Adapted optimal transport in mathematical finance: structure, stability and transport inequalities

Jonghwa Park, Carnegie Mellon University

December 8th 2025
University of Michigan

Motivation: Robustness of Models

Metric for stochastic processes ?

Let's say we have two probabilistic models for a financial market:

Model 1: A law $\mathbb{P}$ of a process $X = (X_1, X_2, \ldots, X_N)$.

Model 2: A law $\mathbb{Q}$ of a process $Y = (Y_1, Y_2, \ldots, Y_N)$.

Question: What is a good metric between $\mathbb{P}$ and $\mathbb{Q}$?

Ideally, if $\mathbb{P}$ and $\mathbb{Q}$ are "close", $\mathbb{P}$ and $\mathbb{Q}$ should yield similar values in optimization problems arising in mathematical finance.

Metric for stochastic processes ?

We seek a metric for which the following map is continuous

$$\mathbb{P} \mapsto F(\mathbb{P})$$

for time-dependent operators F such as

- (optimal stopping problem)

$$F(\mathbb{P}) = \inf\{\mathbb{E}^{\mathbb{P}}[\varphi(X_\tau)] : \tau \text{ is a stopping time}\}$$

- (hedging error)

$$F(\mathbb{P}) = \text{ superhedging error under model } \mathbb{P}$$

- (utility maximization)

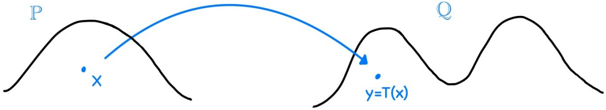
$$F(\mathbb{P}) = \text{ maximum utility under model } \mathbb{P}$$

- . . . many other optimization problems in finance

(Adapted-) Optimal Transport gives a **good** metric on the space of laws of stochastic processes.

Classical optimal transport

Classical optimal transport and the weak topology



$$T \text{ pushes } \mathbb{P} \text{ to } \mathbb{Q} : \ T_{\#}\mathbb{P} = \mathbb{Q}.$$

- The Wasserstein metric is an instance of the optimal transport cost:

$$\mathcal{W}(\mathbb{P}, \mathbb{Q}) := \inf \left\{ \int |x - y| \, \pi(dx, dy) : \pi \text{ has marginals } \mathbb{P}, \mathbb{Q} \right\}$$

- It induces the weak topology: the smallest topology for which

$$\mathbb{P} \mapsto \int \varphi(x) \mathbb{P}(dx)$$

is continuous for all continuous, bounded φ.

Topology for stochastic processes ?

Question: Is the Wasserstein metric suitable for stochastic processes ?

- (optimal stopping problem)

$$F(\mathbb{P}) = \inf\{\mathbb{E}^{\mathbb{P}}[\varphi(X_\tau)] : \tau \text{ is a stopping time}\}$$

- (hedging error)

$$F(\mathbb{P}) = \text{ superhedging error under model } \mathbb{P}$$

- (utility maximization)

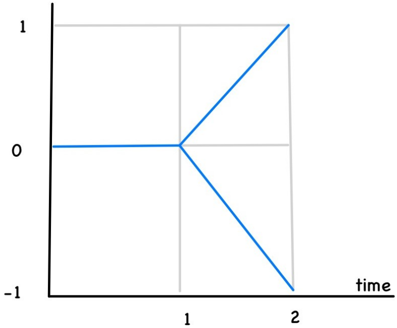
$$F(\mathbb{P}) = \text{ maximum utility under model } \mathbb{P}$$

- ... many other optimization problems in finance

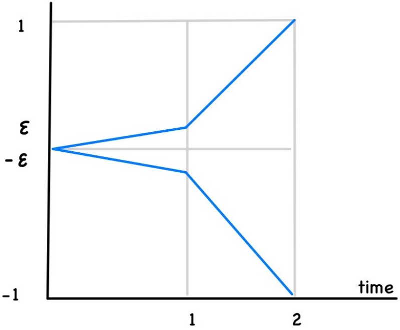
These maps are NOT continuous w.r.t the Wasserstein metric $\mathcal{W}$.

The Wasserstein metric $\mathcal{W}$ does NOT take "filtration" into account.

