Constrained optimal transport problems and stochastic differential equations

Benjamin A. Robinson (Universität Wien)

January 25, 2024 — Assistant Professor Hearing, Universität Klagenfurt

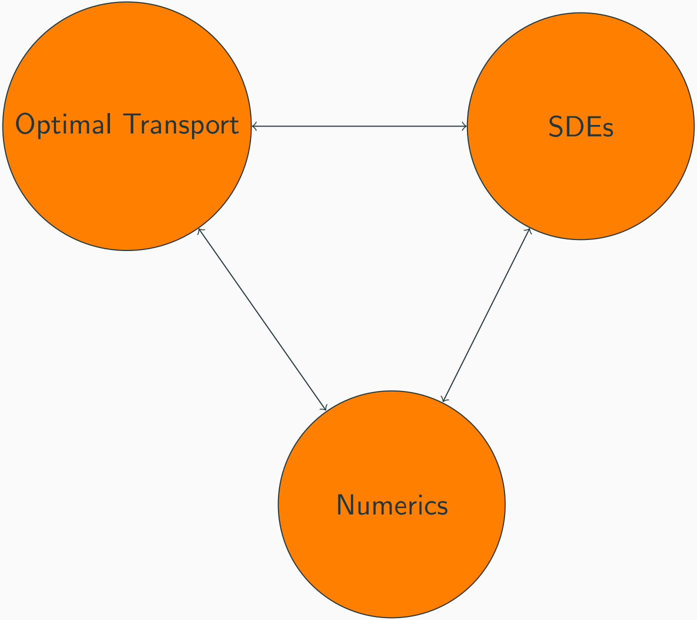
Supported by Austrian Science Fund (FWF) projects Y782-N25, P35519, P34743.

Overview

Current research

1. Computing distances between stochastic processes;
2. Fitting processes to given marginals.

Overview



Comparing stochastic models

Aim: Compute a measure of model uncertainty

E.g.

$$\mathbb{P} \mapsto v(\mathbb{P}) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}^{\mathbb{P}}[\mathcal{J}(\omega, \alpha)]$$

Comparing stochastic models

Aim: Compute a measure of model uncertainty

E.g.

$$\mathbb{P} \mapsto v(\mathbb{P}) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}^{\mathbb{P}}[\mathcal{J}(\omega, \alpha)]$$

Want:

- Appropriate topology on laws of stochastic processes
- Distance we can actually compute

Comparing stochastic models

Aim: Compute a measure of model uncertainty

E.g.

$$\mathbb{P} \mapsto v(\mathbb{P}) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}^{\mathbb{P}}[\mathcal{J}(\omega, \alpha)]$$

Want:

- Appropriate topology on laws of stochastic processes
- Distance we can actually compute

SDEs:

- Good computational methods available
- Rich class of models, beyond Lipschitz coefficients

Distances between stochastic processes

Adapted Wasserstein distance between the laws of SDEs

arXiv:2209.03243, 2022

Julio Backhoff-Veraguas

Universität Wien

Sigrid Källblad

KTH Stockholm

Distances between stochastic processes

Adapted Wasserstein distance between the laws of SDEs
with Julio Backhoff-Veraguas (Universität Wien) and
Sigrid Källblad (KTH Stockholm), arXiv:2209.03243, 2022

Bicausal optimal transport for SDEs with irregular coefficients
Preprint, 2024

Michaela Szölgyenyi
Universität Klagenfurt

Distances between stochastic processes

$b, \bar{b} \colon \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon \mathbb{R} \to [0, \infty)$, $X_0 = \bar{X}_0 = x \in \mathbb{R}$,

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t$$

$\mu = \mathrm{Law}(X)$, $\nu = \mathrm{Law}(\bar{X})$

Distances between stochastic processes

$b, \bar{b} \colon \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon \mathbb{R} \to [0, \infty)$, $X_0 = \bar{X}_0 = x \in \mathbb{R}$,

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t$$

$\mu = \mathrm{Law}(X)$, $\nu = \mathrm{Law}(\bar{X})$

Main result

Under "weak assumptions" on the coefficients, we can compute an "appropriate distance"

Distances between stochastic processes

$b, \bar{b} \colon \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon \mathbb{R} \to [0, \infty)$, $X_0 = \bar{X}_0 = x \in \mathbb{R}$,

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t$$

$\mu = \mathrm{Law}(X)$, $\nu = \mathrm{Law}(\bar{X})$

Main result

Under "weak assumptions" on the coefficients, we can compute an "appropriate distance" by

$$d(\mu,\nu)^2 = \mathbb{E}\left[\int_0^T |X_t - \bar{X}_t|^2 \mathrm{d}t\right], \quad \text{with } B = W.$$

Optimal transport

Probability measures μ, ν on $\mathbb{R}^N$

Find

$$\mathcal{W}_2^2(\mu,\nu) \coloneqq \inf_{T\colon T_\#\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1,\dots,X_N),\dots,T_N(X_1,\dots,X_N))$$

Optimal transport

Probability measures μ, ν on $\mathbb{R}^N$

Find

$$\mathcal{W}_2^2(\mu,\nu) := \inf_{T\colon T_\#\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1,\ldots,X_N),\ldots,T_N(X_1,\ldots,X_N))$$

Optimal transport

Probability measures μ, ν on $\mathbb{R}^N$

Find

$$\mathcal{W}_2^2(\mu,\nu) := \inf_{T\colon T_\# \mu = \nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1,\dots,X_N),\dots,T_N(X_1,\dots,X_N))$$

Monge (1781)

Optimal transport

Probability measures μ, ν on $\mathbb{R}^N$

Find

$$\mathcal{W}_2^2(\mu,\nu) \coloneqq \inf_{T\colon T_{\#}\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1,\ldots,X_N),\ldots,T_N(X_1,\ldots,X_N))$$

