Gaussian rough paths, transportation–cost inequalities and multilevel Monte Carlo

Sebastian Riedel

Technische Universität Berlin

IPAM Workshop "Rough Paths: Theory and Applications",
January 29, 2014

Outline

Gaussian processes and rough paths

Applications

- Concentration of measure on paths spaces
- Multilevel and rough paths

Outline

Goal: Define, solve and analyze solution $Y = I_f(X, \xi)$ to

$$dY_t = f_0(Y_t)\,dt + \sum_{i=1}^{d} f_i(Y_t)\,dX_t^i; \quad Y_0 = \xi \in \mathbb{R}^m$$

where is $X = (X^1, \ldots, X^d)$ is continuous, but not a semimartingale.

Goal: Define, solve and analyze solution $Y = I_f(X, \xi)$ to

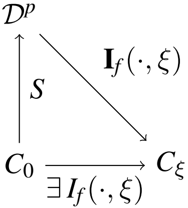
$$dY_t = f_0(Y_t)\,dt + \sum_{i=1}^{d} f_i(Y_t)\,dX_t^i; \quad Y_0 = \xi \in \mathbb{R}^m$$

where is $X = (X^1, \ldots, X^d)$ is continuous, but not a semimartingale.

Motivation

- Models with correlations in the noise (eg. fractional Bm)
- nonlinear SPDE with 1d space variable (eg. Burger's equation, cf. [Hairer; CPAM '11])

One possible way: Use rough paths factorization



where

$$\mathcal{D}^p = \left\{ x, \int_{\Delta^2} dx \otimes dx, \ldots, \int_{\Delta^{[p]}} dx \otimes \cdots \otimes dx \right\}$$

$$S(x) = \left(x, \int_{\Delta^2} dx \otimes dx, \ldots, \int_{\Delta^{[p]}} dx \otimes \cdots \otimes dx \right)$$

Q: Given some process X, what are “natural” candidates for iterated integrals?

Q: Given some process X, what are "natural" candidates for iterated integrals?

- Can show: If B^ε "nice" pathwise smooth approximation of Bm, then iterated integrals converge in rough paths topology:

$$\int_{\Delta^n} dB^\varepsilon \otimes \cdots \otimes dB^\varepsilon \to \int_{\Delta^n} \circ dB \otimes \cdots \otimes \circ dB \qquad (1)$$

  for $\varepsilon \to 0$, $\rightsquigarrow$ "natural" rough path.

Q: Given some process X, what are “natural” candidates for iterated integrals?

- Can show: If B^{ε} “nice” pathwise smooth approximation of Bm, then iterated integrals converge in rough paths topology:

$$\int_{\Delta^n} dB^{\varepsilon} \otimes \cdots \otimes dB^{\varepsilon} \to \int_{\Delta^n} \circ dB \otimes \cdots \otimes \circ dB \qquad (1)$$

  for $\varepsilon \to 0$, $\rightsquigarrow$ “natural” rough path.
- Generalization of (1) to other processes?

For **Gaussian processes** X, the following holds:

Theorem (Friz, Victoir; '10)

If covariance

$$(s,t) \mapsto R(s,t) = \mathbb{E}(X_s \otimes X_t) \in \mathbb{R}^{d \times d}$$

is sufficiently smooth, i.e. is diagonal and has finite ρ-variation for some $\rho < 2$, then pathwise approximations of iterated integrals converge in rough paths space.

What is 2D ρ-variation?

- $R\colon [0,T]^2 \to \mathbb{R}$

What is 2D ρ-variation?

- $R\colon [0,T]^2 \to \mathbb{R}$
- for rectangle $A = [s,t] \times [u,v] \subset [0,T]^2$, define the area increment

$$R(A) := R(t,v) - R(t,u) - R(s,v) + R(s,u)$$

What is 2D ρ-variation?

- $R\colon [0,T]^2 \to \mathbb{R}$
- for rectangle $A = [s,t] \times [u,v] \subset [0,T]^2$, define the area increment

$$R(A) := R(t,v) - R(t,u) - R(s,v) + R(s,u)$$

- For $[s,t] \subseteq [0,T]$, set:

$$\Pi([s,t]^2) := \left\{ \text{all rectangular partitions } D \text{ of } [s,t]^2 \right\}$$

What is 2D ρ-variation?

- $R\colon [0,T]^2 \to \mathbb{R}$
- for rectangle $A = [s,t] \times [u,v] \subset [0,T]^2$, define the area increment

$$R(A) := R(t,v) - R(t,u) - R(s,v) + R(s,u)$$

- For $[s,t] \subseteq [0,T]$, set:

$$\Pi([s,t]^2) := \left\{ \text{all rectangular partitions } D \text{ of } [s,t]^2 \right\}$$

- Define:

$$V_\rho(R;[s,t]) := \sup_{D \in \Pi([s,t]^2)} \left(\sum_{A \in D} |R(A)|^\rho \right)^{\frac{1}{\rho}}$$

Want to make sense of integrals of the form

$$\int_{[0,T]^2} R(s,t)\, dR(s,t) \text{ “} = \int_{[0,T]^2} R(s,t)\, \partial_s \partial_t R(s,t)\, ds\, dt \text{ ”}$$

