Compressive Sensing

Instructor: Lixiang LIAN

Course Information

- Instructor:
  - Lian Lixiang (lianlx@shanghaitech.edu.cn)

- TA:
  - Gao Peng (gaopeng@shanghaitech.edu.cn)
  - Zhang Junbo (zhangjb1@shanghaitech.edu.cn)

- Office Hour
  - Walk in
  - SIST-1A404C

- Rules in Classroom
  - Questions, discussions and suggestions are always welcome

Reference

1. "*High-Dimensional Data Analysis with Low-Dimensional Models: Principles, Computation, and Applications*" by John Wright and Yi Ma

   Can be downloaded through https://book-wright-ma.github.io/

2. A Mathematical Introduction to Compressive Sensing, by Simon Foucart and Holger Rauhut, Applied and Numeric Harmonic Analysis, Springer, New York, 2013, ISBN 978-0-8176-4948-7.

3. Convex Optimization, by Stephen Boyd and Lieven Vandenberghe, Cambridge University Press, 2004, ISBN 0-521-83378-7.

4. Other original research papers

Course Information

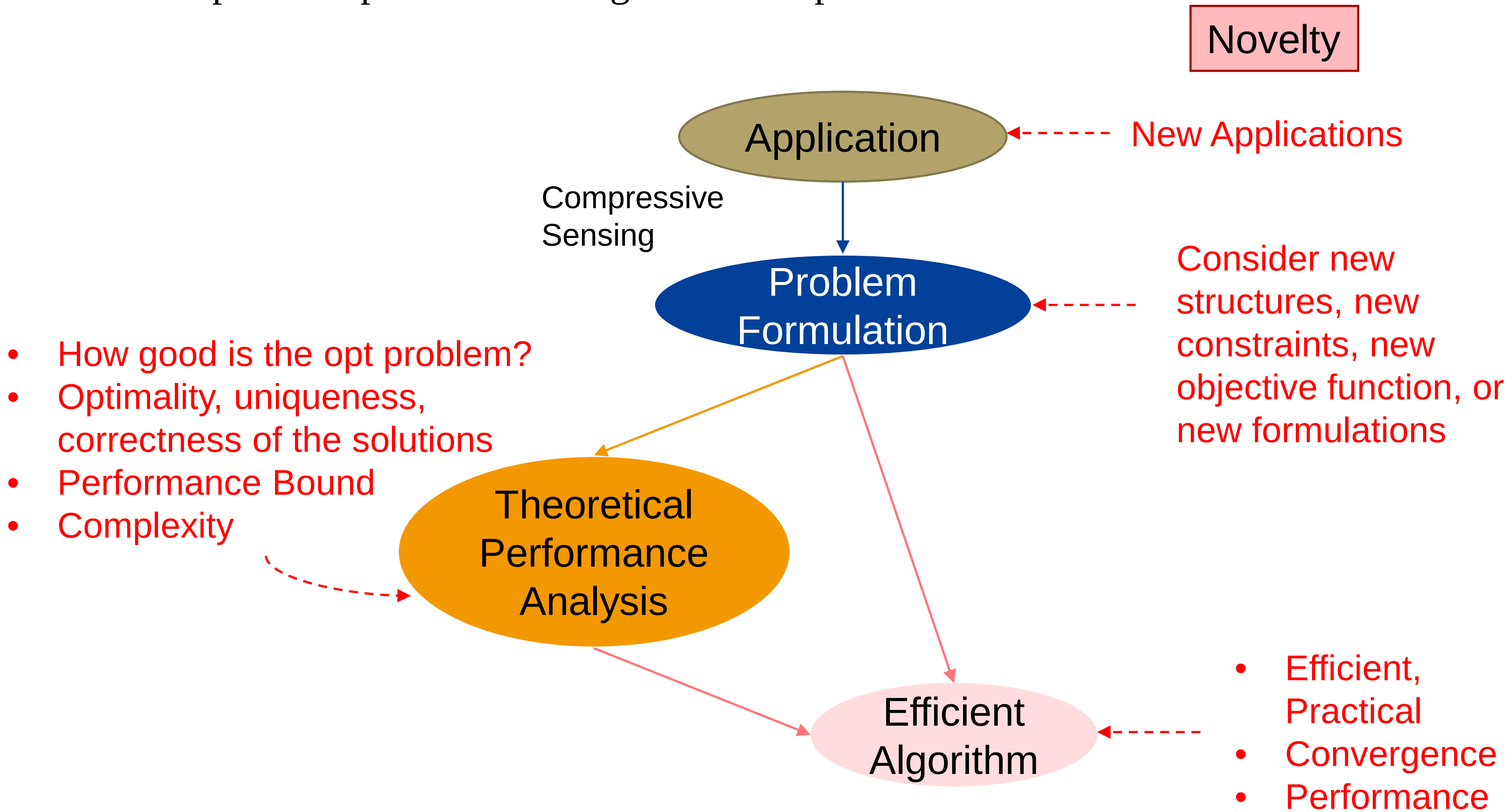
- Course Title:
  - Compressive Sensing
- Course pre-requisites:
  - Linear Algebra, Probability and Statistics
  - Convex Optimization
- Grading
  - **Homework** (4 times): **50%**
  - **Project**
    - **Proposal** before week 8 (1~2 page): **5%**
    - **Presentation: 25%**
      - How long?
        - Presentation + Q&A: 20 min/person
    - **Report: 20%**
      - Length: >=4 pages
      - Language: English
- Honor code:
  - **Plagiarism, zero tolerance**

Course Information

- Project
  - Topic: Compressive Sensing Related Topics
  - No restrictions
    - Computer Science, Signal Processing, Wireless Communication, Life Science, Electronics…
    - Different Spaces: vector space (sparse vector), matrix space (low rankness)
    - Application based (domain specific algorithm design)
    - Mathematical tool based (More general, can be applied to different applications)
  - Key Point:
    - **Novelty (Your Own Contributions)**

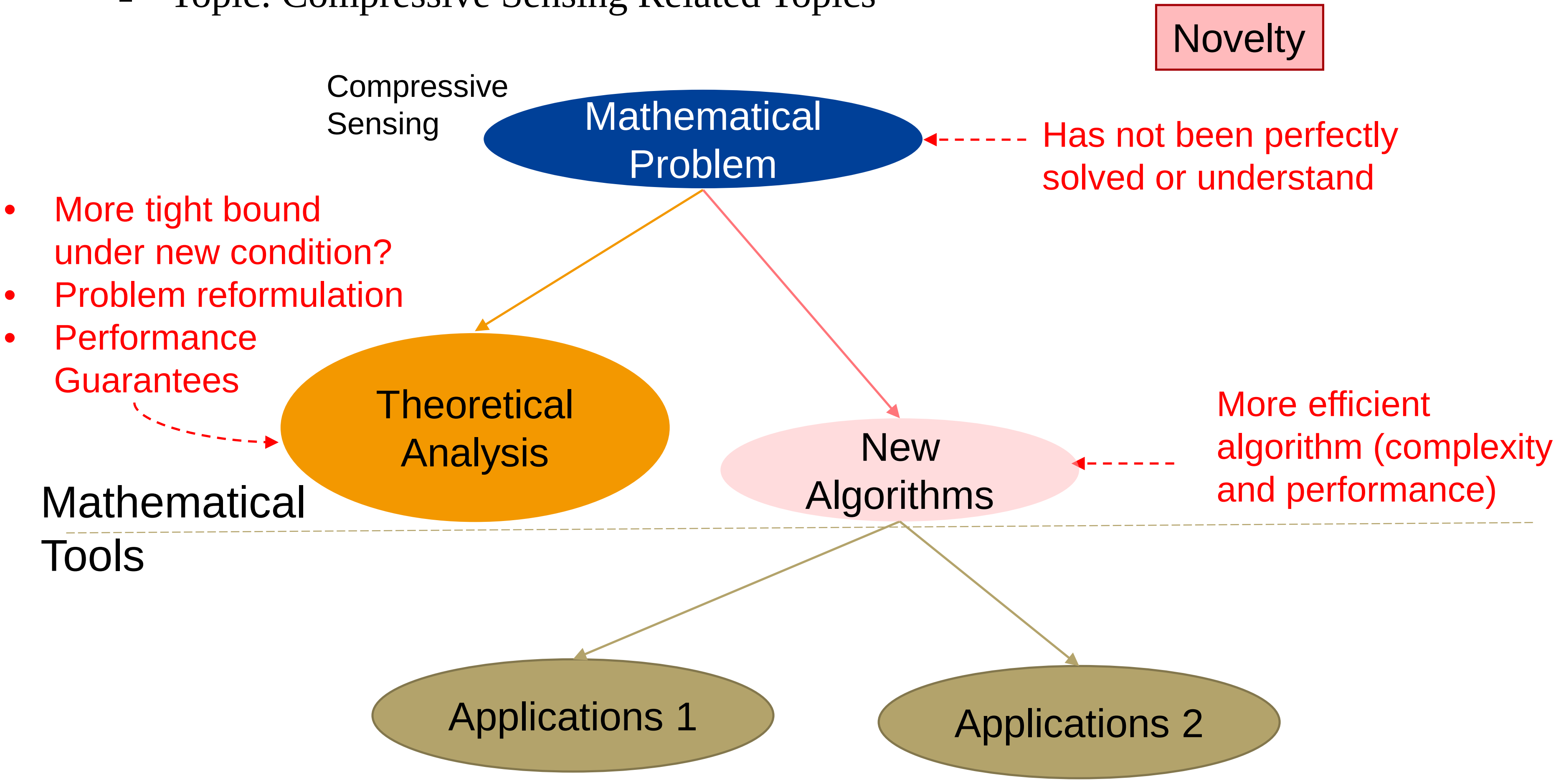
Course Information

- Project
  - Topic: Compressive Sensing Related Topics



Course Information

- Project
  - Topic: Compressive Sensing Related Topics



Syllabus

Week	Topics
1	Chapter 1 Introduction 1.1 What is Compressive Sensing 1.2 Applications of compressive sensing
2	Chapter 2 Sparse Signal Model 2.1 Applications of Sparse Signal Modeling 2.2 Recovering a Sparse Solution 2.3 Relaxing the Sparse Recovery Problem
3-5	Chapter 3 RIP for ℓ_1 Norm Minimization 3.1 Why Does ℓ_1 Norm Minimization Succeed? 3.2 Coherence of a Matrix 3.3 RIP 3.4 Matrices with RIP 3.5 Noisy Observations or Approximate Sparsity 3.6 Phase Transitions in Sparse Recovery
6-8	Chapter 4 Convex Optimization for Sparse Signal Recovery 4.1 Proximal Gradient Methods 4.2 Augmented Lagrange Multipliers 4.3 Alternating Direction Method of Multipliers 4.4 Frank-Wolfe Method for Scalable Optimization

Week	Topics
9-10	Chapter 5 Bayesian Inference for Sparse Signal Recovery 5.1 Bayesian Inference 5.2 Expectation Maximization 5.3 Sparse Bayesian Learning 5.4 Variational Bayesian Inference
11-13	Chapter 6 Structured Sparse Signal Recovery 6.1 What is Factor Graph? 6.2 Message Passing on Factor Graph 6.3 Turbo-AMP 6.4 Turbo-VBI
13-14	Chapter 7 The Applications of Compressive Sensing in Wireless Communication and Image Processing
14-16	Project Presentation

Course Objective

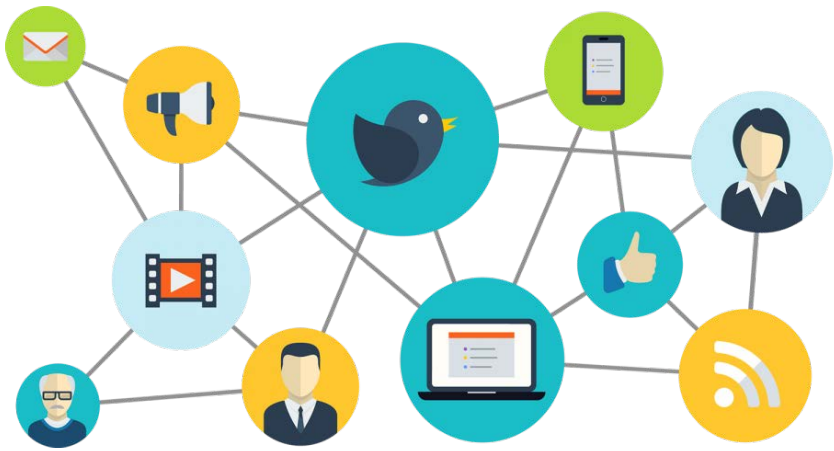
- We focus on the topics in compressive sensing theory
  - Focus on simple signal model
  - Problem Formulation → Performance Analysis → Efficient Algorithm
  - Idealistic Case → More General Case
- After the course, the students are expected to
  - Learn about the basic concepts in compressive sensing: sparse signal, RIP, phase transition, Proximal Gradient, Sparse Bayesian Learning, Message Passing…
  - Learn about the different algorithms is CS
  - (inspiration, grasp new knowledge, stimulate your interests, etc.)

Lecture 1 Introduction

Instructor: Lixiang LIAN

Motivation

- Data deluge



We create 2.5 quintillion bytes of data a day according to IBM.

