Statistical Sciences
UNIVERSITY OF TORONTO

# Optimal Transport Divergences induced by Scoring Functions

Silvana M. Pesenti, U. Toronto

joint with

Steven Vanduffel (Vrije Universiteit Brussel)

https://doi.org/10.48550/arXiv.2311.12183

Women in OT – April 17-19, 2024

# Motivation - Why robustify?

Let $\rho\colon L^\infty \to \mathbb{R}$ be a risk measure. Of interest

$$\rho(X)$$

# Motivation - Why robustify?

Let $\rho\colon L^\infty \to \mathbb{R}$ be a risk measure. Of interest

$$\rho(X)$$

- Distributional uncertainty - missing / incomplete data
- Model uncertainty, e.g., $X = g(Z_1, \dots, Z_n)$
- Dependence uncertainty
- Distributional robust optimisation: "Best action in the worst case"
- Applications: robust decision making, portfolio management, hedging, partial identification, inequality measurement, …

# Worst-Case Risk Measures

Let $\rho\colon L^\infty \to \mathbb{R}$ be a risk measure. A distributional worst-case risk measure can be defined as

$$\sup_{X \in \mathcal{U}} \rho(X)\,,$$

for a suitable uncertainty set $\mathcal{U}$.

# Worst-Case Risk Measures

Let $\rho\colon L^\infty \to \mathbb{R}$ be a risk measure. A distributional worst-case risk measure can be defined as

$$\sup_{X \in \mathcal{U}} \rho(X)\,,$$

for a suitable uncertainty set $\mathcal{U}$.

$\rightarrow$ what are desirable properties of $\mathcal{U}$

$\rightarrow$ trade-off between too small and too large

# A motivating Example – [Bernard, P., Vanduffel 2023]

Distributional robust risk measures

$$\sup_{G \in \mathcal{U}_\varepsilon} \rho(G)$$

# A motivating Example – [Bernard, P., Vanduffel 2023]

Distributional robust risk measures

$$\sup_{G \in \mathcal{U}_\varepsilon} \rho(G)$$

Let $F$ be a reference distribution

$$\mathcal{U}_\varepsilon := \Big\{ G \mid d_W(F, G)^2 \leq \varepsilon \Big\},$$

where $d_W(G, F)$ denotes the Wasserstein distance of order 2, which for $F$, $G$, with finite second moment, has representation

$$d_W(F, G)^2 = \int_0^1 |F^{-1}(u) - G^{-1}(u)|^2 \mathrm{d}u.$$

# A motivating Example – [Bernard, P., Vanduffel 2023]

Distributional robust risk measures

$$\sup_{G \in \mathcal{U}_\varepsilon} \rho(G)$$

Let $F$ be a reference distribution

$$\mathcal{U}_\varepsilon := \Big\{ G \mid d_W(F, G)^2 \leq \varepsilon \Big\},$$

where $d_W(G, F)$ denotes the Wasserstein distance of order 2, which for $F$, $G$, with finite second moment, has representation

$$d_W(F, G)^2 = \int_0^1 |F^{-1}(u) - G^{-1}(u)|^2 \mathrm{d}u.$$

! penalises losses and gains symmetrically

# A motivating Example – [Bernard, P., Vanduffel 2023]

Let $\rho_\gamma$ be a concave distortion (coherent) risk measure

$$\int_0^1 \gamma(u) G^{-1}(u)\, du$$

# A motivating Example – [Bernard, P., Vanduffel 2023]

Let $\rho_\gamma$ be a concave distortion (coherent) risk measure , then

$$\sup_{G \in \mathcal{U}_\varepsilon} \int_0^1 \gamma(u) G^{-1}(u)\, du = \rho(F) + \sqrt{\varepsilon} \sqrt{\int_0^1 \gamma(u)^2\, du}$$

and the worst-case quantile function is

$$F^{-1,*}(u) := F^{-1}(u) + \frac{\sqrt{\varepsilon}}{\sqrt{\int_0^1 \gamma(u)^2\, du}}\, \gamma(u)\,.$$

