Discrete optimal transport

Peter Gladbach

University of Leipzig

joint work with Eva Kopfer and Jan Maas

May 22, 2018

Kantorovich distance

Given probability measures $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$, their **Kantorovich distance** is given by

$$W^2(\mu_0, \mu_1) = \inf_{\gamma \in \mathrm{Cpl}(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, d\gamma(x, y),$$

where $\mathrm{Cpl}(\mu_0, \mu_1) = \{\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \; : \; \pi_0^\sharp \gamma = \mu_0, \pi_1^\sharp \gamma = \mu_1\}$.
Optimal transport problem first formulated by Gaspard Monge in 1781.

The Benamou-Brenier formula

According to Benamou-Brenier (2000) we can write

$$W^2(\mu_0,\mu_1) = \inf_{(\mu_t)_t} \left\{ \int_0^1 \int_{\mathbb{R}^d} |v_t|^2 \, d\mu_t(x) \, dt \; : \; \partial_t \mu_t + \operatorname{div} \mu_t v_t = 0 \right\}$$

$$= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} |V_t|^2 \left(\frac{d\mu_t}{dx} \right)^{-1} dx \, dt \; : \; \partial_t \mu_t + \operatorname{div} V_t = 0 \right\}$$

The Benamou-Brenier formula

According to Benamou-Brenier (2000) we can write

$$
\begin{aligned}
W^2(\mu_0,\mu_1) &= \inf_{(\mu_t)_t} \left\{ \int_0^1 \int_{\mathbb{R}^d} |v_t|^2 \, d\mu_t(x)\, dt \;:\; \partial_t \mu_t + \operatorname{div} \mu_t v_t = 0 \right\} \\
&= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} |V_t|^2 \left(\frac{d\mu_t}{dx} \right)^{-1} dx\, dt \;:\; \partial_t \mu_t + \operatorname{div} V_t = 0 \right\} \\
&= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} |\nabla \phi_t|^2 \, d\mu_t(x)\, dt \;:\; \partial_t \mu_t + \operatorname{div} \mu_t \nabla \phi_t = 0 \right\}
\end{aligned}
$$

The Benamou-Brenier formula

According to Benamou-Brenier (2000) we can write

$$
\begin{aligned}
W^2(\mu_0,\mu_1) &= \inf_{(\mu_t)_t} \left\{ \int_0^1 \int_{\mathbb{R}^d} |v_t|^2 \, d\mu_t(x)\, dt \;:\; \partial_t \mu_t + \operatorname{div} \mu_t v_t = 0 \right\} \\
&= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} |V_t|^2 \left(\frac{d\mu_t}{dx} \right)^{-1} dx\, dt \;:\; \partial_t \mu_t + \operatorname{div} V_t = 0 \right\} \\
&= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} |\nabla \phi_t|^2 \, d\mu_t(x)\, dt \;:\; \partial_t \mu_t + \operatorname{div} \mu_t \nabla \phi_t = 0 \right\} \\
&= 2 \inf \left\{ \int_0^1 \underbrace{\sup_{\phi} \langle \partial_t \mu_t, \phi \rangle - \int_{\mathbb{R}^d} \frac{1}{2} |\nabla \phi|^2 \, d\mu_t(x)}_{A^*(\mu_t, \partial_t \mu_t)} \; dt \right\}.
\end{aligned}
$$

Discretization (first attempt)

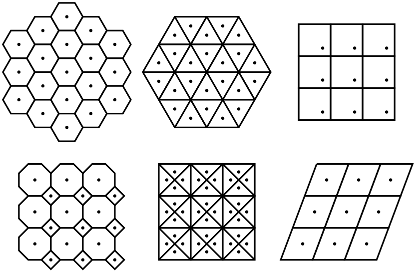
- Closed convex set $\Omega \subset \mathbb{R}^d$.
- Approximate $(\mathcal{P}(\Omega), W)$ using a discrete space.
- Idea: Pick $x_1, \ldots, x_N \in \Omega$.
- $X_N = \{\sum_{i=1}^N \alpha_i \delta_{x_i} \,:\, \sum_i \alpha_i = 1\}$.
- The metric space (X_N, W) Gromov-Hausdorff converges to $(\mathcal{P}(\Omega), W)$ as long as $\lim_{N\to\infty}\{x_i\}_{i=1}^N = \Omega$.