- **Challenges**: data acquisition and interpretation
  - Reduce the amount of information that must be measured
  - without sacrificing the ability to make useful inference from the data
  - Inverse problem→ Compressive Sensing Algorithm

Motivation

Low dimensional structure

(a) a calibration rig

(b) a carpet

(c) windows

(d) a door

(e) a license plate

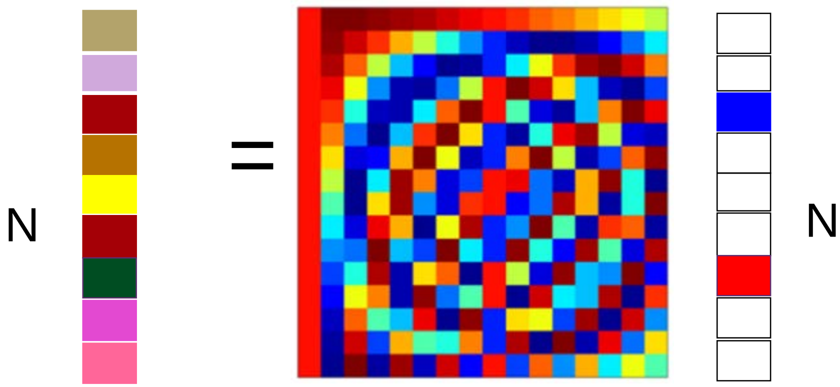
(f) characters

(g) a car

(h) a face

Figure 15.1 Representative examples of structured objects. These images viewed as matrices are all (approximately) low-rank matrices.

Sparse Representation

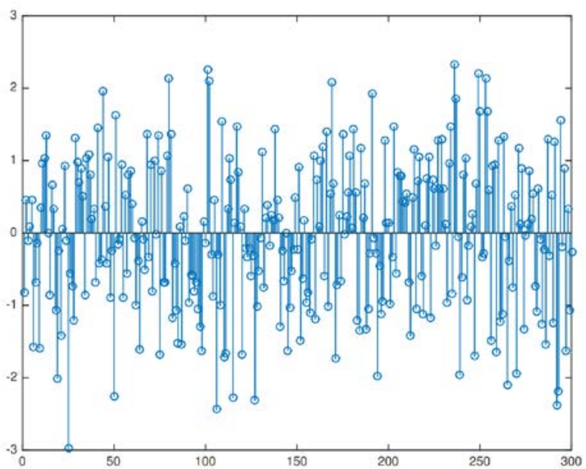


$\boldsymbol{z}$ $\quad\quad$ $\boldsymbol{\Psi}$ $\quad\quad$ $\mathbf{x}$

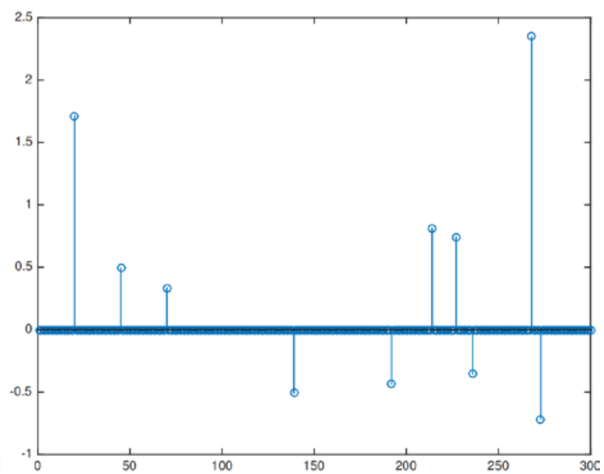
$\boldsymbol{z} = \boldsymbol{\Psi}\boldsymbol{x}$, with $\|\mathbf{x}\|_0 \ll$ N

- N-length signal of interest $\mathbf{z}$ is sparse or compressible in a known basis $\boldsymbol{\Psi}$ (e.g., wavelet, Fourier or identity basis)
- Denote the non-zero indices as the **support of x**

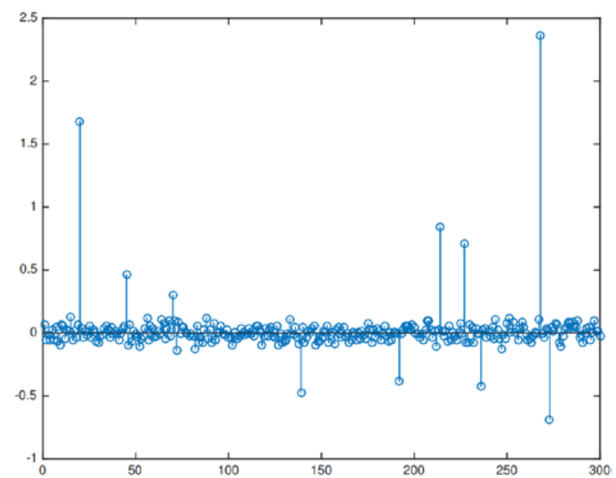
$$\Omega = \{n : \mathbf{x}[n] \neq 0\}$$



dense vector



sparse vector

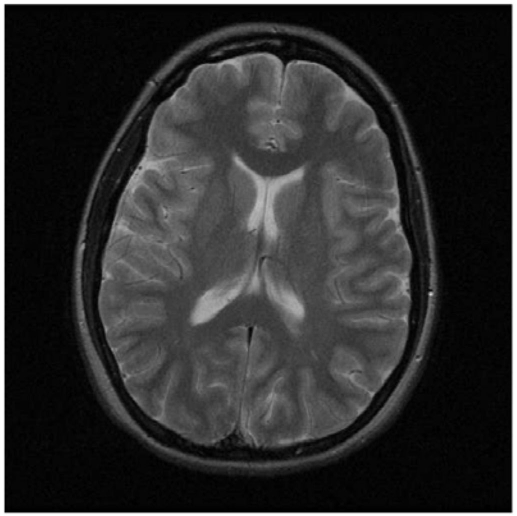


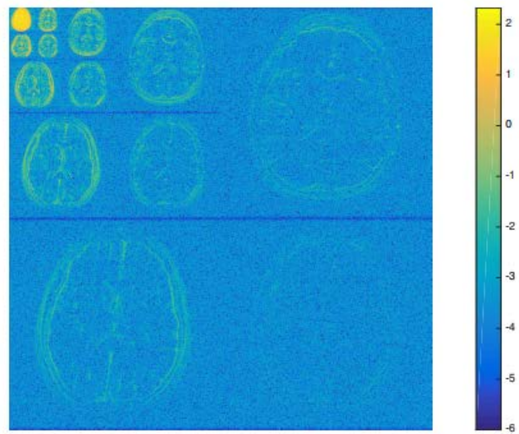
compressible vector

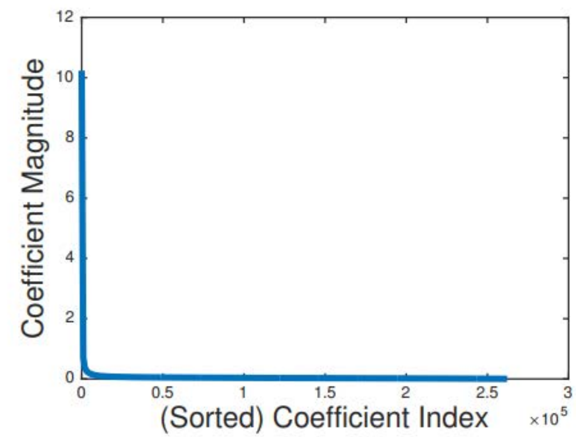
Sparse Representation

Example 1: Magnetic Resonance (MR)

$$\underset{\text{image}}{I} = \sum_{i=1}^{N^2} \underset{i\text{-th basis signal}}{\psi_i} \times \underset{i\text{-th coefficient}}{x_i}.$$







wavelet coefficients x: $I = \Psi[x]$.

Target image of human brain

Wavelet transform

Sparse Representation

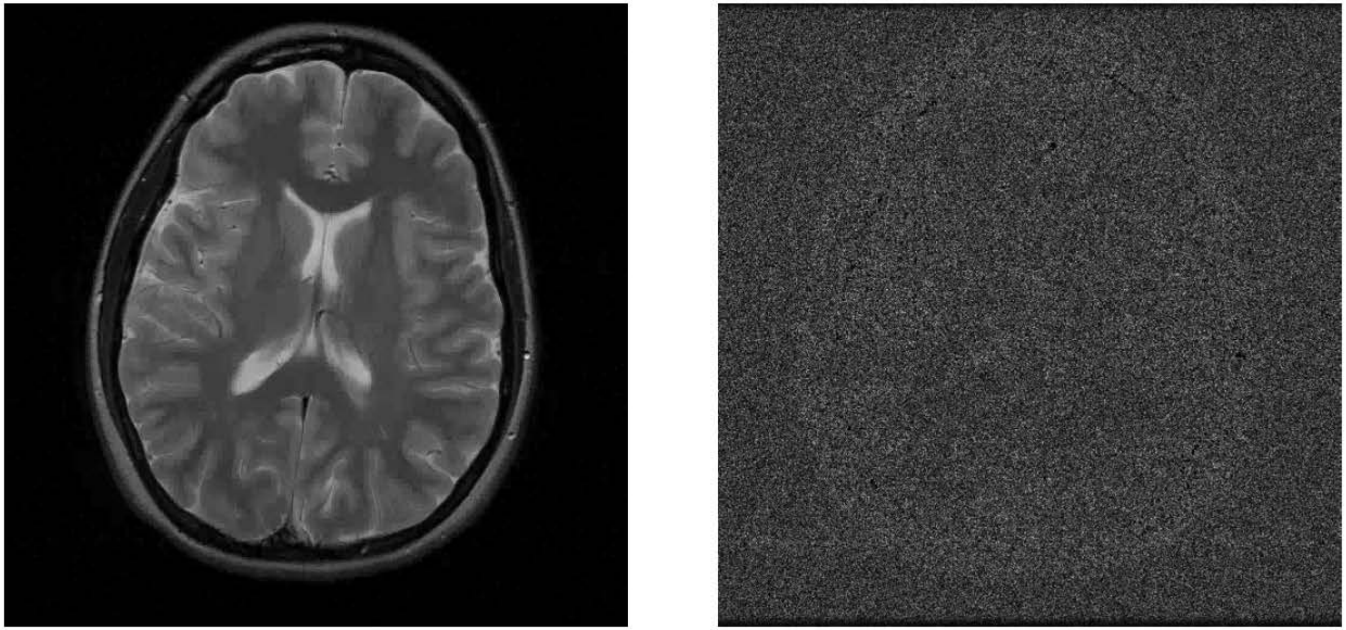
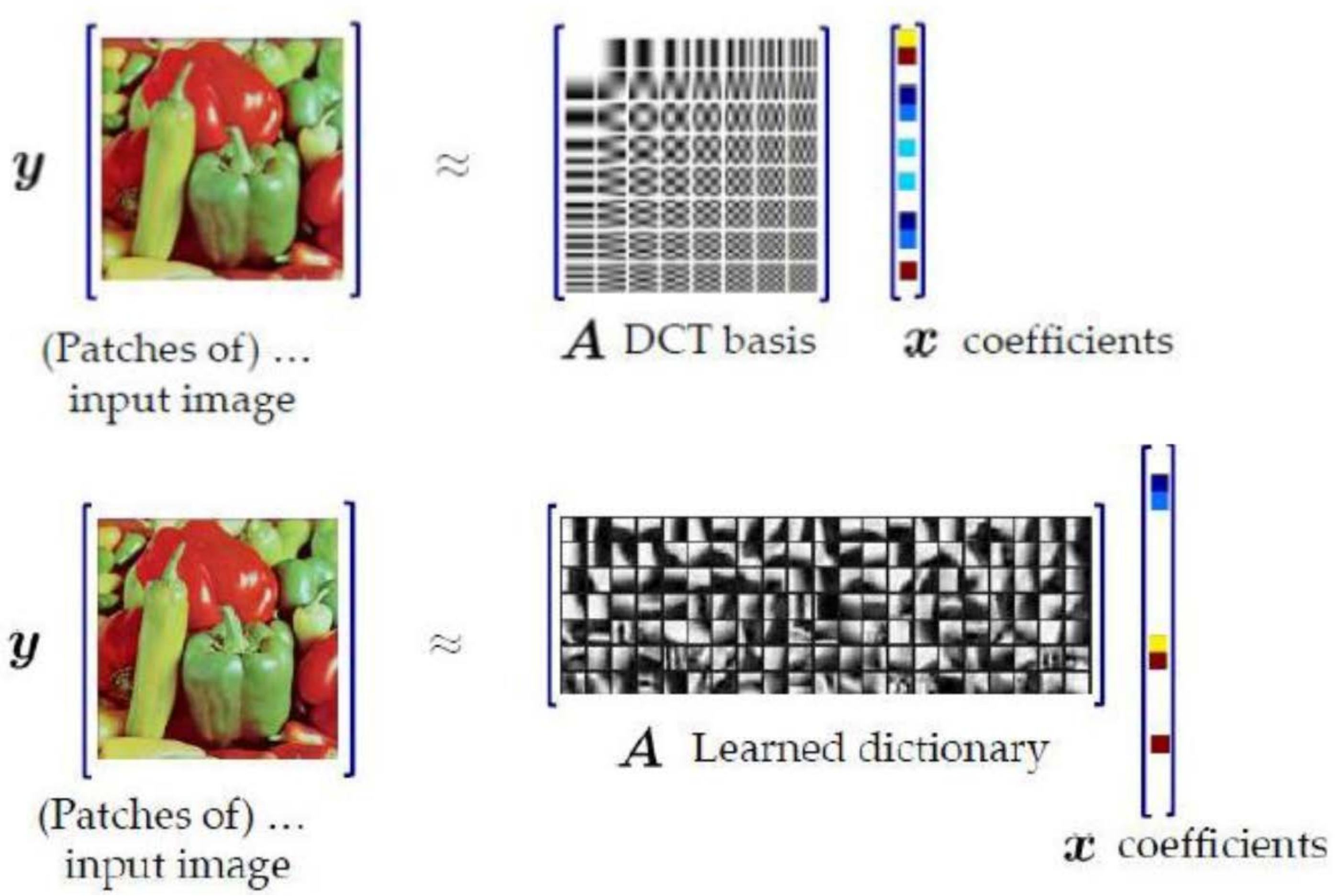


Figure 2.3 Wavelet Approximation $\tilde{I}$ to I and Approximation Error. Left: approximation to the image in Figure 2.2 using the most significant 7% of the wavelet coefficients. **Right:** approximation error $|I - \tilde{I}|$. The error contains mostly noise, suggesting that most of the important structure of the image is captured in the wavelet approximation $\tilde{I}$.