Standard example



$\mathbb{P} = \frac{1}{2}\delta_{(0,-1)} + \frac{1}{2}\delta_{(0,1)}$



$\mathbb{P}^\varepsilon = \frac{1}{2}\delta_{(-\varepsilon,-1)} + \frac{1}{2}\delta_{(\varepsilon,1)}$

- They are close in the Wasserstein metric, $\lim_{\varepsilon\to 0} \mathcal{W}(\mathbb{P}, \mathbb{P}^\varepsilon) = 0$.
- However, they are significantly different as stochastic processes.
- $\mathcal{L}^{\mathbb{P}}(X_2|X_1)$ is "random", $\mathcal{L}^{\mathbb{P}^\varepsilon}(X_2|X_1)$ is "deterministic".
- $\mathbb{P}$ is a martingale. $\mathbb{P}^\varepsilon$ is not a martingale, ...

Adapted optimal transport

Adapted weak topology: history

The Wasserstein distance (or the weak topology) ignores the temporal structure inherent in stochastic processes.

To address this, several approaches have been proposed:

- Aldous '81: extended weak convergence
- Hoover, Keisler '84: Model theory
- Hellwig '96: information topology
- Pflug-Pichler '12: nested distance
- Lasalle '18: bicausal couplings
- Bonnier, Liu, Oberhauser '20: higher rank signatures
- ...

They all define finer topology than the weak topology.

Adapted weak topology

Theorem (Backhoff, Bartl, Beiglböck, Eder, PTRF '20)

All topologies listed above are equivalent.

It is called the adapted weak topology.

- The adapted weak topology is finer than the weak topology.
- It is the smallest topology for which

  $$F(\mathbb{P}) = \inf\{\mathbb{E}^{\mathbb{P}}[\varphi(X_\tau)] : \tau \text{ is a stopping time}\}$$

  is continuous for all continuous, bounded, adapted φ.
- The adapted weak topology is metrizable.
- It extends to the continuous time setting (Bartl, Beiglböck, Pammer, Schrott, Zhang '25).

Adapted Wasserstein metric

All measures are on $\mathbb{R}^N$: a law of $X = (X_1, \cdots, X_N)$. Recall that

$$\mathcal{W}(\mathbb{P}, \mathbb{Q}) := \inf \left\{ \int |x - y| \, \pi(dx, dy) : \pi \text{ has marginals } \mathbb{P}, \mathbb{Q} \right\}.$$

The adapted Wasserstein distance respects the temporal structure:

$$\mathcal{AW}(\mathbb{P}, \mathbb{Q}) := \inf \left\{ \int |x - y| \, \pi(dx, dy) : \pi \text{ has marginals } \mathbb{P}, \mathbb{Q} \text{ and is bicausal} \right\}.$$

π is a bicausal coupling between $\mathbb{P}$ and $\mathbb{Q}$ if π has marginals $\mathbb{P}$ and $\mathbb{Q}$ and

$$Y_1, Y_2, \dots, Y_t \perp\!\!\!\perp_{X_1, \dots, X_t} X_{t+1}, \dots, X_N \text{ for all t},$$
$$X_1, X_2, \dots, X_t \perp\!\!\!\perp_{Y_1, \dots, Y_t} Y_{t+1}, \dots, Y_N \text{ for all t}.$$

At each time t, coupling is independent of future information given the current information.

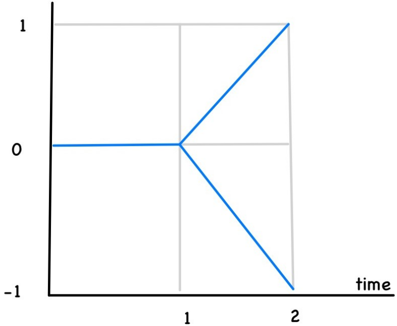
Adapted Wasserstein metric: $N = 2$

For example, when $N = 2$,

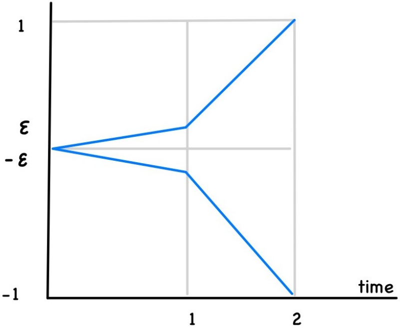
$$\mathcal{AW}(\mathbb{P}, \mathbb{Q})$$
$$= \inf \left\{ \int |x_1 - y_1| + \mathcal{W}(\mathbb{P}_{x_1}, \mathbb{Q}_{y_1})\, \pi^1(dx_1, dy_1) : \pi^1 \text{ has marginals } \mathbb{P}_1, \mathbb{Q}_1 \right\}.$$

- Couple conditional laws $\mathbb{P}_{x_1} = \mathcal{L}^{\mathbb{P}}(X_2 | X_1 = x_1)$ and $\mathbb{Q}_{y_1} = \mathcal{L}^{\mathbb{P}}(Y_2 | Y_1 = y_1)$. how close are the conditional laws ?
- Couple first marginals $\mathbb{P}_1$ and $\mathbb{Q}_1$. how close are first marginals ?

