The Traveling Salesman Problem, Data Parametrization and Multi-resolution Analysis

Raanan Schul

Stony Brook

Motivation

(which I usually give to mathematicians)

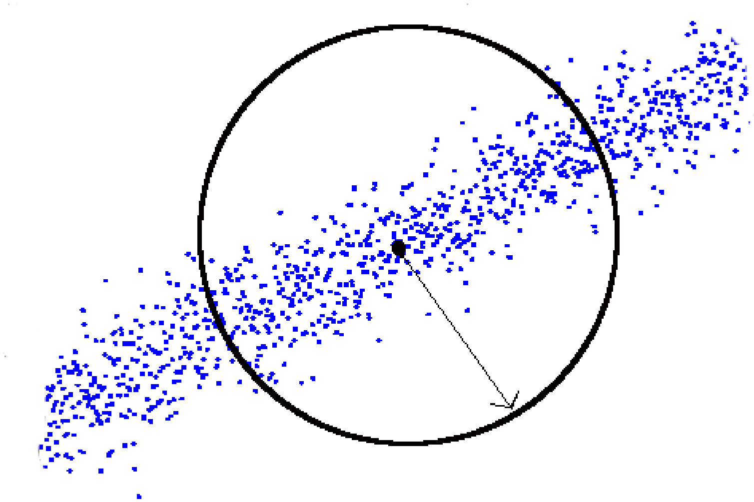
- example:
  - use the web, and collect 1,000,000 grey-scale images, each having 256 by 256 pixels.
  - each picture can be thought of as a point in 65,536 dimensional space ($256 \times 256 = 65536$).
  - you have 1,000,000 points in $\mathbb{R}^{65536}$.
- If this collection of points has *nice* geometric properties then this is useful. (For example, this makes *image recognition* easier).
- One reason to hope for this, is that not all pixel configurations appear in *natural images*.

Motivation

- It is relatively easy to collect large amounts of data.
- Data = a bunch of points $\subset \mathbb{R}^D$, with D being **large.**
- It is useful to learn what the geometry of this data is.
- High dimension $\implies$ hard to analyze.
    - a unit cube in $\mathbb{R}^{10}$ has 2^{10} disjoint sub-cubes of half the sidelength
    - because of this, many algorithms have a complexity (take a time) which grows **exponentially** with dimension.
    - this is often called *the curse of dimensionality*
- Dimensionality Reduction.
- Note: the Euclidean metric may not be the right one!

Some Assumtions

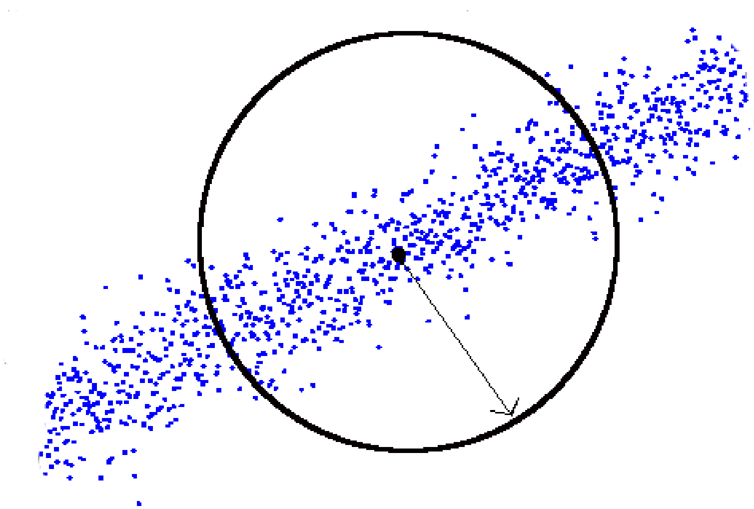
- Many data sets, while living in a high dimensional space, really exhibit low dimensional behavior.



- $\#(\mathrm{Ball}(x_i, r) \cap X) \sim r^m$ (in the picture, $m = 1$or $m = 2$, depending on scale).

The Main Point

- While D (ambient dimension) can be very large (say 50), m can often be very small (1,2,3,...).
- (Note that in different parts of that data, m can be different. Also, relevant r (scale) can be different.)
- For these sets of points we have more tools.
- We will focus on one of these tools.