Sparse Representation

Example 2: Image Compression (JPEG, JPEG2000)

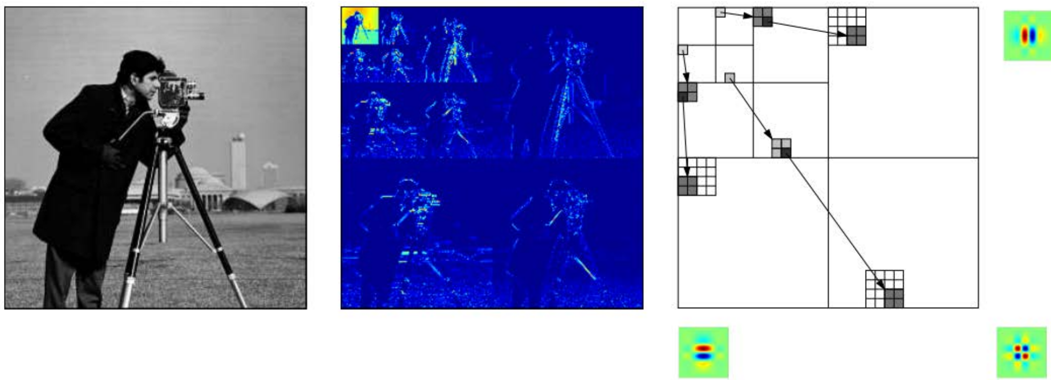
JPEG: DCT(discrete cosine transform): for smooth variations



Sparse Representation

Example 2: Image Compression (JPEG, JPEG2000)

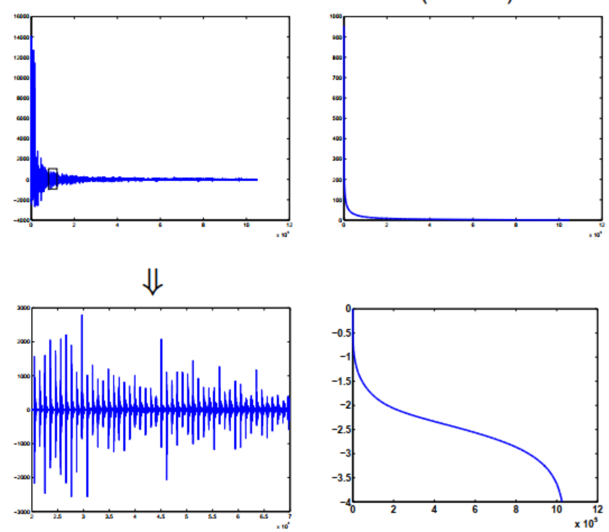
JPEG2000: 2D wavelets, for sharp edges



1 megapixel image

wavelet coeffs

(sorted)



zoom in

($\log_{10}$ sorted)

25k term approximation

Sparse Representation

Example 2: Image Compression (JPEG, JPEG2000)

Only a small number of parameters matter

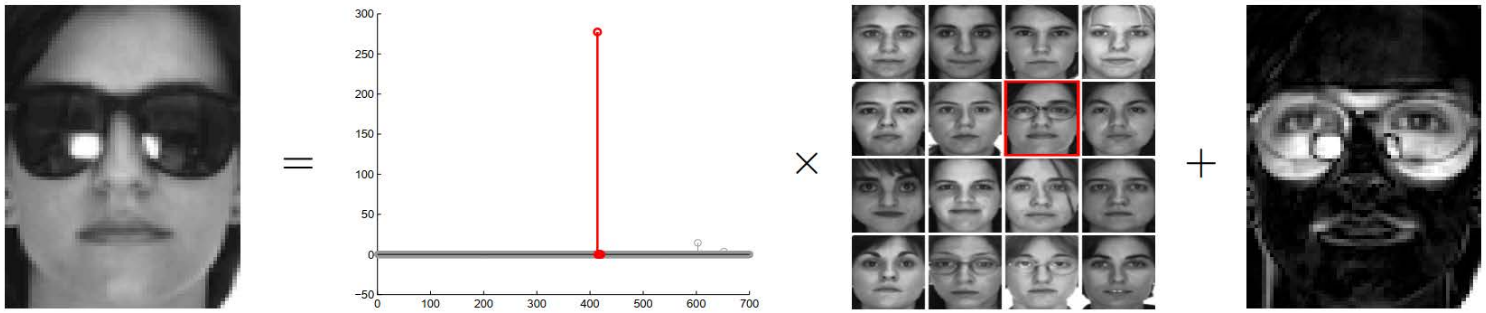
Raw: 15MB

JPEG: 150KB

There is (almost) no loss in quality between the raw image and its JPEG compressed form

Sparse Representation

Example 3: Face Recognition



$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{e}.$$

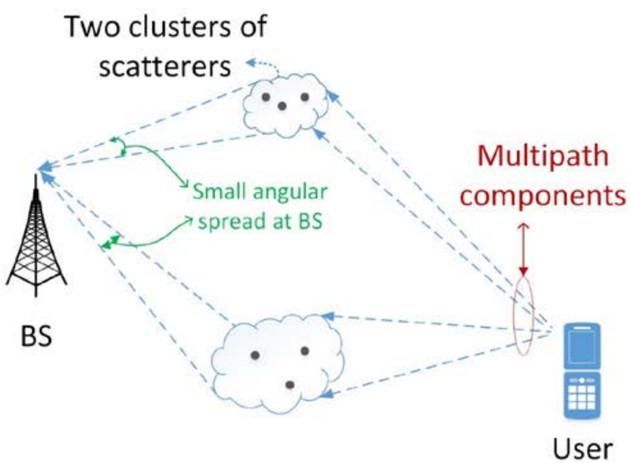
$$\boldsymbol{B} = \underbrace{[\boldsymbol{B}_1 \mid \boldsymbol{B}_2 \mid \cdots \mid \boldsymbol{B}_n]}_{\text{all training images}} \in \mathbb{R}^{m \times n}, \quad n = \sum_i n_i.$$

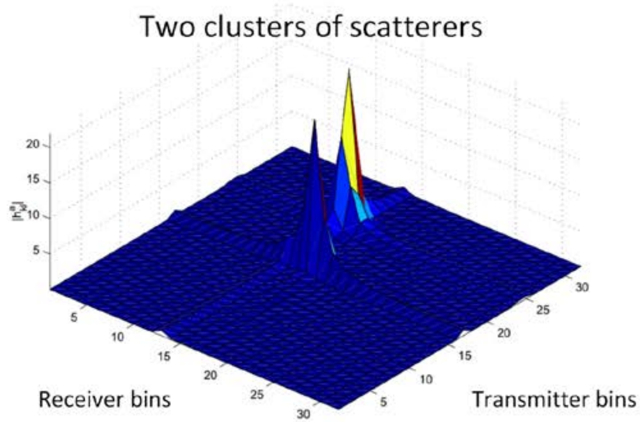
Sparse: pixels that are occluded or corrupted

Sparse: images of the true subject

Sparse Representation

Example 4: Sparse Channel Estimation





Angular domain channel

- Discrete multipath channel model:

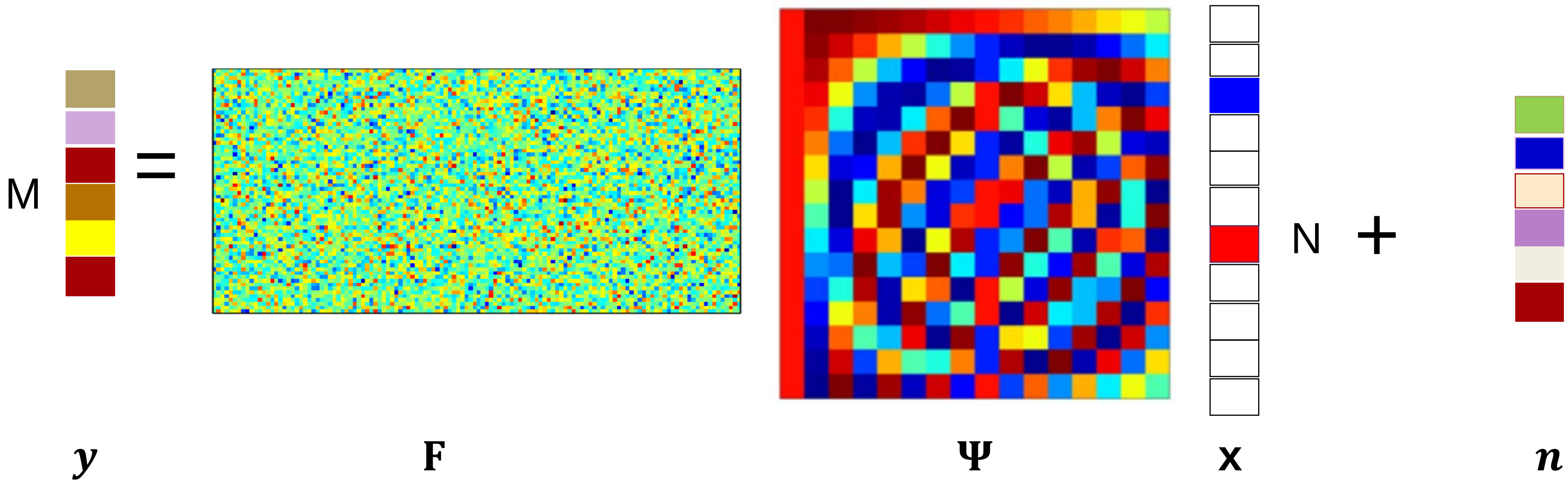
$$\mathbf{H} = \sum_{l=1}^{L} \alpha_l \mathbf{a}_R(\theta_{R,l}) \mathbf{a}_T^H(\theta_{T,l})$$

- Assume the AoA and AoD can only take value from a uniformly spaced discrete set

$$\mathbf{H} = \sum_{n=1}^{\tilde{N}} \sum_{m=1}^{\tilde{M}} x_{n,m} \mathbf{a}_R\left(\tilde{\theta}_{R,n}\right) \mathbf{a}_T^H\left(\tilde{\theta}_{T,m}\right) = \mathbf{A}_R \mathbf{X} \mathbf{A}_T^H$$

Angular domain channel matrix

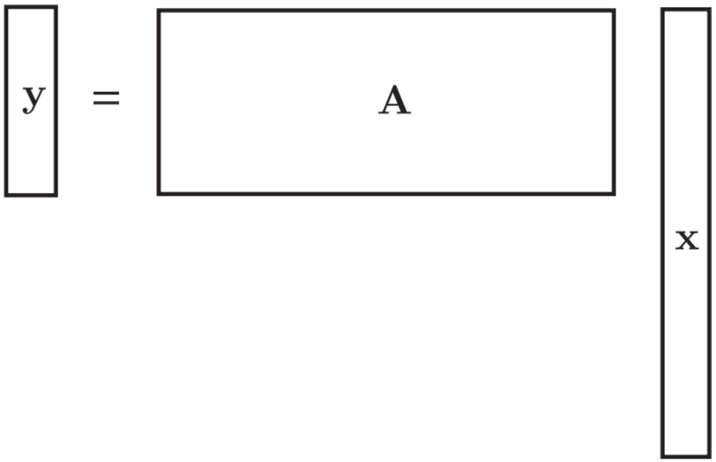
Sparse Signal Recovery

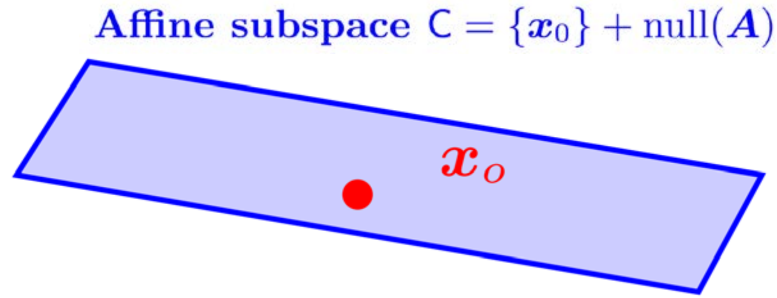


$$\mathbf{y} = \mathbf{Fz} + \mathbf{n} = \mathbf{F\Psi x} + \mathbf{n} = \mathbf{Ax} + \mathbf{n}$$

- We observe **M ≪ N** noisy linear measurements **y**:
  - Underdetermined system
  - Infinitely many feasible solutions to $\boldsymbol{Ax} = \boldsymbol{y}$

Sparse Signal Recovery





Find $\boldsymbol{x} \in \mathbb{C}^n$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$

where $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] \in \mathbb{C}^{m \times n}$ obeys

- underdetermined system: $m < n$
- full-rank: $\mathrm{rank}(\boldsymbol{A}) = m$

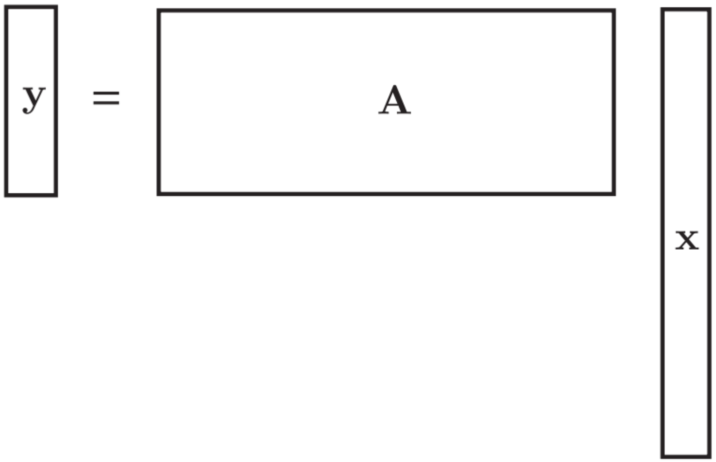
$\boldsymbol{A}$: an *over-complete basis / dictionary*; $\boldsymbol{a}_i$: atom;
$\boldsymbol{x}$: representation in this basis / dictionary

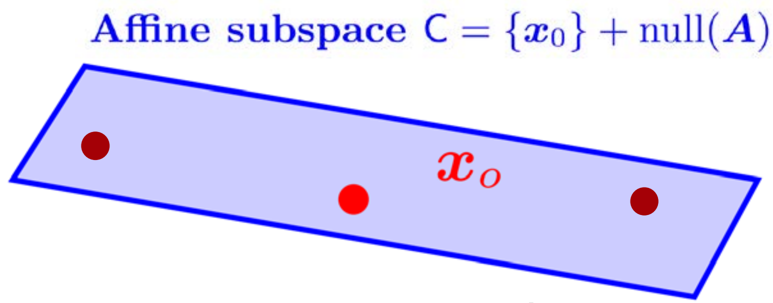
Solution set:

$$\underbrace{\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^*)^{-1}}_{\boldsymbol{A}^\dagger}\boldsymbol{y} + \mathrm{null}(\boldsymbol{A})$$

How can we possibly solve? Need more structures!

