Bicausal optimal transport for SDEs with irregular coefficients

Benjamin A. Robinson (Universität Klagenfurt)

June 6, 2024

Klagenfurt–Berlin Workshop on Multiple Perspectives in Optimisation

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

Bicausal optimal transport for SDEs with irregular coefficients

Benjamin A. Robinson (Universität Klagenfurt)

June 6, 2024

Klagenfurt–Berlin Workshop on Multiple Perspectives in Optimisation

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

Joint work with

Michaela Szölgyenyi
University of Klagenfurt

Bicausal optimal transport for SDEs with irregular coefficients

Benjamin A. Robinson (Universität Klagenfurt)

June 6, 2024

Klagenfurt–Berlin Workshop on Multiple Perspectives in Optimisation

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

Joint work with

Michaela Szölgyenyi
University of Klagenfurt

Julio Backhoff-Veraguas
University of Vienna

Sigrid Källblad
KTH Stockholm

Bicausal optimal transport for SDEs with irregular coefficients

Benjamin A. Robinson (Universität Klagenfurt)

June 6, 2024

Klagenfurt–Berlin Workshop on Multiple Perspectives in Optimisation

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

Joint work with

Michaela Szölgyenyi
University of Klagenfurt

Julio Backhoff-Veraguas
University of Vienna

Sigrid Källblad
KTH Stockholm

Adapted Wasserstein distance between the laws of SDEs
(with J. Backhoff-Veraguas and S. Källblad) — arXiv:2209.03243, 2022

Bicausal optimal transport for SDEs with irregular coefficients
(with M. Szölgyenyi) — arXiv:2403.09941, 2024

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

Main result

$b, \bar{b} \colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\begin{aligned} \mathrm{d}X_t &= b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0, \\ \mathrm{d}\bar{X}_t &= \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0. \end{aligned}$$

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

Main result

$b, \bar{b} \colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0,$$
$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0.$$

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

Theorem [R., Szölgyenyi '24]

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

Main result

$b, \bar{b}\colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma}\colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0,$$
$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0.$$

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

Theorem [R., Szölgyenyi '24]

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

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

Auxiliary result

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

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t. \tag{SDE}$$

Auxiliary result

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

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t. \tag{SDE}$$

Assumption (A)

b satisfies piecewise regularity conditions and exponential growth condition,
σ is Lipschitz and non-zero at the discontinuity points of b.

Auxiliary result

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

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t. \qquad \text{(SDE)}$$

Assumption (A)

b satisfies piecewise regularity conditions and exponential growth condition,

σ is Lipschitz and non-zero at the discontinuity points of b.

Theorem [R., Szölgyenyi '24]

Strong existence, pathwise uniqueness, and moment bounds hold for (SDE) with coefficients satisfying (A).

Auxiliary result

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

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t. \qquad \text{(SDE)}$$

Assumption (A)

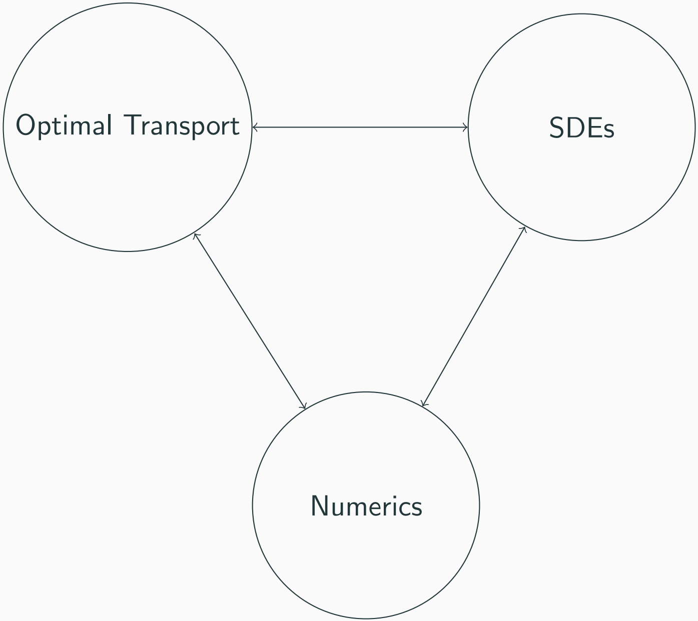
b satisfies piecewise regularity conditions and exponential growth condition,

σ is Lipschitz and non-zero at the discontinuity points of b.

