Metric dimension reduction: A snapshot of the Ribe program

Assaf Naor
Princeton University
ICM 2018
The Ribe program: An organizing principle for the wild (intimidating?) diversity of metric spaces
The Ribe program: An organizing principle for the wild (intimidating?) diversity of metric spaces

\[ d_m : M \times M \rightarrow [0, \infty) \]

\[ d_m(x, y) \leq d_m(x, z) + d_m(z, y) \]
Also,

• Metrics that arise as solutions of optimization problems.
• Spaces of measures equipped with an optimal transport metric.
• Application-driven metrics (e.g. on strings and images).
• Many more...
A more tame world: Normed spaces
A more tame world: Normed spaces

$K \subseteq \mathbb{R}^k$ origin-symmetric convex body.

$x \in K \iff -x \in K$. 
A more tame world: Normed spaces
A more tame world: Normed spaces
A more tame world: Normed spaces
A more tame world: Normed spaces
A more tame world: Normed spaces
A more tame world: Normed spaces

\[ \|x\| = \frac{\alpha}{\beta}. \]
A more tame world: Normed spaces

\[ \| \cdot \| : \mathbb{R}^k \to [0, \infty) \text{ homogeneous of order 1} \]

\[ \| x + y \| \leq \| x \| + \| y \| \]

\[ d_{\| \cdot \|}(x, y) = \| x - y \| \]

\[ \text{Ball}_{\| \cdot \|} = \{ x \in \mathbb{R}^k : \| x \| \leq 1 \} = K. \]
A “Ribe-inspired” correspondence

ANALOGY

Normed space-inspired reasoning/intuition/phenomena
Informal statement: Isomorphic finite-dimensional linear properties of normed spaces are actually metric properties “in disguise.”
Example of isomorphic finite-dimensional linear property of an infinite-dimensional normed space $X$:

There is a constant $c>0$ such that for every integer $n$, for any unit vectors $x_1, \ldots, x_n \in X$ one can find signs $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 1\}$ (a choice of orientation for each of the given vectors) such that

$$\left\| \varepsilon_1 x_1 + \varepsilon_2 x_2 + \ldots + \varepsilon_n x_n \right\|_X \geq c \sqrt[3]{n}.$$ 

([Michel Talagrand, 1992]: deep relation to “Rademacher cotype 3.”)
A property $\mathcal{P}$ of a normed space $(X, \| \cdot \|_X)$ is said to be an isomorphic finite-dimensional linear property if any other normed space $(Y, \| \cdot \|_Y)$, all of whose finite dimensional linear subspaces occur in $X$ up to an arbitrarily large but fixed error, also has the property $\mathcal{P}$.
$Y$ must also have $\mathcal{P}$ if there is $C>0$ such that for any finite dimensional subspace $F$ of $Y$ (of any dimension and any “location”) there is a linear subspace $F'$ of $X$ and a linear isomorphism $T : F \to F'$ such that

$$\forall x \in F, \quad \|x\|_X \leq \|Tx\|_Y \leq C\|x\|_X.$$  

(In previous example, take $F = \text{span}\{x_1, \ldots, x_n\}$.)

Such properties are commonly called “local properties.”
Formal statement of Ribe’s rigidity theorem

Let $X$ and $Y$ be normed that are homeomorphic (as metric spaces, i.e., ignoring the linear structure altogether) via a bijection $f : X \to Y$ such that both $f$ and $f^{-1}$ are uniformly continuous. Then the isomorphic finite-dimensional linear properties of $X$ and $Y$ coincide.
Formal statement of Ribe’s rigidity theorem

Let $X$ and $Y$ be normed that are homeomorphic (as metric spaces, i.e., ignoring the linear structure altogether) via a bijection $f : X \to Y$ such that both $f$ and $f^{-1}$ are uniformly continuous. Then the isomorphic finite-dimensional linear properties of $X$ and $Y$ coincide.

With mere continuity rather than uniform continuity, this is false: Mikhail Kadec (1966): Any two separable infinite dimensional complete normed spaces are homeomorphic (as topological spaces).
Local theory (Grothendieck, James...)

