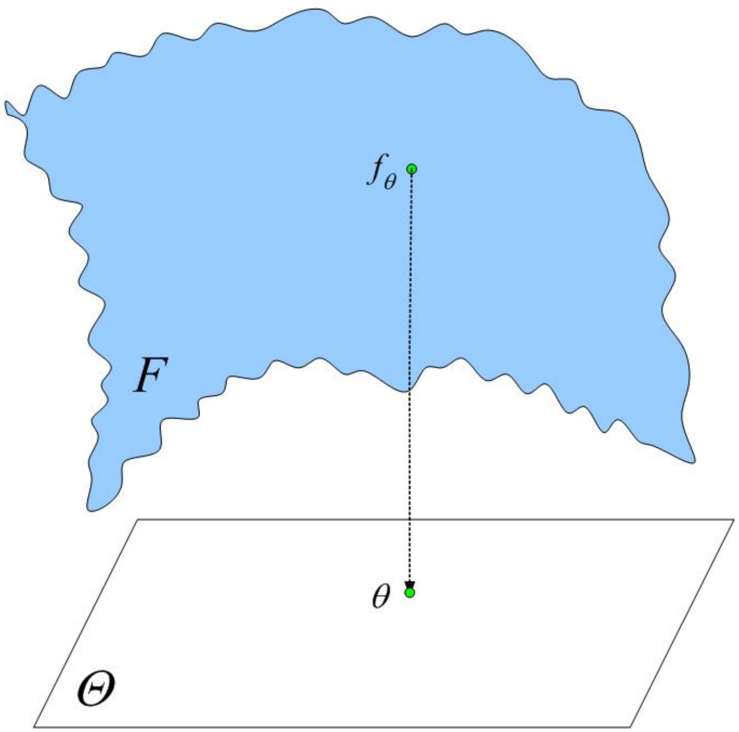
Near-Isometric Linear Embeddings of Manifolds

Chinmay Hegde
Aswin Sankaranarayanan
Richard Baraniuk

RICE UNIVERSITY



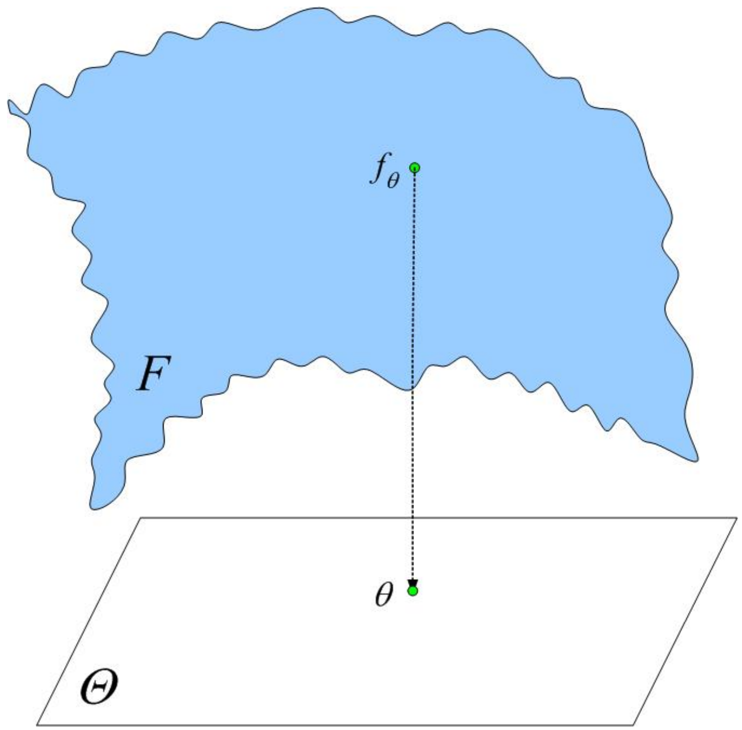
Too Much Information

The Data Deluge

- How to sense?
- How to compress?
- How to process?

The Data Deluge

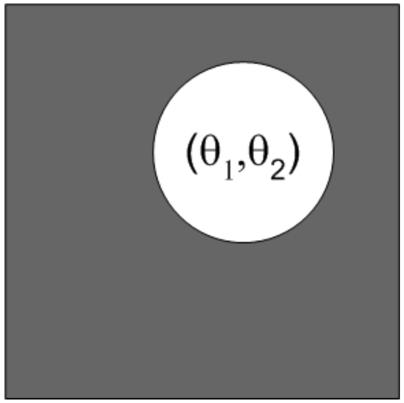
- How to sense?
- How to compress?
- How to process?



- *Answer*: Concise models for the data

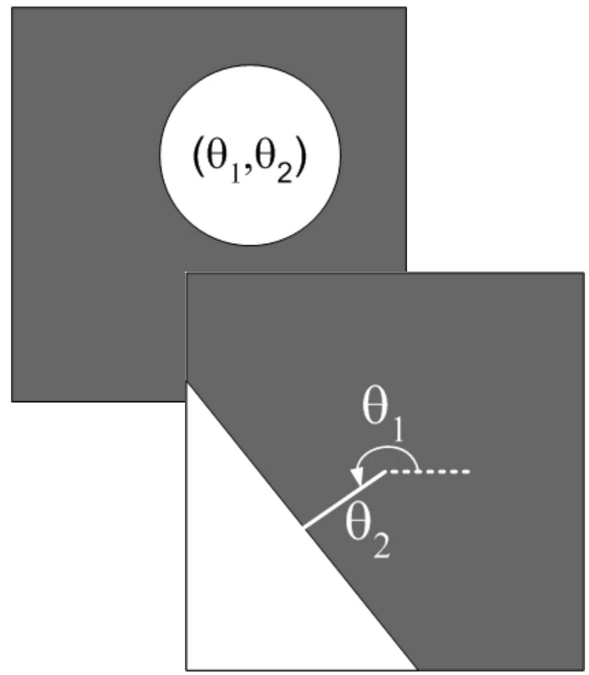
Concise Models

- Our focus: **Collections** of "images" parameterized by $\theta \in \Theta$
  - translations of an object
    θ: x-offset and y-offset



