Exponential decay of reconstruction error from binary measurements of sparse signals

Deanna Needell

CLAREMONT McKENNA COLLEGE

Joint work with R. Baraniuk, S. Foucart, Y. Plan, and M. Wootters

Outline

- Introduction
  - Mathematical Formulation & Methods
- Practical CS
  - Other notions of sparsity
  - Heavy quantization
  - Adaptive sampling

The mathematical problem

1. Signal of interest $f \in \mathbb{C}^d (= \mathbb{C}^{N \times N})$
2. Measurement operator $\mathscr{A} : \mathbb{C}^d \to \mathbb{C}^m \; (m \ll d)$
3. Measurements $y = \mathscr{A} f + \xi$

$$\begin{bmatrix} \\ y \\ \\ \end{bmatrix} = \begin{bmatrix} & & \\ & \mathscr{A} & \\ & & \end{bmatrix} \begin{bmatrix} \\ f \\ \\ \\ \\ \end{bmatrix} + \begin{bmatrix} \\ \xi \\ \\ \end{bmatrix}$$

4. *Problem:* Reconstruct signal f from measurements y

Sparsity

Measurements $y = \mathscr{A} f + \xi$.

$$\begin{bmatrix} \\ y \\ \\ \end{bmatrix} = \begin{bmatrix} & & \\ & \mathscr{A} & \\ & & \end{bmatrix} \begin{bmatrix} \\ f \\ \\ \\ \\ \end{bmatrix} + \begin{bmatrix} \\ \xi \\ \\ \end{bmatrix}$$

Assume f is *sparse*:

- ✧ In the coordinate basis: $\|f\|_0 \stackrel{\text{def}}{=} |\operatorname{supp}(f)| \le s \ll d$
- ✧ In orthonormal basis: $f = Bx$ where $\|x\|_0 \le s \ll d$

In practice, we encounter *compressible* signals.

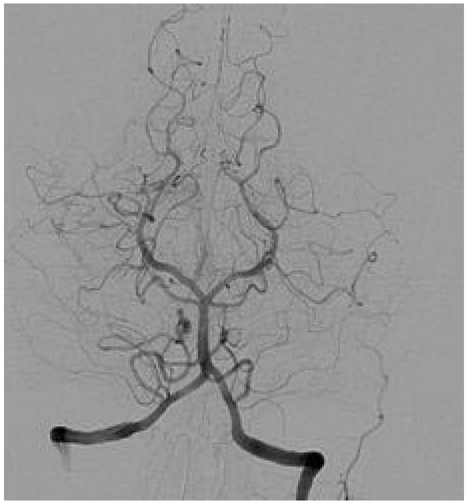
- ✦ f_s is the best s-sparse approximation to f

Many applications...

- Radar, Error Correction
- Computational Biology, Geophysical Data Analysis
- Data Mining, classification
- Neuroscience
- Imaging
- Sparse channel estimation, sparse initial state estimation
- Topology identification of interconnected systems
- ...

Sparsity...

Sparsity in coordinate basis: f=x



Reconstructing the signal f from measurements y

- ℓ_1-minimization [Candès-Romberg-Tao]

Let A satisfy the *Restricted Isometry Property* and set:

$$\hat{f} = \operatorname*{argmin}_g \|g\|_1 \quad \text{such that} \quad \|\mathscr{A} f - y\|_2 \leq \varepsilon,$$

where $\|\xi\|_2 \leq \varepsilon$. Then we can stably recover the signal f:

$$\|f - \hat{f}\|_2 \lesssim \varepsilon + \frac{\|x - x_s\|_1}{\sqrt{s}}.$$

This error bound is optimal.

Restricted Isometry Property

- $\mathscr{A}$ satisfies the Restricted Isometry Property (RIP) when there is $\delta < c$ such that

$$(1-\delta)\|f\|_2 \le \|\mathscr{A} f\|_2 \le (1+\delta)\|f\|_2 \quad \text{whenever } \|f\|_0 \le s.$$

- $m \times d$ Gaussian or Bernoulli measurement matrices satisfy the RIP with high probability when

$$m \gtrsim s \log d.$$

- Random Fourier and others with fast multiply have similar property: $m \gtrsim s \log^4 d$.

Other recovery methods

Greedy Algorithms

- If A satisfies the RIP, then $A^* A$ is “close” to the identity on sparse vectors
- Use proxy $p = A^* y = A^* Ax \approx x$
- Threshold to maintain sparsity: $\hat{x} = H_s(p)$
- Repeat
- (Iterative Hard Thresholding)

The One-Bit Compressive Sensing Problem

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.
- In practice, measurements need to be quantized.

