Approximate Minimum-Weight Partial Matching under Translation

Pankaj Agarwal [1] Haim Kaplan [2] Geva Kipper [2]
Wolfgang Mulzer [3] Günter Rote [3] Micha Sharir [2] Allen Xiao [1]

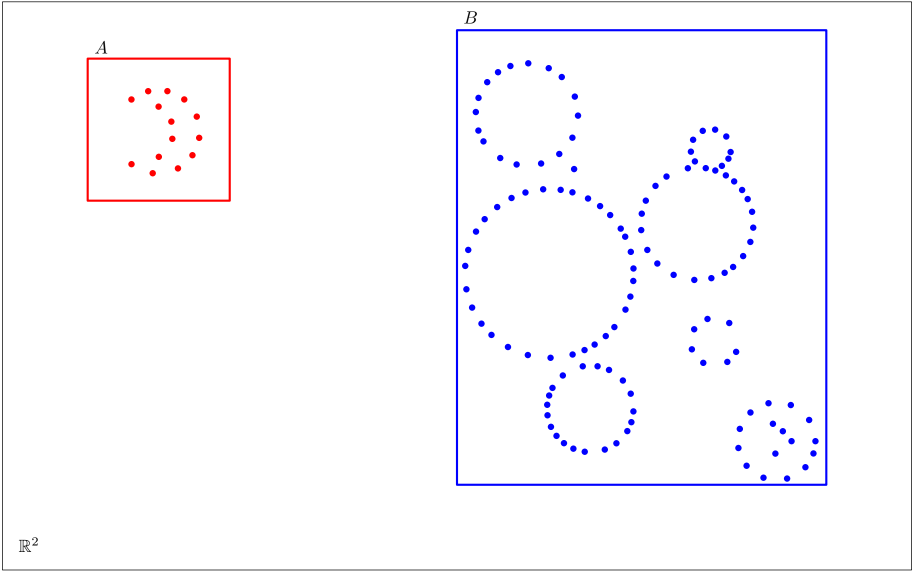
[1] Duke University

[2] Tel Aviv University

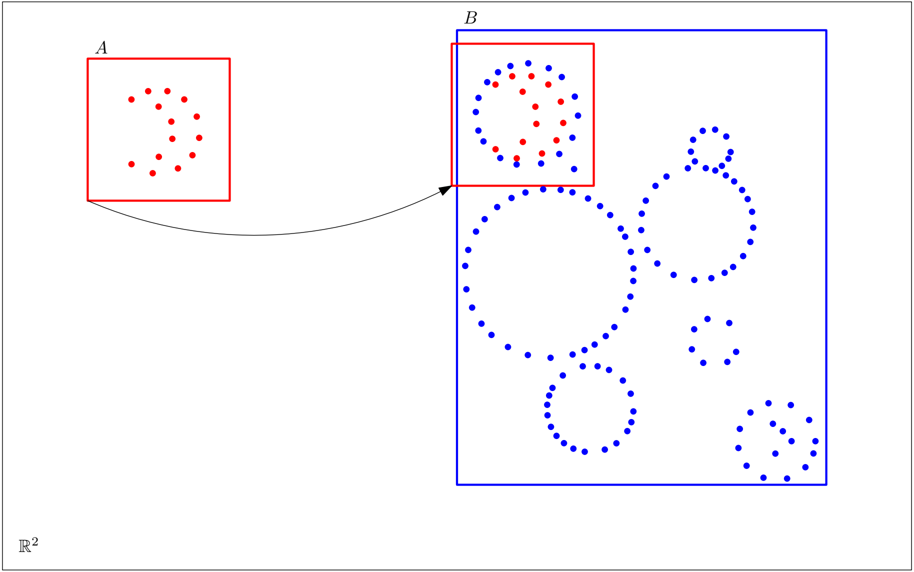
[3] Freie Universität Berlin

December 2018

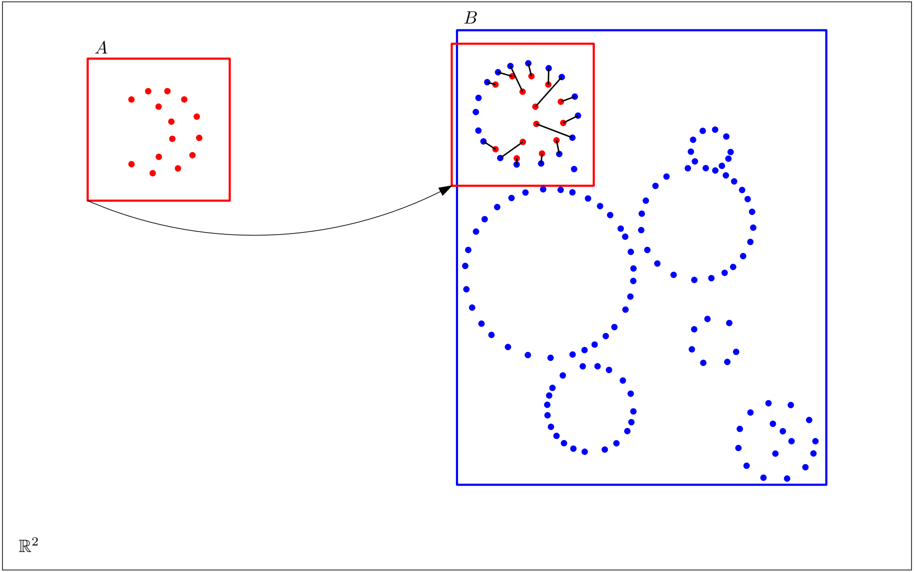
Example



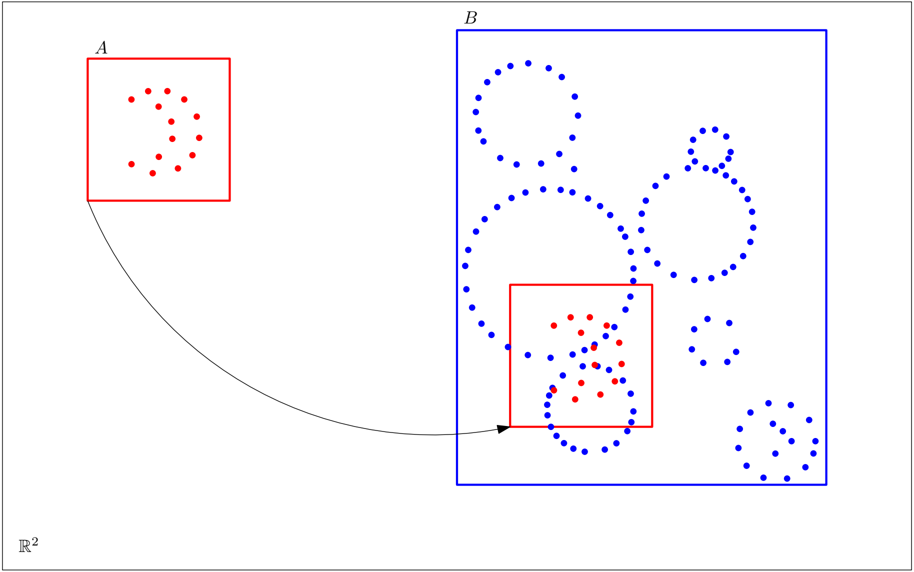
Example



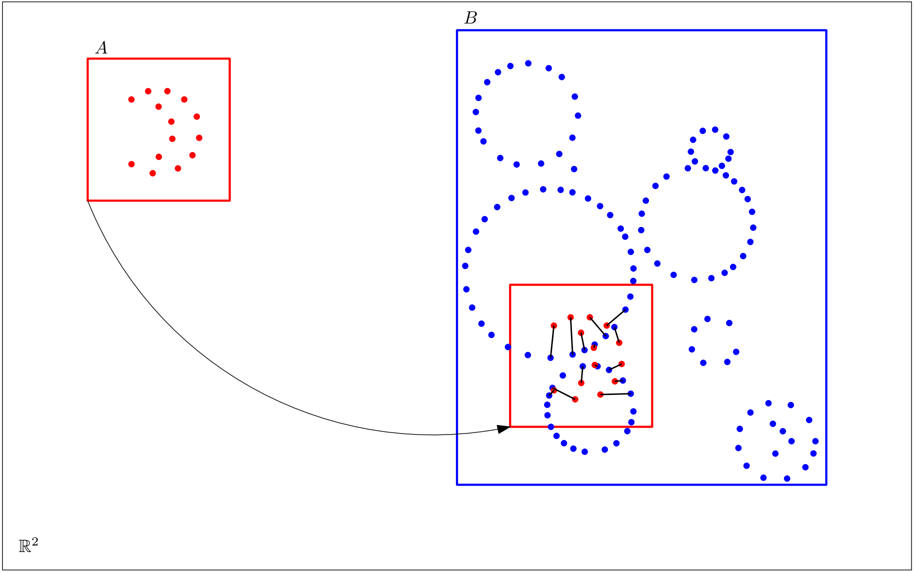
Example



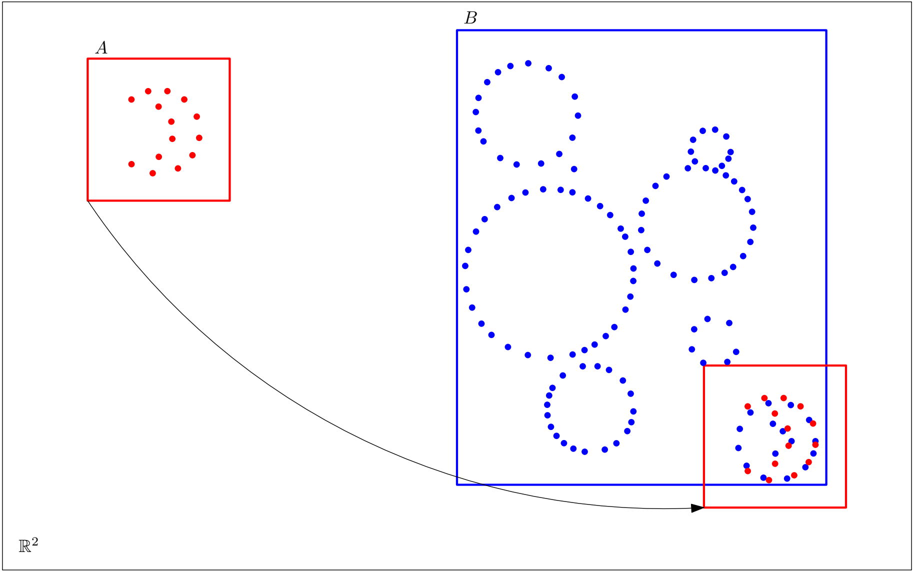
Example



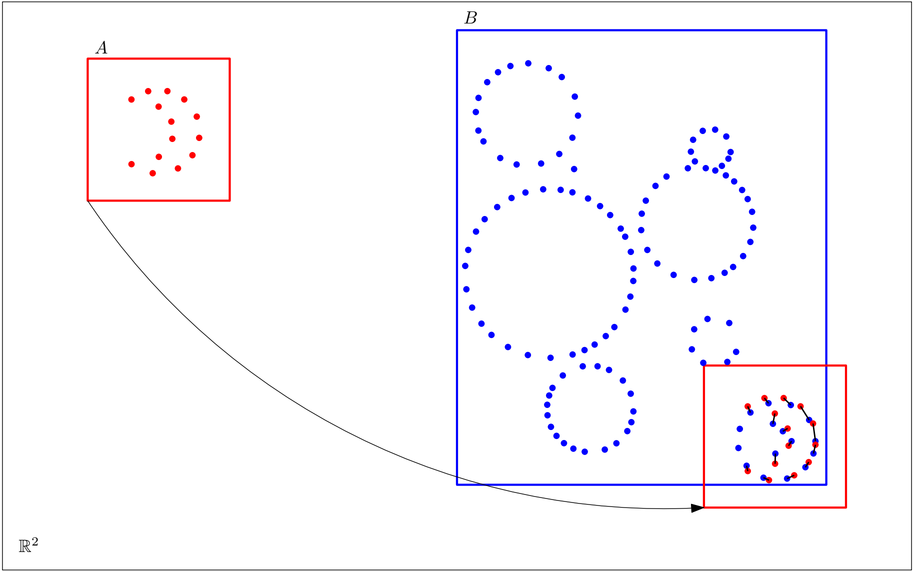
Example



Example



Example



Cost function

- For k-matching M and translation t, define the *L_p-cost*:

$$\text{cost}(M,t) := \left[\frac{1}{k}\sum_{(a,b)\in M}\|a+t-b\|^p\right]^{1/p}$$