# A motivating Example – [Bernard, P., Vanduffel 2023]

Let $\rho_\gamma$ be a concave distortion (coherent) risk measure , then

$$\sup_{G \in \mathcal{U}_\varepsilon} \int_0^1 \gamma(u) G^{-1}(u)\, du = \rho(F) + \sqrt{\varepsilon} \sqrt{\int_0^1 \gamma(u)^2\, du}$$

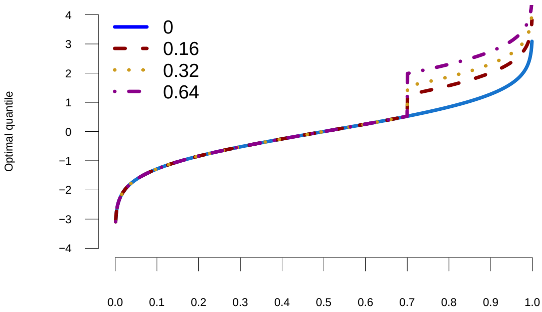
and the worst-case quantile function is

$$F^{-1,*}(u) := F^{-1}(u) + \frac{\sqrt{\varepsilon}}{\sqrt{\int_0^1 \gamma(u)^2\, du}}\, \gamma(u)\,.$$

! robust risk measure: constant shift

! constant is independent of $F$

# Worst-case Quantile Functions



# Motivation

- Distances other than the $p$-Wasserstein distances
- Divergences that:
    - allow for comparison of distributions with differing support
    - are asymmetric, penalising different parts of the distribution
    - constructive
    - interpretation from a statistical and risk management

# Motivation

- Distances other than the $p$-Wasserstein distances
- Divergences that:
  - allow for comparison of distributions with differing support
  - are asymmetric, penalising different parts of the distribution
  - constructive
  - interpretation from a statistical and risk management

→ connecting OT & risk measures & elicitability
→ uncertainty sets induced by the risk to be assess

# Monge-Kantorovich optimal transport problem

**Definition 1**
Let $c\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a cost function. Then the Monge-Kantorovich optimisation problem with respect to the cdfs $F_1$ and $F_2$ is given by

$$\inf_{\pi \in \Pi(F_1,\, F_2)} \left\{ \int_{\mathbb{R}^2} c\big(z_1, z_2\big)\, \pi(\mathrm{d}z_1, \mathrm{d}z_2) \right\}, \tag{1}$$

where $\Pi(F_1, F_2)$ denotes the set of all bivariate cdfs with marginal cdfs $F_1$ and $F_2$, respectively.

# Monge-Kantorovich optimal transport problem

**Definition 1**
Let $c\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a cost function. Then the Monge-Kantorovich optimisation problem with respect to the cdfs $F_1$ and $F_2$ is given by

$$\inf_{\pi \in \Pi(F_1,\, F_2)} \left\{ \int_{\mathbb{R}^2} c\big(z_1, z_2\big)\, \pi(\mathrm{d}z_1, \mathrm{d}z_2) \right\}, \tag{1}$$

where $\Pi(F_1, F_2)$ denotes the set of all bivariate cdfs with marginal cdfs $F_1$ and $F_2$, respectively.

$\rightarrow$ $c(z_1, z_2) = |z_1 - z_2|^p$ gives the $p$-Wasserstein distance.

