Distances between stochastic processes

Benjamin A. Robinson

University of Vienna

May 24, 2023 — Doctoral Seminar, Universität Klagenfurt

Joint work with

Julio Backhoff-Veraguas
University of Vienna

Sigrid Källblad
KTH Stockholm

Michaela Szölgyenyi
Universität Klagenfurt

Distances between stochastic processes

Benjamin A. Robinson

University of Vienna

May 24, 2023 — Doctoral Seminar, Universität Klagenfurt

Joint work with

Julio Backhoff-Veraguas
University of Vienna

Sigrid Källblad
KTH Stockholm

Michaela Szölgyenyi
Universität Klagenfurt

Distances between stochastic processes

$(X_n)_{n\in\{1,\dots N\}}$, $(Y_n)_{n\in\{1,\dots,N\}}$ real-valued stochastic processes

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

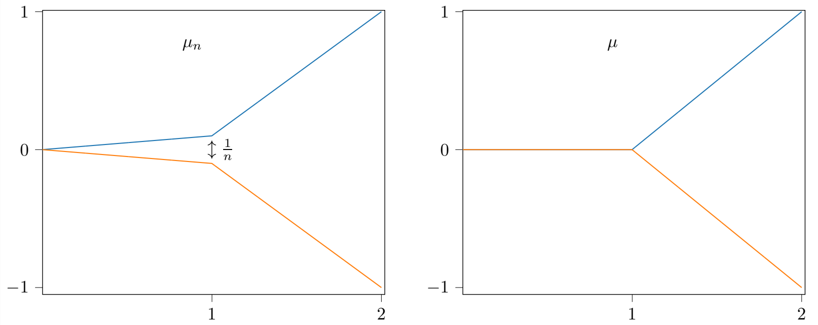
Distances between stochastic processes

$(X_n)_{n\in\{1,\dots N\}}$, $(Y_n)_{n\in\{1,\dots,N\}}$ real-valued stochastic processes

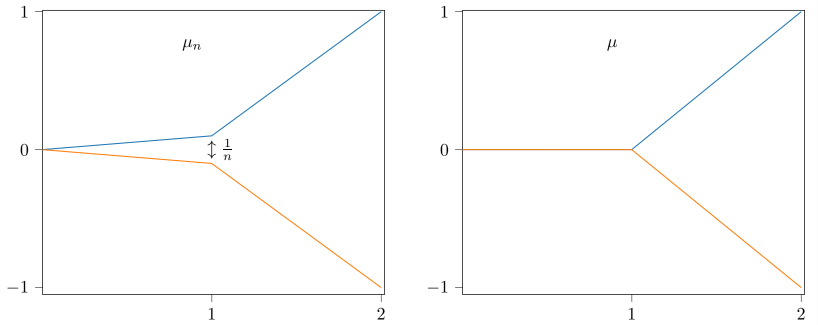
$\rightsquigarrow$ μ,ν probability measures on $\mathbb{R}^N$

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

Distances between stochastic processes

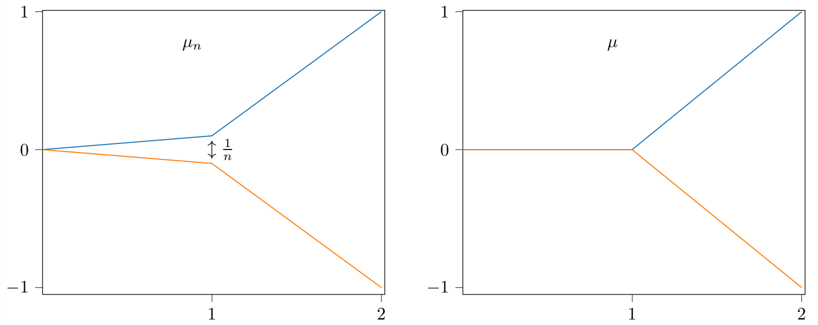


Distances between stochastic processes



"Can get rich"   "Cannot get rich"

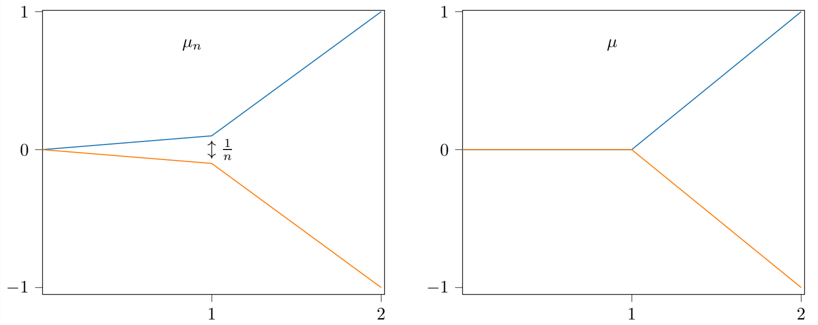
Distances between stochastic processes



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

$$V := \sup_\tau \mathbb{E}^{\mu}[X_\tau] = 0$$

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

$(X_n)_{n\in\{1,\ldots N\}}$, $(Y_n)_{n\in\{1,\ldots,N\}}$ real-valued stochastic processes