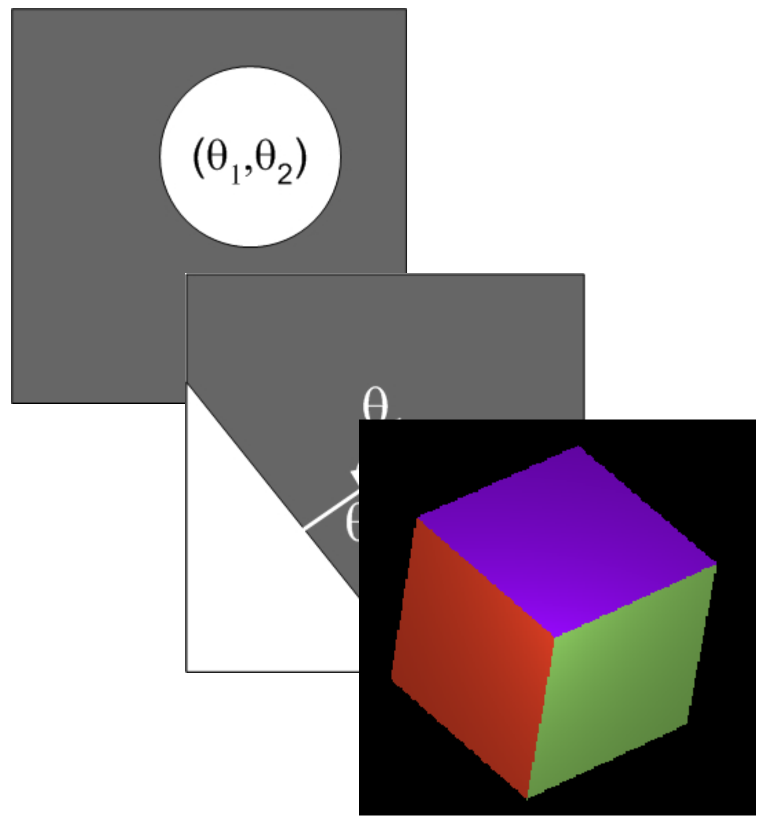
Concise Models

- Our focus: **Collections** of images parameterized by $\theta \in \Theta$
  - translations of an object
    θ: x-offset and y-offset
  - wedgelets
    θ: orientation and offset



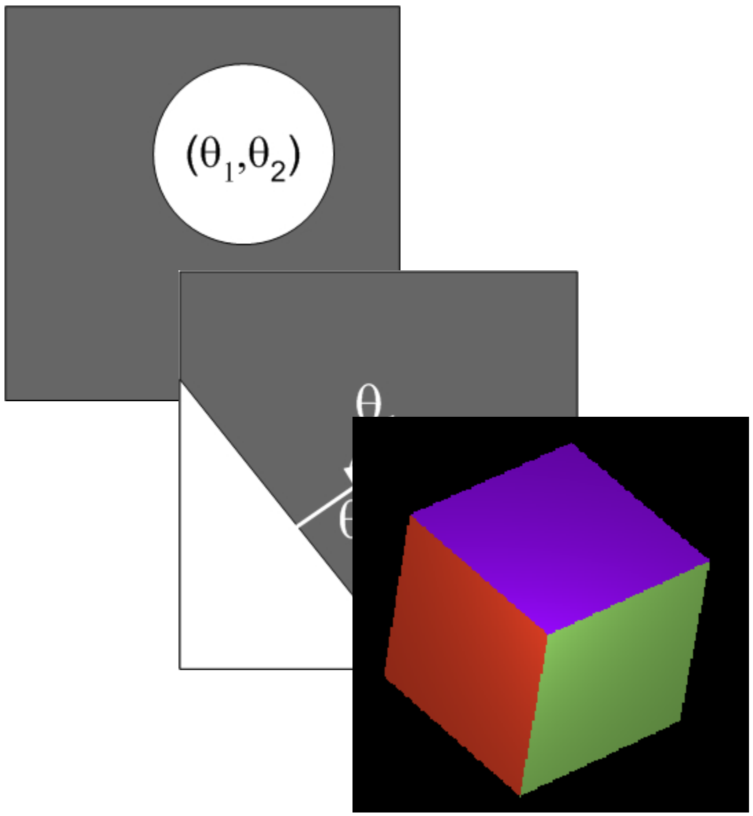
Concise Models

- Our focus: **Collections** of images parameterized by $\theta \in \Theta$
  - translations of an object
    θ: x-offset and y-offset
  - wedgelets
    θ: orientation and offset
  - rotations of a 3D object
    θ: pitch, roll, yaw



Concise Models

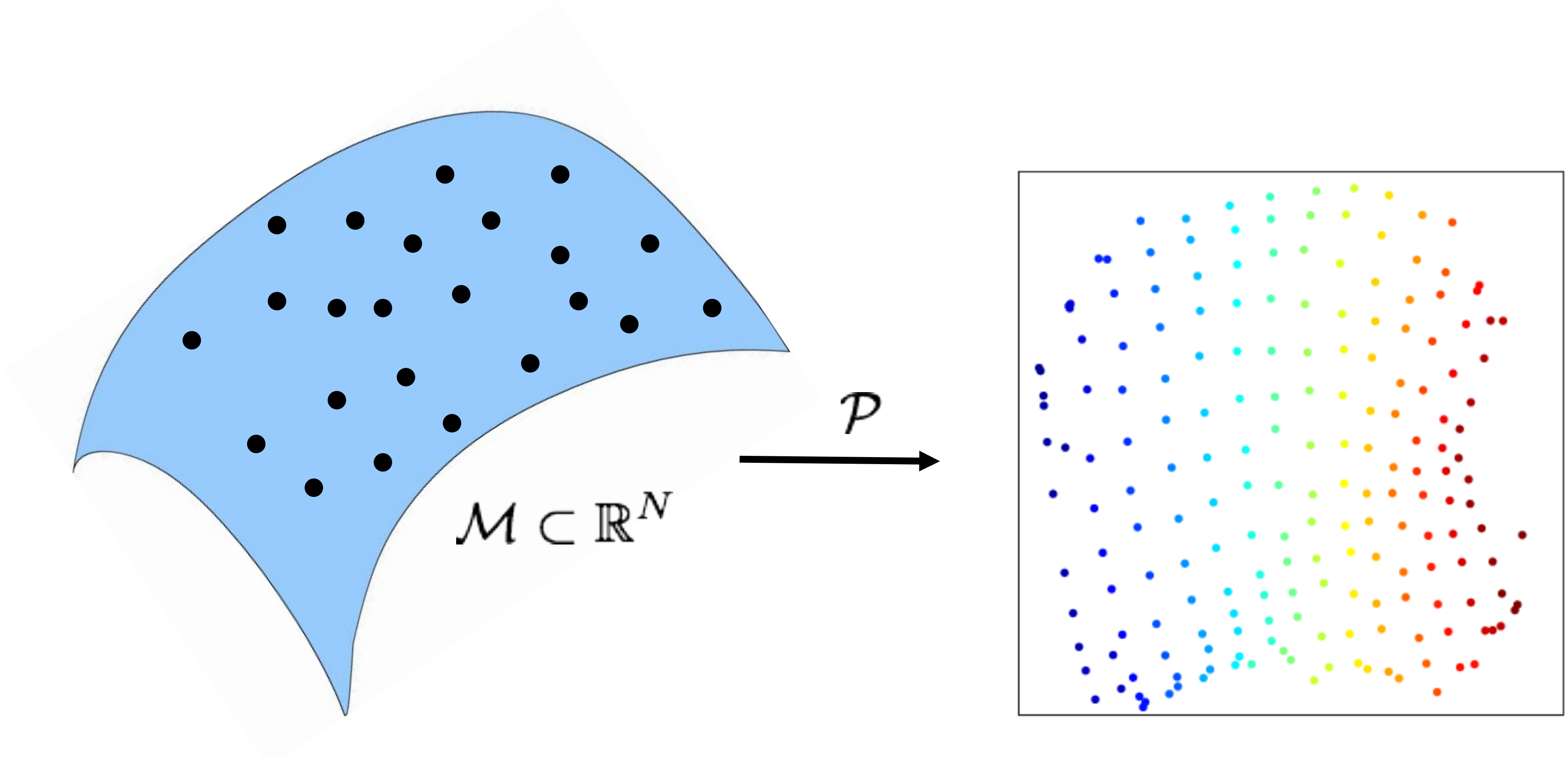
- Our focus: **Collections** of images parameterized by $\theta \in \Theta$
  - translations of an object
    θ: x-offset and y-offset
  - wedgelets
    θ: orientation and offset
  - rotations of a 3D object
    θ: pitch, roll, yaw