Standard example revisited



$\mathbb{P} = \frac{1}{2}\delta_{(0,-1)} + \frac{1}{2}\delta_{(0,1)}$



$\mathbb{P}^{\varepsilon} = \frac{1}{2}\delta_{(-\varepsilon,-1)} + \frac{1}{2}\delta_{(\varepsilon,1)}$

- $\lim_{\varepsilon\to 0} \mathcal{W}(\mathbb{P}^{\varepsilon}, \mathbb{P}) = 0.$
- $\lim_{\varepsilon\to 0} \mathcal{AW}(\mathbb{P}^{\varepsilon}, \mathbb{P}) = 1$ ($\mathbb{P}^{\varepsilon}$ are $\mathbb{P}$ are NOT close).

Adapted Wasserstein metric in finance

The metric $\mathcal{AW}$ is indeed a **good** metric in applications:

Theorem (Backhoff, Bartl, Beiglböck, Eder, F&S '20)

For many time-dependent optimization problems F in finance (e.g., optimal stopping problems, hedging error, risk measurement, stochastic programming and utility maximization),

$$|F(\mathbb{P}) - F(\mathbb{Q})| \lesssim \mathcal{AW}(\mathbb{P}, \mathbb{Q}).$$

Recall that F listed above are NOT continuous w.r.t $\mathcal{W}$.

It has been extensively studied. Key directions include

- Geometric properties (completion, embedding into the Wasserstein space, characterization of compact sets, geodesics, barycenters, ...)
- Statistical estimation (finite-sample guarantees, smoothing, ...)
- Applications (robust finance, DRO, ML, ...)

Quantitative bounds in applications

Regrettably, it is hard to compute or estimate $\mathcal{AW}$. This poses a significant limitation in applications where quantitative properties of $\mathcal{AW}$ become relevant.

- $\mathcal{AW}$ is inherently a nested optimization problem (dynamic programming principle): $\mathcal{AW}(\mathbb{P}, \mathbb{Q}) = \mathcal{W}(\mathrm{ip}(\mathbb{P}), \mathrm{ip}(\mathbb{Q}))$ where

$$\mathrm{ip}(\mathbb{P}) = \mathcal{L}^{\mathbb{P}}(X_1, \mathcal{L}^{\mathbb{P}}(X_2, \mathcal{L}^{\mathbb{P}}(X_3, \cdots \mathcal{L}^{\mathbb{P}}(X_N | X_{1:N-1}) \cdots | X_{1:2}) | X_1)).$$

- Closed-form expressions are rare (only known for Gaussian processes).
- $\mathcal{W}$-consistent estimators are not $\mathcal{AW}$-consistent (e.g., an empirical measure is no longer a consistent estimator).

We know $\mathcal{W} \leq \mathcal{AW}$ but what more can be said?

Main question

Question: can we compare

$$\mathcal{AW}(\mathbb{P}, \mathbb{Q}) \text{ vs. } \mathcal{J}(\mathbb{P}, \mathbb{Q})$$

for some well-understood metric (or functional) $\mathcal{J}$?

- If such a comparison exists, results and intuitions developed for $\mathcal{J}$ can be transferred to $\mathcal{AW}$.

Main question

Question: can we compare

$$\mathcal{AW}(\mathbb{P},\mathbb{Q}) \text{ vs. } \mathcal{J}(\mathbb{P},\mathbb{Q})$$

for some well-understood metric (or functional) $\mathcal{J}$?

- If such a comparison exists, results and intuitions developed for $\mathcal{J}$ can be transferred to $\mathcal{AW}$.