Monge (1781)
Kantorovich (1942) $\rightsquigarrow T$ random:

Optimal transport

Probability measures μ, ν on $\mathbb{R}^N$

Find

$$\mathcal{W}_2^2(\mu,\nu) := \inf_{T\colon T_{\#}\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1,\ldots,X_N),\ldots,T_N(X_1,\ldots,X_N))$$

Monge (1781)

Kantorovich (1942) $\rightsquigarrow$ T random: replace $(X, T(X))$ with coupling (X,Y), $X \sim \mu$, $Y \sim \nu$

Optimal transport

Probability measures μ, ν on $\mathbb{R}^N$

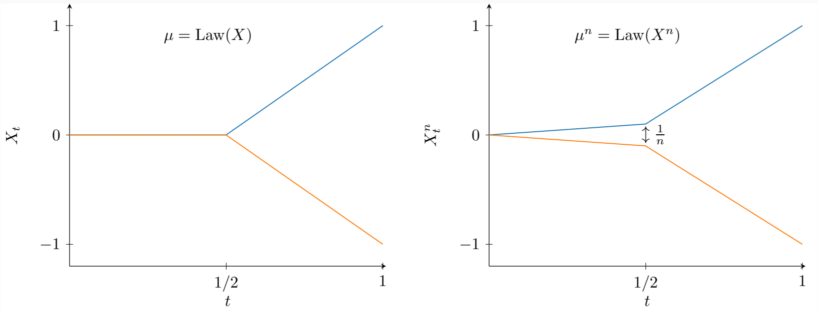
Find

$$\mathcal{W}_2^2(\mu,\nu) := \inf_{T\colon T_\#\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1,\dots,X_N),\dots,T_N(X_1,\dots,X_N))$$

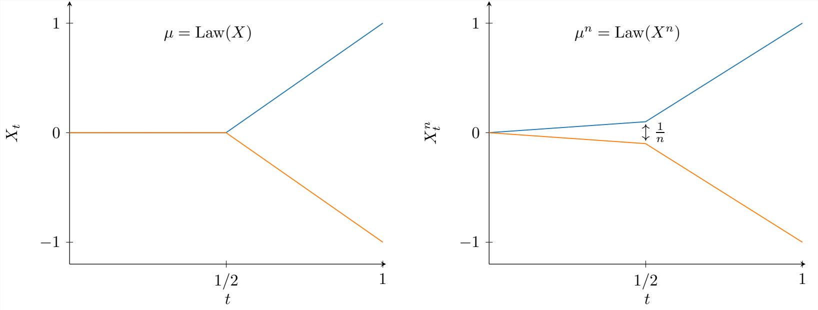
Monge (1781)
Kantorovich (1942) $\rightsquigarrow$ T random: replace $(X, T(X))$ with coupling (X,Y), $X \sim \mu$, $Y \sim \nu$
Wasserstein distance metrises usual weak topology

Example



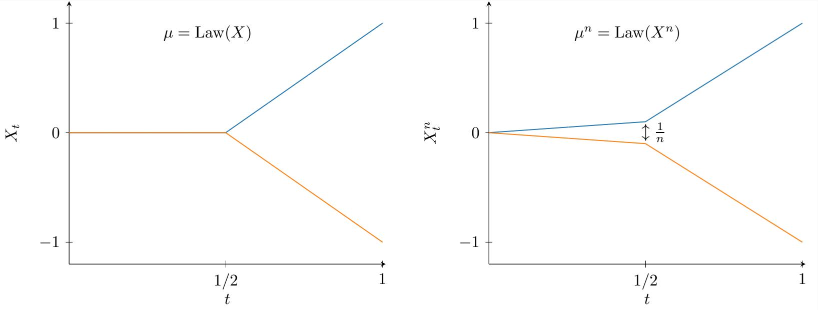
$\mu = \mathrm{Law}(X)$
X_t
1
0
-1
1/2
1
t
$\mu^n = \mathrm{Law}(X^n)$
X_t^n
1
0
-1
$\frac{1}{n}$
1/2
1
t

Example



"Cannot get rich" "Can get rich"

Example



"Cannot get rich" "Can get rich"

But

$$\mathcal{W}_2^2(\mu, \mu_n) = \inf_{T\colon T_\#\mu = \mu_n} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right] \to 0$$

$$\mu_n \rightharpoonup \mu$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{W}_2^2(\mu,\nu) := \inf_{T:\, T_\# \mu = \nu} \mathbb{E}\left[\sum_{n=1}^N |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1, \ldots, X_N), \ldots, T_N(X_1, \ldots, X_N))$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \inf_{\substack{T\colon T_\# \mu = \nu \\ \text{adapted}}} \mathbb{E}\left[\sum_{n=1}^N |T_n(X) - X_n|^2\right].$$

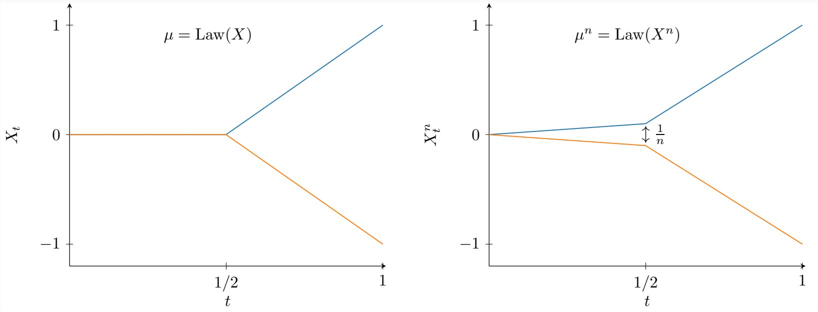
$$T(X) = (T_1(X_1), T_2(X_1, X_2), \dots, T_N(X_1, \dots, X_N))$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_2^2(\mu, \nu) := \inf_{\substack{T:\, T_\# \mu = \nu \\ \text{biadapted}}} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1), T_2(X_1, X_2), \ldots, T_N(X_1, \ldots, X_N))$$