- $p = 2$: root-mean-squared cost
- For fixed t, minimum M computable in $\text{poly}(k,m,n)$ time, e.g. by Hungarian algorithm.

Cost function

- For k-matching M and translation t, define the *L_p-cost*:

$$\mathrm{cost}(M,t) := \left[\frac{1}{k} \sum_{(a,b)\in M} \|a+t-b\|^p \right]^{1/p}$$

- $p = 2$: root-mean-squared cost
- For fixed t, minimum M computable in $\mathrm{poly}(k,m,n)$ time, e.g. by Hungarian algorithm.

Cost function

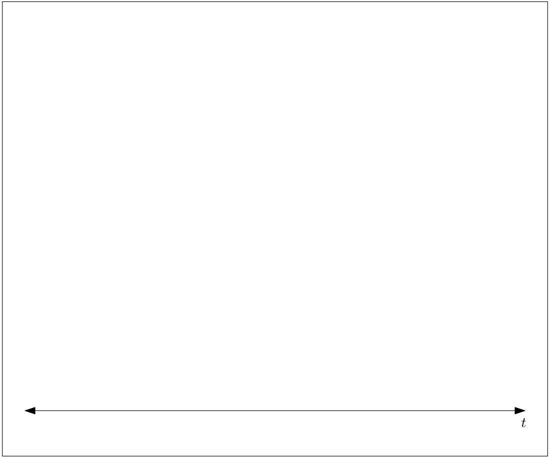
- For k-matching M and translation t, define the *L_p-cost*:

$$\operatorname{cost}(M,t) := \left[\frac{1}{k}\sum_{(a,b)\in M}\|a+t-b\|^p\right]^{1/p}$$

- $p=2$: root-mean-squared cost
- For fixed t, minimum M computable in $\operatorname{poly}(k,m,n)$ time, e.g. by Hungarian algorithm.

