Compressed Sensing & Wireless Communications

Dongning Guo

with Lei Zhang, Jun Luo, Ming Gan

Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, IL

Presented at the 2nd International Workshop on Mathematical Issues in Information Sciences, Xi’an, China
July 1, 2013
1 Compressed Sensing

2 Discovery, Messaging, Ranging and Localization via Compressed Sensing

3 Other Applications of Compressed Sensing in Wireless Systems
1 Compressed Sensing

2 Discovery, Messaging, Ranging and Localization via Compressed Sensing

3 Other Applications of Compressed Sensing in Wireless Systems
Geometry of sparse recovery

\[ y = \underbrace{S}_{M \times N} x \]

\[ M \times 1 \quad M \times N \quad N \times 1 \]

\( K \) sparse
\[ \hat{x} = \arg \min_{y = Sx'} \|x'\|_0 \]

Minimum $L_0$ solution correct if $M \geq 2K$. (w.p. 1 for Gaussian $S$)
Why $L_2$ recovery doesn’t work

$$\hat{x} = \arg\min_{y=Sx'} \|x'\|_2$$

least squares
minimum $L_2$ solution is almost never sparse
\( \hat{x} = \arg \min_{y=Sx'} \|x'\|_1 \)

Minimum \( L_1 \) solution is identical to \( L_0 \) sparsest solution if \( M \geq K \log N \ll N \)
What it takes for $L_1$ recovery to succeed

\[ \hat{x} = \arg \min_{y=Sx'} \|x'\|_1 \]

If we ensure that with high probability that a randomly oriented $(N - M)$-plane, anchored on a $K$-face of the $L_1$ ball, will not intersect the ball.

Need $K$ small and $N - M$ small.
Equivalence of $L_0$ and $L_1$

**Theorem (Donoho, Tanner)**

For Gaussian $S$, it suffices to have

$$M \approx 2eK \log\left(\frac{N}{M\sqrt{\pi}}\right)$$

measurements to recover every $K$-sparse signal using $L_1$ minimization. It suffices to have

$$M \approx 2K \log\left(\frac{N}{M}\right)$$

measurements to recover a large majority of $K$-sparse signals.
Ristricited isometry property

- [Candés, Romberg, Tao] Measurement matrix $S$ has RIP of order $K$ if

$$ (1 - \delta_K) \leq \frac{\|Sx\|_2^2}{\|x\|_2^2} \leq (1 + \delta_K) $$

for all $K$-sparse signals $x$.

- Does not hold for $K > M$; may hold for smaller $K$.

- Implications: tractable, stable, robust $L_1$ recovery.
Bayesian framework

- Early studies aim at recovery of (almost) every sparse signal.
- Noise is addressed through robustness of recovery.
- However, signal and noise are often statistical in engineering problems.
- Bayesian sparse recovery studied in, e.g., [Baron, Sarvotham & Baraniuk '08], [Guo, Baron & Shamai '09], [Wu & Verdú '10], [Donoho, Javanmard & Montanari '12].
- Characterization of noisy CS [Guo, Baron & Shamai '09], [Fletcher, Rangan & Goyal '09], [Donoho, Maleki & Montanari '12].
**Characterization of noisy CS**

\[ Y = \sqrt{\gamma} SX + N \]

**Proposition (Guo, Baron & Shamai ’09)**

\( X_n \text{ i.i.d. } \sim P_X \). \( S \) i.i.d. entries with zero mean and unit variance. As \( N, M \to \infty \) with \( N/M \to \beta \), the quality of estimating \( X_n \) becomes equivalent to estimating \( X \sim P_X \) from

\[ Z = \sqrt{\eta\gamma} X + N(0,1) \]

where

\[ \frac{1}{\eta} = 1 + \beta\gamma \text{mmse}(X|\sqrt{\eta\gamma} X + N(0,1)). \]

Related work: [Fletcher, Rangan & Goyal ’09], [Donoho, Maleki & Montanari ’12]
1. Compressed Sensing

2. Discovery, Messaging, Ranging and Localization via Compressed Sensing

3. Other Applications of Compressed Sensing in Wireless Systems
Discovery, messaging, ranging and localization

- **Discovery** — Nodes discover and acquire one-hop neighbors’ network interface addresses (NIAs), such as
  - access point or UE IDs in a heterogeneous network,
  - machine IDs or MAC addresses in a M2M or WiFi network,
  - sensor IDs in a sensor network.