Discretization (first attempt)

- Closed convex set $\Omega \subset \mathbb{R}^d$.
- Approximate $(\mathcal{P}(\Omega), W)$ using a discrete space.
- Idea: Pick $x_1, \ldots, x_N \in \Omega$.
- $X_N = \{\sum_{i=1}^N \alpha_i \delta_{x_i} \,:\, \sum_i \alpha_i = 1\}$.
- The metric space (X_N, W) Gromov-Hausdorff converges to $(\mathcal{P}(\Omega), W)$ as long as $\lim_{N\to\infty} \{x_i\}_{i=1}^N = \Omega$.
- Problem: (X_N, W) doesn't have geodesics.
- (X_2, W) has no nontrivial curves of finite length.
- $W((1-t)\delta_0 + t\delta_1, (1-s)\delta_0 + s\delta_1) = \sqrt{|t-s|}$.

Discretization

- Finite volume method: **mesh** $\mathcal{T} = \{K \subset \Omega \ : \ K \in \mathcal{T}\}$ finite partition of Ω into convex sets
- Control points $\{x_K \in K \ : \ K \in \mathcal{T}\}$
- interfaces $(K|L) = \overline{K} \cap \overline{L}$ between neighboring cells
- distances $d_{KL} = |x_K - x_L|$
- Admissibility condition: $x_K - x_L \perp (K|L)$
- Piecewise constant probability measures $\mathcal{P}(\mathcal{T}) = \{\mu(dx) = \sum_{K\in\mathcal{T}} \rho(K) 1_K \, dx \ : \ \sum_{K\in\mathcal{T}} |K|\rho(K) = 1\}$



Discrete Kantorovich distance

- Discretize continuity equation:

$$|K|\partial_t \rho_t(K) + \sum_{L\sim K} \frac{|(K|L)|}{d_{KL}} \widehat{\rho}_t(K,L)(\phi_t(L) - \phi_t(K)) = 0 \qquad (*)$$

- Here $\phi : \mathcal{T} \to \mathbb{R}$ is a potential.

Discrete Kantorovich distance

- Discretize continuity equation:

$$|K|\partial_t\rho_t(K) + \sum_{L\sim K}\frac{|(K|L)|}{d_{KL}}\widehat{\rho}_t(K,L)(\phi_t(L)-\phi_t(K)) = 0 \qquad (*)$$

- Here $\phi : \mathcal{T} \to \mathbb{R}$ is a potential.
- $\widehat{\rho}_t(K,L) = \theta(\rho(K),\rho(L))$ is a concave mean.
- Examples:

$$\begin{aligned}
\theta(s,t) =& (s+t)/2 \\
\theta(s.t) =& \sqrt{st} \\
\theta(s,t) =& \frac{2st}{s+t} \\
\theta(s,t) =& \frac{s-t}{\log s - \log t} = \int_0^1 s^p t^{1-p}\, dp
\end{aligned}$$

- Discretize weighted H^1-seminorm:

$$\mathcal{A}_{\mathcal{T}}(\rho_t, \phi) = \frac{1}{4} \sum_{K,L} \frac{|(K|L)|}{d_{KL}} \widehat{\rho}_t(K, L)(\phi_t(L) - \phi_t(K))^2.$$

- Discretize weighted H^1-seminorm:

$$\mathcal{A}_{\mathcal{T}}(\rho_t, \phi) = \frac{1}{4} \sum_{K,L} \frac{|(K|L)|}{d_{KL}} \widehat{\rho}_t(K, L)(\phi_t(L) - \phi_t(K))^2.$$

- Discretize action:

$$\mathcal{A}^*_{\mathcal{T}}(\rho_t, \partial_t \rho_t) = \sup_{\phi} \sum_{K \in \mathcal{T}} |K| \partial_t \rho_t(K) \phi(K) - \mathcal{A}_{\mathcal{T}}(\rho_t, \phi).$$

- Discretize weighted H^1-seminorm:

$$\mathcal{A}_{\mathcal{T}}(\rho_t, \phi) = \frac{1}{4} \sum_{K,L} \frac{|(K|L)|}{d_{KL}} \widehat{\rho_t}(K,L)(\phi_t(L) - \phi_t(K))^2.$$