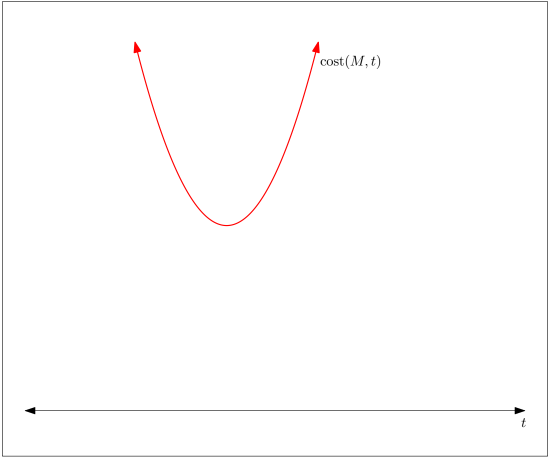
Minimization diagram $\mathcal{M}$

$$\mathrm{cost}^*(t) := \min_{k\text{-matching } M} \mathrm{cost}(M, t)$$



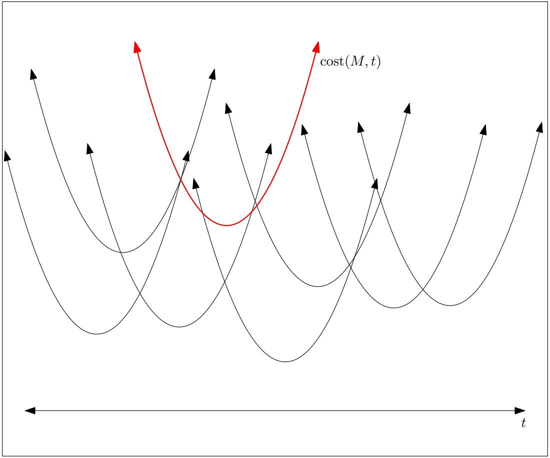
Minimization diagram $\mathcal{M}$

$$\text{cost}^*(t) := \min_{k\text{-matching } M} \text{cost}(M, t)$$



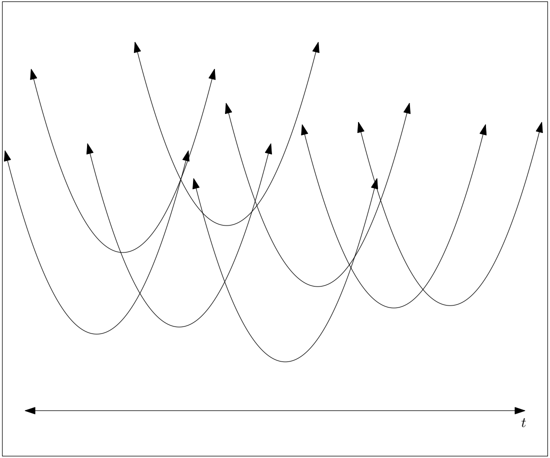
Minimization diagram $\mathcal{M}$

$$\mathrm{cost}^*(t) := \min_{k\text{-matching } M} \mathrm{cost}(M,t)$$



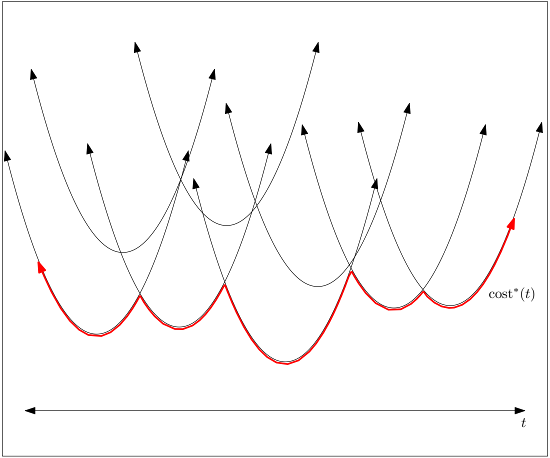
Minimization diagram $\mathcal{M}$

$$\mathrm{cost}^*(t) := \min_{k\text{-matching } M} \mathrm{cost}(M, t)$$



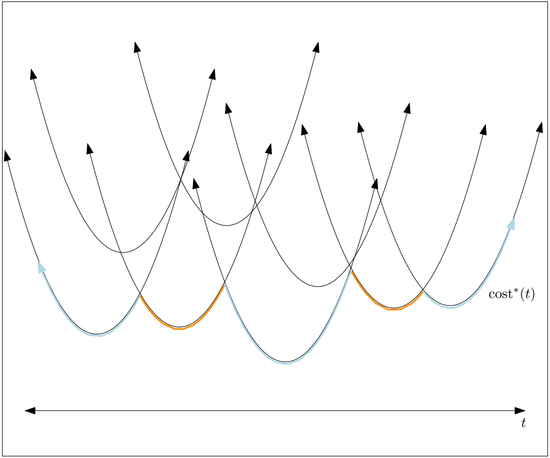
Minimization diagram $\mathcal{M}$

$$\text{cost}^*(t) := \min_{k\text{-matching } M} \text{cost}(M, t)$$



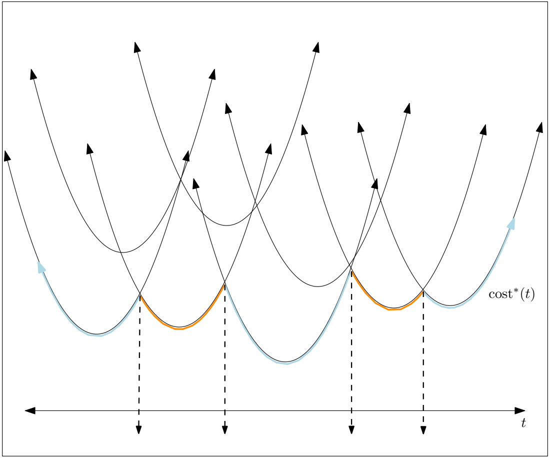
Minimization diagram $\mathcal{M}$

$$\text{cost}^*(t) := \min_{k\text{-matching } M} \text{cost}(M, t)$$



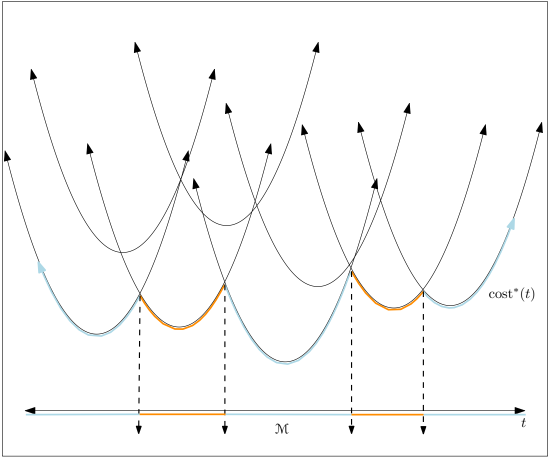
Minimization diagram $\mathcal{M}$

$$\mathrm{cost}^*(t) := \min_{k\text{-matching } M} \mathrm{cost}(M, t)$$



Minimization diagram $\mathcal{M}$

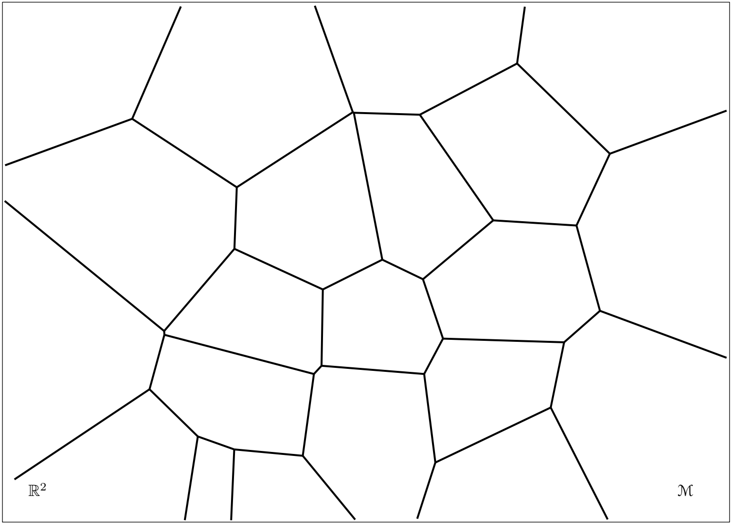
$$\mathrm{cost}^*(t) := \min_{k\text{-matching } M} \mathrm{cost}(M, t)$$



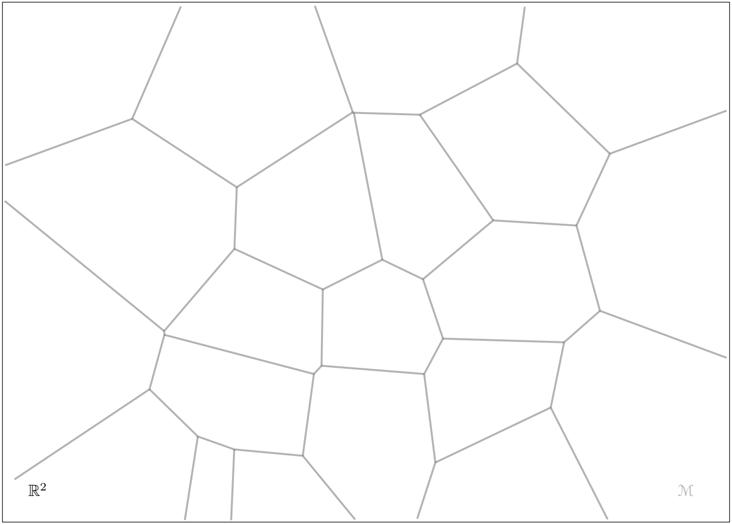
Questions and prior results

1. What is the *combinatorial complexity of $\mathcal{M}$*? Is it polynomial?
   - [Rote 10]: for $p = 2$ and $k = m$,
     a line crosses at most $m(n - m) + 1$ cells
   - [Ben-Avraham *et al.* 14]: for $p = 2$ and $k = m$,
     $O(n^2 m^{3.5}(e \ln m + e)^m)$
2. How quickly can we *compute t^**, a global minimum of $\mathrm{cost}^*(t)$?
   - Explore $\mathcal{M}$, use static algorithm within cells.