More concretely, I will present my answers to the following question:

- $\mathcal{J}(\mathbb{P},\mathbb{Q}) = f(\mathcal{W}(\mathbb{P},\mathbb{Q}))$
  (characterization of $\mathcal{AW}$-compact sets and statistical estimation)
- $\mathcal{J}(\mathbb{P},\mathbb{Q}) = \sqrt{\mathcal{H}(\mathbb{Q}|\mathbb{P})}$
  (concentration of measure)

Compactness in the adapted weak topology: $\mathcal{AW}$ vs. $\mathcal{W}$

Bounding $\mathcal{AW}$ metrics: special case

Recall that $\mathcal{W} \leq \mathcal{AW}$. We aim to find:

$$\mathcal{W}(\mathbb{P},\mathbb{Q}) \leq \mathcal{AW}(\mathbb{P},\mathbb{Q}) \leq f(\mathcal{W}(\mathbb{P},\mathbb{Q}))$$

for some nice continuous function f and $f(0) = 0$.

A special case: the cost $c(x,y) = 2\mathbb{1}_{\{x \neq y\}}$.

$$\mathrm{TV}(\mathbb{P},\mathbb{Q}) = \inf\{2\pi(X \neq Y) : \pi \text{ has marginals } \mathbb{P},\mathbb{Q}\},$$
$$\mathrm{ATV}(\mathbb{P},\mathbb{Q}) = \inf\{2\pi(X \neq Y) : \pi \text{ has marginals } \mathbb{P},\mathbb{Q} \text{ and is bicausal}\}.$$

Theorem (Eckstein, Pammer, AAP '24; Acciaio, Hou, Pammer '25)

Let $\mathbb{P},\mathbb{Q}$ be laws of N-step processes.

$$\mathrm{TV}(\mathbb{P},\mathbb{Q}) \leq \mathrm{ATV}(\mathbb{P},\mathbb{Q}) \leq (2N-1)\mathrm{TV}(\mathbb{P},\mathbb{Q}).$$

$\mathcal{AW}$-compact sets

We aim to find:

$$\mathcal{W}(\mathbb{P}, \mathbb{Q}) \le \mathcal{AW}(\mathbb{P}, \mathbb{Q}) \le f(\mathcal{W}(\mathbb{P}, \mathbb{Q}))$$

for some nice continuous function f and $f(0) = 0$.

- This is ill-posed in general: if this were true,

  $$\text{the weak topology} \;=\; \text{the adapted weak topology}$$

  which is a contradiction.
- If K is a $\mathcal{AW}$-relatively compact set, $\mathcal{AW}$ and $\mathcal{W}$ generate the same topology on K.

Our hope is to find:

$$\mathcal{AW}(\mathbb{P}, \mathbb{Q}) \le f(\mathcal{W}(\mathbb{P}, \mathbb{Q})) \text{ for } \mathbb{P}, \mathbb{Q} \in K.$$

We need to characterize $\mathcal{AW}$-relatively compact sets K.

Standard example revisited

Recall that $\mathbb{P}^\varepsilon = \frac{1}{2}\delta_{(-\varepsilon,-1)} + \frac{1}{2}\delta_{(\varepsilon,1)}$ and $\{\mathbb{P}^\varepsilon : 0 < \varepsilon < 1\}$ is not $\mathcal{AW}$-compact.

Observation: $\mathcal{L}^{\mathbb{P}^\varepsilon}(X_2|X_1)$ becomes increasingly irregular as $\varepsilon \to 0$.

$$\mathcal{L}^{\mathbb{P}^\varepsilon}(X_2|X_1 = x) = \begin{cases} \delta_1 \text{ if } x = \varepsilon, \\ \delta_{-1} \text{ if } x = -\varepsilon. \end{cases}$$

- It exhibits a discontinuous jump near $x = 0$ (heavyside step function)!
- K is $\mathcal{AW}$-compact if and only if conditional laws are **uniformly** controlled.

Arzela–Ascoli theorem

Theorem (Eder '19)

K is a $\mathcal{AW}$-relatively compact set if and only if K is a $\mathcal{W}$-relatively compact set and

$$\lim_{\delta \to 0} \sup_{\mathbb{P} \in K} \omega_{\mathbb{P}}(\delta) \to 0$$

where $\omega_{\mathbb{P}} : (0, \infty) \to [0, \infty)$ is the modulus of continuity of $\mathbb{P}$.

