Adapted Wasserstein distance between the laws of SDEs

Benjamin A. Robinson

University of Vienna

March 9, 2023 — GPSD23, Essen

Joint work with

Julio Backhoff-Veraguas (University of Vienna) and Sigrid Källblad *(KTH Stockholm)*

Adapted Wasserstein distance between the laws of SDEs

Benjamin A. Robinson

University of Vienna

March 9, 2023 — GPSD23, Essen

Joint work with

Julio Backhoff-Veraguas (University of Vienna) and Sigrid Källblad *(KTH Stockholm)*

Adapted Wasserstein distance between the laws of SDEs

Benjamin A. Robinson

University of Vienna

March 9, 2023 — GPSD23, Essen

Joint work with

Julio Backhoff-Veraguas (University of Vienna) and Sigrid Källblad *(KTH Stockholm)*

http://arxiv.org/abs/2209.03243

Distances between stochastic processes

$(X_t)_{t\in[0,1]}$, $(Y_t)_{t\in[0,1]}$ continuous-time real-valued processes

$\rightsquigarrow$ μ, ν probability measures on $\Omega := C([0,1], \mathbb{R})$

Distances between stochastic processes

$(X_t)_{t\in[0,1]}$, $(Y_t)_{t\in[0,1]}$ continuous-time real-valued processes

$\rightsquigarrow$ μ, ν probability measures on $\Omega := C([0,1], \mathbb{R})$

How to choose a “good” distance $d(\mu, \nu)$?

Distances between stochastic processes

$(X_t)_{t\in[0,1]}$, $(Y_t)_{t\in[0,1]}$ continuous-time real-valued processes

$\rightsquigarrow$ μ, ν probability measures on $\Omega := C([0,1], \mathbb{R})$

How to choose a "good" distance $d(\mu, \nu)$?

E.g. Wasserstein distance $\mathcal{W}_p$ — from optimal transport

Distances between stochastic processes

$(X_t)_{t\in[0,1]}$, $(Y_t)_{t\in[0,1]}$ continuous-time real-valued processes

$\rightsquigarrow$ μ, ν probability measures on $\Omega := C([0,1], \mathbb{R})$

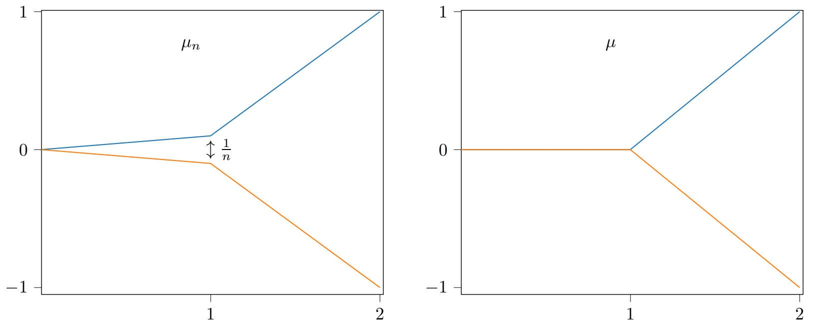
How to choose a "good" distance $d(\mu, \nu)$?

E.g. Wasserstein distance $\mathcal{W}_p$:

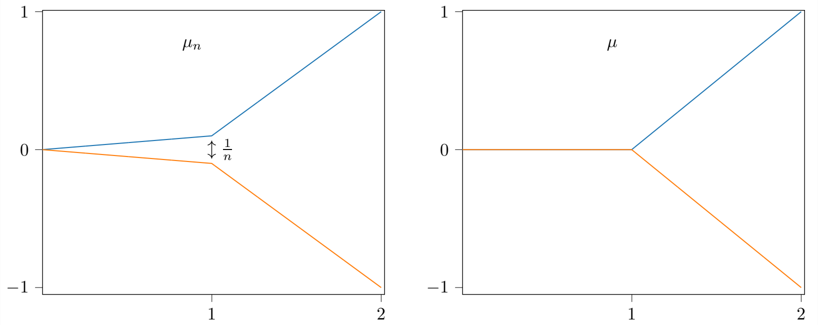
$$\mathcal{W}_p^p(\mu, \nu) := \inf_{\pi \in \mathrm{Cpl}(\mu,\nu)} \mathbb{E}^\pi \left[\int_0^1 |\omega_t - \bar{\omega}_t|^p \, \mathrm{d}t \right]$$

