Model-Based Compressive Sensing for Signal Ensembles

Marco F. Duarte
Volkan Cevher
Richard G. Baraniuk
Concise Signal Structure

- **Sparse** signal: only $K$ out of $N$ coordinates nonzero
  - model: *union of $K$-dimensional subspaces* aligned with coordinate axes
Compressive Sensing

- Replace **samples** by more general **encoder** based on a few linear projections (inner products)
- Recover $x$ from $y$ using **optimization** ($\ell_1$-norm minimization, LPs, QPs) or **greedy algorithms** (OMP, CoSaMP, SP, etc.), $x$ is sparse

$$y = \Phi x$$

$$M \approx K \ll N$$
Restricted Isometry Property (RIP)

- Preserve the structure of sparse signals
- RIP of order $2K$ implies: for all $K$-sparse $x_1$ and $x_2$

\[(1 - \delta_{2K}) \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq (1 + \delta_{2K})\]

[Image of $K$-planes in $\mathbb{R}^N$ mapped to $\mathbb{R}^M$ by $\Phi$]

[Candès and Tao]
Restricted Isometry Property (RIP)

- Preserve the structure of sparse signals
- Random (iid Gaussian, Bernoulli) matrix has the RIP with high probability if

\[ M = O(K \log(N/K)) \]

[Candès and Tao; Baraniuk, Davenport, DeVore and Wakin]
Sensor Networks

- Networks of many *sensor nodes*
  - sensor, microprocessor for computation, wireless communication, networking, battery
  - sensors observe *single event*, acquire *correlated signals*

- Must be energy efficient
  - *minimize communication* at expense of off-site computation
  - motivates *distributed compression*
Distributed
Compressive Sensing (DCS)

\[ y_1 = \Phi_1 x_1 \]
\[ y_2 = \Phi_2 x_2 \]
\[ \vdots \]
\[ y_J = \Phi_J x_J \]

[Baron, Duarte, Wakin, Sarvotham, Baraniuk]
Distributed Compressive Sensing (DCS)

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_J 
\end{bmatrix} =
\begin{bmatrix}
  \Phi_1 & 0 & \cdots & 0 \\
  0 & \Phi_2 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \Phi_J 
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_J 
\end{bmatrix}
\]

Joint Recovery

[Baron, Duarte, Wakin, Sarvotham, Baraniuk]
JSM-2: Common Sparse Supports Model

- Measure $J$ signals, each $K$-sparse
- *Signals share sparse components but with different coefficients*
- Recovery using Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm [Tropp, Gilbert, Strauss]

\[
x_j = \sum_{n \in \Omega} \theta_j(n) \psi_n, \quad |\Omega| = K
\]
Beyond Sparse Models

- Sparse signal model captures only **simplistic primary structure**

- For many signal types, location of nonzero coefficients in sparse representation provide **additional structure**
Beyond Sparse Models

- Sparse signal model captures only **simplistic primary structure**

- For many signal types, location of nonzero coefficients in sparse representation provide **additional structure**
Sparse Signals

- Defn: A $K$-sparse signal lives on the collection of $K$-dim subspaces aligned with coord. axes
Structured Sparse Signals

- Defn: A $K$-structured sparse signal lives on a particular (reduced) collection of $K$-dim canonical subspaces

[Lu and Do]
[Blumensath and Davies]
Sparse Signal Ensemble

Structured Sparse Signal Ensemble

- **Defn:** An *structured ensemble* of $J K$-sparse signals with *common sparse support* lives on a particular (reduced) collection of $JK$-dim canonical subspaces.
RIP for Structured Sparsity Model

• Preserve the structure **only** of sparse signals that **follow the structure**

• Random (i.i.d. Gaussian, Bernoulli) matrix has the JSM-2 RIP with high probability if

\[
M = O(JK + \log m_K)
\]

[Blumensath and Davies]
RIP for Structured Sparsity Model

- Preserve the structure **only** of sparse signals that **follow the structure**

- Random (i.i.d. Gaussian, Bernoulli) matrix has the JSM-2 RIP with high probability if

\[
M = O(JK + \log m_K)
\]

[Blumensath and Davies]
RIP for Common Sparse Support Model

- Random (i.i.d. Gaussian, Bernoulli) matrix has the model-based RIP with high probability if

\[ M = O(KJ + K \log(N/K)) \]

- Distributed settings: measurements from different sensors can be \textit{added together} to effectively obtain dense measurement matrix.
Standard CS Recovery

**CoSaMP**  
[Needell and Tropp]

- calculate current residual
  \[ r = y - \Phi \hat{x} \]
- form residual signal estimate
  \[ e = \Phi^T r \]
- calculate enlarged support
  \[ \Omega = \text{supp}(\hat{x}) \cup \text{supp}(\mathcal{S}(e, 2K)) \]
- estimate signal for enlarged support
  \[ b|_{\Omega} = \Phi|_{\Omega} \dagger y, \quad b|_{\Omega^c} = 0 \]
- shrink support
  \[ \hat{x} = \mathcal{S}(b, K) \]
Model-Based CS Recovery