Want to make sense of integrals of the form

$$\int_{[0,T]^2} R(s,t)\, dR(s,t) \text{ “} = \int_{[0,T]^2} R(s,t)\, \partial_s \partial_t R(s,t)\, ds\, dt \text{ ”}$$

- if $\partial_s \partial_t R$ is “nice” distribution, integral should exist

Want to make sense of integrals of the form

$$\int_{[0,T]^2} R(s,t)\, dR(s,t) \text{ “} = \int_{[0,T]^2} R(s,t)\, \partial_s \partial_t R(s,t)\, ds\, dt \text{ ”}$$

- if $\partial_s \partial_t R$ is “nice” distribution, integral should exist
- example: Brownian motion

$$\mathbb{E} \left| \int_0^T \tilde{B}_s\, dB_s \right|^2 = \int_{[0,T]^2} (s \wedge t)\, \delta_{s,t}\, ds\, dt = \int_0^T s\, ds$$

Want to make sense of integrals of the form

$$\int_{[0,T]^2} R(s,t)\, dR(s,t) \text{ “} = \int_{[0,T]^2} R(s,t)\, \partial_s \partial_t R(s,t)\, ds\, dt \text{ ”}$$

- if $\partial_s \partial_t R$ is “nice” distribution, integral should exist
- example: Brownian motion

$$\mathbb{E} \left| \int_0^T \tilde{B}_s\, dB_s \right|^2 = \int_{[0,T]^2} (s \wedge t)\, \delta_{s,t}\, ds\, dt = \int_0^T s\, ds$$

- typically, R “smooth” outside the diagonal $D = \{(s,t) \,:\, s = t\}$

$$\sigma^2(s,t) = \mathrm{Var}(X_t - X_s) = R([s,t]^2)$$

$$\sigma^2(s,t) = \mathrm{Var}(X_t - X_s) = R([s,t]^2)$$

Theorem (Friz, Gess, Gulisashvili, R.)

Assume: $\mu = \partial_s \partial_t R(s,t)$ *Radon measure on* $(0,T)^2 \setminus D$ *with decomposition* $\mu = \mu_+ - \mu_-$.

$$\sigma^2(s,t) = \mathrm{Var}(X_t - X_s) = R([s,t]^2)$$

Theorem (Friz, Gess, Gulisashvili, R.)

Assume: $\mu = \partial_s \partial_t R(s,t)$ *Radon measure on* $(0,T)^2 \setminus D$ *with decomposition* $\mu = \mu_+ - \mu_-$.

1. *If* μ_- *has finite mass, then* R *has finite (2d)* 1*-variation*

$$\sigma^2(s,t) = \mathrm{Var}(X_t - X_s) = R([s,t]^2)$$

Theorem (Friz, Gess, Gulisashvili, R.)

Assume: $\mu = \partial_s \partial_t R(s,t)$ *Radon measure on* $(0,T)^2 \setminus D$ *with decomposition* $\mu = \mu_+ - \mu_-$.

1. *If* μ_- *has finite mass, then* R *has finite (2d)* 1*-variation*
2. *If*
   - (i) μ_+ *has finite mass,*
   - (ii) $R([u,v] \times [s,t]) \geq 0$ *for all* $[u,v] \subseteq [s,t]$,
   - (iii) σ^2 *has finite (1d)* ρ*-variation,* $\rho \in [1,\infty)$,

   then R *has finite (2d)* ρ*-variation.*

$$\sigma^2(s,t) = \mathrm{Var}(X_t - X_s) = R([s,t]^2)$$

Theorem (Friz, Gess, Gulisashvili, R.)

Assume: $\mu = \partial_s \partial_t R(s,t)$ *Radon measure on* $(0,T)^2 \setminus D$ *with decomposition* $\mu = \mu_+ - \mu_-$.

1. *If* μ_- *has finite mass, then* R *has finite (2d)* 1*-variation*
2. *If*
   (i) μ_+ *has finite mass,*
   (ii) $R([u,v] \times [s,t]) \geq 0$ *for all* $[u,v] \subseteq [s,t]$,
   (iii) σ^2 *has finite (1d)* ρ*-variation,* $\rho \in [1,\infty)$,

   then R *has finite (2d)* ρ*-variation.*

In particular, for $\rho \in [1,2)$, X *can be lifted to process with trajectories in rough paths space (in the sense of Friz–Victoir).*

$$\sigma^2(s,t) = \mathrm{Var}(X_t - X_s) = R([s,t]^2)$$

Theorem (Friz, Gess, Gulisashvili, R.)

Assume: $\mu = \partial_s \partial_t R(s,t)$ Radon measure on $(0,T)^2 \setminus D$ with decomposition $\mu = \mu_+ - \mu_-$.

1. *If μ_- has finite mass, then R has finite (2d) 1-variation*
2. *If*
   - (i) *μ_+ has finite mass,*
   - (ii) *$R([u,v] \times [s,t]) \geq 0$ for all $[u,v] \subseteq [s,t]$,*
   - (iii) *σ^2 has finite (1d) ρ-variation, $\rho \in [1,\infty)$,*

   then R has finite (2d) ρ-variation.

In particular, for $\rho \in [1,2)$, X can be lifted to process with trajectories in rough paths space (in the sense of Friz–Victoir).