$$\mathrm{Cpl}(\mu, \nu) := \{\pi \in \mathcal{P}(\Omega \times \Omega) \colon \ \pi \text{ has marginals } \mu, \nu\}$$

Distances between stochastic processes

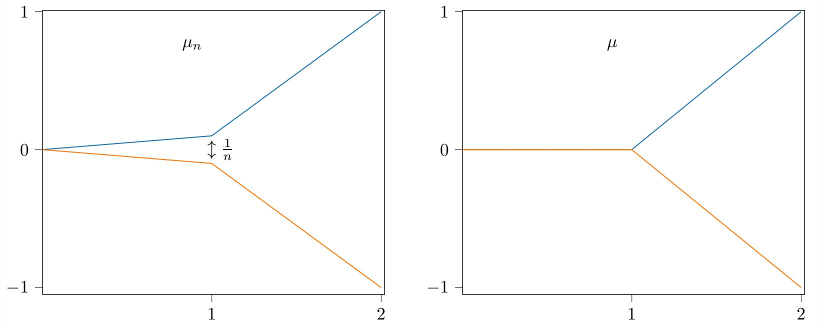


Distances between stochastic processes



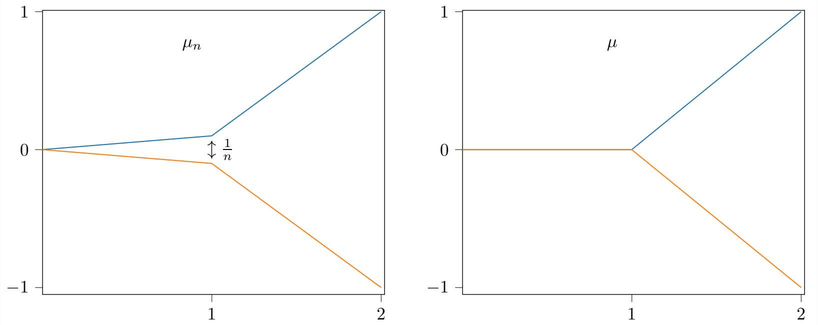
$$V_n := \sup_\tau \mathbb{E}^{\mu_n}[X_\tau] \qquad V := \sup_\tau \mathbb{E}^{\mu}[X_\tau]$$

