Random Projections for Manifold Learning

Chinmay Hegde, Richard G. Baraniuk — Rice University

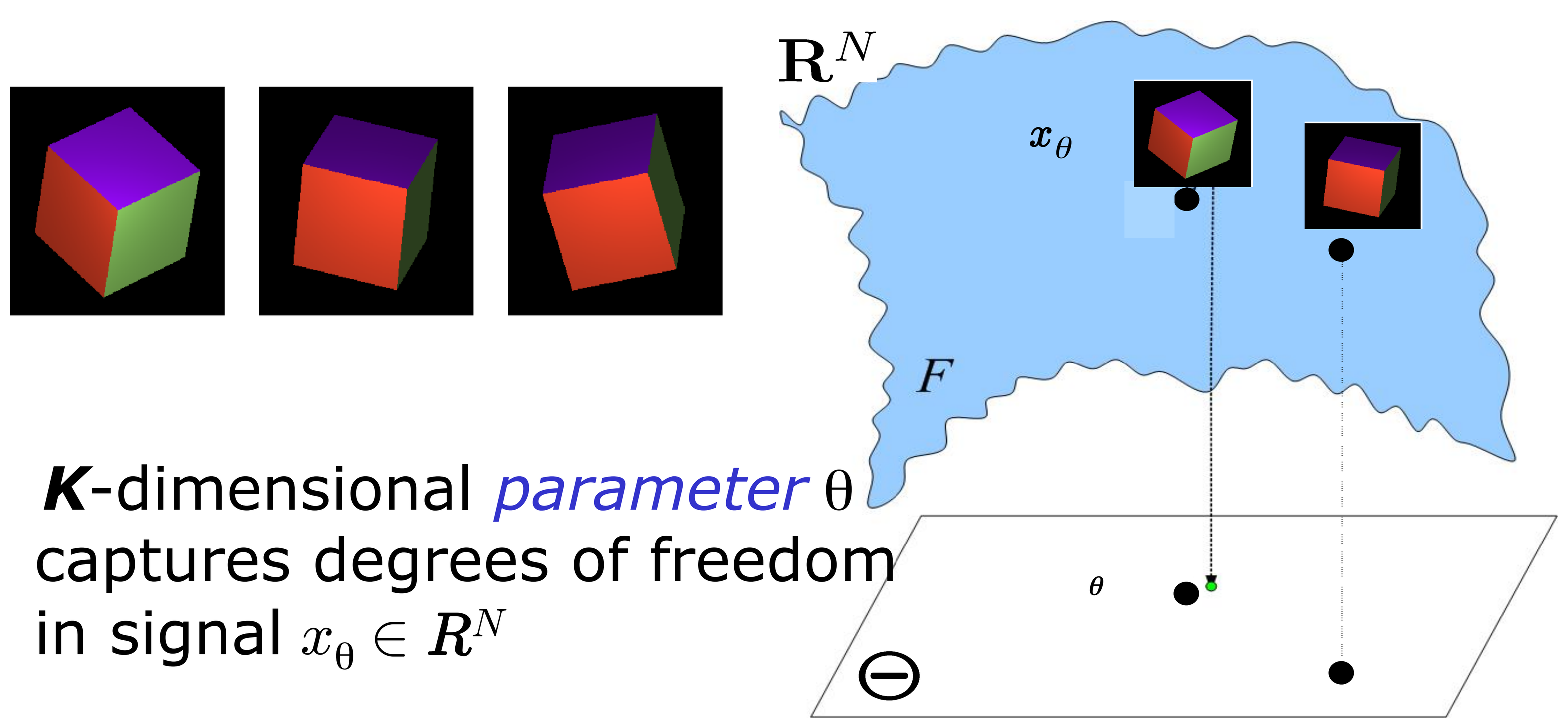
Michael B. Wakin — University of Michigan, Ann Arbor

Background

- Manifold models sought to overcome the "curse of dimensionality"
- Manifold learning: adaptively construct a low-dimensional mapping that preserves *ensemble* structure
- Random projections: information-preserving non-adaptive mappings

Manifold Models

Many high-dimensional signal ensembles possess intrinsic low-dimensional geometric structure