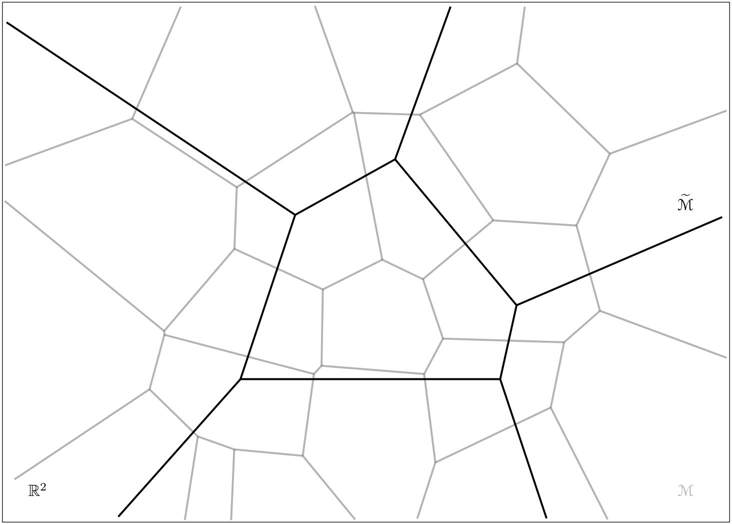
Approximating $\mathcal{M}$



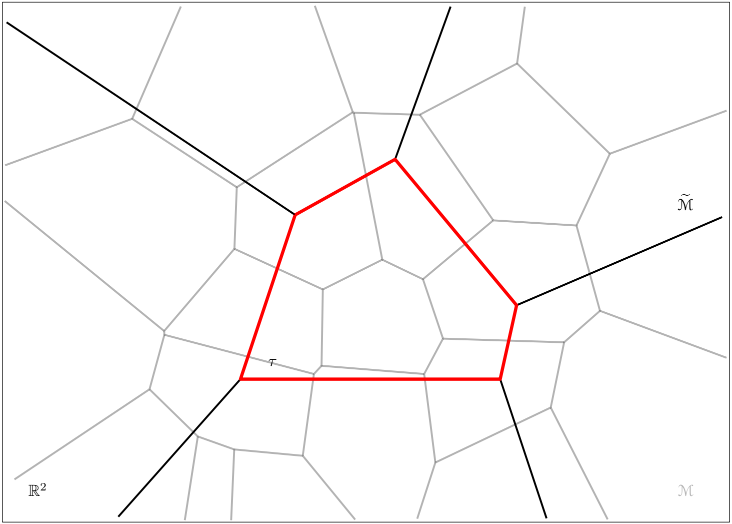
Approximating $\mathcal{M}$



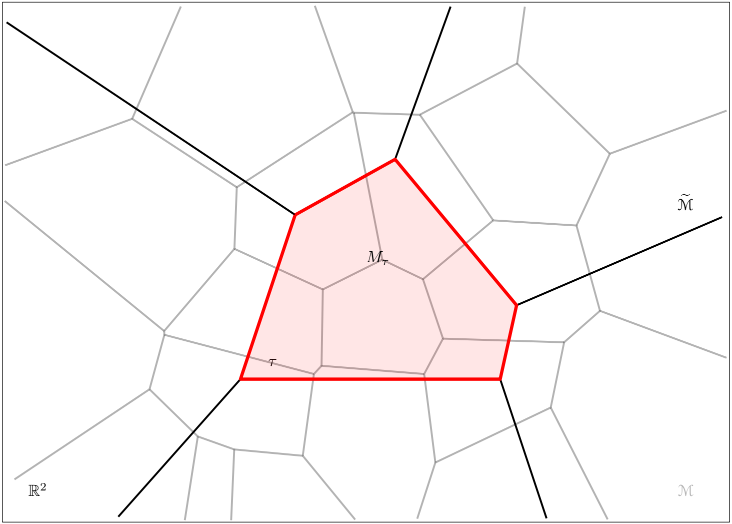
Approximating $\mathcal{M}$



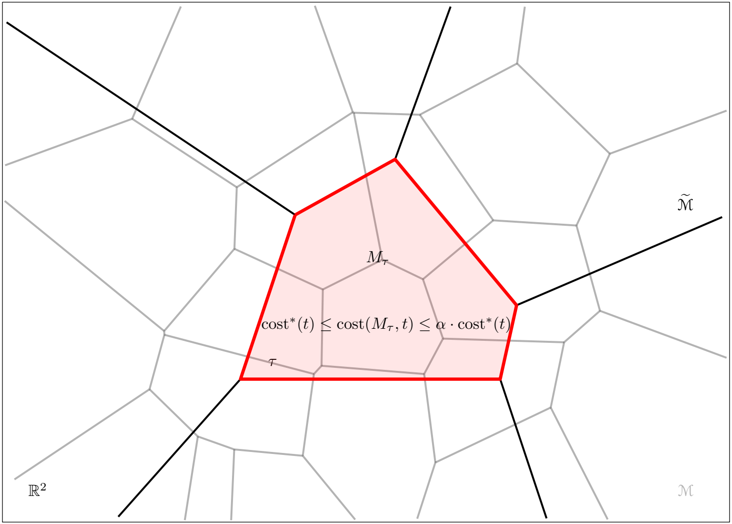
Approximating $\mathcal{M}$



Approximating $\mathcal{M}$



Approximating $\mathcal{M}$



Our results (approximation helps)

1. What is the *combinatorial complexity of* $\mathcal{M}$? Is it polynomial?

Theorem

In $\mathrm{poly}(k, m, n, \varepsilon^{-1})$ *time, can construct a* $(1+\varepsilon)$ *approximate diagram* $\widetilde{\mathcal{M}}$ *of complexity* $O((mn/k)\varepsilon^{-2}\log\varepsilon^{-1})$*.*

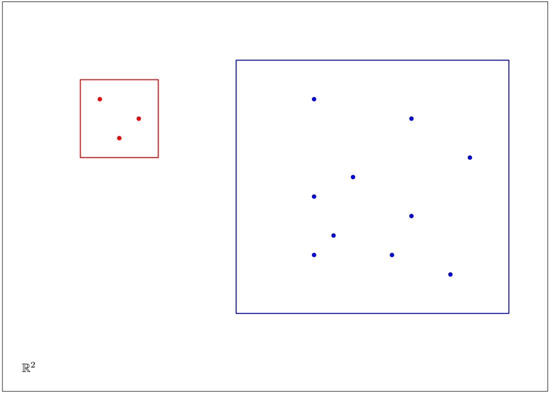
2. How quickly can we *compute* t^*, a global minimum of $\mathrm{cost}^*(t)$?

