Distributed Compressed Sensing with One-Bit Measurements and Hard Thresholding

Lars Palzer

Joint work with Johannes Maly

2nd TUM-EURECOM Workshop on Communications and Machine Learning
May 17, 2018
Compressed Sensing
Compressed Sensing

\[ y = A \cdot x \]

- Reconstruct "undersampled" signal \( x \in \mathbb{R}^n \) from \( m \) linear measurements.
- Basic assumption: \( \|x\|_0 \leq s \ll n \)

Compressed Sensing

Reconstruct "undersampled" signal $x \in \mathbb{R}^n$ from $m$ linear measurements

- Basic assumption: $\|x\|_0 \leq s \ll n$
- Tractable uniform recovery if $m \geq C \cdot s \ln(en/s)$
- Tractable non-uniform recovery if $m > 2s \ln(en/s)$

---

1-Bit Compressed Sensing
1-Bit Compressed Sensing

\[
y = \text{sign}(A \cdot x) = \begin{bmatrix}
\text{sign}(\langle a_1, x \rangle) \\
\vdots \\
\text{sign}(\langle a_m, x \rangle)
\end{bmatrix}
\]

- Very low complexity
- Very coarse quantization
- Amplitude information lost
Signal Approximation

**Theorem (Foucart)**

If $A$ satisfies $\text{RIP}_1(2s, \delta)$, i.e.,

$$(1 - \delta) \|z\|_2 \leq \|Az\|_1 \leq (1 + \delta) \|z\|_2$$

for all $2s$-sparse vectors $z \in \mathbb{R}^n$, every $\ell_2$-normalized $s$-sparse vector $x \in \mathbb{R}^n$ observed via $y = \text{sign}(A \cdot x)$ has

$$\|x - \mathbb{H}_s(A^T y)\|_2 \lesssim \sqrt{\delta},$$

where $\mathbb{H}_s(\cdot)$ keeps only the $s$ largest (in magnitude) entries.

---

S. Foucart, "Flavors of compressive sensing", Int. Conf. Approx. Th., 2016
Signal Approximation

Theorem (Foucart)

If $A$ satisfies $RIP_1(2s, \delta)$, i.e.,

$$(1 - \delta) \|z\|_2 \leq \|Az\|_1 \leq (1 + \delta) \|z\|_2$$

for all $2s$-sparse vectors $z \in \mathbb{R}^n$, every $\ell_2$-normalized $s$-sparse vector $x \in \mathbb{R}^n$ observed via $y = \text{sign}(A \cdot x)$ has

$$\|x - \mathbb{H}_s(A^T y)\|_2 \lesssim \sqrt{\delta},$$

where $\mathbb{H}_s(\cdot)$ keeps only the $s$ largest (in magnitude) entries.

- $RIP_1(2s, \delta)$ satisfied w.h.p. if entries of $A$ are iid Gaussian and $m \gtrsim \delta^{-2}s \ln(en/s)$

---

S. Foucart, "Flavors of compressive sensing", Int. Conf. Approx. Th., 2016
Signal Approximation

**Theorem (Foucart)**

If $A$ satisfies $\text{RIP}_1(2s, \delta)$, i.e.,

$$(1 - \delta) \|z\|_2 \leq \|Az\|_1 \leq (1 + \delta) \|z\|_2$$

for all $2s$-sparse vectors $z \in \mathbb{R}^n$, every $\ell_2$-normalized $s$-sparse vector $x \in \mathbb{R}^n$ observed via $y = \text{sign}(A \cdot x)$ has

$$\|x - \mathbb{H}_s(A^T y)\|_2 \lesssim \sqrt{\delta},$$

where $\mathbb{H}_s(\cdot)$ keeps only the $s$ largest (in magnitude) entries.