- **Messaging** — Nodes exchange messages with one-hop neighbors, such as
  - request-to-send, clear-to-send,
  - power, rate, and modulation information,
  - queue length, scheduling, routing information.

- **Ranging and localization** — Nodes measure pairwise distances and infer about their own location in, e.g.,
  - vehicle-to-vehicle, robot-to-robot networks,
  - sensor networks.
Messing is fundamentally compressed sensing

- Suppose neighbors/peers have been discovered and identified.
- Each node wishes to send data or a (common) message of $l$ bits to all or a subset of one-hop peers.
- Node $j$ has $2^l$ codewords $s_j(1), \ldots, s_j(2^l)$.
- Linear measurements:

$$Y_k = \sum_{j \in \partial k} U_{jk} \sqrt{\gamma_j} s_j(w_j) + W_k$$

$$= SX + W_k$$

- Out of a total of $2^l|\partial k|$ codewords from all neighbors, which $|\partial k|$ codewords, one from each codebook, were transmitted?
- The sparsity of $X$ is $2^{-l}$.
  $l$ is large for long data packets, small for short messaging.
Half duplex constraint

Frame transmissions often scheduled away from reception, via frequency-division duplex (FDD) or time-division duplex (TDD)
State of the art: random access

- Repeat packets interleaved with random delay (ALOHA, CSMA, CSMA/CA)
- Drawback: data and energy loss due to collision
Virtual full duplex

- The idea [Guo–Zhang ’10]: rapid on-off-division duplex (RODD)
- Half duplex ⇔ received signal erased by own transmissions.
- Introduce off-slots in transmission frame to allow reception
- On-off signaling at symbol/slot level
  (a slot interval ≪ a frame interval)
RODD with multiple users

- Multiaccess channel with erasure.
- Random schedule in a microscopic timescale.
- Error control coded over each frame.
- More stable access delay, simpler higher-layer protocols.
Capacity of RODD

Theorem (Guo & Zhang ’10)

A clique of $K+1$ nodes, each link with SNR $\gamma$:

$$Y_{nm} = (1 - s_{nm})\sqrt{\gamma} \sum_j s_{jm} X_{jm} + V_{nm}$$

Everyone broadcasts a message to neighbors over a frame. The best symmetric rate is

$$C = \frac{1 - q}{2K} \sum_{\kappa=1}^{K} \binom{K}{\kappa} q^\kappa (1 - q)^{K-\kappa} \log \left( 1 + \frac{\kappa \gamma}{q} \right)$$
RODD vs. slotted ALOHA

RODD symmetric rate:

\[
C = \frac{1 - q}{2K} \sum_{\kappa=1}^{K} \binom{K}{\kappa} q^{\kappa} (1 - q)^{K-\kappa} \log \left( 1 + \frac{\kappa \gamma}{q} \right) = \frac{1 - q}{2K} \binom{K}{1} q (1 - q)^{K-1} \log \left( 1 + \frac{\gamma}{q} \right)
\]

Total throughput of ALOHA:

\[
\frac{K}{2} q (1 - q)^K \log \left( 1 + \frac{\gamma}{q} \right) < C
\]
Messingg throughput: RODD vs. ALOHA

Throughput (bits/channel use)

\[
t_{\text{RODD}}(q) = \begin{cases} 
20 & \text{for } q = 0 \\
5 & \text{for } q = 1 
\end{cases}
\]

\[
t_{\text{ALOHA}}(q) = \begin{cases} 
20 & \text{for } q = 0 \\
5 & \text{for } q = 1 
\end{cases}
\]
RODD vs. random access

RODD sees an **ergodic multiaccess channel with erasure**. ALOHA/CSMA sees a **nonergodic packet-based multiaccess channel**.
RODD vs. random access