Theorem

In $\mathrm{poly}(k, m, n, \varepsilon^{-1})$ *time, can compute a* $(1+\varepsilon)$ *approximation to* $\mathrm{cost}^*(t^*)$ *by exploring the cells of* $\widetilde{\mathcal{M}}$*.*

Point-to-point translations [Cabello *et al.* 08]

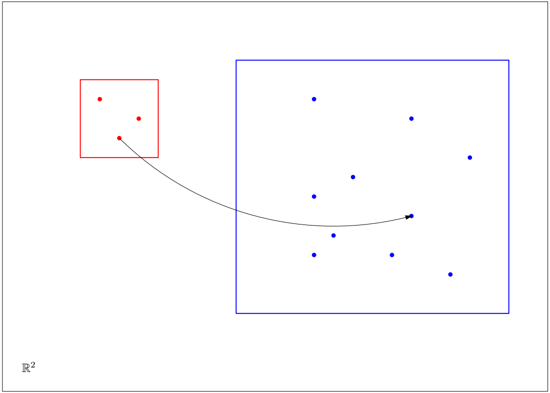
- *Point-to-point translations*: $T := \{t_{ba} = (b - a) \mid a \in A, b \in B\}$



- Claim: Let $\widetilde{\mathcal{M}}$ be $\mathrm{Vor}(T)$, with each $\mathrm{VorRegion}(t_{ba})$ assigned the optimal k-matching at t_{ba}. Then, $\widetilde{\mathcal{M}}$ is a $O(1)$-approximate diagram.

Point-to-point translations [Cabello *et al.* 08]

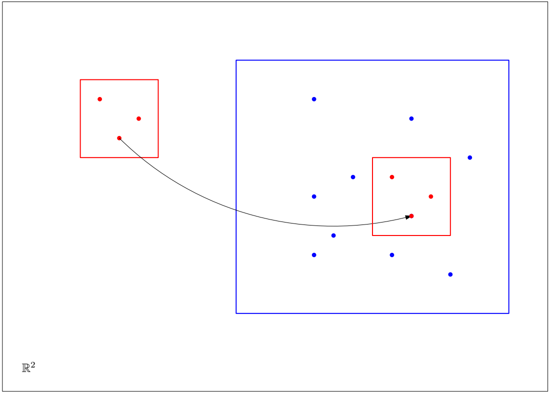
- *Point-to-point translations*: $T := \{t_{ba} = (b - a) \mid a \in A, b \in B\}$



- Claim: Let $\widetilde{\mathcal{M}}$ be $\mathrm{Vor}(T)$, with each $\mathrm{VorRegion}(t_{ba})$ assigned the optimal k-matching at t_{ba}. Then, $\widetilde{\mathcal{M}}$ is a $O(1)$-approximate diagram.

Point-to-point translations [Cabello *et al.* 08]

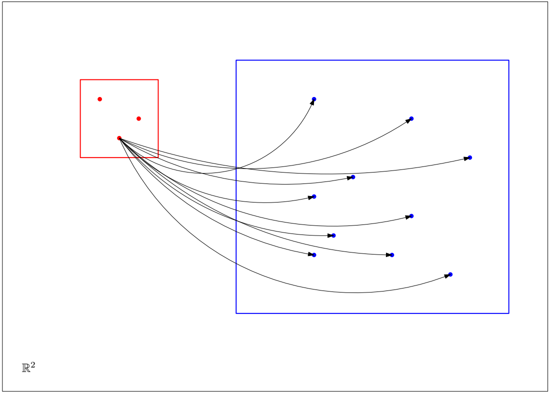
- *Point-to-point translations*: $T := \{t_{ba} = (b - a) \mid a \in A, b \in B\}$



- Claim: Let $\widetilde{\mathcal{M}}$ be $\mathrm{Vor}(T)$, with each $\mathrm{VorRegion}(t_{ba})$ assigned the optimal k-matching at t_{ba}. Then, $\widetilde{\mathcal{M}}$ is a $O(1)$-approximate diagram.

Point-to-point translations [Cabello *et al.* 08]

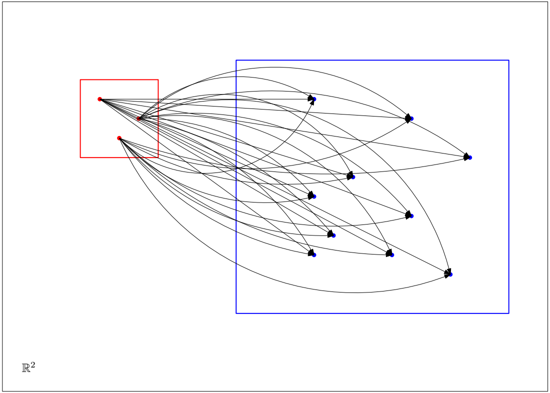
- *Point-to-point translations*: $T := \{t_{ba} = (b - a) \mid a \in A, b \in B\}$



- Claim: Let $\widetilde{\mathcal{M}}$ be $\mathrm{Vor}(T)$, with each $\mathrm{VorRegion}(t_{ba})$ assigned the optimal k-matching at t_{ba}. Then, $\widetilde{\mathcal{M}}$ is a $O(1)$-approximate diagram.

Point-to-point translations [Cabello *et al.* 08]

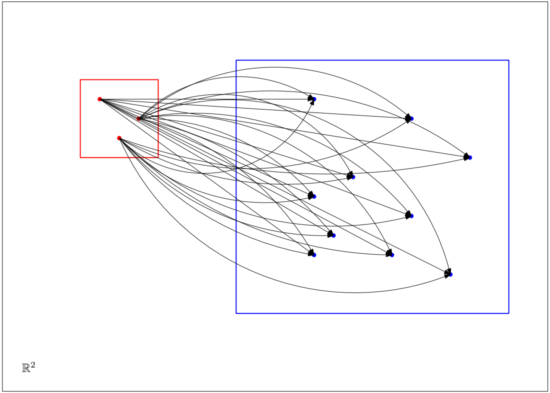
- *Point-to-point translations*: $T := \{t_{ba} = (b - a) \mid a \in A, b \in B\}$



- Claim: Let $\widetilde{\mathcal{M}}$ be $\mathrm{Vor}(T)$, with each $\mathrm{VorRegion}(t_{ba})$ assigned the optimal k-matching at t_{ba}. Then, $\widetilde{\mathcal{M}}$ is a $O(1)$-approximate diagram.

Point-to-point translations [Cabello *et al.* 08]

- *Point-to-point translations*: $T := \{t_{ba} = (b - a) \mid a \in A, b \in B\}$



- Claim: Let $\widetilde{\mathcal{M}}$ be $\mathrm{Vor}(T)$, with each $\mathrm{VorRegion}(t_{ba})$ assigned the optimal k-matching at t_{ba}. Then, $\widetilde{\mathcal{M}}$ is a $O(1)$-approximate diagram.

Lipschitz continuity for L_p-cost

$$\operatorname{cost}(M,t) := \left[\frac{1}{k}\sum_{(a,b)\in M}\|a+t-b\|^p\right]^{1/p}$$

Lemma (Lipschitz condition)

Given $t, \Delta \in \mathbb{R}^2$, let M_t be the optimal k-matching at t.
Then $\operatorname{cost}(M_t, t+\Delta) \le \operatorname{cost}^*(t) + \|\Delta\|$.

Proof of approximation for T

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.

$$
\begin{aligned}
\text{cost}^*(t) &= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|a + t - b\|^p \right]^{1/p} \\
&= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|t - t_{ba}\|^p \right]^{1/p} \\
&\geq \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|t - t_0\|^p \right]^{1/p} \\
&= \|t - t_0\|
\end{aligned}
$$

Proof of approximation for T

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.

$$\begin{aligned}
\mathrm{cost}^*(t) &= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|a + t - b\|^p \right]^{1/p} \\
&= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|t - t_{ba}\|^p \right]^{1/p} \\
&\geq \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|t - t_0\|^p \right]^{1/p} \\
&= \|t - t_0\|
\end{aligned}$$

Proof of approximation for T

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.

$$
\begin{aligned}
\mathrm{cost}^*(t) &= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b) \in M} \|a + t - b\|^p \right]^{1/p} \\
&= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b) \in M} \|t - t_{ba}\|^p \right]^{1/p} \\
&\geq \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b) \in M} \|t - t_0\|^p \right]^{1/p} \\
&= \|t - t_0\|
\end{aligned}
$$

Proof of approximation for T

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.

$$\begin{aligned}
\text{cost}^*(t) &= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|a + t - b\|^p \right]^{1/p} \\
&= \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|t - {\color{red} t_{ba}}\|^p \right]^{1/p} \\
&\geq \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|t - {\color{red} t_0}\|^p \right]^{1/p} \\
&= \|t - t_0\|
\end{aligned}$$

Proof of approximation for T (cont.)