- **Image manifold**

$$\mathcal{M} = \{I_\theta \ : \ \theta \in \Theta\}$$

Dimensionality Reduction



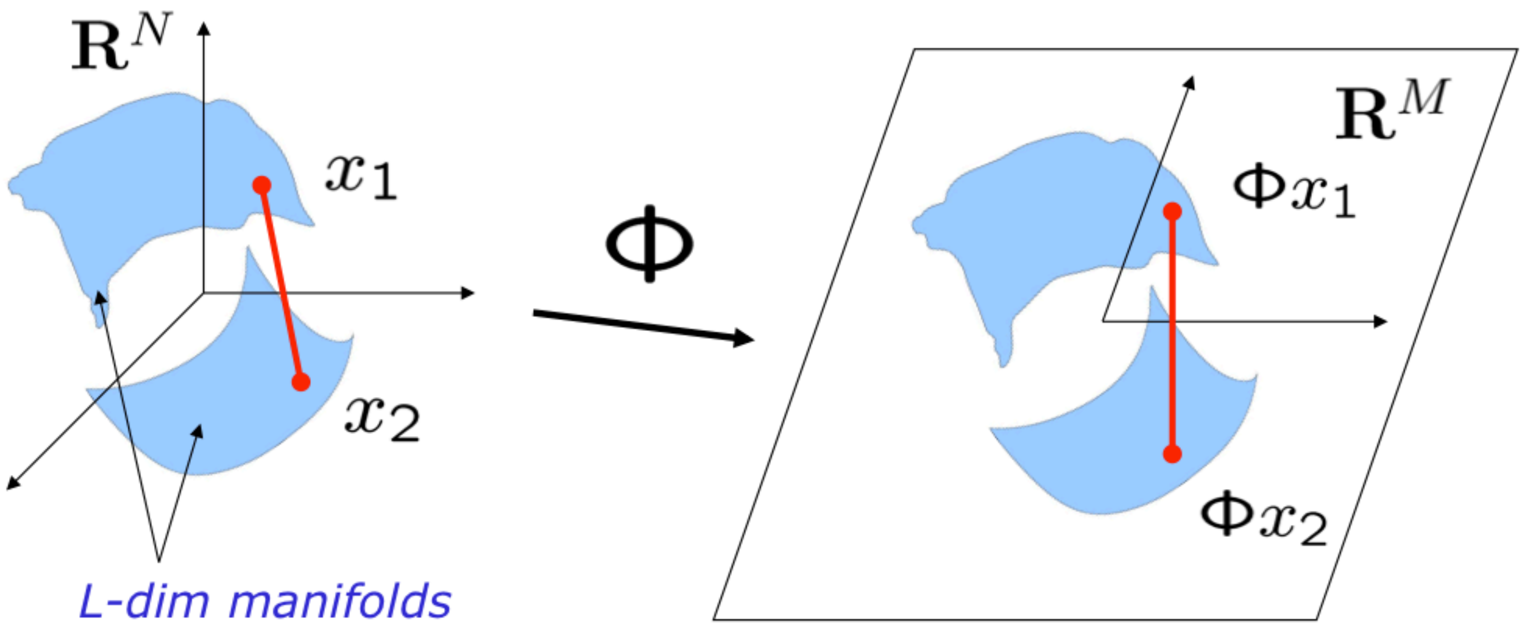
"Manifold Learning" - Isomap, LLE, MVU, ...

- Obtain a low-dimensional representation of the data
- Preserve geometric information

Dimensionality Reduction

- Despite their great promise, nonlinear methods suffer from drawbacks:
  - (often) do not generalize to out-of-sample points
  - (often) unstable
  - nonlinear methods not easy to implement as sensing schemes

Linear Dim. Reduction

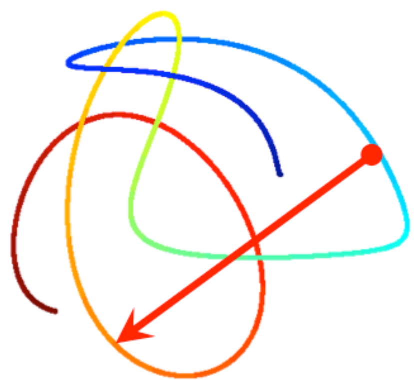


- **Want**: a linear map that preserves model geometry
- Q. Is it even possible? Does such a map exist?

Whitney's Theorem

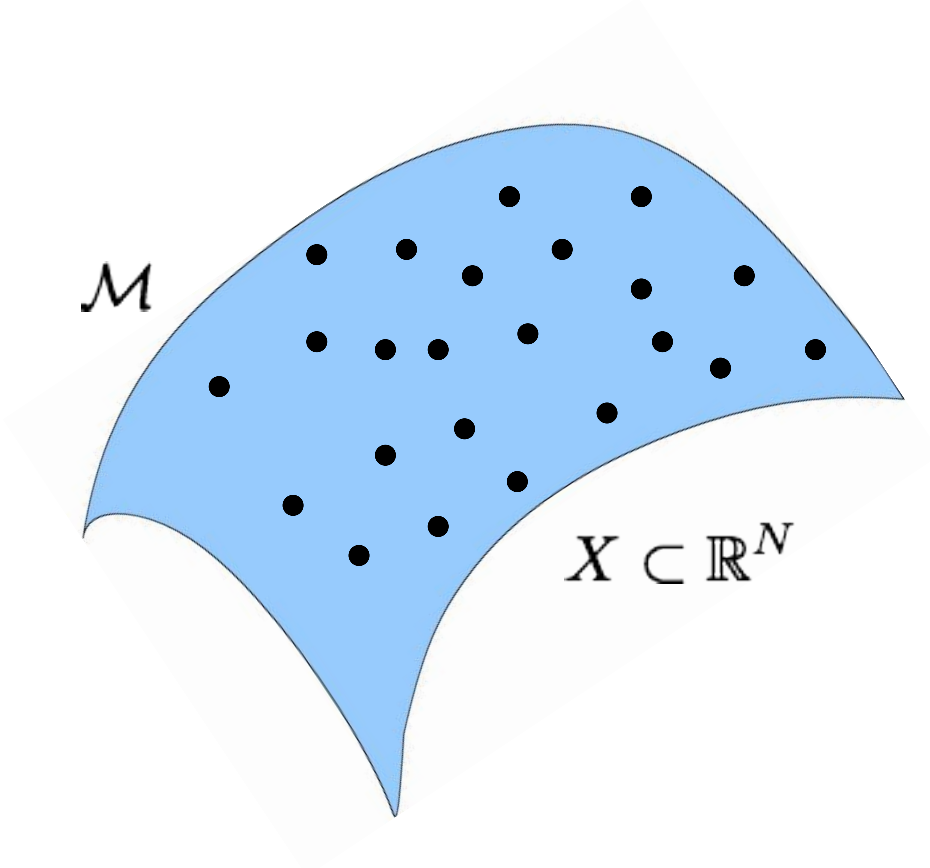
- [W,1936] Let $\mathcal{M} \subset \mathbf{R}^N$ be a compact, smooth, K-dimensional manifold. Then, there exists a smooth embedding of $\mathcal{M}$ into $\mathbf{R}^{2K+1}$.