$\rightarrow$ Asymmetric cost functions? via scoring functions

# Interlude - Scoring Functions

# Scoring rules in statistics

- $y_1, \ldots, y_N$ observations of r.v. $Y \sim F$
- Aim: forecast functional $T(F)$, e.g., mean, quantile, risk measure
- How to compare forecasts of models $A$ & $B$:

  **(A)** $z_1^{(A)}, \ldots, z_N^{(A)} \in \mathsf{A}$

  **(B)** $z_1^{(B)}, \ldots, z_N^{(B)} \in \mathsf{A}$

- $y_1, \ldots, y_N$ observations of r.v. $Y \sim F$
- Aim: forecast functional $T(F)$, e.g., mean, quantile, risk measure
- How to compare forecasts of models $A$ & $B$:

  (A) $z_1^{(A)}, \ldots, z_N^{(A)} \in \mathsf{A}$

  (B) $z_1^{(B)}, \ldots, z_N^{(B)} \in \mathsf{A}$

- Use a loss/scoring function $S \colon \mathsf{A} \times \mathbb{R} \to [0, \infty]$ and compare

$$L^{(A)} := \frac{1}{N} \sum_{i=1}^{N} S\big(z_i^{(A)},\, y_i\big) \overset{?}{\lessgtr} \frac{1}{N} \sum_{i=1}^{N} S\big(z_i^{(B)},\, y_i\big) =: L^{(B)}\,.$$

# Scoring rules in statistics

- $y_1, \ldots, y_N$ observations of r.v. $Y \sim F$
- Aim: forecast functional $T(F)$, e.g., mean, quantile, risk measure
- How to compare forecasts of models $A$ & $B$:

  (A) $z_1^{(A)}, \ldots, z_N^{(A)} \in \mathsf{A}$

  (B) $z_1^{(B)}, \ldots, z_N^{(B)} \in \mathsf{A}$

- Use a loss/scoring function $S$: $\mathsf{A} \times \mathbb{R} \to [0, \infty]$ and compare

$$L^{(A)} := \frac{1}{N} \sum_{i=1}^{N} S\big(z_i^{(A)},\, y_i\big) \overset{?}{\lesseqgtr} \frac{1}{N} \sum_{i=1}^{N} S\big(z_i^{(B)},\, y_i\big) =: L^{(B)}\,.$$

$\to$ meaningful forecast comparison, model selection, regression, M-estimation, ...

# Scoring Functions [Murphy & Daan, 1985, Engelberg et al., 2009]

- A scoring function is a measurable map $S$: A $\times \mathbb{R} \to [0, \infty]$.
- $T$: $\mathcal{M} \to$ A law-invariant functional of interest.
- $\mathcal{M}$ subset of probability measures on $\mathbb{R}$

# Scoring Functions [Murphy & Daan, 1985, Engelberg et al., 2009]

- A scoring function is a measurable map $S\colon \mathsf{A} \times \mathbb{R} \to [0, \infty]$.
- $T\colon \mathcal{M} \to \mathsf{A}$ law-invariant functional of interest.
- $\mathcal{M}$ subset of probability measures on $\mathbb{R}$

For a functional $T\colon \mathcal{M} \to \mathsf{A}$, we say

$(i)$ $S$ is *consistent* for $T$, if for all $F \in \mathcal{M}$ and for all $z \in \mathsf{A}$

$$\int S(T(F), y)\, \mathrm{d}F(y) \le \int S(z, y)\, \mathrm{d}F(y)\,. \tag{2}$$

$(ii)$ $S$ is *strictly* consistent for $T$, if it is consistent for $T$ and if (2) is strict for all $z \neq T(F)$.

# Elicitability – Bayes act

$T$ is *elicitable* on $\mathcal{M}$, if there exists a strictly $\mathcal{M}$-consistent scoring function for $T$. Moreover,

$$T(F) = \arg\min_{z \in \mathbb{R}} \int S(z, y) \, \mathrm{d}F(y)$$
$$= \arg\min_{z \in \mathbb{R}} \mathbb{E}[S(z, Y)], \quad Y \sim F.$$

# Elicitability – Bayes act

$T$ is *elicitable* on $\mathcal{M}$, if there exists a strictly $\mathcal{M}$-consistent scoring function for $T$. Moreover,

$$\begin{aligned} T(F) &= \arg\min_{z\in\mathbb{R}} \int S(z, y)\,\mathrm{d}F(y) \\ &= \arg\min_{z\in\mathbb{R}} \mathbb{E}[S(z, Y)]\,, \qquad Y \sim F. \end{aligned}$$