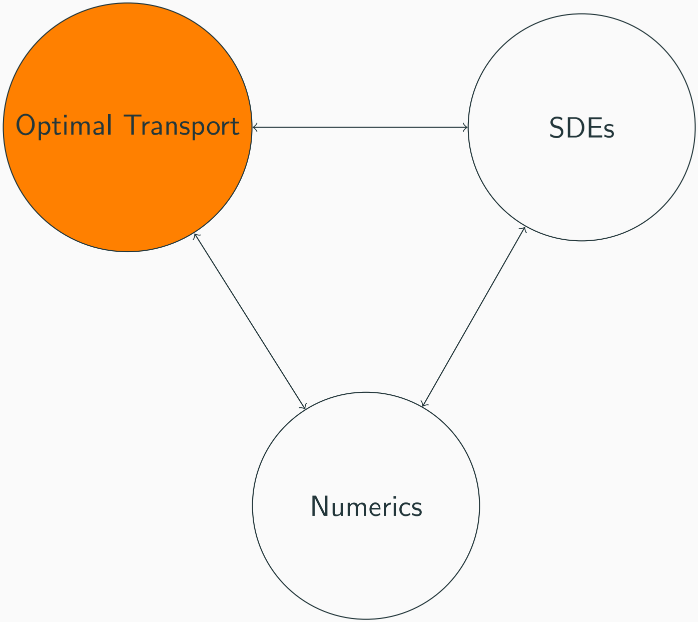
Theorem [R., Szölgyenyi '24]

Strong existence, pathwise uniqueness, and moment bounds hold for (SDE) with coefficients satisfying (A). Moreover, for a transformation-based semi-implicit Euler scheme, we obtain strong convergence rates.

Ingredients



Ingredients



Optimal transport

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

Find

$$\inf_{T\colon T_\#\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^p\right]$$

Optimal transport

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

Find

$$\inf_{T\colon T_\#\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^p\right]$$

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

Optimal transport

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

Find

$$\inf_{T\colon T_{\#}\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^p\right]$$

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

Monge (1781)

Optimal transport

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

Find

$$\inf_{T\colon T_{\#}\mu=\nu} \mathbb{E}\left[\sum_{n=1}^{N} |T_n(X) - X_n|^p\right]$$

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

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

Optimal transport

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

Find

$$\mathcal{W}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

$$\mathrm{Cpl}(\mu,\nu) = \{\pi = \mathrm{Law}(X,Y) \colon \ X \sim \mu, Y \sim \nu\}$$

Monge (1781)

Kantorovich (1942) $\rightsquigarrow$ T random: replace $(X, T(X))$ with coupling $\pi = \mathrm{Law}(X,Y)$, $X \sim \mu$, $Y \sim \nu$

Optimal transport

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

Find

$$\mathcal{W}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

$$\mathrm{Cpl}(\mu,\nu) = \{\pi = \mathrm{Law}(X,Y) \colon\; X \sim \mu, Y \sim \nu\}$$

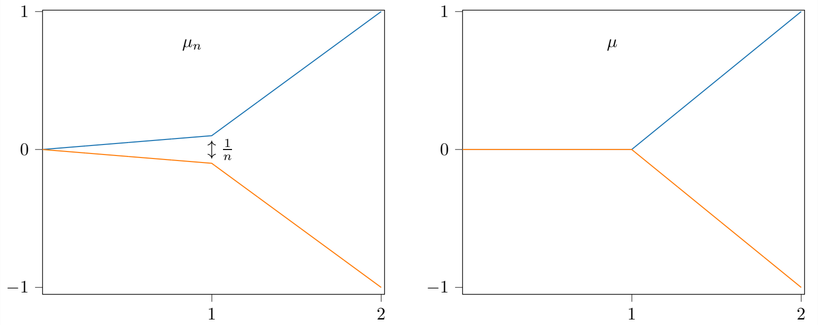
Monge (1781)

Kantorovich (1942) $\rightsquigarrow$ T random: replace $(X, T(X))$ with coupling $\pi = \mathrm{Law}(X,Y)$, $X \sim \mu$, $Y \sim \nu$

Wasserstein distance metrises usual weak topology

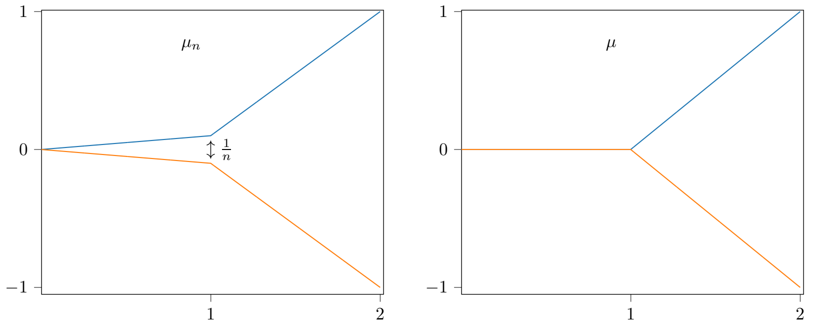
Example

[Backhoff-Veraguas, Bartl, Beiglböck, Eder '20], [Aldous '81]



