Signal recovery from incomplete data

Holger Rauhut
Lehrstuhl C für Mathematik (Analysis)
RWTH Aachen

Data Science: Theory and Applications
RWTH Aachen
October 26, 2015

European Research Council

RWTH AACHEN UNIVERSITY

Mathematics and Data Processing

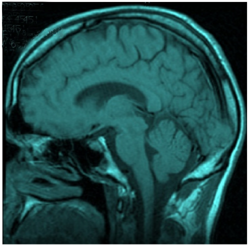
Data processing is a constant source for mathematical problems

Some examples

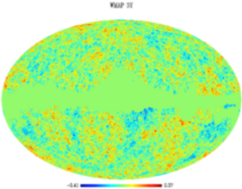
- The Shannon-Nyquist sampling theorem is at the basis of most electronic communication systems
- Computer tomography requires theory of Radon transforms
- Design of WLAN standards (OFDM) uses tools from time-frequency / harmonic analysis
  - Reducing the power consumption of base stations for mobile communication leads to very deep problems in harmonic analysis (peak-to-average power ratio (PAPR) problem)
- Machine learning techniques require a lot of mathematics, both for the design of algorithms as well as for their analysis
- Compressive Sensing: Signal reconstruction from small number of measurements

Goals: Mathematical analysis of basic data processing problems, fundamental limits, development and analysis of algorithms

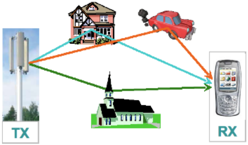
Data, Signal and Image Processing



Medical Imaging

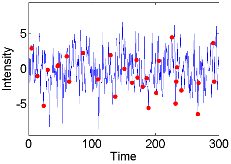


Cosmic Microwave Background

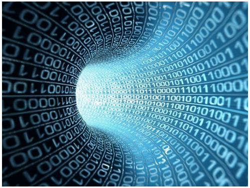


Wireless communication

Image Processing



A/D Conversion



Massive Internet Data

Compressive sensing

Reconstruction of signals from minimal amount of measured data (Candès, Romberg, Tao; Donoho 2004)

Key ingredients

- Compressibility / Sparsity (small complexity of relevant information)
- Efficient algorithms (convex optimization)
- Randomness (random matrices)

Compressive sensing

Reconstruction of signals from minimal amount of measured data (Candès, Romberg, Tao; Donoho 2004)

Key ingredients

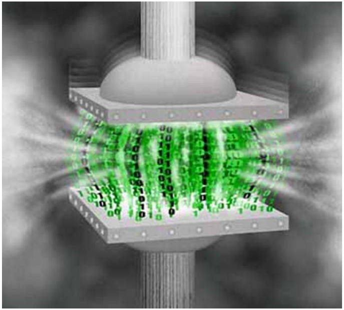
- Compressibility / Sparsity (small complexity of relevant information)
- Efficient algorithms (convex optimization)
- Randomness (random matrices)

Useful whenever it is difficult, expensive, time-consuming or impossible to obtain a large number of measurements.

Example applications:

- Magnetic Resonance Imaging
- Radar
- Wireless communications
- Astronomical signal processing
- (High-dimensional) Statistics
- Numerical solution of (High-dimensional) parametric PDEs

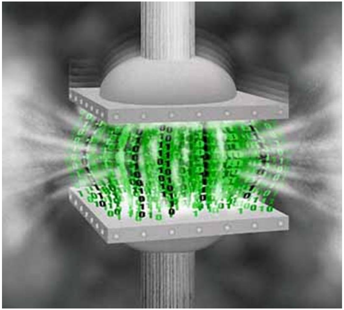
Sparsity / Compressibility



Data Compression

Most types of signals can be represented well by a sparse expansion, i.e., with only few nonzero coefficients in an appropriate basis (JPEG, MPEG, MP3 etc.).

Sparsity / Compressibility



Data Compression

Most types of signals can be represented well by a sparse expansion, i.e., with only few nonzero coefficients in an appropriate basis (JPEG, MPEG, MP3 etc.).