- Discretize action:

$$\mathcal{A}^*_{\mathcal{T}}(\rho_t, \partial_t \rho_t) = \sup_{\phi} \sum_{K \in \mathcal{T}} |K| \partial_t \rho_t(K) \phi(K) - \mathcal{A}_{\mathcal{T}}(\rho_t, \phi).$$

- Discrete distance:

$$\begin{aligned} \frac{1}{2} \mathcal{W}^2_{\mathcal{T}}(\rho_0, \rho_1) &= \inf_{(\rho_t, \phi_t)_t} \left\{ \int_0^1 \mathcal{A}_{\mathcal{T}}(\rho_t, \phi_t) dt \, : \, (*) \right\} \\ &= \inf_{(\rho_t)_t} \left\{ \int_0^1 \mathcal{A}^*_{\mathcal{T}}(\rho_t, \partial_t \rho_t) \, dt \right\} \end{aligned}$$

Gromov-Hausdorff convergence

Show that:

- As $[\mathcal{T}] = \max_{K \in \mathcal{T}} \operatorname{diam} K \to 0$, for uniformly regular meshes, we have

  $$\lim_{[\mathcal{T}] \to 0} \mathcal{W}_{\mathcal{T}}(P_{\mathcal{T}}\mu_0, P_{\mathcal{T}}\mu_1) = W(\mu_0, \mu_1)$$

  uniformly.
- The near-isometry $P_{\mathcal{T}} : \mathcal{P}(\Omega) \to \mathcal{P}(\mathcal{T})$ is the projection $\mu \mapsto (\rho(K))_{K \in \mathcal{T}} = (\mu(K)/|K|)_{K \in \mathcal{T}}$.
- The metric spaces $(\mathcal{P}(\mathcal{T}), \mathcal{W}_{\mathcal{T}})$ Gromov-Hausdorff converge to $(\mathcal{P}(\Omega), W)$.

Gromov-Hausdorff convergence

Show that:

- As $[\mathcal{T}] = \max_{K \in \mathcal{T}} \operatorname{diam} K \to 0$, for uniformly regular meshes, we have
$$\lim_{[\mathcal{T}] \to 0} \mathcal{W}_{\mathcal{T}}(P_{\mathcal{T}}\mu_0, P_{\mathcal{T}}\mu_1) = W(\mu_0, \mu_1)$$
uniformly.
- The near-isometry $P_{\mathcal{T}} : \mathcal{P}(\Omega) \to \mathcal{P}(\mathcal{T})$ is the projection $\mu \mapsto (\rho(K))_{K \in \mathcal{T}} = (\mu(K)/|K|)_{K \in \mathcal{T}}$.
- The metric spaces $(\mathcal{P}(\mathcal{T}), \mathcal{W}_{\mathcal{T}})$ Gromov-Hausdorff converge to $(\mathcal{P}(\Omega), W)$.

Theorem (Gigli-Maas 2013)

Let $(\mathcal{T}_k)_{k \in \mathbb{N}}$ be refining cubic meshes on the torus $\mathbb{T}^d$. Then $(\mathcal{P}(\mathcal{T}_k), \mathcal{W}_{\mathcal{T}_k}) \to (\mathcal{P}(\mathbb{T}^d), W)$ Gromov-Hausdorff with near-isometry $P_{\mathcal{T}}$.

Upper bound for general meshes

- Show that for $\mu_0, \mu_1 \in \mathcal{P}(\Omega)$ there are curves $(\rho_t^{\mathcal{T}})_{t\in[0,1]} \subset \mathcal{P}(\mathcal{T})$ connecting $\rho_0^{\mathcal{T}} = P_{\mathcal{T}}\mu_0$ and $\rho_1^{\mathcal{T}} = P_{\mathcal{T}}\mu_1$, such that

$$\limsup_{[\mathcal{T}]\to 0} \int_0^1 \mathcal{A}_{\mathcal{T}}^*(\rho_t^{\mathcal{T}}, \partial_t \rho_t^{\mathcal{T}})\, dt \le W^2(\mu_0, \mu_1).$$