- The study of isomorphic finite-dimensional linear properties of normed spaces.
Local theory (Grothendieck, James...)

- The study of isomorphic finite-dimensional linear properties of normed spaces.
- All norms on $\mathbb{R}^n$ are equivalent, so this is inherently quantitative and asymptotic in nature.
Local theory (Grothendieck, James...)

• The study of isomorphic finite-dimensional linear properties of normed spaces.

• All norms on $\mathbb{R}^n$ are equivalent, so this is inherently quantitative and asymptotic in nature.

• Over the past >60 years, developed by many mathematicians to a powerful theory (structural results, analytic tools, useful invariants).
Local theory (Grothendieck, James...)

- The study of isomorphic finite-dimensional linear properties of normed spaces.
- All norms on $\mathbb{R}^n$ are equivalent, so this is inherently quantitative and asymptotic in nature.
- Over the past >60 years, developed by many mathematicians to a powerful theory (structural results, analytic tools, useful invariants).
- Deep phenomena and connections to other areas of mathematics.
Local theory (Grothendieck, James...)

• The study of isomorphic finite-dimensional linear properties of normed spaces.
• All norms on $\mathbb{R}^n$ are equivalent, so this is inherently quantitative and asymptotic in nature.
• Over the past >60 years, developed by many mathematicians to a powerful theory (structural results, analytic tools, useful invariants).
• Deep phenomena and connections to other areas of mathematics.
• Decisive impact on several mathematical disciplines as well as application areas (e.g. computer science, statistics).
The Ribe program

• An organizing principle for general metric spaces based on analogy to the local theory.
The Ribe program

• An organizing principle for general metric spaces based on analogy to the local theory.

• The Ribe theorem indicates that there is a hidden dictionary that translates linear properties, phenomena, insights and intuitions that a priori make sense only for normed spaces, to metric spaces.
The Ribe program

• An organizing principle for general metric spaces based on analogy to the local theory.

• The Ribe theorem indicates that there is a hidden dictionary that translates linear properties, phenomena, insights and intuitions that a priori make sense only for normed spaces, to metric spaces.

• Formulated by Jean Bourgain (1986). Key insights, theorems and leadership by Joram Lindenstrauss.
The Ribe program

- An organizing principle for general metric spaces based on analogy to the local theory.
- The Ribe theorem indicates that there is a hidden dictionary that translates linear properties, phenomena, insights and intuitions that a priori make sense only for normed spaces, to metric spaces.
- Formulated by Jean Bourgain (1986). Key insights, theorems and leadership by Joram Lindenstrauss.
- A rich network of questions, many are about how to define fundamental concepts. Wide-ranging implications/applications.
A “Ribe-inspired” correspondence

ANALOGY

Normed space-inspired reasoning/intuition/phenomena
Nonpositive curvature

∀ p, q ∈ ∂Δₓ, |ψ(p) − ψ(q)| ≥ dₓ(p, q).
Coarse non-universality of nonpositive curvature

Theorem [Eskenazis, Mendel, N., 2018]: There exists \((m, d_m)\) such that for no \((X, d_X)\) of nonpositive curvature there is \(f : M \rightarrow X\) satisfying

\[
\forall x, y \in M, \quad \omega(d_m(x, y)) \leq d_X(f(x), f(y)) \leq \Omega(d_m(x, y)),
\]

where \(\lim_{s \rightarrow \infty} \omega(s) = \infty\).

Answers a question of Mikhail Gromov (1993).

Not true for nonnegative curvature (Andoni, N., Neiman, 2015).
Snapshot of Ribe program: Metric dimension reduction

For a **panorama** of the Ribe program, several surveys are available (including the proceedings of this ICM), as well as forthcoming lecture series that will be recorded and available online.
The broader term *dimension reduction* is one of the most important issues that are being tackled in areas such as statistics, machine learning and theoretical computer science. It refers to the desire to decrease the degrees of freedom of a (seemingly) high-dimensional data set while approximately preserving some of its pertinent features. Here we will focus on this topic only from within pure mathematics, and only from the geometric perspective, namely approximate preservation of metrical characteristics of the data set.

