Distributed Compressive Sensing

A Framework for Integrated Sensing and Processing for Signal Ensembles

Rice University
dsp.rice.edu/cs

Marco Duarte
Shriram Sarvotham
Michael Wakin
Dror Baron
Richard Baraniuk
DSP Sensing

• The typical sensing/compression setup
  – compress = transform, sort coefficients, encode
  – most computation at sensor (asymmetrical)
  – lots of work to throw away >80% of the coefficients
Compressive Sensing (CS)

- Measure projections onto *incoherent* basis/frame
  - random “white noise” is *universally incoherent*
- Reconstruct via nonlinear techniques
- Mild oversampling: \( M \geq cK \ll N, \quad c \approx 3 \)
- Highly asymmetrical (most computation at *receiver*)
CS Reconstruction

- Underdetermined
- Possible approaches:

\[ \hat{x} = \arg \min_{y=\Phi x} ||x||_2 \]  
Wrong solution (not sparse)

\[ \hat{x} = \arg \min_{y=\Phi x} ||x||_0 \]  
Right solution, but not tractable

\[ \hat{x} = \arg \min_{y=\Phi x} ||x||_1 \]  
Right solution and tractable if \( M > cK \) (\( c \sim 3 \) or 4)

- Also: efficient greedy algorithms for sparse approximation
Compressive Sensing

• CS changes the rules of the data acquisition game
  – changes what we mean by “sampling”
  – exploits a priori signal/image *sparsity* information
    (that the signal is compressible in some representation)
  – Related to multiplex sampling (D. Brady - DISP)

• Potential next-generation data acquisition
  – new distributed source coding algorithms for multi-sensor applications
  – new A/D converters (sub Nyquist) [Darpa A2I]
  – new mixed-signal analog/digital processing systems
  – new imagers, imaging, and image processing algorithms
  – ...
Permuted FFT (PFFT)

- Longer signals via “random” transforms
- Non-Gaussian measurement scheme

\[ y = T \mathcal{F} \mathcal{P} x \quad \Phi = T \mathcal{F} \mathcal{P} \]

- Low complexity measurement
- (approx O(N) versus O(MN))
  - universally incoherent
- Low complexity reconstruction
  - e.g., Matching Pursuit
    \[ p = y^H \Phi = (\mathcal{P}^H \mathcal{F}^H T^H y)^H \]
    - compute using transforms
      (approx O(N^2) versus O(MN^2))

\[ \text{Reconstruction PSNR, dB} \]
\[ \text{Number of measurements, } M \]
Reconstruction from PFFT Coefficients

- Original 65536 pixels
- Wavelet Thresholding 6500 coefficients
- CS Reconstruction 26000 measurements

- 4x oversampling enables good approximation
- Wavelet encoding requires
  - extra location encoding + fancy quantization strategy
- Random projection encoding requires
  - no location encoding + only uniform quantization
Random Filtering
[with J. Tropp]

• Hardware/software implementation

\[ x \xrightarrow{\mathcal{H}} y = \mathcal{D}\mathcal{H}x \]

“Random”
FIR Filter

Downsample
(keep \(M\) out
of \(N\))

• Structure of \(\Phi\)
  – convolution \(\rightarrow\) Toeplitz/circulant
  – downsampling \(\rightarrow\) keep certain rows
  – if filter has few taps, \(\Phi\) is sparse
  – potential for fast reconstruction

\[ \Phi = \mathcal{D}\mathcal{H} \]

• Can be generalized to analog input \(x\)
Rice CS Camera

- Single photon detector
- Low-cost, fast, sensitive optical detection
- Image encoded by PMM and random basis
- Random pattern on DMD array (see also Coifman et al.)