When $N = 2$,

$$\omega_{\mathbb{P}}(\delta) := \sup \left\{ \mathbb{E}[\mathcal{W}(\mathcal{L}^{\mathbb{P}}(X_2 \mid X_1), \mathcal{L}^{\mathbb{P}}(Y_2 \mid Y_1))] : X, Y \sim \mathbb{P}, \mathbb{E}[|X_1 - Y_1|] < \delta \right\}.$$

- The modulus of continuity gauges the (spatial) regularity of conditional laws.
- In the standard example, $\sup_{0<\varepsilon<1} \omega_{\mathbb{P}^\varepsilon}(\delta) = 2$.

Examples of $\mathcal{AW}$-compact sets

Many measures in applications have α-Hölder kernels:

$$\mathcal{W}(\mathcal{L}^{\mathbb{P}}(X_2|X_1 = x), \mathcal{L}^{\mathbb{P}}(X_2|X_1 = y)) \leq C\,|x - y|^{\alpha}\,.$$

Lemma (Backhoff et al., AAP '22; Larsson, Park, Wiesel, '25)

$\mathbb{P}$ has Lipschitz kernels ($\alpha = 1$) if one of the following holds:

- $\mathbb{P}$ is supported on finitely many points.
- $\mathbb{P}$ has a Lipschitz density that is bounded away from 0.
- $X_{t+1} = F_t(X_1, \ldots, X_t, \varepsilon_{t+1})$ for independent ε_{t+1} and Lipschitz F_t.
- $\mathbb{P} = \mathbb{Q} * \mathcal{N}(0, \mathrm{I})$ where $\mathcal{N}(0, \mathrm{I})$ is an independent Gaussian measure.

Main result

Theorem (Blanchet, Larsson, Park, Wiesel, '24)

There exists $C = C(N) > 0$ such that for any $\delta > 0$ and $R > 0$,

$$\mathcal{AW}(\mathbb{P},\mathbb{Q}) \le C\left(\frac{R}{\delta}\mathcal{W}(\mathbb{P},\mathbb{Q}) + \omega_{\mathbb{P}}(\delta) \vee \omega_{\mathbb{Q}}(\delta) + \int_{\{|x|>R\}} |x|\,(\mathbb{P}+\mathbb{Q})(dx)\right).$$

The RHS is:

$$\frac{R}{\delta}\mathcal{W}(\mathbb{P},\mathbb{Q}) + \underbrace{\omega_{\mathbb{P}}(\delta) \vee \omega_{\mathbb{Q}}(\delta)}_{\text{“regularity of kernels”}} + \underbrace{\int_{\{|x|>R\}} |x|\,(\mathbb{P}+\mathbb{Q})(dx)}_{\text{“tails”}}.$$

- As $\delta \to 0$, $\omega_{\mathbb{P}}(\delta) \vee \omega_{\mathbb{Q}}(\delta) \to 0$ uniformly in a $\mathcal{AW}$-compact set.
- As $R \to \infty$, $\int_{\{|x|>R\}} |x|\,(\mathbb{P}+\mathbb{Q})(dx) \to 0$ uniformly in a $\mathcal{AW}$-compact set.
- Tradeoff: in this case, $\frac{R}{\delta} \to \infty$.

Corollary

Corollary (Blanchet, Larsson, Park, Wiesel, '24)

If $\mathbb{P}, \mathbb{Q}$ are compactly supported and have Lipschitz kernels,

$$\mathcal{AW}(\mathbb{P}, \mathbb{Q}) \lesssim \sqrt{\mathcal{W}(\mathbb{P}, \mathbb{Q})}.$$

Note

$$\mathcal{AW}(\mathbb{P}, \mathbb{Q}) \lesssim \frac{R}{\delta}\mathcal{W}(\mathbb{P}, \mathbb{Q}) + \underbrace{\omega_{\mathbb{P}}(\delta) \vee \omega_{\mathbb{Q}}(\delta)}_{\lesssim \delta} + \underbrace{\int_{\{|x|>R\}} |x| \, (\mathbb{P} + \mathbb{Q})(dx)}_{=0}.$$

Thus,

$$\mathcal{AW}(\mathbb{P}, \mathbb{Q}) \lesssim \frac{1}{\delta}\mathcal{W}(\mathbb{P}, \mathbb{Q}) + \delta \Rightarrow \text{optimize over } \delta.$$

- Further developed by Acciaio, Hou, Pammer '25 (when densities are in the Sobolev space).