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.
- $\|t - t_0\| \leq \operatorname{cost}^*(t)$

$$\begin{aligned}
\operatorname{cost}(M_{t_0}, t) &\leq \operatorname{cost}^*(t_0) + \|t - t_0\| && \text{(Lipschitz cond.)} \\
&\leq \operatorname{cost}^*(t) + 2\|t - t_0\| && \text{(Lipschitz cond.)} \\
&\leq \operatorname{cost}^*(t) + 2\operatorname{cost}^*(t) \\
&= 3\operatorname{cost}^*(t)
\end{aligned}$$

Proof of approximation for T (cont.)

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.
- $\|t - t_0\| \le \operatorname{cost}^*(t)$

$$\begin{aligned}
\operatorname{cost}(M_{t_0}, t) &\le \operatorname{cost}^*(t_0) + \|t - t_0\| && \text{(Lipschitz cond.)} \\
&\le \operatorname{cost}^*(t) + 2\|t - t_0\| && \text{(Lipschitz cond.)} \\
&\le \operatorname{cost}^*(t) + 2\operatorname{cost}^*(t) \\
&= 3\operatorname{cost}^*(t)
\end{aligned}$$

Proof of approximation for T (cont.)

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.
- $\|t - t_0\| \leq \mathrm{cost}^*(t)$

$$\begin{aligned} \mathrm{cost}(M_{t_0}, t) &\leq \mathrm{cost}^*(t_0) + \|t - t_0\| && \text{(Lipschitz cond.)} \\ &\leq \mathrm{cost}^*(t) + 2\|t - t_0\| && \text{(Lipschitz cond.)} \\ &\leq \mathrm{cost}^*(t) + 2\,\mathrm{cost}^*(t) \\ &= 3\,\mathrm{cost}^*(t) \end{aligned}$$

Proof of approximation for T (cont.)

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.
- $\|t - t_0\| \leq \operatorname{cost}^*(t)$

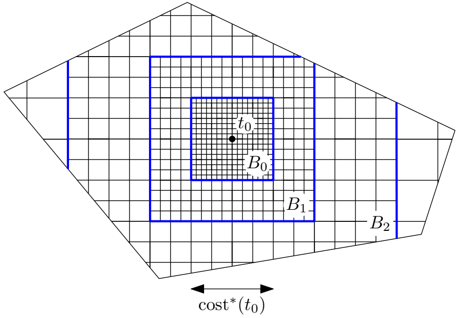
$$\begin{aligned}
\operatorname{cost}(M_{t_0}, t) &\leq \operatorname{cost}^*(t_0) + \|t - t_0\| && \text{(Lipschitz cond.)} \\
&\leq \operatorname{cost}^*(t) + 2\|t - t_0\| && \text{(Lipschitz cond.)} \\
&\leq \operatorname{cost}^*(t) + 2\operatorname{cost}^*(t) \\
&= 3\operatorname{cost}^*(t)
\end{aligned}$$

Proof of approximation for T (cont.)

- Claim: $\widetilde{\mathcal{M}}$ is an $O(1)$-approximate diagram.
- Given $t \in \mathbb{R}^2$, let t_0 be its nearest neighbor in T, and M_{t_0} the optimal matching at t_0.
- $\|t - t_0\| \leq \mathrm{cost}^*(t)$

$$\begin{aligned} \mathrm{cost}(M_{t_0}, t) &\leq \mathrm{cost}^*(t_0) + \|t - t_0\| && \text{(Lipschitz cond.)} \\ &\leq \mathrm{cost}^*(t) + 2\|t - t_0\| && \text{(Lipschitz cond.)} \\ &\leq \mathrm{cost}^*(t) + 2\,\mathrm{cost}^*(t) \\ &= 3\,\mathrm{cost}^*(t) \end{aligned}$$

$O(1) \to (1+\varepsilon)$ approximation

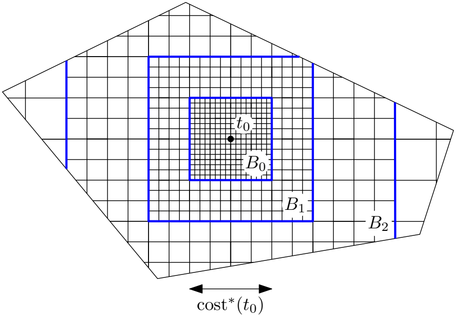


- $(1+\varepsilon)$ approximate diagram of size $O(|T|\varepsilon^{-2}\log\varepsilon^{-1}) = O((mn)\varepsilon^{-2}\log\varepsilon^{-1})$

Theorem

In $\mathrm{poly}(k, m, n, \varepsilon^{-1})$ *time, can construct a* $(1+\varepsilon)$ *approximate diagram* $\widetilde{\mathcal{M}}$ *of complexity* $O((mn/k)\varepsilon^{-2}\log\varepsilon^{-1})$.

$O(1) \to (1+\varepsilon)$ approximation



- $(1+\varepsilon)$ approximate diagram of size $O(|T|\varepsilon^{-2}\log\varepsilon^{-1}) = O((mn)\varepsilon^{-2}\log\varepsilon^{-1})$

Theorem

In $\mathrm{poly}(k, m, n, \varepsilon^{-1})$ *time, can construct a* $(1+\varepsilon)$ *approximate diagram* $\widetilde{\mathcal{M}}$ *of complexity* $O((mn/k)\varepsilon^{-2}\log\varepsilon^{-1})$.

Compressing $\widetilde{\mathcal{M}}$

- Reduce size from $O(mn) \to O(mn/k)$.
- Any $O(1)$ approx. diagram is enough for the $(1+\varepsilon)$ diagram.
- *Cluster the points of T*... while keeping cluster radius small w.r.t. $\text{cost}^*(t)$.

Averaging argument

$$\operatorname{cost}^*(t) := \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b) \in M} \|a + t - b\|^p \right]^{1/p}$$

- At most $k/2$ pairs $(a, b) \in M_t$ have $\|a + t - b\| \geq 2^{1/p} \operatorname{cost}^*(t)$.
- At most $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \geq 2^{1/p} \operatorname{cost}^*(t)$.
- At least $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \leq 2^{1/p} \operatorname{cost}^*(t)$.

Averaging argument

$$\mathrm{cost}^*(t) := \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b)\in M} \|a + t - b\|^p \right]^{1/p}$$

- At most $k/2$ pairs $(a, b) \in M_t$ have $\|a + t - b\| \geq 2^{1/p}\,\mathrm{cost}^*(t)$.
- At most $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \geq 2^{1/p}\,\mathrm{cost}^*(t)$.
- At least $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \leq 2^{1/p}\,\mathrm{cost}^*(t)$.