Tool: Multiscale Geometry

- Use multiscale analysis. *Quantitative rectifiability*.
- Analyze the geometry on a coarse scale...
- ...and then refine over and over.
- Tools come from Harmonic Analysis and Geometric Measure Theory. They are used to keep track of what is happening.
- (the things I discuss are actually part of HA and GMT)

On route we discuss

- quantitative differentiation
- metric embedings
- TSP

Sample Questions:

- When is a set $K \subset \mathbb{R}^D$ contained inside a single connected set of finite length?
- Can we estimate the length of the shortest connected set containing K?
- What do these estimates depend on?
  - Number of points?
  - Ambient dimension (=D for $\mathbb{R}^D$) ?
- Can we build this connected set?
- Does this connected set form an *efficient network*. (Or, can it be made into one)

Related Questions:

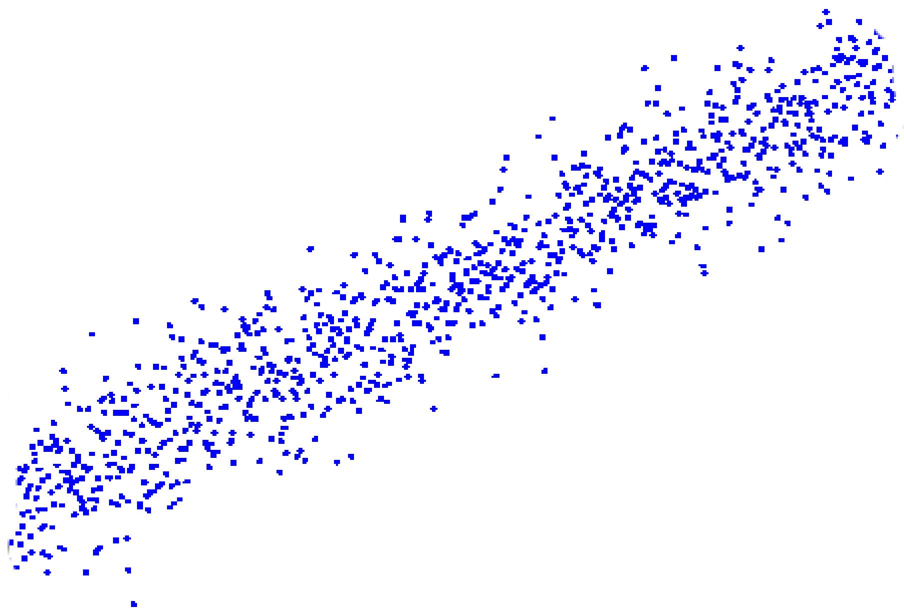
(which we will not discuss today)

- What is a good way to go beyond curves (Lipschitz or biLipschitz surfaces)
- the Traveling Bandit Problem (rob many banks with a car while traveling a short distance)

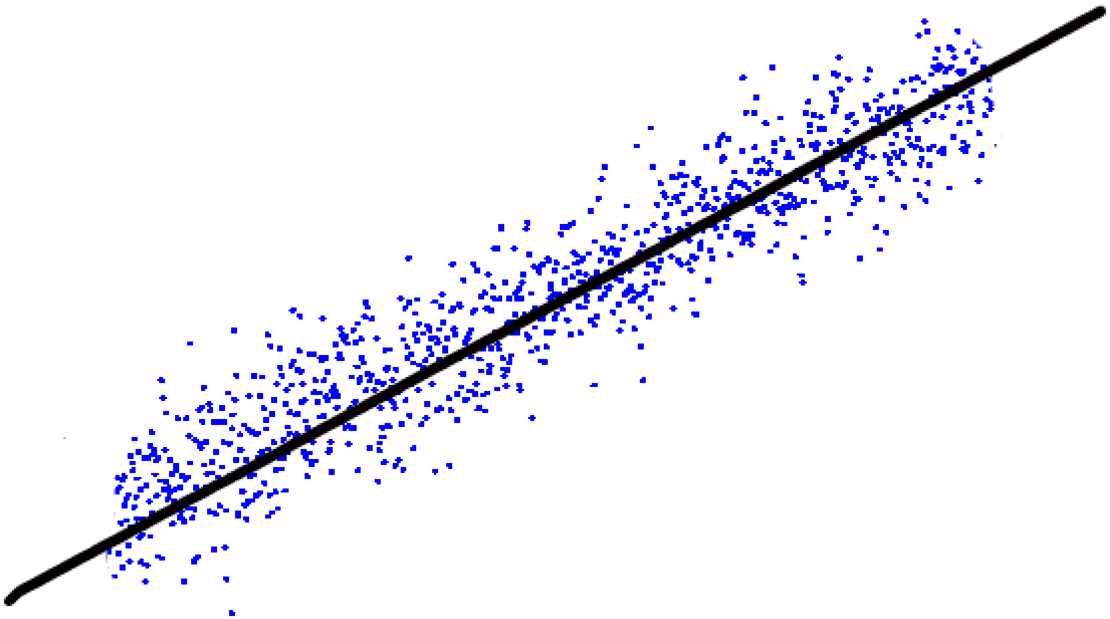
For now, we will discuss

curves, connected sets and efficient networks.