Applications

It can be used in statistical estimation:

- $\widehat{\mathbb{P}}_n = \frac{1}{n}\sum_{j=1}^n \delta_{X^j}$ is an empirical measure of $\mathbb{P}$ where $X^1, \cdots, X^n$ are i.i.d.
- $\widehat{\mathbb{P}}_n$ is NOT a $\mathcal{AW}$-consistent estimator.

Corollary (Blanchet, Larsson, Park, Wiesel, '24)

Let $\varepsilon > 0$. Assume that $\mathbb{P}$ has q-moment for $q > \frac{1}{2\varepsilon}$,

$$\mathbb{E}[\mathcal{AW}(\mathbb{P} * \mathcal{N}(0, \sigma^2 \mathrm{I}), \widehat{\mathbb{P}}_n * \mathcal{N}(0, \sigma^2 \mathrm{I}))] \lesssim n^{-\frac{1}{2}+\varepsilon}.$$

- The sharp rate is $n^{-\frac{1}{2}}$: obtained in Larsson, Park, Wiesel, '25.
- Its non-adapted counterpart is well studied in the literature.

Extensions: continuous time stochastic processes

In ongoing project, we aim to extend this result to diffusion processes:

Theorem (Larsson, Park, Wiesel, in preparation)

Let $\mathbb{P}^{b,\sigma} \in \mathcal{P}(C([0,1]))$ be a law of

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t.$$

If b_n, σ_n are nice enough,

$$\lim_{n\to\infty} \mathcal{W}(\mathbb{P}^{b_n,\sigma_n}, \mathbb{P}^{b,\sigma}) = 0 \;\Leftrightarrow\; \lim_{n\to\infty} \mathcal{AW}(\mathbb{P}^{b_n,\sigma_n}, \mathbb{P}^{b,\sigma}) = 0.$$

Aim to find conditions for

Weak convergence = Convergence in optimization problems.

Transport-entropy inequality: $\mathcal{AW}$ vs. $\mathcal{H}$

Transport-entropy inequality

We aim to study an adapted L^p-**transport-entropy inequality**:

$$\mathcal{AW}_p(\mathbb{P},\mathbb{Q}) \leq C\sqrt{\mathcal{H}(\mathbb{Q}|\mathbb{P})}$$

where $\mathcal{H}(\mathbb{Q}|\mathbb{P})$ is a relative entropy (or KL divergence) and

$$\mathcal{AW}_p(\mathbb{P},\mathbb{Q}) := \inf\left\{\left(\int |x-y|^p\,\pi(dx,dy)\right)^{1/p} : \pi \text{ is bicausal}\right\}$$

Non-adapted case ?

$$\mathcal{W}_p(\mathbb{P},\mathbb{Q}) \leq C\sqrt{\mathcal{H}(\mathbb{Q}\,|\,\mathbb{P})}$$

- It has deep connections with concentration of measure.
- Related to functional inequalities (log-sobolev inequality, Poincare inequality, ...).
- Dembo, Marton, Talagrand, Zeitouni, Ledoux, Bobkov, Götze, ...

Gaussian concentration

Theorem (Villani)

The following are equivalent.

1. $\mathbb{P}$ satisfies a T_1 inequality: for some $C_1 > 0$,

$$\mathcal{W}_1(\mathbb{P}, \mathbb{Q}) \leq C_1 \sqrt{\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})} \text{ for all probability measures } \mathbb{Q}.$$

2. $\mathbb{P}$ exhibits the Gaussian concentration: for some $C_2 > 0$,

$$\mathbb{P}(\varphi(X) > \mathbb{E}^{\mathbb{P}}[\varphi(X)] + \varepsilon) \leq e^{-C_2 \varepsilon^2} \text{ for all bounded, 1-Lipschitz } \varphi.$$

3. $\mathbb{P}$ is subgaussian: for some $C_3 > 0$,

$$\mathbb{E}^{\mathbb{P}}\left[e^{C_3 |X|^2}\right] < \infty.$$

Adapted T_1 inequality

Theorem (Park '25)

The following are equivalent.

1. $\mathbb{P}$ satisfies a T_1 inequality: for some $C_1 > 0$,

$$\mathcal{W}_1(\mathbb{P}, \mathbb{Q}) \leq C_1 \sqrt{\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})} \text{ for all probability measures } \mathbb{Q}.$$

2. $\mathbb{P}$ exhibits the Gaussian concentration: for some $C_2 > 0$,

$$\mathbb{P}(\varphi(X) > \mathbb{E}^{\mathbb{P}}[\varphi(X)] + \varepsilon) \leq e^{-C_2 \varepsilon^2} \text{ for all bounded, 1-Lipschitz } \varphi.$$

3. $\mathbb{P}$ is subgaussian: for some $C_3 > 0$,

$$\mathbb{E}^{\mathbb{P}}\left[e^{C_3 |X|^2}\right] < \infty.$$

Adapted T_1 inequality

Theorem (Park '25)

The following are equivalent.