Example:

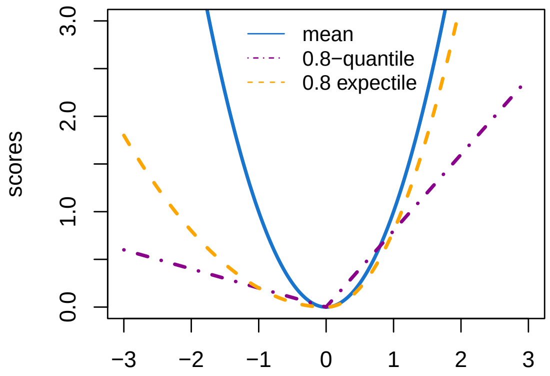
$$\mathbb{E}[Y] = \arg\min_{z\in\mathbb{R}} \mathbb{E}[(z - Y)^2]$$

# Elicitable Functionals [Osband 1985, Weber 2006, Gneiting 2011]

| $T$ | $S(z, y)$ |
|---|---|
| mean | $(x - y)^2$ |
| median | $\|x - y\|$ |
| $\mathrm{VaR}_\alpha$ | $(\mathbb{1}_{\{y \leq z\}} - \alpha)(z - y)$ |
| variance | NO |
| Expected Shortfall (ES) | NO |
| (mean, variance) | YES! |
| $(\mathrm{VaR}_\alpha, \mathrm{ES}_\alpha)$ | YES! |

# Scores for different functionals

$S(0, y)$ as a function of realisations $y$



# Mean - Bregman divergence

Proposition 1 (Elicitability of Mean – [Gneiting, 2011])

*Under technical conditions, the class of (strictly) consistent scoring functions for the mean are*

$$S_\phi(z, y) = B_\phi(y, z)\,, \quad z, y \in \mathbb{R}\,, \tag{3}$$

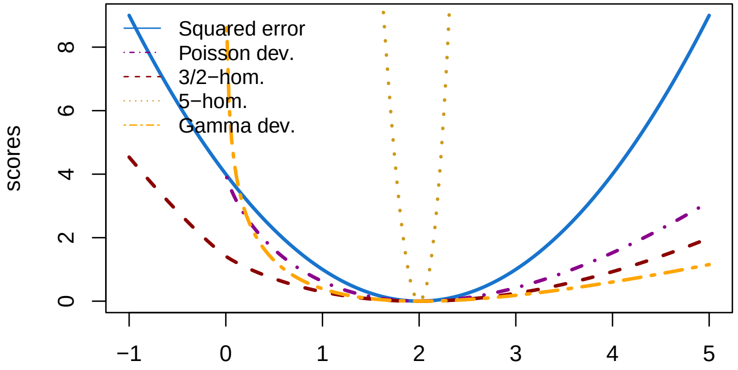
*where* $B_\phi$ *is the Bregman divergence*

$$B_\phi\big(x_1, x_2\big) := \phi(x_1) - \phi(x_2) - \phi'(x_2)(x_1 - x_2)\,, \quad x_1, x_2 \in \mathbb{R}\,,$$

*and* $\phi$ *(strictly) convex.*

# Scores for the mean

$S(2, y)$ as a function of realisations $y$



# Monge-Kantorovich divergences with Scores

Let $S$ be an $\mathcal{M}$-consistent score for $T$.
The Monge-Kantorovich (MK) divergence induced by $S$ from the cdf $F_1 \in \mathcal{M}$ to the cdf $F_2 \in \mathcal{M}$ is

$$\mathscr{S}(F_1, F_2) := \inf_{\pi \in \Pi(F_1,\, F_2)} \left\{ \int_{\mathbb{R}^2} S\big(z_2, z_1\big)\, \pi(\mathrm{d}z_1, \mathrm{d}z_2) \right\}. \qquad (4)$$