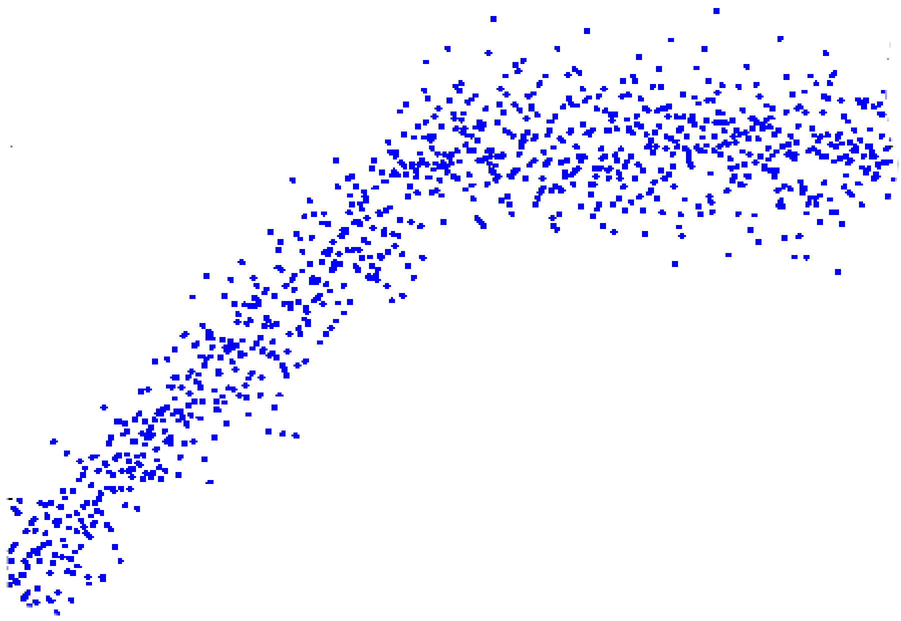
Motivation examples



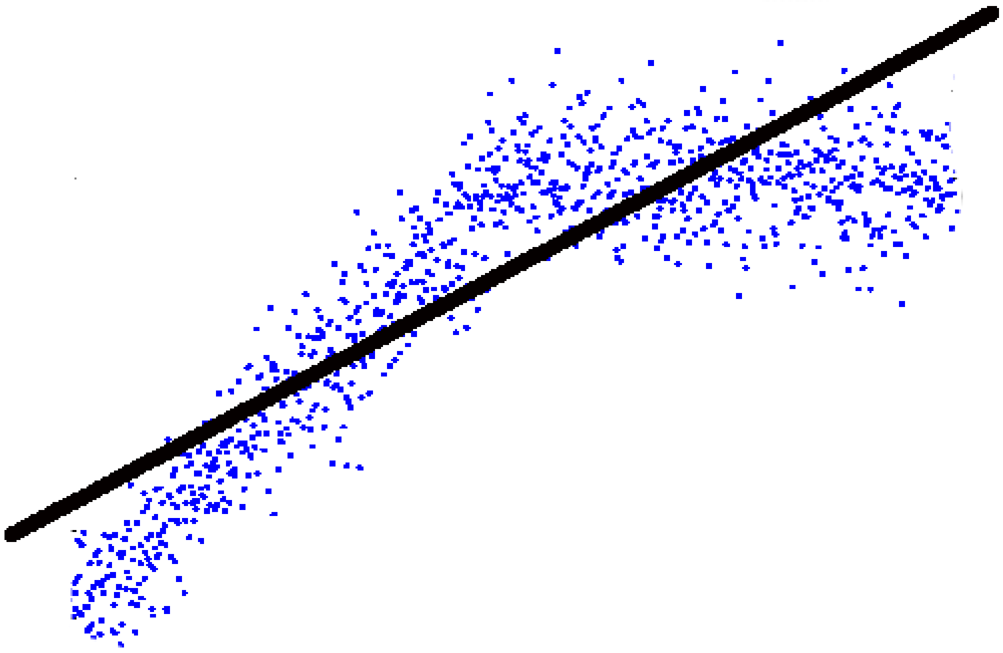
Motivation examples



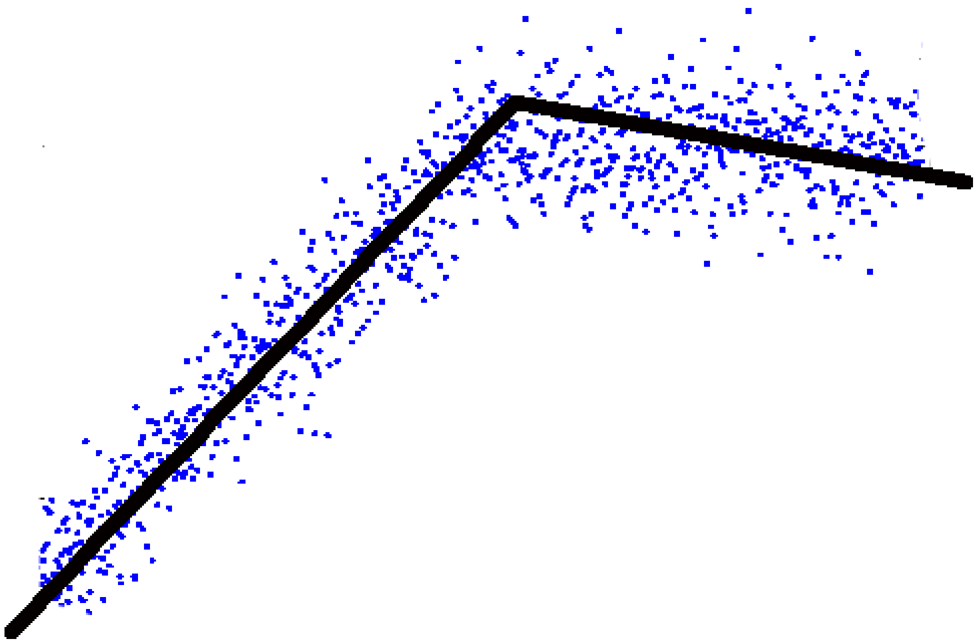
Motivation examples



