $\mathbb{R}^2$ replaced by the hyperbolic plane and z by the hyperbolic swept area. Using (normal) polar coordinates on the hyperbolic plane, the BM is given by

$$\begin{aligned} dr_t &= dW_t^1 + \frac{1}{2}\coth(r_t)\,dt \\ d\theta_t &= \frac{1}{\sinh(r_t)}\,dW_t^2 \\ dz_t &= \tanh\left(\frac{r_t}{2}\right)\,dW_t^2. \end{aligned}$$

Further, z_t can be written as

$$z_t = z_0 + W_{\int_0^t \tanh^2\left(\frac{r_s}{2}\right)ds}$$

where W_t is a 1D Brownian motion independent of r_t (and r_t is the radial process on the hyperbolic plane).

Vertical coupling on $\widetilde{SL(2,\mathbb{R})}$

The coupling from points $(0,0,0)$ and $(0,0,2a)$ on the same fiber proceeds the same way, with the same picture. But the coupling time needs different estimates. No density for z_t this time, and we need to go a bit beyond Baudoin-Demni-Wang ’23+

Vertical coupling on $\widetilde{\mathrm{SL}(2,\mathbb{R})}$

The coupling from points $(0,0,0)$ and $(0,0,2a)$ on the same fiber proceeds the same way, with the same picture. But the coupling time needs different estimates. No density for z_t this time, and we need to go a bit beyond Baudoin-Demni-Wang '23+

Theorem (Luo-N. '24)

There exist $c, T_0 > 0$, independent of a, such that, for $t \geq T_0$,

$$P\left(\sigma_a > t\right) \leq \left(\frac{1}{\sqrt{2\pi}} + 2c\right) \frac{a}{\sqrt{t}} \quad \textit{for all } a > 0.$$

Moreover, this bound is sharp, in the sense that, for any $a > 0$ and $\varepsilon > 0$, there exists $T' > 0$ (which may depend on a and ε), such that

$$P\left(\sigma_a > t\right) \geq \left(\frac{1}{\sqrt{2\pi}} - \varepsilon\right) \frac{a}{\sqrt{t}} \quad \textit{for } t \geq T'.$$

$\mathrm{SL}(2,\mathbb{R})$

From this perspective, $\mathrm{SL}(2,\mathbb{R})$ is $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ with z taken modulo 4π, and $\mathrm{SL}(2,\mathbb{R})$ is topologically $\mathbb{R}^2 \times \mathbb{S}^1$.

So now

$$\sigma_a = \inf\{t > 0 : z_t = a \text{ or } 2\pi - a\}.$$

$\mathrm{SL}(2,\mathbb{R})$

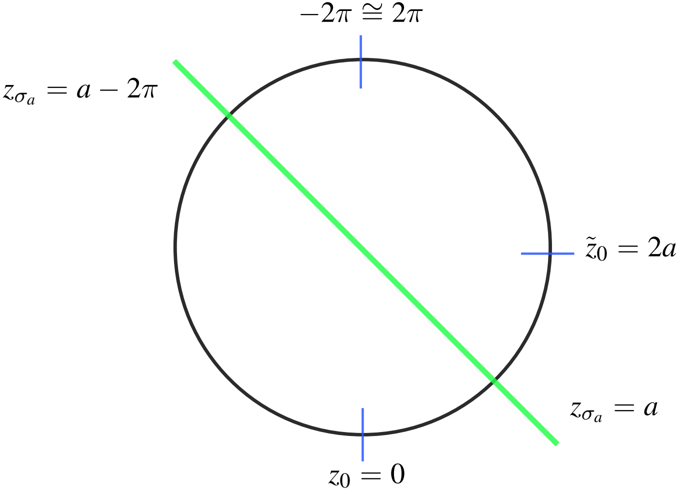
From this perspective, $\mathrm{SL}(2,\mathbb{R})$ is $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ with z taken modulo 4π, and $\mathrm{SL}(2,\mathbb{R})$ is topologically $\mathbb{R}^2 \times \mathbb{S}^1$.

So now

$$\sigma_a = \inf\{t > 0 : z_t = a \text{ or } 2\pi - a\}.$$

But the reflection still works almost the same way, the coupling is maximal, etc.

The vertical fiber picture



σ_a is when the particle from the bottom hits the green line. Reflection in the hyperbolic plane then induces reflection through the green line.