Sparse Signal Recovery





Some sparse solutions

- Explore signal structure: sparsity
- Recovery via efficient algorithms (e.g., convex optimization)

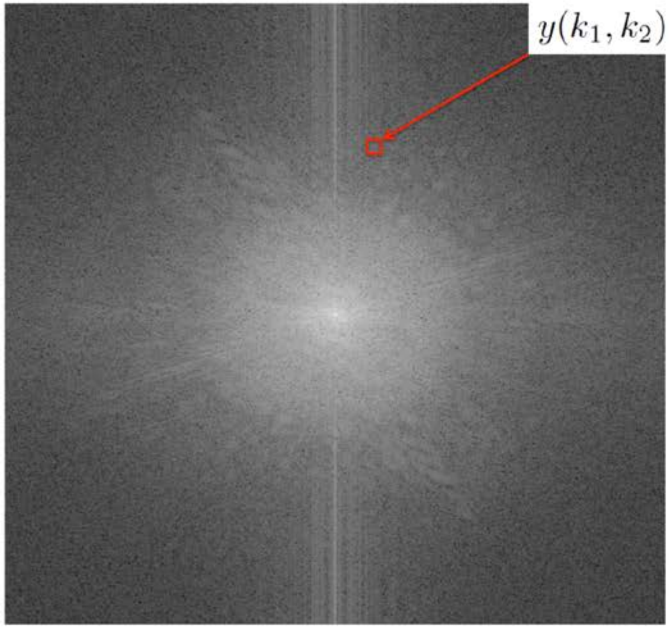
Magnetic Resonance Imaging (MRI)

MR scanner

MR image

Magnetic Resonance Imaging (MRI)

What an MRI machine sees

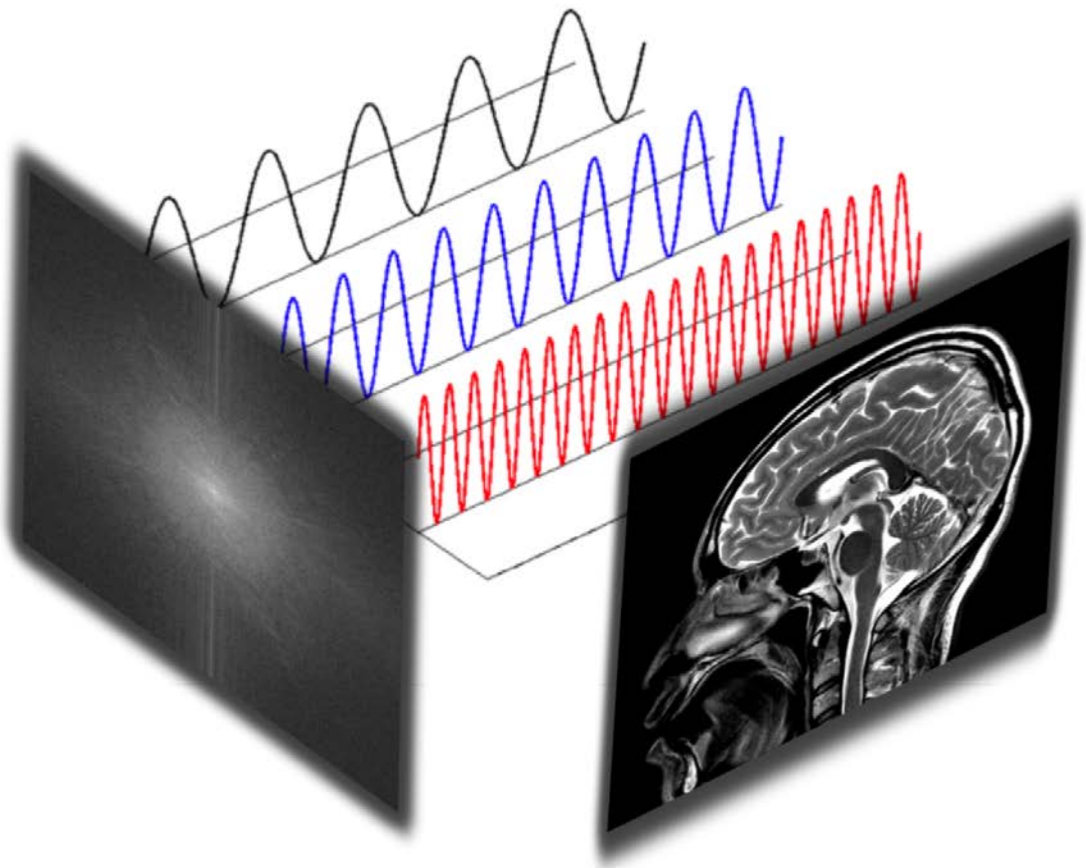


Measured data $y(k_1, k_2)$ $\longleftarrow$ Fourier transform of image $f(x_1, x_2)$

Magnetic Resonance Imaging (MRI)

Fourier Transform

$$y(k_1, k_2) \approx \sum_{x_1} \sum_{x_2} f(x_1, x_2) \mathrm{e}^{-i2\pi(k_1 x_1 + k_2 x_2)}$$



Magnetic Resonance Imaging (MRI)

How do we form an image?

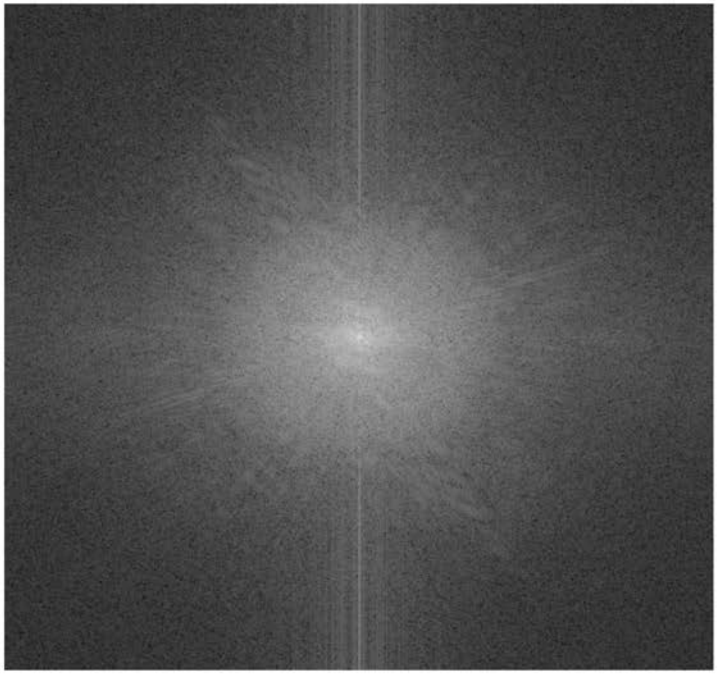
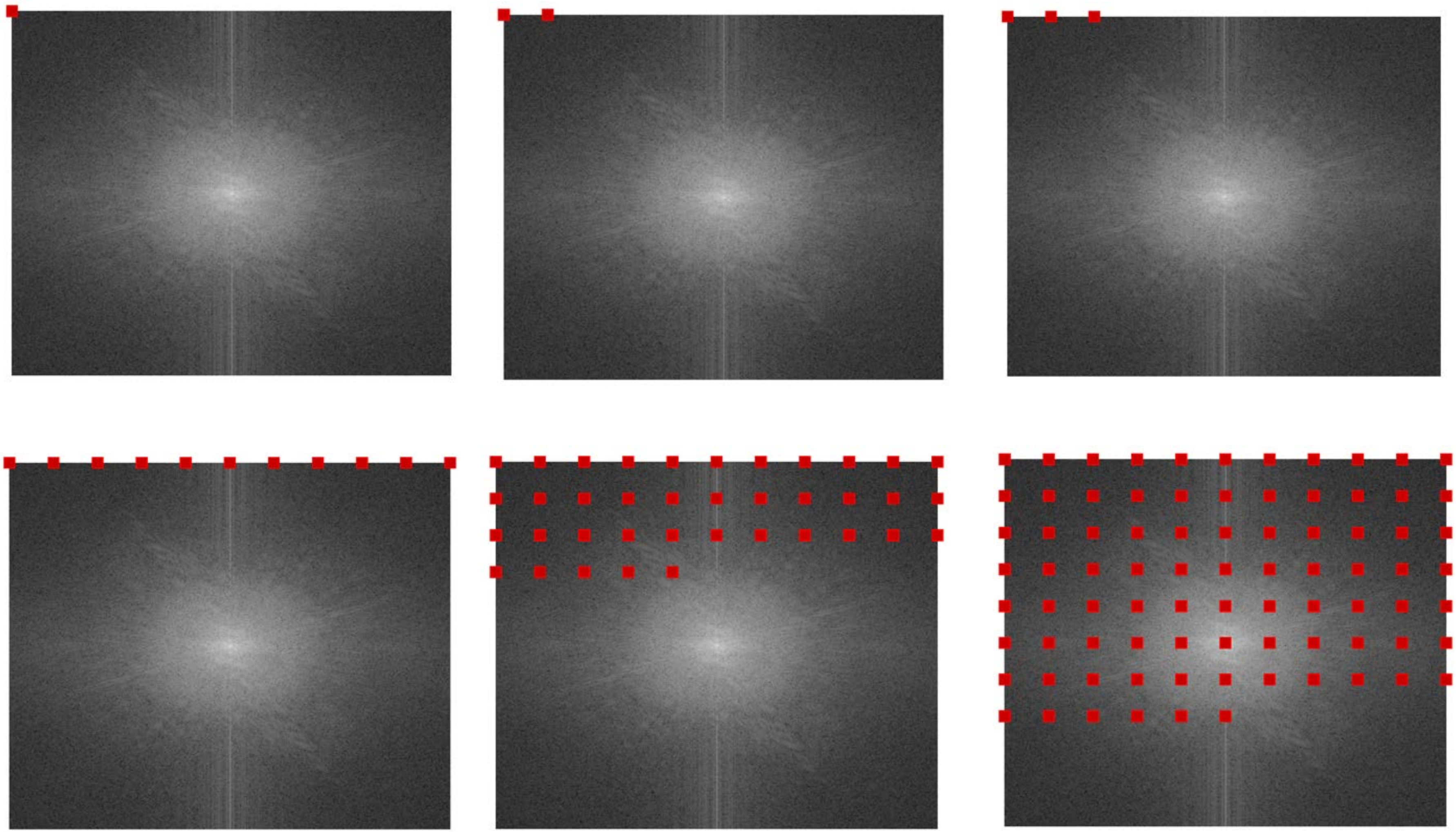


image $f(x_1, x_2)$ $\longleftarrow$ inverse Fourier transform of measurements

$$f(x_1, x_2) \approx \sum \sum y(k_1, k_2) \mathrm{e}^{i2\pi(k_1 x_1 + k_2 x_2)}$$

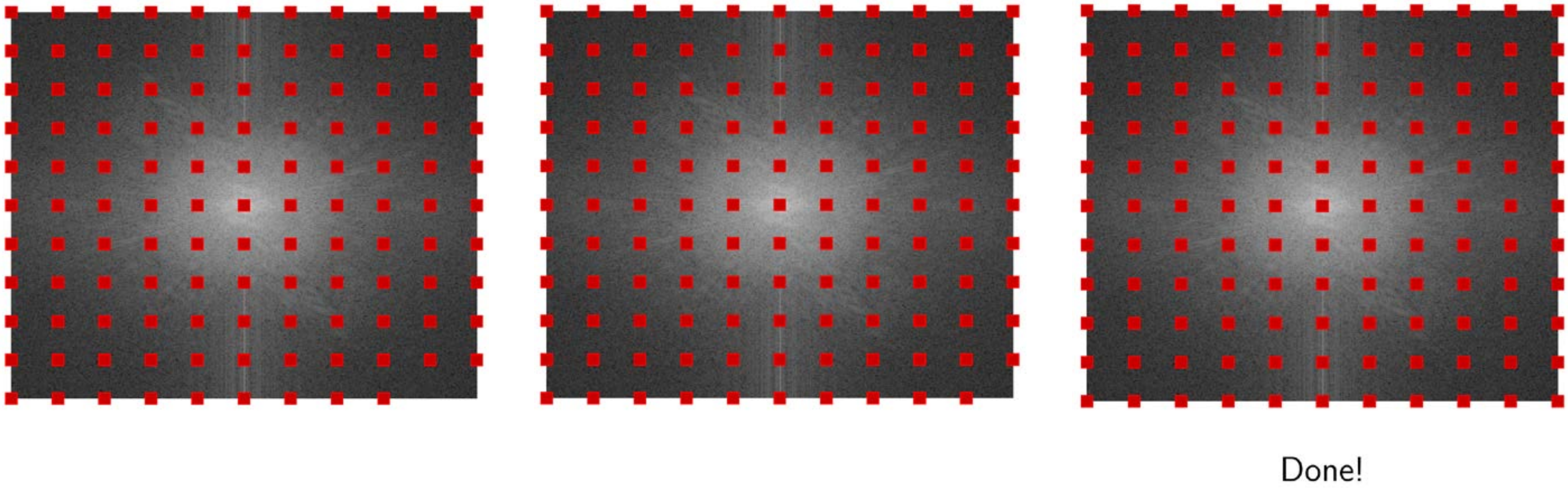
Magnetic Resonance Imaging (MRI)