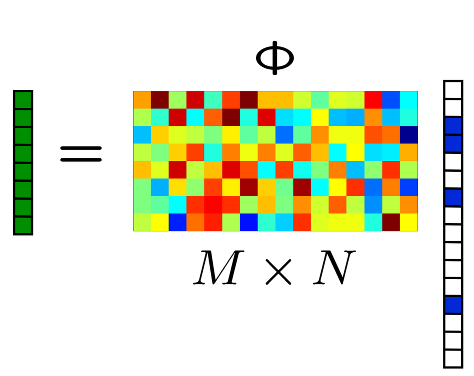


***K*-dimensional *parameter* θ captures degrees of freedom in signal $x_\theta \in R^N$**

The Random Projections Method

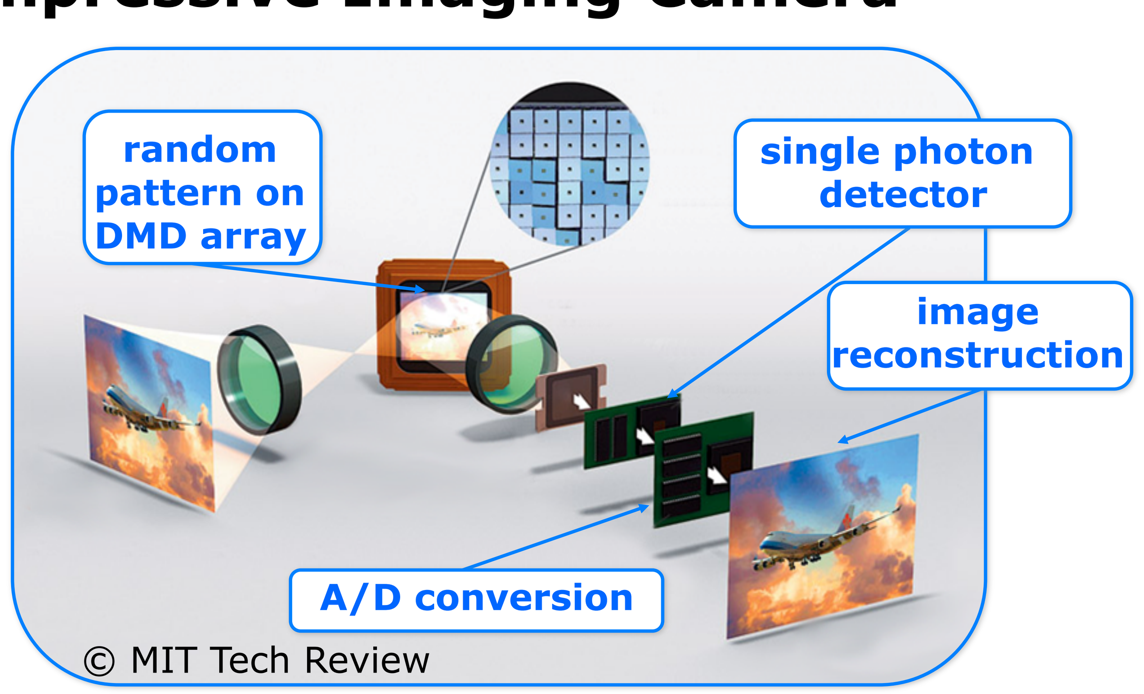
Compute *random linear combinations* of components of high dimensional data

$M \times 1$ **projections**



$N \times 1$ **signal**

Rice Compressive Imaging Camera



Learning from Random Projections

- Theoretical bounds on the performance of learning algorithms
- Number of projections sufficient for reliable inference *M* linear in the intrinsic dimension *K* and only logarithmic in the data length *N*
- Bounds depend on manifold parameters like volume, curvature.

Nonlinear Inference

Random projections preserve enough information for discovering ensemble structure

- ***Intrinsic dimension estimation***
- ***Construction of low-dimensional embedding***

Theoretical Results

Main contribution: bounds on the performance of learning algorithms on randomly projected data

Algorithms studied:

- **The Grassberger-Procaccia Algorithm for estimation of intrinsic dimension**
  - **Involves computing pairwise distances between data points**
  - **Result:**

    If $M \square O \square K \log N / \square^2 \square$, the GP estimate in the projected domain is bounded by a *multiplicative* error quantity:

$$\square 1 - \square \square K \square K_{\square} \square \square 1 \square \square \square K$$

- **Isomap for construction of a low-dimensional mapping of the input data**
  - **Involves estimating geodesic distances between data points**
  - **Result:**