[Zhang–Guo ’13] 9 neighbors on average, each broadcasts 10 bits, SNR=10 dB between all links, no fading
RODD vs. random access

[Zhang–Guo '13] Large Poisson network, 9 neighbors on average, each broadcasts 5 bits, with path loss $\alpha = 3$, Rayleigh fading.
Neighbor discovery: prior art

- Random access discovery:
  Each node sends its NIA repeatedly with random delay.
  [McGlynn & Borbash ’01], [Borbash, Ephremides & McGlynn ’07],
  [Vasudevan, Kurose & Towsley ’05], [Khalili, Goeckel, Towsley &
  Swami ’10], [Ni, Srikant & Wu ’10], [Felemban et al ’10]

- Existing protocols:
  - TND Protocol (IETF MANET Workgroup)
  - WiFi ad hoc mode
  - FlashLinQ (single-tone OFDM, CSMA-like)
Discovery is fundamentally a compressed sensing problem

- Consider the entire discovery period of $M$ symbol intervals.
- $N$ NIAs total. Node $n$ sends signal $s_n$ (including delay).
- Linear measurements via multiaccess channel (w/ fading):

$$Y = \sum_{n \in \partial k} s_n U_n + W$$

$$= \sum_{n=1}^{N} s_n X_n + W$$

$$= SX + W$$

- $X_n \approx 0$ for all but a few neighbors.
Signaling

- Synchronized transmissions;
- Each node transmits a single frame (a unique signature);
- Network-wide (full-duplex) discovery in a single frame interval;
- In case of half-duplex constraint:
  - on-off signatures (RODD);
  - listen through off-slots.
- One key challenge is the decoding complexity (need to scale to $2^{20} - 2^{48}$ NIAs).
Random vs. deterministic signatures

- Random signatures used for
  - discovery w/ group testing decoding in [Luo & Guo ’08, ’09];
  - messaging w/ belief propagation decoding [Zhang & Guo ’12].

- Decoding complexity at least $O(MN)$ because basically all signatures need be visited.

- Too expensive to solve for over $2^{20}$ unknowns.

- Random signatures not suitable for large networks; need structured codes/signatures.
Deterministic signatures

- We use signatures from second-order Reed-Muller codes;
- RM codes known to be efficient for CS [Calderbank, Gilbert & Strauss ’06].
- Signatures are basically discrete chirps.
- Introduce random erasures to create on-off signatures if w/ half-duplex constraint.
- Low-complexity chirp decoding algorithm [Howard, Calderbank & Searle ’08].
For \( x, l \in \mathbb{Z}_2^m \), \( P_{m \times m} \) is a binary symmetric matrix:

- Second-order Reed-Muller, RM(2):

\[
\varphi_{P,l}(x) = j x^T Px + 2l^T x, \quad j = \sqrt{-1}.
\]

- With signature length \( M = 2^m \), codebook size up to \( 2^{m(m+3)/2} \):
  - \( m = 5, M = 2^5 = 32, N \) up to \( 2^{20} \) codewords.
  - \( m = 10, M = 2^{10} = 1,024, N \) up to \( 2^{65} \);
  - \( m = 12, M = 2^{12} = 4,096, N \) up to \( 2^{90} \).

- Introduce about 50% erasures.
Error rate vs. SNR

$2^{20}$ nodes, path loss exponent = 3, Rayleigh fading

- $M = 4,096$, 30 neighbors on average
- $M = 1,024$, 10 neighbors on average

Dongning Guo (Northwestern Univ.)
Compressed Sensing & Wireless Communications
Comparison with random access

\[ N = 2^{20} \text{ nodes, on average 10 neighbors, SNR = 11.5 dB} \]

Target \( P_e = 0.002 \)

<table>
<thead>
<tr>
<th></th>
<th>Random access</th>
<th>RODD</th>
</tr>
</thead>
<tbody>
<tr>
<td># of frames</td>
<td>194</td>
<td>1</td>
</tr>
<tr>
<td># of symbols</td>
<td>( \geq 194 \times 20 = 3,880 )</td>
<td>1,024</td>
</tr>
</tbody>
</table>

In addition, significant reduction of per-frame overhead
Ranging and localization is a compressed sensing problem

- A network of wireless nodes, a fraction of which are anchors who know their exact location.