MRI data collection is inherently slow

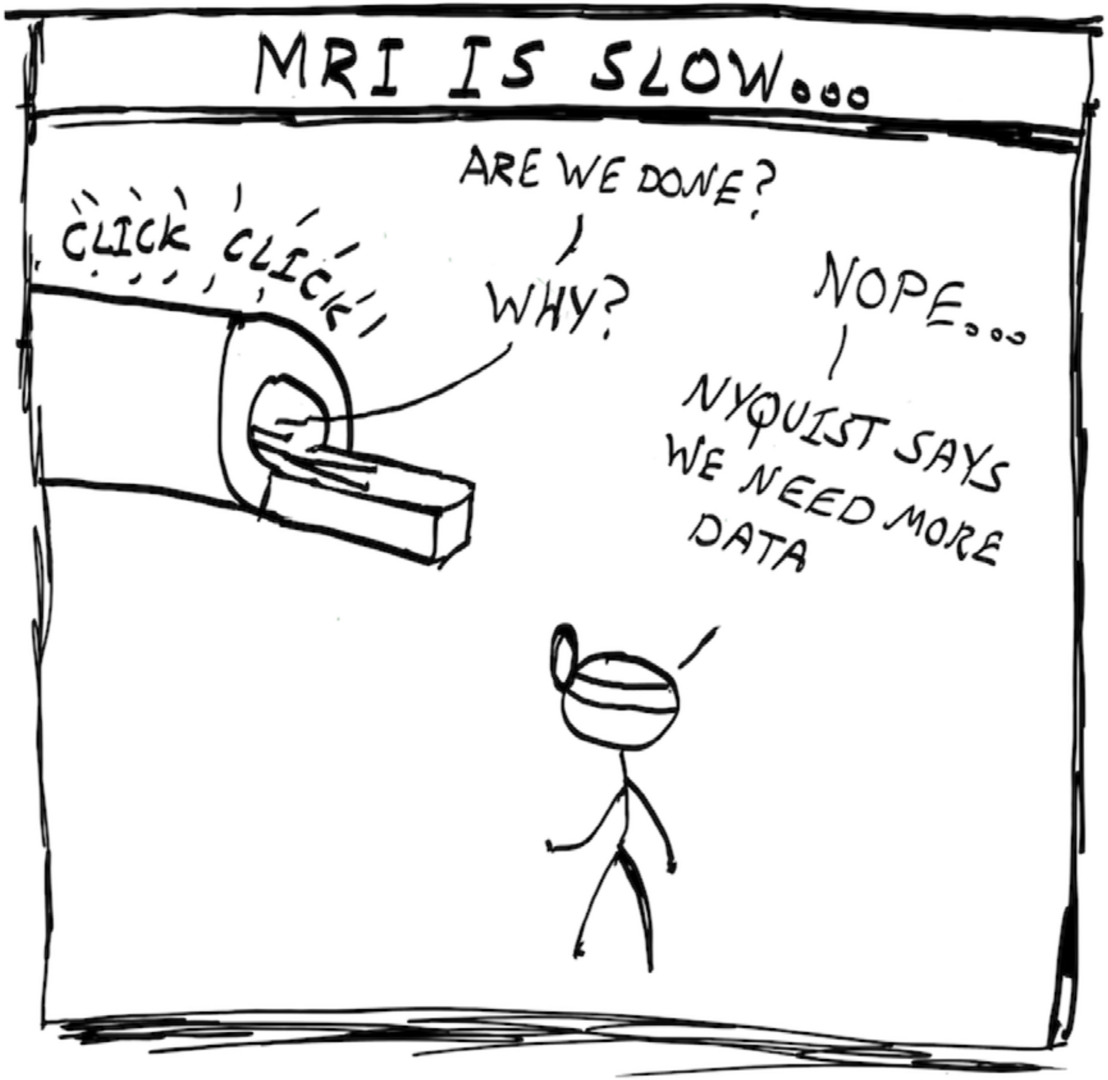


Magnetic Resonance Imaging (MRI)

MRI data collection is inherently slow



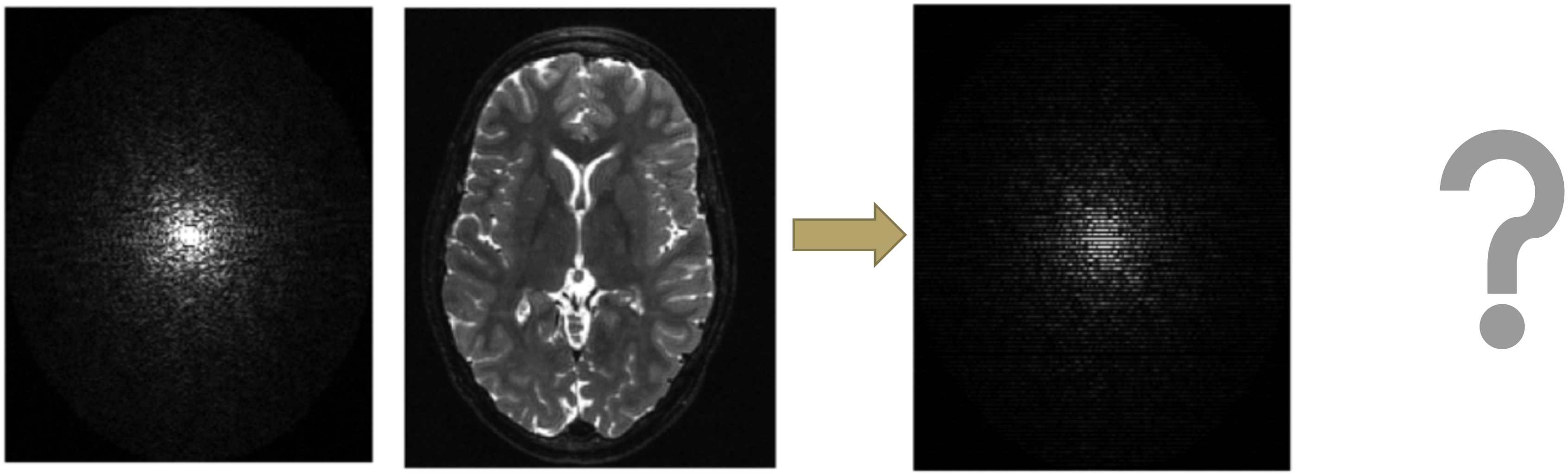
Fact: impact of MRI on children health is limited


MRI IS SLOW...
ARE WE DONE?
CLICK CLICK
WHY?
NOPE...
NYQUIST SAYS WE NEED MORE DATA

M. Lustig

Magnetic Resonance Imaging (MRI)

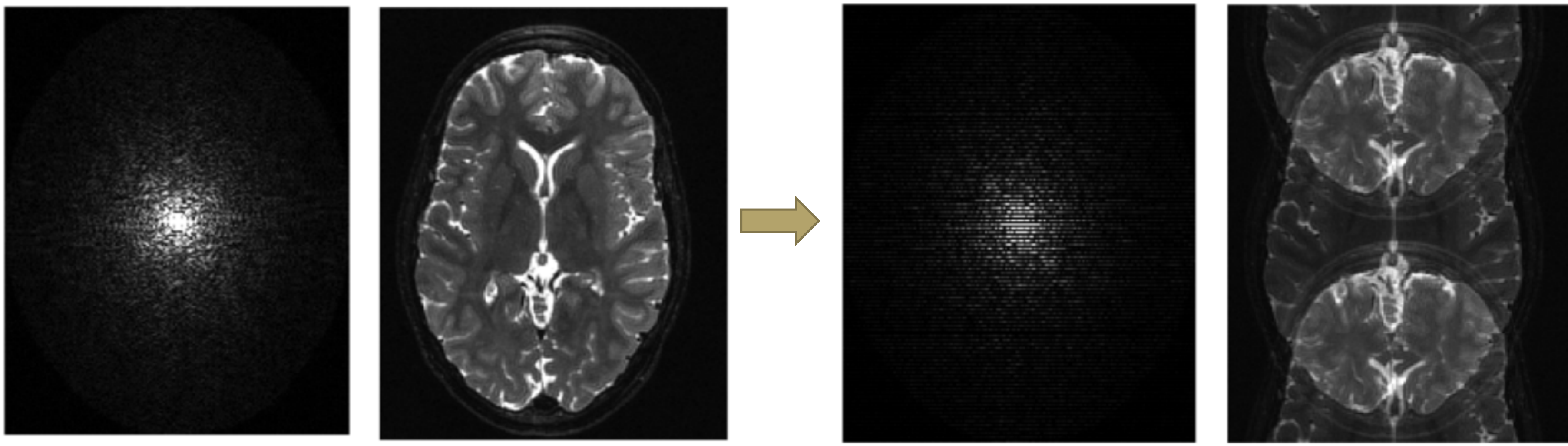
Is it possible to take fewer samples to reduce scan time?



Uniform undersampling by a factor of 2

Magnetic Resonance Imaging (MRI)

Is it possible to take fewer samples to reduce scan time?

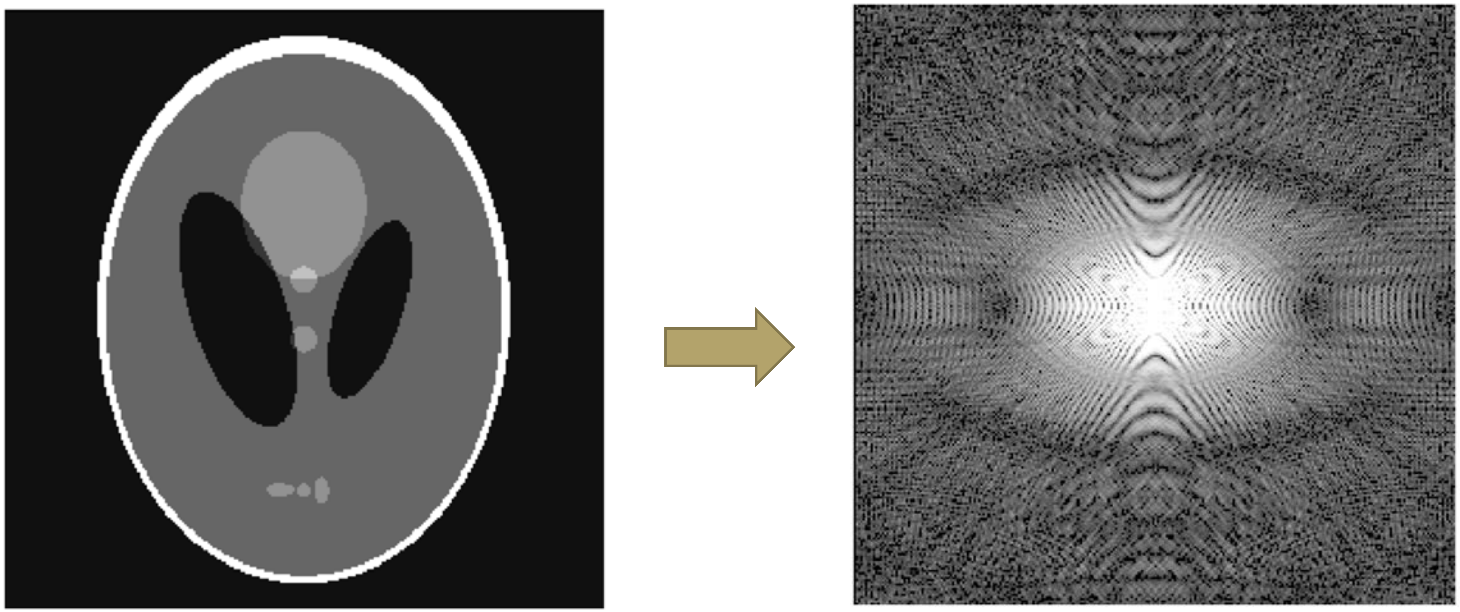


Uniform undersampling by a factor of 2

Magnetic Resonance Imaging (MRI)

Compressive Sensing MRI

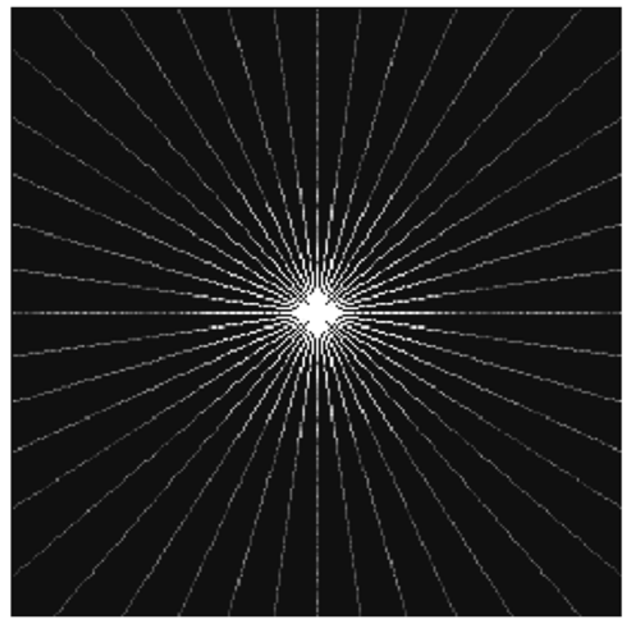
Fourier transform



Magnetic Resonance Imaging (MRI)