Upper bound for general meshes

- Show that for $\mu_0, \mu_1 \in \mathcal{P}(\Omega)$ there are curves $(\rho_t^{\mathcal{T}})_{t\in[0,1]} \subset \mathcal{P}(\mathcal{T})$ connecting $\rho_0^{\mathcal{T}} = P_{\mathcal{T}}\mu_0$ and $\rho_1^{\mathcal{T}} = P_{\mathcal{T}}\mu_1$, such that

$$\limsup_{[\mathcal{T}]\to 0} \int_0^1 \mathcal{A}^*_{\mathcal{T}}(\rho_t^{\mathcal{T}}, \partial_t\rho_t^{\mathcal{T}})\, dt \leq W^2(\mu_0, \mu_1).$$

- Take W-geodesic $(\mu_t)_t$, apply Neumann heat flow for small $a > 0$ to obtain $(H_a\mu_t)_t$.
- Heat flow is W-contraction. (Jordan, Kinderlehrer, Otto 1998)

Upper bound for general meshes

- Show that for $\mu_0, \mu_1 \in \mathcal{P}(\Omega)$ there are curves $(\rho_t^{\mathcal{T}})_{t\in[0,1]} \subset \mathcal{P}(\mathcal{T})$ connecting $\rho_0^{\mathcal{T}} = P_{\mathcal{T}}\mu_0$ and $\rho_1^{\mathcal{T}} = P_{\mathcal{T}}\mu_1$, such that

$$\limsup_{[\mathcal{T}]\to 0} \int_0^1 \mathcal{A}^*_{\mathcal{T}}(\rho_t^{\mathcal{T}}, \partial_t \rho_t^{\mathcal{T}})\, dt \le W^2(\mu_0, \mu_1).$$

- Take W-geodesic $(\mu_t)_t$, apply Neumann heat flow for small $a > 0$ to obtain $(H_a\mu_t)_t$.
- Heat flow is W-contraction. (Jordan, Kinderlehrer, Otto 1998)
- Because $H_a\mu_t$ is smooth with lower bound, the continuity equation

$$\begin{cases} \partial_t H_a\mu_t + \operatorname{div}(H_a\mu_t \nabla\phi_t) = 0 \\ \nabla\phi_t \cdot n = 0 \text{ on } \partial\Omega \end{cases}$$

is elliptic.

- Solve the discrete continuity equation for $\rho_t^{\mathcal{T}} = P_{\mathcal{T}} H_a \mu_t$ and $\partial_t \rho_t^{\mathcal{T}} = P_{\mathcal{T}} H_a \partial_t \mu_t$ to obtain $\phi_t^{\mathcal{T}} : \mathcal{T} \to \mathbb{R}$.
- Elliptic finite volume estimates give for regular meshes

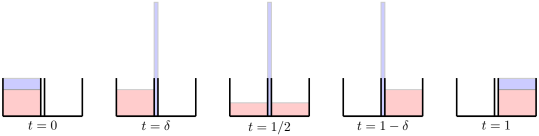
$$\begin{aligned} |\mathcal{A}^*_{\mathcal{T}}(\rho_t^{\mathcal{T}}, \partial_t \rho_t^{\mathcal{T}}) - A^*(H_a \mu_t, \partial_t H_a \mu_t)| &= \left| \mathcal{A}_{\mathcal{T}}(\rho_t^{\mathcal{T}}, \phi_t^{\mathcal{T}}) - \int_\Omega |\nabla \phi_t|^2 \, dH_a \mu_t \right| \\ &\leq C(a) \|\partial_t H_a \mu_t\|^2_{L^2}[\mathcal{T}] \end{aligned}$$

- This gives competing curves between $P_{\mathcal{T}} H_a \mu_0$ and $P_{\mathcal{T}} H_a \mu_1$ after time regularization.
- Connect $P_{\mathcal{T}} \mu_0$ with $P_{\mathcal{T}} H_a \mu_0$ with cost $O(\sqrt{a})$.
- $\Rightarrow$ Uniform upper bound.

Lower bound?