Estimating σ_a for circular vertical fibers

It's easy to see there is some exponential tail bound for the hitting time:
There exist constants $C > 0$, $c > 0$, and $T_0 > 0$ such that, for any $a \in \mathbb{S}^1$,

$$\mathbb{P}\left(\sigma_a > t\right) \leq Ce^{-ct} \quad \text{for all } t > T_0.$$

Estimating σ_a for circular vertical fibers

It's easy to see there is some exponential tail bound for the hitting time:
There exist constants $C > 0$, $c > 0$, and $T_0 > 0$ such that, for any $a \in \mathbb{S}^1$,

$$\mathbb{P}\left(\sigma_a > t\right) \leq Ce^{-ct} \quad \text{for all } t > T_0.$$

But computing/bounded c would require much finer control of the area process…

$SU(2)$

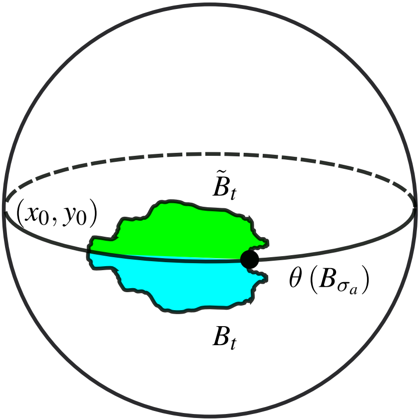
From this perspective, BM on $SU(2)$ is BM on the sphere, with the associated stochastic area, necessarily taken modulo 4π. (And this is nicely compatible with the Hopf fibration.)

$SU(2)$

From this perspective, BM on $SU(2)$ is BM on the sphere, with the associated stochastic area, necessarily taken modulo 4π. (And this is nicely compatible with the Hopf fibration.)

Once again, the construction of the maximal vertical coupling and its qualitative properties are the same as before.

The picture



Blue area = Lévy area of $B_{\sigma_a} = a$
Green area = Lévy area of $\tilde{B}_{\sigma_a} = -a$

Exponential decay of coupling time

We again get a non-explicit exponential bound on the vertical hitting time,

$$\mathbb{P}(\sigma_a > t) \leq Ce^{-ct} \quad \text{for all } t > T_0.$$

You can (probably) extract c from the Fourier analysis of Baudoin-Bonnefont '09, but stochastically, it's not so nice.

Exponential decay of coupling time

We again get a non-explicit exponential bound on the vertical hitting time,

$$\mathbb{P}(\sigma_a > t) \leq Ce^{-ct} \quad \text{for all } t > T_0.$$

You can (probably) extract c from the Fourier analysis of Baudoin-Bonnefont ’09, but stochastically, it’s not so nice.

Also, Bénéfice (’23) considers the two-stage coupling, and gives the corresponding horizontal gradient bounds.

Coda: Parallel-type couplings

On a Riemannian manifold with Ricci curvature bounded below by k, there is a Markov coupling with

$$\operatorname{dist}\left(B_t, \tilde{B}_t\right) \leq e^{-K/2} \operatorname{dist}\left(x_0, \tilde{x}_0\right) \text{ for all } t \geq 0.$$

Coda: Parallel-type couplings

On a Riemannian manifold with Ricci curvature bounded below by k, there is a Markov coupling with

$$\operatorname{dist}\left(B_t, \tilde{B}_t\right) \leq e^{-K/2} \operatorname{dist}\left(x_0, \tilde{x}_0\right) \text{ for all } t \geq 0.$$

For $\mathbb{H}$, an abstract existence result: (Kuwada '10) shows that there is a constant C such that, for any given $t > 0$, there exists a coupling such that

$$\operatorname{dist}\left(B_t, \tilde{B}_t\right) \leq C \operatorname{dist}\left(x_0, \tilde{x}_0\right) \text{ for that } t.$$

Coda: Non-Markovian parallel coupling

Bonnefont-Juillet ’20 prove that no Markovian coupling on $\mathbb{H}$ can stay bounded distance. So even for $\mathbb{H}$, it’s not clear what this coupling “looks like.”

Coda: Non-Markovian parallel coupling

Bonnefont-Juillet '20 prove that no Markovian coupling on $\mathbb{H}$ can stay bounded distance. So even for $\mathbb{H}$, it's not clear what this coupling "looks like."

Driver-Melcher '05 show that $C \geq \sqrt{2}$. It is conjectured that $C = \sqrt{2}$. To the best of my knowledge, nothing else about C is known.