1. $\mathbb{P}$ satisfies a T_1 inequality: for some $C_1 > 0$,

$$\mathcal{W}_1(\mathbb{P}, \mathbb{Q}) \leq C_1 \sqrt{\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})} \text{ for all probability measures } \mathbb{Q}.$$

2. $\mathbb{P}$ exhibits the Gaussian concentration: for some $C_2 > 0$,

$$\mathbb{P}(\varphi(X) > \mathbb{E}^{\mathbb{P}}[\varphi(X)] + \varepsilon) \leq e^{-C_2 \varepsilon^2} \text{ for all bounded, 1-Lipschitz } \varphi.$$

3. $\mathbb{P}$ is subgaussian: for some $C_3 > 0$,

$$\mathbb{E}^{\mathbb{P}}\left[e^{C_3 |X|^2}\right] < \infty.$$

4. $\mathbb{P}$ satisfies an adapted T_1 inequality: for some $C_4 > 0$,

$$\mathcal{AW}_1(\mathbb{P}, \mathbb{Q}) \leq C_4 \sqrt{\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})} \text{ for all probability measures } \mathbb{Q}.$$

Bolley–Villani bound

Theorem (Bolley–Villani)

Let $\varphi \geq 0$ and

$$\mathrm{TV}_\varphi(\mathbb{P}, \mathbb{Q}) := \inf_{\pi \in \mathrm{Cpl}(\mathbb{P}, \mathbb{Q})} \int (\varphi(x) + \varphi(y)) \mathbb{1}_{\{x \neq y\}} \pi(dx, dy).$$

Then

$$\mathrm{TV}_\varphi(\mathbb{P}, \mathbb{Q}) \leq \left(1 + \log \mathbb{E}^{\mathbb{P}}[e^{\varphi(X)^2}]\right)^{1/2} \sqrt{2\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})}.$$

The Bolley–Villani bound propagates in time.

- To pass from $\mathrm{Cpl}(\mathbb{P}, \mathbb{Q})$ to $\mathrm{Cpl}_{\mathrm{bc}}(\mathbb{P}, \mathbb{Q})$, apply the Bolley–Villani bound inductively along time.
- Chain rule of entropy:

$$\mathcal{H}(\mathbb{Q}|\mathbb{P}) = \sum_{t=0}^{N-1} \int \mathcal{H}(\mathbb{Q}_{x_{1:t}} | \mathbb{P}_{x_{1:t}}) \mathbb{Q}(dx_{1:t}).$$

Adapted Bolley–Villani bound

Theorem (Park '25)

Let $\varphi \geq 0$. Then

$$\mathrm{ATV}_{\varphi}(\mathbb{P}, \mathbb{Q}) \leq (2\sqrt{N} + 1) \left(1 + \log \mathbb{E}^{\mathbb{P}}[e^{\varphi(X)^2}]\right)^{1/2} \sqrt{2\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})}.$$

- The $\sqrt{N}$ scaling is sharp.

Corollary (Park '25)

Let $p \geq 1$. If $\mathbb{E}^{\mathbb{P}}[e^{\alpha |X|^{2p}}] < \infty$, then

$$\mathcal{AW}_p(\mathbb{P}, \mathbb{Q})^p \leq \frac{2^{p-1}(2\sqrt{N} + 1)}{\sqrt{\alpha}} \left(1 + \log \mathbb{E}^{\mathbb{P}}[e^{\alpha |X|^{2p}}]\right)^{1/2} \sqrt{2\mathcal{H}(\mathbb{Q} \,|\, \mathbb{P})}.$$

T_2 inequality ?

$$\begin{array}{c} \text{log-Sobolev inequality} \;\Rightarrow\; T_2 \text{ inequality} \Rightarrow\; \text{Poincare inequality} \\ \Updownarrow \\ \textit{dimension-free} \text{ Gaussian concentration} \end{array}$$

Question: What about an adapted T_2 inequality? Work in progress.

Thank you !