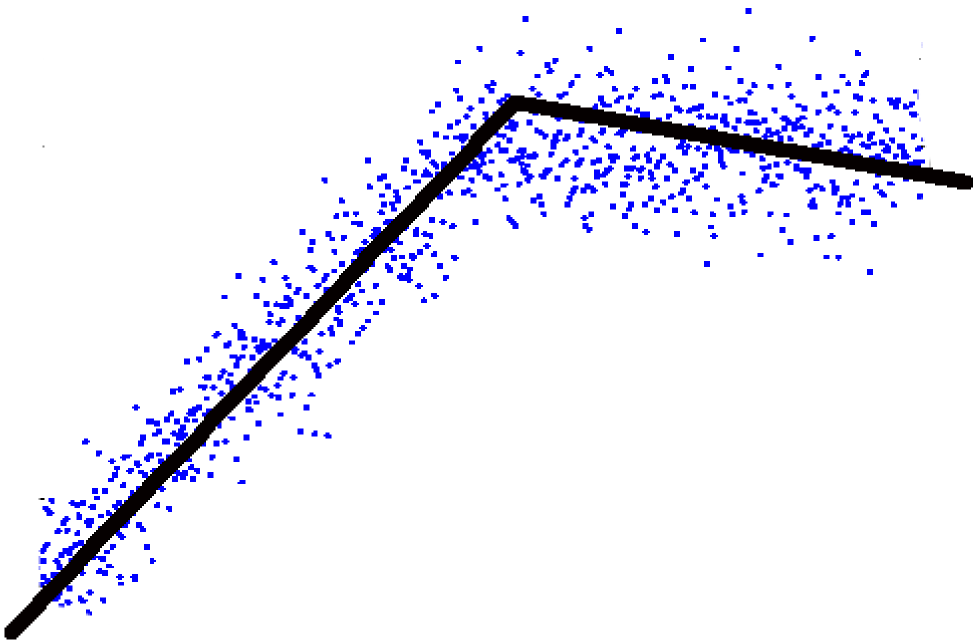
Motivation examples



Motivation examples

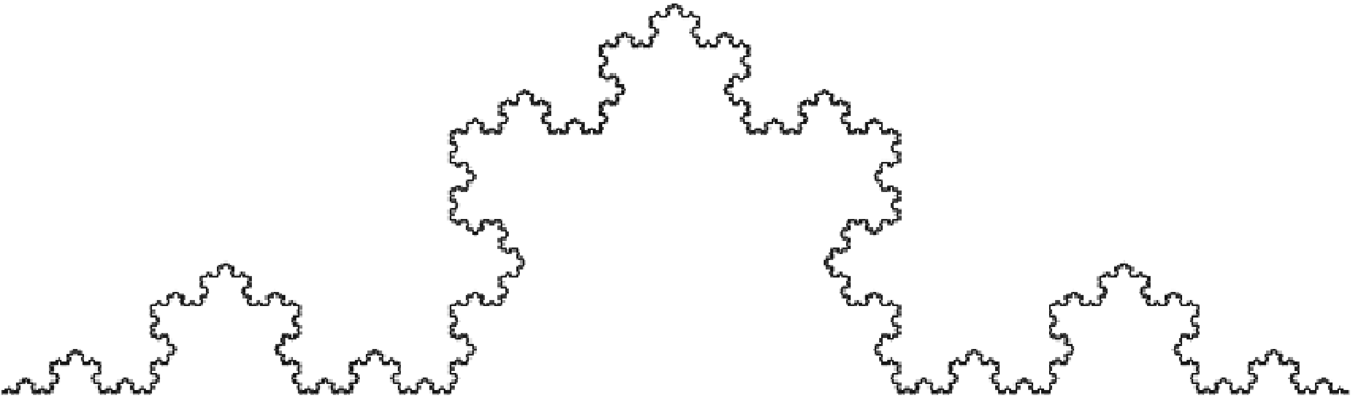


Motivation examples



How much did the length increase by?