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

Distances between stochastic processes

$(X_n)_{n\in\{1,\dots N\}}$, $(Y_n)_{n\in\{1,\dots,N\}}$ real-valued stochastic processes

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

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

Distances between stochastic processes

$(X_n)_{n\in\{1,...N\}}$, $(Y_n)_{n\in\{1,...,N\}}$ real-valued stochastic processes

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

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

E.g. Wasserstein distance $\mathcal{W}_2$ — from optimal transport

Distances between stochastic processes

$(X_n)_{n\in\{1,\dots N\}}$, $(Y_n)_{n\in\{1,\dots,N\}}$ real-valued stochastic processes

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

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

E.g. Wasserstein distance $\mathcal{W}_2$:

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

Distances between stochastic processes

$(X_n)_{n\in\{1,\dots N\}}$, $(Y_n)_{n\in\{1,\dots,N\}}$ real-valued stochastic processes

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

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

E.g. Wasserstein distance $\mathcal{W}_2$:

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

Metrises weak convergence: $\mu_n \rightharpoonup \mu$ iff $\mathcal{W}_2(\mu_n,\mu) \to 0$.

Optimal transport

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

Find

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

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

Optimal transport

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

Find

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

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

Optimal transport

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

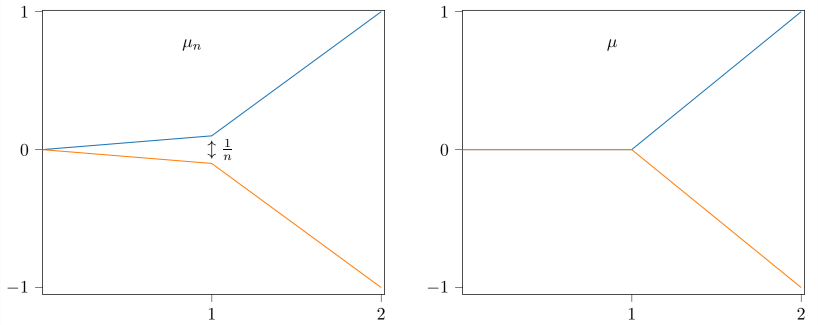
Find

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

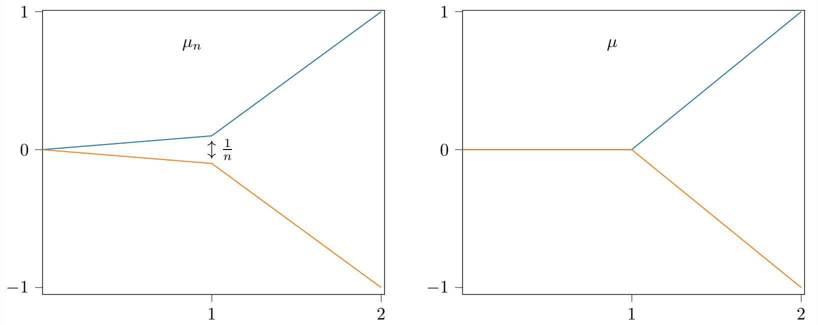
Monge (1781), Kantorovich (1942), ... and many more!