$$\langle \phi_i, x - x' \rangle \neq 0$$

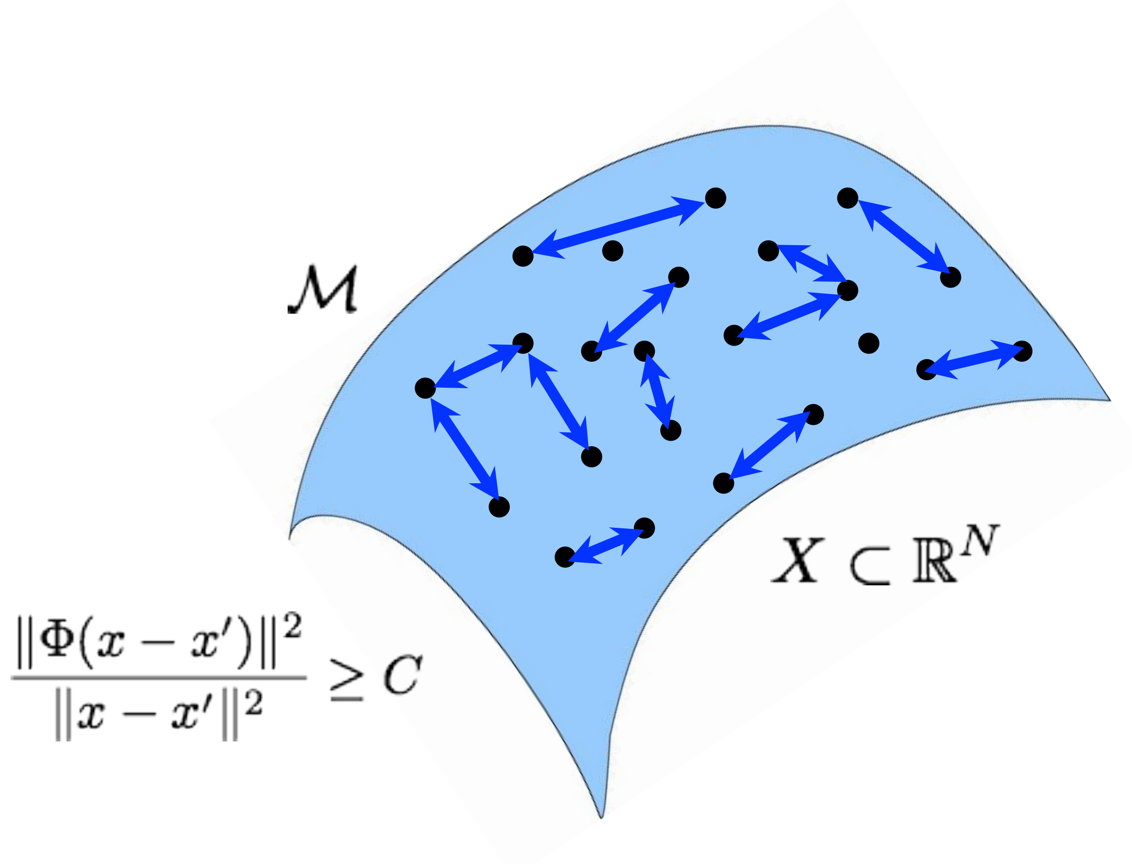


  - Key insight in proof: consider set of **manifold secant directions**. Each secant is a direction *not* to project.

“Good” Linear Maps



"Good" Linear Maps



- Goal: preserve norms of all pairwise secants of X
- If X is a dense enough sampling, then Φ is an "good" mapping for the entire manifold $\mathcal{M}$

Linear Dim. Reduction Methods

- Principal Components Analysis (PCA)
  - Easy to compute

$$X = USV^T$$

  - But **distorts** pairwise distances

- Random Projections
  - Guarantees pairwise distance preservation (JL '84)

$$1 - \delta \leq \frac{\|\Phi(x_1 - x_2)\|^2}{\|x_1 - x_2\|^2} \leq 1 + \delta$$

  - But constants are **poor**
  - Oblivious to structure of data

Designing a "Good" Linear Map

Want: a short, fat matrix Φ , such that

$$1 - \delta \leq \|\Phi v_i\|^2 \leq 1 + \delta$$
$$i = 1, 2, \ldots, Q$$

$$\Updownarrow$$

$$\text{minimize } \operatorname{rank}(\Phi), \quad \text{subject to}$$

$$\left| \|\Phi v_i\|_2^2 - 1 \right| \leq \delta$$
$$i = 1, 2, \ldots, Q$$

Designing a "Good" Linear Map

- Convert quadratic constraints in Φ into *linear* constraints in $P = \Phi^T \Phi$

- Use a nuclear-norm relaxation of the rank

- Simplified problem:

$$\begin{array}{c} \text{minimize } \operatorname{rank}(\Phi) \\ \left| \|\Phi v_i\|_2^2 - 1 \right| \leq \delta \\ i = 1, 2, \ldots, Q \end{array} \quad \Leftrightarrow \quad \begin{array}{c} \text{minimize } \|P\|_* \\ \|\mathcal{A}(P) - \mathbf{1}\|_\infty \leq \delta \\ P \succ 0,\ P = P^T \end{array}$$

[HSB12]

Task Adaptivity

Can prune the secants according to **task**

$$\mathcal{A} : P \mapsto \{v_i^T P v_i\}_{i=1}^{Q}$$

- If goal is classification, preserve norms of only inter-class secants
- If goal is reconstruction / parameter estimation, preserve both inter- and intra-class secants

- Can preferentially weight secants according to importance (akin to *boosting*)

A Fast Algorithm

- Practical considerations: *Q* very large, *N* very large

$$\text{minimize } \|P\|_*$$
$$\|\mathcal{A}(P) - \mathbf{1}\|_\infty \leq \delta$$

- Alternating Direction Method of Multipliers (**ADMM**)