- Ranging via received signal strength

\[ P_r = P_t d^{-\alpha} \quad \Rightarrow \quad d = \left( \frac{P_t}{P_r} \right)^{1/\alpha} \]

- Nodes repeatedly exchange their estimated location with one-hop neighbors so that all nodes can improve their estimates.

- Localization based on distance [Srirangarajan–Tewfik–Luo 2008]:

\[ \min_{x_1, \ldots, x_n} \sum_{(i,j) \in A} \left| \| x_i - x_j \|^2 - d_{ij}^2 \right| \]

Can be relaxed to a convex optimization problem and solved efficiently.
Localization: preliminary numerical results
Compressed Sensing

Discovery, Messaging, Ranging and Localization via Compressed Sensing

Other Applications of Compressed Sensing in Wireless Systems
Sparse multipath channel estimation

- **Ultra-wideband:** The impulse response is a combination of a few atoms from a large dictionary. [Paredes–Arce–Wang ’07] proposes a CS based scheme which outperforms traditional correlator.

![Graphs](left) impulse response with LOS (right) without LOS.

- **Sparse multipath channels:** A CS-based scheme proposed in [Bajwa–Haupt–Sayeed–Nowak ’10, Haupt–Bajwa–Raz–Nowak ’10] with significant gain in energy efficiency and latency.
Sparse doubly selective channel estimation

- Underwater acoustic channel [Berger–Zhou–Preisig–Wilett ’10]:

\[ c(\tau, t) = \sum_{p=1}^{P} A_p \delta(\tau - (\tau_p - a_p t)) \]

After resampling that corresponds to a rough Doppler estimate,

\[ z(t) = \sum_{p=1}^{P} A_p x((1 + b_p)(t - \tau_p')) + w(t) \]

Wish to estimate \((A_p, b_p, \tau_p')\), \(p = 1, \ldots, P\). With multicarriers, an equivalent discrete-time model is

\[ z = Hs + v \]

\(H\) has \(K^2\) entries but depends on \(3P \ll K^2\) parameters. CS outperforms array processing algorithms MUSIC and ESPRIT.

- Related work [Tauböck–Hlawatsch–Eiwen–Rauhut ’10].
Sensor networks

- [Ling–Tian ’11] proposed CS algorithms for recovering a sparse signal that represents the physical field using a small number of sensory measurements.
[Herman–Strohmer ’09] transmits a sufficiently “incoherent” pulse and use CS to reconstruct the target scene. Successful as long as the number of targets \( \ll N^2 \).
Multiuser activity and signal detection

\[ Y = s_1 X_1 + s_2 X_2 + \cdots + s_N X_N + W \]
\[ = \mathbf{S} \mathbf{X} + W \]

If many of the \( N \) users are inactive, so only a few \( X_k \neq 0 \). To detect active users and their symbols is a CS problem.

- Multiuser activity detection was studied in [Lin–Lim '04, Biglieri–Lops '07, Angelosante–Biglieri-Lops '09], but not using CS.
- User activity detection studied in [Luo–Guo '08] in the context of neighbor discovery, using group testing, a special CS algorithm.
- Sparsity exploiting sphere decoder studied in [Zhu–Giannakis '11]
- CS-based coded multiuser detection proposed in [Bockermann–Schepker–Dekorsy '13]
- Reduced-dimension multisuer detection in [Xie–Eldar–Goldsmith '13]
Concluding remarks

- Compressed sensing finds many applications in wireless systems.
- Would be interesting to evaluate the performance of discovery, messaging and localization.
- Practical codes.
- Virtual full duplex:
  - implementation
  - capacity in general
- Other (complementary) means of full-duplex communication:
  
  [Radunovic, Gunawardena, Key, Singh, Balan & Dejean ’09]
  [Choi, Jain, Srinivasan, Levis & Katti ’10]
  [Duarte & Sabharwal ’10]