Example



$$\mathcal{AW}_2^2(\mu, \mu_n) > \frac{1}{2}$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_2^2(\mu, \nu) := \inf_{\substack{T:\, T_\# \mu = \nu \\ \text{biadapted}}} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1), T_2(X_1, X_2), \ldots, T_N(X_1, \ldots, X_N))$$

Similar definition of adapted Wasserstein distance in continuous time w.r.t. L^p norm on $C([0,T], \mathbb{R})$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_2^2(\mu, \nu) := \inf_{\substack{T\colon T_\# \mu = \nu \\ \text{biadapted}}} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^2\right].$$

$$T(X) = (T_1(X_1), T_2(X_1, X_2), \ldots, T_N(X_1, \ldots, X_N))$$

Similar definition of adapted Wasserstein distance in continuous time w.r.t. L^p norm on $C([0,T], \mathbb{R})$

More general cost functions $\rightsquigarrow$ bicausal optimal transport

Adapted topology

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [Backhoff-Veraguas, Källblad, R. '22]

Optimising over biadapted maps $T_\# \mu = \nu$

$\Leftrightarrow$

Optimising over correlations between B, W

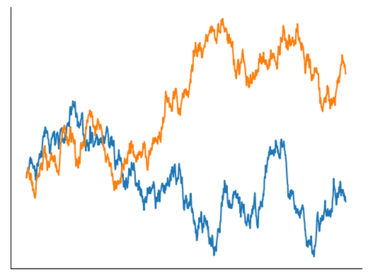
Adapted topology

Theorem [Backhoff-Veraguas, Källblad, R. '22]

$$\text{Optimising over biadapted maps } T_{\#}\mu = \nu$$
$$\Leftrightarrow$$
$$\text{Optimising over correlations between } B, W$$

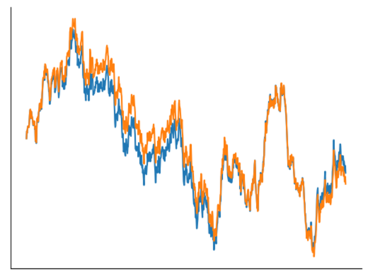
Product coupling

B, W independent



Synchronous coupling

Choose the same driving Brownian motion $B = W$.



Regular SDEs

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow \ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow \ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [Backhoff-Veraguas, Källblad, R. '22]

For Lipschitz coefficients, the adapted Wasserstein distance is

$$\mathcal{AW}_p^p(\mu,\nu) = \mathbb{E}\Big[\int_0^T |X_t - \bar{X}_t|^p \mathrm{d}t\Big], \quad \text{with } B = W.$$

Regular SDEs

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [Backhoff-Veraguas, Källblad, R. '22]

For Lipschitz coefficients, the adapted Wasserstein distance is

$$\mathcal{AW}_p^p(\mu,\nu) = \mathbb{E}\Big[\int_0^T |X_t - \bar{X}_t|^p \mathrm{d}t\Big], \quad \text{with } B = W.$$

1. Discretise SDEs;

Regular SDEs

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow \ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow \ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [Backhoff-Veraguas, Källblad, R. '22]

For Lipschitz coefficients, the adapted Wasserstein distance is

$$\mathcal{AW}_p^p(\mu,\nu) = \mathbb{E}\Big[\int_0^T |X_t - \bar{X}_t|^p \mathrm{d}t\Big], \quad \text{with } B = W.$$

1. Discretise SDEs;
2. Solve discrete-time bicausal optimal transport problem;

Regular SDEs

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [Backhoff-Veraguas, Källblad, R. '22]

For Lipschitz coefficients, the adapted Wasserstein distance is

$$\mathcal{AW}_p^p(\mu,\nu) = \mathbb{E}\Big[\int_0^T |X_t - \bar{X}_t|^p \mathrm{d}t\Big], \quad \text{with } B = W.$$

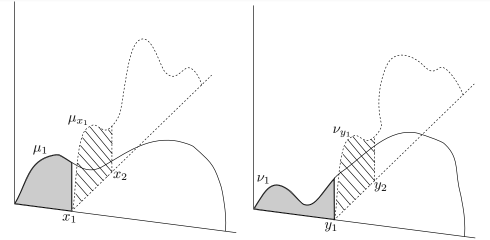
1. Discretise SDEs;
2. Solve discrete-time bicausal optimal transport problem;
3. Pass to a limit.

Discretisation

$$\mu_N, \nu_N \in \mathcal{P}(\mathbb{R}^N) \ \rightsquigarrow \ \mathcal{AW}_p^p(\mu_N, \nu_N) := \inf_{\substack{T:\, T_\# \mu_N = \nu_N \\ \text{biadapted}}} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^p\right]$$

Theorem [Rüschendorf '85]

For μ_N, ν_N stochastically co-monotone, the unique optimiser is the Knothe–Rosenblatt rearrangement



Discretisation

$$\mu_N, \nu_N \in \mathcal{P}(\mathbb{R}^N) \rightsquigarrow \mathcal{AW}_p^p(\mu_N, \nu_N) := \inf_{\substack{T\colon T_\# \mu_N = \nu_N \\ \text{biadapted}}} \mathbb{E}\left[\sum_{n=1}^N |T_n(X) - X_n|^p\right]$$