Suppose that one is given $1,000,000,000$ vectors

$$x_1, \ldots, x_{1,000,000,000} \in \mathbb{R}^{1,000,000,000}.$$  

Then, there are new vectors $y_1, \ldots, y_{1,000,000,000} \in \mathbb{R}^{329}$ such that

$$\forall i, j, \quad |x_i - x_j| \leq |y_i - y_j| \leq 2|x_i - x_j|.$$
Can crude information on boundedly many distances force the ambient dimension to be high?

Let $(X, d_X)$ be a metric space for which a natural notion of “dimension” is defined, e.g. a finite-dimensional normed space, certain manifolds (with geometric restrictions), a space of measures (with optimal transport metric) over a finite geometry...
Can crude information on boundedly many distances force the ambient dimension to be high?

Let \((X, d_X)\) be a metric space for which a natural notion of “dimension” is defined, e.g. a finite-dimensional normed space, certain manifolds (with geometric restrictions), a space of measures (with optimal transport metric) over a finite geometry...

Is there a universal constant $\alpha \geq 1$ such that for any $n$-point metric space $(M, d_M)$ one can find $k_m \in \mathbb{N}$ with

$$k_m = O(\log n),$$

a norm $\| \cdot \|_m$ on $\mathbb{R}^{k_m}$ and an embedding $f : M \to \mathbb{R}^{k_m}$ such that

$$\forall x, y \in M, \quad d_M(x, y) \leq \|f(x) - f(y)\|_m \leq \alpha d_m(x, y)?$$

(Terminology: $(M, d_M)$ embeds into $(\mathbb{R}^{k_m}, \| \cdot \|_m)$ with distortion $\alpha.$)
Why \( \log n \)?

- This is a natural barrier, even when one considers the trivial (equilateral) \( n \)-point metric space \((m, d_m)\) in which all distances between distinct points are equal to 1.

\[
x \neq y \implies 1 \leq \| f(x) - f(y) \|_m \leq \alpha.
\]
\[ f(x_0) \]
\[ f(x_0) \]
$f(y)$

$\bullet$

$\bullet$

$f(z)$
$f(y) \geq 1 \Rightarrow f(\hat{z})$
\[ f(y) \]

\[ f(z) \]

\[ \frac{1}{2} \]

\[ \frac{1}{2} \]
Volume comparison

\[
\left( \alpha + \frac{1}{2} \right)^{k_m} \text{vol} \left( \text{Ball}_{\| \cdot \|_m} \right) = \text{vol} \left( f(x_0) + \left( \alpha + \frac{1}{2} \right) \text{Ball}_{\| \cdot \|_m} \right)
\]

\[
\geq \text{vol} \left( \bigcup_{y \in m} \left( y + \frac{1}{2} \right) \text{Ball}_{\| \cdot \|_m} \right) = n \frac{\text{vol} \left( \text{Ball}_{\| \cdot \|_m} \right)}{2^{k_m}}.
\]

So, necessarily \( k_m \gtrsim_\alpha \log n \).
Why $\log n$?

- William B. Johnson and Joram Lindenstrauss were motivated by a (then conjectural) metric space analogue of an important theorem of Fritz John (1948) in the local theory of Banach spaces, which would have followed from the desired $O(\log n)$. This was later solved differently by the influential Bourgain embedding theorem (1985).
Why $\log n$?

- William B. Johnson and Joram Lindenstrauss were motivated by a (then conjectural) metric space analogue of an important theorem of Fritz John (1948) in the local theory of Banach spaces, which would have followed from the desired $O(\log n)$. This was later solved differently by the influential Bourgain embedding theorem (1985).

- More generally, within the Ribe program one first attempts to replace the dimension of a normed space by the logarithm of the cardinality of a metric space (more nuanced adjustments of this idea are sometimes needed, but this is a meaningful place to start).
Why $\log n$?

• Johnson and Lindenstrauss proved that the answer is \textbf{positive} for Euclidean metrics.