Model-based CoSaMP

$\mathcal{M}_K$: $K$-term structured sparse approximation algorithm

- calculate current residual
  \[ r = y - \Phi \hat{x} \]

- form residual signal estimate
  \[ e = \Phi^T r \]

- calculate enlarged support
  \[ \Omega = \text{supp}(\hat{x}) \cup \text{supp}(\mathcal{M}_{2K}(e)) \]

- estimate signal for enlarged support
  \[ b|_{\Omega} = \Phi|_{\Omega}^\dagger y, \quad b|_{\Omega^c} = 0 \]

- shrink support
  \[ \hat{x} = \mathcal{M}_K(b) \]

[Baraniuk, Cevher, Duarte, Hegde 2008]
Model-Based Recovery for JSM-2

Model-based Distributed CoSaMP → CoSOMP

- calculate current residual \textit{at each sensor} 
  \[
  r_j = y_j - \Phi_j \hat{x}_j
  \]

- form residual signal estimate \textit{at each sensor} 
  \[
  e_j = \Phi_j^T r_j
  \]

- \textit{merge sensor estimates}
  \[
  e = \sum_{j=1}^{J} (e_j \cdot e_j)
  \]

- \textit{calculate enlarged support} 
  \[
  \Omega = \text{supp}(\hat{x}) \cup \text{supp}(\mathcal{T}(e, 2K))
  \]

- estimate signal proxy \textit{at each sensor} 
  \[
  b_j|_\omega = \Phi_j|_\Omega y_j, \quad b_j|_{\omega^c} = 0
  \]

- \textit{merge sensor estimates} 
  \[
  b = \sum_{j=1}^{J} (b_j \cdot b_j)
  \]

- \textit{shrink estimate support} 
  \[
  \Lambda = \text{supp}(\mathcal{T}(b, K))
  \]

- update signal estimates \textit{at each sensor} 
  \[
  \hat{x}_j|_\Lambda = b_j|_\Lambda, \quad \hat{x}_j|_{\Lambda^c} = 0
  \]
Model-Based CS Recovery Guarantees

**Theorem:**
Assume we obtain noisy CS measurements of a signal ensemble \( \mathbf{Y} = \Phi \mathbf{X} + \mathbf{n} \). If \( \Phi \) has the model-based RIP with \( \delta_K < 0.1 \), then we have

\[
\| \mathbf{X} - \hat{\mathbf{X}} \|_2 \leq C_1 \| \mathbf{X} - \mathcal{M}_K(\mathbf{X}) \|_2 + \frac{C_2}{\sqrt{K}} \| \mathbf{X} - \mathcal{M}_K(\mathbf{X}) \|_1 + C_3 \| \mathbf{n} \|_2
\]

In words, **instance optimality** based on **structured sparse approximation**.
Model-Based CS Recovery Guarantees

**Theorem:**
Assume we obtain noisy CS measurements of a signal ensemble $\mathbf{Y} = \Phi \mathbf{X} + \mathbf{n}$. If each $\Phi_j$ has the RIP with $\delta_K < 0.1$, then we have

$$\|\mathbf{X} - \hat{\mathbf{X}}\|_2 \leq C_1 \|\mathbf{X} - \mathcal{M}_K(\mathbf{X})\|_2 + \frac{C_2}{\sqrt{K}} \|\mathbf{X} - \mathcal{M}_K(\mathbf{X})\|_1 + C_3 \|\mathbf{n}\|_2$$

In words, *instance optimality* based on *structured sparse approximation* error.
Real Data Example

• Environmental Sensing in Intel Berkeley Lab
• $J = 49$ sensors, $N = 1024$ samples each

• Compare:
  - independent recovery CoSaMP
  - existing joint recovery SOMP
  - model-based joint recovery CoSOMP
Experimental Results - Brightness Data

(a) Original signal

(c) CoSaMP recovery, distortion = 15.1733 dB

(d) CoSOMP recovery, distortion = 16.3197 dB

\[ N = 1024, \quad J = 48, \quad M = 400 \]
Experimental Results - Humidity Data

![Graph showing experimental results]

- Number of measurements, $M$
- Average recovery distortion, dB

$N = 1024, J = 48$

Legend:
- CoSaMP (dashed blue line)
- CoSOMP (solid red line)
- SOMP (solid black line)
Experimental Results - Network Size

Number of measurements, $M$

Probability of exact recovery

$N = 1024$
Conclusions

• **Intuitive union-of-subspaces model** to encode structure of jointly sparse signal ensembles

• Structure enables *reduction* in number of measurements required for recovery

• Signal recovery algorithms are *easily adapted* to leverage additional structure

• **New structure-based recovery guarantees** such as instance optimality