Motivation examples



Motivation summery

- Approximating the geometry by a line is a way of reducing the dimension.
- This may not be good enough (even for 1-dim. data).
- Repeatedly refining this approximation may get closer.
- This process yields longer curves. (too long?)
- There is an interesting family of data sets where one can make quantitative mathematical statements about this. (And an extensive theory about them)

Quantitative Rectifiability

- Intuitive Picture:
  - A connected set (in $\mathbb{R}^D$) of finite length is 'flat' on most scales and in most locations.
  - This can be used to characterize subsets of finite length connected sets.
  - One can give a quantitative version of this using multiresolutional analysis.
  - This quantitative version also constructs the curve.
  - this quantity is also used to construct efficient networks

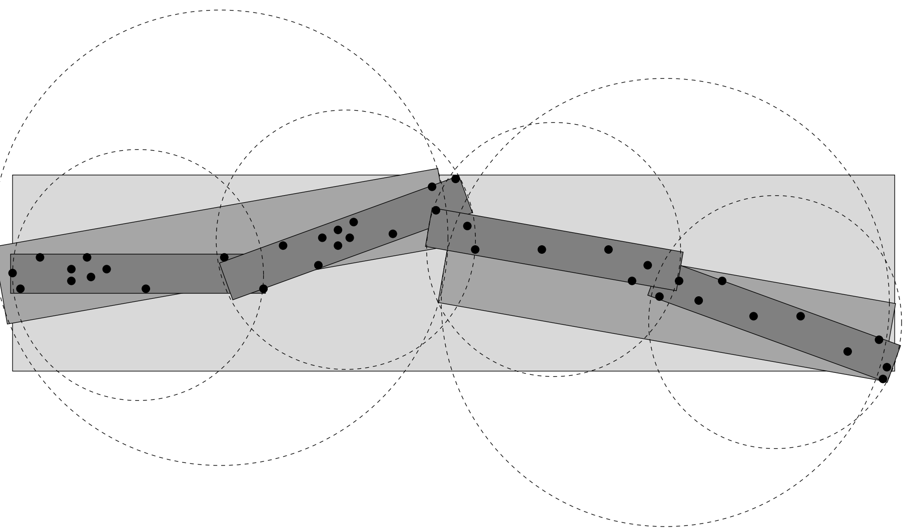
Efficient network

- Let $\Gamma \subset \mathbb{R}^D$ be a connected, finite length set (a road system)
- Define $\mathrm{dist}_\Gamma(x, y)$ as distance along the road system
- For $x, y \in \Gamma$, can we bound $\mathrm{dist}_\Gamma(x, y)$ in terms of $\mathrm{dist}_{\mathbb{R}^d}(x, y)$?
- in general, no... (think of a hair-pin turn)
- **Theorem** [Azzam - S.]: There is a constant $C = C(D)$ such that if we let $\Gamma \subset \mathbb{R}^D$ be a connected, then there exists $\tilde{\Gamma} \supset \Gamma$ such that for $x, y \in \tilde{\Gamma}$,
  - $\mathrm{dist}_{\tilde{\Gamma}}(x, y) \lesssim \mathrm{dist}_{\mathbb{R}^d}(x, y)$ and
  - $\ell(\tilde{\Gamma}) \lesssim \ell(\Gamma)$.
- note that x, y can be taken to be any two points in the **new** road system $\tilde{\Gamma}$

A notion of curvature

Definition: (Jones β number)

$$\beta_K(Q) = \frac{1}{\mathrm{diam}(Q)} \inf_{L \text{ line}} \sup_{x \in K\cap Q} \mathrm{dist}(x, L)$$
$$= \frac{\text{radius of the thinest tube containing } K\cap Q}{\mathrm{diam}(Q)}.$$



Quantitative Rectifiability

- **Theorem 1:**[P. Jones D=2, K. Okikiolu D>2]
  For any connected $\Gamma \subset \mathbb{R}^D$

$$\text{“Total Multiscale Curvature”}(\Gamma) \qquad \sum_{Q \in \text{dyadic grid}} \beta_\Gamma^2(3Q)\,\mathrm{diam}(Q) \lesssim \ell(\Gamma)$$

- **Theorem 2:**[P. Jones] *For any set* $K \subset \mathbb{R}^D$*, there exists*
  $\Gamma_0 \supset K$,
  Γ_0 *connected, such that*

$$\ell(\Gamma_0) \lesssim \text{“Total Multiscale Curvature”}(K) + \mathrm{diam}(K)$$

$$\sum_{Q \in \text{dyadic grid}} \beta_K^2(3Q)\,\mathrm{diam}(Q)$$

Corollary:

