Nonlinear Dimension Reduction via Outer Bi-Lipschitz Extensions

Sepideh Mahabadi
TTIC

Konstantin Makarychev
Northwestern University

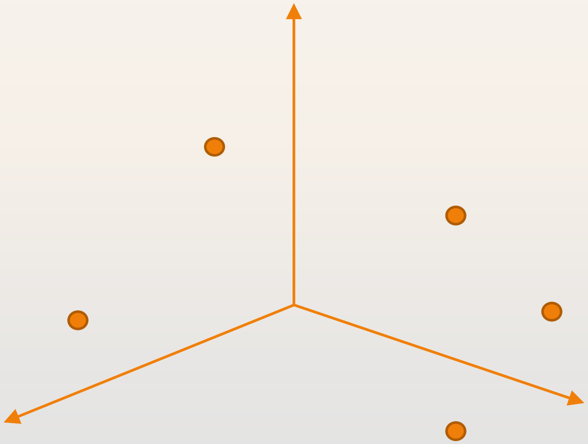
Yury Makarychev
TTIC

Ilya Razenshteyn
MSR Redmond

Dimension Reduction

Given: a set of n points in a d dimensional space

- Can represent feature vectors of a set of objects such as images, documents, etc.



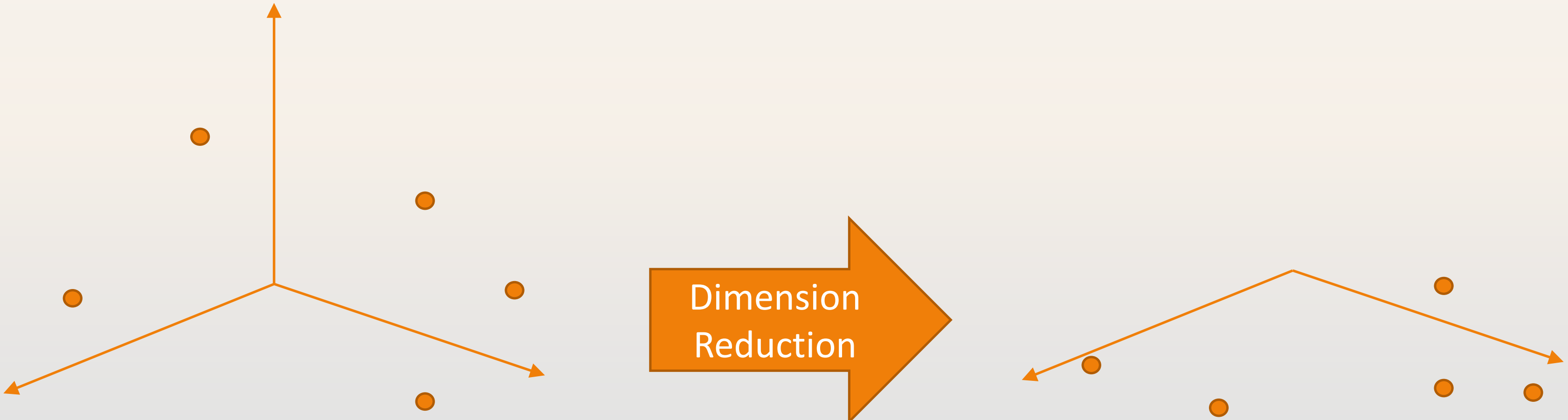
Dimension Reduction

Given: a set of n points in a d dimensional space

- Can represent feature vectors of a set of objects such as images, documents, etc.

Dimension Reduction: reduce dimension of the points, i.e., embed them into a lower dimensional space while preserving pairwise distances

- ➢ Less storage
- ➢ Less communication to transmit the data
- ➢ Less computation



Johnson-Lindenstrauss Lemma

Johnson-Lindenstrauss Lemma: every set $X \subset \mathbb{R}^d$ of size n can be embedded into $\mathbb{R}^{d'}$ where $d' = O(\frac{\log n}{\epsilon^2})$ such that the distances are preserved up to a factor of $(1+\epsilon)$.

Johnson-Lindenstrauss Lemma

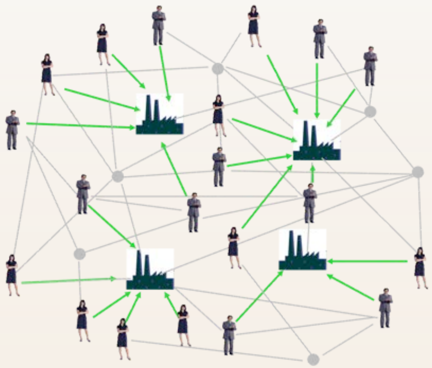
Johnson-Lindenstrauss Lemma: every set $X \subset \mathbb{R}^d$ of size n can be embedded into $\mathbb{R}^{d'}$ where $d' = O(\frac{\log n}{\epsilon^2})$ such that the distances are preserved up to a factor of $(1 + \epsilon)$.

- Applied in a diverse range of areas such as streaming algorithms, nearest neighbor search, graph sparsification, compressed sensing, …

- Known to be tight [Larsen,Nelson’17]

Prioritized Dimension Reduction

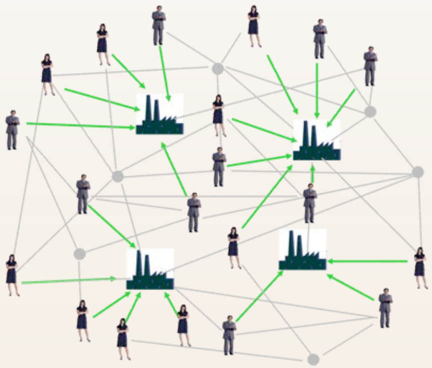
- The points might not be equally important:
  - E.g. They might represent feature vectors of objects (e.g. images, documents) and some are used more often.
  - Or facilities/users where facilities are accessed more frequently
  - High influential vs low influential users in a social media



Prioritized Dimension Reduction

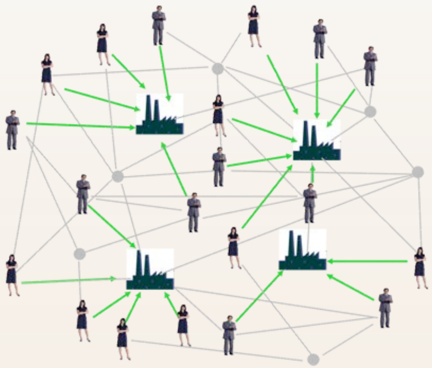
- The points might not be equally important:
  - E.g. They might represent feature vectors of objects (e.g. images, documents) and some are used more often.
  - Or facilities/users where facilities are accessed more frequently
  - High influential vs low influential users in a social media

- For them we want to use even less coordinates.



Prioritized Dimension Reduction

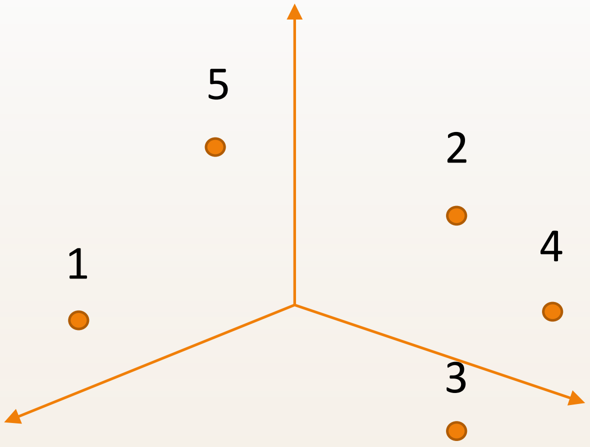
- The points might not be equally important:
  - E.g. They might represent feature vectors of objects (e.g. images, documents) and some are used more often.
  - Or facilities/users where facilities are accessed more frequently
  - High influential vs low influential users in a social media



- For them we want to use even less coordinates.

- **Prioritized/Terminal Dimension Reduction** introduced by **Elkin, Filtser and Neiman'15**.
  - The main motivation of this work.

Prioritized Dimension Reduction



Prioritized Dimension Reduction

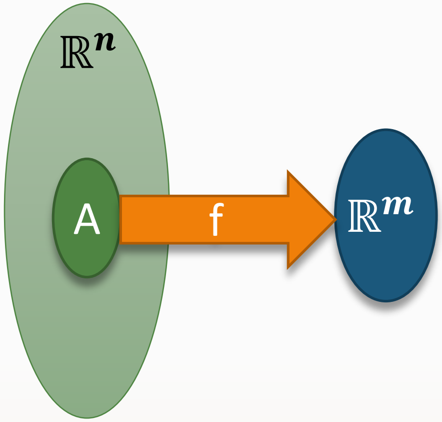
Prioritized Dimension Reduction

Prioritized Dimension Reduction

Plan

1. Definitions and background
2. Introduce Outer Bi-Lipschitz Extension
    - ➢ **New Notion**
3. Present our extension results
    - ➢ **Main Technique**
4. Present its applications to dimension reduction

Extension of Functions

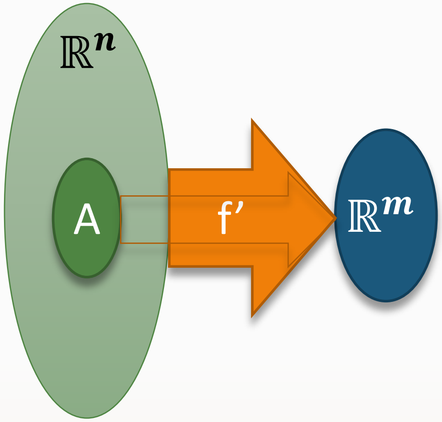


Notation throughout the talk

- We have a function $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$
- Which is defined over a subset $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$

Extensions of the map f to a superset of A.

Extension of Functions



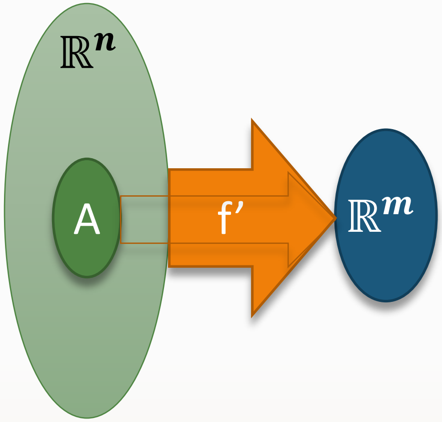
Notation throughout the talk

- We have a function $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$
- Which is defined over a subset $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$

Extensions of the map f to a superset of A.

☐ **Extension to the whole** $\mathbb{R}^n$, i.e., $\boldsymbol{f}': \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ so that
  - $\boldsymbol{f}'(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$ for any $x \in A$
  - Maintaining other properties …

Extension of Functions



Notation throughout the talk

- We have a function $f: A \to \mathbb{R}^m$
- Which is defined over a subset $A \subset \mathbb{R}^n$

Extensions of the map f to a superset of A.

☐ **Extension to the whole** $\mathbb{R}^n$, i.e., $f': \mathbb{R}^n \to \mathbb{R}^m$ so that

  - $f'(x) = f(x)$ for any $x \in A$
  - Maintaining other properties …

1. **Lipschitz Constant**
   (Lipschitz Extension)
2. **Bi-Lipschitz Constant, i.e., distortion**
   (Bi-Lipschitz Extension)

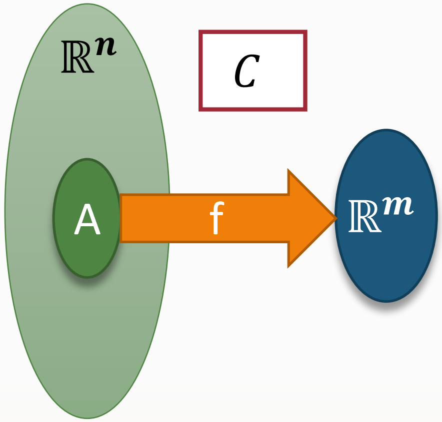
Lipschitz Extension

- A map $f : X \to Y$ is **C-Lipschitz** if for all $x, x' \in X$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq \boldsymbol{C} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

Euclidean

Lipschitz Extension



- ☐ A map $f: X \to Y$ is **C-Lipschitz** if for all $x, x' \in X$:

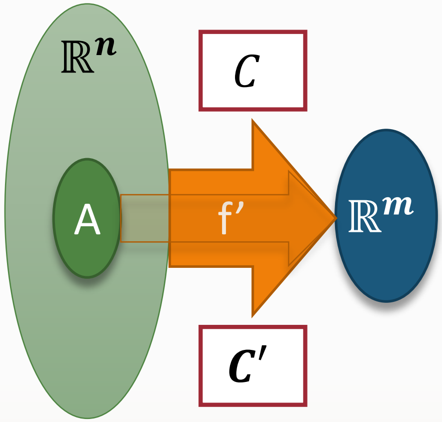
$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq \boldsymbol{C} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

Euclidean

- ☐ **Lipschitz extension:**

Given: a C-Lipschitz map $f: A \to \mathbb{R}^{\boldsymbol{m}}$, where $A \subseteq \mathbb{R}^{\boldsymbol{n}}$

Lipschitz Extension



- ☐ A map $f\colon X \to Y$ is **C-Lipschitz** if for all $x, x' \in X$:

  $$\|f(x) - f(x')\| \le C \cdot \|x - x'\| \rightarrow \boxed{\text{Euclidean}}$$

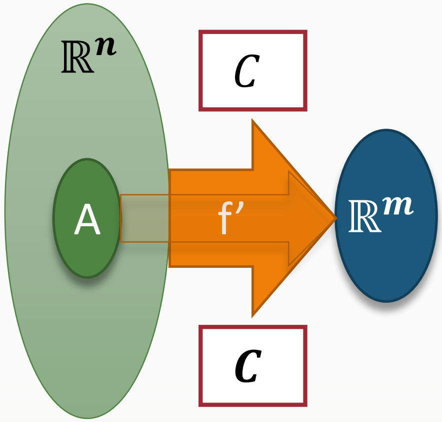
- ☐ **Lipschitz extension:**

  Given: a C-Lipschitz map $f\colon A \to \mathbb{R}^m$, where $A \subseteq \mathbb{R}^n$

  Goal: a map $f'\colon \mathbb{R}^n \to \mathbb{R}^m$ s.t.

  - f' is an extension of f
  - f' is C'-Lipschitz

Lipschitz Extension



- ☐ A map $f: X \to Y$ is **C-Lipschitz** if for all $x, x' \in X$:

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \le \boldsymbol{C} \cdot \|\boldsymbol{x} - \boldsymbol{x}'\| \longrightarrow \boxed{\text{Euclidean}}$$

- ☐ **Lipschitz extension:**

  Given: a C-Lipschitz map $f: A \to \mathbb{R}^{\boldsymbol{m}}$, where $A \subseteq \mathbb{R}^{\boldsymbol{n}}$

  Goal: a map $f': \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ s.t.

  - f' is an extension of f
  - f' is C'-Lipschitz

Kirszbraun extension theorem '34: for $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$, every C-Lipschitz map $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$ can be extended to the whole $\mathbb{R}^n$ keeping the **same** Lipschitz constant, i.e., $\boldsymbol{C}' = \boldsymbol{C}$.

Bi-Lipschitz Counterpart of the Kirszbraun theorem?

Kirszbraun extension theorem ‘34: For $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$, every C-Lipschitz map $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$ can be extended to the whole $\mathbb{R}^n$ keeping the same Lipschitz constant.

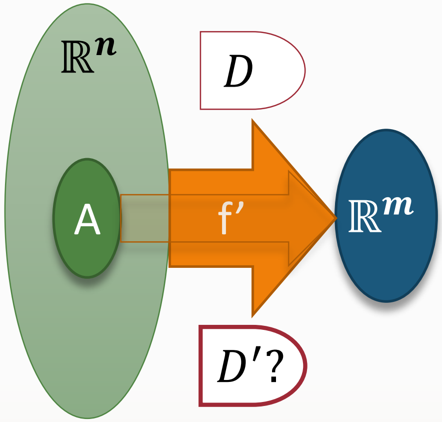
Bi-Lipschitz Extension

- A map $f: X \rightarrow Y$ is **D-bi-Lipschitz** or has **distortion D**

  if for some λ and all $x, x' \in X$:

$$\lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq D \cdot \lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

Bi-Lipschitz Extension



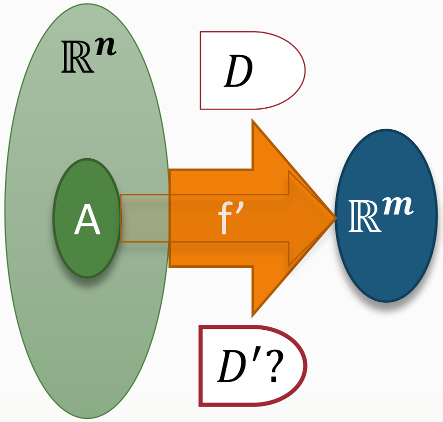
- ☐ A map $f: X \to Y$ is ***D*-bi-Lipschitz** or has **distortion *D***

  if for some λ and all $x, x' \in X$:

  $$\lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}')\| \leq D \cdot \lambda \cdot \|\boldsymbol{x} - \boldsymbol{x}'\|$$

- ☐ **Is there a counterpart of the Kirszbraun theorem for bi-Lipschitz maps?**
  - For $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$, can every map $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$ of distortion D be extended to the whole $\mathbb{R}^n$ keeping the same distortion?

Bi-Lipschitz Extension

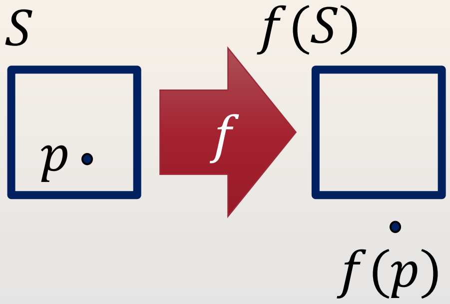


- ❑ A map $f : X \to Y$ is ***D*-bi-Lipschitz** or has **distortion *D***

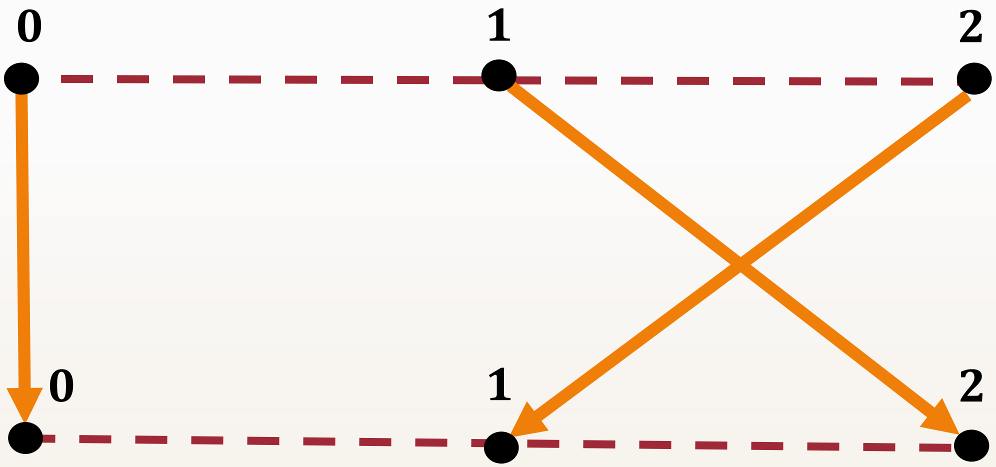
  if for some λ and all $x, x' \in X$:

$$\lambda \cdot \|\boldsymbol{x}-\boldsymbol{x}'\| \leq \|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x}')\| \leq D \cdot \lambda \cdot \|\boldsymbol{x}-\boldsymbol{x}'\|$$

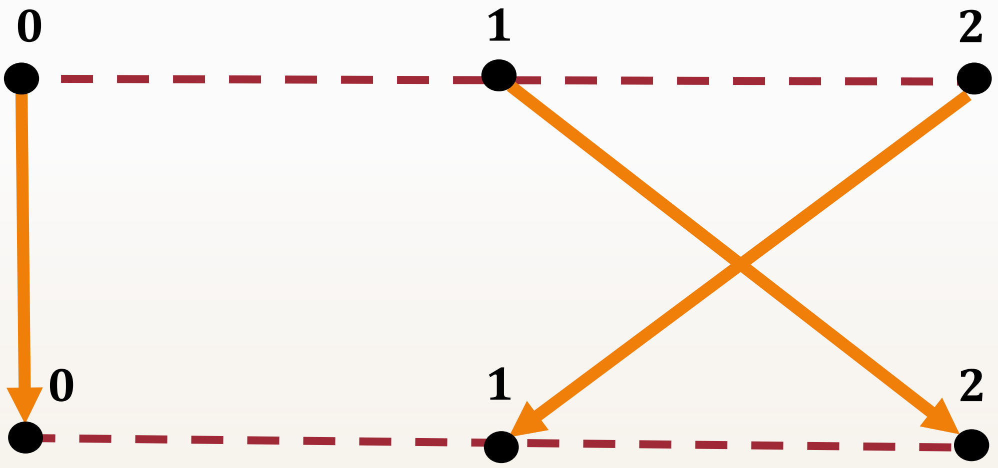
- ❑ **Is there a counterpart of the Kirszbraun theorem for bi-Lipschitz maps?**
  - For $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$, can every map $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$ of distortion D be extended to the whole $\mathbb{R}^{n}$ keeping the same distortion?
    - ❖ **No direct analogue!**
      - ▪ even if we allow $D' > D$
      - ▪ Not even a 1-1 continuous map



Example

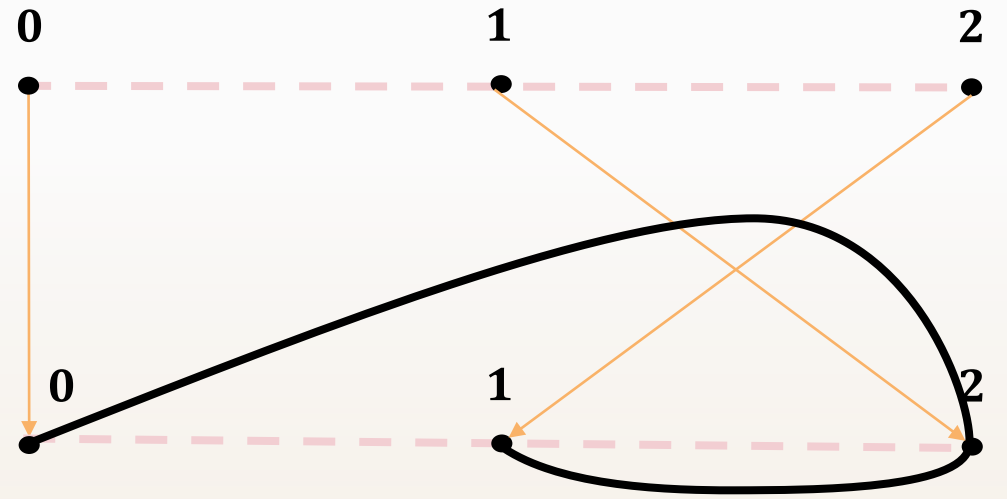


Example



- No 1-1 continuous extension map from $\mathbb{R}$ to $\mathbb{R}$

Example

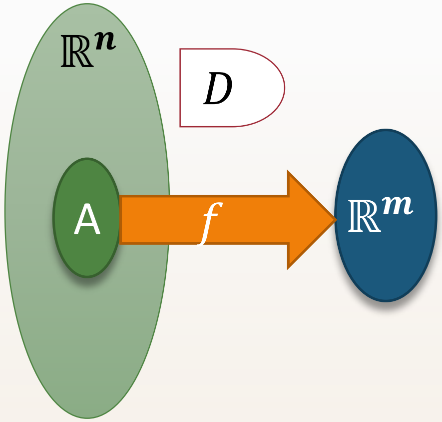


- No 1-1 continuous extension map from $\mathbb{R}$ to $\mathbb{R}$
- There **exists** such a map from $\mathbb{R}$ to $\mathbb{R}^2$
  - **Fix:** Allow additional coordinates

Bi-Lipschitz Outer-Extension

Given: a map $\boldsymbol{f}: \boldsymbol{A} \rightarrow \mathbb{R}^{\boldsymbol{m}}$, where

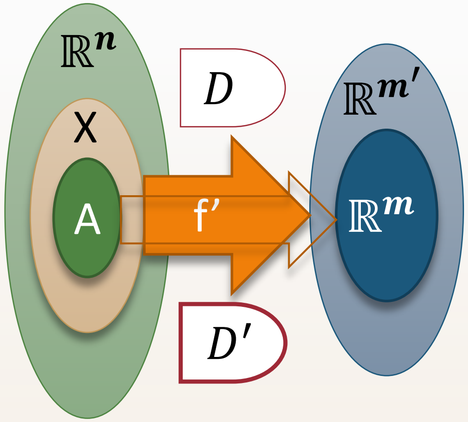
- $A \subseteq X \subset \mathbb{R}^{\boldsymbol{n}}$
- f has distortion D



Bi-Lipschitz Outer-Extension

Given: a map $\boldsymbol{f}: \boldsymbol{A} \rightarrow \mathbb{R}^{\boldsymbol{m}}$, where

- $A \subseteq X \subset \mathbb{R}^{\boldsymbol{n}}$
- f has distortion D



Goal: a map $\boldsymbol{f}': \boldsymbol{X} \rightarrow \mathbb{R}^{\boldsymbol{m}'}$, where

- $m' > m$
- f' has distortion D'
- f' is an **(outer)-extension** of f: for every $x \in A$

$$f'(x) = f(x) \oplus \underbrace{(0, \dots, 0)}_{\boldsymbol{m}' - \boldsymbol{m}}$$

Results

Consider a D-bi-Lipschitz map $f: A \to \mathbb{R}^m$ where $A \subset \mathbb{R}^n$,

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	D	To $\mathbb{R}^n$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta(\sqrt{\epsilon})$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{\log^2 1/\epsilon})$	$\mathbb{R}^2$

❑ Show applications to dimension reduction

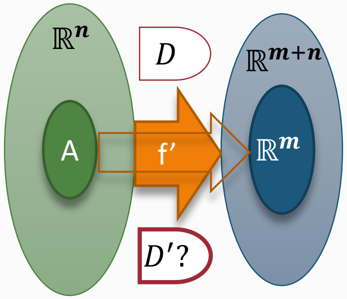
Results

Consider a D-bi-Lipschitz map $f: A \to \mathbb{R}^m$ where $A \subset \mathbb{R}^n$,

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	D	To $\mathbb{R}^n$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta\left(\sqrt{\epsilon}\right)$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{\log^2 1/\epsilon})$	$\mathbb{R}^2$

☐ Show applications to dimension reduction

Counterpart of Kirszbraun



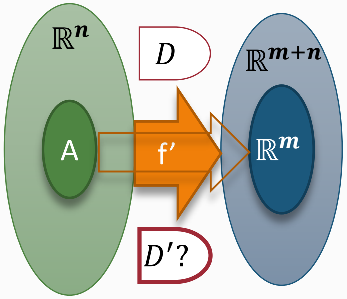
Given a map

- $\boldsymbol{f}(\boldsymbol{x})\colon A \to \mathbb{R}^m$ $(\text{where } A \subset \mathbb{R}^n)$ with distortion D

Come up with a bi-Lipschitz outer-extension of the map

- $\boldsymbol{f}'(\boldsymbol{x})\colon \mathbb{R}^n \to \mathbb{R}^{m'}$

Counterpart of Kirszbraun



Given a map

- $f(x): A \to \mathbb{R}^m$ (where $A \subset \mathbb{R}^n$) with distortion D

Come up with a bi-Lipschitz outer-extension of the map

- $f'(x): \mathbb{R}^n \to \mathbb{R}^{m'}$

New bi-Lipschitz outer extension:

Two applications of the Kirszbraun Lipschitz extension Theorem.

Proof Idea

Let

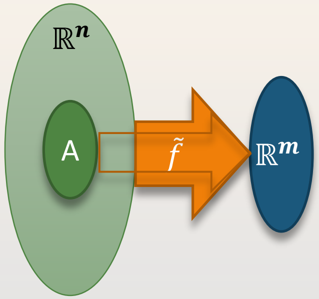
- $f(x): A \to \mathbb{R}^m$ be our map
- $\tilde{f}(x): \mathbb{R}^n \to \mathbb{R}^m$ be its Lipschitz extension

By Kirszbraun

New bi-Lipschitz outer extension:

$$\boldsymbol{f'(x) = \tilde{f}(x)}$$

➢ **The distances might decrease a lot!**



Proof Idea

Let

- $f(x): A \to \mathbb{R}^m$ be our map
- $\tilde{f}(x): \mathbb{R}^n \to \mathbb{R}^m$ be its Lipschitz extension → By Kirszbraun

New bi-Lipschitz outer extension:

$$\boldsymbol{f'(x) = \tilde{f}(x) \oplus h(x)}$$

➢ **The distances might decrease a lot!**

✓ **Add the second component $\boldsymbol{h(x)}$**

- $\boldsymbol{h(x) = 0}$ for all $\boldsymbol{x \in A}$
- When $\boldsymbol{\tilde{f}}$ contracts $\boldsymbol{h}$ should expand

Proof Idea

Let

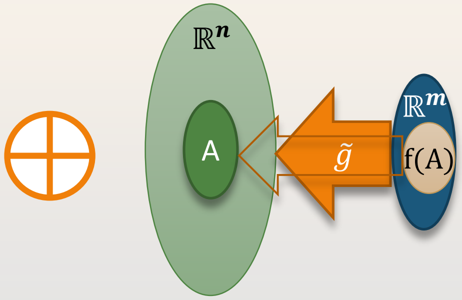
- $f(x): A \to \mathbb{R}^m$ be our map
- $\tilde{f}(x): \mathbb{R}^n \to \mathbb{R}^m$ be its Lipschitz extension → By Kirszbraun
- $g = f^{-1}: f(A) \to \mathbb{R}^n$ be its inverse
- $\tilde{g}(x): \mathbb{R}^m \to \mathbb{R}^n$ be its Lipschitz extension → By Kirszbraun

New bi-Lipschitz outer extension:

$$\boldsymbol{f'(x) = \tilde{f}(x) \oplus h(x)} \quad \textbf{where} \quad \boldsymbol{h(x) = \tilde{g}\left(\tilde{f}(x)\right)}$$

✓ **Add the second component $\boldsymbol{h(x)}$**

  - $\boldsymbol{h(x) = 0}$ for all $\boldsymbol{x \in A}$
  - When $\boldsymbol{\tilde{f}}$ contracts $\boldsymbol{h}$ should expand



Proof Idea

Let

- $f(x): A \to \mathbb{R}^m$ be our map
- $\tilde{f}(x): \mathbb{R}^n \to \mathbb{R}^m$ be its Lipschitz extension → By Kirszbraun
- $g = f^{-1}: f(A) \to \mathbb{R}^n$ be its inverse
- $\tilde{g}(x): \mathbb{R}^m \to \mathbb{R}^n$ be its Lipschitz extension → By Kirszbraun

New bi-Lipschitz outer extension:

$$\boxed{\boldsymbol{f'(x) = \tilde{f}(x) \oplus h(x)} \qquad \textbf{where} \qquad \boldsymbol{h(x) = \tilde{g}\left(\tilde{f}(x)\right) - x}}$$

✓ **Add the second component $\boldsymbol{h(x)}$**

- $\boldsymbol{h(x) = 0}$ for all $\boldsymbol{x \in A}$ → $\boxed{\tilde{g}\left(\tilde{f}(x)\right) - x = g(f(x)) - x = 0}$
- When $\boldsymbol{\tilde{f}}$ contracts $\boldsymbol{h}$ should expand

Proof Idea

Let

- $\boldsymbol{f(x)}: A \to \mathbb{R}^m$ be our map
- $\boldsymbol{\tilde{f}(x)}: \mathbb{R}^n \to \mathbb{R}^m$ be its Lipschitz extension → By Kirszbraun
- $\boldsymbol{g = f^{-1}}: f(A) \to \mathbb{R}^n$ be its inverse
- $\boldsymbol{\tilde{g}(x)}: \mathbb{R}^m \to \mathbb{R}^n$ be its Lipschitz extension → By Kirszbraun

New bi-Lipschitz outer extension:

$$\boldsymbol{f'(x) = \tilde{f}(x) \oplus h(x)} \qquad \textbf{where} \qquad \boldsymbol{h(x) = \frac{\tilde{g}\left(\tilde{f}(x)\right) - x}{\sqrt{2}D}}$$

✓ $\boldsymbol{f'}$ is from $\mathbb{R}^n$ to $\mathbb{R}^{n+m}$

✓ For $x \in A$, $\boldsymbol{f}(x) = \boldsymbol{f'}(x)$

✓ Distortion is at most $3D$

Results

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	$\mathbf{D}$	To $\mathbb{R}^{n}$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta\left(\sqrt{\epsilon}\right)$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{log^{2}1/\epsilon})$	$\mathbb{R}^{2}$

☐ Show applications to dimension reduction

Results

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	$\mathbf{D}$	To $\mathbb{R}^{n}$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta\left(\sqrt{\epsilon}\right)$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta\left(\frac{1}{log^2 1/\epsilon}\right)$	$\mathbb{R}^2$

☐ Show applications to dimension reduction

Prioritized Dimension Reduction

Input:

- a set of n points P in $\mathbb{R}^d$
- a **ranking π** on them: a bijection from P to $[n]$

Goal: reduce the dimension s.t.

$$f(x) \in \mathbb{R}^{g(r)} \subset \mathbb{R}^{c \log n}$$

- where $r = \pi(x)$ is the rank of x and g is polylogarithmic

Prioritized Dimension Reduction

Input:

- a set of n points P in $\mathbb{R}^d$
- a **ranking π** on them: a bijection from P to $[n]$

Goal: reduce the dimension s.t.

$$f(x) \in \mathbb{R}^{g(r)} \subset \mathbb{R}^{c \log n}$$

- where $r = \pi(x)$ is the rank of x and g is polylogarithmic

	Distortion	#Non-zero
[Elkin, Filtser, Neiman, STOC'15]	$O_\epsilon(\log^{4+\epsilon} r)$	$O_\epsilon(\log^4 r)$
This work	$O(\log \log r)$	$O_\epsilon(\log^{3+\epsilon} r)$

Prioritized Dimension Reduction

Input:

- a set of n points P in $\mathbb{R}^d$
- a **ranking π** on them: a bijection from P to $[n]$

Goal: reduce the dimension s.t.

$$f(x) \in \mathbb{R}^{g(r)} \subset \mathbb{R}^{c \log n}$$

- where $r = \pi(x)$ is the rank of x and g is polylogarithmic

	Distortion	#Non-zero
[Elkin, Filtser, Neiman, STOC'15]	$\boldsymbol{O_\epsilon(\log^{4+\epsilon} r)}$	$\boldsymbol{O_\epsilon(\log^{4} r)}$
This work	$\boldsymbol{O(\log\log r)}$	$\boldsymbol{O_\epsilon(\log^{3+\epsilon} r)}$

For two points at rank r and t,

- The time to compute their distance only depends on $\mathrm{polylog}(\max\{r, t\})$
- The distortion of their distance depends on $\log\log(\max\{r, t\})$

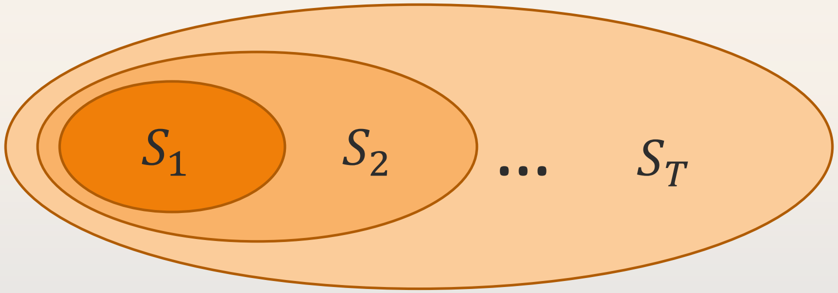
Proof Idea

❑ **Grouping based on priorities**

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

❑ **Iterative Extension:**

- Given $\boldsymbol{f_{i-1}}: \boldsymbol{S_{i-1}} \to \mathbb{R}^{\boldsymbol{d_{i-1}}}$
- Inductively construct $\boldsymbol{f_i}: \boldsymbol{S_i} \to \mathbb{R}^{\boldsymbol{d_i}}$



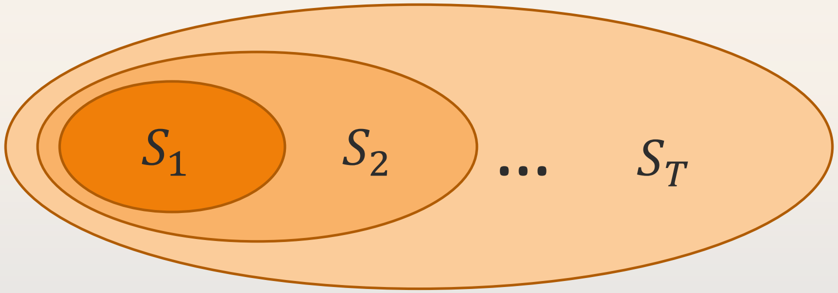
Proof Idea

☐ **Grouping based on priorities**

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

☐ **Iterative Extension:**

- Given $\boldsymbol{f_{i-1}}: \boldsymbol{S_{i-1}} \to \mathbb{R}^{\boldsymbol{d_{i-1}}}$
- Inductively construct $\boldsymbol{f_i}: \boldsymbol{S_i} \to \mathbb{R}^{\boldsymbol{d_i}}$



Construction of $\boldsymbol{f_i}$ from $\boldsymbol{f_{i-1}}$

1. The new bi-Lipschitz outer-extension on f_{i-1}:

$$\boldsymbol{f'_{i-1}(x) = \widetilde{f_{i-1}}(x) \; \oplus \; h(x)}$$

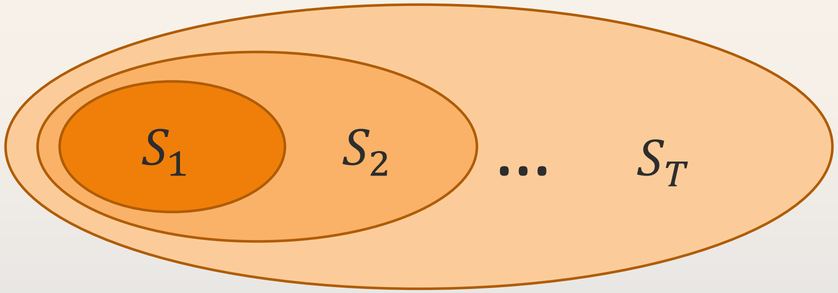
Proof Idea

☐ **Grouping based on priorities**

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

☐ **Iterative Extension:**

- Given $\boldsymbol{f_{i-1}}: \boldsymbol{S_{i-1}} \to \mathbb{R}^{d_{i-1}}$
- Inductively construct $\boldsymbol{f_i}: \boldsymbol{S_i} \to \mathbb{R}^{d_i}$



Construction of $\boldsymbol{f_i}$ from $\boldsymbol{f_{i-1}}$

1. The new bi-Lipschitz outer-extension on f_{i-1}:

$$\underset{\text{Bi-Lip ext.}}{\underset{\uparrow}{\boldsymbol{f'_{i-1}}}}\boldsymbol{(x)} = \underset{\text{Lip ext.}}{\underset{\uparrow}{\widetilde{\boldsymbol{f_{i-1}}}}}\boldsymbol{(x)} \;\oplus\; \boldsymbol{h(x)}$$

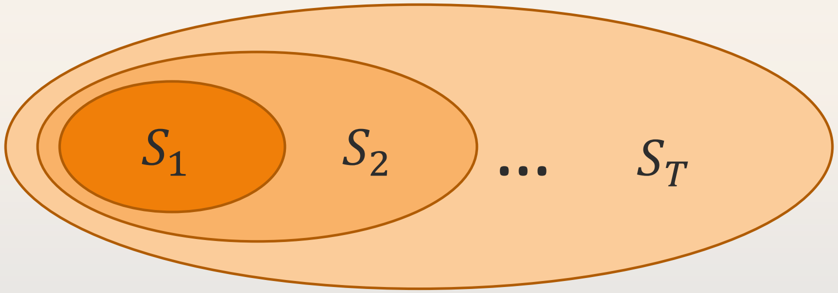
Proof Idea

☐ Grouping based on priorities

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

☐ Iterative Extension:

- Given $\boldsymbol{f_{i-1}}: \boldsymbol{S_{i-1}} \to \mathbb{R}^{\boldsymbol{d_{i-1}}}$
- Inductively construct $\boldsymbol{f_i}: \boldsymbol{S_i} \to \mathbb{R}^{\boldsymbol{d_i}}$



Construction of $\boldsymbol{f_i}$ from $\boldsymbol{f_{i-1}}$

1. The new bi-Lipschitz outer-extension on f_{i-1}:

$$\boldsymbol{f'_{i-1}(x) = \widetilde{f_{i-1}}(x) \ \oplus \ h(x)}$$

We do not change the map on $\boldsymbol{S_{i-1}}$
No new coordinates

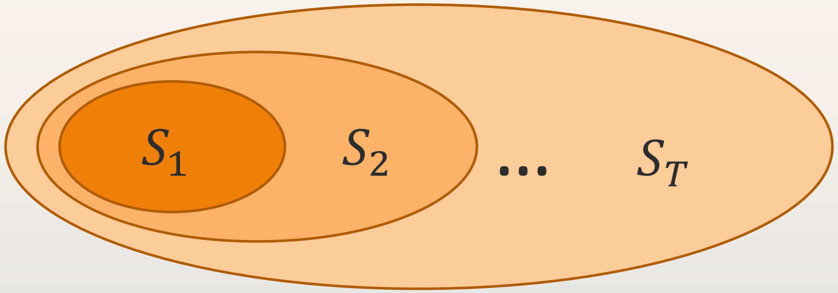
Proof Idea

☐ **Grouping based on priorities**

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

☐ **Iterative Extension:**

- Given $\boldsymbol{f_{i-1}}: \boldsymbol{S_{i-1}} \rightarrow \mathbb{R}^{\boldsymbol{d_{i-1}}}$
- Inductively construct $\boldsymbol{f_i}: \boldsymbol{S_i} \rightarrow \mathbb{R}^{\boldsymbol{d_i}}$



Construction of $\boldsymbol{f_i}$ from $\boldsymbol{f_{i-1}}$

1. The new bi-Lipschitz outer-extension on f_{i-1}:

$$\boldsymbol{f'_{i-1}(x) = \widetilde{f_{i-1}}(x) \;\oplus\; h(x)}$$

- **We do not change the map on $\boldsymbol{S_{i-1}}$**
- **No new coordinates**

- $\boldsymbol{f_{i-1}}: \mathbb{R}^{\boldsymbol{d}} \rightarrow \mathbb{R}^{\boldsymbol{d_{i-1}}}$
- $\boldsymbol{d_i = d_{i-1} + d \qquad > d}$ ☹

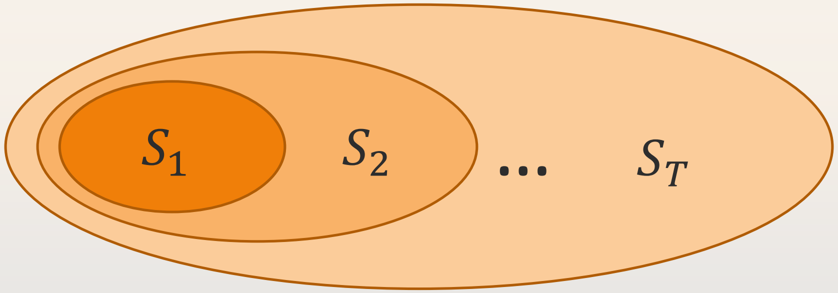
Proof Idea

☐ **Grouping based on priorities**

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

☐ **Iterative Extension:**

- Given $\boldsymbol{f}_{i-1}: \boldsymbol{S}_{i-1} \rightarrow \mathbb{R}^{d_{i-1}}$
- Inductively construct $\boldsymbol{f}_i: \boldsymbol{S}_i \rightarrow \mathbb{R}^{d_i}$



Construction of $\boldsymbol{f}_i$ from $\boldsymbol{f}_{i-1}$

1. The new bi-Lipschitz outer-extension on f_{i-1}:

$$\boldsymbol{f'}_{i-1}(\boldsymbol{x}) = \widetilde{\boldsymbol{f}_{i-1}}(\boldsymbol{x}) \oplus \boldsymbol{h}(\boldsymbol{x})$$

- **Use JL**
  - **Reduce dimension**
  - **No longer an extension of $\boldsymbol{f}_{i-1}$**

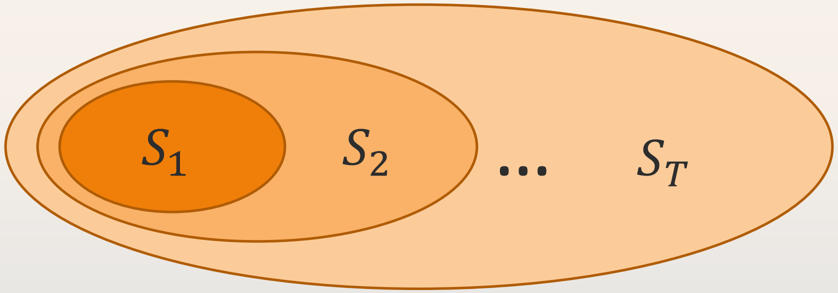
Proof Idea

❑ **Grouping based on priorities**

- $S_1 \subset S_2 \subset \cdots S_T$
- S_1 have the highest priority points.

❑ **Iterative Extension:**

- Given $\boldsymbol{f}_{i-1}\colon \boldsymbol{S}_{i-1} \to \mathbb{R}^{d_{i-1}}$
- Inductively construct $\boldsymbol{f}_i\colon \boldsymbol{S}_i \to \mathbb{R}^{d_i}$



Construction of $\boldsymbol{f}_i$ from $\boldsymbol{f}_{i-1}$

1. The new bi-Lipschitz outer-extension on f_{i-1}:

$$\boldsymbol{f'}_{i-1}(\boldsymbol{x}) = \widetilde{\boldsymbol{f}_{i-1}}(\boldsymbol{x}) \oplus \boldsymbol{h}(\boldsymbol{x})$$

2. Compose partially with the JL mapping J to reduce dimension

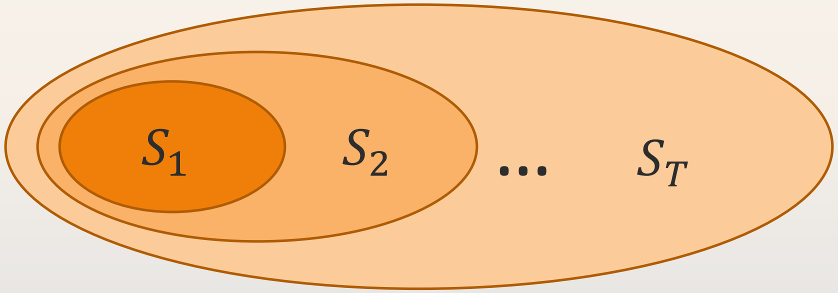
Final Map:

$$\boldsymbol{f}_i(\boldsymbol{x}) = \widetilde{\boldsymbol{f}_{i-1}}(\boldsymbol{x}) \oplus \boldsymbol{J}(\boldsymbol{h}(\boldsymbol{x}))$$

Details

☐ **Group sizes:**

- S_i: first $2^{2^{C^i}}$ points for a constant $C \approx 4$



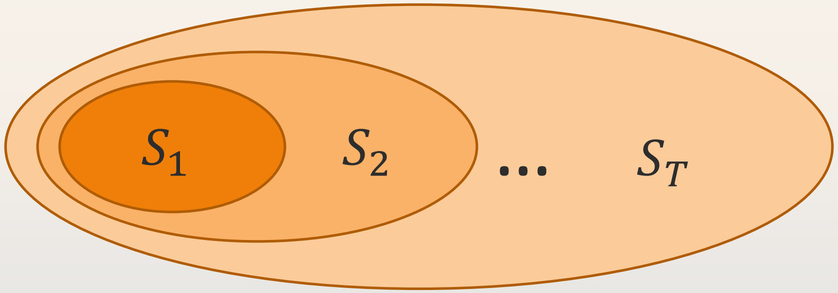
Details

❑ **Group sizes:**

- S_i: first $2^{2^{C^i}}$ points for a constant $C \approx 4$

❑ **Dimension:**

- A point at rank $r = 2^{2^{C^i}} + 1$:
- Dimension is $\log(2^{2^{C^{i+1}}}) = 2^{C^i \cdot C} = \left(2^{C^i}\right)^C = (\log r)^C$



Details

☐ **Group sizes:**

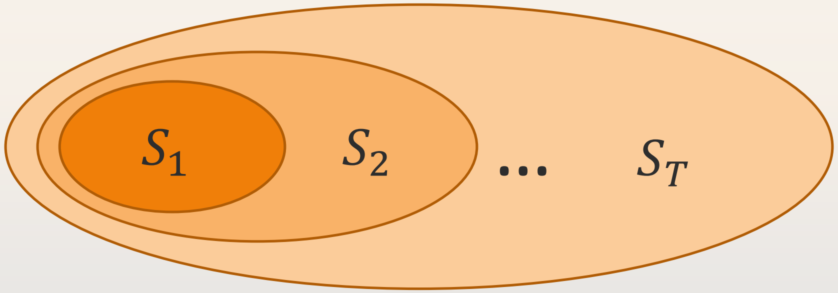
- S_i: first $2^{2^{C^i}}$ points for a constant $C \approx 4$

☐ **Dimension:**

- A point at rank $r = 2^{2^{C^i}} + 1$:
- Dimension is $\log(2^{2^{C^{i+1}}}) = 2^{C^i \cdot C} = \left(2^{C^i}\right)^C = (\log r)^C$

☐ **Distortion:**

- Distortion is $3^{\#groups} = 3^i = 3^{\log\log\log r} = O(\log\log r)$



Prioritized Dimension Reduction

	Distortion	#Non-zero
[Elkin, Filtser, Neiman, STOC'15]	$O_\epsilon(log^{4+\epsilon} r)$	$O_\epsilon(\log^4 r)$
This work	$O(\log\log r)$	$O(\frac{\log^{3+\epsilon} r}{\epsilon^2})$
Setting parameters differently	$O((3+\epsilon)^t)$	$O(\frac{\log r \log^{1/t} n}{\epsilon^2})$
Open Problem	$(1+\epsilon)$	$O(\frac{\log r}{\epsilon^2})$

Results

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	$\mathbf{D}$	To $\mathbb{R}^n$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta(\sqrt{\epsilon})$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{log^2 1/\epsilon})$	$\mathbb{R}^2$

☐ Show applications to dimension reduction

Extension by One Point

Given

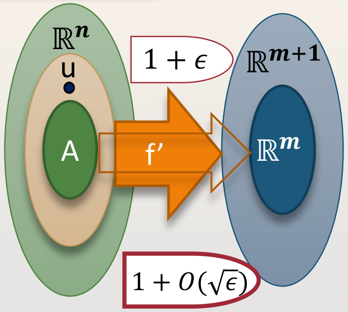
- a map $\boldsymbol{f}: \boldsymbol{A} \to \mathbb{R}^{\boldsymbol{m}}$ (where $\boldsymbol{A} \subset \mathbb{R}^{\boldsymbol{n}}$) that has distortion $1 + \epsilon$
- and any point $u \in \mathbb{R}^{n}$,

we can always extend the map to that point, i.e.,

$$\boldsymbol{f}': \boldsymbol{A} \cup \{\boldsymbol{u}\} \to \mathbb{R}^{\boldsymbol{m}+\boldsymbol{1}}$$

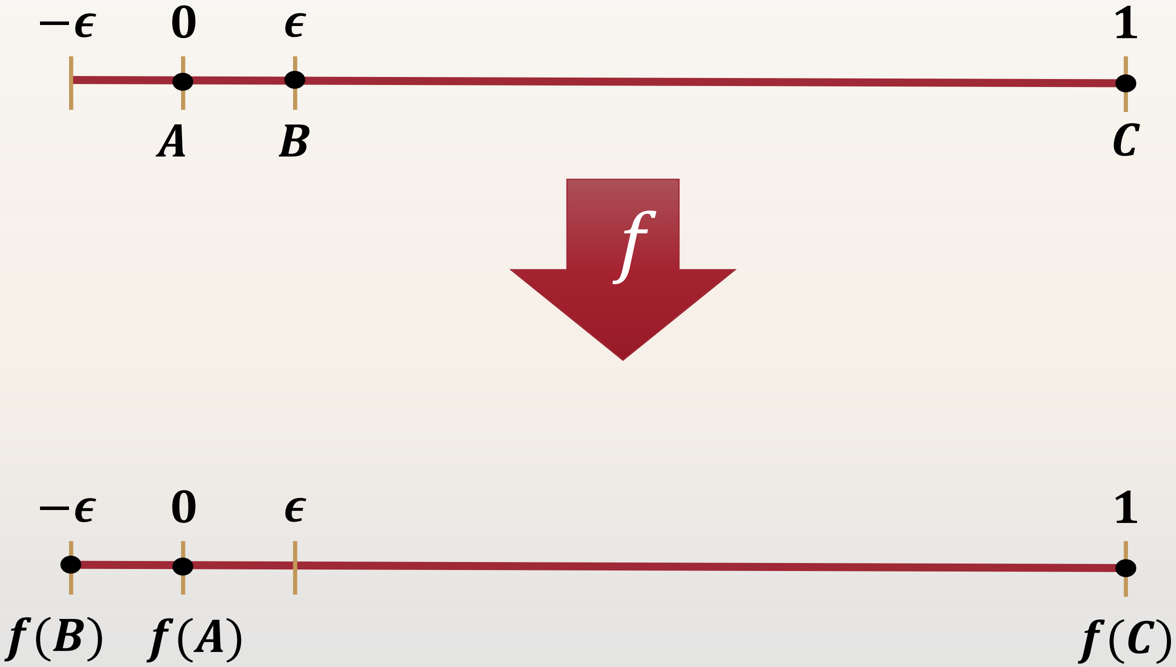
Increasing the distortion to $1 + \sqrt{\epsilon}$

Using Minimax Theorem



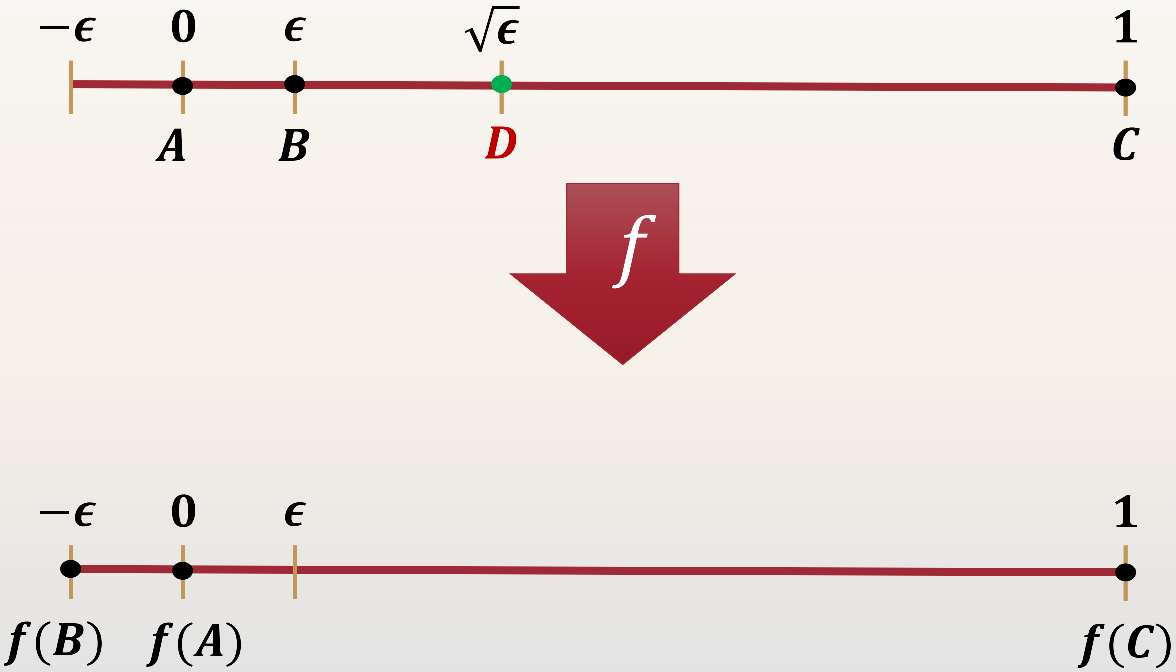
Lower Bound

- $f:\{A,B,C\} \to \mathbb{R}$ has distortion $(1+\epsilon)$



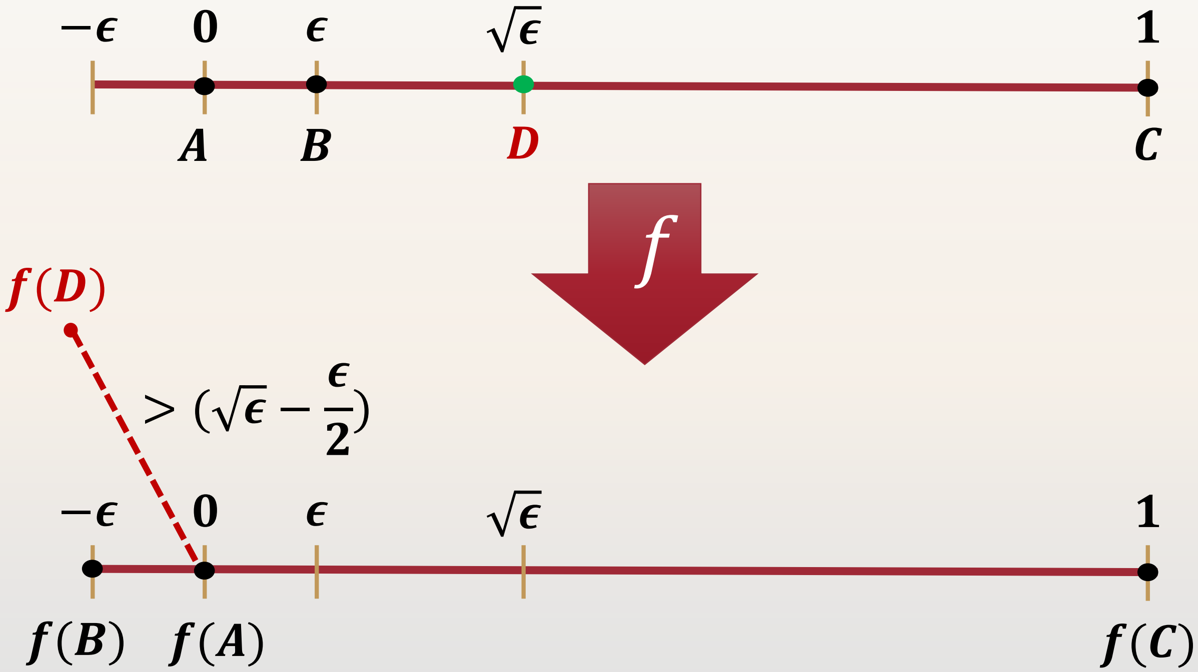
Lower Bound

- $f:\{A,B,C\} \to \mathbb{R}$ has distortion $(1+\epsilon)$
- Its extension to D increases the distortion to $1+\Omega(\sqrt{\epsilon})$



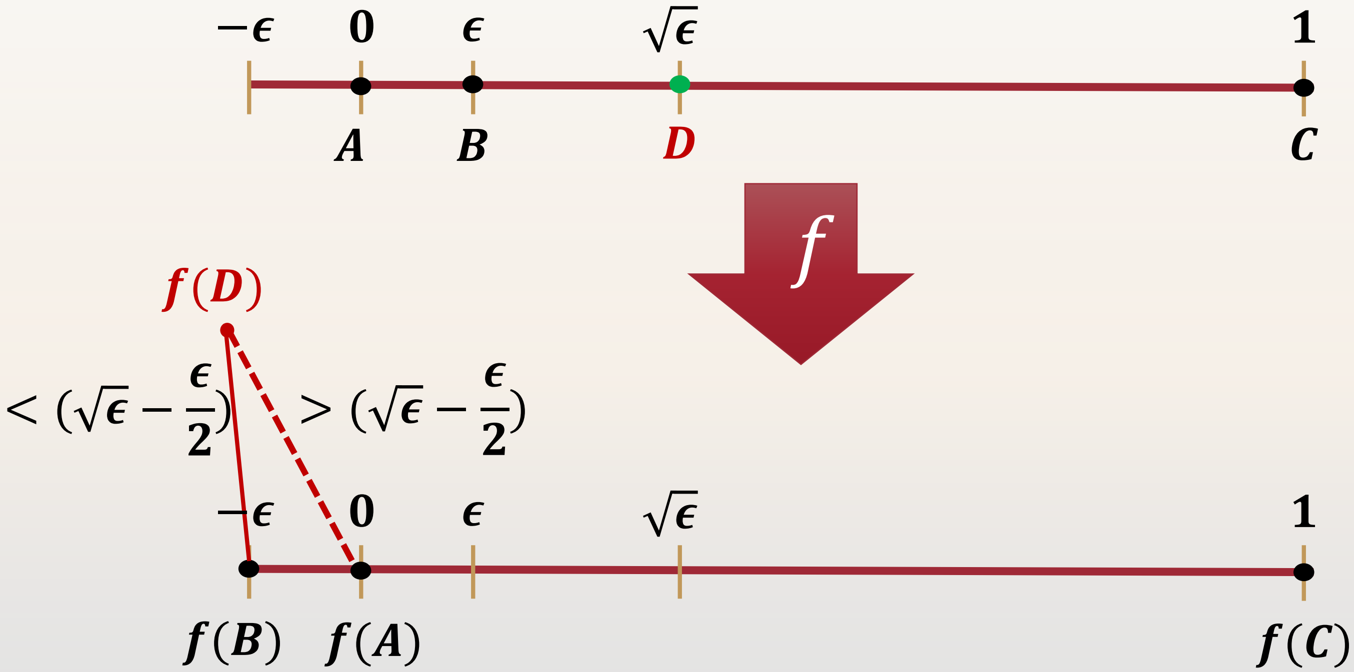
Lower Bound

- $f:\{A,B,C\} \rightarrow \mathbb{R}$ has distortion $(1+\epsilon)$
- Its extension to D increases the distortion to $1+\Omega\left(\sqrt{\epsilon}\right)$
- $\|f(D)-f(A)\| \geq \sqrt{\epsilon}\left(1-\frac{\sqrt{\epsilon}}{2}\right) = (\sqrt{\boldsymbol{\epsilon}}-\frac{\boldsymbol{\epsilon}}{\mathbf{2}})$



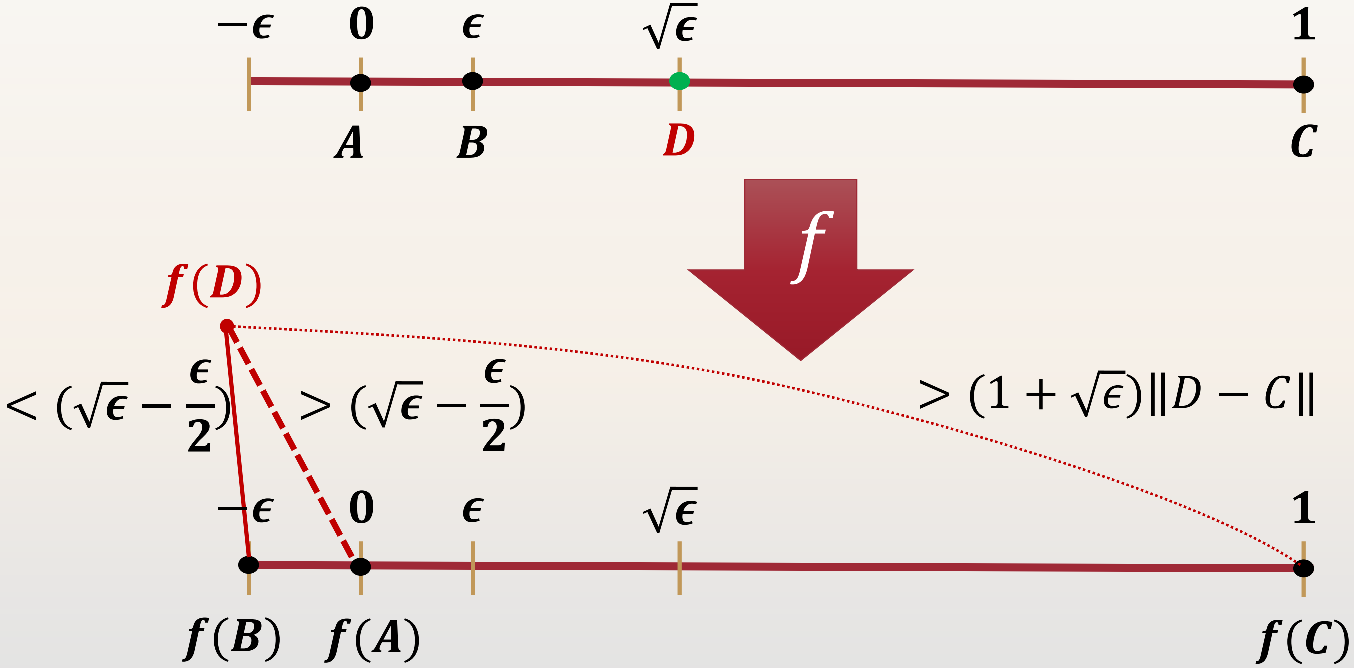
Lower Bound

- $f: \{A, B, C\} \to \mathbb{R}$ has distortion $(1 + \epsilon)$
- Its extension to D increases the distortion to $1 + \Omega\left(\sqrt{\epsilon}\right)$
- $\|f(D) - f(B)\| < \left(\sqrt{\epsilon} - \epsilon\right)\left(1 + \frac{\sqrt{\epsilon}}{2}\right) = \left(\sqrt{\epsilon} - \frac{\epsilon}{2}\right)$



Lower Bound

- $f\colon \{A, B, C\} \to \mathbb{R}$ has distortion $(1+\epsilon)$
- Its extension to D increases the distortion to $1 + \Omega\left(\sqrt{\epsilon}\right)$
- $\frac{\|f(D)-f(C)\|}{\|D-C\|} > \frac{1}{1-\sqrt{\epsilon}} > 1 + \sqrt{\epsilon}$



Terminal Dimension Reduction

Input: a set $X \subset \mathbb{R}^d$ of n terminals

Goal: find a map $f: \mathbb{R}^d \to \mathbb{R}^{d'}$ s.t. for any $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$ and any terminal $\boldsymbol{x} \in \boldsymbol{X}$,

$$\|\boldsymbol{x} - \boldsymbol{p}\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{p})\| \leq \boldsymbol{D} \cdot \|\boldsymbol{x} - \boldsymbol{p}\|$$

Terminal Dimension Reduction

Input: a set $X \subset \mathbb{R}^d$ of n terminals

Goal: find a map $f: \mathbb{R}^d \to \mathbb{R}^{d'}$ s.t. for any $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$ and any terminal $\boldsymbol{x} \in \boldsymbol{X}$,

$$\|\boldsymbol{x}-\boldsymbol{p}\| \leq \|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{p})\| \leq \boldsymbol{D} \cdot \|\boldsymbol{x}-\boldsymbol{p}\|$$

	Distortion	Dimension d'
[Elkin, Filtser, Neiman, '17]	$\boldsymbol{O}(\mathbf{1})$	$\boldsymbol{O}(\log \boldsymbol{n})$
This work	$\mathbf{1}+\boldsymbol{\epsilon}$	$\boldsymbol{O}(\frac{\log \boldsymbol{n}}{\boldsymbol{\epsilon}^4})$
[Narayanan, Nelson, '19]	$\mathbf{1}+\boldsymbol{\epsilon}$	$\boldsymbol{O}(\frac{\log \boldsymbol{n}}{\boldsymbol{\epsilon}^2})$

Terminal Dimension Reduction

Input: a set $X \subset \mathbb{R}^d$ of n terminals

Goal: find a map $f: \mathbb{R}^d \to \mathbb{R}^{d'}$ s.t. for any $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$ and any terminal $\boldsymbol{x} \in \boldsymbol{X}$,

$$\|\boldsymbol{x} - \boldsymbol{p}\| \leq \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{p})\| \leq \boldsymbol{D} \cdot \|\boldsymbol{x} - \boldsymbol{p}\|$$

- Apply JL on the set of terminals X to get a $(1+\epsilon^2)$-distortion embedding

$\frac{log\ n}{\epsilon^4}$ dimensions

	Distortion	Dimension d'
[Elkin, Filtser, Neiman, '17]	$\boldsymbol{O}(\mathbf{1})$	$\boldsymbol{O}(\log \boldsymbol{n})$
This work	$\mathbf{1} + \boldsymbol{\epsilon}$	$\boldsymbol{O}(\frac{\log \boldsymbol{n}}{\boldsymbol{\epsilon}^4})$
[Narayanan, Nelson, '19]	$\mathbf{1} + \boldsymbol{\epsilon}$	$\boldsymbol{O}(\frac{\log \boldsymbol{n}}{\boldsymbol{\epsilon}^2})$

Terminal Dimension Reduction

Input: a set $X \subset \mathbb{R}^d$ of n terminals

Goal: find a map $f: \mathbb{R}^d \to \mathbb{R}^{d'}$ s.t. for any $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$ and any terminal $\boldsymbol{x} \in \boldsymbol{X}$,

$$\|\boldsymbol{x}-\boldsymbol{p}\| \leq \|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{p})\| \leq \boldsymbol{D} \cdot \|\boldsymbol{x}-\boldsymbol{p}\|$$

- Apply JL on the set of terminals X to get a $(1+\epsilon^2)$-distortion embedding

$\frac{log\ n}{\epsilon^4}$ dimensions

- Use an extra dimension to simultaneously extend the map to all non-terminal points independently using our single point extension.

$(1+\epsilon)$-distortion

	Distortion	**Dimension d'**
[Elkin, Filtser, Neiman, '17]	$\boldsymbol{O}(\mathbf{1})$	$\boldsymbol{O}(\mathbf{log}\ \boldsymbol{n})$
This work	$\mathbf{1}+\boldsymbol{\epsilon}$	$\boldsymbol{O}(\frac{\mathbf{log}\ \boldsymbol{n}}{\boldsymbol{\epsilon}^{\mathbf{4}}})$
[Narayanan, Nelson, '19]	$\mathbf{1}+\boldsymbol{\epsilon}$	$\boldsymbol{O}(\frac{\mathbf{log}\ \boldsymbol{n}}{\boldsymbol{\epsilon}^{\mathbf{2}}})$

Results

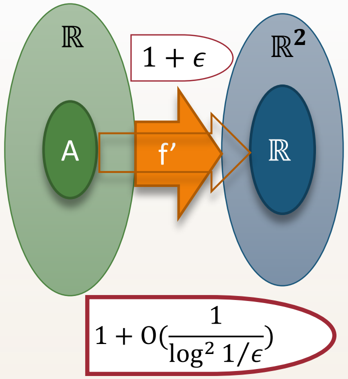
n	m	Initial distortion	Type of extension	New distortion	New image
any	any	$\mathbf{D}$	To $\mathbb{R}^{n}$	$\mathbf{3D}$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta\left(\sqrt{\epsilon}\right)$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{log^2 1/\epsilon})$	$\mathbb{R}^2$

☐ Show applications to dimension reduction

Extension to the Line

Given: a $(1+\epsilon)$-distortion map $f: A \to \mathbb{R}$ where $A \subset \mathbb{R}$

Goal: extend it to the whole line $\mathbb{R}$, i.e., $f': \mathbb{R} \to \mathbb{R}^2$



Extension to the Line

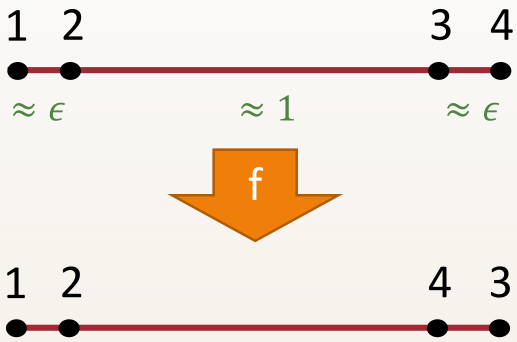
Given: a $(1+\epsilon)$-distortion map $f: A \to \mathbb{R}$ where $A \subset \mathbb{R}$

- such a map should be very structured.

Extension to the Line

Given: a $(1+\epsilon)$-distortion map $f: A \to \mathbb{R}$ where $A \subset \mathbb{R}$

- such a map should be very structured.

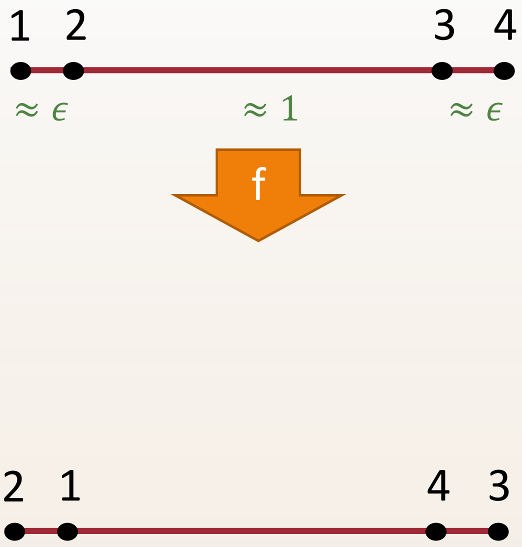


Permutations: permutation corresponding to the ordering defined by the map: $(1,2,4,3)$

Extension to the Line

Given: a $(1+\epsilon)$-distortion map $f: A \to \mathbb{R}$ where $A \subset \mathbb{R}$

- such a map should be very structured.

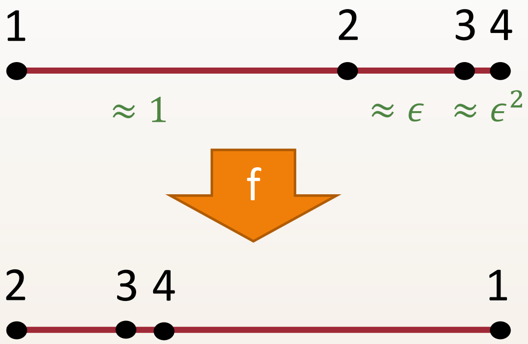


Permutations: permutation corresponding to the ordering defined by the map: $(1,2,4,3)$, $(2,1,4,3)$

Extension to the Line

Given: a $(1+\epsilon)$-distortion map $f\colon A \to \mathbb{R}$ where $A \subset \mathbb{R}$

- such a map should be very structured.

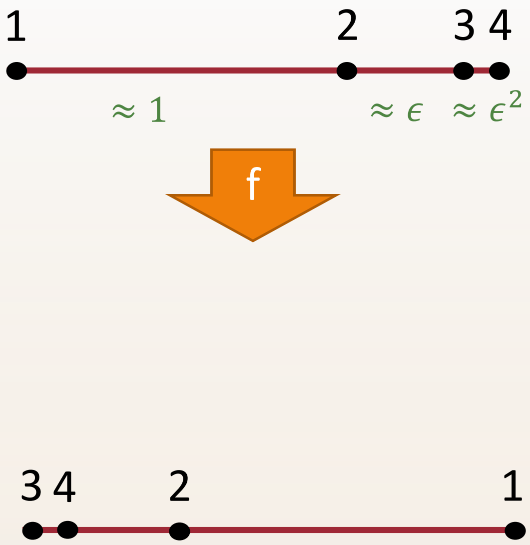


Permutations: permutation corresponding to the ordering defined by the map: $(1,2,4,3)$, $(2,1,4,3)$, $(2,3,4,1)$

Extension to the Line

Given: a $(1+\epsilon)$-distortion map $f: A \to \mathbb{R}$ where $A \subset \mathbb{R}$

- such a map should be very structured.



Permutations: permutation corresponding to the ordering defined by the map: $(1,2,4,3)$, $(2,1,4,3)$, $(2,3,4,1)$, $(3,4,2,1)$, …

Permutations

Consider the permutation corresponding to the ordering defined by the map. ***Are all permutations possible?***

Permutations

Consider the permutation corresponding to the ordering defined by the map. ***Are all permutations possible?***

- **Lemma 1:** a permutation is **valid** iff it **excludes** **(3,1,4,2)** and **(2,4,1,3)** as a “sub-permutation”

Permutation: 4 6 2 5 7 3 1

Permutations

Consider the permutation corresponding to the ordering defined by the map. ***Are all permutations possible?***

- **Lemma 1:** a permutation is **valid** iff it **excludes** **(3,1,4,2)** and **(2,4,1,3)** as a "sub-permutation"

Permutation: 4 6 2 5 7 3 1

Not valid: 4 **6 2** 5 **7 3** 1

6 2 7 3

3 1 4 2

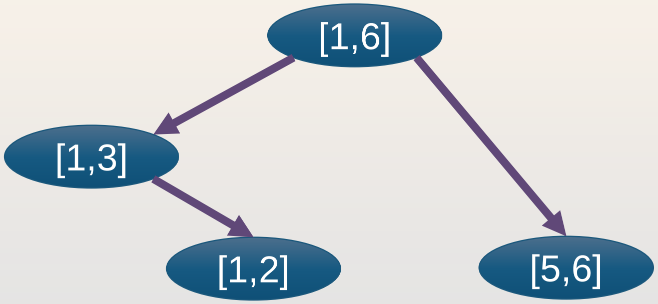
Permutations

Consider the permutation corresponding to the ordering defined by the map. *Are all permutations possible?*

- **Lemma 1:** a permutation is **valid** iff it **excludes** **(3,1,4,2)** and **(2,4,1,3)** as a “sub-permutation”

- **Lemma 2:** such a permutation can be decomposed into a sequence of **“laminar flips”** (reversing an interval)

$$(1,2,3,4,5,6) \rightarrow (\mathbf{3,2,1},4,5,6) \rightarrow (3,\mathbf{1,2},4,5,6) \rightarrow (3,1,2,4,\mathbf{6,5})$$



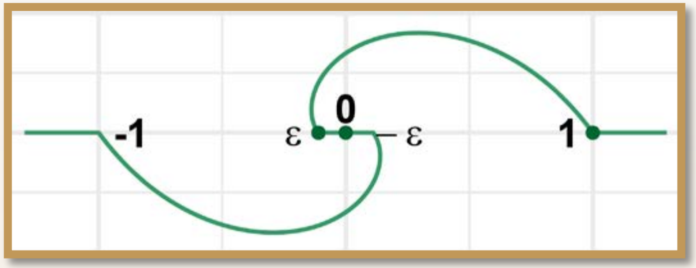
Spirals

- **Basic case:** consider f which maps $(0, \epsilon, 1)$ to $(0, -\epsilon, 1)$; extend it to the segment [0,1]

Spirals

☐ **Basic case:** consider f which maps $(0, \epsilon, 1)$ to $(0, -\epsilon, 1)$; extend it to the segment [0,1]

- Map using a single **spiral**
  - Map $[0, \epsilon]$ to $[0, -\epsilon]$ linearly
  - For $\epsilon \leq x \leq 1$ map x to $\boldsymbol{g(x) = (r(x), \phi(x))}$ in polar coordinates
  - $\boldsymbol{r(x) = x}$ and $\boldsymbol{\phi(x) = \frac{\pi \ln 1/x}{\ln 1/\epsilon}}$

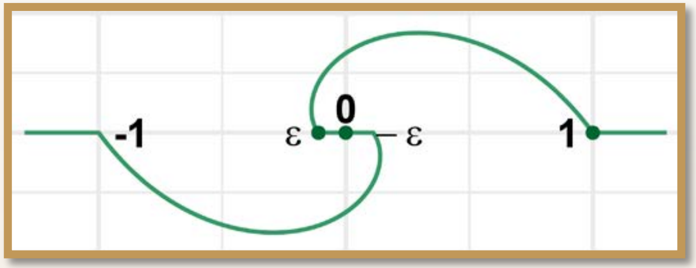


- **Distortion** is $1 + O(1/\ln^2(1/\epsilon))$
- This is **tight!**

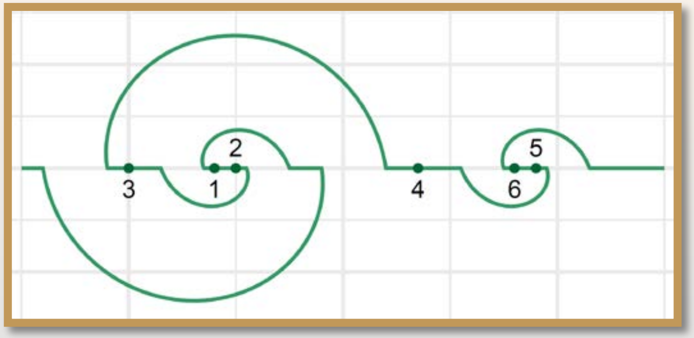
Spirals

☐ **Basic case:** consider f which maps $(0, \epsilon, 1)$ to $(0, -\epsilon, 1)$; extend it to the segment [0,1]

- Map using a single **spiral**
  - Map $[0, \epsilon]$ to $[0, -\epsilon]$ linearly
  - For $\epsilon \leq x \leq 1$ map x to $\boldsymbol{g(x) = (r(x), \phi(x))}$ in polar coordinates
  - $\boldsymbol{r(x) = x}$ and $\boldsymbol{\phi(x) = \frac{\pi \ln 1/x}{\ln 1/\epsilon}}$



- **Distortion** is $1 + O(1/\ln^2(1/\epsilon))$
- This is **tight!**



☐ **General case:** for each flip

- we add a spiral of the “right” scale

Open Problems

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	D	To $\mathbb{R}^n$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta(\sqrt{\epsilon})$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{log^2 1/\epsilon})$	$\mathbb{R}^2$
any	any	$1+\epsilon$	To $\mathbb{R}^n$	$1+g(\epsilon)$	?

❑ **Prioritized Dimension Reduction**

Distortion	#Non-zero
$1+\epsilon$	$(\log r)/\epsilon^2$

Open Problems

Thanks!

Questions?

n	m	Initial distortion	Type of extension	New distortion	New image
any	any	$\mathbf{D}$	To $\mathbb{R}^n$	$3D$	$\mathbb{R}^{n+m}$
any	any	$1+\epsilon$	one point	$1+\Theta\left(\sqrt{\epsilon}\right)$	$\mathbb{R}^{m+1}$
1	1	$1+\epsilon$	To $\mathbb{R}$	$1+\Theta(\frac{1}{log^2 1/\epsilon})$	$\mathbb{R}^2$
any	any	$1+\epsilon$	To $\mathbb{R}^n$	$1+g(\epsilon)$	?

❑Prioritized Dimension Reduction

Distortion	#Non-zero
$1+\epsilon$	$(\log r)/\epsilon^2$