Diagram:
- Image input
- PD (photo detector)
- A/D
- Xmtr (transmitter)
- Compressed, encoded image data sent via RF for reconstruction
- Rcvr (receiver)
- DSP (digital signal processor)
- Image reconstruction
Correlation in Signal Ensembles

- Sensor networks: *intra*-sensor and *inter*-sensor correlation dictated by physical phenomena
- Can we exploit these to *jointly compress*?
- Popular approach: *collaboration*
  - inter-sensor *communication overhead*
  - complexity at sensors
- *Ongoing challenge* in information theory community
Distributed Compressive Sensing (DCS)

Joint sparsity models and algorithms for different physical settings

**Benefits:**
- **Distributed Source Coding:**
  - exploit intra- and inter-sensor correlations
  - fewer measurements necessary
  - zero inter-sensor communication overhead
- **Compressive Sensing:**
  - universality (random projections)
  - “future-proof”
  - encryption
  - robustness to noise, packet loss
  - scalability
  - low complexity at sensors
JSM-1: Common + Innovations Model

- **Motivation:** sampling signals in a smooth field

- **Joint sparsity model:**
  - length-$N$ sequences $x_1$ and $x_2$
  - $z$ is length-$N$ *common* component
  - $z_1, z_2$ length-$N$ *innovation* components
  - $z$ has sparsity $K$
  - $z_1, z_2$ have sparsity $K_1, K_2$

- **Measurements**
  \[ y_1 = \Phi_1 x_1 \]
  \[ y_2 = \Phi_2 x_2 \]

- **Intuition:** Sensors should be able to “share the burden” of measuring $z$
JSM-2: Common Sparse Supports

- measure $J$ signals, each $K$-sparse
- signals share sparse components, different coefficients

\[ x_j = \sum_{\omega \in \Omega} x_{j,\omega} \psi_\omega, \]
\[ |\Omega| = K \]
Joint sparsity model #3 (JSM-3):
- generalization of JSM-1,2: length-$N$ sequences $x_j$
  ⇒ each signal is incompressible
- signals may (DCS-2) or may not (DCS-1) share sparse supports

Intuition: each measurement vector contains clues about the common component $z$

$$x_j = z + z_j$$
$$z_j = \sum_{\omega \in \Omega} z_{j,\omega} \psi_\omega,$$
$$|\Omega| = K$$
DCS Reconstruction

- Measure each $x_j$ independently with $M_j$ random projections $y_j = \Phi_j x_j$
- Reconstruct jointly at central receiver
  - "What is the sparsest joint representation that could yield all measurements $y_j$?"
  - linear programming: use concatenation of measurements $y_j$
  - greedy pursuit: iteratively select elements of support set $\Omega$
  - similar to single-sensor case, but more clues available
Theoretical Results

- $M_j \ll N$ (standard CS results); further reductions from joint reconstruction
  - JSM-1: Slepian-Wolf like bounds for linear programming
  - JSM-2: $c = 1$ with greedy algorithm as $J$ increases. Can recover $\Omega$ with $M_j = 1$!
  - JSM-3: Can measure at $M_j = cK_j$, essentially neglecting $z$; use iterative estimation of $z$ and $z_j$. Would otherwise require $M_j = N$!
JSM-1: Recovery via Linear Programming

\( K_C = 11, K_1 = K_2 = 2, N = 50, \gamma_C = 0.905 \)

\( K_C = 9, K_1 = K_2 = 3, N = 50, \gamma_C = 1.01 \)

\( K_C = 5, K_1 = K_2 = 3, N = 50, \gamma_C = 1.27 \)

\( K_C = 3, K_1 = K_2 = 6, N = 50, \gamma_C = 1.425 \)
Recovery via Linear Programming

For $K_C = 11, K_1 = K_2 = 2, N = 50, \gamma_C = 0.905$

For $K_C = 9, K_1 = K_2 = 3, N = 50, \gamma_C = 1.01$

For $K_C = 5, K_1 = K_2 = 5, N = 50, \gamma_C = 1.27$

For $K_C = 3, K_1 = K_2 = 6, N = 50, \gamma_C = 1.425$
$K=5$
$N=50$

JSM-2 SOMP Results

![Graph showing the probability of exact reconstruction versus the number of measurements per signal, with distinct curves for different numbers of measurements and labels for separate and joint reconstruction.]
Experiment: Sensor Network Light data with JSM-2

Wavelet approximation, 100 coefficients, mean SNR = 26.4842 dB

OMP separate reconstruction, mean SNR = 23.1682 dB

SOMP separate reconstruction, mean SNR = 29.4283 dB
$K = 5$
$N = 50$

JSM-3 ACIE/SOMP Results
(same supports)

Impact of $\varpropto$ vanishes as $J \to \infty$

Probability of Exact Reconstruction

Number of Measurements per Signal, $M$