Distances between stochastic processes



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

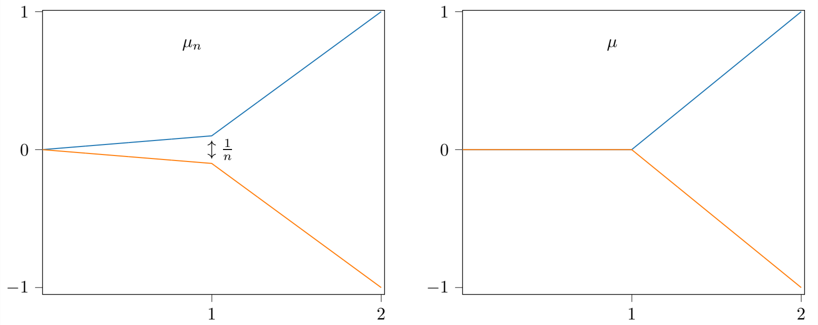
Distances between stochastic processes



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

$$V_n \not\to V$$

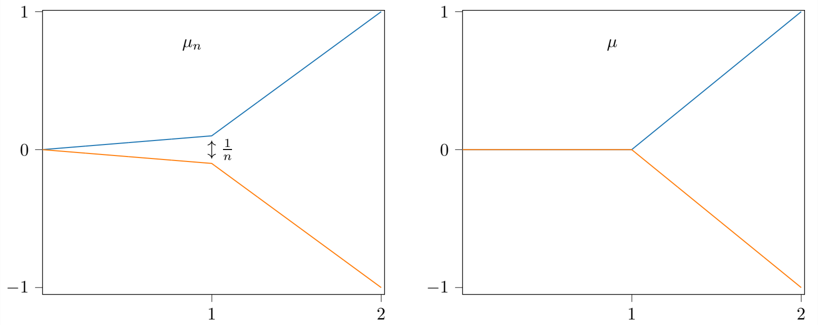
Distances between stochastic processes



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

$$V_n \not\to V \quad \text{but} \quad \mathcal{W}_p(\mu_n, \mu) \to 0$$

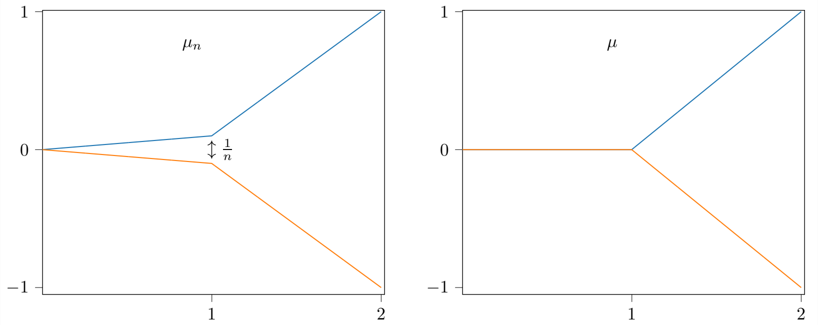
Distances between stochastic processes



Want

$$d(\mu_n, \mu) \not\to 0$$

Distances between stochastic processes



Want

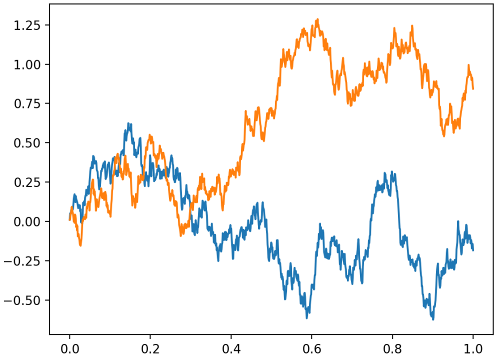
$$d(\mu_n, \mu) \not\to 0$$

E.g. Acciaio, Aldous, Backhoff-Veraguas, Bartl, Beiglböck, Bion-Nadal, Eder, Hellwig, Lassalle, Pammer, Pflug, Pichler, Talay, among others ...

Coupling SDEs

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

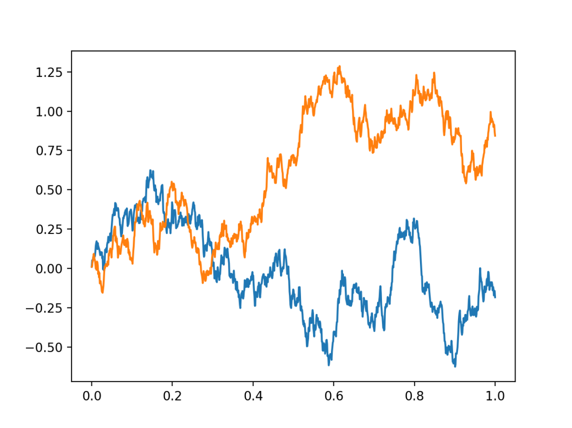
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t \qquad \rightsquigarrow \qquad \nu.$$