- Condition (iii) first appeared in [Jain, Monrad; '83] in the context of sample path regularity

Cameron–Martin embedding

Theorem (Friz, Gess, Gulisashvili, R.)

Under the same assumptions, we have the continuous embedding

$$\mathcal{H} \hookrightarrow C^{q-var}; \qquad q = \frac{1}{\frac{1}{2\rho} + \frac{1}{2}} \quad (\leq \rho)$$

for the Cameron–Martin space $\mathcal{H}$.

Cameron–Martin embedding

Theorem (Friz, Gess, Gulisashvili, R.)

Under the same assumptions, we have the continuous embedding

$$\mathcal{H} \hookrightarrow C^{q-var}; \qquad q = \frac{1}{\frac{1}{2\rho} + \frac{1}{2}} \quad (\leq \rho)$$

for the Cameron–Martin space $\mathcal{H}$.

Recall: For fBm,

$$\mathcal{H}^H \hookrightarrow C^{q-\text{var}}; \qquad q = \frac{1}{H + \frac{1}{2}}$$

If X has stationary increments, $\sigma^2(s,t) =: \sigma^2(t-s)$ and we obtain

Corollary

1. *If σ^2 is convex, R has finite (2d) 1-variation*
2. *If*
   (i) *σ^2 is concave,*
   (ii) *σ^2 is increasing on $[0,T]$,*
   (iii) *σ^2 has finite (1d) ρ-variation (or is $1/\rho$ Hölder continuous),*

   then R has finite (2d) ρ-variation.

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K}\left((s^{2H} + t^{2H})^K - |t-s|^{2HK}\right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K}\left((s^{2H} + t^{2H})^K - |t-s|^{2HK}\right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K} \left((s^{2H} + t^{2H})^K - |t-s|^{2HK} \right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm
- self–similar process with index HK

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K}\left((s^{2H} + t^{2H})^K - |t-s|^{2HK}\right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm
- self–similar process with index HK
- no stationary increments

Example I: bifractional Bm

centered Gaussian process $B^{H,K}\colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K}\left((s^{2H} + t^{2H})^K - |t-s|^{2HK} \right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm
- self–similar process with index HK
- no stationary increments
- no semimartingale

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K} \left((s^{2H} + t^{2H})^K - |t-s|^{2HK} \right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm
- self–similar process with index HK
- no stationary increments
- no semimartingale
- no Markov process

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K}\left((s^{2H} + t^{2H})^K - |t-s|^{2HK}\right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm
- self–similar process with index HK
- no stationary increments
- no semimartingale
- no Markov process
- no Volterra process

Example I: bifractional Bm

centered Gaussian process $B^{H,K} \colon [0,T] \to \mathbb{R}$ with covariance

$$R(s,t) = \frac{1}{2^K}\left((s^{2H} + t^{2H})^K - |t-s|^{2HK} \right),$$

$H \in (0,1)$, $K \in (0,1]$, is called **bifractional Brownian motion**

- $K = 1$: fBm; $K = 1$, $H = 1/2$: Bm
- self–similar process with index HK
- no stationary increments
- no semimartingale
- no Markov process
- no Volterra process

Can show: covariance of bfBm has finite ρ-variation for $\rho = \frac{1}{2HK}$, CM paths have finite q-variation, $q = (HK + 1/2)^{-1}$.

Example II: fractional SHE

Ψ solution of the fractional stochastic heat equation

$$d\Psi_t = (-(-\Delta)^\alpha)\Psi_t \, dt + dW_t; \quad \Psi_0 = 0$$

where

- dW space-time white noise
- spatial variable $x \in [0, 2\pi]$, time variable $t \in [0, T]$
- $(-(-\Delta)^\alpha)$ fractional Laplacian with periodic (**P**), Dirichlet (**D**) or Neumann (**N**) ∂-conditions ($\alpha = 1 \rightsquigarrow$ usual Laplacian Δ)

For fixed t, $x \mapsto \Psi_t(x)$ is Gaussian process; can be written as random Fourier series.

Example II: fractional SHE

Ψ solution of the fractional stochastic heat equation

$$d\Psi_t = (-(-\Delta)^{\alpha})\Psi_t \, dt + dW_t; \quad \Psi_0 = 0$$

where

- dW space-time white noise
- spatial variable $x \in [0, 2\pi]$, time variable $t \in [0, T]$
- $(-(-\Delta)^{\alpha})$ fractional Laplacian with periodic (**P**), Dirichlet (**D**) or Neumann (**N**) ∂-conditions ($\alpha = 1 \rightsquigarrow$ usual Laplacian Δ)

For fixed t, $x \mapsto \Psi_t(x)$ is Gaussian process; can be written as random Fourier series.

Can show: Covariance of $x \mapsto \Psi_t(x)$ has finite ρ-variation for $\rho = \frac{1}{2\alpha - 1}$.