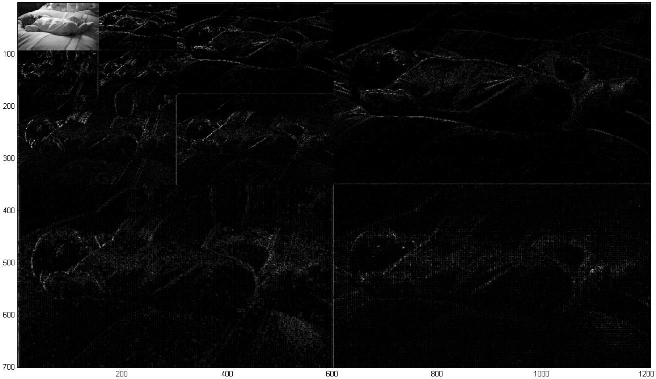
Compressive Sensing / Sparse Recovery

Sparse / Compressible signals can be recovered from only few linear measurements via efficient algorithms

Sparse Representations of Images

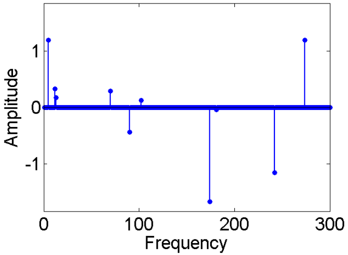
Niels

Wavelet Coefficients

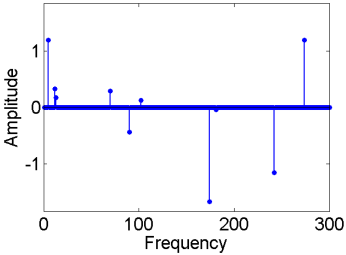


Wavelet compression

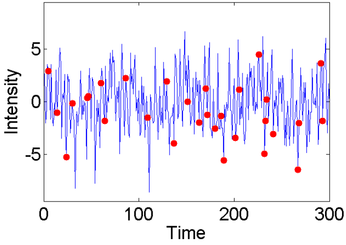
98% of wavelet coefficients are set to zero; only largest coefficients are retained.



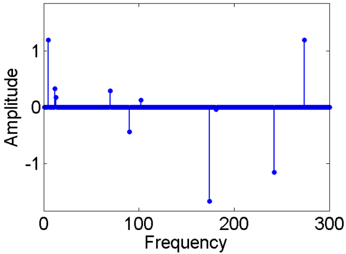
Fourier-Coefficients



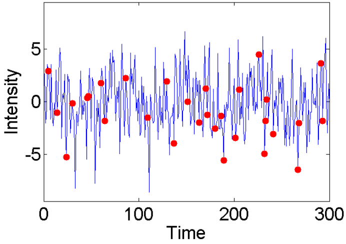
Fourier-Coefficients



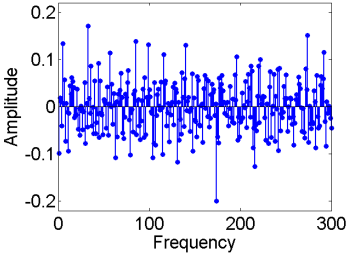
Time-Domain Signal with 30 Samples



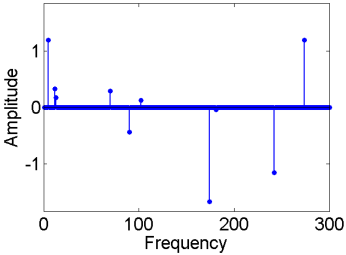
Fourier-Coefficients



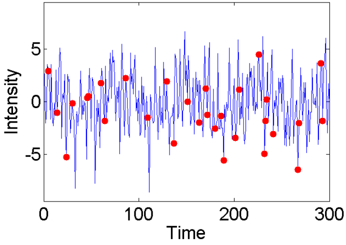
Time-Domain Signal with 30 Samples



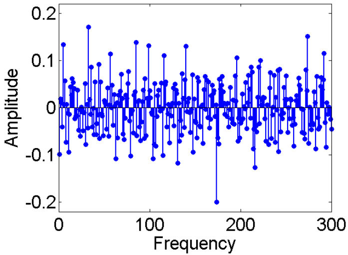
Traditional Reconstruction



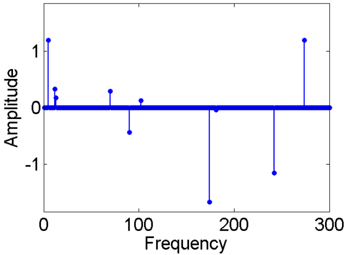
Fourier-Coefficients



Time-Domain Signal with 30 Samples



Traditional Reconstruction



Sparse Recovery Method

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Key finding of compressive sensing:
Recovery is possible if $\mathbf{x}$ belongs to a set of low complexity.