Coupling SDEs

Usual couplings

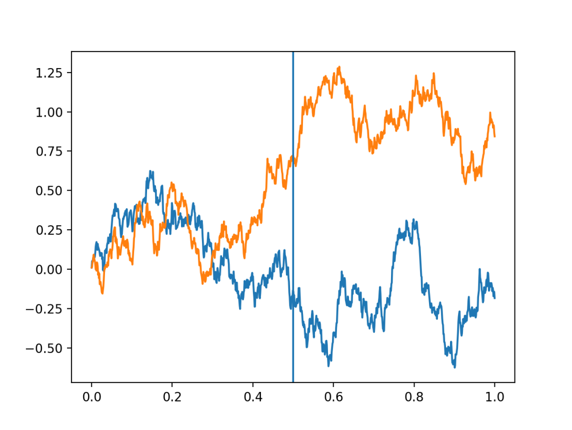
$$\mathrm{Cpl}(\mu,\nu) := \{\pi \in \mathcal{P}(\Omega \times \Omega) \colon \ \pi \text{ has marginals } \mu,\nu\}.$$



Coupling SDEs

Usual couplings

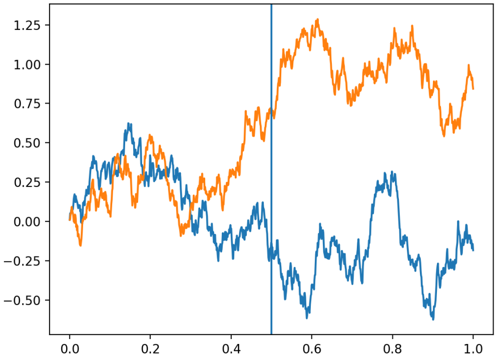
$$\mathrm{Cpl}(\mu, \nu) := \{\pi \in \mathcal{P}(\Omega \times \Omega) \colon \ \pi \text{ has marginals } \mu, \nu\}.$$



Coupling SDEs

Introduce

$$\mathrm{Cpl_{bc}}(\mu,\nu) := \{\pi \in \mathrm{Cpl}(\mu,\nu) \colon \; \pi \text{ is bi-causal}\}$$



Adapted Wasserstein distance

Introduce

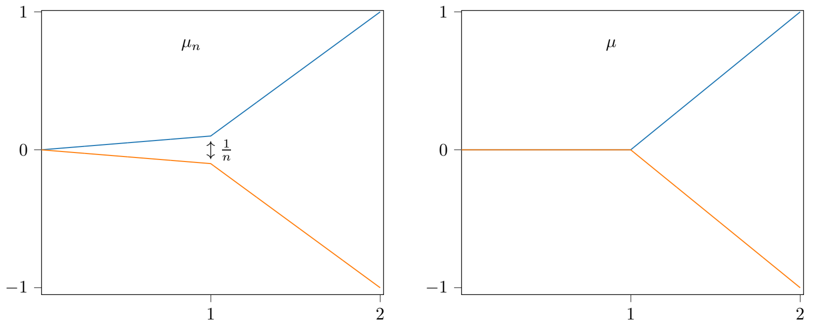
$$\mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu) := \{\pi \in \mathrm{Cpl}(\mu,\nu) \colon\ \pi \text{ is bi-causal}\}$$

The problem:

Find adapted Wasserstein distance:

$$\mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\int_0^1 |\omega_t - \bar{\omega}_t|^p \, \mathrm{d}t \right].$$

Adapted Wasserstein distance



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

Adapted Wasserstein distance

$$\begin{aligned} \mathrm{d}X_t &= b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \qquad &\rightsquigarrow \qquad &\mu \\ \mathrm{d}\bar{X}_t &= \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t \qquad &\rightsquigarrow \qquad &\nu. \end{aligned}$$

The problem:

Find

$$\mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\int_0^1 |\omega_t - \bar{\omega}_t|^p \, \mathrm{d}t \right].$$

Adapted Wasserstein distance

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \qquad \rightsquigarrow \qquad \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t \qquad \rightsquigarrow \qquad \nu.$$

The problem:

Find

$$\mathcal{AW}_p^p(\mu,\nu) := \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu,\nu)} \mathbb{E}^\pi \left[\int_0^1 |\omega_t - \bar{\omega}_t|^p \, \mathrm{d}t \right].$$

Theorem 1 [Backhoff-Veraguas, Källblad, R. '22]

Optimising over bi-causal couplings

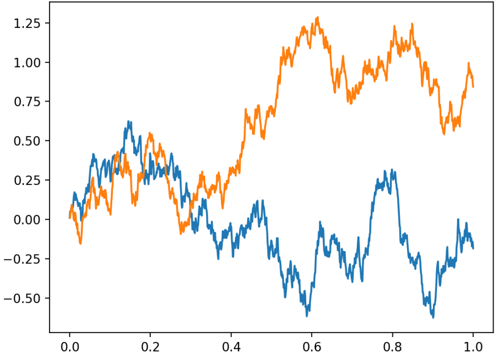
$\Leftrightarrow$

Optimising over correlations between $W, \bar{W}$.

Adapted Wasserstein distance

Example

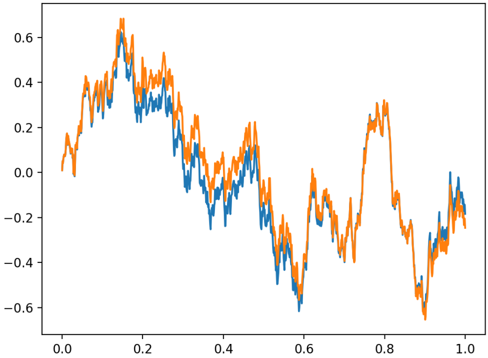
Product coupling — $W, \bar{W}$ independent



Adapted Wasserstein distance

Synchronous coupling

Choose the same driving Brownian motion $W = \bar{W}$.



Adapted Wasserstein distance

$$\begin{aligned} \mathrm{d}X_t &= b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t &\quad \rightsquigarrow \quad& \mu \\ \mathrm{d}\bar{X}_t &= \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t &\quad \rightsquigarrow \quad& \nu. \end{aligned}$$

Theorem 2 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that the coefficients are continuous with linear growth and that pathwise uniqueness holds.

Then the synchronous coupling is optimal for $\mathcal{AW}_p(\mu, \nu)$.

Adapted Wasserstein distance

$$\begin{aligned} \mathrm{d}X_t &= b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t && \rightsquigarrow && \mu \\ \mathrm{d}\bar{X}_t &= \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t && \rightsquigarrow && \nu. \end{aligned}$$

Theorem 2 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that the coefficients are continuous with linear growth and that pathwise uniqueness holds.

Then the synchronous coupling is optimal for $\mathcal{AW}_p(\mu, \nu)$.

[Bion-Nadal, Talay '19]

For smooth coefficients using PDE methods

Adapted Wasserstein distance

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \qquad \rightsquigarrow \qquad \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t \qquad \rightsquigarrow \qquad \nu.$$

Theorem 2 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that the coefficients are continuous with linear growth and that pathwise uniqueness holds.

Then the synchronous coupling is optimal for $\mathcal{AW}_p(\mu,\nu)$.

Example

If all coefficients are Lipschitz, the synchronous coupling is optimal.

Discretisation

$$\mu_n, \nu_n \in \mathcal{P}(\mathbb{R}^n) \quad \rightsquigarrow \quad \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu_n, \nu_n)} \mathbb{E}^\pi \left[\sum_{k=1}^n |x_k - y_k|^p \right].$$

Discretisation

$$\mu_n, \nu_n \in \mathcal{P}(\mathbb{R}^n) \quad \rightsquigarrow \quad \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu_n, \nu_n)} \mathbb{E}^\pi \left[\sum_{k=1}^n |x_k - y_k|^p \right].$$