- *For any connected set* $\Gamma \subset \mathbb{R}^D$

$$\mathrm{diam}(\Gamma) + \text{“Total Multiscale Curvature”}(\Gamma) \sim \ell(\Gamma)$$

$$\sum_{Q \in \textit{dyadic grid}} \beta_\Gamma^2(3Q)\,\mathrm{diam}(Q)$$

More generally:

- *For any set* $K \subset \mathbb{R}^D$

$$\mathrm{diam}(K) + \text{“Total Multiscale Curvature”}(K) \sim \ell(\Gamma_{MST})$$

where Γ_{MST} *is the shortest curve containing* K.

$$\sum_{Q \in \textit{dyadic grid}} \beta_K^2(3Q)\mathrm{diam}(Q)$$

- This solves the problem in $\mathbb{R}^D$ of how to parameterize data by a curve.

Two words about why we care

- After all, one can construct $\Gamma \supset K$ with a greedy algorithm
- This coarse version of curvature (β numbers) can be used (was used!) to understand the behavior of various mathematical objects.
- One example of how this can be useful which is very geometric: the “shortcuts" or “bridges" that were added when we turned a network into an ‘efficient’ one, were constructed based on a certain stopping rule which summed up β numbers.

Hilbert Space

Thm 1: $\forall$ connected $\Gamma \subset \mathbb{R}^d$

$$\sum_Q \beta_\Gamma^2(3Q)\mathrm{diam}(Q) \lesssim \ell(\Gamma)$$

Thm 2: $\forall K \subset \mathbb{R}^d$, $\exists$ connected $\Gamma_0 \supset K$, s.t.

$$\ell(\Gamma_0) \lesssim \mathrm{diam}(K) + \sum_Q \beta_K^2(3Q)\mathrm{diam}(Q)$$

- *"Theorem" :*

  One can reformulate theorems 1 and 2 in a way which will give constants **independent of dimension**
- (Actually, reformulated theorems are true for Γ or K in **Hilbert space**).
- Many properties of the dyadic grid are used in Jones' and Okikiolu's proofs, but in order to go to Hilbert space one needs to give them up and change to a different multiresolution.

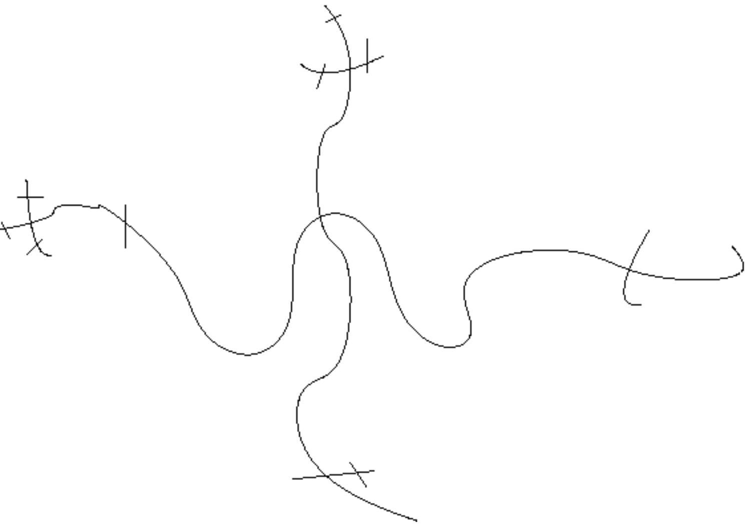
Definitions

- let $K \subset \mathbb{R}^D$ be a subset with $\mathrm{diam}(K) = 1$.
- $X_n \subset K$ is 2^{-n} net for K means
  - $x, y \in X_n$ then $\mathrm{dist}(x, y) \geq 2^{-n}$
  - For any $y \in K$, exists an $x \in X_n$ with $\mathrm{dist}(x, y) < 2^{-n}$
- Take $X_n \subset K$ a 2^{-n} net for K, with $X_n \supset X_{n-1}$
- Define the multiresolution

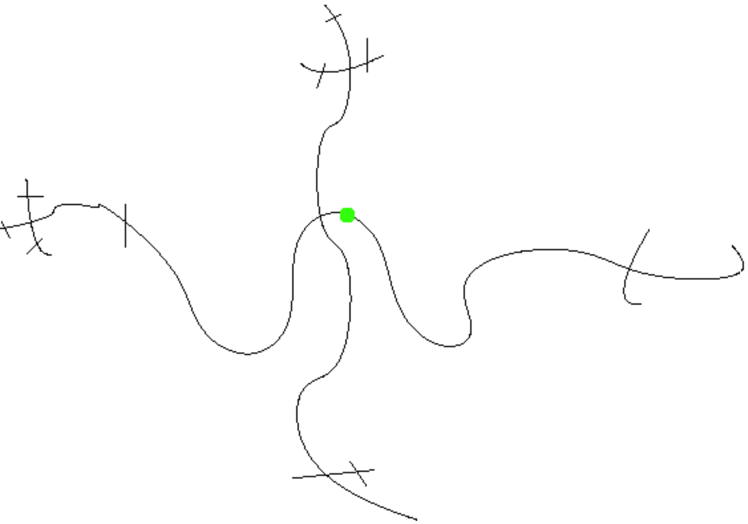
$$\mathcal{G}^K = \{B(x, A2^{-n}) : x \in X_n; n \geq 0\}$$