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Key finding of compressive sensing:
Recovery is possible if $\mathbf{x}$ belongs to a set of low complexity.

- Standard compressive sensing: Sparsity (small number of nonzero coefficients)

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Key finding of compressive sensing:
Recovery is possible if $\mathbf{x}$ belongs to a set of low complexity.

- Standard compressive sensing: Sparsity (small number of nonzero coefficients)
  - Refinements: block sparsity, joint sparsity, group sparsity

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Key finding of compressive sensing:
Recovery is possible if $\mathbf{x}$ belongs to a set of low complexity.

- Standard compressive sensing: Sparsity (small number of nonzero coefficients)
  - Refinements: block sparsity, joint sparsity, group sparsity
- Low rank matrix recovery (matrix completion)

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Key finding of compressive sensing:
Recovery is possible if $\mathbf{x}$ belongs to a set of low complexity.

- Standard compressive sensing: Sparsity (small number of nonzero coefficients)
  - Refinements: block sparsity, joint sparsity, group sparsity
- Low rank matrix recovery (matrix completion)
  - Phase retrieval

Mathematical formulation

Recover a vector $\mathbf{x} \in \mathbb{C}^N$ from underdetermined linear measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N},$$

where $m \ll N$.

Key finding of compressive sensing:
Recovery is possible if $\mathbf{x}$ belongs to a set of low complexity.

- Standard compressive sensing: Sparsity (small number of nonzero coefficients)
  - Refinements: block sparsity, joint sparsity, group sparsity
- Low rank matrix recovery (matrix completion)
  - Phase retrieval
- Low rank tensor recovery
  - Only partial results for tensor recovery available so far.
- Combinations of sparsity and low rank assumptions

Sparsity and Compressibility

- coefficient vector: $\mathbf{x} \in \mathbb{C}^N$, $N \in \mathbb{N}$
- support of $\mathbf{x}$: $\operatorname{supp} \mathbf{x} := \{j, x_j \neq 0\}$
- s- sparse vectors: $\|\mathbf{x}\|_0 := |\operatorname{supp} \mathbf{x}| \leq s$.

Sparsity and Compressibility

- coefficient vector: $\mathbf{x} \in \mathbb{C}^N$, $N \in \mathbb{N}$
- support of $\mathbf{x}$: $\operatorname{supp} \mathbf{x} := \{j, x_j \neq 0\}$
- s- sparse vectors: $\|\mathbf{x}\|_0 := |\operatorname{supp} \mathbf{x}| \leq s$.

s-term approximation error

$$\sigma_s(\mathbf{x})_q := \inf\{\|\mathbf{x} - \mathbf{z}\|_q, \mathbf{z} \text{ is } s\text{-sparse}\}, \quad 0 < q \leq \infty.$$

$\mathbf{x}$ is called compressible if $\sigma_s(\mathbf{x})_q$ decays quickly in s.

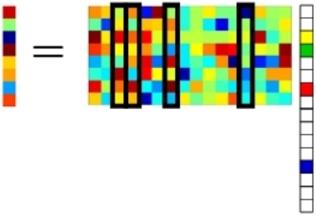
Here $\|\mathbf{x}\|_q = (\sum_{j=1}^N |x_j|^q)^{1/q}$

Compressive Sensing Problem

Reconstruct an s-sparse vector $\mathbf{x} \in \mathbb{C}^N$ (or a compressible vector) from its vector $\mathbf{y}$ of m measurements

$$\mathbf{y} = A\mathbf{x}, \qquad A \in \mathbb{C}^{m \times N}.$$

Interesting case: $s < m \ll N$.



Preferably fast reconstruction algorithm!