Theorem [Rüschendorf '85]

For μ_N, ν_N stochastically co-monotone, the unique optimiser is the Knothe–Rosenblatt rearrangement

Theorem [Backhoff-Veraguas, Källblad, R. '22]

In the case of Lipschitz coefficients, the Euler–Maruyama scheme is stochastically increasing when the Brownian increments are truncated

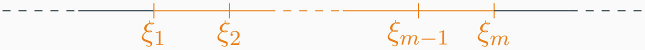
$$X_{k+1}^h = X_k^h + b(X_k^h) \cdot h + \sigma(X_k^h) \Delta W_{k+1}^h$$

Irregular SDEs

Assumption (A)

Drift $b\colon \mathbb{R} \to \mathbb{R}$ satisfies the following conditions piecewise:

- absolute continuity
- one-sided Lipschitz condition
- two-sided local Lipschitz condition
- exponential growth

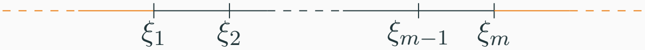


Irregular SDEs

Assumption (A)

Drift $b\colon \mathbb{R} \to \mathbb{R}$ satisfies the following conditions piecewise:

- absolute continuity
- one-sided Lipschitz condition
- two-sided local Lipschitz condition
- exponential growth



Irregular SDEs

Assumption (A)

Drift $b\colon \mathbb{R} \to \mathbb{R}$ satisfies the following conditions piecewise:

– absolute continuity

– one-sided Lipschitz condition

– two-sided local Lipschitz condition

– exponential growth

Diffusion $\sigma\colon \mathbb{R} \to [0, \infty)$ satisfies

– global Lipschitz condition

– $\sigma(\xi_k) \neq 0$, for $k \in \{1, \dots, m\}$ — no uniform ellipticity

Irregular SDEs

Assumption (A)

Drift b piecewise regular with at most exponential growth,
σ Lipschitz and non-zero at discontinuity points of b.

Transformation-based semi-implicit Euler scheme

1. Transform the SDE $Z = G(X)$, with G increasing Lipschitz,
2. Apply a semi-implicit Euler scheme with truncated Brownian increments to Z,
3. Transform back $X^h = G^{-1}(Z^h)$.

Irregular SDEs

Assumption (A)

Drift b piecewise regular with at most exponential growth,
σ Lipschitz and non-zero at discontinuity points of b.

Transformation-based semi-implicit Euler scheme

1. Transform the SDE $Z = G(X)$, with G increasing Lipschitz,
2. Apply a semi-implicit Euler scheme with truncated Brownian increments to Z,
3. Transform back $X^h = G^{-1}(Z^h)$.

Theorem [R., Szölgyenyi '24+]

Let (b, σ) satisfy Assumption (A). Then for all $p \geq 1$, there exists $C_p \geq 0$ such that

$$\mathbb{E}\Big[|X_T - X_T^h|^p\Big]^{\frac{1}{p}} \leq \begin{cases} C_p h^{\frac{1}{2}}, & p \in [1,2], \\ C_p h^{\frac{1}{p(p-1)}}, & p \geq 2. \end{cases}$$

Irregular SDEs

Assumptions
(A) discontinuous drift with exponential growth;

Irregular SDEs

Assumptions
(A) discontinuous drift with exponential growth;

(B) bounded measurable drift, α-Hölder and uniformly elliptic σ;

Irregular SDEs

Assumptions

(A) discontinuous drift with exponential growth;

(B) bounded measurable drift, α-Hölder and uniformly elliptic σ;

(C) “regular” coefficients.

Irregular SDEs

Assumptions

(A) discontinuous drift with exponential growth;

(B) bounded measurable drift, α-Hölder and uniformly elliptic σ;

(C) "regular" coefficients.

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$

$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Irregular SDEs

Assumptions
(A) discontinuous drift with exponential growth;
(B) bounded measurable drift, α-Hölder and uniformly elliptic σ;
(C) "regular" coefficients.

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [R., Szölgyenyi '24+]

Let (b, σ) and $(\bar{b}, \bar{\sigma})$ each satisfy one of assumptions (A), (B), (C).
Then, for $p \in [1, \infty)$, the adapted Wasserstein distance is given by

$$\mathcal{AW}_p^p(\mu, \nu) = \mathbb{E}\Big[\int_0^T |X_t - \bar{X}_t|^p \mathrm{d}t\Big], \text{ with } B = W$$

Irregular SDEs

Assumptions
(A) discontinuous drift with exponential growth;
(B) bounded measurable drift, α-Hölder and uniformly elliptic σ;
(C) "regular" coefficients.

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [R., Szölgyenyi '24+]

Let (b, σ) and $(\bar{b}, \bar{\sigma})$ each satisfy one of assumptions (A), (B), (C). Then, for $p \in [1, \infty)$, the adapted Wasserstein distance is given by

$$\mathcal{AW}_p^p(\mu, \nu) = \mathbb{E}\Big[\int_0^T |X_t - \bar{X}_t|^p \mathrm{d}t\Big], \text{ with } B = W$$

Synchronous coupling solves general bicausal transport problem

Future research directions

- Extension to higher dimensions
  - Examples in [Backhoff-Veraguas, Källblad, R. '22] show that the synchronous coupling is not always optimal
- Extension to jump-diffusions
- Extension to neural SDEs, McKean–Vlasov SDEs
- Convergence of optimisers
  - Use density estimates for SDEs from [Backhoff-Veraguas, Unterberger 2023]
- Application to uniqueness of mimicking martingales

Mimicking martingales

Given probability measures μ, ν on $\mathbb{R}^d$ do there exist random variables M_0, M_1 such that

- $\mathbb{E}[M_1 | M_0] = M_0$ (martingale property)

Mimicking martingales

Given probability measures μ, ν on $\mathbb{R}^d$ do there exist random variables M_0, M_1 such that

- $\mathbb{E}[M_1 | \, M_0] = M_0$ (martingale property) and
- $\mathrm{Law}(M_0) = \mu$ and $\mathrm{Law}(M_1) = \nu$ (mimicking property)?