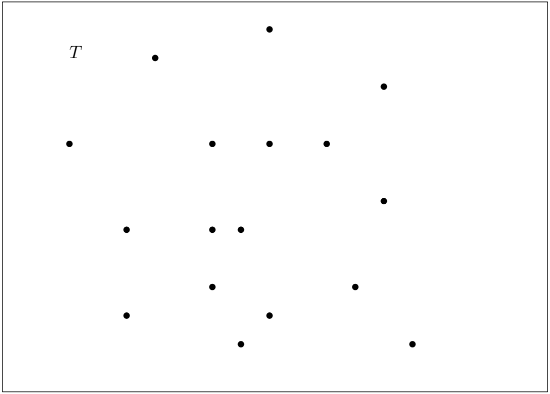
Averaging argument

$$\text{cost}^*(t) := \min_{k\text{-matching } M} \left[\frac{1}{k} \sum_{(a,b) \in M} \|a + t - b\|^p \right]^{1/p}$$

- At most $k/2$ pairs $(a, b) \in M_t$ have $\|a + t - b\| \geq 2^{1/p} \text{cost}^*(t)$.
- At most $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \geq 2^{1/p} \text{cost}^*(t)$.
- At least $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \leq 2^{1/p} \text{cost}^*(t)$.

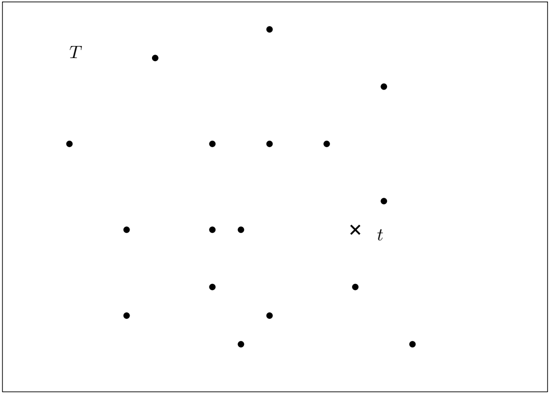
Averaging argument (cont.)

- At least $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \leq 2^{1/p} \operatorname{cost}^*(t)$.



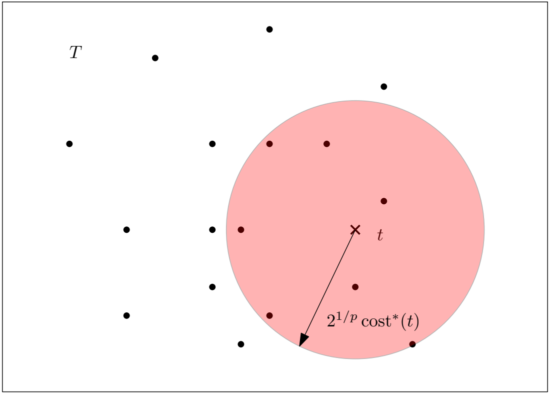
Averaging argument (cont.)

- At least $k/2$ pairs $(a, b) \in M_t$ have $\|t - t_{ab}\| \leq 2^{1/p} \operatorname{cost}^*(t)$.



Averaging argument (cont.)

- At least $k/2$ pairs $(a,b) \in M_t$ have $\|t - t_{ab}\| \leq 2^{1/p}\,\mathrm{cost}^*(t)$.



Clustering T greedily

1. Let $D_i = B(x_i, r_i)$ be the smallest disk containing at least $k/2$ remaining points of T.
2. Remove from T the points covered by D_i.
3. Repeat.

- *Quorum clustering*
  2-approx. algorithm in $O(|T| \operatorname{polylog} n)$ time [Carmi *et al.* 05]
- $O(|T|/k) = O(mn/k)$ clusters.

Clustering T greedily

1. Let $D_i = B(x_i, r_i)$ be the smallest disk containing at least $k/2$ remaining points of T.
2. Remove from T the points covered by D_i.
3. Repeat.

- *Quorum clustering*
  2-approx. algorithm in $O(|T| \operatorname{polylog} n)$ time [Carmi *et al.* 05]
- $O(|T|/k) = O(mn/k)$ clusters.

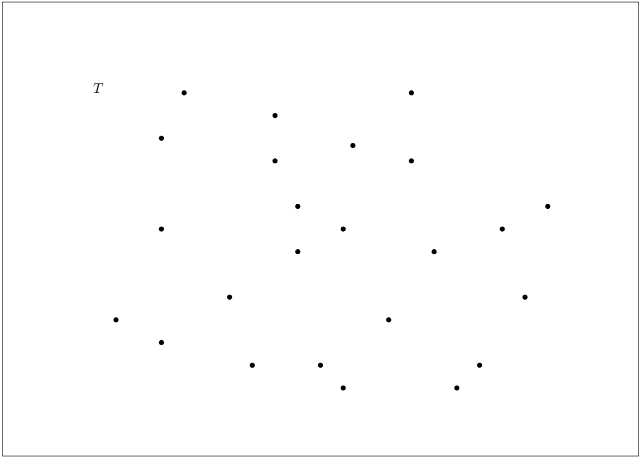
Clustering T greedily

1. Let $D_i = B(x_i, r_i)$ be the smallest disk containing at least $k/2$ remaining points of T.
2. Remove from T the points covered by D_i.
3. Repeat.

- *Quorum clustering*
  2-approx. algorithm in $O(|T| \operatorname{polylog} n)$ time [Carmi *et al.* 05]
- $O(|T|/k) = O(mn/k)$ clusters.

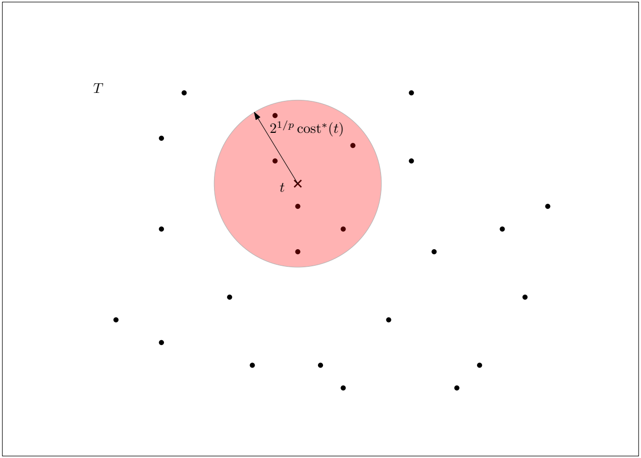
Greedy clustering proof of approximation

- Let X be the set of centers x_i from the greedy clustering, and $\widetilde{\mathcal{M}} = \mathrm{Vor}(X)$ with each $\mathrm{VorRegion}(x_i)$ assigned M_{x_i}.



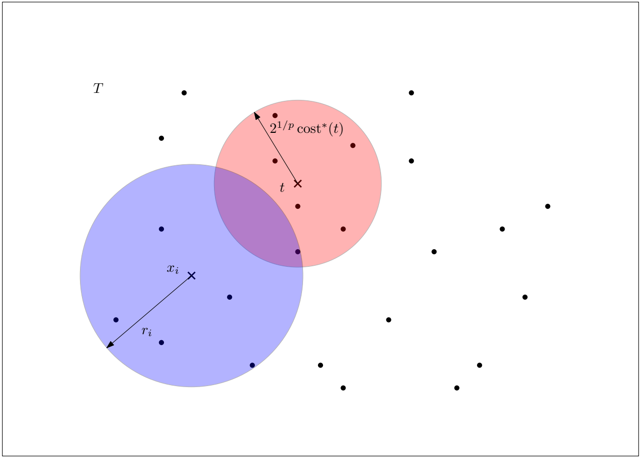
Greedy clustering proof of approximation

- Let X be the set of centers x_i from the greedy clustering, and $\widetilde{\mathcal{M}} = \mathrm{Vor}(X)$ with each $\mathrm{VorRegion}(x_i)$ assigned M_{x_i}.



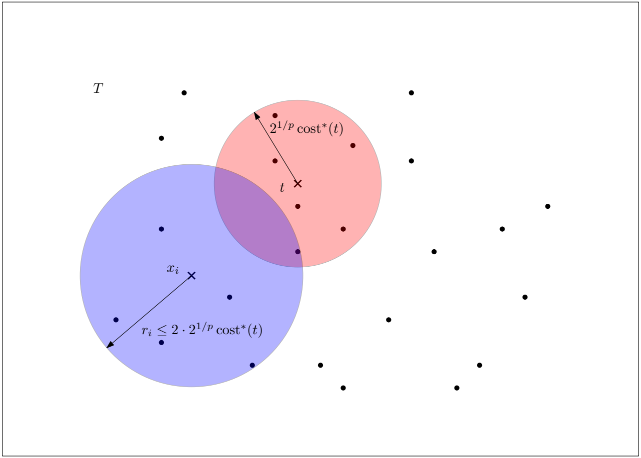
Greedy clustering proof of approximation

- Let X be the set of centers x_i from the greedy clustering, and $\widetilde{\mathcal{M}} = \mathrm{Vor}(X)$ with each $\mathrm{VorRegion}(x_i)$ assigned M_{x_i}.