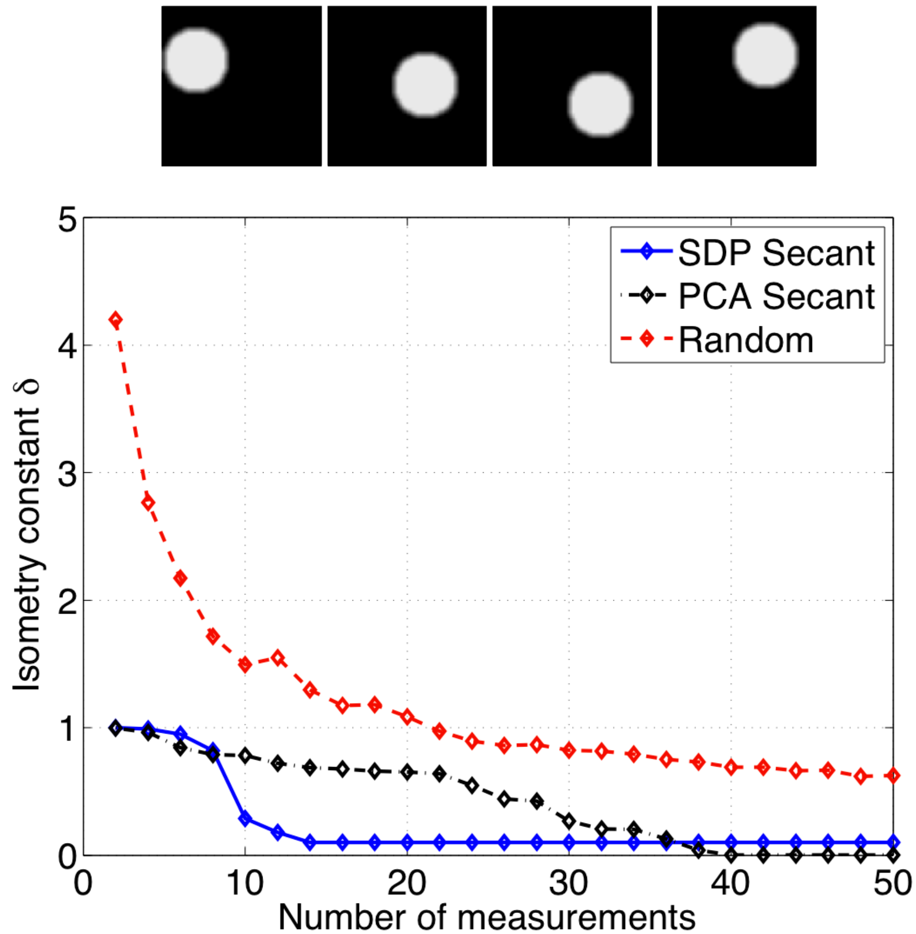
$$\text{minimize } \|P\|_*$$
$$P = L,\ \mathcal{A}(L) = q,\ \|q - \mathbf{1}\|_\infty \leq \delta$$

Iterate:

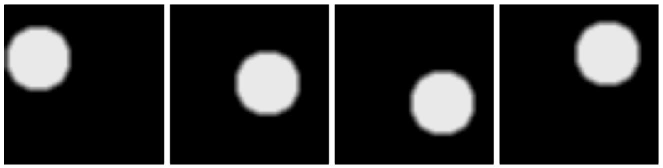
- solve for *P* using SVD + thresholding
- solve for *L* using least-squares
- solve for *q* using a truncation step

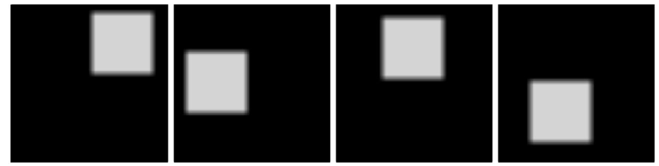
Isometry constants

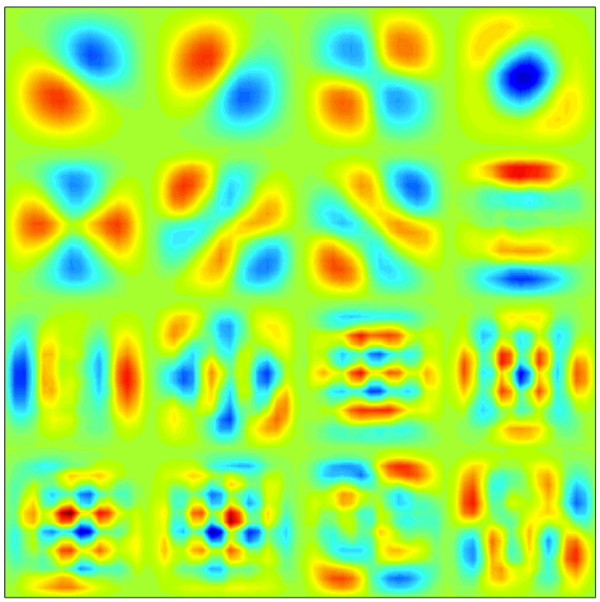
- Training data: Q = 900 test secants, 1000 test secants; measure worst case distortion in norms