ℓ_1-minimization

ℓ_0-minimization is NP-hard:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}.$$

ℓ_1-minimization

ℓ_0-minimization is NP-hard:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}.$$

ℓ_1 minimization

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

ℓ_1-minimization

ℓ_0-minimization is NP-hard:

$$\min_{\mathbf{x}\in\mathbb{C}^N} \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}.$$

ℓ_1 minimization

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

Convex relaxation of ℓ_0-minimization problem.

Efficient minimization methods available.

ℓ_1-minimization

ℓ_0-minimization is NP-hard:

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}.$$

ℓ_1 minimization

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

Convex relaxation of ℓ_0-minimization problem.

Efficient minimization methods available.

Alternatives:
Greedy Algorithms (Matching Pursuits)
Iterative hard thresholding
Iteratively reweighted least squares

Mathematical Questions

- Which $m \times N$ matrices A are suitable?
- How many measurements m (in terms of sparsity s and signal length N) are needed for recovery?

Mathematical Questions

- Which $m \times N$ matrices A are suitable?
- How many measurements m (in terms of sparsity s and signal length N) are needed for recovery?

So far only random matrices are known to work provably well for sparse recovery.

Open to provide deterministic matrices A with rigorous recovery guarantees in the optimal parameter regime.

A typical result in compressive sensing

For a draw of a Gaussian random matrix $A \in \mathbb{R}^{m \times N}$ an s-sparse vector $x \in \mathbb{R}^N$ can be recovered exactly via ℓ_1-minimization (and other algorithms) with high probability from $y = Ax$ provided

$$m \geq Cs \ln(eN/s).$$

Bound optimal;

A typical result in compressive sensing

For a draw of a Gaussian random matrix $A \in \mathbb{R}^{m \times N}$ an s-sparse vector $x \in \mathbb{R}^N$ can be recovered exactly via ℓ_1-minimization (and other algorithms) with high probability from $y = Ax$ provided

$$m \geq Cs \ln(eN/s).$$

Bound optimal; Recovery stable under passing to approximately sparse vectors and under adding noise on the measurements.

A typical result in compressive sensing

For a draw of a Gaussian random matrix $A \in \mathbb{R}^{m \times N}$ an s-sparse vector $x \in \mathbb{R}^N$ can be recovered exactly via ℓ_1-minimization (and other algorithms) with high probability from $y = Ax$ provided

$$m \geq Cs \ln(eN/s).$$

Bound optimal; Recovery stable under passing to approximately sparse vectors and under adding noise on the measurements.

Similar results for certain structured random matrices:

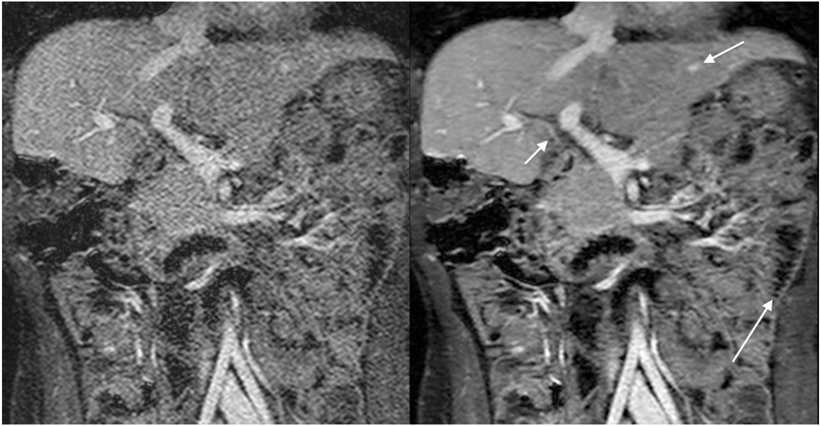
- Randomly sampled Fourier transform of sparse vectors (Candès, Tao '06; Rudelson, Vershynin '08; Rauhut '07, '10, '14; Bourgain '14; Haviv, Regev '15)

$$m \geq Cs \log^2(s) \log(N)$$

- Subsampled random convolution of sparse vectors (Rauhut '09, '10; Rauhut, Romberg Tropp '12; Krahmer, Mendelson, Rauhut '14)