Compressive Sensing MRI

Fourier transform

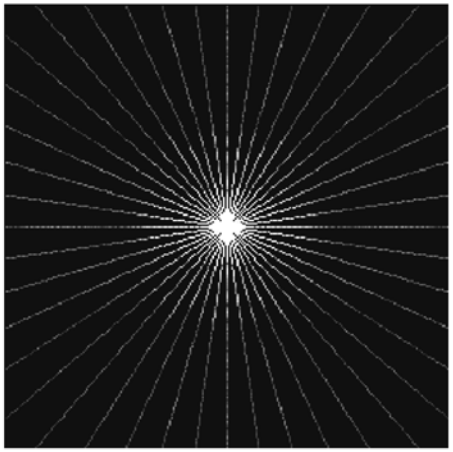


highly subsampled

Magnetic Resonance Imaging (MRI)

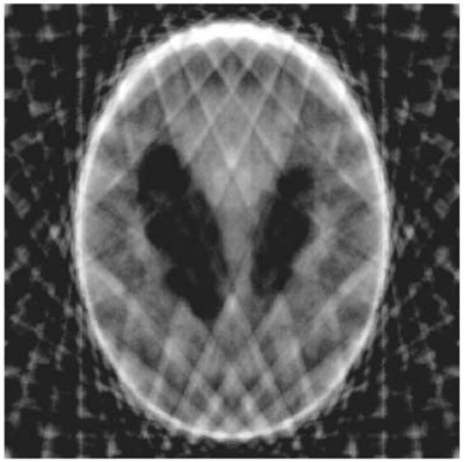
Compressive Sensing MRI

Fourier transform



highly subsampled

classical reconstruction



compressed sensing reconstruction

Magnetic Resonance Imaging (MRI)

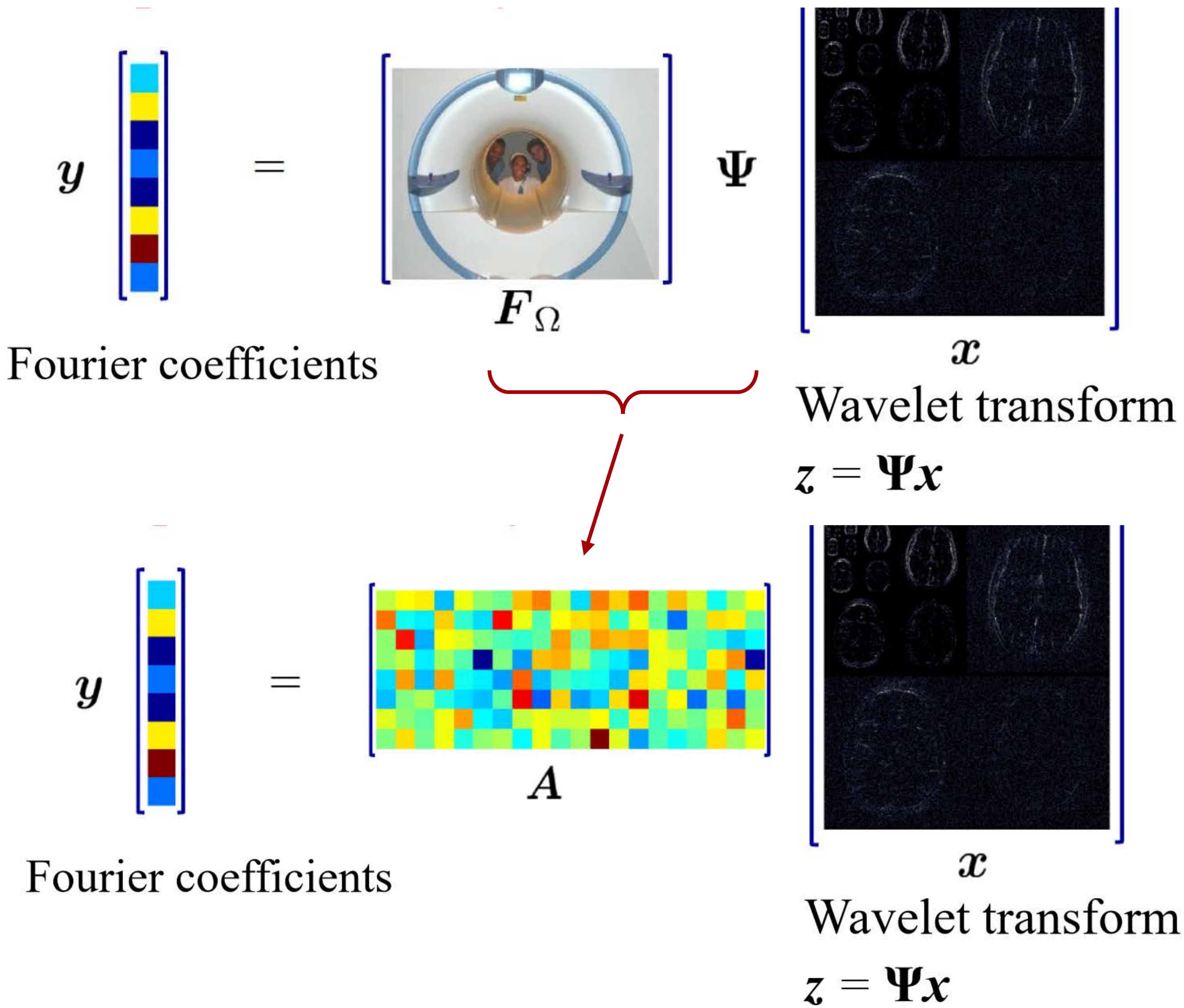


Image Denoising

Image denoising, image inpainting, image super-resolution
(reconstructing original picture from incomplete or corrupted observations)

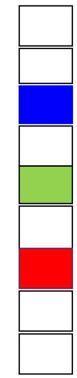
$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$

$\hat{\mathbf{y}}$

Denoised patch

$\widehat{\mathbf{A}}$

Learned dictionary



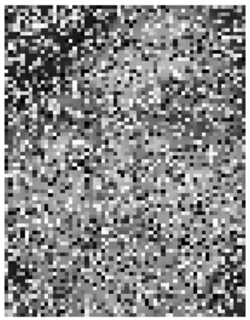
$\hat{\mathbf{x}}$

Sparse Coefficient

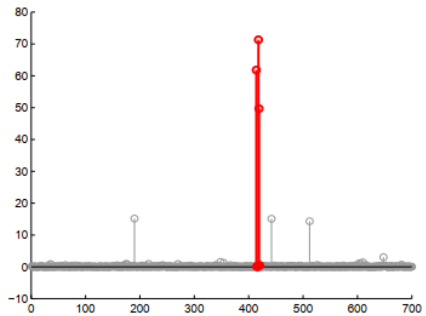
Results of Mairal, Saprio and Elad

Face Recognition

identify the subject under nuisance factors, such as occlusion



=



×

+

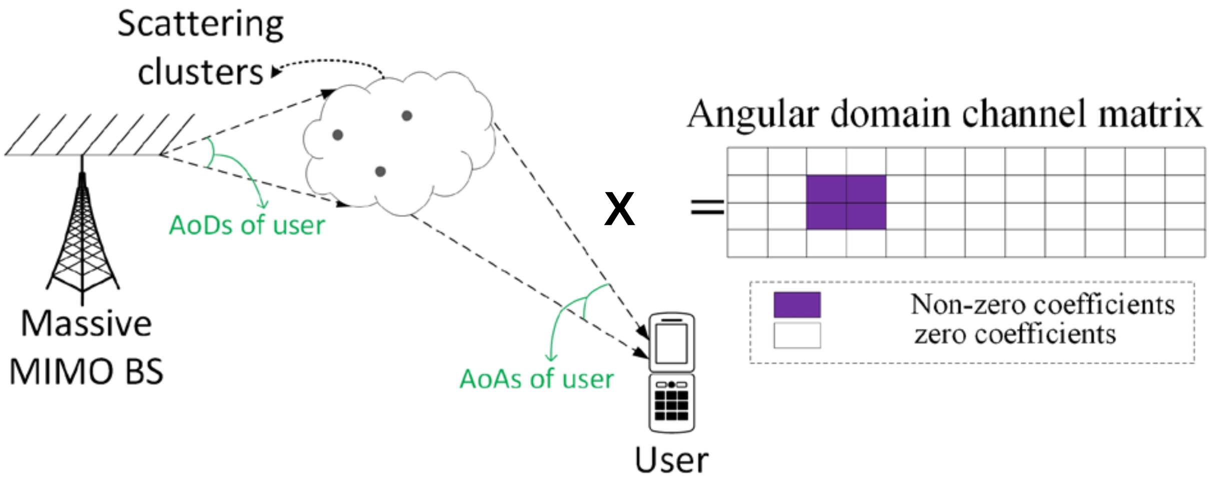
$\boldsymbol{y}$ $\boldsymbol{x}$ $\boldsymbol{A}$ $\boldsymbol{n}$

50% pixels are corrupted

Training dictionary

noise

Sparse Channel Estimation



$$\boldsymbol{y} = \mathbf{F}\,\boldsymbol{\Psi}\,\boldsymbol{x} + \boldsymbol{n}$$

Received Measurements at BS

Pilot Matrix

Angular Steering Basis

Sparse Angular Channel

Noise in Communication System

Compressive Sensing Theory

Problem Formulation

1. How to measure the sparsity?
    - Norm of vector
    - Statistical model
2. How to formulate the opt problem?

Performance Analysis

1. Noiseless case: uniqueness, correctness
2. Noisy case: stable recovery
    - distance to physical ground truth

Tools: High-dimensional Geometry and Statistics

Practical Solution

Practical: Efficient for high dimensional and large-scale problem

1. Convex → solver, iterative algorithm
2. Nonconvex → exhaustive search…

Compressive Sensing Theory

Problem Formulation

Nonconvex (NP hard)

convex relaxation

Convex Problem (efficient solutions)

Chapter 2

Performance Analysis

Behavior of the relaxed problem

Chapter 3

Practical Solution

- Interior Point Method:
  - ✓ efficient;
  - ✓ second order information of the obj;
  - ✓ intractable for high dimension problem
- First order optimization methods
- Acceleration techniques (improve the convergence rate)

Chapter 4

Noiseless CS

Noiseless Case $\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}$

Problem Formulation

ℓ_0 norm

$$\min \quad \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}.$$

ℓ_1 norm

Basis Pursuit

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x},$$

LASSO Regression

$$\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{x}\|_1 \leq k.$$

$$\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_1,$$

Performance Analysis

$\boldsymbol{A}$ satisfy the RIP property & $m = \varphi(k, n)$ → exact recovery

Practical Solution

Noiseless CS

Noiseless Case $y = Ax$

Problem Formulation

ℓ_1 norm

Basis Pursuit

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x},$$

LASSO Regression

$$\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{x}\|_1 \leq k.$$

$$\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_1,$$

Practical Solution

1. Greedy Algorithms
   - Matching pursuit
   - OMP & variant of OMP
2. BP: Augmented Lagrange Multiplier(enforce equality)
3. Frank-Wolfe Methods
4. LASSO Regression:
   - Subgradient method: slow convergence
   - Proximal Gradient: faster convergence rate

Noisy CS

Noisy Case $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}$

Problem Formulation

Basis Pursuit Denoising

ℓ_1 norm

$$\min \; \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2 \le \varepsilon.$$

LASSO

$$\min \; \lambda \|\boldsymbol{x}\|_1 + \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 .$$

Performance Analysis

Estimation error bound: $\|\hat{\boldsymbol{x}} - \boldsymbol{x}_0\|_2^2$

- Deterministic Noise: $\boldsymbol{A}$ satisfy the RIP
- Random Noise: $\boldsymbol{A}$ is i.i.d. Gaussian random matrix

Practical Solution

1. Greedy Algorithms
   - Matching pursuit
   - OMP
2. LASSO: iterative Shrinkage-Thresholding Algorithm (ISTA),FISTA
3. **Bayesian Inference-Based Algorithm**

Structured CS

$$\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x} + \boldsymbol{n}$$

- **Measurement Matrix:**
  - Well-behaved
    - Satisfies RIP, i.i.d. Gaussian, partial orthogonal.
  - Ill-conditioned
    - Non-zero mean, column-correlated, sparse, has a general form.
    - May contain uncertain parameters

- **Structure of the Sparse Signal:**
  - Support of x
    - Block sparse, Markov model, Markov tree model.
  - Values of large coefficients
    - Correlations, smooth transitions.
  - Improve the CS performance

Bayesian Inference-Based Algorithm

1. VBI or SBI
2. AMP (OAMP,VAMP,GAMP,Turbo-AMP, Turbo-OAMP)
3. Turbo-VBI

Chapter 5 and Chapter 6

Going Beyond Sparsity

Low Rank Matrix Recovery

Predict unseen ratings

Low Rank Matrix Recovery

Can we infer the missing entries?

$$\begin{bmatrix} \checkmark & ? & ? & ? & \checkmark & ? \\ ? & ? & \checkmark & \checkmark & ? & ? \\ \checkmark & ? & ? & \checkmark & ? & ? \\ ? & ? & \checkmark & ? & ? & \checkmark \\ \checkmark & ? & ? & ? & ? & ? \\ ? & \checkmark & ? & ? & \checkmark & ? \\ ? & ? & \checkmark & \checkmark & ? & ? \end{bmatrix}$$

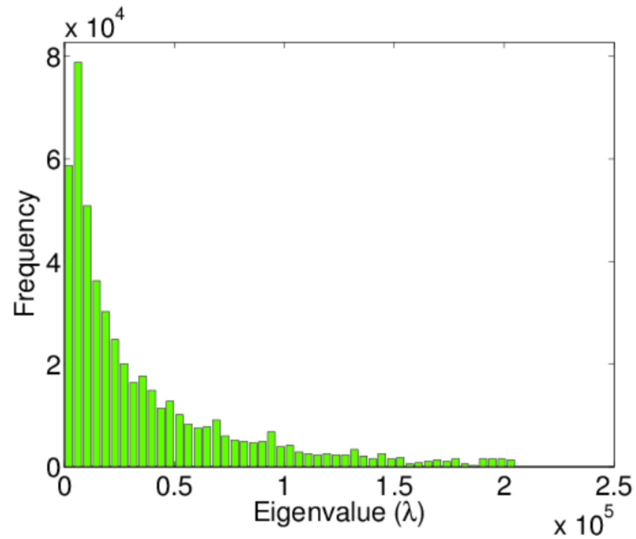
- Underdetermined system (more unknowns than revealed entries)
- Seems hopeless

$$\underset{\text{Observed ratings}}{\boldsymbol{Y}} = \mathcal{P}_\Omega \left[\underset{\text{Complete ratings}}{\boldsymbol{X}} \right]$$

Low Rank Matrix Recovery

What if unknown matrix has structures?