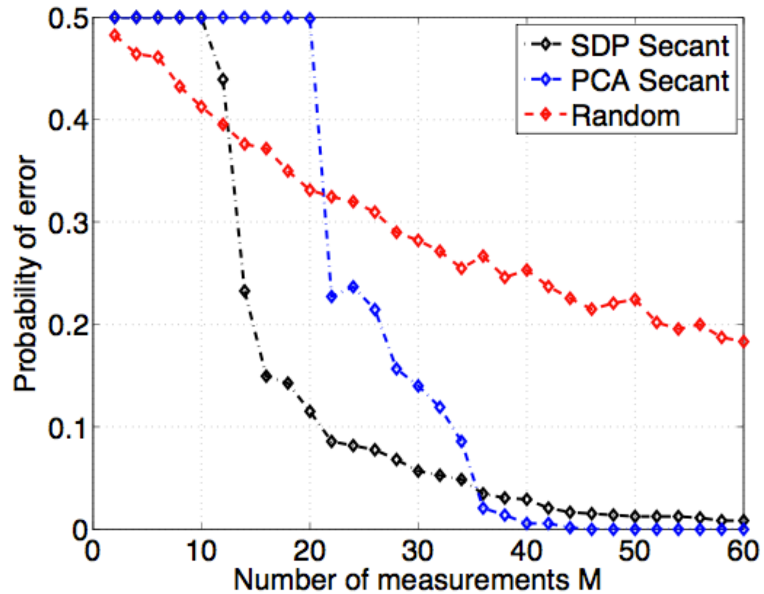


Circles and Squares

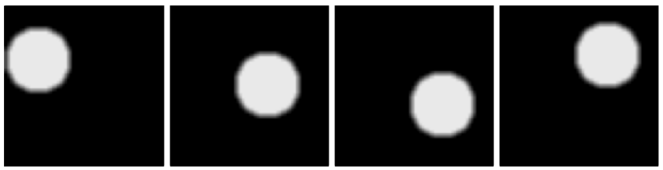


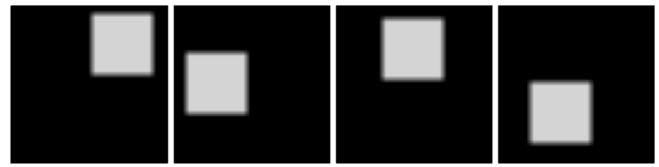


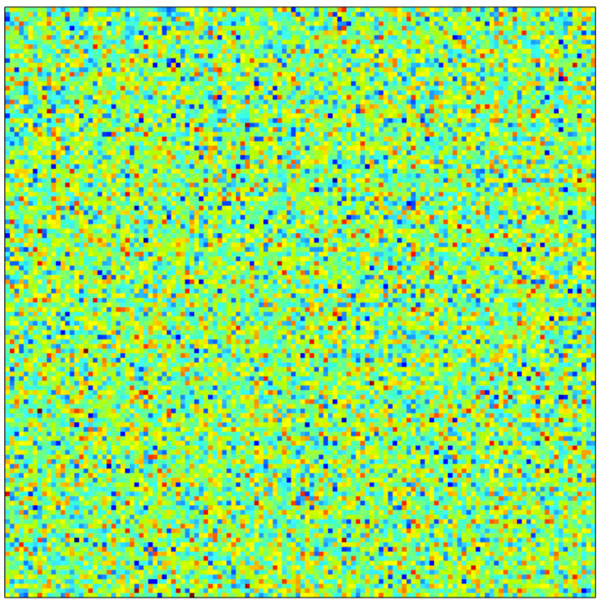


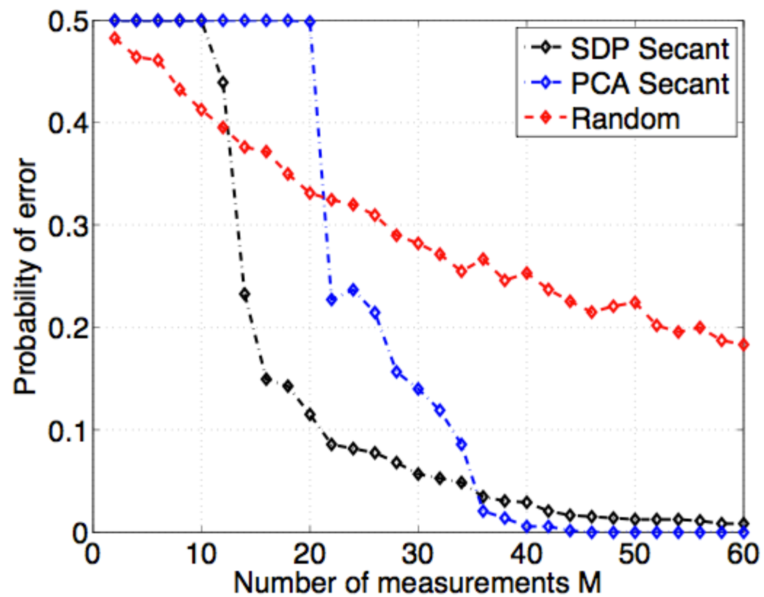


Circles and Squares

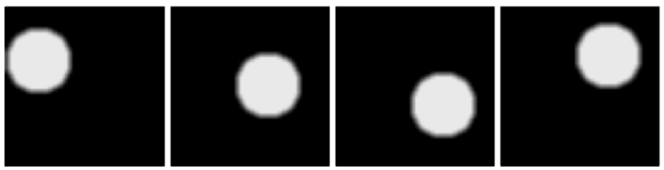