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

Distances between stochastic processes

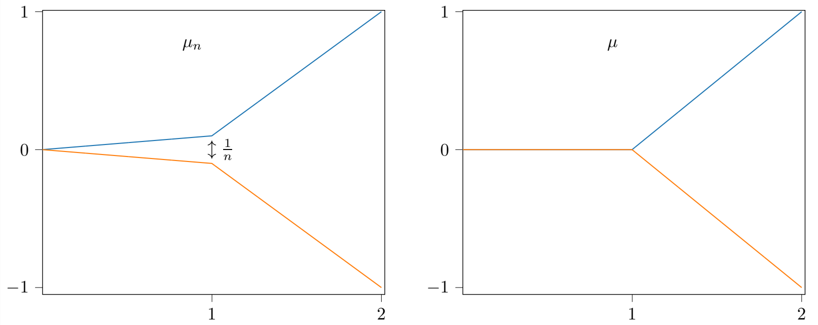


Distances between stochastic processes



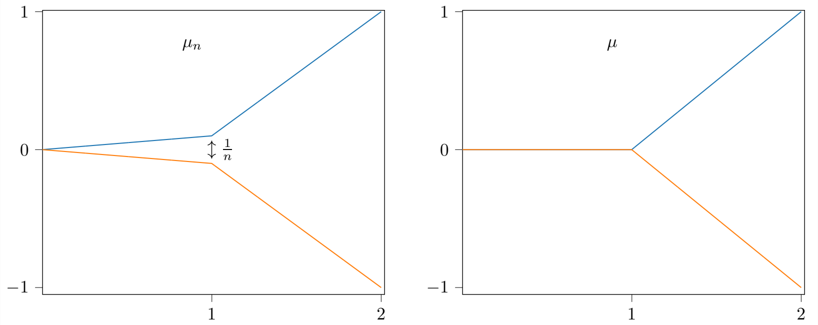
“Can get rich” “Cannot get rich”

Distances between stochastic processes



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

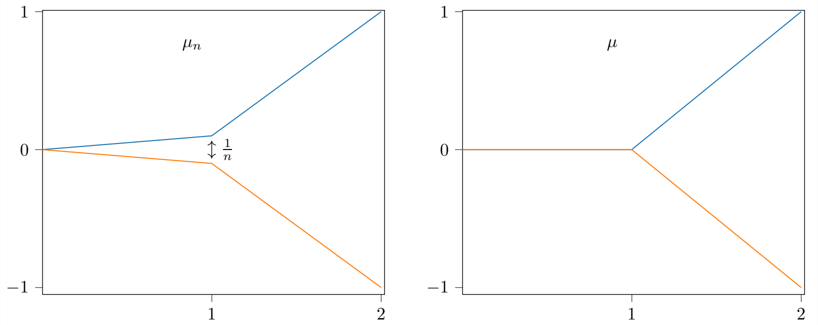
Distances between stochastic processes



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

$$V_n \not\to V$$

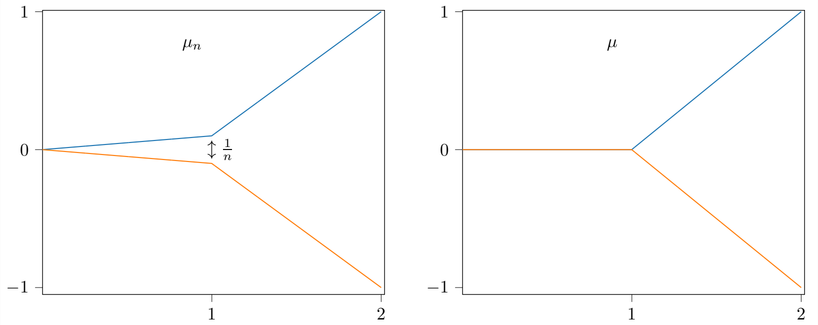
Distances between stochastic processes



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

$V_n \not\to V$ but $\mu_n \rightharpoonup \mu$

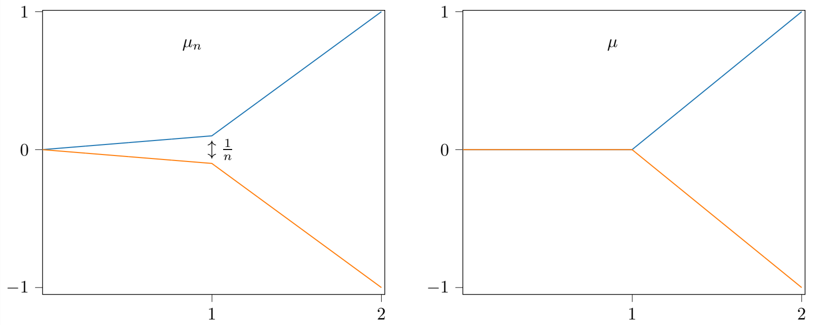
Distances between stochastic processes



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, Posch, Talay, among others ...