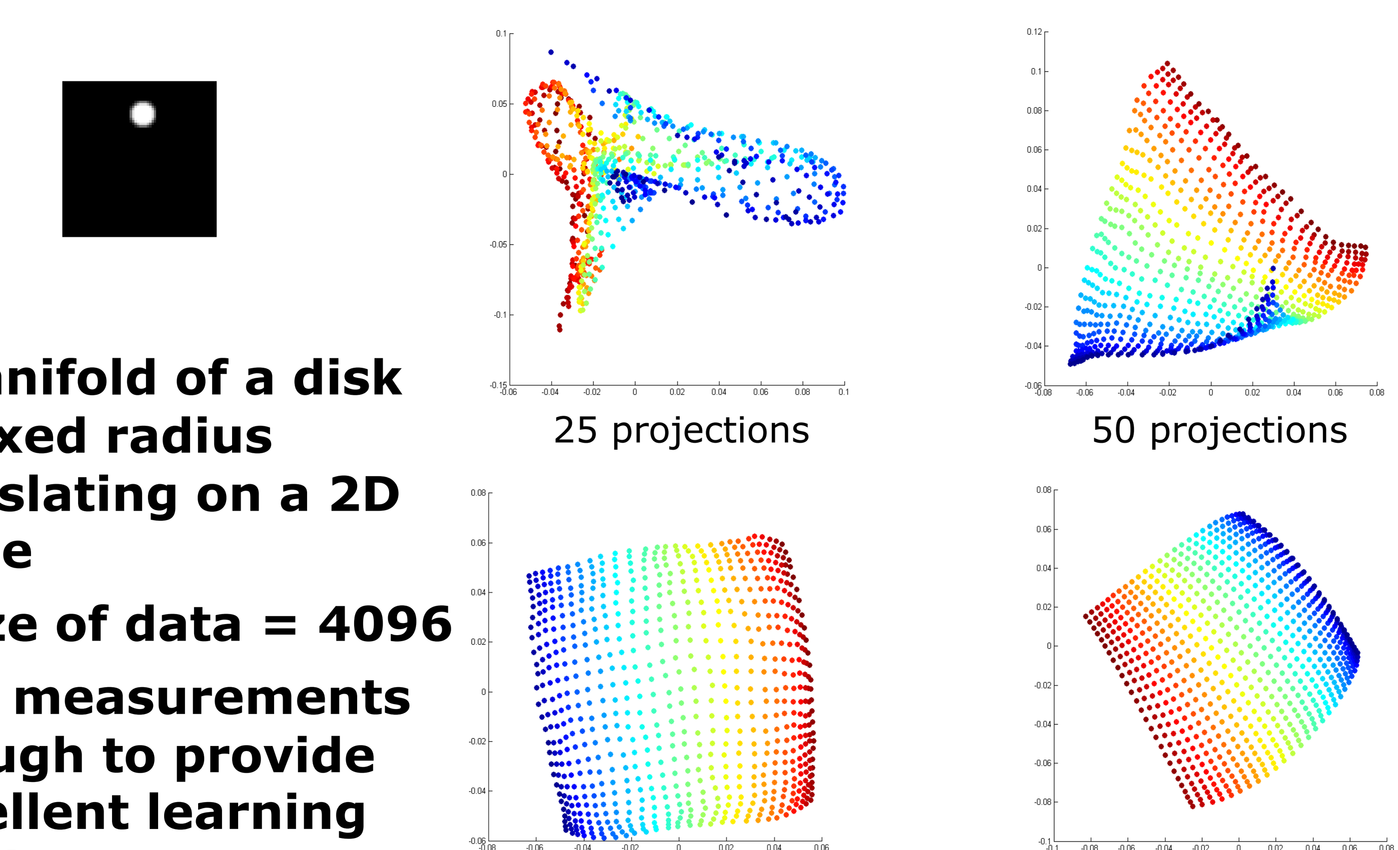
    If $M \square O \square K \log N / \square^2 \square$, the Isomap *residual variance* in the projected domain is bounded by an *additive* error quantity.

$$R_{\square} \square R \square C \square$$

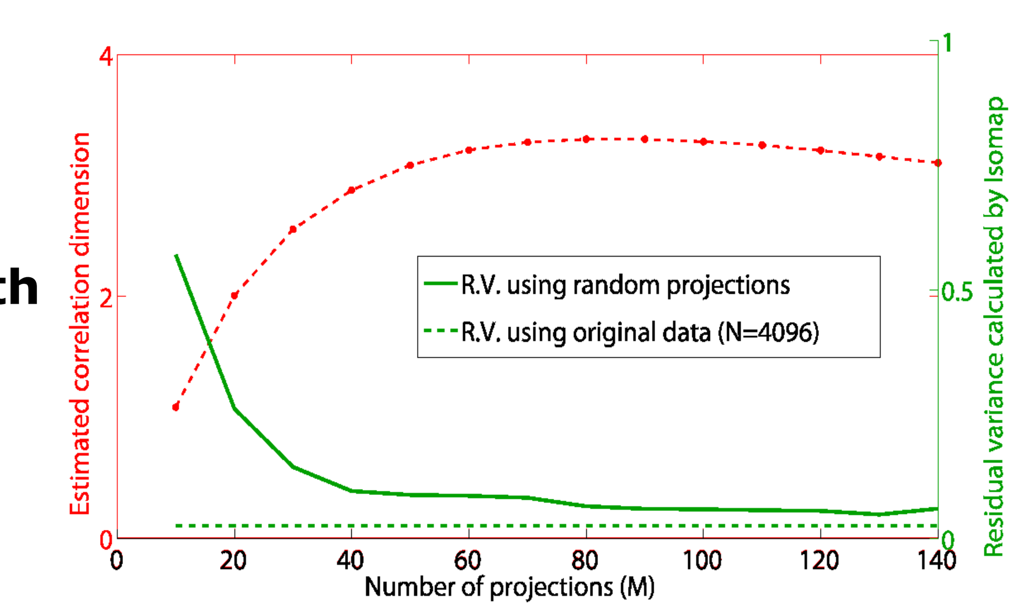
Experiments

- Bounds are pessimistic, difficult to compute explicitly
- Solution: start small, progressively acquire more measurements
- Testing done on synthetic and real databases used in learning

Learning with projected data



- **Manifold of a disk of fixed radius translating on a 2D plane**
- **Size of data = 4096**
- **50 measurements enough to provide excellent learning results**



- **Stanford face database (*N* = 4096)**
- **Effective learning with 60 projections per image**

Observations

- **Size of projected data < 2% original dimension in experiments**
- **Dimensionality reduction step simple, linear**
- **Significant cost savings in data transmission, processing**