Example

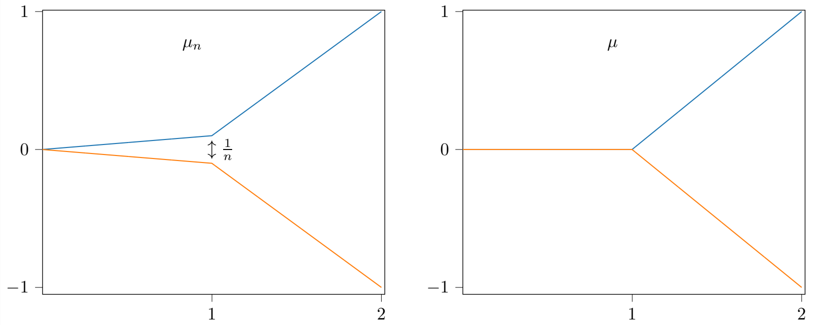
[Backhoff-Veraguas, Bartl, Beiglböck, Eder '20], [Aldous '81]



"Can get rich" "Cannot get rich"

Example

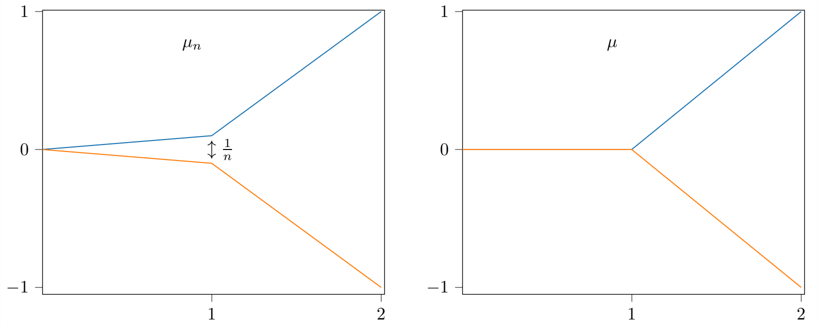
[Backhoff-Veraguas, Bartl, Beiglböck, Eder '20], [Aldous '81]



$$V_n := \sup_{\tau} \mathbb{E}^{\mu_n}[X_\tau] \approx \frac{1}{2} \qquad V := \sup_{\tau} \mathbb{E}^{\mu}[X_\tau] = 0$$

Example

[Backhoff-Veraguas, Bartl, Beiglböck, Eder '20], [Aldous '81]

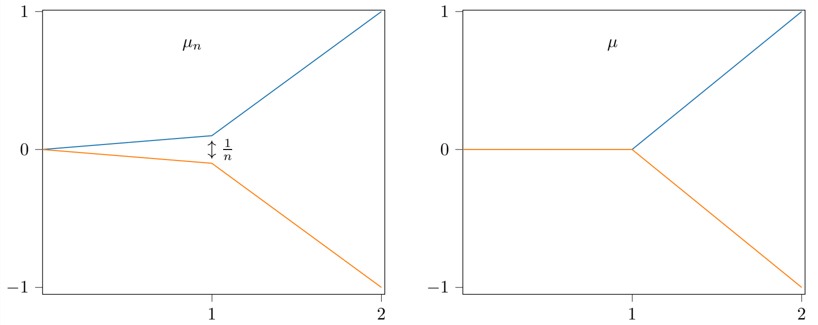


$$V_n := \sup_\tau \mathbb{E}^{\mu_n}[X_\tau] \approx \frac{1}{2} \qquad V := \sup_\tau \mathbb{E}^{\mu}[X_\tau] = 0$$

$$V_n \not\to V$$

Example

[Backhoff-Veraguas, Bartl, Beiglböck, Eder '20], [Aldous '81]



$$V_n \not\to V \quad \text{but} \quad \mu_n \rightharpoonup \mu$$

Adapted topology

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

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

Adapted topology

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

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

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

$$\mathrm{Cpl}_{\mathrm{bc}}(\mu, \nu) = \{\pi \in \mathrm{Cpl}(\mu, \nu) \colon \; \pi \text{ bicausal}\}$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

$$\mathrm{Cpl}_{\mathrm{bc}}(\mu, \nu) = \{\pi \in \mathrm{Cpl}(\mu, \nu) \colon \ \pi \text{ bicausal}\}$$