Mimicking martingales

Given probability measures μ, ν on $\mathbb{R}^d$ do there exist random variables M_0, M_1 such that

- $\mathbb{E}[M_1 | \, M_0] = M_0$ (martingale property) and
- $\mathrm{Law}(M_0) = \mu$ and $\mathrm{Law}(M_1) = \nu$ (mimicking property)?

Necessary condition $\mu \preceq \nu$ in convex order

For any convex function $v : \mathbb{R}^d \to \mathbb{R}$,

$$\int v \mathrm{d}\mu \leq \int v \mathrm{d}\nu.$$

Mimicking martingales

Given probability measures μ, ν on $\mathbb{R}^d$ do there exist random variables M_0, M_1 such that

- $\mathbb{E}[M_1 | \, M_0] = M_0$ (martingale property) and
- $\mathrm{Law}(M_0) = \mu$ and $\mathrm{Law}(M_1) = \nu$ (mimicking property)?

Necessary condition $\mu \preceq \nu$ in convex order

For any convex function $v : \mathbb{R}^d \to \mathbb{R}$,

$$\int v \mathrm{d}\mu \leq \int v \mathrm{d}\nu.$$

... also sufficient [Strassen '65]

Mimicking martingales

Given a family of probability measures $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Mimicking martingales

A regularized Kellerer theorem in arbitrary dimension
Annals of Applied Probability, to appear, with

Gudmund Pammer
ETH Zürich

Walter Schachermayer
Universität Wien

Peacocks

Assume that μ is a peacock; i.e. for any convex function $v : \mathbb{R}^d \to \mathbb{R}$,

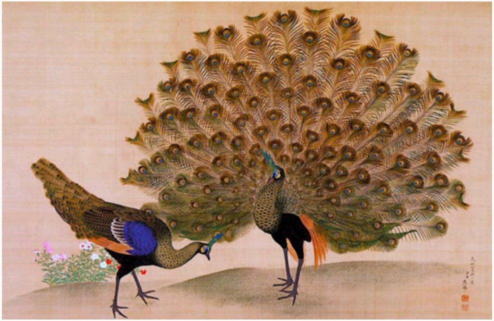
$$\int v \mathrm{d}\mu_s \leq \int v \mathrm{d}\mu_t, \quad s \leq t.$$

Peacocks

Assume that μ is a peacock; i.e. for any convex function $v : \mathbb{R}^d \to \mathbb{R}$,

$$\int v \mathrm{d}\mu_s \leq \int v \mathrm{d}\mu_t, \quad s \leq t.$$

Processus Croissant pour l'Ordre Convexe



[Hirsch, Profetta, Roynette, Yor '11]

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Yes – [Strassen '65], [Doob '68], [Hirsch, Roynette '13]

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Desirable properties

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Desirable properties

- Markovianity

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Desirable properties

- strong Markovianity

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Desirable properties

- strong Markovianity
- continuity of paths

Mimicking martingale

Given a peacock $(\mu_t)_{t\in I}$ on $\mathbb{R}^d$, does there exist a mimicking martingale M such that

$$\mathrm{Law}(M_t) = \mu_t, \quad \forall t \in I?$$

Desirable properties

- strong Markovianity
- continuity of paths
- uniqueness

Mimicking martingales

Discrete time, $d \in \mathbb{N}$ [Strassen '65]

Given a discrete-time peacock $(\mu_n)_{n \in N}$ on $\mathbb{R}^d$,

Mimicking martingales

Discrete time, $d \in \mathbb{N}$ [Strassen '65]

Given a discrete-time peacock $(\mu_n)_{n \in N}$ on $\mathbb{R}^d$, there exists a mimicking Markov martingale M.

Continuous time, $d = 1$ [Kellerer '72]

Given a continuous-time peacock $(\mu_t)_{t \in [0,1]}$ on $\mathbb{R}$,

Mimicking martingales

Discrete time, $d \in \mathbb{N}$ [Strassen '65]

Given a discrete-time peacock $(\mu_n)_{n \in N}$ on $\mathbb{R}^d$, there exists a mimicking Markov martingale M.

Continuous time, $d = 1$ [Kellerer '72]

Given a continuous-time peacock $(\mu_t)_{t \in [0,1]}$ on $\mathbb{R}$, there exists a mimicking strong Markov martingale M.

Mimicking martingales

Discrete time, $d \in \mathbb{N}$ [Strassen '65]

Given a discrete-time peacock $(\mu_n)_{n\in N}$ on $\mathbb{R}^d$, there exists a mimicking Markov martingale M.

Continuous time, $d = 1$ [Kellerer '72], [Lowther '08–'10]

Given a continuous-time peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}$, there exists a mimicking strong Markov martingale M.

- M is the unique strong Markov mimicking martingale

Mimicking martingales

Discrete time, $d \in \mathbb{N}$ [Strassen '65]

Given a discrete-time peacock $(\mu_n)_{n \in N}$ on $\mathbb{R}^d$, there exists a mimicking Markov martingale M.

Continuous time, $d = 1$ [Kellerer '72], [Lowther '08–'10]

Given a continuous-time peacock $(\mu_t)_{t \in [0,1]}$ on $\mathbb{R}$, there exists a mimicking strong Markov martingale M.

- M is the unique strong Markov mimicking martingale
- $t \mapsto M_t$ is continuous, if $t \mapsto \mu_t$ is weakly continuous with convex support.

Mimicking martingales

Discrete time, $d \in \mathbb{N}$ [Strassen '65]

Given a discrete-time peacock $(\mu_n)_{n \in N}$ on $\mathbb{R}^d$, there exists a mimicking Markov martingale M.

Continuous time, $d = 1$ [Kellerer '72], [Lowther '08–'10]

Given a continuous-time peacock $(\mu_t)_{t \in [0,1]}$ on $\mathbb{R}$, there exists a mimicking strong Markov martingale M.

- M is the unique strong Markov mimicking martingale
- $t \mapsto M_t$ is continuous, if $t \mapsto \mu_t$ is weakly continuous with convex support.

Continuous time, $d \geq 2$

No previous results with Markov property

Mimicking martingales

Weakly continuous $\mathbb{R}^d$-valued square-integrable peacock $(\mu_t)_{t\in[0,1]}$.

Regularize with a Gaussian $\mu_t^{\mathrm{r}} := \mu_t * \gamma^{\varepsilon(t+\delta)}$

Mimicking martingales

Weakly continuous $\mathbb{R}^d$-valued square-integrable peacock $(\mu_t)_{t\in[0,1]}$.
Regularize with a Gaussian $\mu_t^{\mathrm{r}} := \mu_t * \gamma^{\varepsilon(t+\delta)}$

Theorem [Pammer, R., Schachermayer '24]

There exists a measurable $(t,x) \mapsto \sigma_t^{\mathrm{r}}(x)$ that is locally Lipschitz in x and non-degenerate, uniformly in $t \in [0,1]$, and a Brownian motion B such that $\mathrm{Law}(M_t^{\mathrm{r}}) = \mu_t^{\mathrm{r}}$, for all $t \in [0,1]$, where

$$\mathrm{d}M_t^{\mathrm{r}} = \sigma_t^{\mathrm{r}}(M_t^{\mathrm{r}})\mathrm{d}B_t.$$

Mimicking martingales

Weakly continuous $\mathbb{R}^d$-valued square-integrable peacock $(\mu_t)_{t\in[0,1]}$.
Regularize with a Gaussian $\mu_t^{\mathrm{r}} := \mu_t * \gamma^{\varepsilon(t+\delta)}$

Theorem [Pammer, R., Schachermayer '24]

There exists a measurable $(t,x) \mapsto \sigma_t^{\mathrm{r}}(x)$ that is locally Lipschitz in x and non-degenerate, uniformly in $t \in [0,1]$, and a Brownian motion B such that $\mathrm{Law}(M_t^{\mathrm{r}}) = \mu_t^{\mathrm{r}}$, for all $t \in [0,1]$, where

$$\mathrm{d}M_t^{\mathrm{r}} = \sigma_t^{\mathrm{r}}(M_t^{\mathrm{r}})\mathrm{d}B_t.$$

Moreover:

- M^{r} is a strong Markov martingale with continuous paths;

Mimicking martingales

Weakly continuous $\mathbb{R}^d$-valued square-integrable peacock $(\mu_t)_{t\in[0,1]}$.
Regularize with a Gaussian $\mu_t^{\mathrm{r}} := \mu_t * \gamma^{\varepsilon(t+\delta)}$

Theorem [Pammer, R., Schachermayer '24]

There exists a measurable $(t, x) \mapsto \sigma_t^{\mathrm{r}}(x)$ that is locally Lipschitz in x and non-degenerate, uniformly in $t \in [0, 1]$, and a Brownian motion B such that $\mathrm{Law}(M_t^{\mathrm{r}}) = \mu_t^{\mathrm{r}}$, for all $t \in [0, 1]$, where

$$\mathrm{d}M_t^{\mathrm{r}} = \sigma_t^{\mathrm{r}}(M_t^{\mathrm{r}})\mathrm{d}B_t.$$

Moreover:

- M^{r} is a strong Markov martingale with continuous paths;
- The result does not hold without regularization;

Mimicking martingales

Weakly continuous $\mathbb{R}^d$-valued square-integrable peacock $(\mu_t)_{t\in[0,1]}$.
Regularize with a Gaussian $\mu_t^{\mathrm{r}} := \mu_t * \gamma^{\varepsilon(t+\delta)}$

Theorem [Pammer, R., Schachermayer '24]

There exists a measurable $(t,x) \mapsto \sigma_t^{\mathrm{r}}(x)$ that is locally Lipschitz in x and non-degenerate, uniformly in $t \in [0,1]$, and a Brownian motion B such that $\mathrm{Law}(M_t^{\mathrm{r}}) = \mu_t^{\mathrm{r}}$, for all $t \in [0,1]$, where

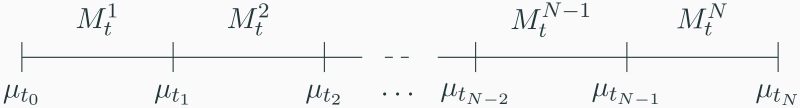
$$\mathrm{d}M_t^{\mathrm{r}} = \sigma_t^{\mathrm{r}}(M_t^{\mathrm{r}})\mathrm{d}B_t.$$

Moreover:

- M^{r} is a strong Markov martingale with continuous paths;
- The result does not hold without regularization;
- There is no uniqueness for $d \geq 2$.

Discretize and take Bass martingales from μ_{t_k} to $\mu_{t_{k+1}}$ to get a "nice" diffusion process
[Backhoff-Veraguas, Beiglböck, Huesmann, Källblad 2020]
[Backhoff-Veraguas, Beiglböck, Schachermayer, Tschiderer 2023]

Bass martingales (Martingale Optimal Transport)



$$M_t = \mathbb{E}[\nabla v(B_1) \mid \mathcal{F}_t] = \nabla v_t(B_t), \quad v_t \colon \mathbb{R}^d \to \mathbb{R} \text{ convex}, \quad B_0 \sim \alpha.$$

Prove compactness of such "nice" diffusions.