Outline

Measure concentration on metric spaces

- (E, d) metric space, $\mathcal{B}(E)$ Borel sets

Measure concentration on metric spaces

- (E, d) metric space, $\mathcal{B}(E)$ Borel sets
- if $A \subset E$,

$$A^r = \{x \in E \,:\, \exists \bar{x} \in A \text{ s.t. } d(x, \bar{x}) \leq r\}$$

Measure concentration on metric spaces

- (E, d) metric space, $\mathcal{B}(E)$ Borel sets
- if $A \subset E$,

$$A^r = \{x \in E \,:\, \exists \bar{x} \in A \text{ s.t. } d(x, \bar{x}) \leq r\}$$

- $\mu \in P(E)$ has Gaussian concentration property if

$$\mu(A^r) \geq 1 - be^{-a(r-r_0)^2}$$

for all $r \geq r_0$, $A \in \mathcal{B}(E)$ s.t. $\mu(A) \geq 1/2$.

Functional approach?

Q: "Functional approach" to concentration of measure phenomenon?

Functional approach?

Q: "Functional approach" to concentration of measure phenomenon?

- Recall: If $\mu, \nu \in P(E)$, $p \geq 1$,

$$\mathcal{W}_p(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left(\int_{E \times E} d(x,y)^p \, d\pi(x,y) \right)^{\frac{1}{p}}$$

Wasserstein p-metric

Functional approach?

Q: "Functional approach" to concentration of measure phenomenon?

- Recall: If $\mu, \nu \in P(E)$, $p \geq 1$,

$$\mathcal{W}_p(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left(\int_{E\times E} d(x,y)^p \, d\pi(x,y) \right)^{\frac{1}{p}}$$

Wasserstein p-metric

-

$$H(\nu \,|\, \mu) = \begin{cases} \int_E \log\left(\frac{d\nu}{d\mu}\right) d\nu & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise} \end{cases}$$

relative entropy of ν w.r.t. μ

Transportation–cost inequalities

p-transportation–cost inequality holds for $\mu \in P(E)$ iff

$$\mathcal{W}_p(\nu, \mu) \leq \sqrt{CH(\nu \mid \mu)} \qquad \text{for all } \nu \in P(E) \qquad (T_p(C))$$

Relations to Gaussian concentration

Theorem (Marton; '96)

$T_1(C)$ implies Gaussian concentration property.

Relations to Gaussian concentration

Theorem (Marton; ’96)

$T_1(C)$ implies Gaussian concentration property.

Theorem (Djellout, Guillin, Wu; ’04)

Gaussian concentration implies $T_1(C)$.

Relations to Gaussian concentration

Theorem (Marton; '96)

$T_1(C)$ implies Gaussian concentration property.

Theorem (Djellout, Guillin, Wu; '04)

Gaussian concentration implies $T_1(C)$.

Theorem (Gozlan; '09)

$T_2(C)$ is equivalent to dimension free Gaussian concentration property.

Gaussian context

$(E, \mathcal{H}, \gamma)$ Gaussian Banach space,

$$d_{\mathcal{H}}(x,y) = \begin{cases} \sqrt{\langle x-y, x-y\rangle_{\mathcal{H}}} & \text{if } x-y \in \mathcal{H} \\ +\infty & \text{otherwise} \end{cases}$$

Gaussian context

$(E, \mathcal{H}, \gamma)$ Gaussian Banach space,

$$d_{\mathcal{H}}(x,y) = \begin{cases} \sqrt{\langle x-y, x-y\rangle_{\mathcal{H}}} & \text{if } x-y \in \mathcal{H} \\ +\infty & \text{otherwise} \end{cases}$$

Theorem

$$\inf_{\pi \in \Pi(\nu,\gamma)} \int_{E\times E} d_{\mathcal{H}}(x,y)^2 \, d\pi(x,y) \leq 2H(\nu \,|\, \gamma)$$

for all $\nu \in P(E)$.

Possible proofs

- finite dimensional case: [Talagrand; ’96] (via tensorization)
- infinite dimensional case: [Feyel, Üstünel; ’03] (via Girsanov transform)
- **our proof:** combines large deviation approach from Gozlan with the Borell-Sudakov–Tsirelson inequality

Transportation inequalities stable under Lipschitz maps

Key observation:

Lemma (Djellout, Guillin, Wu; '04)

- $\Psi\colon (E, d_E) \to (F, d_F)$ *Lipschitz with constant* L
- $T_p(C)$ *holds for some* $\mu \in P(E)$

Then $T_p(L^2 C)$ *holds for* $\tilde{\mu}$ *where* $\tilde{\mu} = \mu \circ \Psi^{-1}$.

The additive noise case I

Consider

$$dY_t = dX_t + b(Y_t)\,dt; \quad Y_0 = \xi \in \mathbb{R}^d,$$

$X \colon [0,T] \to \mathbb{R}^d$ continuous Gaussian process,

$$I_b(\cdot,\xi) \colon C_0([0,T],\mathbb{R}^d) \to C_\xi([0,T],\mathbb{R}^d)$$

solution map.

The additive noise case I

Consider

$$dY_t = dX_t + b(Y_t)\,dt; \quad Y_0 = \xi \in \mathbb{R}^d,$$

$X\colon [0,T] \to \mathbb{R}^d$ continuous Gaussian process,

$$I_b(\cdot,\xi)\colon C_0([0,T],\mathbb{R}^d) \to C_\xi([0,T],\mathbb{R}^d)$$

solution map.
Assume continuous embedding of $\mathcal{H}$ into fractional Sobolev space $W^{p,\delta}$:

$$\mathcal{H} \hookrightarrow W^{\delta,p}$$

for some $\delta \in (0,1]$ and $p \in (1/\delta, \infty)$.