Theorem (First counterexample, G-Kopfer-Maas 2017)

For the alternating large-small mesh with ratio $b \in (0,1)$ on $[0,1]$, we have $\limsup_{[\mathcal{T}]\to 0} \mathcal{W}_{\mathcal{T}}(P_{\mathcal{T}}\mu_0, P_{\mathcal{T}}\mu_1) < W(\mu_0,\mu_1)$ whenever $\mu_0 \neq \mu_1$. In fact, $\lim_{b\to 0} \limsup_{[\mathcal{T}]\to 0} \mathcal{W}_{\mathcal{T}}(P_{\mathcal{T}}\delta_0, P_{\mathcal{T}}\delta_1) = 0$.



$$\mathcal{A}^*_{\mathcal{T}}(\rho,\partial_t\rho) = \sup_{\phi}\langle \partial_t\rho, \phi\rangle - \frac{N}{2}\sum_{k=1}^{N} \widehat{\rho}(k-1,k)(\phi(k)-\phi(k-1))^2$$

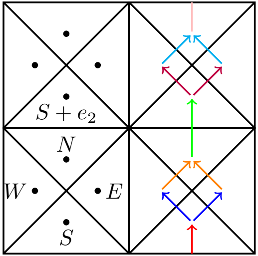
Idea: $\widehat{\rho}(K,L)$ is always larger than it should be.

$\Rightarrow \mathcal{A}^*_{\mathcal{T}}$ is always smaller than it should be.

Lower bound with symmetry?

Theorem (Second counterexample, G-Kopfer-Maas 2018)

For the periodic mesh consisting of 45-45-90 *triangles, there are* $\mu_0, \mu_1 \in \mathcal{P}([0,1]^2)$ *such that* $\limsup_{[\mathcal{T}]\to 0} \mathcal{W}_{\mathcal{T}}(P_{\mathcal{T}}\mu_0, P_{\mathcal{T}}\mu_1) < W(\mu_0, \mu_1)$*.*
In fact, this happens whenever mass is transported in cardinal directions.



$$\mathcal{A}^*_{\mathcal{T}}(\rho, \partial_t \rho) = \sup_{\phi} \langle \partial_t \rho, \phi \rangle - \frac{1}{4} \sum_{K \sim L} \widehat{\rho}(K, L)(\phi(L) - \phi(K))^2$$

- Lower bound exists for cubic mesh.
- No lower bound for right isosceles triangular mesh.
- What is a geometric condition on the mesh to guarantee convergence?

- Lower bound exists for cubic mesh.
- No lower bound for right isosceles triangular mesh.
- What is a geometric condition on the mesh to guarantee convergence?
- (COM) Center-of-mass condition: For all $K \sim L$,

$$\fint_{(K|L)} x \, d\mathcal{H}^{d-1} = (x_K + x_L)/2.$$

- Lower bound exists for cubic mesh.
- No lower bound for right isosceles triangular mesh.
- What is a geometric condition on the mesh to guarantee convergence?
- (COM) Center-of-mass condition: For all $K \sim L$,

$$\fint_{(K|L)} x \, d\mathcal{H}^{d-1} = (x_K + x_L)/2.$$

- (ISO) Isometry condition: For all K,

$$\sum_{L \sim K} n_{KL} \otimes n_{KL} d_{KL} |(K|L)| = 2|K| Id.$$

- Lower bound exists for cubic mesh.
- No lower bound for right isosceles triangular mesh.
- What is a geometric condition on the mesh to guarantee convergence?
- (COM) Center-of-mass condition: For all $K \sim L$,

$$\fint_{(K|L)} x \, d\mathcal{H}^{d-1} = (x_K + x_L)/2.$$

- (ISO) Isometry condition: For all K,

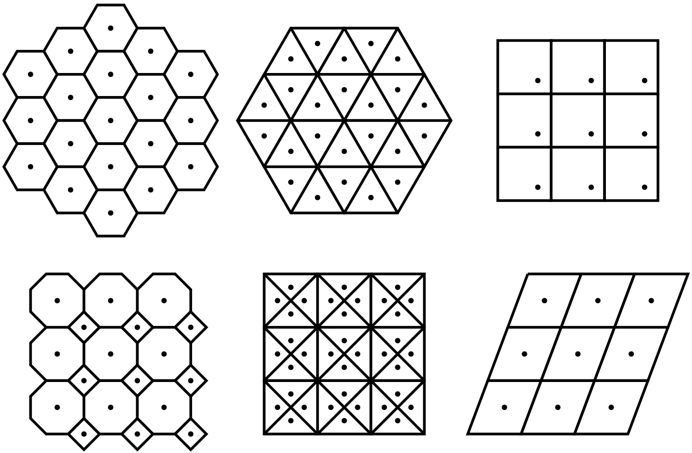
$$\sum_{L \sim K} n_{KL} \otimes n_{KL} d_{KL} |(K|L)| = 2|K| Id.$$