More general cost functions $\rightsquigarrow$ bicausal optimal transport

$$\inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N c_n(X_n, Y_n) \right]$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

$$\mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu) = \{\pi \in \mathrm{Cpl}(\mu,\nu) \colon \ \pi \text{ bicausal}\}$$

More general cost functions $\rightsquigarrow$ bicausal optimal transport

$$\inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N c_n(X_n, Y_n) \right]$$

c_n continuous, polynomial growth, quasi-monotone

$$c_n(x,y) + c_n(x',y') - c_n(x,y') - c_n(x',y) \geq 0, \quad \forall\, x \leq x', y \leq y'$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

$$\mathrm{Cpl}_{\mathrm{bc}}(\mu, \nu) = \{\pi \in \mathrm{Cpl}(\mu, \nu) \colon \ \pi \text{ bicausal}\}$$

More general cost functions $\rightsquigarrow$ bicausal optimal transport

$$\inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N c_n(X_n, Y_n) \right]$$

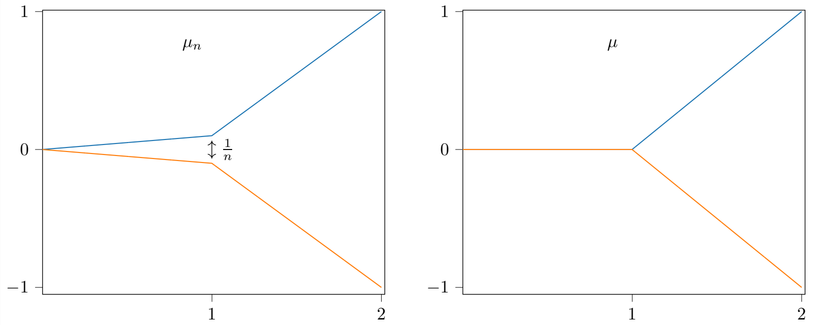
c_n continuous, polynomial growth, quasi-monotone

$$c_n(x, y) + c_n(x', y') - c_n(x, y') - c_n(x', y) \geq 0, \quad \forall\, x \leq x', y \leq y'$$

Acciaio, Aldous, Backhoff-Veraguas, Bartl, Beiglböck, Bion-Nadal, Eder, Hellwig, Källblad, Pammer, Pflug, Pichler, Talay, Zalaschko, ...

Example revisited

[Backhoff-Veraguas, Bartl, Beiglböck, Eder '20], [Aldous '81]



$$V_n := \sup_\tau \mathbb{E}^{\mu_n}[X_\tau] \approx \frac{1}{2} \qquad V := \sup_\tau \mathbb{E}^{\mu}[X_\tau] = 0$$

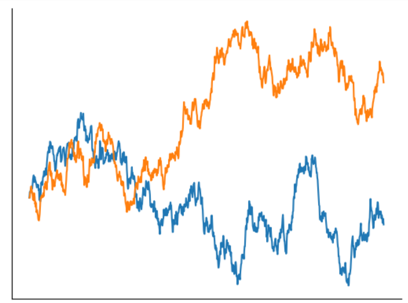
$V_n \not\to V$ and $\mathcal{AW}_p(\mu_n, \mu) \not\to 0$

Continuous time

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

$$\mu,\nu \in \mathcal{P}(\Omega) \quad \rightsquigarrow \quad \mathcal{W}_p^p(\mu,\nu) := \inf_{\pi\in\mathrm{Cpl}(\mu,\nu)} \mathbb{E}^\pi\left[\int_0^T |\omega_t - \bar{\omega}_t|^p \mathrm{d}t\right]$$

$$\mathrm{Cpl}(\mu,\nu) = \{\pi = \mathrm{Law}(X,Y)\colon\ X\sim\mu, Y\sim\nu\}$$

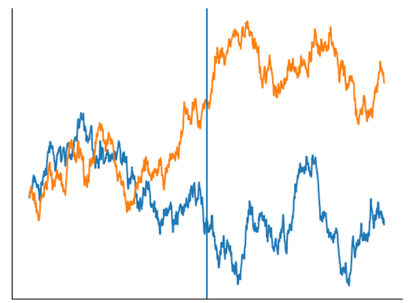


Continuous time

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

$$\mu,\nu \in \mathcal{P}(\Omega) \quad \rightsquigarrow \quad \mathcal{W}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}(\mu,\nu)} \mathbb{E}^\pi \left[\int_0^T |\omega_t - \bar{\omega}_t|^p \mathrm{d}t \right]$$

$$\mathrm{Cpl}(\mu,\nu) = \{\pi = \mathrm{Law}(X,Y) \colon\; X \sim \mu, Y \sim \nu\}$$