# Monge-Kantorovich divergences with Scores

Let $S$ be an $\mathcal{M}$-consistent score for $T$.
The Monge-Kantorovich (MK) divergence induced by $S$ from the cdf $F_1 \in \mathcal{M}$ to the cdf $F_2 \in \mathcal{M}$ is

$$\mathscr{S}(F_1, F_2) := \inf_{\pi \in \Pi(F_1, F_2)} \left\{ \int_{\mathbb{R}^2} S\big(z_2, z_1\big)\, \pi(\mathrm{d}z_1, \mathrm{d}z_2) \right\}. \tag{4}$$

→ non-negative; satisfies $\mathscr{S}(F, F) = 0$

→ $\mathscr{S}(F_1, F_2) = 0$ need not imply $F_1 = F_2$

→ depends on the choice of $S$

→ What is the optimal coupling?

# Bregman-Wasserstein divergence

Let $S$ be a consistent score for the mean.

Then the MK divergence is the Bregman-Wasserstein divergence [Rankin & Wong, 2023]

$$\mathscr{B}_\phi(F_1, F_2) := \inf_{\pi \in \Pi(F_1,\, F_2)} \left\{ \int_{\mathbb{R}^2} B_\phi(z_1,\ z_2)\, \pi(\mathrm{d}z_1, \mathrm{d}z_2) \right\}, \tag{5}$$

reduces to 2-Wasserstein distance for $\phi(x) = x^2$.

# Bregman-Wasserstein divergence

Let $S$ be a consistent score for the mean.

Then the MK divergence is the Bregman-Wasserstein divergence [Rankin & Wong, 2023]

$$\mathscr{B}_\phi(F_1, F_2) := \inf_{\pi \in \Pi(F_1,\, F_2)} \left\{ \int_{\mathbb{R}^2} B_\phi(z_1,\, z_2)\, \pi(\mathrm{d}z_1, \mathrm{d}z_2) \right\}, \tag{5}$$

reduces to 2-Wasserstein distance for $\phi(x) = x^2$.

**Theorem**
The comonotonic coupling $(F_1^{-1}(U), F_2^{-1}(U))$, $U \sim U(0,1)$ is optimal, equivalently, the optimal transport map is $\alpha(x) = F_2^{-1}\big(F_1(x)\big)$.

# Bregman-Wasserstein divergence

Let $S$ be a consistent score for the mean.

Then the MK divergence is the Bregman-Wasserstein divergence [Rankin & Wong, 2023]

$$\mathscr{B}_\phi(F_1, F_2) := \mathbb{E}\Big[B_\phi\left(F_2^{-1}(U)\,, F_1^{-1}(U)\right)\Big] \tag{5}$$

reduces to 2-Wasserstein distance for $\phi(x) = x^2$.

**Theorem**
The comonotonic coupling $(F_1^{-1}(U), F_2^{-1}(U))$, $U \sim U(0,1)$ is optimal, equivalently, the optimal transport map is $\alpha(x) = F_2^{-1}\big(F_1(x)\big)$.

$\rightarrow$ For all choices of consistent scores for the mean.

# (Generalised) Quantiles

Proposition 2 (Elicitability of Quantiles – [Gneiting, 2011])

*Under technical conditions, the class of (strictly) consistent scores for the $\alpha$-quantile are*

$$S_g(z, y) = \big(\mathbb{1}_{\{y \leq z\}} - \alpha\big)\big(g(z) - g(y)\big)\,, \qquad z, y \in \mathbb{R}\,, \tag{6}$$

*where $g$ is a (strictly) increasing function.*

# (Generalised) Quantiles

Proposition 2 (Elicitability of Quantiles – [Gneiting, 2011])

*Under technical conditions, the class of (strictly) consistent scores for the $\alpha$-quantile are*

$$S_g(z, y) = \left(\mathbb{1}_{\{y \leq z\}} - \alpha\right)\left(g(z) - g(y)\right), \qquad z, y \in \mathbb{R}, \tag{6}$$