- $\mathcal{G}^K$ replaces the dyadic grid

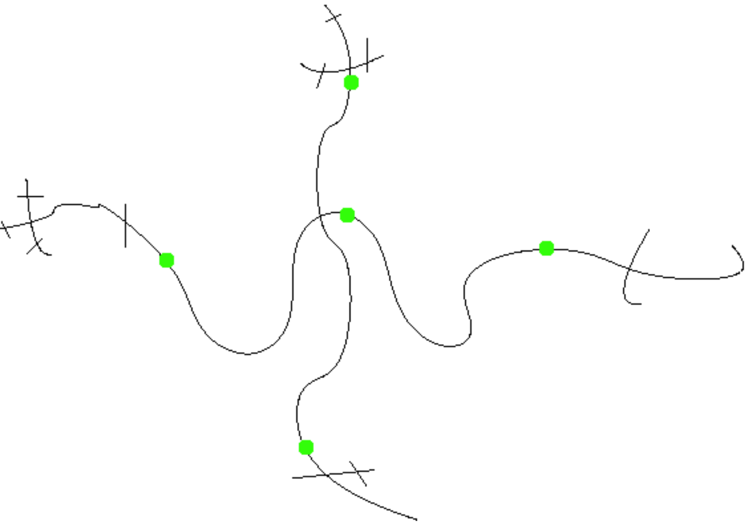
K



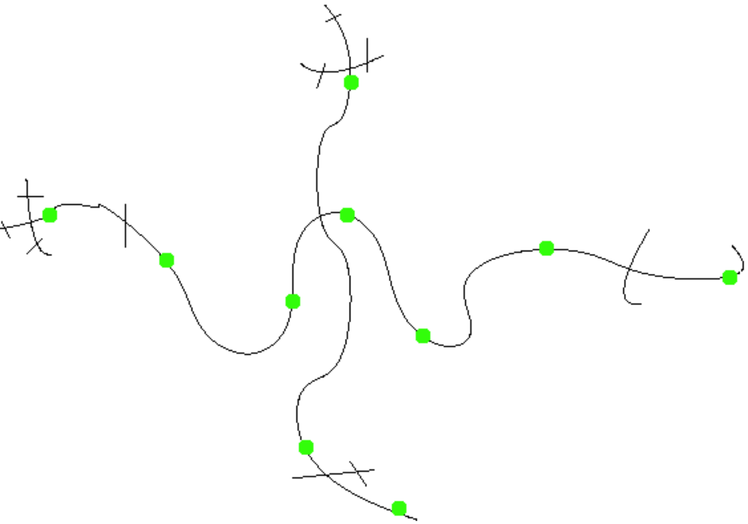
K and X_0



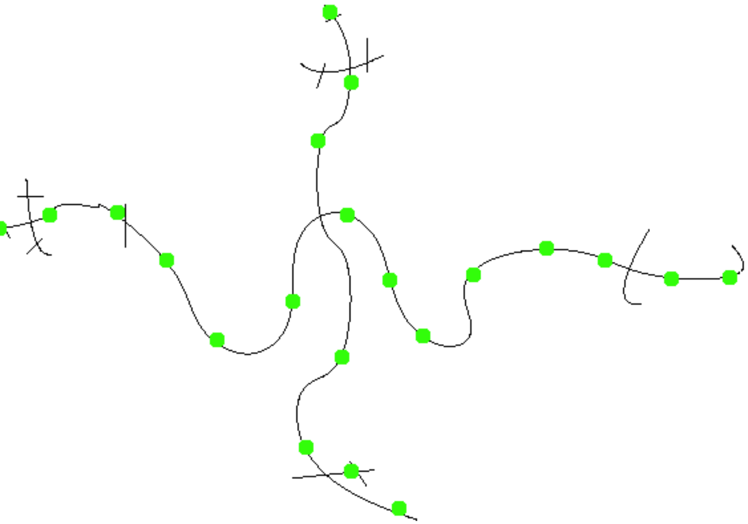
K and X_1



K and X_2



K and X_3



Hilbert Space

- Constants that make inequalities true are **independent of dimension** D (Theorems hold in Hilbert Spaces.)