The additive noise case II

Corollary

Let $\mu \in P(C_\xi)$ *be the law of* Y. *Then*

$$\inf_{\pi \in \Pi(\nu,\mu)} \int_{C_\xi \times C_\xi} d_{W^{\delta,p}}(x,y)^2 \, d\pi(x,y) \leq CH(\nu \,|\, \mu)$$

for all $\nu \in P(C_\xi)$.

The additive noise case II

Corollary

Let $\mu \in P(C_\xi)$ be the law of Y. Then

$$\inf_{\pi \in \Pi(\nu,\mu)} \int_{C_\xi \times C_\xi} d_{W^{\delta,p}}(x,y)^2 \, d\pi(x,y) \leq CH(\nu \,|\, \mu)$$

for all $\nu \in P(C_\xi)$.

Proof. Map $I_b(\cdot,\xi)$ Lipschitz w.r.t. $d_{W^{\delta,p}}$ (Gronwall).

The additive noise case II

Corollary

Let $\mu \in P(C_\xi)$ *be the law of* Y. *Then*

$$\inf_{\pi \in \Pi(\nu,\mu)} \int_{C_\xi \times C_\xi} d_{W^{\delta,p}}(x,y)^2 \, d\pi(x,y) \leq CH(\nu \mid \mu)$$

for all $\nu \in P(C_\xi)$.

Proof. Map $I_b(\cdot,\xi)$ Lipschitz w.r.t. $d_{W^{\delta,p}}$ (Gronwall).

Examples

The additive noise case II

Corollary

Let $\mu \in P(C_\xi)$ *be the law of* Y. *Then*

$$\inf_{\pi \in \Pi(\nu,\mu)} \int_{C_\xi \times C_\xi} d_{W^{\delta,p}}(x,y)^2 \, d\pi(x,y) \leq CH(\nu \mid \mu)$$

for all $\nu \in P(C_\xi)$.

Proof. Map $I_b(\cdot,\xi)$ Lipschitz w.r.t. $d_{W^{\delta,p}}$ (Gronwall).

Examples

- Brownian motion: $\mathcal{H} = W^{1,2}$

The additive noise case II

Corollary

Let $\mu \in P(C_\xi)$ be the law of Y. Then

$$\inf_{\pi \in \Pi(\nu,\mu)} \int_{C_\xi \times C_\xi} d_{W^{\delta,p}}(x,y)^2 \, d\pi(x,y) \leq CH(\nu \mid \mu)$$

for all $\nu \in P(C_\xi)$.

Proof. Map $I_b(\cdot,\xi)$ Lipschitz w.r.t. $d_{W^{\delta,p}}$ (Gronwall).

Examples

- Brownian motion: $\mathcal{H} = W^{1,2}$
- fractional Bm: $\mathcal{H} \hookrightarrow W^{\delta,2}$ where $\delta < H + 1/2$

The multiplicative noise case I

Consider

$$dY_t = \sum_{i=1}^{d} f_i(Y_t) \circ dX_t^i; \quad Y_0 = \xi \in \mathbb{R}^m,$$

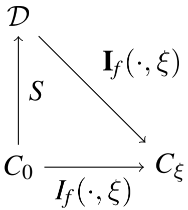
$X \colon [0,T] \to \mathbb{R}^d$ continuous Gaussian process with law γ and trajectories of finite p-variation.

The multiplicative noise case I

Consider

$$dY_t = \sum_{i=1}^{d} f_i(Y_t) \circ dX_t^i; \quad Y_0 = \xi \in \mathbb{R}^m,$$

$X \colon [0,T] \to \mathbb{R}^d$ continuous Gaussian process with law γ and trajectories of finite p-variation.
Solution defined via



The multiplicative noise case II

Assume

$$\mathcal{H} \hookrightarrow C^{q-\mathrm{var}}$$

with $\frac{1}{p} + \frac{1}{q} > 1$ ("complementary Young regularity").

The multiplicative noise case II

Assume

$$\mathcal{H} \hookrightarrow C^{q-\text{var}}$$

with $\frac{1}{p} + \frac{1}{q} > 1$ ("complementary Young regularity").

Lemma

$$\|I_f(x,\xi) - I_f(y,\xi)\|_\infty \leq L(x,y)(d_{\mathcal{H}}(x,y) \vee d_{\mathcal{H}}(x,y)^q)$$

where L has every moment for every $\pi \in \Pi(\nu,\gamma)$ with $\nu \ll \gamma$.

The multiplicative noise case III

Theorem

Let μ *be law of* Y *on* $C_\xi([0,T],\mathbb{R}^m)$*. Assume* $q=1$*. Then for every* $\varepsilon>0$*,*

$$\inf_{\pi\in\Pi(\nu,\mu)}\left(\int_{C_\xi\times C_\xi}\|x-y\|_\infty^{2-\varepsilon}\,d\pi(x,y)\right)^{\frac{1}{2-\varepsilon}}\leq\sqrt{CH(\nu\,|\,\mu)}$$

for all $\nu\in P(C_\xi)$*.*

The multiplicative noise case III

Theorem

Let μ be law of Y on $C_\xi([0,T],\mathbb{R}^m)$. Assume $q=1$. Then for every $\varepsilon>0$,

$$\inf_{\pi\in\Pi(\nu,\mu)}\left(\int_{C_\xi\times C_\xi}\|x-y\|_\infty^{2-\varepsilon}\,d\pi(x,y)\right)^{\frac{1}{2-\varepsilon}}\leq\sqrt{CH(\nu\,|\,\mu)}$$

for all $\nu\in P(C_\xi)$.