$$m \geq Cs \log^2(s) \log^2(N)$$

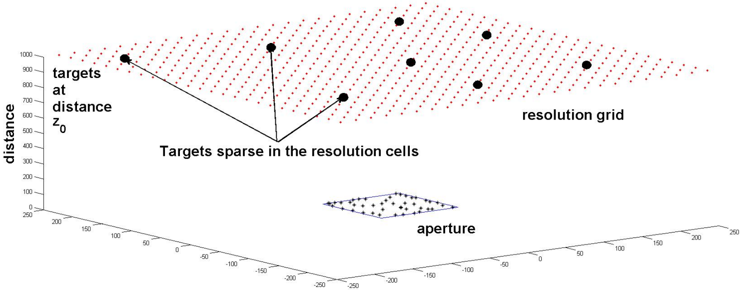
Application: Magnetic Resonance Imaging



Comparison of a traditional MRI reconstruction (left) and a compressive sensing reconstruction (right). Acquisition accelerated by a factor of 7.2 by random subsampling of the frequency domain

Image courtesy of Michael Lustig and Shreyas Vasanawala, Stanford University

Remote sensing (radar imaging)



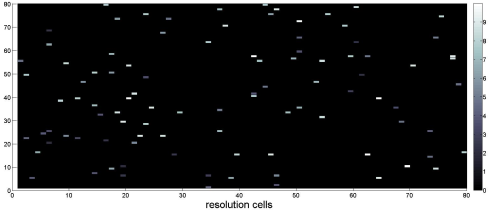
n antenna elements on square $[0, B]^2$ in plane $z = 0$.
Targets in the plane $z = z_0$ on grid of resolution cells
$r_j \in [-L, L]^2 \times \{z_0\}$, $j = 1, \ldots, N$ with mesh size h.

$\mathbf{x} \in \mathbb{C}^N$: vector of reflectivities in resolution cells $(r_j)_{j=1,\ldots,N}$.

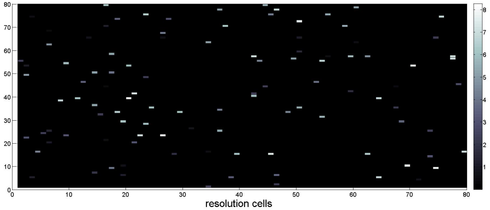
Often sparse scene!
$m = n^2$ with n antennas

Reconstruction via ℓ_1-minimization

Sparse scene (sparsity $s = 100$, 6400 grid points):



Reconstruction ($n = 30$ antennas, 900 noisy measurements, SNR 20dB)



Recovery if $m \geq Cs \log^2(N)$ (Hügel, Rauhut, Strohmer 2014)

Low Rank Matrix Recovery

Recover $X \in \mathbb{C}^{n_1 \times n_2}$ of low rank from $y = \mathcal{A}(X) \in \mathbb{C}^m$, where $m \ll n_1 n_2$!

Low Rank Matrix Recovery

Recover $X \in \mathbb{C}^{n_1 \times n_2}$ of low rank from $y = \mathcal{A}(X) \in \mathbb{C}^m$, where $m \ll n_1 n_2$!

Rank minimization problem $\min_{Z:\mathcal{A}(Z)=y} \operatorname{rank}(X)$ is NP-hard.

Low Rank Matrix Recovery

Recover $X \in \mathbb{C}^{n_1 \times n_2}$ of low rank from $y = \mathcal{A}(X) \in \mathbb{C}^m$, where $m \ll n_1 n_2$!

Rank minimization problem $\min_{Z:\mathcal{A}(Z)=y} \operatorname{rank}(X)$ is NP-hard.

Observation: $\operatorname{rank}(X) = \|\sigma(X)\|_0$ where $\sigma(X)$ is vector of singular values of X

Nuclear norm minimization

$$\min \|X\|_* \quad \text{subject to } \mathcal{A}(X) = y$$

with $\|X\|_* = \sum_\ell \sigma_\ell(X)$.

Recovery of rank r matrix X from m subgaussian random measurements (Fazel, Parrilo, Recht; Candès, Plan) when

$$m \geq Cr(n_1 + n_2).$$

Low Rank Matrix Recovery

Recover $X \in \mathbb{C}^{n_1 \times n_2}$ of low rank from $y = \mathcal{A}(X) \in \mathbb{C}^m$, where $m \ll n_1 n_2$!
Rank minimization problem $\min_{Z:\mathcal{A}(Z)=y} \operatorname{rank}(X)$ is NP-hard.