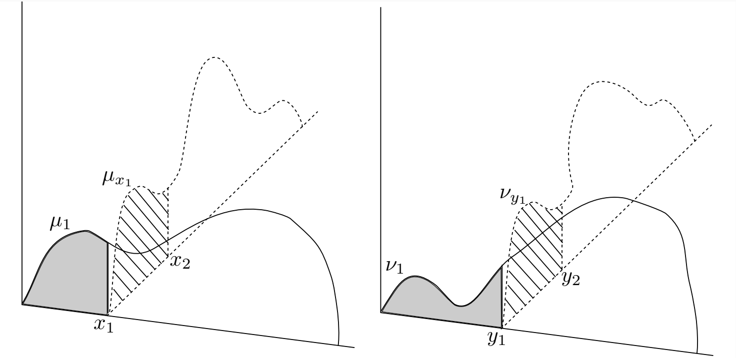
Knothe–Rosenblatt rearrangement

Discretisation

$$\mu_n, \nu_n \in \mathcal{P}(\mathbb{R}^n) \quad \rightsquigarrow \quad \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu_n, \nu_n)} \mathbb{E}^\pi \left[\sum_{k=1}^n |x_k - y_k|^p \right].$$

Knothe–Rosenblatt rearrangement

— generalisation of monotone rearrangement



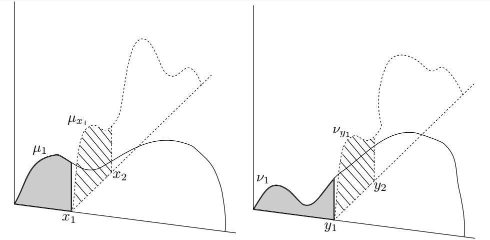
Discretisation

$$\mu_n, \nu_n \in \mathcal{P}(\mathbb{R}^n) \quad \rightsquigarrow \quad \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu_n,\nu_n)} \mathbb{E}^\pi \left[\sum_{k=1}^n |x_k - y_k|^p \right].$$

Knothe–Rosenblatt rearrangement

$$U_k \overset{iid}{\sim} \mathrm{Unif}(0,1), \quad X_k = F^{-1}_{\mu_{X_1,\dots,X_{k-1}}}(U_k), \quad Y_k = F^{-1}_{\nu_{Y_1,\dots,Y_{k-1}}}(U_k),$$

$$\pi^{\mathrm{KR}}(\mu_n, \nu_n) := \mathrm{Law}(X, Y).$$



Discretisation

$$\mu_n, \nu_n \in \mathcal{P}(\mathbb{R}^n) \quad \rightsquigarrow \quad \inf_{\pi \in \mathrm{Cpl}_{\mathrm{bc}}(\mu_n, \nu_n)} \mathbb{E}^\pi \left[\sum_{k=1}^n |x_k - y_k|^p \right].$$

Knothe–Rosenblatt rearrangement

$$U_k \overset{iid}{\sim} \mathrm{Unif}(0,1), \quad X_k = F^{-1}_{\mu_{X_1,\dots,X_{k-1}}}(U_k), \quad Y_k = F^{-1}_{\nu_{Y_1,\dots,Y_{k-1}}}(U_k),$$

$$\pi^{\mathrm{KR}}(\mu_n, \nu_n) := \mathrm{Law}(X, Y).$$

Theorem [Rüschendorf ’85]

If μ_n and ν_n are both stochastically increasing, then the unique optimiser is the Knothe–Rosenblatt coupling $\pi^{\mathrm{KR}}(\mu_n, \nu_n)$.

Adapted Wasserstein distance

$$\begin{aligned} \mathrm{d}X_t &= b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t && \rightsquigarrow && \mu \\ \mathrm{d}\bar{X}_t &= \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t && \rightsquigarrow && \nu. \end{aligned}$$

Theorem 2 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that the coefficients are continuous with linear growth and that pathwise uniqueness holds. Then the synchronous coupling is optimal for $\mathcal{AW}_p(\mu, \nu)$.

Under additional conditions, $\mathcal{AW}_p(\mu_n, \nu_n) \to \mathcal{AW}_p(\mu, \nu)$.