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.
- In practice, measurements need to be quantized.
- One-Bit CS: extreme quantization as $\mathbf{y} = \mathrm{sign}(\mathbf{A}\mathbf{x})$, i.e.,

$$y_i = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle, \qquad i = 1, \ldots, m.$$

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.
- In practice, measurements need to be quantized.
- One-Bit CS: extreme quantization as $\mathbf{y} = \mathrm{sign}(\mathbf{A}\mathbf{x})$, i.e.,

$$y_i = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle, \qquad i = 1, \ldots, m.$$

- Goal: find reconstruction maps $\Delta : \{\pm 1\}^m \to \mathbb{R}^n$ such that,

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.
- In practice, measurements need to be quantized.
- One-Bit CS: extreme quantization as $\mathbf{y} = \mathrm{sign}(\mathbf{A}\mathbf{x})$, i.e.,

$$y_i = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle, \qquad i = 1, \ldots, m.$$

- Goal: find reconstruction maps $\Delta : \{\pm 1\}^m \to \mathbb{R}^n$ such that, assuming the ℓ_2-normalization of $\mathbf{x}$ (why?),

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.
- In practice, measurements need to be quantized.
- One-Bit CS: extreme quantization as $\mathbf{y} = \mathrm{sign}(\mathbf{Ax})$, i.e.,

$$y_i = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle, \qquad i = 1, \ldots, m.$$

- Goal: find reconstruction maps $\Delta : \{\pm 1\}^m \to \mathbb{R}^n$ such that, assuming the ℓ_2-normalization of $\mathbf{x}$ (why?),

$$\|\mathbf{x} - \Delta(\mathbf{y})\| \le \gamma$$

provided the oversampling factor satisfies

$$\lambda := \frac{m}{s \ln(n/s)} \ge f(\gamma)$$

for f slowly increasing when γ decreases to zero

The One-Bit Compressive Sensing Problem

- Standard CS: vectors $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$ acquired via nonadaptive linear measurements $\langle \mathbf{a}_i, \mathbf{x} \rangle$, $i = 1, \ldots, m$.
- In practice, measurements need to be quantized.
- One-Bit CS: extreme quantization as $\mathbf{y} = \operatorname{sign}(\mathbf{A}\mathbf{x})$, i.e.,

$$y_i = \operatorname{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle, \qquad i = 1, \ldots, m.$$

- Goal: find reconstruction maps $\Delta : \{\pm 1\}^m \to \mathbb{R}^n$ such that, assuming the ℓ_2-normalization of $\mathbf{x}$ (why?),

$$\|\mathbf{x} - \Delta(\mathbf{y})\| \leq \gamma$$

provided the oversampling factor satisfies

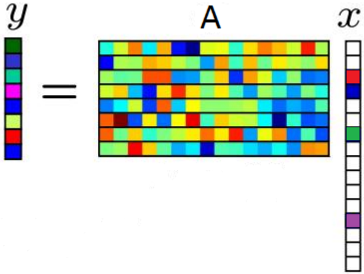
$$\lambda := \frac{m}{s \ln(n/s)} \geq f(\gamma)$$

for f slowly increasing when γ decreases to zero, equivalently

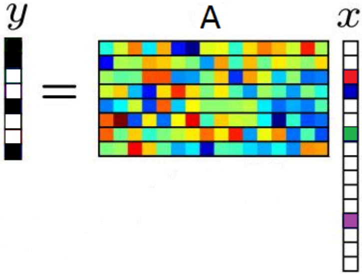
$$\|\mathbf{x} - \Delta(\mathbf{y})\| \leq g(\lambda)$$

for g rapidly decreasing to zero when λ increases.

A visual



A visual



Existing Theoretical Results

Existing Theoretical Results

- Convex optimization algorithms [Plan–Vershynin 13a, 13b].

Existing Theoretical Results

- Convex optimization algorithms [Plan–Vershynin 13a, 13b].
- Uniform, nonadaptive, no quantization error:
  If $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a Gaussian matrix, then w/hp

$$\left\| \mathbf{x} - \frac{\Delta_{\mathrm{LP}}(\mathbf{y})}{\|\Delta_{\mathrm{LP}}(\mathbf{y})\|_2} \right\|_2 \lesssim \lambda^{-1/5} \quad \text{whenever } \|\mathbf{x}\|_0 \le s, \|\mathbf{x}\|_2 = 1.$$

Existing Theoretical Results

- Convex optimization algorithms [Plan–Vershynin 13a, 13b].
- Uniform, nonadaptive, no quantization error:
  If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a Gaussian matrix, then w/hp

$$\left\| \mathbf{x} - \frac{\Delta_{\mathrm{LP}}(\mathbf{y})}{\|\Delta_{\mathrm{LP}}(\mathbf{y})\|_2} \right\|_2 \lesssim \lambda^{-1/5} \quad \text{whenever } \|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 = 1.$$

- Nonuniform, nonadaptive, random quantization error:
  Fix $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 = 1$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a Gaussian matrix, then w/hp

$$\|\mathbf{x} - \Delta_{\mathrm{SOCP}}(\mathbf{y})\|_2 \lesssim \lambda^{-1/4}.$$

Existing Theoretical Results

- Convex optimization algorithms [Plan–Vershynin 13a, 13b].
- Uniform, nonadaptive, no quantization error:
  If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a Gaussian matrix, then w/hp

$$\left\| \mathbf{x} - \frac{\Delta_{\mathrm{LP}}(\mathbf{y})}{\|\Delta_{\mathrm{LP}}(\mathbf{y})\|_2} \right\|_2 \lesssim \lambda^{-1/5} \quad \text{whenever } \|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 = 1.$$

- Nonuniform, nonadaptive, random quantization error:
  Fix $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$, $\|\mathbf{x}\|_2 = 1$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a Gaussian matrix, then w/hp

$$\|\mathbf{x} - \Delta_{\mathrm{SOCP}}(\mathbf{y})\|_2 \lesssim \lambda^{-1/4}.$$

- Uniform, nonadaptive, adversarial quantization error:
  If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a Gaussian matrix, then w/hp

$$\|\mathbf{x} - \Delta_{\mathrm{SOCP}}(\mathbf{y})\|_2 \lesssim \lambda^{-1/12} \quad \text{whenever } \|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 = 1.$$

Limitations of the Framework

Limitations of the Framework

- Power decay is optimal since

$$\|\mathbf{x} - \Delta_{\mathrm{opt}}(\mathbf{y})\|_2 \gtrsim \lambda^{-1}$$

even if $\mathrm{supp}(\mathbf{x})$ known in advance [Goyal–Vetterli–Thao 98].

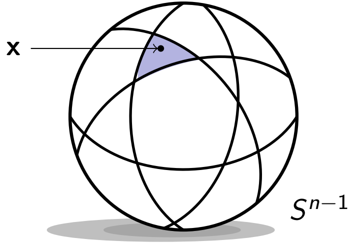
Limitations of the Framework

- Power decay is optimal since

$$\|\mathbf{x} - \Delta_{\mathrm{opt}}(\mathbf{y})\|_2 \gtrsim \lambda^{-1}$$

  even if $\mathrm{supp}(\mathbf{x})$ known in advance [Goyal–Vetterli–Thao 98].

- Geometric intuition



`http://dsp.rice.edu/1bitCS/choppyanimated.gif`

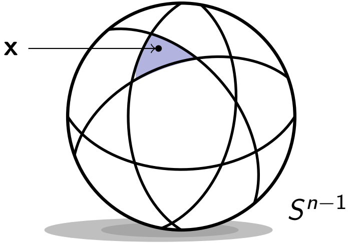
Limitations of the Framework

- Power decay is optimal since

$$\|\mathbf{x} - \Delta_{\mathrm{opt}}(\mathbf{y})\|_2 \gtrsim \lambda^{-1}$$

even if $\mathrm{supp}(\mathbf{x})$ known in advance [Goyal–Vetterli–Thao 98].

- Geometric intuition



http://dsp.rice.edu/1bitCS/choppyanimated.gif

- Remedy: adaptive choice of dithers $\tau_1, \ldots, \tau_m$ in

$$y_i = \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i), \qquad i = 1, \ldots, m.$$

Exponential Decay: General Strategy

Exponential Decay: General Strategy

- Rely on an *order-one quantization/recovery scheme*: for any $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$, $\|\mathbf{x}\|_2 \leq R$, take $q \asymp s \ln(n/s)$ one-bit measurements and estimate both the direction and the magnitude of $\mathbf{x}$ by producing $\widehat{\mathbf{x}}$ such that

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 \leq R/4.$$

Exponential Decay: General Strategy

- Rely on an *order-one quantization/recovery scheme*: for any $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$, $\|\mathbf{x}\|_2 \leq R$, take $q \asymp s \ln(n/s)$ one-bit measurements and estimate both the direction and the magnitude of $\mathbf{x}$ by producing $\widehat{\mathbf{x}}$ such that

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 \leq R/4.$$

- Let $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$, $\|\mathbf{x}\|_2 \leq R$. Start with $\mathbf{x}^0 = \mathbf{0}$.

Exponential Decay: General Strategy

- Rely on an *order-one quantization/recovery scheme*: for any $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$, $\|\mathbf{x}\|_2 \le R$, take $q \asymp s \ln(n/s)$ one-bit measurements and estimate both the direction and the magnitude of $\mathbf{x}$ by producing $\widehat{\mathbf{x}}$ such that

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 \le R/4.$$

- Let $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$, $\|\mathbf{x}\|_2 \le R$. Start with $\mathbf{x}^0 = \mathbf{0}$.
- For $t = 0, 1, \ldots$, estimate $\mathbf{x} - \mathbf{x}^t$ by $\widehat{\mathbf{x} - \mathbf{x}^t}$, then set

$$\mathbf{x}^{t+1} = \quad \mathbf{x}^t + \widehat{\mathbf{x} - \mathbf{x}^t}\;, \quad \text{so that} \quad \|\mathbf{x} - \mathbf{x}^{t+1}\|_2 \le R/4^{t+1}.$$

Exponential Decay: General Strategy

- Rely on an *order-one quantization/recovery scheme*: for any $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$, $\|\mathbf{x}\|_2 \le R$, take $q \asymp s \ln(n/s)$ one-bit measurements and estimate both the direction and the magnitude of $\mathbf{x}$ by producing $\widehat{\mathbf{x}}$ such that
$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 \le R/4.$$
- Let $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$, $\|\mathbf{x}\|_2 \le R$. Start with $\mathbf{x}^0 = \mathbf{0}$.
- For $t = 0, 1, \ldots$, estimate $\mathbf{x} - \mathbf{x}^t$ by $\widehat{\mathbf{x} - \mathbf{x}^t}$, then set
$$\mathbf{x}^{t+1} = H_s(\mathbf{x}^t + \widehat{\mathbf{x} - \mathbf{x}^t}), \quad \text{so that} \quad \|\mathbf{x} - \mathbf{x}^{t+1}\|_2 \le R/2^{t+1}.$$

Exponential Decay: General Strategy

- Rely on an *order-one quantization/recovery scheme*: for any $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$, $\|\mathbf{x}\|_2 \le R$, take $q \asymp s \ln(n/s)$ one-bit measurements and estimate both the direction and the magnitude of $\mathbf{x}$ by producing $\widehat{\mathbf{x}}$ such that

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 \le R/4.$$

- Let $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \le s$, $\|\mathbf{x}\|_2 \le R$. Start with $\mathbf{x}^0 = \mathbf{0}$.
- For $t = 0, 1, \ldots$, estimate $\mathbf{x} - \mathbf{x}^t$ by $\widehat{\mathbf{x} - \mathbf{x}^t}$, then set

$$\mathbf{x}^{t+1} = H_s(\mathbf{x}^t + \widehat{\mathbf{x} - \mathbf{x}^t}), \quad \text{so that} \quad \|\mathbf{x} - \mathbf{x}^{t+1}\|_2 \le R/2^{t+1}.$$

- After T iterations, number of measurements is $m = qT$, and

$$\|\mathbf{x} - \mathbf{x}^T\|_2 \le R\, 2^{-T} = R\, 2^{-\frac{m}{q}} = R \exp\left(-c\lambda\right).$$

Exponential Decay: General Strategy

- Rely on an *order-one quantization/recovery scheme*: for any $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$, $\|\mathbf{x}\|_2 \leq R$, take $q \asymp s \ln(n/s)$ one-bit measurements and estimate both the direction and the magnitude of $\mathbf{x}$ by producing $\widehat{\mathbf{x}}$ such that

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 \leq R/4.$$

- Let $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$, $\|\mathbf{x}\|_2 \leq R$. Start with $\mathbf{x}^0 = \mathbf{0}$.
- For $t = 0, 1, \ldots$, estimate $\mathbf{x} - \mathbf{x}^t$ by $\widehat{\mathbf{x} - \mathbf{x}^t}$, then set

$$\mathbf{x}^{t+1} = H_s(\mathbf{x}^t + \widehat{\mathbf{x} - \mathbf{x}^t}), \quad \text{so that} \quad \|\mathbf{x} - \mathbf{x}^{t+1}\|_2 \leq R/2^{t+1}.$$

- After T iterations, number of measurements is $m = qT$, and

$$\|\mathbf{x} - \mathbf{x}^T\|_2 \leq R\,2^{-T} = R\,2^{-\frac{m}{q}} = R\exp\left(-c\lambda\right).$$

- Software step needed to compute the thresholds $\tau_i = \langle \mathbf{a}_i, \mathbf{x}^t \rangle$.

Order-One Scheme Based on Convex Optimization

Order-One Scheme Based on Convex Optimization

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.

Order-One Scheme Based on Convex Optimization

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Dithers $\tau_1, \ldots, \tau_q$: independent $\mathcal{N}(0, R^2)$.

Order-One Scheme Based on Convex Optimization

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Dithers $\tau_1, \ldots, \tau_q$: independent $\mathcal{N}(0, R^2)$.
- $\widehat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{z}\|_1$ subject to $\|\mathbf{z}\|_2 \leq R$, $y_i(\langle \mathbf{a}_i, \mathbf{z} \rangle - \tau_i) \geq 0$.

Order-One Scheme Based on Convex Optimization

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Dithers $\tau_1, \ldots, \tau_q$: independent $\mathcal{N}(0, R^2)$.
- $\widehat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{z}\|_1$ subject to $\|\mathbf{z}\|_2 \leq R$, $y_i(\langle \mathbf{a}_i, \mathbf{z} \rangle - \tau_i) \geq 0$.
- If $q \geq c\delta^{-4} s \ln(n/s)$, then w/hp

$$\|\mathbf{x} - \widehat{\mathbf{x}}\| \leq \delta R \quad \text{whenever } \|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 \leq R.$$

Order-One Scheme Based on Convex Optimization

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Dithers $\tau_1, \ldots, \tau_q$: independent $\mathcal{N}(0, R^2)$.
- $\widehat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{z}\|_1$ subject to $\|\mathbf{z}\|_2 \leq R$, $y_i(\langle \mathbf{a}_i, \mathbf{z} \rangle - \tau_i) \geq 0$.
- If $q \geq c\delta^{-4} s \ln(n/s)$, then w/hp

$$\|\mathbf{x} - \widehat{\mathbf{x}}\| \leq \delta R \quad \text{whenever } \|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 \leq R.$$

- Pros: dithers are nonadaptive.

Order-One Scheme Based on Convex Optimization

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Dithers $\tau_1, \ldots, \tau_q$: independent $\mathcal{N}(0, R^2)$.
- $\widehat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{z}\|_1$ subject to $\|\mathbf{z}\|_2 \leq R$, $y_i(\langle \mathbf{a}_i, \mathbf{z} \rangle - \tau_i) \geq 0$.
- If $q \geq c\delta^{-4} s \ln(n/s)$, then w/hp

$$\|\mathbf{x} - \widehat{\mathbf{x}}\| \leq \delta R \quad \text{whenever } \|\mathbf{x}\|_0 \leq s, \|\mathbf{x}\|_2 \leq R.$$

- Pros: dithers are nonadaptive.
- Cons: slow, post-quantization error not handled.

Order-One Scheme Based on Hard Thresholding

Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.

Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Use half of them to estimate the direction of $\mathbf{x}$ as

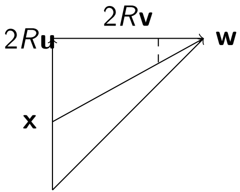
$$\mathbf{u} = H_s'(\mathbf{A}^* \mathrm{sign}(\mathbf{A}\mathbf{x})).$$

Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Use half of them to estimate the direction of $\mathbf{x}$ as

$$\mathbf{u} = H_s'(\mathbf{A}^* \mathrm{sign}(\mathbf{A}\mathbf{x})).$$

- Construct sparse vectors $\mathbf{v}, \mathbf{w}$ $(\mathrm{supp}(\mathbf{v}) \subset \mathrm{supp}(\mathbf{u}))$ according to

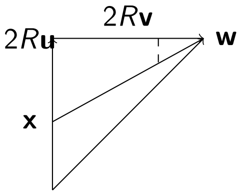


Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \dots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Use half of them to estimate the direction of $\mathbf{x}$ as

$$\mathbf{u} = H_s'(\mathbf{A}^* \mathrm{sign}(\mathbf{A}\mathbf{x})).$$

- Construct sparse vectors $\mathbf{v}, \mathbf{w}$ $(\mathrm{supp}(\mathbf{v}) \subset \mathrm{supp}(\mathbf{u}))$ according to



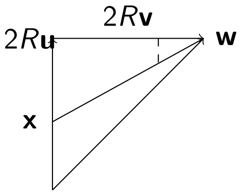
- Use other half to estimate the direction of $\mathbf{x} - \mathbf{w}$ applying hard thresholding again.

Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Use half of them to estimate the direction of $\mathbf{x}$ as

$$\mathbf{u} = H_s'(\mathbf{A}^* \mathrm{sign}(\mathbf{A}\mathbf{x})).$$

- Construct sparse vectors $\mathbf{v}, \mathbf{w}$ $(\mathrm{supp}(\mathbf{v}) \subset \mathrm{supp}(\mathbf{u}))$ according to



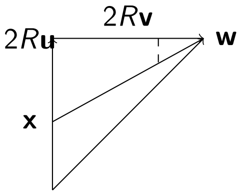
- Use other half to estimate the direction of $\mathbf{x} - \mathbf{w}$ applying hard thresholding again.
- Plane geometry to estimate direction and magnitude of $\mathbf{x}$.

Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Use half of them to estimate the direction of $\mathbf{x}$ as

$$\mathbf{u} = H_s'(\mathbf{A}^* \mathrm{sign}(\mathbf{A}\mathbf{x})).$$

- Construct sparse vectors $\mathbf{v}, \mathbf{w}$ $(\mathrm{supp}(\mathbf{v}) \subset \mathrm{supp}(\mathbf{u}))$ according to



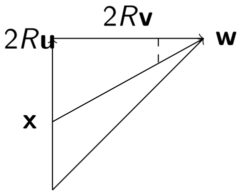
- Use other half to estimate the direction of $\mathbf{x} - \mathbf{w}$ applying hard thresholding again.
- Plane geometry to estimate direction and magnitude of $\mathbf{x}$.
- Cons: dithers $\langle \mathbf{a}_i, \mathbf{w} \rangle$ are adaptive.

Order-One Scheme Based on Hard Thresholding

- Measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$: independent $\mathcal{N}(0, \mathrm{I}_q)$.
- Use half of them to estimate the direction of $\mathbf{x}$ as

$$\mathbf{u} = H_s'(\mathbf{A}^* \mathrm{sign}(\mathbf{A}\mathbf{x})).$$

- Construct sparse vectors $\mathbf{v}, \mathbf{w}$ $(\mathrm{supp}(\mathbf{v}) \subset \mathrm{supp}(\mathbf{u}))$ according to



- Use other half to estimate the direction of $\mathbf{x} - \mathbf{w}$ applying hard thresholding again.
- Plane geometry to estimate direction and magnitude of $\mathbf{x}$.
- Cons: dithers $\langle \mathbf{a}_i, \mathbf{w} \rangle$ are adaptive.
- Pros: deterministic, fast, handles pre/post-quantization errors.

Measurement Errors

Measurement Errors

- Pre-quantization error $\mathbf{e} \in \mathbb{R}^m$ in

$$y_i = \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i + e_i).$$

Measurement Errors

- Pre-quantization error $\mathbf{e} \in \mathbb{R}^m$ in
$$y_i = \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i + e_i).$$
- If $\|\mathbf{e}\|_\infty \leq \varepsilon R\, 2^{-T}$ (or $\|\mathbf{e}^t\|_2 \leq \varepsilon \sqrt{q} \|\mathbf{x} - \mathbf{x}^t\|_2$ throughout), then
$$\|\mathbf{x} - \mathbf{x}^T\|_2 \leq R\, 2^{-T} = R \exp(-c\lambda)$$
for the convex-optimization and hard-thresholding schemes.

Measurement Errors

- Pre-quantization error $\mathbf{e} \in \mathbb{R}^m$ in

$$y_i = \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i + e_i).$$

- If $\|\mathbf{e}\|_\infty \leq \varepsilon R \, 2^{-T}$ (or $\|\mathbf{e}^t\|_2 \leq \varepsilon \sqrt{q} \|\mathbf{x} - \mathbf{x}^t\|_2$ throughout), then

$$\|\mathbf{x} - \mathbf{x}^T\|_2 \leq R \, 2^{-T} = R \exp(-c\lambda)$$

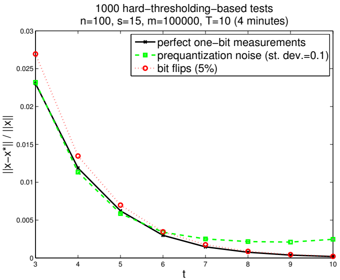
  for the convex-optimization and hard-thresholding schemes.
- Post-quantization error $\mathbf{f} \in \{\pm 1\}^m$ in

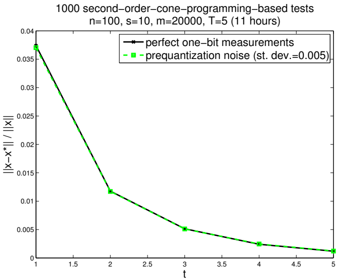
$$y_i = f_i \, \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i).$$

Measurement Errors

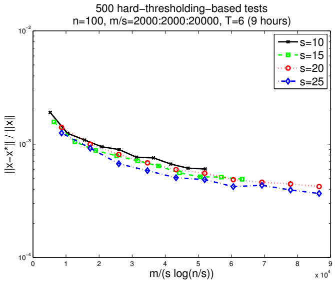
- Pre-quantization error $\mathbf{e} \in \mathbb{R}^m$ in
$$y_i = \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i + e_i).$$
- If $\|\mathbf{e}\|_\infty \leq \varepsilon R\, 2^{-T}$ (or $\|\mathbf{e}^t\|_2 \leq \varepsilon \sqrt{q} \|\mathbf{x} - \mathbf{x}^t\|_2$ throughout), then
$$\|\mathbf{x} - \mathbf{x}^T\|_2 \leq R\, 2^{-T} = R \exp(-c\lambda)$$
for the convex-optimization and hard-thresholding schemes.
- Post-quantization error $\mathbf{f} \in \{\pm 1\}^m$ in
$$y_i = f_i \,\mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i).$$
- If $\mathrm{card}(\{i : f_i^t = -1\}) \leq \eta q$ throughout, then
$$\|\mathbf{x} - \mathbf{x}^T\|_2 \leq R\, 2^{-T} = R \exp(-c\lambda)$$
for the hard-thresholding scheme.

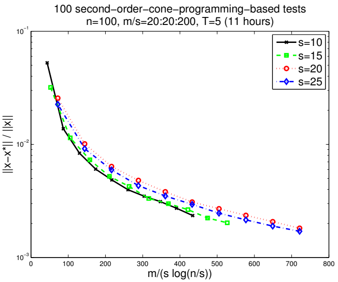
Numerical Illustration





Numerical Illustration, ctd





Ingredients for the Proofs

Ingredients for the Proofs

- Let $\mathbf{A} \in \mathbb{R}^{q \times n}$ with independent $\mathcal{N}(0,1)$ entries.

Ingredients for the Proofs

- Let $\mathbf{A} \in \mathbb{R}^{q \times n}$ with independent $\mathcal{N}(0,1)$ entries.
- Sign Product Embedding Property: if $q \geq C\delta^{-6} s \ln(n/s)$, then with w/hp

$$\left| \frac{\sqrt{\pi/2}}{q} \langle \mathbf{A}\mathbf{w}, \mathrm{sign}(\mathbf{A}\mathbf{x}) \rangle - \langle \mathbf{w}, \mathbf{x} \rangle \right| \leq \delta$$

for all $\mathbf{w}, \mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{w}\|_0, \|\mathbf{x}\|_0 \leq s$ and $\|\mathbf{w}\|_2 = \|\mathbf{x}\|_2 = 1$.

Ingredients for the Proofs

- Let $\mathbf{A} \in \mathbb{R}^{q \times n}$ with independent $\mathcal{N}(0,1)$ entries.
- Sign Product Embedding Property: if $q \geq C\delta^{-6} s \ln(n/s)$, then with w/hp

$$\left| \frac{\sqrt{\pi/2}}{q} \langle \mathbf{A}\mathbf{w}, \mathrm{sign}(\mathbf{A}\mathbf{x}) \rangle - \langle \mathbf{w}, \mathbf{x} \rangle \right| \leq \delta$$

  for all $\mathbf{w}, \mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{w}\|_0, \|\mathbf{x}\|_0 \leq s$ and $\|\mathbf{w}\|_2 = \|\mathbf{x}\|_2 = 1$.
- Simultaneous (ℓ_2, ℓ_1)-Quotient Property: w/hp, every $\mathbf{e} \in \mathbb{R}^q$ can be written as

$$\mathbf{e} = \mathbf{A}\mathbf{u} \quad \text{with} \quad \begin{cases} \|\mathbf{u}\|_2 & \leq & d\|\mathbf{e}\|_2/\sqrt{q}, \\ \|\mathbf{u}\|_1 & \leq & d'\sqrt{s_*}\|\mathbf{e}\|_2/\sqrt{q}, \end{cases}$$

  where $s_* = q/\ln(n/q)$.

Ingredients for the Proofs

- Let $\mathbf{A} \in \mathbb{R}^{q \times n}$ with independent $\mathcal{N}(0,1)$ entries.
- Sign Product Embedding Property: if $q \geq C\delta^{-6} s \ln(n/s)$, then with w/hp

$$\left| \frac{\sqrt{\pi/2}}{q} \langle \mathbf{A}\mathbf{w}, \operatorname{sign}(\mathbf{A}\mathbf{x}) \rangle - \langle \mathbf{w}, \mathbf{x} \rangle \right| \leq \delta$$

  for all $\mathbf{w}, \mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{w}\|_0, \|\mathbf{x}\|_0 \leq s$ and $\|\mathbf{w}\|_2 = \|\mathbf{x}\|_2 = 1$.
- Simultaneous (ℓ_2, ℓ_1)-Quotient Property: w/hp, every $\mathbf{e} \in \mathbb{R}^q$ can be written as

$$\mathbf{e} = \mathbf{A}\mathbf{u} \quad \text{with} \quad \begin{cases} \|\mathbf{u}\|_2 & \leq & d\|\mathbf{e}\|_2/\sqrt{q}, \\ \|\mathbf{u}\|_1 & \leq & d'\sqrt{s_*}\|\mathbf{e}\|_2/\sqrt{q}, \end{cases}$$

  where $s_* = q/\ln(n/q)$.
- Restricted Isometry Property: if $q \geq C\delta^{-2} s \ln(n/s)$, then with w/hp

$$\left| \frac{1}{q}\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2 \right| \leq \delta \|\mathbf{x}\|_2^2$$

  for all $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_0 \leq s$.

Ingredients for the Proofs, ctd

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap S^{n-1}$ with $\mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}\rangle = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}'\rangle$, $i = 1, \ldots, q$, satisfy

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap S^{n-1}$ with $\mathrm{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}' \rangle$, $i = 1, \ldots, q$, satisfy

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap B_2^n$:

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \dots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap S^{n-1}$ with $\mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}\rangle = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}'\rangle$, $i = 1, \dots, q$, satisfy

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap B_2^n$:
  - $\mathbf{a}_1, \dots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$,

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap S^{n-1}$ with $\mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}\rangle = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}'\rangle$, $i = 1, \ldots, q$, satisfy
$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap B_2^n$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$,
  - $\tau_1, \ldots, \tau_q \in \mathbb{R}$ independent $\mathcal{N}(0, 1)$,

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap S^{n-1}$ with $\mathrm{sign}\langle \mathbf{a}_i, \mathbf{x} \rangle = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}' \rangle$, $i = 1, \ldots, q$, satisfy

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap B_2^n$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$,
  - $\tau_1, \ldots, \tau_q \in \mathbb{R}$ independent $\mathcal{N}(0, 1)$,
  - apply the previous results to $[\mathbf{a}_i, -\tau_i]$, $[\mathbf{x}, 1]$, $[\mathbf{x}', 1]$.

Ingredients for the Proofs, ctd

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap S^{n-1}$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap S^{n-1}$ with $\mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}\rangle = \mathrm{sign}\langle \mathbf{a}_i, \mathbf{x}'\rangle$, $i = 1, \ldots, q$, satisfy

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

- Random hyperplane tessellations of $\sqrt{s}B_1^n \cap B_2^n$:
  - $\mathbf{a}_1, \ldots, \mathbf{a}_q \in \mathbb{R}^n$ independent $\mathcal{N}(0, \mathrm{I}_q)$,
  - $\tau_1, \ldots, \tau_q \in \mathbb{R}$ independent $\mathcal{N}(0, 1)$,
  - apply the previous results to $[\mathbf{a}_i, -\tau_i]$, $[\mathbf{x}, 1]$, $[\mathbf{x}', 1]$.
  - If $q \geq C\delta^{-4} s \ln(n/s)$, then w/hp all $\mathbf{x}, \mathbf{x}' \in \sqrt{s}B_1^n \cap B_2^n$ with $\mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x}\rangle - \tau_i) = \mathrm{sign}(\langle \mathbf{a}_i, \mathbf{x}'\rangle - \tau_i)$, $i = 1, \ldots, q$, satisfy

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \delta.$$

Thank you!

E-mail:

✧ dneedell@cmc.edu

Web:

✧ www.cmc.edu/pages/faculty/DNeedell

References:

- R. Baraniuk, S. Foucart, D. Needell, Y. Plan, M. Wootters. Exponential decay of reconstruction error from binary measurements of sparse signal, submitted.
- E. J. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. Communications on Pure and Applied Mathematics, 59(8):1207Ű1223, 2006.
- E. J. Candès, Y. C. Eldar, D. Needell and P. Randall. Compressed sensing with coherent and redundant dictionaries. Applied and Computational Harmonic Analysis, 31(1):59-73, 2010.
- M. A. Davenport, D. Needell and M. B. Wakin. Signal Space CoSaMP for Sparse Recovery with Redundant Dictionaries, submitted.
- D. Needell and R. Ward. Stable image reconstruction using total variation minimization. J. Fourier Analysis and Applications, to appear.
- Y. Plan and R. Vershynin. One-bit compressed sensing by linear programming, Comm. Pure Appl. Math., to appear.