- COM $\Rightarrow$ ISO:

$$\begin{aligned}\sum_{L\sim K} n_{KL} \otimes n_{KL} d_{KL} |(K|L)| &\underset{COM}{=} 2 \sum_{L\sim K} n_{KL} \otimes (x - x_K) |(K|L)| \\ &= 2 \int_{\partial K} n \otimes (x - x_K) \, d\mathcal{H}^{d-1} \\ &= 2|K| Id.\end{aligned}$$

Theorem (G-Kopfer-Maas 2018)

Under ISO condition, for uniformly regular admissible meshes on convex closed full-dimensional $\Omega \subset \mathbb{R}^d$, $(\mathcal{P}(\mathcal{T}), \mathcal{W}_{\mathcal{T}})) \to (\mathcal{P}(\Omega), W)$ *Gromov-Hausdorff with near-isometry* $P_{\mathcal{T}}$.



Proof of lower bound

First show that for $\phi \in C^1(\Omega)$ with $\nabla\phi \cdot n = 0$ on $\partial\Omega$, the discretized version $\phi^{\mathcal{T}}(K) = \phi(x_K)$ satisfies $\limsup_{[\mathcal{T}]\to 0} \mathcal{A}_{\mathcal{T}}(\rho^{\mathcal{T}}, \phi^{\mathcal{T}}) \leq \int_\Omega |\nabla\phi|^2\, d\mu$ whenever $\rho^{\mathcal{T}} \stackrel{*}{\rightharpoonup} \mu$.

$$
\begin{aligned}
&\frac{1}{2}\sum_{K,L} \frac{|(K|L)|}{d_{KL}} \widehat{\rho^{\mathcal{T}}}(K,L)(\phi(x_L)-\phi(x_K))^2 \\
\leq&\frac{1}{4}\sum_{K,L} \frac{|(K|L)|}{d_{KL}} (\rho^{\mathcal{T}}(K)+\rho^{\mathcal{T}}(L))(\phi(x_L)-\phi(x_K))^2 \\
=&\frac{1}{2}\sum_{K} \rho^{\mathcal{T}}(K) \sum_{L\sim K} \frac{|(K|L)|}{d_{KL}} (\phi(x_L)-\phi(x_K))^2 \\
\approx&\frac{1}{2}\sum_{K} \rho^{\mathcal{T}}(K)\nabla\phi(x_K)\otimes\nabla\phi(x_K) : \sum_{L\sim K} |(K|L)| d_{KL} n_{KL}\otimes n_{KL} \\
=&\sum_{K} \rho^{\mathcal{T}}(K)|K||\nabla\phi(x_K)|^2 \approx \langle \rho^{\mathcal{T}}, |\nabla\phi|^2\rangle \to \int_\Omega |\nabla\phi|^2\, d\mu.
\end{aligned}
$$

Less symmetric meshes

- We can treat some meshes without the isotropy condition by using weighted means
- In $\mathcal{A}^*$, use weighted mean $\widehat{\rho}(K,L) \approx \lambda_{KL}\rho(K) + (1-\lambda_{KL})\rho(L)$
- (WCOM) Weighted center-of-mass condition: For all $K \sim L$,

$$\fint_{(K|L)} x \, d\mathcal{H}^{d-1} = \lambda_{KL} x_K + (1-\lambda_{KL}) x_L.$$

- (WISO) Weighted isometry condition: For all $K \in \mathcal{T}$,

$$\sum_{L\sim K} n_{KL} \otimes n_{KL} d_{KL} |(K|L)| \lambda_{KL} = |K| Id.$$

- WCOM $\Rightarrow$ WISO.

Theorem (G-Kopfer-Mass 2018)

Under WISO condition, for uniformly regular admissible meshes on convex closed full-dimensional $\Omega \subset \mathbb{R}^d$, $(\mathcal{P}(\mathcal{T}), \mathcal{W}_{\mathcal{T}}) \to (\mathcal{P}(\Omega), W)$ *Gromov-Hausdorff with near-isometry* $P_{\mathcal{T}}$.