Adapted Wasserstein distance

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \qquad \rightsquigarrow \qquad \mu$$
$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}\bar{W}_t \qquad \rightsquigarrow \qquad \nu.$$

Theorem 2 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that the coefficients are continuous with linear growth and that pathwise uniqueness holds. Then the synchronous coupling is optimal for $\mathcal{AW}_p(\mu, \nu)$.

Under additional conditions, $\mathcal{AW}_p(\mu_n, \nu_n) \to \mathcal{AW}_p(\mu, \nu)$.

"Synchronous is continuous-time limit of Knothe–Rosenblatt"

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, $h \ll 1$

Corollary

The unique discrete-time bi-causal optimal coupling between μ^h, ν^h is the Knothe–Rosenblatt coupling.

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

$X_0^h = X_0,$

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

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

$X_0^h = X_0,$

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

Write $\quad X_k^h := X_{kh}^h \quad$ 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

$X_0^h = X_0,$

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

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, σ 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

$X_0^h = X_0,$

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

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, σ is Lipschitz, $h \ll 1$

Corollary

The unique discrete-time bi-causal optimal coupling between μ^h, ν^h is the Knothe–Rosenblatt coupling.

Stability

Theorem 3 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that $(W, \bar{W})$ ρ-correlated induces an optimal coupling for $\mathcal{AW}_p(\mu^h, \nu^h)$, for all $h > 0$.

Stability

Theorem 3 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that $(W, \bar{W})$ ρ-correlated induces an optimal coupling for $\mathcal{AW}_p(\mu^h, \nu^h)$, for all $h > 0$.

Then $(W, \bar{W})$ also induces an optimal coupling for the limiting problem $\mathcal{AW}_p(\mu, \nu)$.

Stability

Theorem 3 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that $(W, \bar{W})$ ρ-correlated induces an optimal coupling for $\mathcal{AW}_p(\mu^h, \nu^h)$, for all $h > 0$.

Then $(W, \bar{W})$ also induces an optimal coupling for the limiting problem $\mathcal{AW}_p(\mu, \nu)$.

Corollary

$$\mathcal{AW}_p(\mu^h, \nu^h) \to \mathcal{AW}_p(\mu, \nu).$$

Stability

Theorem 3 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that $(W, \bar{W})$ ρ-correlated induces an optimal coupling for $\mathcal{AW}_p(\mu^h, \nu^h)$, for all $h > 0$.

Then $(W, \bar{W})$ also induces an optimal coupling for the limiting problem $\mathcal{AW}_p(\mu, \nu)$.

Corollary

$$\mathcal{AW}_p(\mu^h, \nu^h) \to \mathcal{AW}_p(\mu, \nu).$$

Theorem 2 [Backhoff-Veraguas, Källblad, R. '22]

Suppose that the coefficients are continuous with linear growth and that pathwise uniqueness holds.

Then the synchronous coupling is optimal for $\mathcal{AW}_p(\mu, \nu)$.

Additional results

- Stability of (degenerate) correlated SDEs;
- Equivalence of topologies on a compact set;
- Extension to SDEs with irregular drifts — *work in progress with Michaela Szölgyenyi*
- Convergence of optimisers — *work in progress with Julio Backhoff and Sigrid Källblad*

Additional results

- Stability of (degenerate) correlated SDEs;
- Equivalence of topologies on a compact set;
- Extension to SDEs with irregular drifts — *work in progress with Michaela Szölgyenyi*
- Convergence of optimisers — *work in progress with Julio Backhoff and Sigrid Källblad*

Open questions in discrete and continuous time:

- Non-Markovianity
- Higher dimensions
- ...

Summary

- We prove optimality of the synchronous coupling;
- We introduce a *monotone* numerical scheme;
- We show a stability result for bi-causal transport.

Julio Backhoff-Veraguas, Sigrid Källblad, and Benjamin A Robinson, *Adapted Wasserstein distance between the laws of SDEs*, arXiv:2209.03243 [math] (2022).