- **Theorem 1':(S.)** *For any connected* $\Gamma \subset H, \Gamma \supset K$

  "Total Multiscale Curvature" (Γ) → $$\sum_{Q \in \mathcal{G}^K} \beta_\Gamma^2(Q)\mathrm{diam}(Q) \lesssim \ell(\Gamma)$$

- **Theorem 2':(S.)** *For any set* $K \subset H$*, there exists* $\Gamma_0 \supset K$*,* Γ_0 *connected, such that*

$$\ell(\Gamma_0) \lesssim \text{"Total Multiscale Curvature"}(K) + \mathrm{diam}(K)$$

"Total Multiscale Curvature" ← $$\sum_{Q \in \mathcal{G}^K} \beta_K^2(Q)\mathrm{diam}(Q)$$

Hilbert Space

- Corollary:
  - For any set $K \subset$ Hilbert Space

$$\mathrm{diam}(K) + \text{“Total Multiscale Curvature”}(K) \sim \ell(\Gamma_{MST})$$

where Γ_{MST} is the shortest curve containing K.

  - This solves the problem in Hilbert space of how to parameterize data by a curve.

Non-parametric vs. parametric

- Non-Parametric: you are given data, and you know (or hope) that a curve can go through it, but you do not know how to draw such a curve
- Parametric: You are given such a curve (and your data is then the image of the curve)
- 1-dim case: curves and connected sets of finite length. Go back and forth between the param. and non-param.:
  - parametric $\to$ non-parametric:
    $f : [0,1] \to \mathbb{R}^D$ is given , so consider the image, $f[0,1]$.
  - non-parametric $\to$ parametric:
    Given Γ, construct $f : [0,1] \to \mathbb{R}^D$ such that $\Gamma = f[0,1]$.
    You can do so with $\|f\|_{Lip} \lesssim \ell(\Gamma)$.

continued

- - non-parametric $\rightarrow$ parametric:
    Given Γ, construct $f : [0,1] \rightarrow \mathbb{R}^D$ such that
    $\Gamma = f[0,1]$.
    You can do so with $\|f\|_{Lip} \lesssim \ell(\Gamma)$.
- As said before, you don't need much to do this (e.g. greedy algorithm).
- Keeping track of β numbers helps you do other things like add shortcuts in the "efficient network" result)
- β numbers are an analogue to wavelet coefficients. They allow analysis of a set.

Some obvious questions

- Can you have this discussion about sets of higher intrinsic dimension?
- You have parametrized using Lipschitz curves. Isn't bi-Lipschitz curves a more natural category? Can you say something about that?

The answer to all of the above questions is yes.

Lip vs biLip

- **Theorem**[Jones, David, S.] Let $\delta > 0$ and $n \geq 1$ be given. There constants $M = M(\delta, n)$, and $c = c(n)$ such that if $\mathcal{M}$ is a metric space and $f : [0,1]^n \to \mathcal{M}$ is a 1-Lipschitz function satisfying $\mathcal{H}^n_\infty(f[0,1]^n) \geq \delta$, then there is a set $E \subset [0,1]^n$ such that the following hold
  - $\mathcal{H}^n(E) > \frac{\delta}{M}$
  - for all $x, y \in E$ we have

$$c\delta|x-y| < \mathrm{dist}(f(x), f(y)) < |x-y|$$

Notes

- Jones, David (80's): $\mathcal{M} = \mathbb{R}^D$.
- S.: $\mathcal{M}$ metric space (faking wavelet coeficients!!)
- $\mathcal{H}^n_\infty(K) = \inf\{\sum \mathrm{diam}(B_i)^n : \cup B_i \supset K\}$

References

[1] J. Azzam and R. Schul. *How to take shortcuts in Euclidean space: making a given set into a short quasi-convex set.* Proc. London Math. Soc. (2012) `arxiv:0912.1356`.

[2] G. David, *Morceaux de graphes lipschitziens et integrales singulières sur une surface.*, Revista matemática iberoamericana **4** (1988), no. 1, 73.

[3] G. David and S. Semmes, *Analysis of and on uniformly rectifiable sets.* Mathematical Surveys and Monographs, **38**. American Mathematical Society, Providence, RI, 1993.

[4] P. W. Jones, *Lipschitz and bi-Lipschitz functions*, Rev. Mat. Iberoamericana **4** (1988), no. 1, 115–121.

[5] K. Okikiolu, *Characterization of subsets of rectifiable curves in* $\mathbf{R}^n$, Journal of the London Mathematical Society, **2**, 46(2):336–348, 1992.

[6] R. Schul, *Bi-Lipschitz decomposition of Lipschitz functions into a metric space*, Rev. Mat. Iberoam. **25** (2009), no. 2, 521–531.

[7] R. Schul. Analyst's traveling salesman theorems. A survey. *In the tradition of Ahlfors and Bers, IV*, volume 432 of *Contemp. Math.*, pages 209–220. Amer. Math. Soc., Providence, RI, 2007.

[8] R. Schul. Subsets of rectifiable curves in Hilbert space. *Journal d'Analyse Mathématique* 103 (2007), 331-375..