Examples:

- bifractional Bm $B^{H,K}$ with $HK=\frac{1}{2}$
- solution $x\mapsto\Psi_t(x)$ to the heat equation ($\alpha=1$)

Theorem

V linear polish, $\mathcal{B} \subseteq V$ normed space,

$$d_{\mathcal{B}}(x,y) = \begin{cases} \|x-y\|_{\mathcal{B}} & \text{if } x-y \in \mathcal{B} \\ +\infty & \text{otherwise.} \end{cases}$$

Let $\mu \in P(V)$ and assume

$$\inf_{\pi \in \Pi(\nu,\mu)} \left(\int_{V \times V} d_{\mathcal{B}}(x,y)^p \, d\pi(x,y) \right)^{\frac{1}{p}} \leq \sqrt{H(\nu \,|\, \mu)} \quad \forall \nu \in P(V).$$

Theorem

V linear polish, $\mathcal{B} \subseteq V$ normed space,

$$d_{\mathcal{B}}(x,y) = \begin{cases} \|x-y\|_{\mathcal{B}} & \text{if } x-y \in \mathcal{B} \\ +\infty & \text{otherwise.} \end{cases}$$

Let $\mu \in P(V)$ and assume

$$\inf_{\pi \in \Pi(\nu,\mu)} \left(\int_{V \times V} d_{\mathcal{B}}(x,y)^p \, d\pi(x,y) \right)^{\frac{1}{p}} \leq \sqrt{H(\nu \,|\, \mu)} \quad \forall \nu \in P(V).$$

Let (E,d) metric space, $f \colon V \to (E,d)$ such that μ - a.s. for all $h \in \mathcal{B}$,

$$d(f(x+h), e) \leq c(x) \left(g(x) + \|h\|_{\mathcal{B}} \right)$$

for some $e \in E$ where $c \in L^q(\mu)$, $\frac{1}{p} + \frac{1}{q} = 1$ and $\int cg \, d\mu < \infty$.

Theorem

V linear polish, $\mathcal{B} \subseteq V$ normed space,

$$d_{\mathcal{B}}(x,y) = \begin{cases} \|x-y\|_{\mathcal{B}} & \text{if } x-y \in \mathcal{B} \\ +\infty & \text{otherwise.} \end{cases}$$

Let $\mu \in P(V)$ and assume

$$\inf_{\pi \in \Pi(\nu,\mu)} \left(\int_{V \times V} d_{\mathcal{B}}(x,y)^p \, d\pi(x,y) \right)^{\frac{1}{p}} \leq \sqrt{H(\nu \,|\, \mu)} \quad \forall \nu \in P(V).$$

Let (E,d) metric space, $f : V \to (E,d)$ such that μ - a.s. for all $h \in \mathcal{B}$,

$$d(f(x+h), e) \leq c(x) \left(g(x) + \|h\|_{\mathcal{B}} \right)$$

for some $e \in E$ where $c \in L^q(\mu)$, $\frac{1}{p} + \frac{1}{q} = 1$ and $\int cg \, d\mu < \infty$.
Then $f : x \to d(f(x), e)$ has Gaussian tails.

Examples

Examples in the Gaussian framework:

Examples

Examples in the Gaussian framework:

- $f = Id$, $\rightsquigarrow$ Fernique's theorem: $\mathbb{E}[e^{\eta \|X\|^2}] < \infty$

Examples

Examples in the Gaussian framework:

- $f = Id$, $\rightsquigarrow$ Fernique's theorem: $\mathbb{E}[e^{\eta \|X\|^2}] < \infty$
- $f = S$, $\rightsquigarrow$ Fernique's theorem for Gaussian rough paths ([Friz, Oberhauser; '10]): $\mathbb{E}[e^{\eta \|\mathbf{X}\|^2_{p-\mathrm{var}}}] < \infty$

Examples

Examples in the Gaussian framework:

- $f = Id$, $\rightsquigarrow$ Fernique's theorem: $\mathbb{E}[e^{\eta\|X\|^2}] < \infty$
- $f = S$, $\rightsquigarrow$ Fernique's theorem for Gaussian rough paths ([Friz, Oberhauser; '10]): $\mathbb{E}[e^{\eta\|\mathbf{X}\|^2_{p-\mathrm{var}}}] < \infty$
- $f = I_f(\cdot, \xi)$, $\rightsquigarrow$ Gaussian integrability for solutions of RDEs ([Cass, Litterer, Lyons; '13] and [Friz, R.; '13]): $\mathbb{E}[e^{\eta\|Y\|^2_\infty}] < \infty$

Examples

Examples in the Gaussian framework:

- $f = Id$, $\rightsquigarrow$ Fernique's theorem: $\mathbb{E}[e^{\eta \|X\|^2}] < \infty$
- $f = S$, $\rightsquigarrow$ Fernique's theorem for Gaussian rough paths ([Friz, Oberhauser; '10]): $\mathbb{E}[e^{\eta \|\mathbf{X}\|^2_{p-\mathrm{var}}}] < \infty$
- $f = I_f(\cdot, \xi)$, $\rightsquigarrow$ Gaussian integrability for solutions of RDEs ([Cass, Litterer, Lyons; '13] and [Friz, R.; '13]): $\mathbb{E}[e^{\eta \|Y\|^2_\infty}] < \infty$
- $f = N_\alpha(S(\cdot), [0, T])$, $\rightsquigarrow$ integrabililty results from [Cass, Litterer, Lyons; '13] for "greedy partitions"

Multilevel and rough paths

Multilevel and rough paths

Goal: Find efficient way to evaluate

$$\mathbb{E}[g(Y_\cdot)]$$

where Y solves

$$dY_t = \sum_{i=1}^{d} f_i(Y_t)\, dX_t^i; \qquad Y_0 = \xi \in \mathbb{R}^m$$

and

$$g\colon C_\xi([0,T],\mathbb{R}^m) \to \mathbb{R}.$$

Numerical schemes

Simple step-3 Milstein scheme ([Deya, Neuenkirch, Tindel; '12]):

Numerical schemes

Simple step-3 Milstein scheme ([Deya, Neuenkirch, Tindel; '12]):

$$\begin{aligned}\bar{Y}_0 &= Y_0 \\ \bar{Y}_{t_{m+1}} &= \bar{Y}_{t_m} + f_i(\bar{Y}_{t_m}) X^i_{t_m,t_{m+1}} \\ &\quad + \frac{1}{2} f_i f_j (\bar{Y}_{t_m}) X^i_{t_m,t_{m+1}} X^j_{t_m,t_{m+1}} \\ &\quad + \frac{1}{6} f_{i_1} f_{i_2} f_{i_3} (\bar{Y}_{t_m}) X^{i_1}_{t_m,t_{m+1}} X^{i_2}_{t_m,t_{m+1}} X^{i_3}_{t_m,t_{m+1}}\end{aligned}$$

where $t_m = \frac{mT}{k}$.

Assume $|R|_{\rho-\mathrm{var};[s,t]} \lesssim |t-s|^{\frac{1}{\rho}}$.

Theorem (Friz, R; ’14 and Bayer, Friz, R. Schoenmakers)

f smooth, then for all $\eta < \frac{1}{\rho} - \frac{1}{2}$ and $q \geq 1$ there exists a constant C s.t.

$$|\|Y - \bar{Y}\|_\infty|_{L^q} \leq C \left(\frac{1}{k}\right)^\eta \quad \forall\, k \in \mathbb{N}.$$

Multilevel approach 1/3

- For simplicity:
  - $P := g(Y)$
  - $\hat{P}_l :=$ some Milstein-approximation of P with step size $h_l = \frac{T}{M_0 M^l}$

Multilevel approach 1/3

- For simplicity:
  - $P := g(Y)$
  - $\hat{P}_l :=$ some Milstein-approximation of P with step size $h_l = \frac{T}{M_0 M^l}$
- **Basic idea multilevel:**

$$\begin{aligned} \mathbb{E}[P] &\approx \mathbb{E}[\hat{P}_L] \quad \text{for } L \text{ large} \\ &= \mathbb{E}[\hat{P}_0] + \sum_{l=1}^{L} \mathbb{E}[\hat{P}_l - \hat{P}_{l-1}], \end{aligned}$$

now MC for every summand with N_l samples, $l = 0, \ldots, L$

Multilevel approach 2/3

- Set $\hat{P}_{-1} = 0$ and

$$\hat{Y}_l = \frac{1}{N_l} \sum_{i=1}^{N_l} \left(\hat{P}_l^{(i)} - \hat{P}_{l-1}^{(i)} \right)$$

Multilevel approach 2/3

- Set $\hat{P}_{-1} = 0$ and

$$\hat{Y}_l = \frac{1}{N_l} \sum_{i=1}^{N_l} \left(\hat{P}_l^{(i)} - \hat{P}_{l-1}^{(i)} \right)$$

- Then, if $\mathbb{E}\left[\left| \hat{P}_l - P \right|^2 \right] \lesssim h_l^\beta$ for $l = 1, \ldots, L$,

$$\mathbb{V}(\hat{Y}_0) \lesssim \frac{\mathbb{V}(\hat{P}_0)}{N_0} \quad \text{and} \quad \mathbb{V}(\hat{Y}_l) \lesssim \frac{h_l^\beta}{N_l}$$

Multilevel approach 2/3

- Set $\hat{P}_{-1} = 0$ and

$$\hat{Y}_l = \frac{1}{N_l} \sum_{i=1}^{N_l} \left(\hat{P}_l^{(i)} - \hat{P}_{l-1}^{(i)} \right)$$

- Then, if $\mathbb{E}\left[\left| \hat{P}_l - P \right|^2 \right] \lesssim h_l^\beta$ for $l = 1, \ldots, L$,

$$\mathbb{V}(\hat{Y}_0) \lesssim \frac{\mathbb{V}(\hat{P}_0)}{N_0} \quad \text{and} \quad \mathbb{V}(\hat{Y}_l) \lesssim \frac{h_l^\beta}{N_l}$$