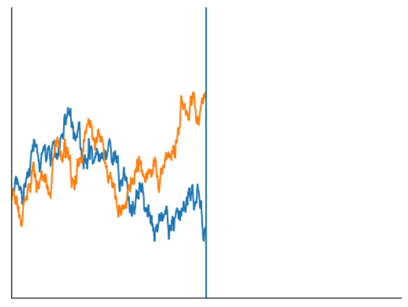


Continuous time

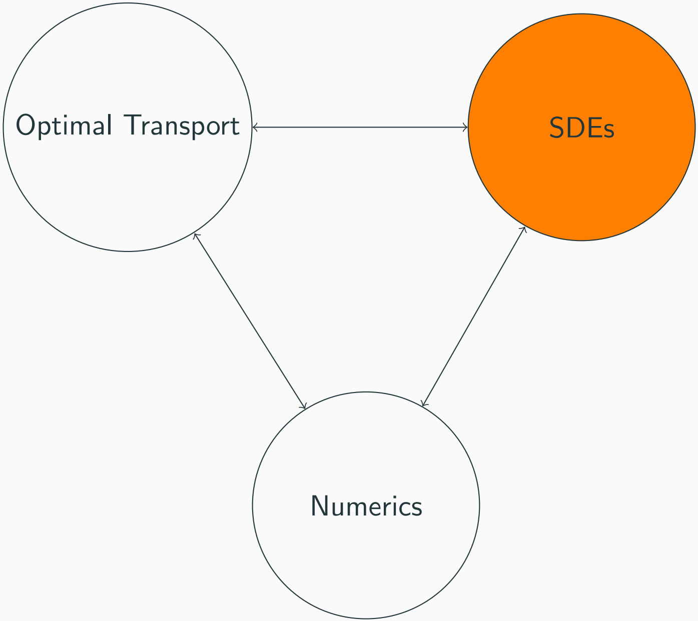
Similar definition of adapted Wasserstein distance in continuous time w.r.t. L^p norm on $\Omega := C([0,T],\mathbb{R})$

$$\mu,\nu \in \mathcal{P}(\Omega) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\int_0^T |\omega_t - \bar{\omega}_t|^p \mathrm{d}t \right]$$

$$\mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu) = \{\pi \in \mathrm{Cpl}(\mu,\nu) \colon \ \pi \text{ bicausal}\}$$



Ingredients



Main result

$b, \bar{b}\colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma}\colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0,$$
$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0.$$

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

Main result

$b, \bar{b} \colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0,$$
$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0.$$

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

Theorem [R., Szölgyenyi '24]

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

Main result

$b, \bar{b} \colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\begin{aligned} \mathrm{d}X_t &= b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0, \\ \mathrm{d}\bar{X}_t &= \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0. \end{aligned}$$

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

Theorem [R., Szölgyenyi ’24]

Under “weak assumptions” on the coefficients, we can compute the adapted Wasserstein distance

Main result

$b, \bar{b} \colon [0,T] \times \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} \colon [0,T] \times \mathbb{R} \to [0,\infty)$,

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_t, \quad X_0 = x_0,$$
$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}W_t, \quad \bar{X}_0 = x_0.$$

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

Coupling SDEs

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

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

Optimising over bicausal couplings $\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu, \nu)$

$\Leftrightarrow$

Optimising over correlations between B, W

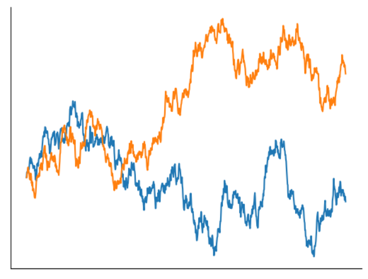
Coupling SDEs

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

Optimising over bicausal couplings $\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu, \nu)$
$\Leftrightarrow$
Optimising over correlations between B, W

