Compressive Sensing and Structured Random Matrices

Holger Rauhut
Hausdorff Center for Mathematics
& Institute for Numerical Simulation
University of Bonn

Probability & Geometry in High Dimensions
Marne la Vallée
May 19, 2010

Overview

- Compressive Sensing
- Random Sampling in Bounded Orthonormal Systems
- Partial Random Circulant Matrices
- Random Gabor Frames

- Many types of signals, images are sparse, or can be well-approximated by sparse ones.

Key Ideas of compressive sensing

- Many types of signals, images are sparse, or can be well-approximated by sparse ones.
- Question: Is it possible to recover such signals from only a small number of (linear) measurements, i.e., without measuring all entries of the signal?

- 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: $\|\mathbf{x}\|_0 := |\operatorname{supp} \mathbf{x}| \leq s$.

Sparse Vectors in Finite Dimension

- 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.

hausdorff CENTER FOR MATHEMATICS

Reconstruct a 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: $s < m \ll N$.

Underdetermined linear system of equations with a sparsity constraint.

Reconstruct a 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: $s < m \ll N$.

Underdetermined linear system of equations with a sparsity constraint.

Preferably we would like to have a fast algorithm that performs the reconstruction.

ℓ_0-minimization:

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

ℓ_0-minimization:

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

Problem: ℓ_0-minimization is NP hard!

ℓ_1 minimization:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 = \sum_{j=1}^{N} |x_j| \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

hausdorff CENTER FOR MATHEMATICS

ℓ_1 minimization:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 = \sum_{j=1}^{N} |x_j| \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

Convex relaxation of ℓ_0-minimization problem.

Efficient minimization methods available.

hausdorff CENTER FOR MATHEMATICS

Definition

The restricted isometry constant δ_s of a matrix $A \in \mathbb{C}^{m \times N}$ is defined as the smallest δ_s such that

$$(1 - \delta_s)\|\mathbf{x}\|_2^2 \leq \|A\mathbf{x}\|_2^2 \leq (1 + \delta_s)\|\mathbf{x}\|_2^2$$

for all s-sparse $\mathbf{x} \in \mathbb{C}^N$.

Requires that all s-column submatrices of A are well-conditioned.

RIP implies recovery by ℓ_1-minimization

Theorem (Candès, Romberg, Tao 2004 – Candès 2008 – Foucart, Lai 2009 – Foucart 2009)

Assume that the restricted isometry constant δ_{2s} of $A \in \mathbb{C}^{m\times N}$ satisfies

$$\delta_{2s} < \frac{2}{3 + \sqrt{7/4}} \approx 0.4627.$$

Then ℓ_1-minimization reconstructs every s-sparse vector $\mathbf{x} \in \mathbb{C}^N$ from $y = Ax$.

hausdorff CENTER FOR MATHEMATICS

Theorem (Candès, Romberg, Tao 2004 – Candès 2008 – Foucart, Lai 2009 – Foucart 2009)

Let $A \in \mathbb{C}^{m \times N}$ *with* $\delta_{2s} < \frac{2}{3+\sqrt{7/4}} \approx 0.4627$. *Let* $x \in \mathbb{C}^N$, *and assume that noisy data are observed,* $y = Ax + \eta$ *with* $\|\eta\|_2 \leq \sigma$. *Let* $x^{\#}$ *by the solution of*

$$\min_z \|z\|_1 \quad \textit{such that} \quad \|Az - y\|_2 \leq \sigma.$$

Then

$$\|x - x^{\#}\|_2 \leq C_1\sigma + C_2 \frac{\sigma_s(x)_1}{\sqrt{s}}$$

for constants $C_1, C_2 > 0$, *that depend only on* δ_{2s}.

hausdorff CENTER FOR MATHEMATICS

Open problem: Give explicit matrices $A \in \mathbb{C}^{m \times N}$ with small $\delta_{2s} \leq 0.46$ for "large" s.

Open problem: Give explicit matrices $A \in \mathbb{C}^{m \times N}$ with small $\delta_{2s} \leq 0.46$ for "large" s.

Goal: $\delta_s \leq \delta$, if

$$m \geq C_\delta s \log^\alpha(N),$$

for constants C_δ and α.

Deterministic matrices known, for which $m \geq C_\delta s^2$ suffices.

Open problem: Give explicit matrices $A \in \mathbb{C}^{m \times N}$ with small $\delta_{2s} \leq 0.46$ for "large" s.

Goal: $\delta_s \leq \delta$, if

$$m \geq C_\delta s \log^\alpha(N),$$

for constants C_δ and α.

Deterministic matrices known, for which $m \geq C_\delta s^2$ suffices.

Way out: consider random matrices.

Gaussian: entries of A independent standard normal distributed random rv

Bernoulli : entries of A independent Bernoulli ± 1 distributed rv

Theorem

Let $A \in \mathbb{R}^{m \times N}$ be a Gaussian or Bernoulli random matrix and assume

$$m \geq C\delta^{-2}(s \ln(N/s) + \ln(\varepsilon^{-1}))$$

for a universal constant $C > 0$. Then with probability at least $1 - \varepsilon$ the restricted isometry constant of $\frac{1}{\sqrt{m}}A$ satisfies $\delta_s \leq \delta$.

Consequence

Gaussian or Bernoulli matrices $A \in \mathbb{R}^{m\times}$ allow (stable) sparse recovery using ℓ_1-minimization with probability at least $1-\varepsilon = 1-\exp(-cm)$, $c = 1/(2C)$, provided

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

hausdorff
CENTER FOR MATHEMATICS

Gaussian or Bernoulli matrices $A \in \mathbb{R}^{m\times}$ allow (stable) sparse recovery using ℓ_1-minimization with probability at least $1 - \varepsilon = 1 - \exp(-cm)$, $c = 1/(2C)$, provided

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

No quadratic bottleneck!

Gaussian or Bernoulli matrices $A \in \mathbb{R}^{m\times}$ allow (stable) sparse recovery using ℓ_1-minimization with probability at least $1 - \varepsilon = 1 - \exp(-cm)$, $c = 1/(2C)$, provided

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

No quadratic bottleneck!

Bound is optimal as follows from bounds for Gelfand widths of ℓ_p^N-balls ($0 < p \leq 1$),
Kashin (1977), Gluskin – Garnaev (1984), Carl – Pajor (1988), Vybiral (2008), Foucart – Pajor – Rauhut – Ullrich (2010).

Why structure?

- Applications impose structure due to physical constraints, limited freedom to inject randomness.
- Fast matrix vector multiplies (FFT) in recovery algorithms, unstructured random matrices impracticable for large scale applications.
- Storage problems for unstructured matrices.

Bounded orthonormal systems (BOS)

$\mathcal{D} \subset \mathbb{R}^d$ endowed with probability measure ν.

$\psi_1, \ldots, \psi_N : \mathcal{D} \to \mathbb{C}$ function system on $\mathcal{D}$.

hausdorff CENTER FOR MATHEMATICS

Bounded orthonormal systems (BOS)

$\mathcal{D} \subset \mathbb{R}^d$ endowed with probability measure ν.
$\psi_1, \ldots, \psi_N : \mathcal{D} \to \mathbb{C}$ function system on $\mathcal{D}$.
Orthonormality

$$\int_{\mathcal{D}} \psi_j(t)\overline{\psi_k(t)} d\nu(t) = \delta_{j,k} = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

hausdorff CENTER FOR MATHEMATICS

Bounded orthonormal systems (BOS)

$\mathcal{D} \subset \mathbb{R}^d$ endowed with probability measure ν.
$\psi_1, \ldots, \psi_N : \mathcal{D} \to \mathbb{C}$ function system on $\mathcal{D}$.
Orthonormality

$$\int_{\mathcal{D}} \psi_j(t)\overline{\psi_k(t)} d\nu(t) = \delta_{j,k} = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Uniform bound in L^∞:

$$\|\psi_j\|_\infty = \sup_{t \in \mathcal{D}} |\psi_j(t)| \leq K \quad \text{for all } j \in [N].$$

hausdorff CENTER FOR MATHEMATICS

Bounded orthonormal systems (BOS)

$\mathcal{D} \subset \mathbb{R}^d$ endowed with probability measure ν.
$\psi_1, \ldots, \psi_N : \mathcal{D} \to \mathbb{C}$ function system on $\mathcal{D}$.
Orthonormality

$$\int_{\mathcal{D}} \psi_j(t) \overline{\psi_k(t)} d\nu(t) = \delta_{j,k} = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Uniform bound in L^∞:

$$\|\psi_j\|_\infty = \sup_{t \in \mathcal{D}} |\psi_j(t)| \leq K \quad \text{for all } j \in [N].$$

It always holds $K \geq 1$:

$$1 = \int_{\mathcal{D}} |\psi_j(t)|^2 d\nu(t) \leq \sup_{t \in \mathcal{D}} |\psi_j(t)|^2 \int_{\mathcal{D}} 1 d\nu(t) \leq K^2.$$

hausdorff CENTER FOR MATHEMATICS

Consider functions

$$f(t) = \sum_{k=1}^{N} x_k \psi_k(t), \quad t \in \mathcal{D}.$$

Consider functions

$$f(t) = \sum_{k=1}^{N} x_k \psi_k(t), \quad t \in \mathcal{D}.$$

f is called s-sparse if $\mathbf{x}$ is s-sparse.

Consider functions

$$f(t) = \sum_{k=1}^{N} x_k \psi_k(t), \quad t \in \mathcal{D}.$$

f is called s-sparse if $\mathbf{x}$ is s-sparse.
Sampling points $t_1, \ldots, t_m \in \mathcal{D}$. Sample values:

$$y_\ell = f(t_\ell) = \sum_{k=1}^{N} x_k \psi_k(t_\ell)\ , \quad \ell \in [m].$$

Consider functions

$$f(t) = \sum_{k=1}^{N} x_k \psi_k(t), \quad t \in \mathcal{D}.$$

f is called s-sparse if $\mathbf{x}$ is s-sparse.
Sampling points $t_1, \dots, t_m \in \mathcal{D}$. Sample values:

$$y_\ell = f(t_\ell) = \sum_{k=1}^{N} x_k \psi_k(t_\ell)\ , \quad \ell \in [m].$$

Sampling matrix $A \in \mathbb{C}^{m \times N}$ with entries

$$A_{\ell,k} = \psi_k(t_\ell), \qquad \ell \in [m], k \in [N].$$

hausdorff
CENTER FOR MATHEMATICS

Consider functions

$$f(t) = \sum_{k=1}^{N} x_k \psi_k(t), \quad t \in \mathcal{D}.$$

f is called s-sparse if $\mathbf{x}$ is s-sparse.
Sampling points $t_1, \ldots, t_m \in \mathcal{D}$. Sample values:

$$y_\ell = f(t_\ell) = \sum_{k=1}^{N} x_k \psi_k(t_\ell)\ , \quad \ell \in [m].$$

Sampling matrix $A \in \mathbb{C}^{m \times N}$ with entries

$$A_{\ell,k} = \psi_k(t_\ell), \qquad \ell \in [m], k \in [N].$$

Then

$$\mathbf{y} = A\mathbf{x}.$$

Problem: Reconstruct s-sparse f — equivalently $\mathbf{x}$ — from its sample values $\mathbf{y} = A\mathbf{x}$.

Problem: Reconstruct s-sparse f — equivalently $\mathbf{x}$ — from its sample values $\mathbf{y} = A\mathbf{x}$.

We consider ℓ_1-minimization as recovery method.

Behavior of A as measurement matrix?

Choose sampling points $t_1, \dots, t_\ell$ independently at random according to the measure ν, that is,

$$\mathbb{P}(t_\ell \in B) = \nu(B), \qquad \text{for all measurable } B \subset \mathcal{D}.$$

Choose sampling points $t_1, \ldots, t_\ell$ independently at random according to the measure ν, that is,

$$\mathbb{P}(t_\ell \in B) = \nu(B), \qquad \text{for all measurable } B \subset \mathcal{D}.$$

The sampling matrix A is then a structured random matrix.

Examples of Bounded Orthonormal Systems

Trigonometric System. $\mathcal{D} = [0, 1]$ with Lebesgue measure.

$$\psi_k(t) = e^{2\pi ikt}, \quad t \in [0, 1].$$

The trigonometric system is orthonormal with $K = 1$.

Examples of Bounded Orthonormal Systems

Trigonometric System. $\mathcal{D} = [0,1]$ with Lebesgue measure.

$$\psi_k(t) = e^{2\pi i k t}, \quad t \in [0,1].$$

The trigonometric system is orthonormal with $K = 1$. Take subset $\Gamma \subset \mathbb{Z}$ of cardinality N; trigonometric polynomials

$$f(t) = \sum_{k \in \Gamma} x_k \psi_k(t) = \sum_{k \in \Gamma} x_k e^{2\pi i k t}.$$

Examples of Bounded Orthonormal Systems

Trigonometric System. $\mathcal{D} = [0,1]$ with Lebesgue measure.

$$\psi_k(t) = e^{2\pi ikt}, \quad t \in [0,1].$$

The trigonometric system is orthonormal with $K = 1$. Take subset $\Gamma \subset \mathbb{Z}$ of cardinality N; trigonometric polynomials

$$f(t) = \sum_{k \in \Gamma} x_k \psi_k(t) = \sum_{k \in \Gamma} x_k e^{2\pi ikt}.$$

Samples $t_1, \dots, t_m$ are chosen indendepently and uniformly at random from $[0,1]$.

Examples of Bounded Orthonormal Systems

Trigonometric System. $\mathcal{D} = [0,1]$ with Lebesgue measure.

$$\psi_k(t) = e^{2\pi i k t}, \quad t \in [0,1].$$

The trigonometric system is orthonormal with $K = 1$. Take subset $\Gamma \subset \mathbb{Z}$ of cardinality N; trigonometric polynomials

$$f(t) = \sum_{k \in \Gamma} x_k \psi_k(t) = \sum_{k \in \Gamma} x_k e^{2\pi i k t}.$$

Samples $t_1, \dots, t_m$ are chosen indendepently and uniformly at random from $[0,1]$. Nonequispaced random Fourier matrix

$$A_{\ell,k} = e^{2\pi i k t_\ell}, \quad \ell \in [m], k \in \Gamma.$$

Examples of Bounded Orthonormal Systems

Trigonometric System. $\mathcal{D} = [0, 1]$ with Lebesgue measure.

$$\psi_k(t) = e^{2\pi ikt}, \quad t \in [0, 1].$$

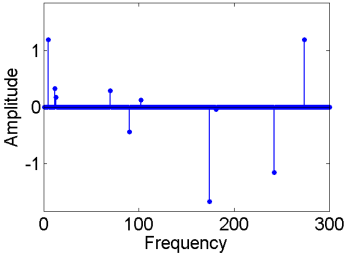
The trigonometric system is orthonormal with $K = 1$. Take subset $\Gamma \subset \mathbb{Z}$ of cardinality N; trigonometric polynomials

$$f(t) = \sum_{k \in \Gamma} x_k \psi_k(t) = \sum_{k \in \Gamma} x_k e^{2\pi ikt}.$$

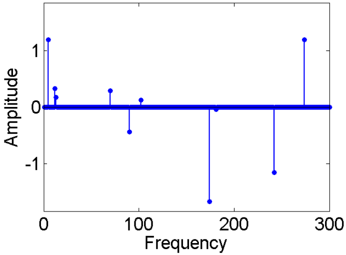
Samples $t_1, \ldots, t_m$ are chosen indendepently and uniformly at random from $[0, 1]$. Nonequispaced random Fourier matrix

$$A_{\ell,k} = e^{2\pi ikt_\ell}, \quad \ell \in [m], k \in \Gamma.$$

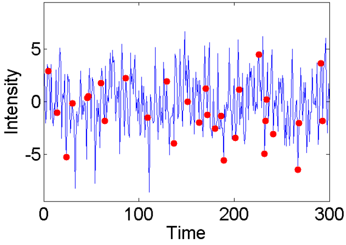
Fast matrix vector multiply using the nonequispaced fast Fourier transform (NFFT).



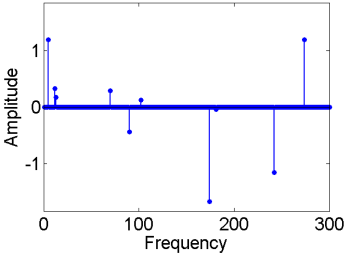
Fourier coefficients



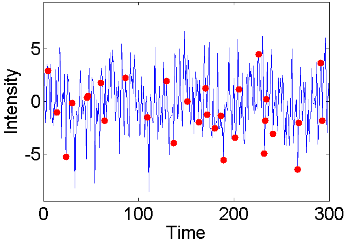
Fourier coefficients



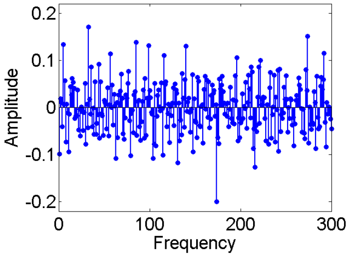
Time domain signal with 30 samples



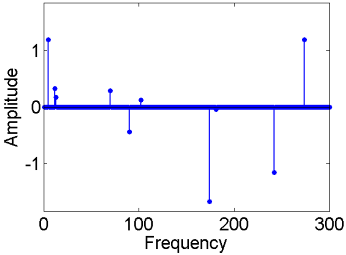
Fourier coefficients



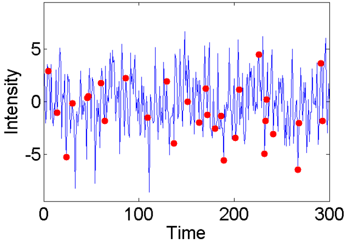
Time domain signal with 30 samples



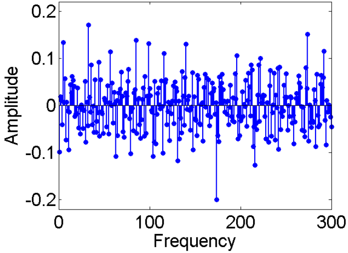
ℓ_2-minimization



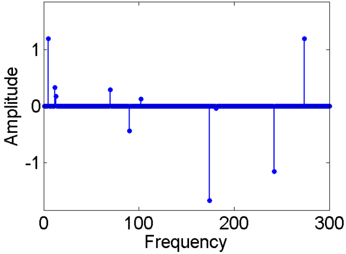
Fourier coefficients



Time domain signal with 30 samples



ℓ_2-minimization



ℓ_1-minimization

Further examples

- Real trigonometric polynomials
- Discrete systems – Random rows of bounded orthogonal matrices
- Random partial Fourier matrices
- Legendre polynomials (needs a “twist”, see below)

hausdorff
CENTER FOR MATHEMATICS

Theorem (Rauhut 2006, 2009)

Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a bounded orthonormal system with constant $K \geq 1$. Suppose

$$\frac{m}{\ln(m)} \geq CK^2\delta^{-2}s\ln^2(s)\ln(N).$$

Then with probability at least $1 - N^{-\gamma\ln^2(s)\ln(m)}$ the restricted isometry constant of $\frac{1}{\sqrt{m}}A$ satisfies $\delta_s \leq \delta$.

Improvement of previous results for the discrete case due to Candès – Tao (2005) and Rudelson – Vershynin (2006).
Explicit (but bad) constants.
Simplified condition

$$s \geq CK^2 s\ln^4(N)$$

for uniform s-sparse recovery with probability $\geq 1 - N^{-\gamma\ln^3(N)}$.

Consider $\mathcal{D} = [-1, 1]$ with normalized Lebesgue measure and orthonormal system of Legendre polynomials $\phi_j = P_j$, $j = 0, \ldots, N-1$.

hausdorff
CENTER FOR MATHEMATICS

Consider $\mathcal{D} = [-1, 1]$ with normalized Lebesgue measure and orthonormal system of Legendre polynomials $\phi_j = P_j$, $j = 0, \ldots, N-1$.

It holds $\|P_j\|_\infty = \sqrt{2j+1}$, so $K = \sqrt{2N-1}$.

Legendre Polynomials

Consider $\mathcal{D} = [-1, 1]$ with normalized Lebesgue measure and orthonormal system of Legendre polynomials $\phi_j = P_j$, $j = 0, \ldots, N-1$.

It holds $\|P_j\|_\infty = \sqrt{2j+1}$, so $K = \sqrt{2N-1}$.

The previous result yields the (almost) trivial bound

$$m \geq C N s \log^2(s) \log(m) \log(N) > N.$$

hausdorff CENTER FOR MATHEMATICS

Consider $\mathcal{D} = [-1, 1]$ with normalized Lebesgue measure and orthonormal system of Legendre polynomials $\phi_j = P_j$, $j = 0, \ldots, N-1$.

It holds $\|P_j\|_\infty = \sqrt{2j+1}$, so $K = \sqrt{2N-1}$.

The previous result yields the (almost) trivial bound

$$m \geq C N s \log^2(s) \log(m) \log(N) > N.$$

Can we do better?

hausdorff CENTER FOR MATHEMATICS

Do not sample uniformly, but with respect to the “Chebyshev” probability measure

$$\nu(dx) = \frac{1}{\pi}(1 - x^2)^{-1/2} dx \quad \text{on } [-1, 1].$$

The functions

$$g_j(x) = \sqrt{\frac{\pi}{2}}(1 - x^2)^{1/4} P_j(x)$$

are orthonormal with respect to ν.

Do not sample uniformly, but with respect to the "Chebyshev" probability measure

$$\nu(dx) = \frac{1}{\pi}(1-x^2)^{-1/2}dx \quad \text{on } [-1,1].$$

The functions

$$g_j(x) = \sqrt{\frac{\pi}{2}}(1-x^2)^{1/4}P_j(x)$$

are orthonormal with respect to ν.
A classical estimate for Legendre polynomials states that

$$\sup_{x\in[-1,1]} |g_j(x)| \leq \sqrt{2} \quad \text{for all } j \in N_0.$$

hausdorff CENTER FOR MATHEMATICS

Random sampling of sparse Legendre expansions

Theorem (Rauhut – Ward 2010)

Let P_j, $j = 0, \ldots, N-1$, be the normalized Legendre polynomials, and let x_ℓ, $\ell = 1, \ldots, m$, be sampling points in $[-1,1]$ which are chosen independently at random according to Chebyshev probability measure $\pi^{-1}(1-x^2)^{-1/2}dx$ on $[-1,1]$. Assume

$$m \geq Cs \log^4(N).$$

Then with probability at least $1 - N^{-\gamma \log^3(N)}$ every s-sparse Legendre expansion

$$f(x) = \sum_{j=0}^{N-1} x_j P_j(x)$$

can be recovered from $y = (f(x_\ell))_{\ell=1}^m$ via ℓ_1-minimization.

Let $D = \sqrt{\pi/2}\,\mathrm{diag}\{(1-x_\ell^2)^{1/4}, \ell = 1,\ldots,m\} \in \mathbb{R}^{m\times m}$
and $A \in \mathbb{R}^{m\times N}$, $B \in \mathbb{R}^{m\times N}$ with entries

$$A_{\ell,j} = P_j(x_\ell), \qquad B_{\ell,j} = g_j(x_\ell).$$

Then $B = DA$. Hence,

$$\ker B = \ker A.$$

Since the constant $K \leq C$ for the system $\{g_\ell\}$, the matrix B satisfies RIP under the stated condition.

Circulant matrix: For $\mathbf{b} = (b_0, b_1, \ldots, b_{N-1}) \in \mathbb{C}^N$ let $\Phi = \Phi(\mathbf{b}) \in \mathbb{C}^{N \times N}$ be the matrix with entries $\Phi_{i,j} = b_{j-i \mod N}$,

$$\Phi(\mathbf{b}) = \begin{pmatrix} b_0 & b_1 & \cdots & \cdots & b_{N-1} \\ b_{N-1} & b_0 & b_1 & \cdots & b_{N-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_1 & b_2 & \cdots & b_{N-1} & b_0 \end{pmatrix}.$$

Let $\Theta \subset [N]$ arbitrary of cardinality m.
R_Θ: operator that restricts a vector $\mathbf{x} \in \mathbb{C}^N$ to its entries in Θ.

Restrict $\Phi(\mathbf{b})$ to the rows indexed by Θ:
Partial circulant matrix: $\Phi^\Theta(\mathbf{b}) = R_\Theta \Phi(\mathbf{b}) \in \mathbb{C}^{m \times N}$

Partial random circulant matrices

Let $\Theta \subset [N]$ arbitrary of cardinality m.
R_Θ: operator that restricts a vector $\mathbf{x} \in \mathbb{C}^N$ to its entries in Θ.

Restrict $\Phi(\mathbf{b})$ to the rows indexed by Θ:
Partial circulant matrix: $\Phi^\Theta(\mathbf{b}) = R_\Theta \Phi(\mathbf{b}) \in \mathbb{C}^{m \times N}$

Convolution followed by subsampling:
$\mathbf{y} = R_\Theta \Phi(\mathbf{b})\mathbf{x} = R_\Theta(\mathbf{b} * \mathbf{x})$

Let $\Theta \subset [N]$ arbitrary of cardinality m.
R_Θ: operator that restricts a vector $\mathbf{x} \in \mathbb{C}^N$ to its entries in Θ.

Restrict $\Phi(\mathbf{b})$ to the rows indexed by Θ:
Partial circulant matrix: $\Phi^\Theta(\mathbf{b}) = R_\Theta \Phi(\mathbf{b}) \in \mathbb{C}^{m \times N}$

Convolution followed by subsampling:
$\mathbf{y} = R_\Theta \Phi(\mathbf{b})\mathbf{x} = R_\Theta(\mathbf{b} * \mathbf{x})$
Matrix vector multiplication via the FFT!

Let $\Theta \subset [N]$ arbitrary of cardinality m.
R_Θ: operator that restricts a vector $\mathbf{x} \in \mathbb{C}^N$ to its entries in Θ.

Restrict $\Phi(\mathbf{b})$ to the rows indexed by Θ:
Partial circulant matrix: $\Phi^\Theta(\mathbf{b}) = R_\Theta \Phi(\mathbf{b}) \in \mathbb{C}^{m \times N}$

Convolution followed by subsampling:
$\mathbf{y} = R_\Theta \Phi(\mathbf{b})\mathbf{x} = R_\Theta(\mathbf{b} * \mathbf{x})$
Matrix vector multiplication via the FFT!

We choose the vector $\mathbf{b} \in \mathbb{C}^N$ at random, in particular, as Rademacher sequence $\mathbf{b} = \epsilon$, that is, $\epsilon_\ell = \pm 1$.

Performance of $\Phi^\Theta(\epsilon)$ in compressive sensing?

Nonuniform recovery result for circulant matrices

Theorem (Rauhut 2009)

Let $\Theta \subset [N]$ be an arbitrary (deterministic) set of cardinality m. Let $\mathbf{x} \in \mathbb{C}^N$ be s-sparse such that the signs of its non-zero entries form a Rademacher or Steinhaus sequence. Choose $\epsilon \in \mathbb{R}^N$ to be a Rademacher sequence. Let $\mathbf{y} = \Phi^{\Theta}(\epsilon)\mathbf{x} \in \mathbb{C}^m$. If

$$m \geq 57 s \ln^2(17N^2/\varepsilon)$$

then $\mathbf{x}$ can be recovered from $\mathbf{y}$ via ℓ_1-minimization with probability at least $1 - \varepsilon$.

Theorem (Rauhut – Romberg – Tropp 2010)

Let $\Theta \subset [N]$ be an arbitrary (deterministic) set of cardinality m. Let $\epsilon \in \mathbb{R}^N$ be a Rademacher sequence. Assume that

$$m \geq C\delta^{-1} s^{3/2} \log^{3/2}(N),$$

and, for $\varepsilon \in (0,1)$,

$$m \geq C\delta^{-2} s \log^2(s) \log^2(N) \log(\varepsilon^{-1})$$

Then with probability at least $1-\varepsilon$ the restricted isometry constants of $\frac{1}{\sqrt{m}}\Phi^\Theta(\epsilon)$ satisfy $\delta_s \leq \delta$.

Theorem is also valid for Steinhaus or Gaussian sequence.

Proof Idea

With translation operators $S_\ell : \mathbb{C}^N \to \mathbb{C}^N, (S_\ell h)_k = h_{k-\ell \mod N}$ we can write

$$A = \frac{1}{\sqrt{m}} \Phi^\Theta(\epsilon) = \frac{1}{\sqrt{m}} \sum_{\ell=1}^{N} \epsilon_\ell R_\Theta S_\ell.$$

Proof Idea

With translation operators $S_\ell : \mathbb{C}^N \to \mathbb{C}^N, (S_\ell h)_k = h_{k-\ell \mod N}$ we can write

$$A = \frac{1}{\sqrt{m}} \Phi^\Theta(\epsilon) = \frac{1}{\sqrt{m}} \sum_{\ell=1}^{N} \epsilon_\ell R_\Theta S_\ell.$$

Denote $T_s := \{x \in \mathbb{R}^N, \|x\|_2 \leq 1, \|x\|_0 \leq s\}$. Then

$$\delta_s = \sup_{x \in T_s} |\langle (A^* A - I)x, x \rangle| = \sup_{x \in T_s} \frac{1}{m} |\sum_{k \neq j} \epsilon_j \epsilon_k x^* Q_{j,k} x|$$

with $Q_{j,k} = S_j^* P_\Theta S_k$ and $P_\Theta = R_\Theta^* R_\Theta$ is the projection of a vector in $\mathbb{R}^N$ onto its entries in Θ.

We arrive at estimating the supremum of a Rademacher chaos process of order 2.

Dudley type inequality for chaos processes

Theorem (Talagrand)

Let $Y_x = \sum_{k,j} \epsilon_j \epsilon_k Z_{jk}(x)$ *be a scalar Rademacher chaos process indexed by* $x \in T$, *with* $Z_{jj}(x) = 0$ *and* $Z_{jk}(x) = Z_{kj}(x)$. *Introduce two metrics on* T, *with* $(Z(x))_{j,k} = Z_{jk}(x)_{j,k}$,

$$\begin{aligned} d_1(x,y) &= \|Z(x) - Z(y)\|_F, \\ d_2(x,y) &= \|Z(x) - Z(y)\|_{2\to 2}. \end{aligned}$$

Let $N(T, d_i, u)$ *denote the minimal number of balls of radius* u *in the metric* d_i *necessary to cover* T. *There exists a universal constant* K *such that, for an arbitrary* $x_0 \in T$,

$$\mathbb{E} \sup_{x \in T} |Y_x - Y_{x_0}| \leq$$
$$K \max\left\{ \int_0^\infty \sqrt{\log(N(T,d_1,u))}du, \int_0^\infty \log(N(T,d_2,u))du \right\}.$$

In our situation,

$$\int_0^\infty \sqrt{\log(N(T_s, d_1, u))}du \leq C\sqrt{\frac{s \log^2(s)\log^2(N)}{m}},$$

and

$$\int_0^\infty \log(N(T_s, d_2, u))du \leq C\frac{s^{3/2}\log^{3/2}(N)}{m}.$$

Technique: Pass to Fourier transform, and use estimates due to Rudelson and Vershynin.

Probability estimate:
Concentration inequality due to Talagrand (1996),
with improvements due to Boucheron, Lugosi, Massart (2003).

Translation and Modulation on $\mathbb{C}^n$

$$(S_p h)_q = h_{(p+q) \mod n} \quad \text{and} \quad (M_\ell h)_q = e^{2\pi i \ell q/n} h_q.$$

For $h \in \mathbb{C}^n$ define Gabor system (Gabor synthesis matrix)

$$A_h = (M_\ell S_p h)_{\ell,p=0,\dots,n-1} \in \mathbb{C}^{n \times n^2}$$

Motivation: Wireless communications and sonar.

Translation and Modulation on $\mathbb{C}^n$

$$(S_p h)_q = h_{(p+q) \mod n} \quad \text{and} \quad (M_\ell h)_q = e^{2\pi i \ell q/n} h_q.$$

For $h \in \mathbb{C}^n$ define Gabor system (Gabor synthesis matrix)

$$A_h = (M_\ell S_p h)_{\ell,p=0,\ldots,n-1} \in \mathbb{C}^{n \times n^2}$$

Motivation: Wireless communications and sonar.

Choose $h \in \mathbb{C}^n$ at random, more precisely as a Steinhaus sequence: All entries h_q, $q = 0, \ldots, n-1$, are chosen independently and uniformly at random from the torus $\{z \in \mathbb{C}, |z| = 1\}$.

Performance of $A_h \in \mathbb{C}^{n \times n^2}$ for sparse recovery?

Theorem (Pfander – Rauhut 2007)

Let $x \in \mathbb{C}^{n^2}$ be s-sparse. Choose $A_h \in \mathbb{C}^{n \times n^2}$ at random (that is, let h be a Steinhaus sequence). Assume that

$$s \leq C \frac{n}{\log(n/\varepsilon)}.$$

Then with probability at least $1 - \varepsilon$ ℓ_1-minimization recovers x from $y = A_h x$.

Theorem (June 2010)

Choose $A_h \in \mathbb{C}^{n \times n^2}$ at random (this is, let h be a Steinhaus sequence). Assume that

$$n \geq C\delta^{-1} s^{3/2} \log^{3/2}(n),$$

and, for $\varepsilon \in (0,1)$,

$$n \geq C\delta^{-2} s \log^2(s) \log^2(n) \log(\varepsilon^{-1}).$$

Then with probability at least $1 - \varepsilon$ the restricted isometry constants of $\frac{1}{\sqrt{n}} A_h$ satisfy $\delta_s \leq \delta$.

Result is valid also for Rademacher or Gaussian generator h.

THAT'S ALL

THANKS!