- **Observe:**
  - **low levels:** computational cost for Milstein-approximation small, but variance large
  - **high levels:** computational cost for approximation large, but variance small

Multilevel approach 3/3

- Set $\hat{Y} = \sum_{l=0}^{L} \hat{Y}_l$

Multilevel approach 3/3

- Set $\hat{Y} = \sum_{l=0}^{L} \hat{Y}_l$
- **Strategy:** for given

$$\mathrm{MSE} = \mathbb{E}\left[\left(\hat{Y} - \mathbb{E}[P]\right)^2\right] \leq \varepsilon^2,$$

optimize over $L, N_0, \dots, N_L$ in order to minimize computational cost C

Theorem (Bayer, Friz, R., Schoenmakers)

Assume $\exists\, \alpha < 1/2,\ \beta \leq 2\alpha < 1$ *s.t.*

(i) $\mathbb{E}[\hat{P}_l - P] \lesssim h_l^{\alpha}$,

(ii) $\mathbb{E}[|\hat{P}_l - P|^2] \lesssim h_l^{\beta}$.

Theorem (Bayer, Friz, R., Schoenmakers)

Assume $\exists\, \alpha < 1/2$, $\beta \leq 2\alpha < 1$ *s.t.*

(i) $\mathbb{E}[\hat{P}_l - P] \lesssim h_l^{\alpha}$,

(ii) $\mathbb{E}[|\hat{P}_l - P|^2] \lesssim h_l^{\beta}$.

Then $\forall\, \varepsilon > 0$, *there are* $L, N_0, \ldots, N_L \in \mathbb{N}$ *s.t.*

$$MSE = \mathbb{E}\left[\left(\hat{Y} - \mathbb{E}[P]\right)^2\right] \leq \varepsilon^2,$$

and for the computational complexity C *one has*

$$C = \mathcal{O}\left(\varepsilon^{-\frac{1+2\alpha-\beta}{\alpha}}\right).$$

Theorem (Bayer, Friz, R., Schoenmakers)

Assume $\exists\, \alpha < 1/2$, $\beta \le 2\alpha < 1$ *s.t.*

(i) $\mathbb{E}[\hat{P}_l - P] \lesssim h_l^{\alpha}$,

(ii) $\mathbb{E}[|\hat{P}_l - P|^2] \lesssim h_l^{\beta}$.

Then $\forall\, \varepsilon > 0$, *there are* $L, N_0, \ldots, N_L \in \mathbb{N}$ *s.t.*

$$MSE = \mathbb{E}\left[\left(\hat{Y} - \mathbb{E}[P]\right)^2\right] \le \varepsilon^2,$$

and for the computational complexity C *one has*

$$C = \mathcal{O}\left(\varepsilon^{-\frac{1+2\alpha-\beta}{\alpha}}\right).$$

Recall: complexity for usual MC would be $C = \mathcal{O}\left(\varepsilon^{-(2+1/\alpha)}\right)$

Remarks:

Remarks:

- Extension of Giles' Complexity theorem ([Giles; '08]) to the case $\alpha < 1/2$

Remarks:

- Extension of Giles' Complexity theorem ([Giles; '08]) to the case $\alpha < 1/2$
- For $\beta \approx 1$ (BM case), $C \approx \mathcal{O}(\varepsilon^{-2-})$ ((almost) sharp!)

Remarks:

- Extension of Giles' Complexity theorem ([Giles; '08]) to the case $\alpha < 1/2$
- For $\beta \approx 1$ (BM case), $C \approx \mathcal{O}(\varepsilon^{-2-})$ ((almost) sharp!)
- If $\alpha = \beta/2$ (always true by Jensen), then $C = \mathcal{O}(\varepsilon^{-2/\beta})$

Remarks:

- Extension of Giles' Complexity theorem ([Giles; '08]) to the case $\alpha < 1/2$
- For $\beta \approx 1$ (BM case), $C \approx \mathcal{O}(\varepsilon^{-2-})$ ((almost) sharp!)
- If $\alpha = \beta/2$ (always true by Jensen), then $C = \mathcal{O}(\varepsilon^{-2/\beta})$
- All constants can be calculated explicitly

Theorem (Bayer, Friz, R., Schoenmakers)

Assume

- *g Lipschitz*
- *X Gaussian, usual assumptions hold.*

Theorem (Bayer, Friz, R., Schoenmakers)

Assume

- *g Lipschitz*
- *X Gaussian, usual assumptions hold.*

Then the MC evaluation of $\mathbb{E}[g(Y)]$ *under condition*

$$MSE \leq \varepsilon^2$$

can be achieved with computational complexity

$$\mathcal{O}(\varepsilon^{-\theta}) \quad \forall \theta > \frac{2\rho}{2-\rho}.$$

References

C. Bayer, P. Friz, S. Riedel, J. Schoenmakers.
From rough path estimates to multilevel Monte Carlo.
arXiv preprint, 2013.

P. Friz, B. Gess, A. Gulisashvili, S. Riedel.
Jain–Monrad criterion for rough paths and applications.
arXiv preprint, 2013.

P. Friz, S. Riedel.
Convergence rates for the full Gaussian rough paths.
AIHP, 2014.

S. Riedel.
Talagrand's transportation–cost inequality and applications to (rough) paths spaces.
Preprint, 2014.

Thank you for your attention.