*where $g$ is a (strictly) increasing function.*

Theorem 2

*The optimal coupling of the MK divergence for any score $S_g$ is the comonotonic coupling.*

**Proposition 2 (Elicitability of Quantiles – [Gneiting, 2011])**

*Under technical conditions, the class of (strictly) consistent scores for the $\alpha$-quantile are*

$$S_g(z, y) = \left(\mathbb{1}_{\{y \leq z\}} - \alpha\right)\left(g(z) - g(y)\right), \qquad z, y \in \mathbb{R}, \tag{6}$$

*where $g$ is a (strictly) increasing function.*

**Theorem 2**

*The optimal coupling of the MK divergence for any score $S_g$ is the comonotonic coupling.*

$\rightarrow$ Generalisable to $\Lambda$-quantiles

# Expectiles – [Newey & Powell 1987]

$$e_\alpha(Y) := \underset{z \in \mathbb{R}}{\operatorname{argmin}} \; \alpha \, \mathbb{E}\big[\big(Y - z\big)_+^2\big] + (1-\alpha) \, \mathbb{E}\big[\big(Y - z\big)_-^2\big]$$

# Expectiles – [Newey & Powell 1987]

$$e_\alpha(Y) := \underset{z \in \mathbb{R}}{\operatorname{argmin}} \; \alpha \, \mathbb{E}\big[\big(Y - z\big)_+^2\big] + (1-\alpha)\, \mathbb{E}\big[\big(Y - z\big)_-^2\big]$$

Proposition 3 (Elicitability of Expectiles – [Gneiting, 2011])

*Under technical conditions, the class of (strictly) consistent scores for the $\alpha$-expectile are*

$$S(z, y) = \big|\mathbb{1}_{\{y \leq z\}} - \alpha\big| \, B_\phi(y, z)\,, \quad z, y \in \mathbb{R}\,, \tag{7}$$

*where $\phi$ is (strictly) convex.*

# Expectiles – [Newey & Powell 1987]

$$e_\alpha(Y) := \operatorname*{argmin}_{z \in \mathbb{R}} \; \alpha \, \mathbb{E}\big[\big(Y - z\big)_+^2\big] + (1 - \alpha) \, \mathbb{E}\big[\big(Y - z\big)_-^2\big]$$

Proposition 3 (Elicitability of Expectiles – [Gneiting, 2011])

*Under technical conditions, the class of (strictly) consistent scores for the $\alpha$-expectile are*

$$S(z, y) = \big|\mathbb{1}_{\{y \le z\}} - \alpha\big| \, B_\phi(y, z) \,, \quad z, y \in \mathbb{R} \,, \tag{7}$$

*where $\phi$ is (strictly) convex.*

Theorem 3

*The optimal coupling of the MK divergence for any scores in (7) is the comonotonic coupling.*

# Divergences where the comonotonic coupling is not optimal?

Proposition 4 (Osband's principle for OT)

- *$\tilde{T}$ elicitable with (strictly) consistent score $\tilde{S}$*
- *MK divergence of $\tilde{S}$ is attained by the comonotonic coupling*
- *$T := g \circ \tilde{T}, \quad g \colon \mathbb{R} \to \mathbb{R}$ strictly monotone*

*Then $T$ is elicitable with (strictly) consistent score*

$$S(z, y) := \tilde{S}\big(g^{-1}(z), y\big)\,. \tag{8}$$

# Divergences where the comonotonic coupling is not optimal?

Proposition 4 (Osband's principle for OT)

- $\tilde{T}$ *elicitable with (strictly) consistent score* $\tilde{S}$
- *MK divergence of* $\tilde{S}$ *is attained by the comonotonic coupling*
- $T := g \circ \tilde{T}, \quad g\colon \mathbb{R} \to \mathbb{R}$ *strictly monotone*

*Then* $T$ *is elicitable with (strictly) consistent score*

$$S(z, y) := \tilde{S}\big(g^{-1}(z), y\big)\,. \tag{8}$$

*If* $g$ *is increasing (decreasing), then the optimal coupling of the MK divergence of the score* (8) *is the comonotonic (antitonic) coupling.*