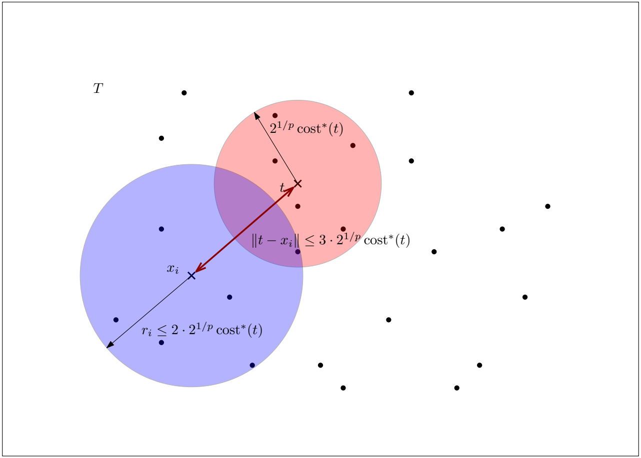
Greedy clustering proof of approximation

- Let X be the set of centers x_i from the greedy clustering, and $\widetilde{\mathcal{M}} = \mathrm{Vor}(X)$ with each $\mathrm{VorRegion}(x_i)$ assigned M_{x_i}.

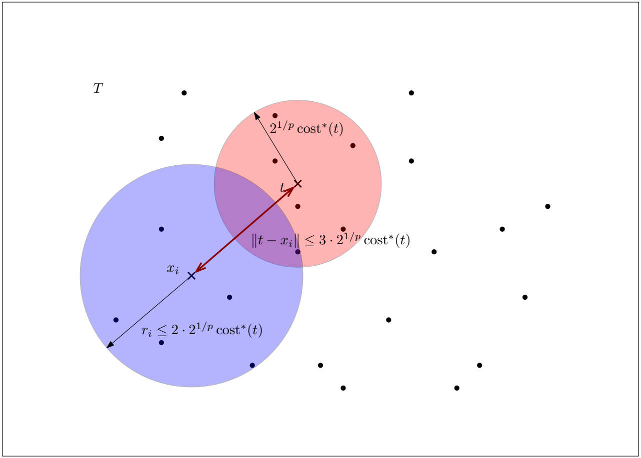


Greedy clustering proof of approximation

- Let X be the set of centers x_i from the greedy clustering, and $\widetilde{\mathcal{M}} = \mathrm{Vor}(X)$ with each $\mathrm{VorRegion}(x_i)$ assigned M_{x_i}.



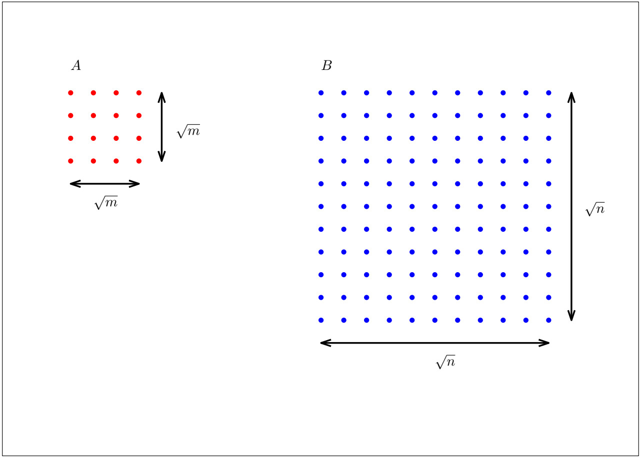
Greedy clustering proof of approximation



- $\mathrm{cost}(M_{x_i}, t) \leq (1 + 6 \cdot 2^{1/p})\,\mathrm{cost}^*(t)$
- Using grids, $(1 + \varepsilon)$ approx. diagram of size $O((mn/k)\varepsilon^{-2} \log \varepsilon^{-1})$.

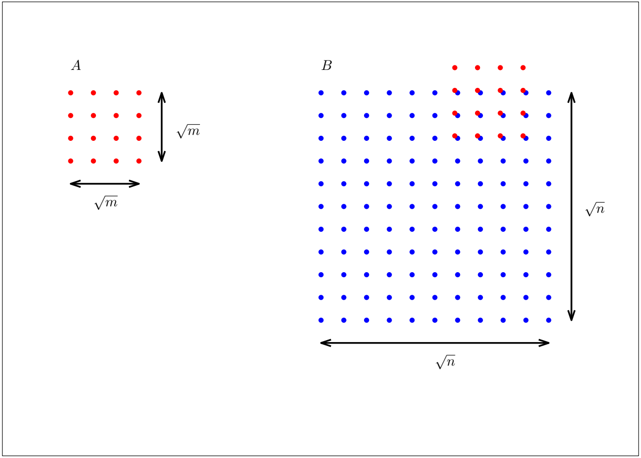
Lower bound: $O(mn/k)$ tight for large k

- For $k = c \cdot \min(m, n)$, constant c:



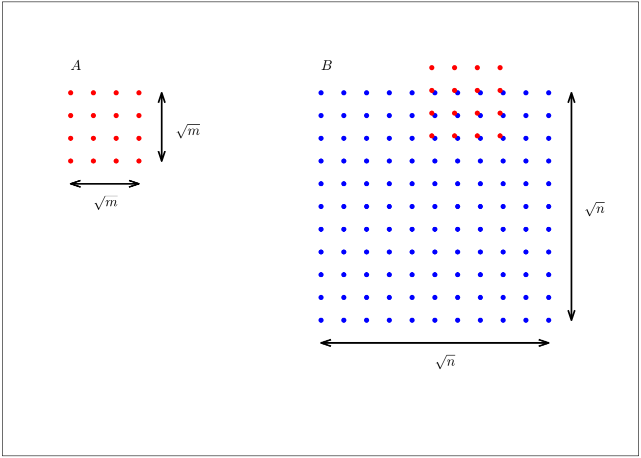
Lower bound: $O(mn/k)$ tight for large k

- For $k = c \cdot \min(m, n)$, constant c:



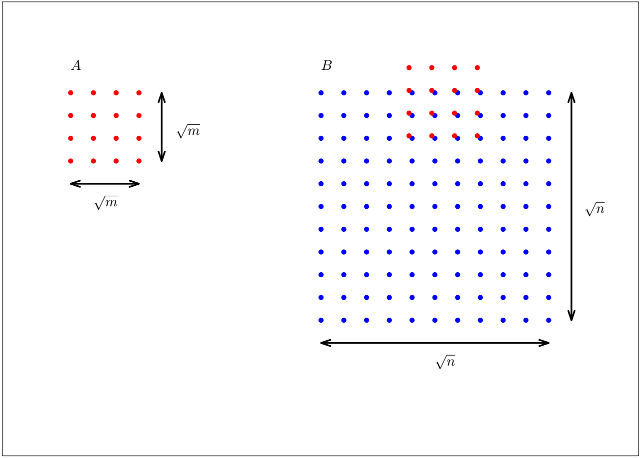
Lower bound: $O(mn/k)$ tight for large k

- For $k = c \cdot \min(m, n)$, constant c:



Lower bound: $O(mn/k)$ tight for large k

- For $k = c \cdot \min(m, n)$, constant c:



Open questions

1. Is the complexity of $\mathcal{M}$ polynomial, for any p?
2. Approximate diagrams for rotations? Rigid transforms?

The End

Thank you.