The ratings of distinct users (or distinct products) are correlated, and hence the target matrix $\boldsymbol{X}$ is low-rank



$$\begin{array}{ll} \min & \operatorname{rank}(\boldsymbol{X}) \\ \text{subject to} & \mathcal{A}[\boldsymbol{X}] = \boldsymbol{y}. \end{array}$$

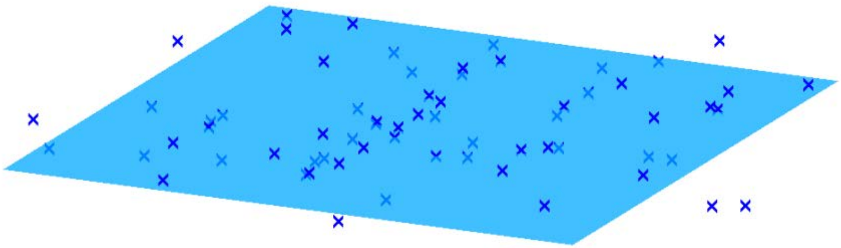
Principal Component Analysis

Project the high-dimensional random vector $\boldsymbol{y}$ onto much fewer directions.

$$\boldsymbol{Y} \doteq [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n] \in \mathbb{R}^{m \times n}$$

$$\min_{\boldsymbol{X}} \|\boldsymbol{Y} - \boldsymbol{X}\|_2^2 \quad \text{subject to} \quad \operatorname{rank}(\boldsymbol{X}) \leq d.$$

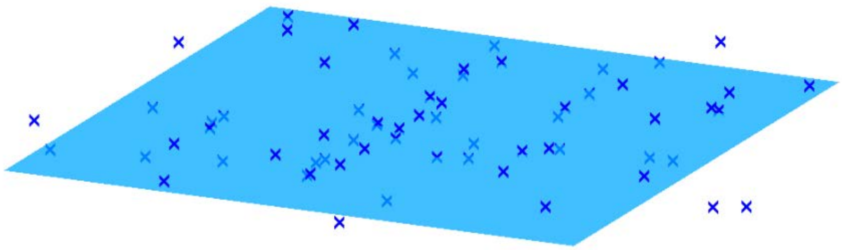
$$\boldsymbol{Y} \doteq [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n] \in \mathbb{R}^{m \times n}$$



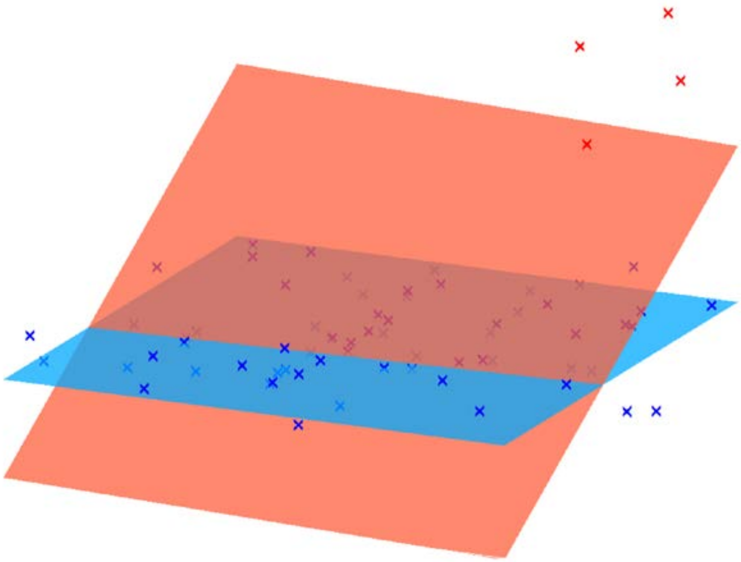
Low rank matrix X

Principal Component Analysis

$$\boldsymbol{Y} \doteq [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n] \in \mathbb{R}^{m \times n}$$



Low rank matrix X



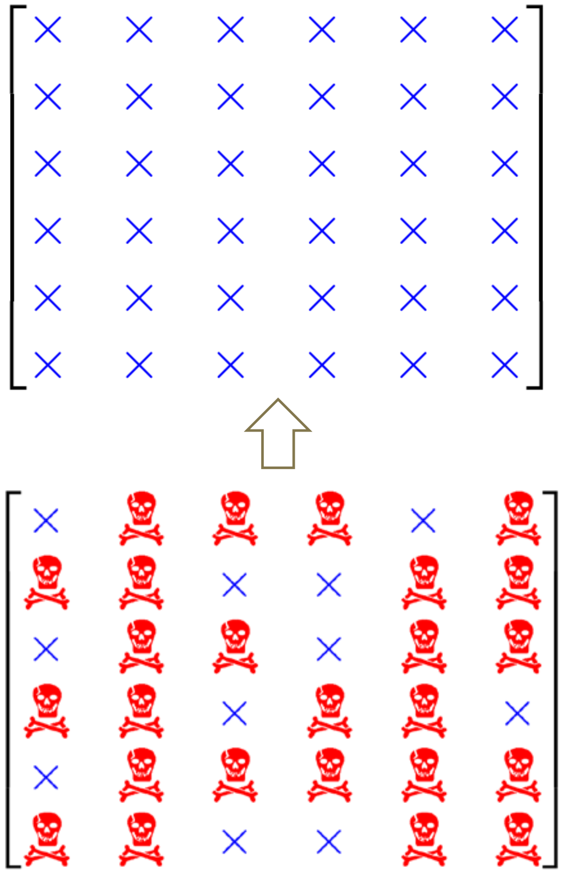
Robust Principal Component Analysis

Also called Sparse Low-rank Seperation (robust PCA)

Recover low-dimensional structure from corrupted data

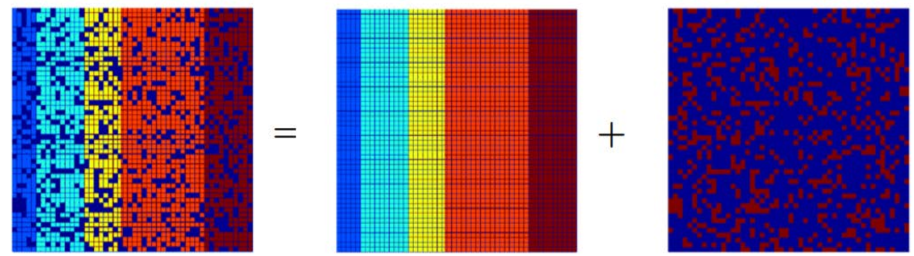
$$\boldsymbol{Y} = \boldsymbol{L} + \boldsymbol{S}$$

- $\boldsymbol{Y}$: data matrix (observed)
- $\boldsymbol{L}$: low-rank component (unobserved)
- $\boldsymbol{S}$: sparse outliers (unobserved)

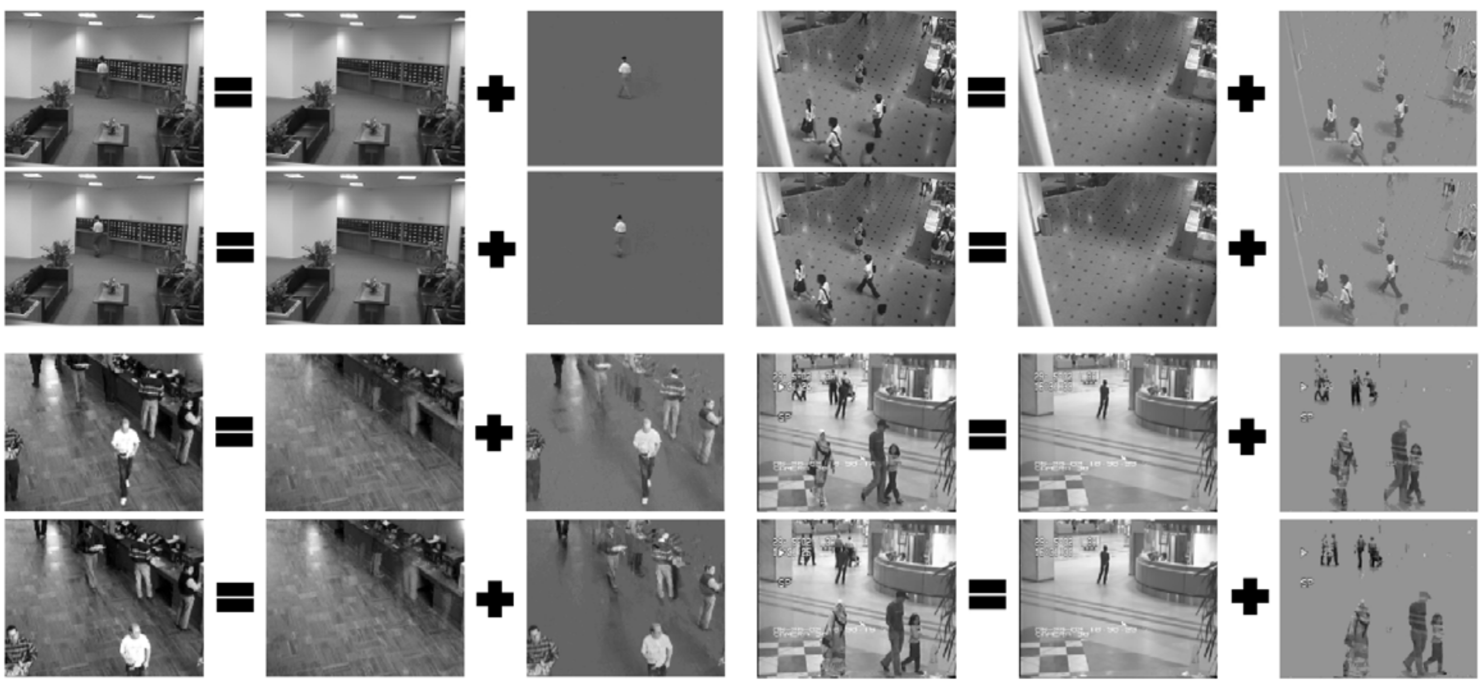


Robust Principal Component Analysis

$$
\begin{array}{ll}
\text{minimize} & \|L\|_* + \lambda \|S\|_1 \\
\text{subject to} & L + S = Y.
\end{array}
$$



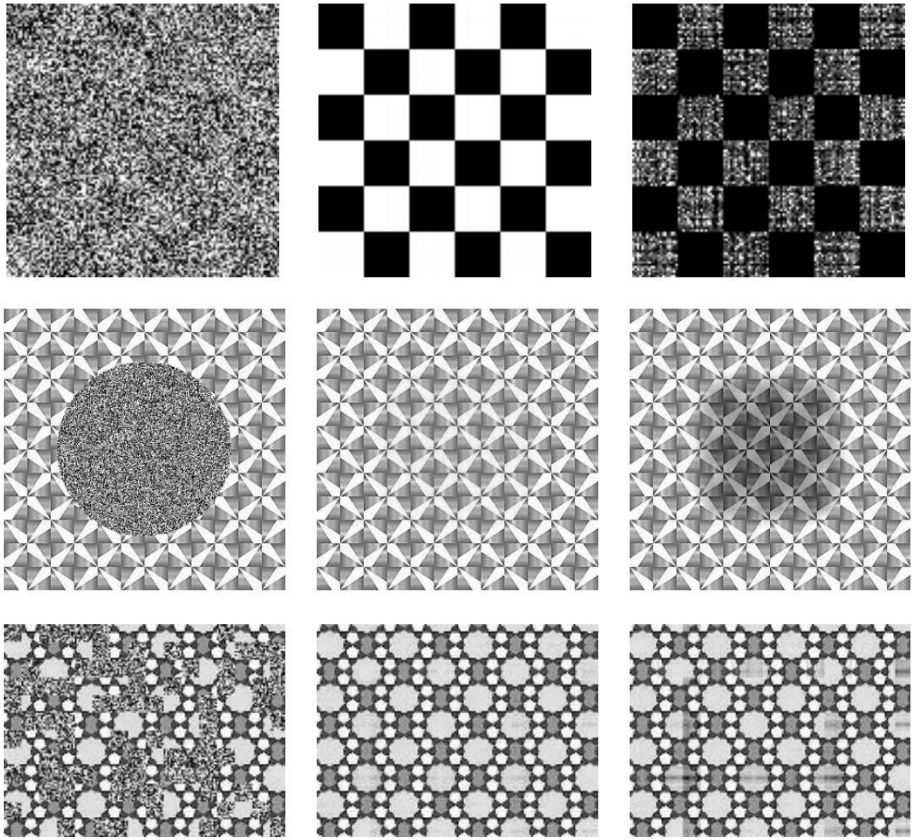
Example: separation of background (low-rank) and foreground (sparse) in videos