# Divergences where the comonotonic coupling is not optimal?

Proposition 4 (Osband's principle for OT)

- *$\tilde{T}$ elicitable with (strictly) consistent score $\tilde{S}$*
- *MK divergence of $\tilde{S}$ is attained by the comonotonic coupling*
- *$T := g \circ \tilde{T}, \quad g\colon \mathbb{R} \to \mathbb{R}$ strictly monotone*

*Then $T$ is elicitable with (strictly) consistent score*

$$S(z, y) := \tilde{S}\big(g^{-1}(z), y\big)\,. \tag{8}$$

*If $g$ is increasing (decreasing), then the optimal coupling of the MK divergence of the score (8) is the comonotonic (antitonic) coupling.*
$\rightarrow$ *Example: $T(F) = \frac{1}{\tilde{T}(F)}$.*

# Law invariant risk measures

A risk measure $\rho$ is coherent, if for all r.v. $X$

i) monotone: $\rho(X) \leq \rho(Y)$, if $X \leq Y$ a.s
ii) translation invariant: $\rho(X + m) = \rho(X) + m$, for all $m \in \mathbb{R}$
iii) positive homogeneous: $\rho(\lambda X) = \lambda \rho(X)$, $\lambda \geq 0$
iv) subadditive: $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

# Law invariant risk measures

A risk measure $\rho$ is coherent, if for all r.v. $X$

i) monotone: $\rho(X) \leq \rho(Y)$, if $X \leq Y$ a.s
ii) translation invariant: $\rho(X + m) = \rho(X) + m$, for all $m \in \mathbb{R}$
iii) positive homogeneous: $\rho(\lambda X) = \lambda \rho(X)$, $\lambda \geq 0$
iv) subadditive: $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Theorem 4 (Coherent Risk Measures)

*Let $T$ be an elicitable coherent risk measure satisfying $T(0) = 0$, and let $S$ be any strictly consistent score for $T$. Then, the optimal coupling of the MK divergence induced by the score $S$ is the comonotonic coupling.*

# Summary & Outlook

- Introduced asymmetric OT divergences & their optimal coupling
- Divergences that penalise different parts of the distribution asymmetrically
- Uncertainty set induced by the criterion to be optimised
- How to choose the MK divergence?
- How do uncertainty balls induced MK divergences look like?

# Thank you!

https://pesenti.utstat.utoronto.ca/

silvana.pesenti@utoronto.ca

# References


Engelberg, J., Manski, C. F., & Williams, J. (2009).
**Comparing the point predictions and subjective probability distributions of professional forecasters.**
*Journal of Business & Economic Statistics*, 27(1), 30–41.

Gneiting, T. (2011).
**Making and Evaluating Point Forecasts.**
*Journal of the American Statistical Association*, 106(494), 746–762.

Murphy, A. H. & Daan, H. (1985).
**Forecast Evaluation.**
In A. H. Murphy & R. W. Katz (Eds.), *Probability, Statistics and Decision Making in the Atmospheric Sciences* (pp. 379–437). Westview Press, Boulder, Colorado.

Rankin, C. & Wong, T.-K. L. (2023).
**Bregman-wasserstein divergence: geometry and applications.**
*arXiv preprint arXiv:2302.05833.*

# Assumptions on scores

**Assumption**

Let $S$ be an $\mathcal{M}$-consistent score for $T$ and denote by $\delta_y$, $y \in \mathbb{R}$, point measures. Then it holds that

$(i)$ $S\big(T(\delta_y), y\big) < S(z, y)$ for all $z \neq T(\delta_y)$ and $y \in \mathbb{R}$, and

$(ii)$ $S\big(T(\delta_y), y\big) = 0$ for all $y \in \mathbb{R}$.

$(i)$ means strict consistency on Dirac measures

$(ii)$ normalisation.