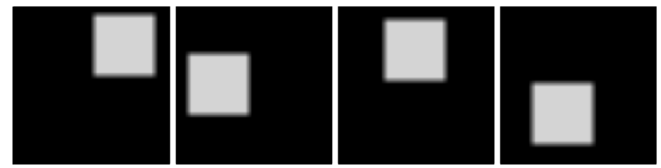


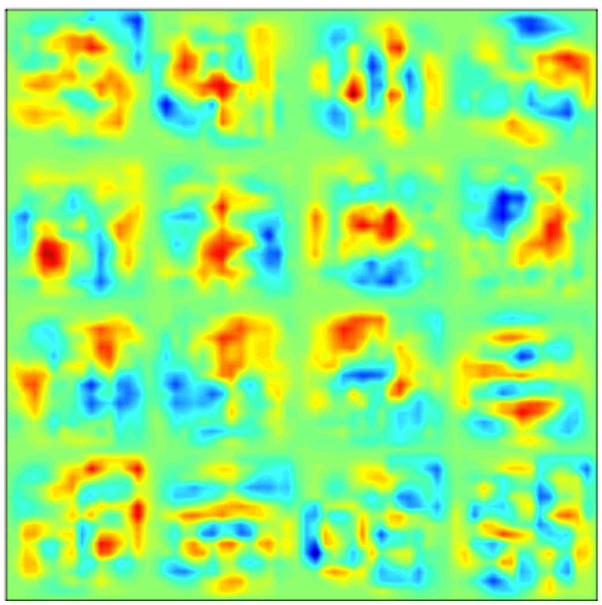


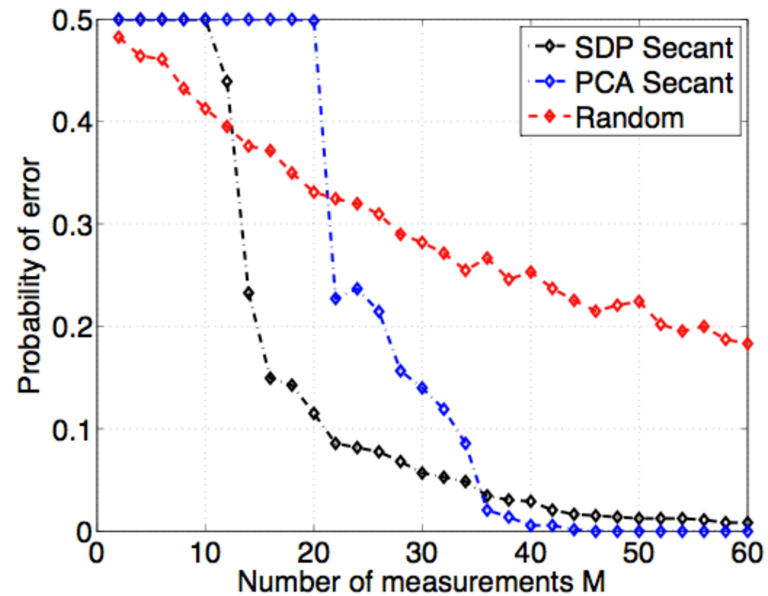


Circles and Squares



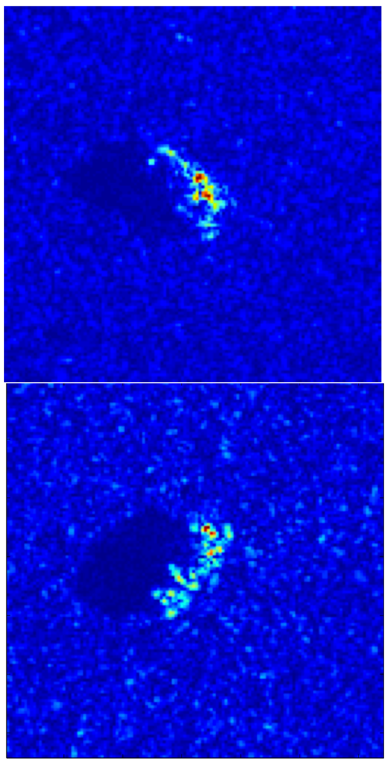


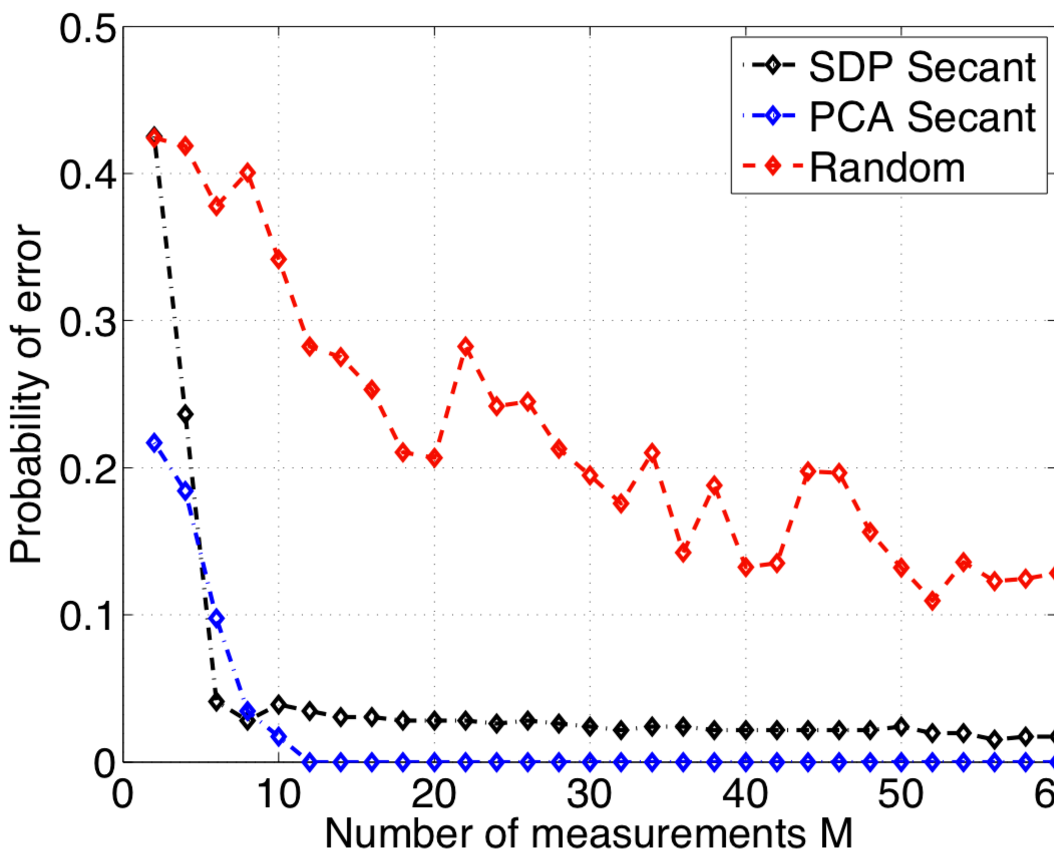




MSTAR

- Training data: 230 radar images per class, test data: 180 radar images per class