Adapted topology

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

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

Adapted topology

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

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

Adapted topology

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

Adapted topology

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

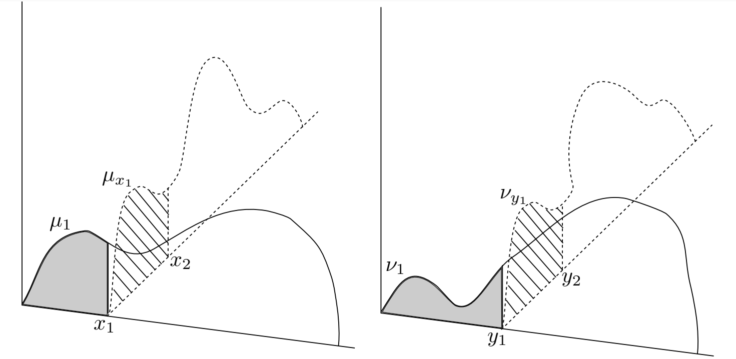
Knothe–Rosenblatt rearrangement

Adapted topology

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

Knothe–Rosenblatt rearrangement

— generalisation of monotone rearrangement

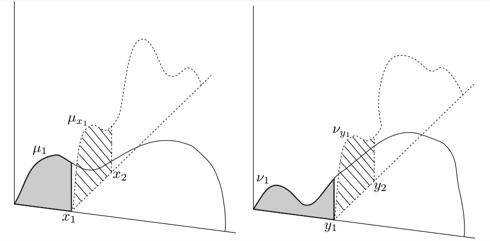


Adapted topology

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

Knothe–Rosenblatt rearrangement

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



Adapted topology

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

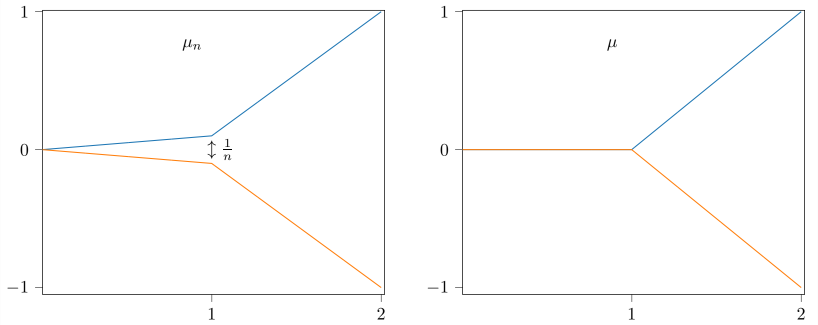
Knothe–Rosenblatt rearrangement

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

Theorem [Rüschendorf '85] [Posch '23+]

Under a monotonicity condition, the unique optimiser is the Knothe–Rosenblatt map T^{KR}. This induces the adapted weak toology.

Distances between stochastic processes



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

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

Coupling SDEs

ODE

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

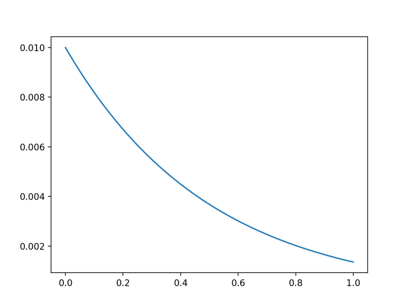
Coupling SDEs

ODE

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

Example

$$\mathrm{d}X_t = -2X_t\mathrm{d}t$$



Coupling SDEs

SDE

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

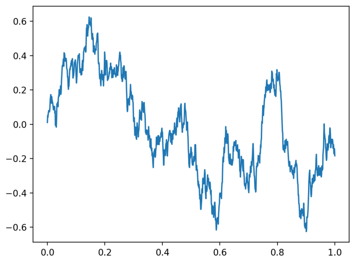
Coupling SDEs

SDE

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

Example

$$\mathrm{d}X_t = -2X_t\mathrm{d}t + \mathrm{d}W_t$$

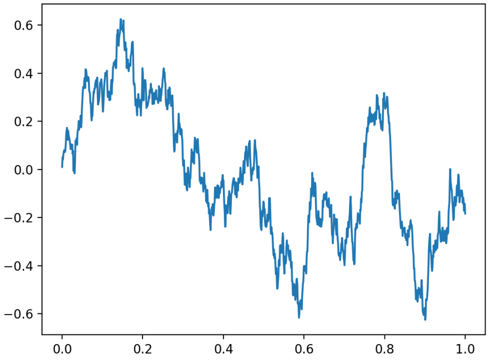


Coupling SDEs

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

$$\rightsquigarrow$$

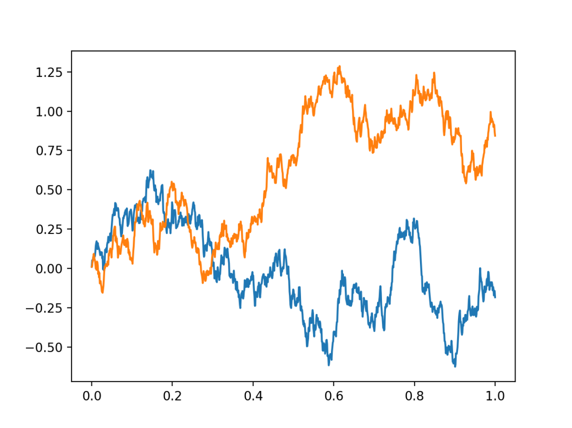
$$\mu \in \mathcal{P}(\Omega), \quad \Omega := C([0,1], \mathbb{R})$$