Observation: $\operatorname{rank}(X) = \|\sigma(X)\|_0$ where $\sigma(X)$ is vector of singular values of X

Nuclear norm minimization

$$\min \|X\|_* \quad \text{subject to } \mathcal{A}(X) = y$$

with $\|X\|_* = \sum_\ell \sigma_\ell(X)$.

Recovery of rank r matrix X from m subgaussian random measurements (Fazel, Parrilo, Recht; Candès, Plan) when

$$m \geq Cr(n_1 + n_2).$$

Subgaussian assumption can be relaxed: four finite moments are sufficient (Kabanava, Kueng, Rauhut, Terstiege '15).

Matrix completion

Complete missing entries of a low rank matrix:

$$\begin{pmatrix} ? & 10 & ? & 2 & ? & ? \\ 3 & ? & ? & ? & 3 & ? \\ ? & ? & 14 & ? & ? & 14 \\ ? & 15 & 6 & ? & ? & ? \\ 6 & ? & 4 & ? & 6 & 4 \end{pmatrix}$$

Recovery via nuclear norm minimization under certain assumptions on the singular vectors of X when

$$m \geq Cr(n_1 + n_2) \ln^2(n_1 + n_2).$$

Candès, Recht, Gross, ...
Application: Consumer taste prediction (Netflix prize),...

Quantum state tomography

The state of a (finite-dimensional) quantum system is described by symmetric positive semidefinite matrix $A \in \mathbb{C}^{n \times n}$ with $\operatorname{tr} A = 1$.

Quantum measurements often of the form

$$y_j = \mathcal{A}(X)_j := a_j^* X a_j = \operatorname{tr}(X a_j a_j^*), \quad j = 1, \dots, m$$

Pure states: $\operatorname{rank}(X) = 1$
mixed states: $\operatorname{rank}(X) = r \ll n$

Quantum state tomography

The state of a (finite-dimensional) quantum system is described by symmetric positive semidefinite matrix $A \in \mathbb{C}^{n \times n}$ with $\operatorname{tr} A = 1$.

Quantum measurements often of the form

$$y_j = \mathcal{A}(X)_j := a_j^* X a_j = \operatorname{tr}(X a_j a_j^*), \quad j = 1, \dots, m$$

Pure states: $\operatorname{rank}(X) = 1$
mixed states: $\operatorname{rank}(X) = r \ll n$

Recovery via nuclear norm minimization (Kabanava, Kueng, Rauhut, Terstiege '15):

- $a_j \in \mathbb{R}^N$ independent Gaussian random vectors:

$$m \geq Crn$$

- $a_j \in \mathbb{C}^N$ chosen at random from a (weighted, approximative) 4-design:

$$m \geq Crn \log(n)$$

Applications: quantum optical circuits, quantum computing?

Reflections on Data Science

What is data science?

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.
- Extracting (relevant) information from data.

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.
- Extracting (relevant) information from data.
- Applications: what can be done with (large, small) data sets?

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.
- Extracting (relevant) information from data.
- Applications: what can be done with (large, small) data sets?

Contributions of my group:

- Analysis and development of data recovery algorithms

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.
- Extracting (relevant) information from data.
- Applications: what can be done with (large, small) data sets?

Contributions of my group:

- Analysis and development of data recovery algorithms
- Fundamental performance limits (upper and lower bounds)

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.
- Extracting (relevant) information from data.
- Applications: what can be done with (large, small) data sets?

Contributions of my group:

- Analysis and development of data recovery algorithms
- Fundamental performance limits (upper and lower bounds)
- Methods: Convex optimization, approximation theory, (structured) random matrices (via asymptotic geometric analysis), computational harmonic analysis

Reflections on Data Science

What is data science?

- Science of measuring, processing, manipulating, transmitting and understanding of data.
- Extracting (relevant) information from data.
- Applications: what can be done with (large, small) data sets?

Contributions of my group:

- Analysis and development of data recovery algorithms
- Fundamental performance limits (upper and lower bounds)
- Methods: Convex optimization, approximation theory, (structured) random matrices (via asymptotic geometric analysis), computational harmonic analysis

Thank you!

Questions?