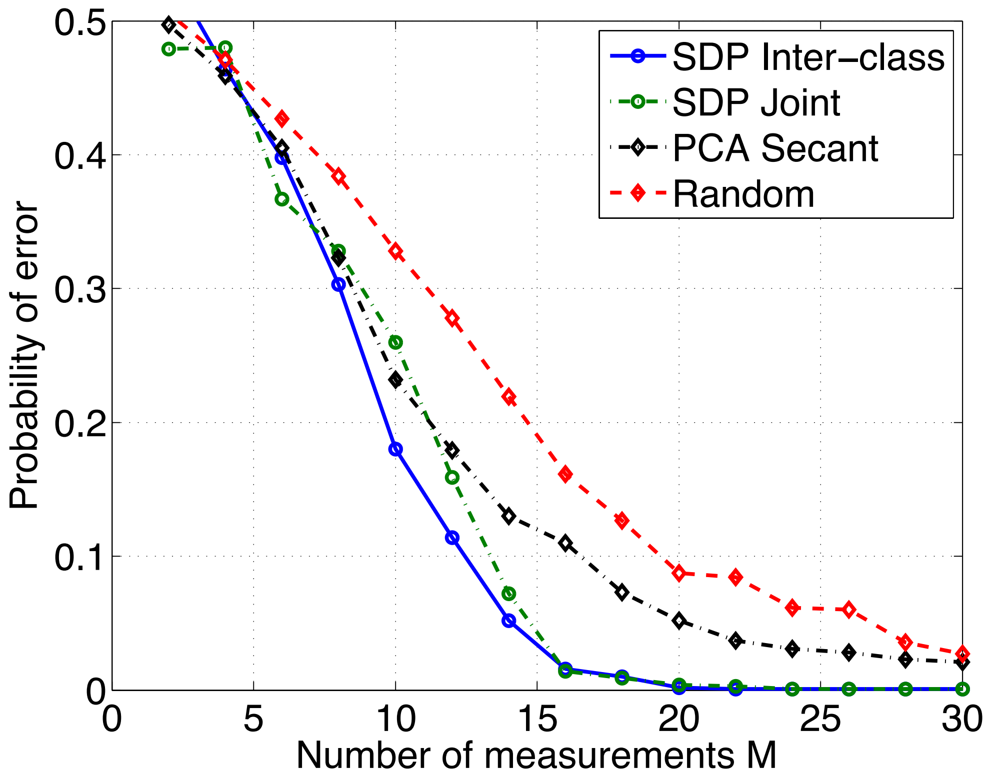


CVDomes

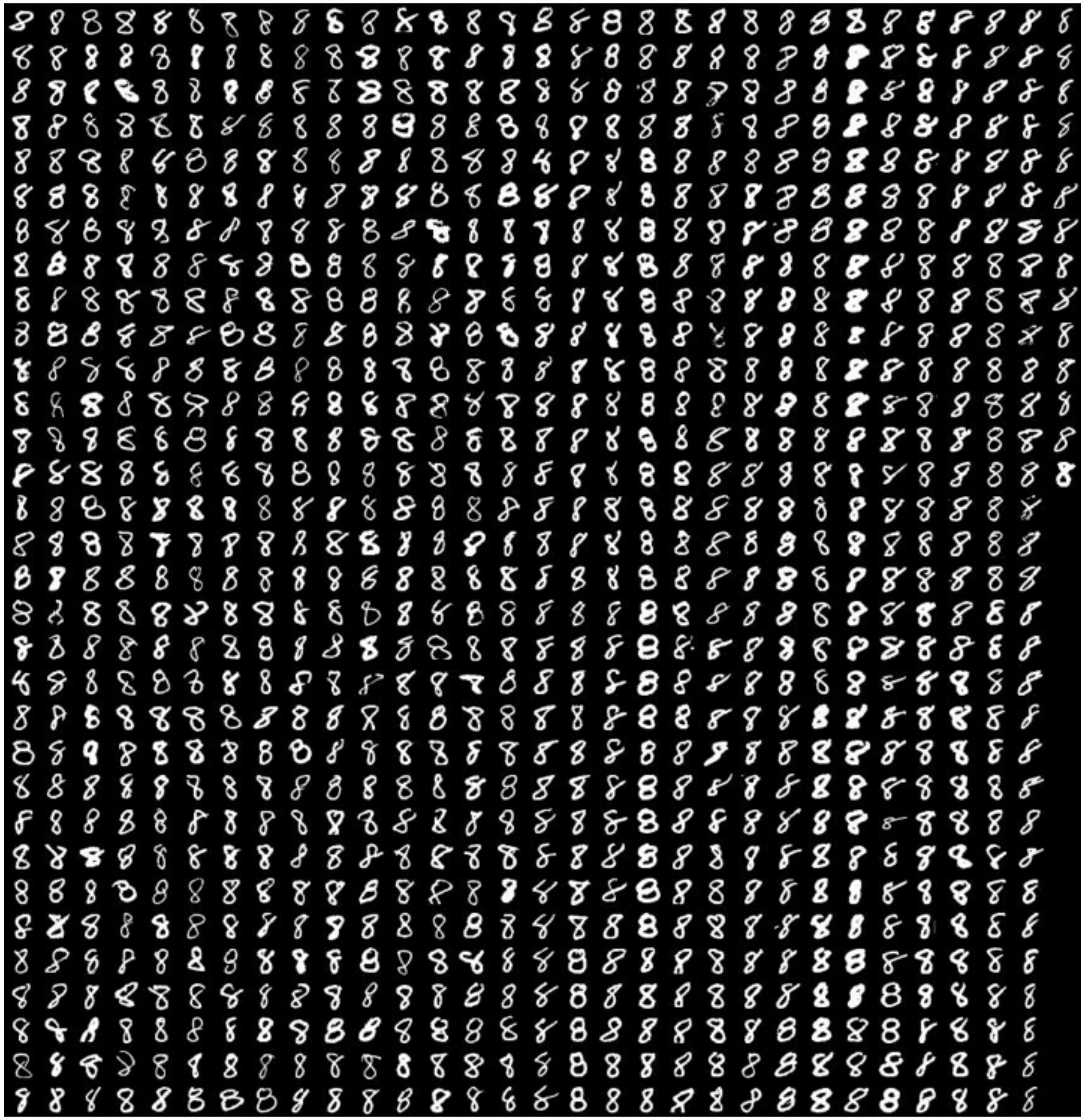
- Training data: 2000 secants (inter-class, joint)
- Test data: 100 signatures from each class

Camry:

Maxima:

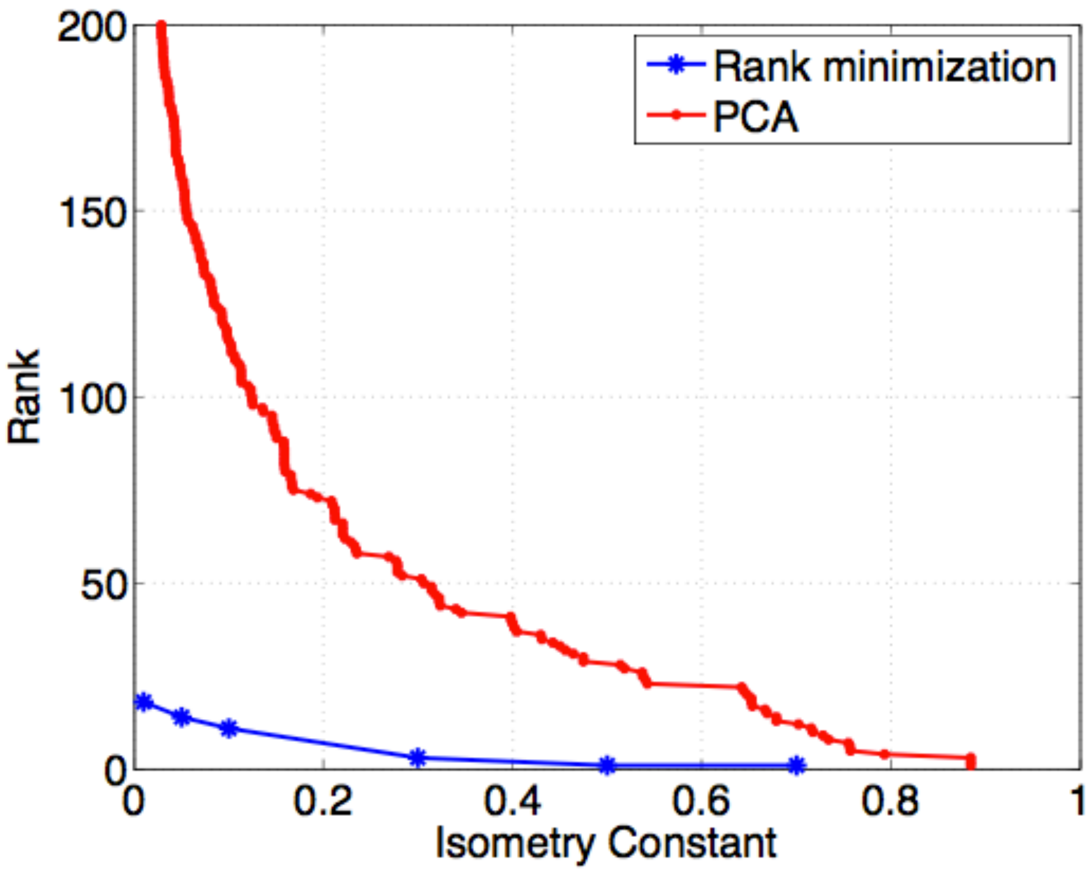


MNIST Dataset



Compression of MNIST

- Size of images: 28x28, $Q >= 1000$
- Excellent dim. reduction using merely M=20 linear projections



Summary

- Goal: develop a manifold embedding representation that is ***linear, isometric***

- Inspiration: Whitney's Theorem (preserve secants)

- Can be posed as a rank-minimization problem
  - Semi-definite program (SDP) achieves this efficiently

- **Applications**: manifold embedding, classification, compression

Directions

- Incorporate block-Toeplitz / circulant structure
  - How to cope with loss of #degrees of design freedom?

- Establish (rigorous) equivalence between rank and trace minimization problems
  - What is the relation b/w isometry constant, rank

- Secant-based approach can be linked to existing theory of *optimal transmit-receive radar* (Guerci *et al.*, 2001)