The Johnson-Lindenstrauss lemma (1984): Suppose that $x_1, \ldots, x_n$ are vectors in a Hilbert space $H$ and $\alpha > 1$. There are $k \in \mathbb{N}$ and new vectors $y_1, \ldots, y_n$ in $\mathbb{R}^k$ such that

\[ k \lesssim_\alpha \log n, \]

\[ \forall i, j \in \{1, \ldots, n\}, \quad \|x_i - x_j\|_H \leq |y_i - y_j| \leq \alpha \|x_i - x_j\|_H. \]
Johnson-Lindenstrauss flattening/compression

By tracking the bounds in the JL proof, one gets that if one is given $1,000,000,000$ vectors $x_1, \ldots, x_{1,000,000,000}$ in $\mathbb{R}^{1,000,000,000}$, and one wishes to preserve all pairwise distances up to a factor 2, then this can be achieved using a configuration of vectors

$$y_1, \ldots, y_{1,000,000,000} \in \mathbb{R}^{329}.$$
Johnson-Lindenstrauss flattening/compression

Numerous applications in many areas. Original use was as a lemma towards a basic theorem about extension of Lipschitz functions.
Bourgain’s solution/question

Problem (Johnson-Lindenstrauss): Does any $n$-point metric space embed with distortion $O(1)$ into some normed space of dimension $O(\log n)$?
Bourgain’s solution/question

Problem (Johnson-Lindenstrauss): Does any $n$-point metric space embed with distortion $O(1)$ into some normed space of dimension $O(\log n)$?

Theorem (Bourgain, 1985): For any $n$ there exists an $n$-point metric space $(M, d_M)$ such that for every norm $\| \cdot \|$ on $\mathbb{R}^k$, if $(M, d_M)$ embeds into $(\mathbb{R}^k, \| \cdot \|)$ with distortion $O(1)$, then necessarily

$$k \gtrsim \left( \frac{\log n}{\log \log n} \right)^2.$$
Proof uses randomness in two ways: Random graphs and the Johnson-Lindenstrauss lemma itself (whose proof is probabilistic).

Not quite as good as what Johnson and Lindenstrauss hoped for, but still a slowly growing lower bound.

**Question (Bourgain, 1985)**: What is the true asymptotic behavior?

Perhaps one could always embed into some space of dimension $(\log n)^{O(1)}$?
Theorem (Nathan Linial, Eran London, Yuri Rabinovich, 1995): The conclusion of Bourgain’s theorem can be improved to

\[ k \gtrsim (\log n)^2. \]
Jiří Matoušek’s solution (1996)

**Theorem**: The conclusion of Bourgain’s theorem can be improved to

\[ k \geq n^c, \]

for some universal constant \( c > 0 \).
Matoušek’s elegant solution of the Johnson-Lindenstrauss/Bourgain problem is a beautiful combination of combinatorics, functional analysis, and real algebraic geometry.

Impossibility of average dimension reduction

Theorem (N., 2016): For every \( n \) there exists an \( n \)-point metric space \((\mathcal{M}, d_{\mathcal{M}})\) such that for every norm \( \| \cdot \| \) on \( \mathbb{R}^k \), if there were an embedding \( f : \mathcal{M} \to \mathbb{R}^k \) such that

\[
\forall x, y \in \mathcal{M}, \quad \| f(x) - f(y) \| \lesssim d_{\mathcal{M}}(x, y),
\]

and

\[
\frac{1}{n^2} \sum_{x, y \in \mathcal{M}} \| f(x) - f(y) \| \gtrsim \frac{1}{n^2} \sum_{x, y \in \mathcal{M}} d_{\mathcal{M}}(x, y),
\]

then necessarily \( k \geq n^c \) for some universal constant \( c > 0 \).
Expanders

An $n$-vertex regular graph $G = (V, E)$ is a $O(1)$-expander if its degree is $O(1)$ and every $S \subseteq V$ with $|S| \leq n/2$ satisfies

$$\left| \left\{ \{u, v\} \in E : \{u, v\} \cap S \neq \emptyset \land \{u, v\} \setminus S \neq \emptyset \right\} \right| \geq |S|.$$
Expanders

An $n$-vertex regular graph $G = (V, E)$ is a $O(1)$-expander if its degree is $O(1)$ and every $S \subseteq V$ with $|S| \leq n/2$ satisfies

$$\left| \left\{ \{u, v\} \in E : \{u, v\} \cap S \neq \emptyset \land \{u, v\} \setminus S \neq \emptyset \right\} \right| \geq |S|.$$
Expanders

An $n$-vertex regular graph $G = (V, E)$ is a $O(1)$-expander if its degree is $O(1)$ and every $S \subseteq V$ with $|S| \leq n/2$ satisfies

$$\left| \{\{u, v\} \in E : \{u, v\} \cap S \neq \emptyset \land \{u, v\} \setminus S \neq \emptyset \} \right| \geq |S|.$$
Expanders

An $n$-vertex regular graph $G = (V, E)$ is a $O(1)$-expander if its degree is $O(1)$ and every $S \subseteq V$ with $|S| \leq n/2$ satisfies

$$\left| \left\{ \{u, v\} \in E : \{u, v\} \cap S \neq \emptyset \land \{u, v\} \setminus S \neq \emptyset \right\} \right| \geq |S|.$$ 

Always equipped with the shortest-path metric

$$d_G : V \times V \to \mathbb{N} \cup \{0\}.$$
Geometric meaning of expander

Puzzle: What is the average (Euclidean) distance between the audience in this lecture hall?
Geometric meaning of expander
Geometric meaning of expander

(TRIVIAL) Puzzle: What is the average (Euclidean) distance between the audience in this lecture hall?

Answer: join each pair of audience members by a straight line, calculate its length, and compute the average of the resulting numbers.

\[
\left( \frac{\# \text{ audience members}}{2} \right)
\]

numbers.
Geometric meaning of expander

Puzzle: What is the average (Euclidean) distance between the audience in this lecture hall?
\[ \binom{5,350}{2} = 14,308,575. \]
A faster (approximate) way?

Suppose that we are happy with knowing the answer up to a universal constant factor.

(Concretely, factor 8 works below; with more effort $1 + \varepsilon$.)
Impose arbitrarily the structure of a 3-regular graph on the audience members, i.e., a graph

\[ G = (\{1, \ldots, 5350\}, E) \]

with

\[ |E| = \frac{3 \times 5350}{2} = 8,025, \]

and an arbitrary bijection \( \{1, \ldots, 5350\} \leftrightarrow \{\text{seats in lecture hall}\} \).
Suggestion for an approximate answer: Evaluate the distances only between those audience members that are joined by an edge.
Geometric meaning of expander
Geometric meaning of expander
Geometric meaning of expander
Geometric meaning of expander
Once this recipe (pairs of distances to be computed) is declared, we demand that it will always work, with 100% certainly.
Once this recipe (pairs of distances to be computed) is declared, we demand that it will always work \textit{with 100\% certainty}.  

E.g. it must output the correct estimate regardless of where the audience members will be, say, \textit{tomorrow at 7 PM}. 
Geometric meaning of expander

**Fact**: The average of the resulting 8,035 numbers will be within a $O(1)$ factor of the actual average of all the 14,308,575 distances, regardless of the bijection and the location of the 5,350 points, if and only if the “template graph” $G$ is a $O(1)$ expander.
Definition: Let \((M, d_M)\) be a metric space. An \(n\)-vertex \(O(1)\)-regular graph \(G = (V, E)\) is a \(O(1)\)-expander with respect to \((M, d_M)\) if every configuration of \(n\) points \(\{x_u\}_{u \in V} \subseteq M\) satisfies

\[
\frac{1}{n^2} \sum_{u,v \in V} d_M(x_u, x_v) \approx \frac{1}{|E|} \sum_{\{u,v\} \in E} d_M(x_u, x_v).
\]
Expanders with respect to metric spaces

Definition: Let \((M, d_M)\) be a metric space. An \(n\)-vertex \(O(1)\)-regular graph \(G = (V, E)\) is a \(O(1)\)-expander with respect to \((M, d_M)\) if every configuration of \(n\) points \(\{x_u\}_{u \in V} \subseteq M\) satisfies

\[
\frac{1}{n^2} \sum_{u,v \in V} d_M(x_u, x_v) \leq \frac{1}{|E|} \sum_{\{u,v\} \in E} d_M(x_u, x_v).
\]

Generally, for \(C > 1\) it is a \(C\)-expander with respect to \((M, d_M)\) if

\[
\frac{1}{n^2} \sum_{u,v \in V} d_M(x_u, x_v) \leq \frac{C}{|E|} \sum_{\{u,v\} \in E} d_M(x_u, x_v).
\]
Expanders with respect to metric spaces

**Definition**: Let \((M, d_M)\) be a metric space. An \(n\)-vertex \(O(1)\)-regular graph \(G = (V, E)\) is a \(O(1)\)-expander with respect to \((M, d_M)\) if every configuration of \(n\) points \(\{x_u\}_{u \in V} \subseteq M\) satisfies

\[
\frac{1}{n^2} \sum_{u,v \in V} d_M(x_u, x_v) \gtrsim \frac{1}{|E|} \sum_{\{u,v\} \in E} d_M(x_u, x_v).
\]

Generally, for \(C > 1\) it is a \(C\)-expander with respect to \((M, d_M)\) if

\[
\frac{1}{|E|} \sum_{\{u,v\} \in E} d_M(x_u, x_v) \lesssim \frac{1}{n^2} \sum_{u,v \in V} d_M(x_u, x_v) \leq \frac{C}{|E|} \sum_{\{u,v\} \in E} d_M(x_u, x_v).
\]

always valid
• The classical notion of expander coincide with being an expander with respect to Hilbert space (even with respect to \{0,1\}).
• The classical notion of expander coincide with being an expander with respect to Hilbert space (even with respect to \{0,1\}).

• In general, the situation is much more subtle, and there are many unsolved important mysteries about expanders with respect to metric spaces; we have been investigating this systematically (mostly with Manor Mendel) over the past decade.
• The classical notion of expander coincide with being an expander with respect to Hilbert space (even with respect to $\{0,1\}$).

• In general, the situation is much more subtle, and there are many unsolved important mysteries about expanders with respect to metric spaces; we have been investigating this systematically (mostly with Manor Mendel) over the past decade.

• There are “nice” metric spaces that do not admit any expander.
• The classical notion of expander coincide with being an expander with respect to Hilbert space (even with respect to \(\{0,1\}\)).
• In general, the situation is much more subtle, and there are many unsolved important mysteries about expanders with respect to metric spaces; we have been investigating this systematically (mostly with Manor Mendel) over the past decade.
• There are “nice” metric spaces that do not admit any expander.
• There are metric spaces with respect to which uniformly random regular graphs are not expanders (with overwhelming probability), yet one can find specially-crafted expanders with respect to them.
• The classical notion of expander coincide with being an expander with respect to Hilbert space (even with respect to \(\{0,1\}\)).

• In general, the situation is much more subtle, and there are many unsolved important mysteries about expanders with respect to metric spaces; we have been investigating this systematically (mostly with Manor Mendel) over the past decade.

• There are “nice” metric spaces that do not admit any expander.

• There are metric spaces with respect to which uniformly random regular graphs are not expanders (with overwhelming probability), yet one can find specially-crafted expanders with respect to them.

• There are “superexpanders” that are simultaneously expanders with respect to a large class of spaces (Vincent Lafforgue, Mendel-N.).
Theorem (N., 2016): Any $O(1)$-expander (in the classical sense) is also an $O(\log k)$-expander with respect to any norm $\| \cdot \|$ on $\mathbb{R}^k$.

This is sharp.
So, if $G = (V, E)$ is an $n$-vertex $O(1)$-expander, and $f : V \to \mathbb{R}^k$ satisfies

\[
\frac{1}{|E|} \sum_{\{u,v\} \in E} \| f(u) - f(v) \| \lesssim \frac{1}{|E|} \sum_{\{u,v\} \in E} d_G(u,v) = 1,
\]

\[
\frac{1}{n^2} \sum_{u,v \in V} \| f(u) - f(v) \| \gtrsim \frac{1}{n^2} \sum_{u,v \in V} d_G(u,v) \asymp \log n.
\]

Then necessarily

\[
\log k \gtrsim \log n \iff k \geq n^c.
\]
Theorem (N., 2016): Any $O(1)$-expander does not admit a Lipschitz embedding into dimension $n^{o(1)}$ that preserves the average distance.
Theorem (N., 2016): Any $O(1)$-expander does not admit a Lipschitz embedding into dimension $n^{o(1)}$ that preserves the average distance.

- Entirely different solution of the Johnson-Lindenstrauss/Bourgain problem.
Theorem (N., 2016): Any $O(1)$-expander does not admit a Lipschitz embedding into dimension $n^o(1)$ that preserves the average distance.

- Entirely different solution of the Johnson-Lindenstrauss/Bourgain problem.
- Provides a criterion for intrinsic high dimensionality (expansion). In contrast, Matoušek’s theorem is an existential statement: An algebraic proof by contradiction that all members of a certain specially-crafted family of metric spaces cannot reside with $O(1)$ distortion in a low-dimensional normed space.
Theorem (N., 2016): Any $O(1)$-expander does not admit a Lipschitz embedding into dimension $n^{o(1)}$ that preserves the average distance.

- Entirely different solution of the Johnson-Lindenstrauss/Bourgain problem.
- Provides a criterion for intrinsic high dimensionality (expansion). In contrast, Matoušek’s theorem is an existential statement: An algebraic proof by contradiction that all members of a certain specially-crafted family of metric spaces cannot reside with $O(1)$ distortion in a low-dimensional normed space.
- The distortion of any low-dimensional embedding now occurs on average.
Theorem (N., 2016): Any $O(1)$-expander does not admit a Lipschitz embedding into dimension $n^{o(1)}$ that preserves the average distance.

- Entirely different solution of the Johnson-Lindenstrauss/Bourgain problem.
- Provides a criterion for intrinsic high dimensionality (expansion). In contrast, Matoušek’s theorem is an existential statement: An algebraic proof by contradiction that all members of a certain specially-crafted family of metric spaces cannot reside with $O(1)$ distortion in a low-dimensional normed space.
- The distortion of any low-dimensional embedding now occurs on average.
- Answers question of Andoni-Nguyen-Nikolov-Razenshteyn-Waingarten (2016). They proved that if the opposite conclusion were true (namely, some expander does reside in a low-dimensional normed space), then this could be used to confirm a belief that a certain algorithmic task is impossible.
Bad luck, or an opportunity?
Wait! Clearly any normed space contains many “geometric” graphs

Let $\| \cdot \|$ be a norm on $\mathbb{R}^k$. Take any $V \subseteq \mathbb{R}^k$ and join two elements $x, y$ of $V$ by an edge iff

$$\| u - v \|_X \leq 1.$$
Provided \( k \) isn’t very large, the presence of \( G \) in \((\mathbb{R}^k, \| \cdot \|)\) can be reconciled with the previous theorem only if at least one of the following two scenarios occurs.

**Scenario 1:** The average distance in \( V \) is not large. This implies that there is a substantially smaller ball that contains a large fraction of the points of \( V \), i.e., \( V \) has a large “clustered” subset.

**Scenario 2:** \( G \) is not an expander. By definition of expansion, this means that \( V \) can be partitioned into two parts with few edges crossing the boundary, i.e., the parts are “mostly” far apart.
An intrinsic partitioning scheme

Andoni-N.-Nikolov-Razenshteyn-Waingarten (2018): This dichotomy can be iterated to provide a hierarchical partition that yields a new structural description of finite-dimensional normed spaces.

The intrinsic geometry governs the partitioning procedure. Each “cutting step” comes from either a dense geometric ball (and its complement) or a “spectral partition” that is imposed (Cheeger’s inequality) by the failure of expansion of the geometric graph.
• Need a version of the theorem for weighted graphs of unbounded degree (the proof gives that automatically).
• Actual implementation is randomized, resulting in a random partition.
• Substantial algorithmic issues; “metric Rayleigh quotient version” needed.
A nearest-neighbor data structure circumventing the “curse of dimensionality” for general norms
Approximate nearest neighbor search

- A data set $D$ consisting of $n$ points in a normed space $(\mathbb{R}^k, \| \cdot \|)$ is an array of $kn$ numbers.
Approximate nearest neighbor search

• A data set $D$ consisting of $n$ points in a normed space $(\mathbb{R}^k, \| \cdot \|)$ is an array of $kn$ numbers.

• A new point $q \in \mathbb{R}^k$ comes along. What is its (approximately) closest point $p$ in $D$?
Approximate nearest neighbor search

- A data set $D$ consisting of $n$ points in a normed space $(\mathbb{R}^k, \| \cdot \|)$ is an array of $kn$ numbers.
- A new point $q \in \mathbb{R}^k$ comes along. What is its (approximately) closest point $p$ in $D$?
- Can be answered exactly in time $n$ (store $D$ as a “pile of books”).
Approximate nearest neighbor search

• A data set $D$ consisting of $n$ points in a normed space $(\mathbb{R}^k, \| \cdot \|)$ is an array of $kn$ numbers.

• A new point $q \in \mathbb{R}^k$ comes along. What is its (approximately) closest point $p$ in $D$?

• Can be answered exactly in time $n$ (store $D$ as a “pile of books”).

• Build from $D$ a “library” such that there is an approximate answer in sublinear time, i.e., $o(n)$. 
• Query time (\# of distance computations) $n$ is immediate.
• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with $k$. 
• Query time (# of distance computations) $n$ is immediate.
• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with $k$. 
• Query time (\# of distance computations) $n$ is immediate.
• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with $k$. 
• Query time (\# of distance computations) $n$ is immediate.
• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with $k$. 

![Diagram of curse of dimensionality](image)
• Query time (\# of distance computations) \( n \) is immediate.
• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with \( k \).
• Query time (number of distance computations) $n$ is immediate.
• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with $k$. 
• Query time (number of distance computations) $n$ is immediate.

• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with $k$.

• [Indyk-Motwani, 1998], [Kushilevitz-Ostrovsky-Rabani, 1998]: Constructed a data structure of size $(kn)^c$, using which one can output in time that is **sublinear in $n$** (and **polynomial in $k$**) a data point that is within a factor $k^c$ the closest one.
• Query time (\# of distance computations) \( n \) is immediate.

• **Curse of dimensionality**: Query time and data structure size deteriorating exponentially with \( k \).

• [Indyk-Motwani, 1998], [Kushilevitz-Ostrovsky-Rabani, 1998]: Constructed a data structure of size \( (kn)^c \), using which one can output in time that is **sublinear in \( n \)** (and **polynomial in \( k \)**) a data point that is within a factor \( k^c \) the closest one.

(Using the Johnson-Lindenstrauss lemma and John’s theorem.)
A nearest-neighbor data structure circumventing the “curse of dimensionality” for general norms

Andoni-N.-Nikolov-Razenshteyn-Waingarten (2018): Given $n$ points in a $k$-dimensional normed space $X$, there is a data structure of size $k^C n^{1+\varepsilon}$ such that given a point $q$ in $X$, it outputs in time $k^C n^\varepsilon$ a data point $p$ which is guaranteed to be within a factor $k^\delta$ the closest point to $q$. 
Spaces admitting metric dimension reduction

A normed space $X$ admits metric dimension reduction if every $n$ points in $X$ embed with distortion $O(1)$ into some linear subspace of $X$ of dimension $n^{o(1)}$.

- **Johnson-Lindenstrauss**: $\ell_2$ admits metric dimension reduction.
- **Matoušek**: $\ell_\infty$ does not admit metric dimension reduction.
- **Johnson-N. (2010)**: There exists a Banach space that admits metric dimension reduction yet isn’t isomorphic to a Hilbert space.
- **Brinkman-Charikar (2005)**: $\ell_1$ does not admit metric dimension reduction. [N.-Young, 2018]: Conceptually different new example.
Open questions

• What happens in $\ell_p$ when $p \notin \{1, 2, \infty\}$?
• Positive results in $\ell_1$?
• Metric dimension reduction in spaces that are not normed spaces?

*Metric dimension reduction is fertile ground for deep interactions between diverse areas of mathematics, powerful applications. Despite many achievements, the present state of the field awaits new ideas and methods for addressing a remarkable wealth of important longstanding mysteries.*