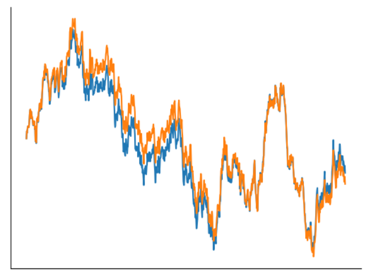
Product coupling

B, W independent

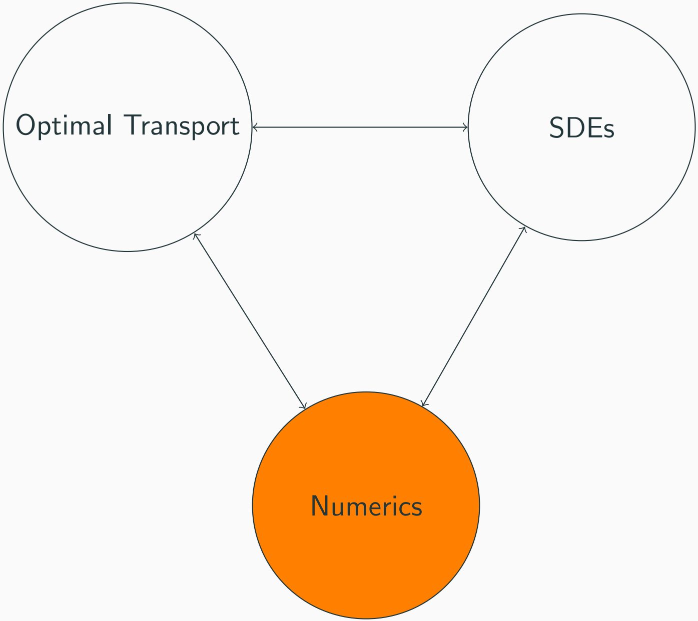


Synchronous coupling

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



Ingredients



Proof of main result

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

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

Proof of main result

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

1. Discretise SDEs;

Proof of main result

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

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

Proof of main result

$$dX_t = b_t(X_t)dt + \sigma_t(X_t)dB_t,\ X_0 = x \ \rightsquigarrow\ \mathrm{Law}(X) = \mu$$
$$d\bar{X}_t = \bar{b}_t(\bar{X}_t)dt + \bar{\sigma}_t(\bar{X}_t)dW_t,\ \bar{X}_0 = x \ \rightsquigarrow\ \mathrm{Law}(\bar{X}) = \nu$$

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

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

Proof of main result

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

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

Key result in discrete time

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

Key result in discrete time

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

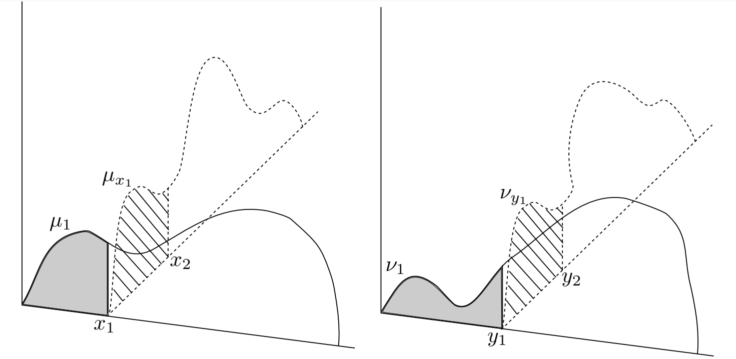
Knothe–Rosenblatt rearrangement

Key result in discrete time

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

Knothe–Rosenblatt rearrangement

— generalisation of monotone rearrangement

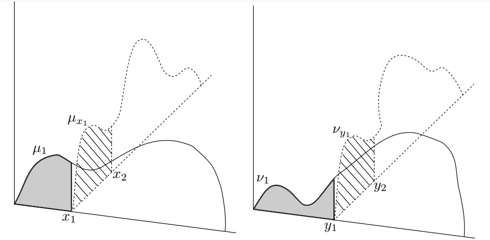


Key result in discrete time

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl_{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

Knothe–Rosenblatt rearrangement

$$Y_k = T_k^{\mathrm{KR}}(X_1, \ldots, X_k) = F^{-1}_{\nu_{Y_1,\ldots,Y_{k-1}}} \circ F_{\mu_{X_1,\ldots,X_{k-1}}}(X_k),$$



Key result in discrete time

$$\mu, \nu \in \mathcal{P}(\mathbb{R}^N) \quad \rightsquigarrow \quad \mathcal{AW}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\sum_{n=1}^N |X_n - Y_n|^p \right]$$

Knothe–Rosenblatt rearrangement

$$Y_k = T_k^{\mathrm{KR}}(X_1, \ldots, X_k) = F^{-1}_{\nu_{Y_1,\ldots,Y_{k-1}}} \circ F_{\mu_{X_1,\ldots,X_{k-1}}}(X_k),$$

Theorem [Rüschendorf '85] [Posch '23]

For μ, ν stochastically co-monotone, the unique optimiser is the Knothe–Rosenblatt rearrangement.

This induces the adapted weak topology.