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists a weakly continuous square-integrable peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}^4$ such that, for the peacock $(\mu_t * \gamma^t)_{t\in[0,1]}$, there exists no mimicking Markov martingale.

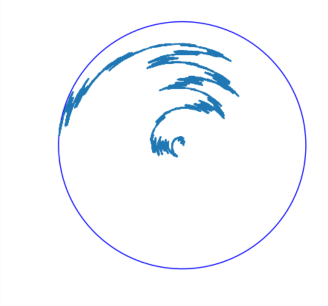
Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists a weakly continuous square-integrable peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}^4$ such that, for the peacock $(\mu_t * \gamma^t)_{t\in[0,1]}$, there exists no mimicking Markov martingale.

Optimal control of martingales in a radially symmetric environment
Alexander Cox, B.R., *Stochastic Processes and their Applications* 2023

SDEs with no strong solutions arising from a problem of stochastic control
Alexander Cox, B.R., *Electronic Journal of Probability* 2023



[Cox, R. '23]

Mimicking martingales

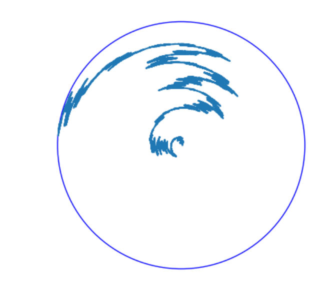
Theorem [Pammer, R., Schachermayer '24]

There exists a weakly continuous square-integrable peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}^4$ such that, for the peacock $(\mu_t * \gamma^t)_{t\in[0,1]}$, there exists no mimicking Markov martingale.

Theorem [Cox, R. '23]

There is a unique weak solution but no strong solution of

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix} -X_t^2 \\ X_t^1 \end{bmatrix}\mathrm{d}W_t, \quad X_0 = 0.$$



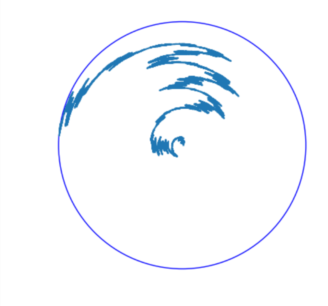
[Cox, R. '23]

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists a weakly continuous square-integrable peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}^4$ such that, for the peacock $(\mu_t * \gamma^t)_{t\in[0,1]}$, there exists no mimicking Markov martingale.

1. No continuous Markov martingale mimicking μ;



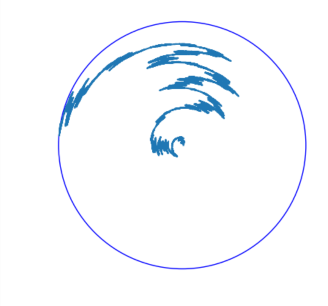
[Cox, R. '23]

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists a weakly continuous square-integrable peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}^4$ such that, for the peacock $(\mu_t * \gamma^t)_{t\in[0,1]}$, there exists no mimicking Markov martingale.

1. No continuous Markov martingale mimicking μ;
2. No Markov martingale mimicking μ;



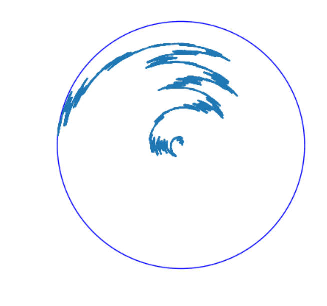
[Cox, R. '23]

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists a weakly continuous square-integrable peacock $(\mu_t)_{t\in[0,1]}$ on $\mathbb{R}^4$ such that, for the peacock $(\mu_t * \gamma^t)_{t\in[0,1]}$, there exists no mimicking Markov martingale.

1. No continuous Markov martingale mimicking μ;
2. No Markov martingale mimicking μ;
3. No Markov martingale mimicking $(\mu * \gamma^t)_{t\in[0,1]}$.



[Cox, R. '23]

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists an $\mathbb{R}^2$-valued strong Markov martingale diffusion with Brownian marginals, which is not a Brownian motion.

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

There exists an $\mathbb{R}^2$-valued strong Markov martingale diffusion with Brownian marginals, which is not a Brownian motion.

Fake Brownian motion in 1D

There exists a *continuous Markov martingale with Brownian marginals*, which is *not* a Brownian motion.
[Albin 2008, Beiglböck, Lowther, Pammer, Schachermayer 2023, Hamza, Klebaner 2007, Oleszkiewicz 2008]

Mimicking martingales

Theorem [Pammer, R., Schachermayer '24]

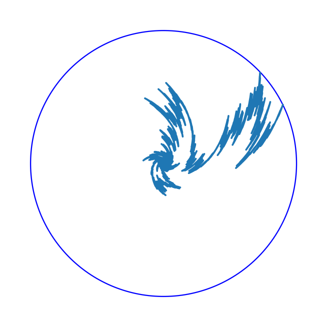
There exists an $\mathbb{R}^2$-valued strong Markov martingale diffusion with Brownian marginals, which is not a Brownian motion.

Theorem [Cox, R. '23]

Let X be a weak solution of

$$\mathrm{d}X_t = \frac{1}{|X_t|}(X_t + X_t^\perp)\mathrm{d}W_t,\ X_0 \sim \eta.$$

Then X is a continuous strong Markov fake Brownian motion.



[Pammer, R., Schachermayer '22]

Future research directions

- Find minimal conditions for the existence of Markov mimicking martingales
- Find the appropriate sense in which a mimicking martingale is canonical

Future research directions

- Find minimal conditions for the existence of Markov mimicking martingales
- Find the appropriate sense in which a mimicking martingale is canonical
  - Bass martingale: unique projection of Brownian motion onto

$$\{M \text{ martingale}: \ M_0 \sim \mu_0, M_1 \sim \mu_1\}$$

with respect to adapted Wasserstein distance

[Backhoff-Veraguas, Beiglböck, Huesmann, Källblad 2020]
[Backhoff-Veraguas, Beiglböck, Schachermayer, Tschiderer 2023]

Future research directions

- Find minimal conditions for the existence of Markov mimicking martingales
- Find the appropriate sense in which a mimicking martingale is canonical
  - Bass martingale: unique projection of Brownian motion onto

    $$\{M \text{ martingale: } M_0 \sim \mu_0, M_1 \sim \mu_1\}$$

    with respect to adapted Wasserstein distance
  - Idea I: project Brownian motion onto

    $$\{M \text{ martingale: } M_t \sim \mu_t, \ \forall t \in [0,1]\}$$

Future research directions

- Find minimal conditions for the existence of Markov mimicking martingales
- Find the appropriate sense in which a mimicking martingale is canonical
  - Bass martingale: unique projection of Brownian motion onto

    $$\{M \text{ martingale}\colon\ M_0 \sim \mu_0, M_1 \sim \mu_1\}$$

    with respect to adapted Wasserstein distance
  - Idea I: project Brownian motion onto

    $$\{M \text{ martingale}\colon\ M_t \sim \mu_t,\ \forall t \in [0,1]\}$$

  - Idea II: project M onto

    $$\text{“}\{\mathrm{d}M_t = \nabla^2 v_t(M_t)\mathrm{d}B_t\colon\ v \text{ convex},\ M_t \sim \mu_t,\ \forall t \in [0,1]\}\text{”}$$

[Conforti, Lacker, Pal 2023, Lacker 2023] — Hessian projection

Summary

Main contributions:

- Computation of adapted distance between laws of SDEs
- Strong approximation of SDEs with discontinuous and exponentially growing drift
- Proved first multidimensional Kellerer theorem

Future directions:

- Understand adapted distance for more general processes
- Prove full Kellerer theorem with uniqueness
- Applications in robust optimisation

References

J. Albin. A continuous non-Brownian motion martingale with Brownian motion marginal distributions. *Statist. Probab. Lett.*, 78(6):682–686, 2008. ISSN 0167-7152. doi: 10.1016/j.spl.2007.09.031. URL http://dx.doi.org/10.1016/j.spl.2007.09.031.

J. Backhoff-Veraguas, C. Unterberger. On the specific relative entropy between martingale diffusions on the line. *Electronic Communications in Probability*, 28(none):1 – 12, 2023. doi: 10.1214/23-ECP548. URL https://doi.org/10.1214/23-ECP548.

J. Backhoff-Veraguas, M. Beiglböck, M. Huesmann, S. Källblad. Martingale Benamou-Brenier: a probabilistic perspective. *Ann. Probab.*, 48(5):2258–2289, 2020. ISSN 0091-1798. doi: 10.1214/20-AOP1422. URL https://doi.org/10.1214/20-AOP1422.

J. Backhoff-Veraguas, S. Källblad, B. A. Robinson. Adapted Wasserstein distance between the laws of SDEs. *arXiv:2209.03243*, sep 2022.

J. Backhoff-Veraguas, M. Beiglböck, W. Schachermayer, B. Tschiderer. The structure of martingale Benamou–Brenier in multiple dimensions. *ArXiv e-prints*, 2306.11019, 2023.

M. Beiglböck, G. Lowther, G. Pammer, W. Schachermayer. Faking Brownian motion through continuous Markov martingales. *Finance Stoch., to appear*, 2023.

G. Conforti, D. Lacker, S. Pal. Projected Langevin dynamics and a gradient flow for entropic optimal transport. *arXiv preprint arXiv:2309.08598*, 2023.

A. M. G. Cox, B. A. Robinson. Optimal control of martingales in a radially symmetric environment. *Stoch. Proc. Appl.*, 159:149–198, 2023a. ISSN 0304-4149. doi: https://doi.org/10.1016/j.spa.2023.01.016. URL https://www.sciencedirect.com/science/article/pii/S030441492300025X.

A. M. G. Cox, B. A. Robinson. SDEs with no strong solution arising from a problem of stochastic control. *Electron. J. Probab.*, 28:1–24, 2023b. ISSN 1083-6489. doi: 10.1214/23-EJP995.

K. Hamza, F. Klebaner. A family of non-Gaussian martingales with Gaussian marginals. *J. Appl. Math. Stoch. Anal.*, pages Art. ID 92723, 19, 2007. ISSN 1048-9533.

D. Lacker. Independent projections of diffusions: Gradient flows for variational inference and optimal mean field approximations. *arXiv preprint arXiv:2309.13332*, 2023.

K. Oleszkiewicz. On fake Brownian motions. *Statist. Probab. Lett.*, 78(11):1251–1254, 2008. ISSN 0167-7152. doi: 10.1016/j.spl.2007.11.026. URL `http://dx.doi.org/10.1016/j.spl.2007.11.026`.

G. Pammer, B. A. Robinson, W. Schachermayer. A regularized Kellerer theorem in arbitrary dimension. *Annals of Applied Probability*, 2024. To appear.

B. A. Robinson, M. Szölgyenyi. Adapted optimal transport for SDEs with irregular coefficients. *Preprint*, 2024.