Coupling SDEs

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

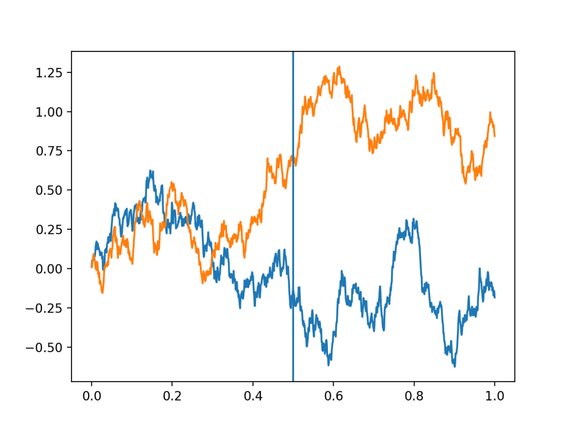
$b, \bar{b} : \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} : \mathbb{R} \to \mathbb{R}_+$



Coupling SDEs

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

$b, \bar{b} : \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} : \mathbb{R} \to \mathbb{R}_+$



Adapted topology

$$\mu, \nu \in \mathcal{P}(\Omega) \quad \rightsquigarrow \quad \mathcal{AW}_2^2(\mu, \nu) := \inf_{\substack{T \colon T_\# \mu = \nu \\ \text{adapted}}} \mathbb{E}\left[\int_0^1 |T(X_t) - X_t|^2 \mathrm{d}t\right].$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\Omega) \quad \rightsquigarrow \quad \mathcal{AW}_2^2(\mu, \nu) := \inf_{\substack{T:\, T_\# \mu = \nu \\ adapted}} \mathbb{E}\left[\int_0^1 |T(X_t) - X_t|^2 \mathrm{d}t\right].$$

Synchronous coupling
Continuous-time version of Knothe–Rosenblatt coupling

$$W = \bar{W}$$

Adapted topology

$$\mu, \nu \in \mathcal{P}(\Omega) \quad \rightsquigarrow \quad \mathcal{AW}_2^2(\mu, \nu) := \inf_{\substack{T\colon T_\# \mu = \nu \\ adapted}} \mathbb{E}\left[\int_0^1 |T(X_t) - X_t|^2 \mathrm{d}t\right].$$

Synchronous coupling
Continuous-time version of Knothe–Rosenblatt coupling

$$W = \bar{W}$$

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

Optimising over adapted maps T

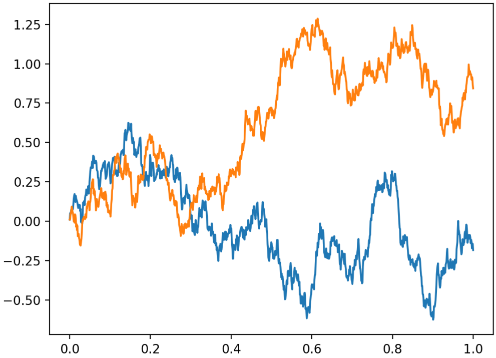
$$\Leftrightarrow$$

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

Adapted topology

Example

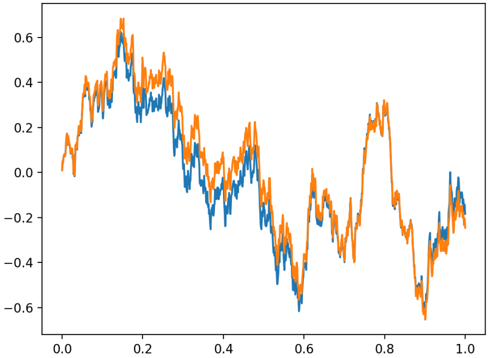
Product coupling — $W, \bar{W}$ independent



Adapted Wasserstein distance

Synchronous coupling

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



Adapted Wasserstein distance

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

$b, \bar{b} : \mathbb{R} \to \mathbb{R}$, $\sigma, \bar{\sigma} : \mathbb{R} \to \mathbb{R}_+$

Theorem 2 **[Backhoff-Veraguas, Källblad, R. '22] [R., Szölgyenyi '23]**

Under very mild conditions, the synchronous coupling is optimal.

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

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.

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).

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).