Sparse PCA

Recover a matrix which is simultaneously sparse and low-rank

$$\hat{\boldsymbol{X}}(\boldsymbol{Y}) = \arg\min_{\boldsymbol{X}} \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{X}\|_F^2 + \lambda_1\|\boldsymbol{X}\|_1 + \lambda_2\|\boldsymbol{X}\|_*.$$



(a) Input image (b) Sparse + Low-rank (c) Low-rank only

regular texture repairing

Nonlinear CS

1. Nonlinear measurements $\boldsymbol{y} = f(\boldsymbol{Ax})$

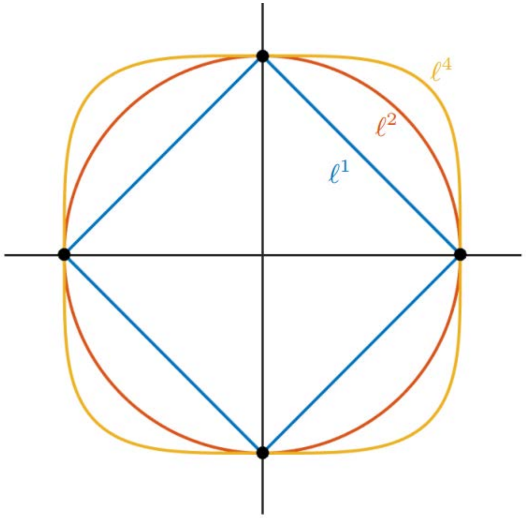
$$\begin{cases} \boldsymbol{y} = |\boldsymbol{Ax}| & \text{Phase retrieval} \\ \boldsymbol{y} = Quantization(\boldsymbol{Ax}) \\ \boldsymbol{y} = Quadratic(\boldsymbol{Ax}) \end{cases}$$

2. Nonlinear Models $\boldsymbol{Y} = \boldsymbol{AX}$

Dictionary Learning (sparse matrix factorization, bilinear CS)

$$\boldsymbol{Y} \approx \underbrace{\boldsymbol{A}_o}_{\text{dictionary}} \underbrace{\boldsymbol{X}_o}_{\text{sparse coefficients}}$$

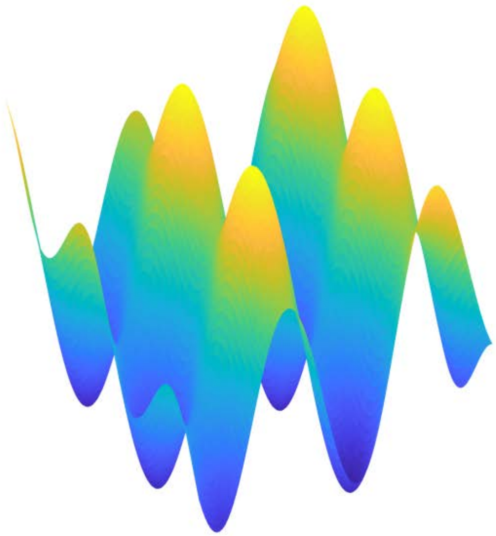
$$\max \|\boldsymbol{A}^*\bar{\boldsymbol{Y}}\|_4^4 \quad \text{subject to} \quad \boldsymbol{A}^* \in \mathsf{O}(n).$$



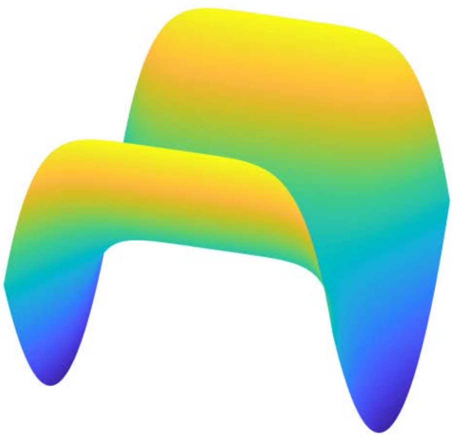
Nonconvex Problem

Extensions for Nonconvex Problem

Nonconvex (no global optimality or algorithm efficiency)



Spurious local minimizers



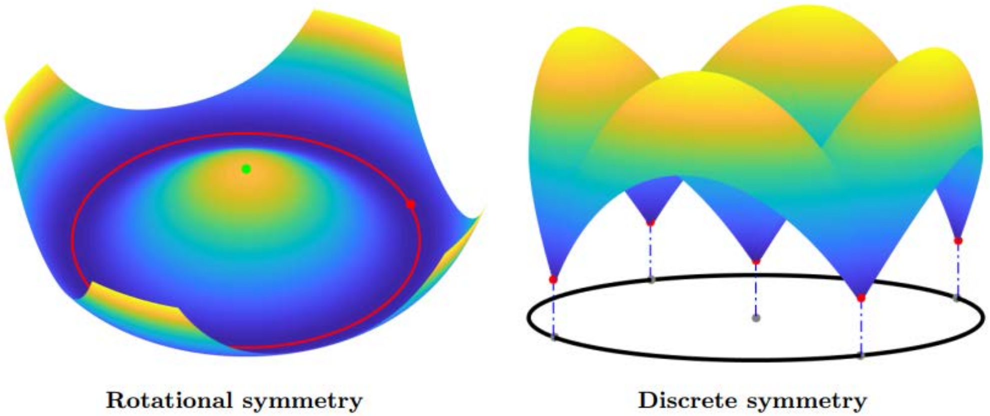
Flat saddle points

Descent methods can become trapped near local minimizers (left) or stagnate near flat saddle points (right).

Brute-force solution: random search or discretization of the space
1. Complexity (high dimensional space)
2. Convergence: Converge to critical points Or converge to some local minimizer

Extensions for Nonconvex Problem

Nonconvex (no global optimality or algorithm efficiency)



Symmetry and the *Global* Geometry of Optimization.
Model problems with continuous (left) and discrete (right) symmetry. For these particular problems, every local minimizer is global.

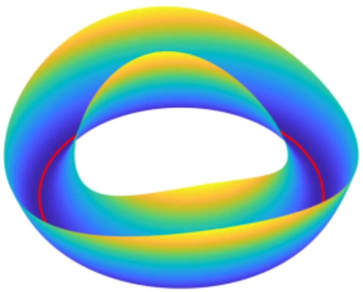
Simple and efficient algorithm→ gradient descent → global optimal solution

Extensions for Nonconvex Problem

Nonconvex Problems with Rotational Symmetries

Eigenspace Computation

Compute the principal subspace of a symmetric matrix.



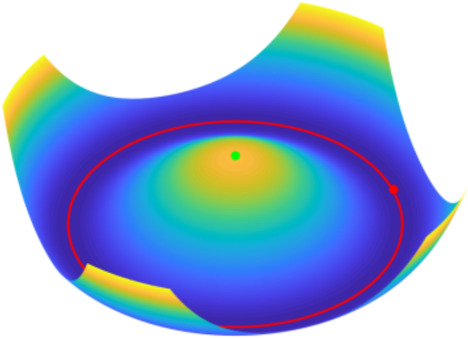
$$\min_{X^* X = I} -\tfrac{1}{2}\operatorname{trace}\left[X^* A X\right].$$

Symmetry: $X \mapsto XR$

$\mathbb{G} = \mathsf{O}(r)$

Generalized Phase Retrieval

Recover a complex vector x_o from magnitude measurements $y = |Ax_o|$.



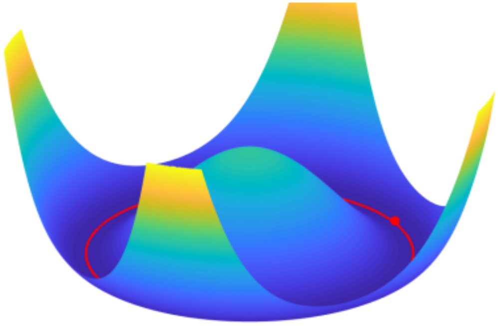
$$\min_{x} \tfrac{1}{2}\left\| y^2 - |Ax|^2 \right\|_2^2.$$

Symmetry: $x \mapsto x e^{i\phi}$

$\mathbb{G} = \mathbb{S}^1 \cong \mathsf{O}(2)$

Matrix Recovery

Recover a low-rank matrix $X = UV^*$ from incomplete/corrupted observations



$$\min_{U,V} \mathcal{L}(Y - \mathcal{A}[UV^*]) \; + \; \rho(U,V).$$

Symmetry: $(U, V) \mapsto (U\Gamma, V\Gamma^{-*})$

$\mathbb{G} = \mathsf{GL}(r)$ or $\mathbb{G} = \mathsf{O}(r)$

Thanks!

References

Image Superresolution

[1] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” vol. 19, no. 11, pp. 2861–2873, 2010.

Image Compression

[1] G. Wallace, “The JPEG still picture compression standard,” *Communications of the ACM*, vol. 34, no. 4, pp. 30-44, 1991.

[2] D. Taubman and M. Marcellin, *JPEG 2000: Image Compression Fundamentals, Standards and Practice*. Norwell, MA: Kluwer Academic Publishers, 2001.

Image Denoising

[1] J. Mairal, M. Elad, and G. Sapiro, \Sparse representation for color image restoration," *Image Processing, IEEE Transactions on*, vol. 17, no. 1, pp. 53-69, 2008.

Face Recognition

[1] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, \Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 200.

References

LASSO regression
[1] R. Tibshirani, \Regression shrinkage and selection via the LASSO," *Journal of the Royal Statistical Society B*, vol. 58, no. 1, pp. 267-288, 1996.

Basis Pursuit
[1] S. Chen, D. Donoho, and M. Saunders, \Atomic decomposition for basis pursuit,“ *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33-61, 1998.

OMP
[1] J. Tropp and A. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655-4666, 2007.
[2] D. Needell and J. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301-321, 2009.

References

ISTA

[1] Daubechies, M. Defrise, and C. Mol, “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Math*, vol. 57, pp. 1413-1457, 2004.

[2] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Science*, vol. 2, no. 1, pp. 183-202, 2009.

ALM

[1] M. R. Hestenes, “Multiplier and gradient methods," *Journal of optimization theory and applications*, vol. 4, no. 5, pp. 303-320, 1969.

[2] R. T. Rockafellar, “The multiplier method of Hestenes and Powell applied to convex programming," *Journal of Optimization Theory and Applications*, vol. 12, no. 6, pp. 555-562, 1973.

References

AMP:
[1] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” Proceedings of the National Academy of Sciences, vol. 106, no. 45, pp. 18 914–18 919, 2009.

OAMP:
[1] J. Ma and L. Ping, “Orthogonal AMP,” IEEE Access, vol. 5, pp. 2020–2033, 2017 .

Turbo-AMP:
[1] P. Schniter, “Turbo reconstruction of structured sparse signals,” in 2010 44th Annual Conference on Information Sciences and Systems (CISS), March 2010, pp. 1–6.
[2] S. Som and P. Schniter, “Compressive imaging using approximate message passing and a Markov-tree prior,” *IEEE transactions on signal processing*, vol. 60, no. 7, pp. 3439–3448, 2012.

SBL:
[1] M. E. Tipping, “Sparse Bayesian learning and the relevance vector machine,” Journal of machine learning research, vol. 1, no. Jun, pp. 211–244, 2001.

References

VBI:

[1] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, “The variational approximation for Bayesian inference,” IEEE Signal Processing Magazine, vol. 25, no. 6, pp. 131–146, 2008.

Dictionary Learning

[1] Y. Zhai, Z. Yang, Z. Liao, J. Wright, and Y. Ma, “Complete dictionary learning via ‘4-norm maximization over the orthogonal group," *Journal of Machine Learning Research*, 2020.

Phase Retrivial

[1] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, “Phase retrieval with application to optical imaging: a contemporary overview,“ *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87-109, 2015.

[2] E. J. Cand`es, Y. C. Eldar, T. Strohmer, and V. Voroninski, “Phase retrieval via matrix completion," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, 2013.

[3] E. J. Candes, X. Li, and M. Soltanolkotabi, “Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985-2007, 2015.

References

Phase Retrivial
[4] J. Sun, Q. Qu, and J. Wright, “A geometric analysis of phase retrieval," *Foundations of Computational Mathematics*, vol. 18, no. 5, pp. 1131-1198, 2018.
[5] A. Fannjiang and T. Strohmer, “The numerics of phase retrieval," 2020.

Low rank factorization
[1] Y. Chi, Y. M. Lu, and Y. Chen, “Nonconvex optimization meets low-rank matrix factorization: An overview," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5239-5269, 2019.
[2] R. Ge, J. D. Lee, and T. Ma, “Matrix completion has no spurious local minimum,“ *arXiv preprint arXiv:1605.07272*, 2016.
[3] R. Ge, C. Jin, and Y. Zheng, “No spurious local minima in nonconvex low rank problems: A unified geometric analysis," in *Proceedings of the 34th International Conference on Machine Learning*, 2017, pp. 1233-1242.

References

Dictionary Learning
[1] D. Gilboa, S. Buchanan, and J. Wright, “Efficient dictionary learning with gradient descent," *ICML*, 2019.
[2] J. Sun, Q. Qu, and J. Wright, “Complete dictionary recovery over the sphere I: Overview and geometric picture,” *IEEE Transactions on Information Theory*, vol. 63, no. 2, 2017.
[3] ——,”Complete dictionary recovery over the sphere II: Recovery by Riemannian trust-region method," *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 885-914, 2017.