- $\text{RIP}_1(2s, \delta)$ satisfied w.h.p. if entries of $A$ are iid Gaussian and $m \gtrsim \delta^{-2} s \ln(en/s)$

**Tractable and accurate** sparse approximation possible from 1-bit measurements!

---

S. Foucart, "Flavors of compressive sensing", Int. Conf. Approx. Th., 2016


Distributed Compressed Sensing
Motivation

- Distributed sensing, joint reconstruction
- Signals share common structure
- Applications such as sensor networks, MRI or MIMO systems
Motivation

- Distributed sensing, joint reconstruction
- Signals share common structure
- Applications such as sensor networks, MRI or MIMO systems

**Idea:** use joint signal structure to reduce measurement rates
Distributed CS with 1-Bit Measurements

(our work)
Model

- Signal $X = [x_1, \ldots, x_L] \in S_{s,L}$ with

$$S_{s,L} = \left\{ Z \in \mathbb{R}^{n \times L} : \left| \text{supp}(Z) \right| \leq s, \|Z\|_2 = \frac{\|Z\|_F}{\sqrt{L}} \quad \forall \ell \in [L] \right\}.$$
Model

- Signal \( X = [x_1, \ldots, x_L] \in S_{s,L} \) with

\[
S_{s,L} = \left\{ Z \in \mathbb{R}^{n \times L} : \text{supp}(Z) \leq s, \|z_\ell\|_2 = \frac{\|Z\|_F}{\sqrt{L}} \forall \ell \in [L] \right\}.
\]

- Distributed 1-bit sensing with properly scaled Gaussian \( A_\ell \):

\[
y_\ell = \text{sign} \left( A_\ell x_\ell \right), \quad \ell \in [L]
\]
Model

- Signal $X = [x_1, \ldots, x_L] \in S_{s,L}$ with
  
  $$S_{s,L} = \left\{ Z \in \mathbb{R}^{n \times L} : |\text{supp}(Z)| \leq s, \|z_\ell\|_2 = \frac{\|Z\|_F}{\sqrt{L}} \forall \ell \in [L] \right\}. $$

- Distributed 1-bit sensing with properly scaled Gaussian $A_\ell$:
  
  $$y_\ell = \text{sign}(A_\ell x_\ell), \quad \ell \in [L]$$

- Reconstruction:
  
  $$\hat{X} = [\hat{x}_1, \ldots, \hat{x}_L] = \mathbb{H}_s(A_1^T y_1, \ldots, A_L^T y_L)$$

  where $\mathbb{H}_s(\cdot)$ keeps the $s$ largest (in $\|\cdot\|_2$) rows.
Model

- Signal $X = [x_1, \ldots, x_L] \in S_{s,L}$ with
  
  $$S_{s,L} = \left\{ Z \in \mathbb{R}^{n \times L} : |\text{supp}(Z)| \leq s, \|z_\ell\|_2 = \frac{\|Z\|_F}{\sqrt{L}} \quad \forall \ell \in [L] \right\}. $$

- Distributed 1-bit sensing with properly scaled Gaussian $A_\ell$:
  
  $$y_\ell = \text{sign}(A_\ell x_\ell), \quad \ell \in [L]$$

- Reconstruction:
  
  $$\hat{X} = [\hat{x}_1, \ldots, \hat{x}_L] = \mathbb{H}_s(A_1^T y_1, \ldots, A_L^T y_L)$$

  where $\mathbb{H}_s(\cdot)$ keeps the $s$ largest (in $\|\cdot\|_2$) rows.

**Question**: how many measurements per signal are necessary?
Main Result

**Theorem**

Fix $X \in S_{s,L}$ with $\|X\|_F = 1$. Choose

$$mL \gtrsim \delta^{-2} s \left( \ln(en/s) + L \right)$$

and let $y_\ell = \text{sign}(A_\ell x_\ell)$ for $\ell \in [L]$. Then, with probability (over the $A_\ell$s) at least $1 - \exp\left(-c\delta^2 mL\right)$,

$$\|X - \hat{X}\|_F \lesssim \sqrt{\delta}$$

where $\hat{X}$ is the result of a single hard thresholding step.
Main Result

Theorem

Fix $X \in S_{s,L}$ with $\|X\|_F = 1$. Choose

$$mL \gtrsim \delta^{-2} s \left( \ln \left( \frac{en}{s} \right) + L \right)$$

and let $y_\ell = \text{sign}(A_\ell x_\ell)$ for $\ell \in [L]$. Then, with probability (over the $A_\ell$s) at least $1 - \exp \left( -c\delta^2 mL \right)$,

$$\|X - \hat{X}\|_F \lesssim \sqrt{\delta}$$

where $\hat{X}$ is the result of a single hard thresholding step.

Remarks:

- For $L \gtrsim \ln(en/s)$, we need effectively $m \gtrsim \delta^{-2} s$ measurements per signal
- Non-uniform recovery result
- Bounds total error of the $L$ signals
Numerical Experiments
Numerical Experiments (1)

$n = 100$, $s = 5$, 500 trials, $\|X\|_F = 1$, $x_1, \ldots, x_L$ uniform on sphere in a random $s$-dim subspace of $\mathbb{R}^n$
Numerical Experiments (2)

Average error $\|X - \hat{X}\|_F$ over 500 trials, $n = 100$
Proof Strategy
Proof Ingredients

- Derive\(^1\) appropriate RIP for \(A_1, \ldots, A_L\) and \(S_{s,L}\)

---

RIP_{2,1}

Denote

\[ A = \begin{pmatrix} A_1 \\ \vdots \\ A_L \end{pmatrix} \]

and recall

\[ S_{s,L} = \left\{ Z \in \mathbb{R}^{n \times L} : \right. \begin{align*}
&|\text{supp}(Z)| \leq s, \\
&\|z_\ell\|_2 = \frac{\|Z\|_F}{\sqrt{L}} \quad \forall \ell \in [L] \end{align*} \right\}. \]

Lemma (RIP_{2,1})

Let the entries of \( A_\ell, \ell \in [L] \), be iid \( \mathcal{N}(0, \pi/(2Lm^2)) \). For

\[ mL \gtrsim \delta^{-2} s \left( \ln(en/s) + L \right) \]

the \( A \) satisfies

\[ (1 - \delta) \|X\|_F \leq \|A \cdot \text{vec}(X)\|_1 \leq (1 + \delta) \|X\|_F \]

for all \( X \in S_{s,L} \) with probability at least \( 1 - \exp\left(-c\delta^2 mL\right) \).
Proof Ingredients

- Derive\(^2\) appropriate RIP for \(A_1, \ldots, A_L\) and \(S_{S,L}\)

- Show concentration of \(\|[A_1^T y_1, \ldots, A_L^T y_L]_S\|_F\)

Concentration of $\|[A_1^T y_1, \ldots, A_L^T y_L]_S\|_F$

Lemma

Let $B = \left[ A_1^T y_1, \ldots, A_L^T y_L \right]_{S \restrictedto \text{supp}(X)} \in \mathbb{R}^{n \times L}$. Then,

$$
\mathbb{P}\left( \left\| B \right\|_F - \sqrt{\mathbb{E}[\|B\|_F^2]} \geq \delta \right) \leq 2 \exp\left(-c \min\left\{ \frac{\delta^4 m^2}{s^2}, \frac{\delta^2 m}{s} \right\} \right).
$$

- Concentration for a single $X$ if $m \gtrsim \delta^{-2} s$
- Since $\|A_\ell^T y_\ell\|_2$ is independent of $\ell$, we can find $\tilde{B} \in S_{s,L}$ close to $B$
Proof Ingredients

- Derive\textsuperscript{3} appropriate RIP for $A_1, \ldots, A_L$ and $S_{s,L}$
- Show concentration of $\| [A_1^T y_1, \ldots, A_L^T y_L ]_S \|_F$
- Conclude that\textsuperscript{4} $\| X - \hat{X} \|_F \leq 2\sqrt{10\delta}$


\textsuperscript{4} inspired by S. Foucart, "Flavors of compressive sensing", Int. Conf. Approx. Th., 2016
Summary
Summary / Remarks

- Non-uniform approximation result for distributed 1-bit CS with hard thresholding
- For $L \gtrsim \ln(en/s)$, measurements per signal scale linearly in the row-sparsity
- Theoretical scaling clearly visible in numerical experiments
Thanks!

arXiv paper:
"Analysis of Hard-Thresholding for Distributed Compressed Sensing with One-Bit Measurements",
https://arxiv.org/abs/1805.03486