Proof of main result

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

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

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t$$

Euler scheme

$$X_0^h = X_0,$$
$$X_t^h = X_{kh}^h + b(X_{kh})(t - kh), \quad t \in (kh, (k+1)h].$$

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t$$

Euler–Maruyama scheme

$$X_0^h = X_0,$$

$$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + W_t - W_{kh}, \quad t \in (kh, (k+1)h].$$

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t$$

Euler–Maruyama scheme

$$X_0^h = X_0,$$
$$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + W_t - W_{kh}, \quad t \in (kh, (k+1)h].$$

Write $X_k^h := X_{kh}^h$ and $\mu^h = \mathrm{Law}((X_k^h)_k)$.

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t$$

Euler–Maruyama scheme

$$X_0^h = X_0,$$
$$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + W_t - W_{kh}, \quad t \in (kh, (k+1)h].$$

Write $X_k^h := X_{kh}^h$ and $\mu^h = \mathrm{Law}((X_k^h)_k)$.

Remark

$X_k^h \mapsto X_{(k+1)}^h$ is increasing if b is Lipschitz, $h \ll 1$

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t$$

Euler–Maruyama scheme

$$X_0^h = X_0,$$
$$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + W_t - W_{kh}, \quad t \in (kh, (k+1)h].$$

Write $X_k^h := X_{kh}^h$ and $\mu^h = \mathrm{Law}((X_k^h)_k)$.

Remark

$X_k^h \mapsto X_{(k+1)}^h$ is increasing if b is Lipschitz, $h \ll 1$

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

Monotone Euler–Maruyama scheme

$X_0^h = X_0,$

$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + \sigma(X_{kh})(W_t^h - W_{kh}^h),\ t \in (kh, (k+1)h].$

$W_t^h - W_{kh}^h = W_{t \wedge \tau_k^h} - W_{kh}, \quad \tau_k^h \coloneqq \inf\{t > kh \colon\ |W_t - W_{kh}| > A_h|\}$

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

Monotone Euler–Maruyama scheme

$X_0^h = X_0,$

$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + \sigma(X_{kh})(W_t^h - W_{kh}^h),\ t \in (kh, (k+1)h].$

$W_t^h - W_{kh}^h = W_{t\wedge\tau_k^h} - W_{kh}, \quad \tau_k^h \coloneqq \inf\{t > kh\colon\ |W_t - W_{kh}| > A_h|\}$

Write $\quad X_k^h \coloneqq X_{kh}^h \quad$ and $\quad \mu^h = \mathrm{Law}((X_k^h)_k)$.

A monotone numerical scheme

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

Monotone Euler–Maruyama scheme

$X_0^h = X_0,$

$X_t^h = X_{kh}^h + b(X_{kh})(t - kh) + \sigma(X_{kh})(W_t^h - W_{kh}^h),\ t \in (kh, (k+1)h].$

$W_t^h - W_{kh}^h = W_{t \wedge \tau_k^h} - W_{kh}, \quad \tau_k^h := \inf\{t > kh \colon\ |W_t - W_{kh}| > A_h|\}$

Write $\quad X_k^h := X_{kh}^h \quad$ and $\quad \mu^h = \mathrm{Law}((X_k^h)_k).$

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

For b, σ Lipschitz, the monotone Euler–Maruyama scheme is stochastically increasing.

Hence the Knothe–Rosenblatt rearrangement is optimal for μ^h, ν^h.

Proof of main result

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

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

Proof of main result

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

Theorem [R., Szölgyenyi '24]

Under "weak assumptions" on the coefficients, we can compute the adapted Wasserstein distance by

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

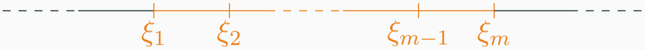
1. Discretise SDEs;
2. Solve discrete-time bicausal optimal transport problem;
3. Pass to a limit.

Transformation-based semi-implicit Euler scheme

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

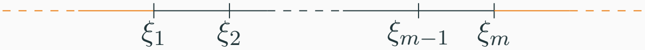


Transformation-based semi-implicit Euler scheme

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



Transformation-based semi-implicit Euler scheme

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 $k \in \{1, \ldots, m\}$ — no uniform ellipticity

Transformation-based semi-implicit Euler scheme

Under Assumption (A) , the scheme is constructed as follows:

1. Apply the transformation G from [Leobacher, Szölgyenyi '17] to (SDE),

$$Z = G(X)$$

$$\mathrm{d}Z_t = \tilde{b}(Z_t)\mathrm{d}t + \tilde{\sigma}(Z_t)\mathrm{d}W_t$$

$\tilde{b}$ one-sided Lipschitz, exponential growth, locally Lipschitz, a.c.

$\tilde{\sigma}$ Lipschitz

Transformation-based semi-implicit Euler scheme

Under Assumption (A) , the scheme is constructed as follows:

1. Apply the transformation G from [Leobacher, Szölgyenyi '17] to (SDE),
2. Apply a semi-implicit Euler scheme with truncated Brownian increments to the transformed SDE,

$$Z = G(X)$$

$$\mathrm{d}Z_t = \tilde{b}(Z_t)\mathrm{d}t + \tilde{\sigma}(Z_t)\mathrm{d}W_t$$

$$Z^h_{(k+1)h} = Z^h_{kh} + \tilde{b}(Z^h_{(k+1)h}) \cdot h + \tilde{\sigma}(Z^h_{kh})(W^h_{(k+1)h} - W^h_{kh})$$

Transformation-based semi-implicit Euler scheme

Under Assumption (A) , the scheme is constructed as follows:

1. Apply the transformation G from [Leobacher, Szölgyenyi '17] to (SDE),
2. Apply a semi-implicit Euler scheme with truncated Brownian increments to the transformed SDE,
3. Transform back.

$$Z = G(X)$$

$$\mathrm{d}Z_t = \tilde{b}(Z_t)\mathrm{d}t + \tilde{\sigma}(Z_t)\mathrm{d}W_t$$

$$Z^h_{(k+1)h} = Z^h_{kh} + \tilde{b}(Z^h_{(k+1)h}) \cdot h + \tilde{\sigma}(Z^h_{kh})(W^h_{(k+1)h} - W^h_{kh})$$

$$X^h_{kh} = G^{-1}(Z^h_{kh})$$

Transformation-based semi-implicit Euler scheme

Under Assumption (A) , the scheme is constructed as follows:

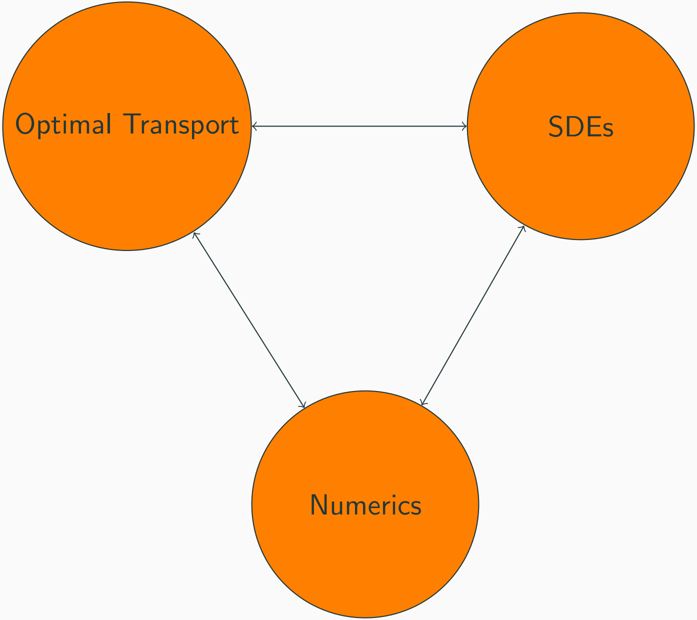
1. Apply the transformation G from [Leobacher, Szölgyenyi '17] to (SDE),
2. Apply a semi-implicit Euler scheme with truncated Brownian increments to the transformed SDE,
3. Transform back.

Theorem [R., Szölgyenyi '24]

Let (b, σ) satisfy Assumption (A). Then (SDE) admits a unique strong solution and, for all $p \geq 1$, there exists $C_p \geq 0$ such that

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

Ingredients



Main result

Assumptions
(A) discontinuous drift with exponential growth (time-homog.);

Main result

Assumptions

(A) discontinuous drift with exponential growth (time-homog.);

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

Main result

Assumptions

(A) discontinuous drift with exponential growth (time-homog.);

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

(C) continuous coefficients, linear growth, pathwise uniqueness.

Main result

Assumptions

(A) discontinuous drift with exponential growth (time-homog.);

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

(C) continuous coefficients, linear growth, pathwise uniqueness.

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

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

Main result

Assumptions

(A) discontinuous drift with exponential growth (time-homog.);

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

(C) continuous coefficients, linear growth, pathwise uniqueness.

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

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

Main 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

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

Main result

Assumptions

(A) discontinuous drift with exponential growth (time-homog.);

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

(C) continuous coefficients, linear growth, pathwise uniqueness.

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

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

Main 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 '23]
- Application to uniqueness of mimicking martingales

Summary

- We compute adapted Wasserstein distance between SDEs with irregular coefficients
- We prove strong convergence rates for a numerical scheme for SDEs with discontinuous and exponentially growing drift

References: